

**Sentiment Analysis in Social Media:
A Deep Learning Framework for Analyzing Public
Sentiment toward Policies during Natural Disasters**

تحليل المشاعر والعواطف في وسائل التواصل الاجتماعي:
منهج متكامل لتحليل مشاعر العامة تجاه السياسات خلال الكوارث الطبيعية
باستخدام طرق التعلم العميق

by

MOHAMED ABDULLA FOOLAD

**A thesis submitted in partial fulfilment
of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE
at
The British University in Dubai**

July 2023



Sentiment Analysis in Social Media:
A Deep Learning Framework for Analyzing Public Sentiment toward
Policies during Natural Disasters

تحليل المشاعر والعواطف في وسائل التواصل الاجتماعي:
منهج متكامل لتحليل مشاعر العامة تجاه السياسات خلال الكوارث الطبيعية باستخدام طرق التعلم
العميق

by
MOHAMED ABDULLA FOOLAD

A thesis submitted to the Faculty of Engineering and IT
in partial fulfilment of the degree of
DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE
at
The British University in Dubai

July 2023

Thesis Supervisor: Professor Sherief Abdallah

Approved for award:

Name
Designation

Name
Designation

Name
Designation

Name
Designation

Date:_____

Declaration

I warrant that the content of this research is the direct result of my own work and that any use made in it of published or unpublished copyright material falls within limits permitted by international copyright conventions.

I understand that a copy of my research will be deposited in the University Library for permanent retention.

I hereby agree that the material mentioned above, for which I am the author and copyright holder, may be copied and distributed by The British University in Dubai for the purposes of research, private study, or education and that The British University in Dubai may recover from purchasers the costs incurred in such copying and distribution, where appropriate.

I understand that The British University in Dubai may make a digital copy available in the institutional repository.

I understand that I may apply to the University to retain the right to withhold or to restrict access to my thesis for a period which shall not normally exceed four calendar years from the congregation at which the degree is conferred, the length of the period to be specified in the application, together with the precise reasons for making that application.



Signature of the student

Copyright and Information to users

The author whose copyright is declared on the title page of the work has granted to the British University in Dubai the right to lend his/her research work to users of its library and to make partial or single copies for educational and research use.

The author has also granted permission to the University to keep or make a digital copy for similar use and for the purpose of preservation of the work digitally.

Multiple copying of this work for scholarly purposes may be granted by either the author, the Registrar, or the Dean only.

Copying for financial gain shall only be allowed with the author's express permission.

Any use of this work in whole or in part shall respect the moral rights of the author to be acknowledged and to reflect in good faith and without detriment to the meaning of the content and the original authorship.

Abstract

This study aims to produce and present a robust, reliable, and accurate framework to measure and analyze public sentiments for decision-makers and policy legislators while managing and administrating the spread of a pandemic or a natural disaster in general. The study utilizes COVID-19 as a case study to evaluate the proposed framework and test relevant hypotheses. The study used word embedding as a feature vector and then tested the accuracy of different models such as LSTM, Transformers, Logistic regression, Random Forest Classifier, KNN, and Multinomial Naïve Bayes. This work used a deep learning model –LSTM model on the collected tweets to classify the emotions and reactions. The models applied in the classification of COVID-19 global data performed differently with regard to accuracy and precision. Transformers emerged as the most accurate model, with a precision score of (90.18%) and an accuracy of (90.17%). LSTM was also superior with regard to accuracy, with an (88.45%) precision score and an accuracy of (88.44%). The Random Forest classifier performed fairly, with a (52.43%) precision score and (51.26%). The other models performed poorly in predicting sentiments; for instance, Multinomial Naïve Bayes recorded (43.54%) precision and (41.23%) prediction accuracy, the logistic regression model recorded a (43.54%) precision and (41.23%) prediction accuracy, and the KNN model had (41.23%) precision and (41.11%) prediction accuracy. Overall, LSTM was established as the most suitable model for the selected dataset since it could fit the dataset and be generalized to the current and new datasets, unlike transformers.

Keywords: Sentiment Analysis, Opinion Mining, Emojis, Social Media, Twitter, LSTM, Transformers, Random Forest, Logistic Regression, K-Nearest Neighbor, Naïve Bayes Classifiers, COVID-19.

الخلاصة

الهدف الرئيسي لهذه الدراسة هو إنتاج وتقديم منهج متكامل ودقيق لقياس وتحليل المشاعر العامة لصناع القرار ومشري السياسات أثناء إدارة انتشار جائحة أو كارثة طبيعية بشكل عام. تستخدم الدراسة جائحة كوفيد-19 كمثال فعلي لتقييم المنهج المقترح. استخدمت الدراسة تضمين الكلمات (Word Embedding) لتحديد واستخراج الميزات، واختبرت دقة نماذج مختلفة مثل نموذج الذاكرة طويلة المدى (LSTM) ونموذج المحولات (Transformers) والانحدار اللوجستي (Logistic Regression) و (Random Forest) و (KNN) وتصنيف (Multinomial Naïve Bayes). استخدمت الدراسة نموذج التعلم العميق (LSTM) على التغريدات التي تم جمعها لتصنيف المشاعر والردود. أداء النماذج المستخدمة في تصنيف بيانات كوفيد-19 العالمية تباين من حيث الدقة. ظهرت نماذج (Transformers) كالنموذج الأكثر دقة بنتيجة (90.18%) و (90.17%). كما كان نموذج (LSTM) متفوقاً من حيث الدقة بدرجة دقة قدرها (88.45%) و (88.44%). أداء (Random Forest) كان مقبولاً بدرجة دقة قدرها (52.43%) و (51.26%). أما النماذج الأخرى فقد قدمت أداءً سيئاً في توقع المشاعر؛ على سبيل المثال، سجل نموذج (Multinomial Naïve Bayes) درجة دقة قدرها (43.54%) و (41.23%)، وسجل نموذج (Logistic Regression) دقة قدرها (43.54%) و (41.23%)، وكان لدى نموذج (KNN) درجة دقة قدرها (41.23%) و (41.11%). بشكل عام، تم تأكيد أن نموذج (LSTM) هو النموذج الأكثر مناسبة لأية مجموعة من البيانات حيث كان قادراً على تحقيق نتائج دقيقة للغاية مع مختلف البيانات ويمكن تعميمه على أية بيانات حالية أو جديدة، بخلاف نموذج (Transformers).

Dedication

I dedicate this work, all my work, and all my life, with every aspect of it and minute detail, with the deepest spot in my heart and every single conscious emotion and subconscious feeling to the very one single human being who has been near me, besides me, by me, and with me throughout my entire life, what passed and what is to come. The one individual who is not my best friend, brother, soul mate, or family. A word is yet to be invented to define our relationship. To me, the word will simply always be his name, Zubair.

Without Zubair, I would have been nothing, and only by him, I become everything, only because he truly is everyone and everything. Without Zubair, I won't continue, I won't achieve, I won't succeed, I won't accomplish, I won't have interest, I won't have passion, I won't have motivation, I won't experience joy, I won't feel happiness, I won't understand life, and most importantly, I won't exist.

It is only this relationship that defines me, flourishes me, improves me, thrives me, and perfects me. It is only this relationship that wakes me up in the morning, keeps me going, makes me happy, makes me appreciate life, makes me thankful for everything, and never wants anything more.

It is such a unique phenomenon that only a very tiny percentage of the entire human race, if any, has or will one day experience. Whatever it is, I'm certain that I don't want anything but it, I will do absolutely anything it takes to keep it, I will remain forever faithful to it, and beyond doubt, I won't survive a single second without it.

Acknowledgments

First and foremost, Alhamdulillah, for granting me the success to accomplish what I have accomplished. Alhamdulillah, with whose grace good deeds are accomplished, and by whose grace goodness and blessings descend, and by whose grace goals and objectives are achieved. All praise be to ALLAH for all blessings, praise befitting His Majesty, and the greatness of His power. All praise be to ALLAH until praise reaches its end.

Alhamdulillah, for granting me a remarkable Supervisor with outstanding knowledge and endless support. Alhamdulillah, for awarding me a loving family who's without nothing would be possible. Alhamdulillah for surrounding me with caring friends who are eager to see me succeed and flourish only for the sake of success. Alhamdulillah, for granting me thoughtful co-workers and colleagues whose unconditional support I will never forget.

I ask Almighty ALLAH for ultimate success in what is to come.

Table of Contents

Chapter 1: Introduction	1
1.1 Social Media & Twitter	5
1.2 Emoji-Based Sentiment Analysis on Twitter.....	6
1.3 Emoji-Based Sentiment Analysis Using Deep Learning.....	9
1.4 Sentiment Analysis & Opinion Mining	11
1.5 COVID-19 Precautionary Lockdowns and Public Sentiments.....	12
1.6 The Undesired Impacts of the Precautionary Policies	21
1.7 Research Aims & Significance	22
1.8 Problem Statement and Gap in literature	25
1.8.1 Problem Statement.....	25
1.8.2 Gap in Literature.....	26
1.9 Thesis Organization	26
2 Chapter 2: Literature Review	28
2.1 Introduction.....	28
2.2 Theoretical Framework.....	29
2.2.1 The Rhetorical Structure Theory (RST)	29
2.2.2 The Network between RST and LSTM	30
2.3 Societal Impact of the COVID-19 Pandemic	31
2.4 The Psychological and Emotional Effects of COVID-19 Lockdowns	34
2.5 Public Sentiments, Opinions, and Emergencies	36
2.6 Precautionary Lockdowns on Public Sentiments and Emotions	39
2.7 Sentiment Analysis and How it is Used.....	41
2.8 Related Works.....	43
2.9 Impact of Sentiment Analysis among Various Societal Groups.....	44
2.10 Varied Impacts of Covid19 Containment Measures.....	47
2.10.1 Impact on Children and Teens	55
2.10.2 Workers.....	56
2.10.3 Adults.....	56
2.10.4 Elderly.....	57
2.10.5 Ethnic Minority Groups	57
2.11 Emoticons and Emojis in Social Media.....	58
2.11.1 Emoticons	58

2.11.2	Emojis	60
2.11.3	Emojis/Emoticons and Sentiment.....	61
2.11.4	Potentially Promising Techniques to Satisfy Hidden Annotation ..	70
2.12	Applications of Sentiment Analysis in Different Contexts	72
2.12.1	Sentiment Detection Approaches.....	72
2.12.2	Application Context of Sentiment Analysis	80
2.12.3	Sentiment and Sentiment Analysis during Crises.....	82
2.13	Techniques to Successfully Identify Emojis' Hidden Annotations.....	90
2.13.1	Manual Annotation	91
2.13.2	Automatic Annotation	91
2.14	Machine Learning Models for Sentiment Analysis	92
2.14.1	Rule- or Lexicon-Based Strategy.....	95
2.14.2	Methods based on Automation or Machine Learning	96
2.15	Summary	99
3	Chapter 3: Research Methodology	102
3.1	Introduction.....	102
3.1.1	Research Approach & Justification	103
3.1.2	Data Acquisition & Initialization.....	104
3.1.3	Model Training & Algorithm Development.....	109
3.1.4	Validation & Evaluation	111
3.2	Ethical Considerations	113
4	Chapter 4: Experimental Results and Analysis.....	115
4.1	Data Acquisition and Initialization.....	115
4.1.1	Data Collection	115
4.1.2	Dataset	116
4.1.3	Experimental Setup.....	118
4.1.4	Data Pre-Processing.....	121
4.1.5	Feature Extraction.....	130
4.1.6	Summary of Methods of Feature Extraction.....	132
4.1.7	Challenges of Data Collection in Sentiment Analysis and How to Solve Them	143
4.2	Model Training & Algorithm Development.....	145
4.2.1	Algorithm Development	145
4.2.2	Deep Learning Neural Networks	154

4.2.3	Sentiment Classification	169
4.2.4	Polarity Detection	173
4.3	Validation & Evaluation	174
4.3.1	Data Splitting and Data Labelling	174
4.3.2	Accuracy Benchmarking.....	176
4.3.3	Adapting to the Use of Semantic Lexicons	178
4.3.4	Supervised training	179
4.3.5	Model Evaluation.....	179
4.3.6	Users' Reactions During the Pandemic	184
4.3.7	Techniques for Precise Outcomes.....	194
4.3.8	Approaches Used in Benchmarking	196
4.4	Chapter Summary	202
5	Chapter 5: Discussion and Evaluation	204
5.1	Chapter Structure	204
5.2	Overview.....	205
5.2.1	Applications of Sentiment Analysis in Different Contexts	212
5.3	Q1 and 2: Public Sentiments Patterns During Natural Disasters.....	216
5.4	Q3: Reliability of Emojis.....	219
5.4.1	Sentiment Analysis on Social Media	233
5.4.2	Efficacy of Emoji Sentiment Analysis and Prediction	245
5.5	Q4: Promising Technique to Identify Emojis' Annotations.....	247
5.5.1	Unsupervised	248
5.5.2	Supervised ML.....	254
5.5.3	Convolutional Neural Network (CNN).....	256
5.5.4	ULMFiT Sentiment Analysis	257
5.5.5	Technique for Identifying Hidden Messages.....	263
5.5.6	Technique to Identify Emojis' Hidden Meaning	267
5.6	Q5: Most Robust Approaches to Achieve Precise Outcomes.....	270
5.7	Q6: Performance Comparison with the Start-of-the-Art Benchmarks	279
5.8	Challenges of Sentiment Analysis in Natural Disasters	286
6	Chapter 6: Summary, Conclusion, and Recommendations	287
6.1	Summary.....	287
6.2	Conclusion	296

6.3 Recommendations.....	298
6.4 Limitations and Suggestions for Future Research	302
6.4.1 Limitations	304
References.....	306
Appendices.....	358

List of Tables

Table 1: Instances of How Emoticons are Used to Relay Sentiments (Hogenboom et al., 2013)	62
Table 2: COVID-19 Twitter Dataset Sample.....	118
Table 3: Replacement of Emojis by Polarity (Siddharth et al., 2019)	123
Table 4: Cleaned Sentiments	125
Table 5: Emoji Translation	143
Table 6: The Bayes Theorem	152
Table 7: Input Features	159
Table 8: Nodes for LSTM Model	162
Table 9: Nodes in the Transformer Model.....	169
Table 10: Distribution of Sentiments across the Dataset	176
Table 11: Models Performance	181
Table 12: Spikes per Sentiment	186
Table 13: Model Performance after the Inclusion of Emojis in the Sentiment Analysis	270

List of Figures

Figure 1: Summary of the Multiple Perspectives of Sentimental Analysis Research (Yue et al., 2019)	12
Figure 2: The Quality Assessment Procedure.....	29
Figure 3: Relation Structure of RST Theory (Fu et al., 2016).....	30
Figure 4: The RST-LSTM Model (Fu et al., 2016)	31
Figure 5: The Research Aim & Methodology	103
Figure 6: The Research Paradigm (Angiani et al., 2016)	104
Figure 7: Most Popular COVID-19 Related Words within the Corpus.....	126
Figure 8: KNN Classifier	147
Figure 9: Random Forest Classifier	149
Figure 10: Logistic Regression Classifier.....	151
Figure 11: Multinomial Naive Bayes.....	154
Figure 12: LSTM Training Accuracy vs. Validation Accuracy	160
Figure 13: LSTM Training Loss vs. Validation Loss	161
Figure 14: LSTM Model.....	163
Figure 15: The Transformer Model Architecture	165
Figure 16: The Transformer Training Accuracy vs. Validation Accuracy	166
Figure 17: The Transformer Training Loss vs. Validation Loss	167
Figure 18: Sentiment Distribution	172
Figure 19: Sentiment Distribution of Three Class Polarity	173
Figure 20: Models Prediction Accuracy	178
Figure 21: Tweet Classification	184
Figure 22: Variation of Sentiments in Relation to Keywords.....	187
Figure 23: Spatial Distribution.....	188

Figure 24: Common Keywords in India (top left), China (top right), the US (bottom left), and Italy (bottom right).....	189
Figure 25: Variation of Sentiments Across a 6-Month Period in China (Chen et al., 2020)..	191
Figure 26: Variation of Sentiments across a 6-Month Period in the US (Shekerdemian et al., 2020)	192
Figure 27: Variation of Sentiments across a 6-Month Period in India (Saigal et al., 2020)..	193
Figure 28: Sentiment-Related Keywords across Different Countries.....	194
Figure 29: Comparison of Emojis before and After the Pandemic (Liu et al., 2022).....	246
Figure 30: Classification of Emojis based on their Sentiments (Liu et al., 2021)	269

List of Abbreviations

ASR	Automatic Speech Recognition
Bi-LSTM	Bidirectional Long Short-Term Memory Recurrent Neural Networks
BOW	Bag of Words
BRT	Bayesian Rose Tree
CNN	Convolutional Neural Networks
DL	Deep learning
EI	Emotional Intelligence
ELSTM	Emotion-Enhanced LSTM
GPUs	Graphical Processing Units
KNN	K-nearest neighbor
LDA	Latent Dirichlet Allocation
LMC	Liberal Market Economy
LSTM	Long Short-Term Memory Networks
ML	Machine Learning
NES	Named Entity Segmentation
NEC	Named Entity Classification
NER	Named Entity Recognition
NLP	Natural language processing
OOV	Out of Vocabulary
POS	Parts-of-Speech Tagging
RAE	Recursive Auto-Encoders
RF	Random Forest
RNNs	Regularized Neural Networks
RST	The Rhetorical Structure Theory

SA	Sentiment Analysis
SVMs	Support Vector Machines
TPUs	Tensor Processing Units
ULMFiT	Universal Language Model Fine-Tuning
ASR	Automatic speech recognition

Chapter 1: Introduction

Linguistically, sentiment refers to feelings, emotions, or sometimes opinions. It is a highly subjective phenomenon and an emotional experience that an individual develops toward incidents, thoughts, or ideas (VandenBos, 2007). The analysis of sentiments is also identified as Opinion Mining simply because it focuses on areas studying an individual's opinion, attitude, and emotions towards a particular entity. In numerous studies, the two are used interchangeably to denote the same area of research, as will be thoroughly discussed later in the study. However, few researchers argued that the two terms present slightly different concepts. For instance, Cambira et al. (2013) claimed that sentiment analysis and opinion mining focus on detecting emotion and polarity, respectively. Medhat et al. (2014) stated that opinion mining investigated opinion, an individual's opinion on a certain entity, while sentiment analysis extracted and assessed the sentiments hidden in the message. With the increased spread of social media networks and platforms, their roles continue to develop and grow exponentially beyond their initial purpose. Today, major industries heavily employ social media to exploit its enormous data. Sentiment Analysis using social media has been extensively investigated in different fields, namely, text, web mining, and NLP (Liu, 2012). The survey by Medhat et al. (2014) presented a broader overview of the existing workings. In the work, the researchers demonstrate a detailed categorization of the numerous studies that illustrate the recent trends in the currently investigated field.

Nonetheless, there are still significant potentials within the massive and highly powerful social media communities and platforms that very few researchers in the field have addressed, including Twitter, microblogs, and blogs, some of which will be discussed in the later chapters and sections (Kouloumpis et al., 2011). However, due to the colloquial and informal linguistic style and the high frequency of slang and misspells, analysing sentiments on a social media platform such as Twitter becomes difficult from the conventional sentence

level (Li et al., 2016). According to many studies highlighted in Chapter 2, it remains an open question as to the extent to which the techniques and used features on more well-formed information will get transferred to the microblogging domain. Additionally, what are the techniques and features that will increase accuracy and speed? In the past years, sentiment analysis of the Twitter platform has attracted attention in the field, with an increasing number of researchers attempting to suggest a novel Twitter-based sentiment analysis technique to assess public sentiment on textual information. Pak and Paroubek (2010) exhibited a methodology that automatically collected a corpus to train the different sentiment classifiers.

Koto and Adriani (2015) conducted a comparative study to establish the best features for performing a sentiment analysis on Twitter. Similarly, Mansour et al. (2015) presented a combination of two Maximum classifiers for the Twitter platform. In the research, they exhausted numerous extracted features in experimenting to discover a compact feature of subsets with the capability of reducing the computational complexity without having to scarify the classification process or, at the same time, accuracy. To categorize the sentiments in the tweets, Barbosa and Feng (2010) constructed the classifiers with various features, including tagging of the parts of speech, polarity lexicons, and special and lexical microblogging features such as hashtags, character repetition, emoticons, and punctuations. Kouloumpis et al. (2011) equally explored the effectiveness of the existing lexical resources and various microblogging features and abbreviations for sentiment analysis of the Twitter platform.

Most researchers in the field have exhibited devotion to the English text since other languages have been discovered to attract less attention. Due to the multilingual nature of data recovered from social media platforms, the analysis of sentiments focused on a single language is less likely to capture the overall sentiment in the content. Recently, sentiment analysis based on multiple languages has also been suggested to deal with the issue.

Dashtipour et al. (2016) review of the multilingual methods of sentiment analysis points out the lack of lexical resources as the main challenge in multilingual sentiment analysis.

Similarly, Lo et al. (2017) provided an overview of the different present tools and approaches useful in multilingual sentiment analysis and established challenges in the same research direction. Since tools used in sentiment analysis specifically possess useful applications in reputation and marketing management (Heerschop et al., 2011), these tools are often analysed on gathered reviews that typically contain the opinions of individuals often expressed in a natural language and with a related score (numeric) that quantifies judgment.

In a similar light, a broadly applied corpus for the evaluation of analysis techniques uses a collection of 2,000 English movie reviews, all categorized for sentiments (Pang and Lee, 2005). Among the popular approaches, a binary representation of the featured text appears, indicating the presence of particular words. According to the researchers, the method has proven effective, resulting in an 87.2% accuracy. Researchers later highlighted the various vectors representing texts, including representation with more features representing semantic distinctions between featured words or the representation of vectors with weights based on TF-IDF for word features (Paltoglou and Thelwall, 2010). Typically, these techniques have been discovered to result in high accuracy based on the data from more than 90% of movie reviews. The alternative, a lexicon-based approach, shows a lower accuracy based on the dataset from the movie review; however, there is a tendency for it to become more robust across various domains (Taboada et al., 2008). A dictionary can also generalize a lexicon-based approach relatively easily than other languages.

A simplified lexicon-based framework of the process has been discovered to show an accuracy of up to 59.5% from the full movie review (Heerschop et al., 2011). Taboada et al. (2011) claimed that an even more complex lexicon-based approach to the process has an average accuracy of 68.0% on 1900 documents from the datasets of a movie. An even deeper

linguistic analysis that focuses on distinguishing the rhetorical roles of the segments of a text has been proven to perform well (Heerschop et al., 2013) comparatively. When investigated on 1000 documents from the datasets from the movie, the approach resulted in a 72.0% accuracy; this improvement rated at 4.5% without accounting for the structural aspects of the different contents. Despite the lexicon-based sentiment analysis approaches exploring a more promising direction of integrating both the semantic and structural content aspects (Hogenboom et al., 2010), the researchers failed to acquire data from cues considered potentially significant for sentiments in the emojis and emoticons generated by the user. However, it is essential to note that emoticons have been investigated to a limited extent, specifically for automated data categorization. For example, a crude difference between numerous negative and positive emoticons has been utilized to automatically generate datasets with negative and positive samples of texts in natural language to train and test the techniques on polarity classification (Read, 2005).

The above findings imply that emoticons' information on polarity relayed is independent of domain and topic. The results have also been applied in the later works of researchers to automatically create a set of negative and positive tweets (Pak and Paroubek, 2010). Recently, a relatively small class of emoticons has been applied to classify polarity (Thelwall et al., 2011). However, the findings of the current research work fail to show that treating emoticons as though they are everyday words that carry sentiment yields significant growth over ignoring emoticons that once classified the polarity of natural language. As long as an emoticon remains slightly less significant for sentiments in the content generated by users, the success behind acquiring information from an emoticon lies in comprehending how they are associated with the overall sentiment of a text. To the best of each researcher's knowledge, current studies have failed to establish how an emoticon impacts a text's sentiments in NL or investigate how the phenomenon can be assessed in a lexicon-based

approach. To enable researchers to focus on this, one must first establish how emoticons are useful in any communication mediated by a computer.

1.1 Social Media & Twitter

Social media, a concept that promotes public interactions and engagements through online platforms and technologies (Edosomwan et al., 2011), has existed since the early 2000s (Efimova et al., 2007). However, at that time, social platforms were exclusive to academic institutions and large enterprises, serving as internal communication channels or traditional knowledge-exchanging forums (Leonardi, 2013). During the mid-2000s, social platforms started to massively blast into people's daily activities and personal lives until they became the absolute mainstream channels for everyday communications. Since then, social media has taken various forms and models, ranging from textual interaction to audio and video sharing (Avram, 2005). Nevertheless, the birth of genuinely legitimate artificial intelligence applications has uncovered the enormous potential of social media platforms as an essential source of big data. From personal biographies, daily trivia, official statistics, verified facts, or the latest news, social media is possibly the most substantial data source on the entire planet. Public figures, distinguished officials, respected academics, content creators, celebrities from every field, and other people with various interests own verified profiles and share a gigantic amount of content daily (Abdulla, 2020a).

On the other hand, Twitter is a prominent social media platform that adopts the concept of microblogging (Kwak et al., 2010). As the term implies, microblogging is a special form of traditional blogging that enables users to share micro or small content similar to SMS (Honey et al., 2009). Founded in 2006, Twitter allows users to post microblogs, referred to as "Tweets," using its diverse web services (Arceneaux et al., 2010). Furthermore, Twitter users can share various forms of media content, such as photos and videos, in addition to various continuously added features and options such as voice notes, fleets (an ephemeral form of

content sharing), and Periscope (a live streaming service). However, as a microblogging service, all content has limited capacities. For instance, tweets are currently limited to 280 characters (originally 140), whereas videos cannot exceed 140 seconds (Gligorić et al., 2018). In addition, Twitter allows users to interact with each other in various ways, such as likes, retweets, quote tweets, and comments.

Hashtag, a means to group tweets from several users using relevant keywords, is a powerful yet straightforward tagging tool first adopted by Twitter in 2007 and swiftly appeared in every other social media platform, such as Facebook and Instagram (Chang, 2010). Users can type the hash symbol (#) before any desired key phrase in a post to categorize it and display it with every tweet or post that includes the hashtag. Hashtags are an extremely effective tool for easily accessing and searching for relevant information, data, or subjects. As of January 2021, Twitter is ranked 16th among the most popular social networks worldwide, possessing over 300 million monthly active users (Tankovska, 2021). Nevertheless, Twitter remains one of the most favourable data sources for data scientists and statistical analysts due to its humongous users' diversity, content simplicity, and the ability to effortlessly access and collect raw data through Twitter's official API (Arias et al., 2014).

1.2 Emoji-Based Sentiment Analysis on Twitter

Twitter is a tremendous source of data for sentiment analysis or opinion mining. It contains numerous extended discussions, distinct opinions, and extensive emotions (Wolny, 2016). According to many studies and observations, emotion tokens known as emojis are widely used by Twitter users to convey emotions regardless of the user's language, thus presenting a potent and compelling universal signal for opinion mining or sentiment analysis. Wolny attempted to investigate sentiment analysis beyond the traditional text-based approach to surpass the barriers of Multilanguage in social media platforms. Using symbol analysis, the author discovered that the accuracy in recognizing various types and classes of

emotions could be significantly enhanced. The existing binary emotions categorization can be notably extended using a multi-way sentiments classification. Tweets have been classified using an unsupervised machine learning method; however, only the most basic and commonly used emojis, which accounted for around 90% of the total 96 million tweets, were successfully identified and interpreted. Therefore, integrating natural language processing approaches and emoji analysis is probably the most promising algorithm for analyzing emotions on Twitter (Wolny, 2016).

Machine learning approaches have recently gained immense popularity in extracting meaningful information from social media emojis apart from text-based interactions. Hallsmar and Palm (2016) investigated the practicality of emoji-based training sets in sentiment classification. The study confined the sentiment categories into four groups, and emojis were restricted to iOS icons. In addition, a limited number of emojis were chosen to represent each category to achieve a decent uniqueness. Although emojis are distinctly language-independent, the research only concentrated on analyzing tweets written in English since the performance evaluation will be manually executed. By applying Multinomial Naïve Bayes Classifiers on training sets of 4,000 to 400,000 tweets, emojis' accuracy surpassed hashtag-based interpretation and even some text-based annotations that utilized traditional natural language processing techniques (Hallsmar et al., 2016). Nevertheless, careful selection of sentiment classes and emoji descriptions is crucial in maximizing the accuracy and performance and preventing frequent and challenging overlaps.

Social media sentiment analysis is an undeniably powerful tool for measuring public feelings and opinions, thus providing worldwide authorities with valuable political insights. Similar to the previous study, Li and colleagues (2016) employed a Multinomial Naïve Bayes classifier to develop a sentiment categorization system. The classifier uses expressive emojis in tweets to analyze public sentiments and political perspectives toward the 2016 U.S.

presidential election. Using Twitter's API and certain keywords, the authors accumulated two months of tweets covering the most active election period. The congregation process yielded a total of 79 million tweets, with only around 2 million tweets containing emojis.

In addition, N-grams were exclusively used as the main features, which resulted in an abundance of noisy, ineffective, and redundant features. Thus, excessive pre-processing and feature assortment techniques have been carried out to eliminate undesirable features, improve computation performance, and enhance the overall quality (Li et al., 2016). The model accomplished promising results in classifying tweets and interactions as positive, negative, or neutral. However, the sentiment score for the infrequently used emojis was undetermined and entailed further investigation. Moreover, sarcasm detection has always been a challenge in text-based sentiment analysis and is definitely an inevitable dispute in the field of emoji analysis, especially when analyzing political interactions that are known to be highly sarcastic.

Sentiment analysis on wide-scale data obtained from various social platforms is not just a privilege anymore but a necessity to closely examine and capture real-life activities and public reactions toward diverse incidents or commercial products (Chen et al., 2018). Unlike text-based sentiment analysis or opinion mining, emoji-based analysis is still considered original and highly underestimated. This study proposes a unique Twitter sentiment analysis scheme with an exclusive emphasis on emojis. First of all, bi-sense emoji embeddings were deployed by assigning each emoji to two different embedding vectors. The first token is for particular cases of positive sentimental frameworks, whereas the other token is for negative sentiment environments. Afterward, an attention-based long short-term memory network sentiment classifier was trained on the previously generated embeddings (Chen et al., 2018). The research demonstrated that the integration of multi-vector embeddings with a multi-level attention-based LSTM could, in fact, achieve state-of-the-art performance in understanding

various sentiments and complicated semantics.

1.3 Emoji-Based Sentiment Analysis Using Deep Learning

Opinion mining has recently gained immense popularity in various sectors, like marketing and journalism (Karthik et al., 2018). It is widely utilized to study the public's approbation of a particular product or to measure people's reactions toward specific news. Karthik and colleagues (2018) have examined various social media platforms to perform opinion mining comparatively. Opinion mining has been limited to textual posts and wordy sentences. Researchers have commonly overlooked emojis, despite their ability to convey plenty of meanings with extremely straightforward expressions. The authors attempted to explore the interpretative differences between opinion mining using textual tweets and emojis separately within the same Twitter post. Using various Deep Learning techniques, the researchers discovered that many sarcastic tweets were overlooked and misinterpreted due to neglecting the emojis within the same tweets (Karthik et al., 2018). Emojis are distinguishingly superior to other forms of expression due to the ability of a single emoji to imitate multiple emotions and reactions depending on the context in which it is used. Yet, this could be the most challenging aspect when it comes to training a Neural Network model since a single emoji can be deciphered into multiple intended expressions. Zhao and Zeng (2017) have carefully investigated this presumable issue by attempting to link and match textual tweets to their consequent expressive ideograms. Utilizing Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN) and Convolutional Neural Networks (CNN) with 1.5 million Twitter tweets as their training set, they achieved decent precision, especially with the CNN model.

The authors admitted that tweets containing texts and emojis could be quite noisy and challenging. In addition, people are usually unstructured with their emojis, particularly in combining several emojis to convey or imply multiple meanings (Zhao et al., 2017). These

kinds of tweets or social media posts generally represent massive noise within the dataset and require extensive cleaning. Finally, the authors probably encountered the most pressing challenge in emojis, thus, their ability to have different interpretations. They affirmed the need for further future investigation to achieve more precise predictions. Human emotions are another substantially challenging aspect in the field of opinion mining or sentiment analysis using social media content. Analyzing such expressions and detecting feelings or thoughts on a specific matter can result in many essential life applications such as healthcare, education, commerce, and politics (Jabreel et al., 2019). A single tweet on Twitter or a Facebook post can contain multiple emotions concurrently. Unlike most single-classification studies, Jabreel and Moreno (2019) aimed to explore and discern social media content that involves mixed feelings using a novel approach that converts Twitter content into binary representations and categorizations. Using a Deep Learning approach, the binary transformation is then analyzed. The authors have succeeded in achieving remarkable accuracy. Nevertheless, the model failed to identify implications and word manipulation, such as sarcasm.

Emojis have progressed significantly during the past years. Notably, today, almost every emoji resembles every expression, emotion, and interaction (Coman et al., 2018). As a result, social media content is becoming more complex to comprehend. The study by Coman et al. (2018) focused on developing a Machine Learning model that can analyze a textual tweet and consequently associate it with the most accurate emoji based on the context and the overall sentiment. Using Bidirectional Long Short-Term Memory Recurrent Neural Networks (Bi-LSTM) and Convolutional Neural Networks (CNN), the authors achieved a reasonable prediction efficiency, especially with the CNN model. The study has proved that simple ML models have the potential to produce equivalent precision as those more complex ones with fine-tuning hyper-parameters (Coman et al., 2018).

1.4 Sentiment Analysis & Opinion Mining

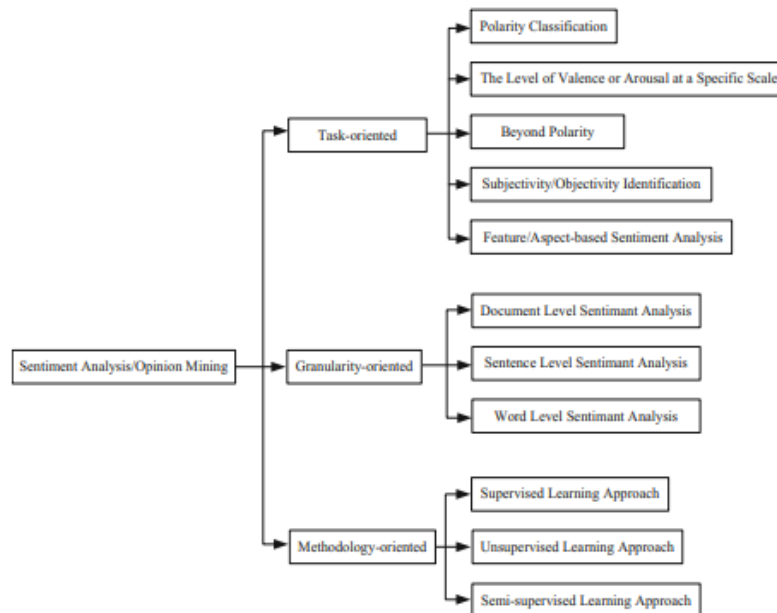
Sentiment analysis or opinion mining is a highly trending research topic that utilizes artificial intelligence techniques such as deep learning (DL) and natural language processing (NLP) to analyze and evaluate individuals' opinions, emotions, sentiments, and attitudes toward a particular product, incident, or topic (Liu, 2012). The concept first appeared in the early 2000s (Nasukawa et al., 2003); nevertheless, it started attracting researchers' attention in computer science and artificial intelligence only recently. This was mainly due to the lack of data required to achieve such sophisticated evaluations and produce reliable outcomes (Abdulla et al., 2019). However, the emergence of social media made the theory behind opinion mining possible. Sentiment analysis has been addressed mainly as a natural language processing (NLP) task due to the textual nature of most social media content (Agarwal et al., 2011). However, text mining has been significantly enhanced since the evolution of emojis (Shiha et al., 2017). Yet, emojis have been considerably overlooked by researchers and data scientists (Abdulla et al., 2020). Many attempts in the literature have found that neglecting emojis in a social media post that contains both text and emojis while performing a sentiment analysis results in a complete misinterpretation (Karthik et al., 2018). Emojis contain hidden emotions and sentiments that are very challenging to express using pure text. Performing a text-only analysis and discarding expressive images will eventually result in a sarcastic tweet or post being taken seriously.

Emojis' potential and abilities to convey mixed feelings and sophisticated emotions could be the most challenging aspect when it comes to training a Neural Network model since a single emoji can be deciphered into multiple intended expressions (Zhao et al., 2017). In other words, tweets containing texts and emojis could be quite noisy and challenging. Additionally, people are usually unstructured with their emojis, particularly in combining several emojis to convey or imply multiple meanings (Abdulla, 2020b). Human emotions are

another substantially challenging aspect in the field of opinion mining or sentiment analysis using social media content.

Figure 1

Summary of the Multiple Perspectives of Sentimental Analysis Research (Yue et al., 2019)



1.5 COVID-19 Precautionary Lockdowns and Public Sentiments

Nshakira-Rukundo & Whitehead (2021) point out that even though COVID-19 preventive measures such as lockdowns were necessary for curtailing the spread of the virus, they had significant socioeconomic implications ranging from loss of income to mental health issues. The balance between the supposed benefits of the lockdown measures and the resulting negative effects shaped the public's perception of these measures and willingness to adhere to them. This view is further supported by Nguyen et al. (2022), who argue that the public's intention to follow the established preventive measures is determined by the population's perception of the disease. In essence, the risk of infection perception positively correlates with the public's willingness to adhere to the set requirement. Individuals with a high perceived risk of contracting COVID-19 demonstrate attitudes characterized by anxiety, insecurity, and nervousness, which in turn increases behavioral control. As the Pandemic

dragged on, the benefits of the lockdown measures continued to be pitched against their negative effects, which meant variations in risk perception and prevention mechanisms (Nshakira-Rukundo & Whitehead, 2021).

Since curbing the spread of the virus depends on a series of personal choices and decisions, understanding the public's sentiment is vital in modifying and implementing pandemic preventive measures (Tong et al., 2020). According to Brown et al. (2021), communication of preventive measures such as lockdowns to the public plays a significant role in shaping perception. For this reason, the researchers suggest that the messages of lockdown measures highlight the virus's threat to life and demonstrate the measures' efficacy. Brown and their colleagues argue that highlighting the virus's threat to life could increase adherence. However, this tactic's downside is the reduction of health-promoting behavior. Any measure's effectiveness depends on the individual's attitude towards the requirement, which is shaped by factors such as the perceived risk of contracting and spreading the virus. Siew et al. (2022) explain that the public's sentiments are shaped by several factors, including the masses' perceptions that the adopted pandemic-related measures are effective, timely, sufficient, and safe.

Additionally, the public's trust in the government, demonstrated by the accessibility of the related information related to the pandemic, contributes to shaping the public's perception, subsequently determining individuals' level of engagement in protective behavior. Nguyen et al. (2022) and Aljameel et al. (2021) argue that the effectiveness of adopted pandemic measures depends on the targeted population's awareness of the risk of the virus and attitude towards the undertaken precautionary procedure by analyzing social media posts on platforms like Twitter, the relevant authority can gauge the overall awareness levels and determine the best alternative. Using three techniques: Support Vector Machine, k-nearest neighbors, and Naïve Bayes, Aljameel et al. (2021) classified tweets from four regions of

Saudi Arabia into negative, positive, and neutral categories. Negative, neutral, and positive attitudes were interested in carelessness, mixed, and high awareness of the adopted precautionary measures.

A similar approach has been used by Jalil et al. (2022), who analyzed the various emotions of the public towards the COVID-19 virus and the adopted containment measures through the classification and analysis of identified features on the collected tweet. Like Nguyen et al. (2022) and Aljameel et al. (2021), Jalil et al. (2022) hold the view that understanding the public's perception of the virus and undertaking precautionary measures provide insights that guide better handling of the Pandemic. Jalil et al. (2022) also categorized the tweets into three groups: positive, neutral, and negative. The response to COVID-19 became what was referred to as the new normal with lockdown measures that were characterized by staying at home and isolation time. For this reason, the Pandemic and the undertaken measures adopted many emotions and perspectives shared on multiple social media platforms, including Twitter. According to Jalil et al. (2022), public sentiment was largely negative once they were implemented, but after a month, the sentiment became increasingly positive, implying that most people had begun to favor the adopted lockdown measure. However, the researchers warn that even though social media is a crucial source of getting the public's perception, most data retrieved from social media are irrelevant to understanding the public's sentiment. This problem could be resolved using the COVIDSenti dataset, which would aid in revealing the public's perception of undertaken measures, which in turn serves as a predictor of the population's response to the established measures.

Agarwal et al. (2022) also analyze public sentiment towards lockdown measures from five worst-hit nations to understand the effects of the containment measure on the population's perceptions. The investigators concluded that the public's perception of the lockdown measures varied across the region. For instance, the researchers found that

Americans' sentiments ranged from neutral to negative, while Italy and India depicted mixed sentiments. Similar studies were carried out by Tsao et al. (2022), focusing on Ontario, Canada. According to the survey, the public's sentiment towards lockdown measures was slightly positive, but the daily sentiments ranged from neutral to positive. Lwin et al. (2022) add to this growing body of knowledge by providing a different perspective. Targeting five nations: South Korea, India, Indonesia, the US, and the UK, the researchers analyzed social media posts and concluded that even though all the selected nations demonstrated negative sentiments towards lockdown measures, positive ones rose simultaneously. For example, in the three Asian countries, the escalation of COVID-19 cases despite the lockdown measures resulted in a rise in negative sentiments. However, after a while, the researchers noticed that positive sentiments depicting hope and resilience rose simultaneously. The same results were recorded in the US and UK, where negative sentiments towards the containment measures were witnessed, while positive ones arose simultaneously.

Researchers interested in this study area have a common interpretation of their findings. For instance, Agarwal et al. (2022) explained that in the first month of implementing the lockdown measures, the public's perceptions were primarily negative or neutral. Still, as time progressed, they became more positive. This transition could be attributed to most people favouring the optimistic approach toward the implemented measures. However, Tsao et al. (2022) uphold that the change in attitudes and perceptions toward lockdown could be explained by the public's attempt to balance the risk of contracting the virus and the disruptions resulting from the restrictions. For instance, the public's sentiments towards COVID-19 lockdown requirements became overly optimistic during the second wave of the Pandemic, indicating increased awareness of the virus, which meant the population's willingness to abide by the set measures. Alternatively, the government's crisis communication approach could explain the nature of the public's sentiment toward COVID-

19 lockdown measures. For instance, Agarwal et al. (2022) found that public sentiments were predominantly negative at the beginning of lockdown measures in five nations: South Korea, Singapore, India, the US, and the UK. But positive sentiments rose simultaneously. The researchers attributed these findings to the nature of the government communication approach.

In essence, a concerted approach to routing communication prevented possible volatility over the evolution of the Pandemic. Mohd Hanafiah et al. (2021) describe a concerted communication strategy as coordinating all the involved parties to ensure alignment of the communicated message. In a time of crisis, the public's normalcy is disrupted, which means that people who are fearful and anxious will also be susceptible to fake news; therefore, communicating in a single voice allows for credibility and trust. Parajuli et al. (2020) support Agarwal et al. (2022) position by noting that communication is an integral part of crisis response as it sets the scene for the public's acceptance of the government's adopted health interventions at the start of the crisis. Concerted communication is essential at all phases of the Pandemic. This approach ensures that all government response actions align, which is crucial in establishing the public's trust in the adopted lockdown measure. This view is also highlighted by Brown et al. (2021), who explain that communication of preventive measures such as lockdowns to the public significantly shapes perception. Therefore, the investigators suggest that the messages of lockdown measures should highlight the virus's threat to life and demonstrate the measures' efficacy. Brown and their colleagues argue that highlighting the virus's threat to life could effectively increase overall adherence.

As mentioned, the public's sentiment towards COVID-19 lockdown measures varied across regions and throughout the Pandemic. A study focusing on China as the target population found that the Pandemic affected the overall sentiment of the public on all social media platforms (Tan et al., 2021). The strict measures that included lockdowns were found

to have induced a long-term negative perception even after the recovery period. These findings seem to counter the supposed positive correlation between the severity of COVID-19, as depicted by the number of cases, and the public's willingness to abide by the set rules specified by Nguyen et al. (2022). Tan et al., 2021 explain that the negative sentiment toward government containment measures in China is neither associated with the severity of the Pandemic nor the post-pandemic economic recession. Instead, the author attributes the prolonged negative sentiments towards the pandemic countermeasures to the Chinese government's failure to consider the downsides of the lockdown measures, such as the public's mental health. Even though studies focusing on other regions do not refute the findings of Tan et al. (2021), they offer varying perspectives. For instance, Meier et al. (2020) compared the public's perceptions of lockdown containment in the Netherlands, Italy, and Germany population cohorts. The researchers' findings justified Nguyen et al.'s. (2022) view that a positive correlation exists between the masses' intent to adhere to the set regulations and the perceived burden of the Pandemic. For example, the perceptions towards the lockdowns were favorable in Italy, which had the highest pandemic burden compared to the Netherlands and Germany. Additionally, Meier et al. (2020) agree with Tan et al., 2021 that a concerted communication approach is crucial during the crisis as the public is likely to adhere to the set regulations if they have high levels of awareness and vice versa. Jalil et al. (2022) also add that the unprecedented nature of the virus, as well as the measures undertaken to combat it, coupled with the popularity of social media, means that the public is susceptible to much fake news. For this reason, Tan et al. (2021) emphasize the need for scientists to find accurate, standardized viral detection measures and the effectiveness of social distance to minimize confusion that could undermine the credibility of the undertaken measures (Jalil et al., 2022).

According to Charquero-Ballester et al. (2021), determining the actual public perception of the lockdown measures could be challenging because the public's social media posts are reactions to multiple elements of the Pandemic, and lockdown is only one. Regardless, the study indicated that a good proportion of the tweets do not demonstrate negative or positive valence; instead, the tweets slightly tended towards a negative valence. The authors attributed their findings to increased misinformation concerning virus characteristics and reported cases. As such, the government should tackle this misinformation to increase the public's trust in the centrally-disseminated information, which would, in turn, justify the countermeasure undertaken. Agosto et al. (2021) studied public sentiment towards lockdown measures in six countries: the US, Brazil, Italy, Australia, South Africa, and the UK. They concluded that the population's online views captured the trend in the Pandemic's curve trajectory.

Furthermore, the authors demonstrated a positive correlation between public sentiments toward the undertaken measures and the community's disease burden, justifying the need to consider public sentiment in making decisions concerning crisis containment measures. Ahmed et al. (2020) note that analyzing users' views on social media offers insight into their emotional concerns about the lockdown. The researchers compared Americans' emotions concerning when the nation resorted to lockdown in March and the reopening decisions. They found that the sentiments towards lockdown were generally negative, with the dominant emotion being fear. However, the emotions shifted to positive when reopening was announced despite the rising number of positive cases; at this stage, the researchers concluded that the dominant emotion was less fear of the prevailing situation.

Tsao et al. (2022) and Dangi et al. (2022) point out that the public generates negative and positive sentiments about unprecedented decisions such as lockdowns. This finding could be attributed to the fact that the public's response to government decisions could be multifaceted, with people agreeing with certain aspects while disagreeing with others. For

instance, lockdown meant having to work at home, which was perceived as appealing by a good proportion of the population. At the same time, individuals were concerned about the downsides of the lockdown, such as widespread unemployment. Dangi et al. (2022) explain that the nature of generated social media content depends on the prevailing situation. The isolation caused by the lockdown meant that social media served as the only opportunity to air their emotion. As such, during the lockdown, people were fearful and turned to social media for reassurance and to offer others their support. Therefore, the generated posts ranged from neutral to positive.

However, the sentiments shifted to negative after the lockdown was lifted; Dangi et al. (2022) have linked these observations to the fact that people had to face multiple restrictions coupled with the lockdown's effect. Similar studies in Wuhan, China, revealed similar patterns, implying that the public sentiments about the lockdown were affected by different factors and never remained constant for an extended duration. Lu et al. (2021) focused on the rising number of cases between January and February 2020, which coincided with the Wuhan lockdown and Dr. Li Wenliang's death on 23rd January and 7th February, respectively. Therefore, the move to have Wuhan under lockdown was applauded, as indicated by the rise of positive sentiments, which further increased when reported cases began dropping. However, as the pandemic evolved, aggressive measures undertaken by the government received varied interpretations, with people becoming overly critical and pointing out the shortcomings in the measures undertaken. These results align with (Tan et al., 2021) that the strict measures undertaken by the Chinese government to contain the spread of the virus, including lockdowns, induced long-term negative perceptions even after the recovery period. These findings counter the supposed positive correlation between the severities of COVID-19, as depicted in other studies.

Suratnoaji et al. (2020) note that social media content analysis revealed that Indonesians held varying views about the lockdown to curtail the spread of the virus. The researchers concluded that the sentiments of 69.67, 14.8, and 17.5 percent of Indonesians on Twitter were neutral, positive, and negative, respectively. The Indonesian government enforced regional lockdowns rather than general ones. Regardless of this decision, the country continued to record high rates of infections and death daily, yet the public's sentiments are neither categorized as wholly positive nor negative. Throughout the study period, most Indonesians were in the non-categorized, that is, the neutral cohort, implying the uncertainty that most people were dealing with at the time.

Kausar et al. (2021) extended the research on public sentiment toward COVID-19 lockdown measures to include 11 nations. The study concluded that the Pandemic elicited similar emotions across the 11 nations included in the study. In essence, the researchers argued that support directed toward the government's lockdown decisions was largely supported by the fact that the option could curb infections and reduce the community's pandemic burden. Besides spending more time on social media, experts are convinced that the Pandemic elicited significant negative emotions, including fear and anxiety. For instance, Lee et al. (2022) concluded that lockdowns that were put in place to contain the spread of the virus meant that people lived in isolation for extended periods. Social media offered the only option for socialization. However, time spent on social media correlates positively with the danger of developing anxiety and depression. Davis et al. (2020) note that the public's sentiments following the Indian government's lockdown decision reflected negative emotions like fear and anxiety, among other mental disorders. Additionally, the impact of the Pandemic on the population's mental well-being was found to be the highest at the onset of the Pandemic. Still, it leveled off, indicating the development of resilience.

1.6 The Undesired Impacts of the Precautionary Policies

Natural disasters and pandemic outbreaks impose certain temporary or, in some cases, permanent lifestyle transformations. Such major modifications are undeniably necessary to sustain and preserve nations' existence, resources, and humanity. Governments, disaster management authorities, and disease control centers around the globe are compelled to deal with extremely challenging decisions and dilemmas when confronting such circumstances. Negative outcomes, hidden impacts, and undesirable side effects are almost inevitable and strongly tied to sudden alterations, restrictive measures, and policies. The concerned authorities have indubitably been struggling to find the optimal technique to analyse and predict the potential outcomes of a particular measure or decision. Artificial Intelligence, in general, has proven to be extraordinarily effective when it comes to addressing such challenging prediction problems. Nevertheless, unlike marketing and politics, the subject of natural disaster management has not been adequately investigated with respect to AI techniques, as will be shown later in the literature review section.

COVID-19 is an ideal and live example to measure the undesired side-effects of the precautionary policies imposed by governments and decision-makers worldwide to suppress the pandemic outbreak. Measures such as partial and complete lockdowns, mandatory masks, quarantines and isolations, border lockdowns, business shutdowns, and distance learning and working were globally adopted to confront the epidemic. These measures have undoubtedly and massively aided in surpassing and controlling the virus's outbreak. Nevertheless, such policies came with great costs and severely impacted many essential economic, societal, psychological, as well as other fundamental life aspects. For instance, a study in Italy examined the psychological, emotional, and social impacts on groups most affected by the pandemic preventive measures, such as students, frontline workers, and children (Saladino et al., 2020). Psychological disorders such as anxiety, stress, and depression have been

evidently observed in the impacted groups. Moreover, behavioral transformations such as concentration difficulties and nervousness were also apparent in children and distance learning students. Another study focusing on the economic outcomes of the pandemic outbreak in the United States has demonstrated a massive spike in unemployment insurance claims. These alarming potentially and intensely severe side effects can lead to catastrophic outcomes that could be worse than the pandemic itself (Baek et al., 2020).

1.7 Research Aims & Significance

Literature observations and close examinations provide valuable insights into the most dominant gaps within a certain field. It has been observed in the literature, as will be heavily discussed in Chapter 2, that very few researchers have attempted to present a framework to aid policy makers in measuring the impact of their policies while administering a natural disaster. In addition, almost none of the studies has presented a complete and comprehensive framework demonstrating and incorporating every single step, from announcing a certain policy to data collection and analyzing the public sentiment, especially using state-of-the-art deep learning methods. Thus, this thesis aims to establish a complete and reliable framework to detect, measure, and analyze public sentiments for decision-makers and policy legislators while managing a natural disaster or pandemic. The presented framework will enable decision-makers to employ massive social media data to promptly measure a new policy's impact during a natural disaster. The research explores COVID-19 as a very recent and practical case study. Moreover, the study will employ multiple machine-learning models to analyze the public sentiments toward official preventive strategies and procedures such as compulsory leave-home permits, restricting religious meetings and congregations, public utilities capacity reduction, international or border lockdowns, and national and complete internal lockdowns. The research will target major announcements and policies will be identified and selected, considering the following criteria:

- A predefined high-engagement threshold indicating the minimum number of interactions per announcement (e.g., 10,000 tweets)
- Social media trending topics and significant events hashtags within a specific location and timeframe.

In other words, the study will closely examine the COVID-19 timeline and carefully analyze preventive restrictions and soft returning procedures and their controversial positive or negative impacts on business owners, students, frontline workers, regular employees, and the general public. Furthermore, the study might extend beyond COVID-19 policies and investigate other governmental or public announcements and newly established guidelines. The research was primarily directed toward governmental organizations and legislative authorities as the main targeted sectors and domains. The proposed model will attempt to detect and analyze public reactions and sentiments toward specific and recently announced procedures, restrictions, or regulations. In addition to opinion mining, the algorithm will seek to predict or forecast potential sentiments and public responses toward a planned or upcoming policy or announcement through comprehensive training on the obtained historical data. The anticipated result shall tremendously support government agencies and authorities in preparing and modifying their communication techniques, delivery approaches, and timing plans ahead of the final release and publication as a proactive measure.

Furthermore, the research will attempt to verify the claimed advantages of emoji-based sentiment analysis over text-based, including its universality, accuracy, and simplicity compared to pure textual statements. In principle, the study pursued the following primary research question:

What is the most reliable and accurate approach or technique to both measure and predict public sentiments for decision-makers while managing and administrating the spread of a pandemic or a natural disaster in general?

The study tests different deep learning techniques, including the transformers model in sentiment analysis, which, to the best knowledge of the researcher, is relatively unresearched in this area. One aspect that makes the transformers model unique when comparing it with other deep learning models in sentiment analysis is the fact that it adopts self-attention and differentially weights the significance of each part of the input (which includes the recursive output) data. This study compares various high-performing models as suggested by existing research to answer the above research question.

As previously stated, the study utilized COVID-19 outbreak-related data/sentiments across the globe as a solid case study and a recent example of a pandemic. Additionally, the research will investigate multiple subsidiary questions concerning the policies' impact on various societal groups, the most promising accuracy enhancement procedures, emojis' sentiments practicality, and the abilities of state-of-the-art-based hybrid methods. The targeted beneficiaries, referred to as decision-makers in the study, are governmental agencies and enactment and executive authorities. In summary, the study's secondary research questions are as follows:

- i. In general, is there a notable and detectable pattern in public sentiments and emotions toward precautionary measures during natural disasters?
- ii. How distinct or diverse was the impact of these measures among various societal groups and categories?
- iii. How reliable and effective are emojis in enhancing the accuracy of sentiment analysis and predictions?
- iv. What is the most potentially promising technique to successfully identify emojis' hidden annotations and vague implications?
- v. Overall, what are the most robust machine-learning approaches or hybrid methods to achieve significantly precise outcomes in sentiment prediction and

forecasting?

- vi. What is the performance of the developed model(s) compared with the state-of-the-art benchmarks in the industry?

In terms of novelty, as addressed before, the study attempts to produce a step-by-step guide utilizing state-of-the-art methods and techniques within every stage in pursuit of producing and achieving an enhanced and highly reliable framework. Thus, the research shall make a scientific contribution to the computer science board of knowledge as well as a societal contribution by providing decision-makers and authorities with a practical framework to measure and predict the impact of their precautionary policies during a natural disaster effectively. In other words, the study addresses the problems of inadequate incorporation among the recent promising techniques within the sentiment analysis paradigm and overlooking disaster management as a significant application for this particular field. Both issues have been thoroughly elaborated on in the introduction and the literature review sections. Finally, the study will test two hypotheses:

- i. Natural disasters' precautionary policies, measures, and restrictions considerably impact public sentiments in general, as well as having a noticeably diverse effect on various societal groups.
- ii. LSTM is the most reliable and accurate approach to both measure and predicts public sentiments for decision-makers while managing and administering a natural disaster or a pandemic.

1.8 Problem Statement and Gap in literature

1.8.1 Problem Statement

Studies surrounding emoji-based sentiment analysis and opinion mining have only been a fragment of what has been accomplished with textual data. Sentiment analysis of textual data is an important area in many industries, including for policymakers, because apart from

analyzing opinions, it puts brands and policymakers (for example, in government) to understand and decipher what people are talking about on a certain topic. Some existing studies have carried out sentiment analysis to understand opinions on various topics to identify potential difficulties and hidden possibilities in people's perceptions. The most pertinent question, however, has always been the accuracy and quality of models that have been used to predict such sentiments on different platforms, such as social media. So far, there is no scholarly consensus on the most reliable and accurate approach or technique to measure and predict public sentiments for decision-makers, especially during natural disasters.

1.8.2 *Gap in Literature*

The COVID-19 phenomenon disrupted not just businesses but also people's ways of life, and in an effort to curb the spread, governments across the globe put in place restrictions on people's movements and ways of life. Being a relatively new phenomenon, the few studies that have carried out a sentiment analysis on COVID-19 were mainly limited in scope and focused on smaller samples. There are also very few studies that have incorporated emojis and their propitious potential in sentiment analyses. Therefore, a gap in the literature exists that the current study sought to fill. By investigating the most reliable and accurate approach or technique to both measure and predict public sentiments for decision-makers while managing and administrating the spread of a pandemic, this study explores an array of machine learning models in an effort to suggest the latest, most accurate prediction model. In doing this, the current study sought to suggest the best ways of analyzing sentiments on Twitter using data from across the globe.

1.9 Thesis Organization

The current research is structured into six consecutive sections. While chapter 1 provided an introductory background, rationale, the study's significance and the study's key

objectives, chapter 2 describes the literature review, where existing research on the research topic is examined. This chapter also focuses on the theoretical framework apart from the empirical literature review. Chapter 3 overviews the research methodology, framework, and paradigm and explains the detailed approach. Data collection and processing methods and approaches are discussed in this chapter. Chapter 4 focuses on the study's results and experimental setup, Chapter 5 discusses and evaluates

Chapter 2: Literature Review

Carefully observing and accurately analyzing the associated literature provides compelling intuitions into the most prevailing discrepancies and gaps within a particular field of interest. This section highlights the most distinct and notable contributions and associated work within the field of sentiment analysis in social media, with specific attention toward emojis.

2.1 Introduction

The bibliographic review will incorporate (1) an extensive overview of prior research, (2) a systematic evaluation of the preceding contributions, (3) a concise discussion of previous observations, and (4) a coherent description of the former results and outcomes (Schwarz et al., 2006). Figure 1 demonstrates the particular methodology and framework adopted to achieve the desired literature review (Rowe, 2014). The examination was initiated based on the abstract research questions, which will be comprehensively elaborated on in the research methodology section. Consequently, former publicly published studies have been indexed as part of an inventorying procedure. The indexing practice involved a vast collection of highly reputed databases and resources, such as Science Direct, Engineering Village, Research Gate, and Google Scholars.

Various databases were thoroughly scanned based on predefined vital phrases appearing in headings or keywords. The search phrases merged data science, sentiment analysis, opinion mining, machine learning, and social media terms using Boolean operators, for example: "sentiment analysis using social media emojis," "emojis interpretation using deep learning," "opinion mining on Twitter emojis," among others. The primary search resulted in numerous studies, of which many were eliminated by establishing a quality assessment procedure. The filtering methodology excluded outdated findings, unpublished papers, and repetitive concepts. Moreover, the inventory was further refined using citations, rankings, and

overall reputation.

Figure 2

The Quality Assessment Procedure

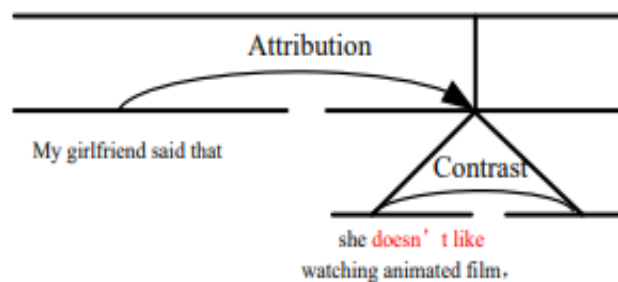


2.2 Theoretical Framework

This subsection will overview the foundational roadmap adopted to develop this study.

2.2.1 The Rhetorical Structure Theory (RST)

Unlike Long Short-Term Memory (LSTM), this theory (RST) definitively highlights the significance associated with the different segments in a text. The RS theory was proposed in 1987 by Thompson and Man (Pang and Lee, 2005). The theory is termed descriptive since it explains the organization of the text, particularly the association between the different sections of a text. The theory proposes 23 types of rhetorical associations, such as condition, contrast, and attribution. The relations, in theory, divide a text into peripheral and core segments, named satellites and nuclei. The relation structure of the theory in a sentence is described by Fu et al. (2016) in Figure 3 (next page) with the sentence being *My girlfriend says she doesn't like watching an animated film, but very much likes this movie*. The horizontal line indicates the layers' boundaries in the theory's roles in a hierarchical rhetorical structure. The arrow shows the pattern from the satellite to the nucleus with definitions over the horizontal line showing the particular relation between the satellites and the nuclei. It includes a few of the 23 types of rhetorical relations mentioned above. Emotional words are indicated in red, and the contrast association is identified with the line lacking arrows on both ends, indicating parallel relations.

Figure 3*Relation Structure of RST Theory (Fu et al., 2016)*

2.2.2 The Network between RST and LSTM

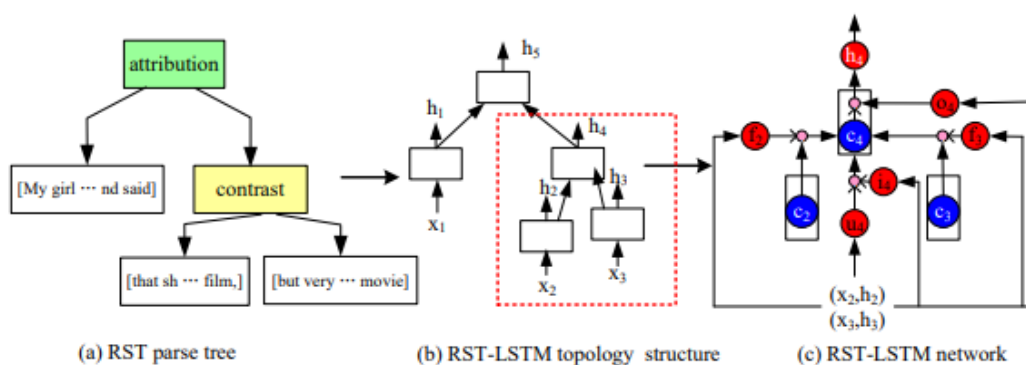
The structure of a text is considered a key research direction for tasks associated with sentiment analysis. According to Hogenboom et al. (2015), recent research has used the RST to parse texts to differentiate significant segments from less significant ones regarding their contribution to the overall sentiment. The proposed model by Fu et al. (2016) includes three sections: the model's topology structure, the theory's parse tree, and the network of the model. As mentioned by researchers, the model's parse tree considers the complete hierarchical rhetorical structure of a sentence instead of the rhetorical functions of a segment. The researchers used two parsers for the rhetorical structure that was proposed by Surdeanu et al. (2015), where the two parsers used a similar underlying framework with the characteristics of the syntax differentiating. One parser is based on the dependency parse tree (D), and the other is based on the constituent parse tree (C). The parse trees used in the sentence were the same. As discovered, it remains possible to parse another sentence using the constituent and dependency RST tools.

Compared to other discourse analysis tools of the RS theory, the featured discourse parsers utilize a simple modular structure. Not only is their speed in parsing fast, but they also influence achieving better results. Depending on the dependency or constituent syntax, the RST parse tree describes the master-slave association in the nodes of each layer, the satellite and nucleus segment, or the parallel association. The two sections placed at the bottom right

of Figure 4 (next page) are of equal significance, concluding that all are nucleus segments. The relations data is found in their parent node. Fu et al. (2016) also acknowledged that previous scholars had used the RS theory parse tree; however, the combination technique for the sentences was a simple weighted method rather than automatic learning. Fu et al. (2016), motivated by the works of Zhu et al. (2015), use the LSMT method as a replacement for the weighted method. The rectangles in Figure 4 are a representation of an LSMT unit.

Figure 4

The RST-LSTM Model (Fu et al., 2016)



The proposed model by the researchers not only uses the theory's parse tree to differentiate the segments of a sentence instead of considering sentences as a temporal sequence, but it also uses the LSTM structural features to automatically filter the satellite data of a text and improve the nucleus information instead of the simple weighted of the distinct rhetorical roles of segments. The proposed model combined LSMT and RST, making it possible for researchers to realize their advantages to the greatest degree.

2.3 Societal Impact of the COVID-19 Pandemic

Coronavirus disease 2019 (COVID-19) was first tested in December 2019 in Hubei province in China. Later, it was discovered that a SARS-CoV-2, termed COVID-19, was responsible for causing the infection. The virus outbreak compromised the stability of

government systems, generating unprecedented economic and social disruptions in countries worldwide. According to the World Health Organization (WHO), one year after the detection of COVID-19, 2% of the world's population was infected, with 3.5 million of the 2% losing their lives. Bearing in mind the rise of infected people, the government, through health institutions, introduced preventive measures and policies to mitigate the spread of the virus. These safety precautions include wearing safe masks and social distancing, and with a further surge of infected people, stay-at-home and lockdowns were adopted by some countries. Through this, social media platforms became a crucial element in crisis communication, allowing people to connect, work virtually, and share personal feelings, expectations, aspects, and the prevalence of the disease. Later, social media assisted policymakers in overseeing public perceptions and opinions in responding to the COVID-19 pandemic through sentiment analysis.

The coronavirus resulted in significant societal unrest and unprecedented changes in work, lifestyle, and social interactions. This precipitated the introduction of lockdown precautionary policies such as social distancing and closure of interaction and gathering centers, leading to increased social consequences. Studies show that, in a patriarchal lifestyle, an elongated stay at home leads to increased pressure on women to perform household tasks. The pandemic's psychological and economic dimensions also substantially impact families (Hosseinzadeh et al., 2022). As a consequence of the crisis, cases of domestic violence such as elder, child, spousal, and disability abuse were on the rise. The closure of universities and schools deprived many children and young people of social and educational activities for a long period, which exhibited difficulties in compensating. Travel restrictions impounded by different nations lowered social relations in external spheres and contributed to the isolation of people. Since the pandemic influenced every sense of human life, divorce rates have increased, and marriages have reduced in many countries. In the context of couples, a

substantial drop in marital associations was witnessed due to the fear of transmitting or contracting the virus. This significantly altered people's opinions online, which ranged from positive to negative.

Individuals across the globe were directly or indirectly affected by the severe consequences of the crisis. As a result of isolation and cessation of specific social affairs, the pandemic amplified the problems related to economic recession, social anxiety, insecurity, and increased psychological stress, which needed coordinated and integrated measures to control and prevent them. Individuals were expected to follow WHO and government guidelines to make the necessary, which contradicted the internal desire of their daily plans. Studies highlight different groups and categories significantly affected by the virus, including older persons, people with disabilities, indigenous persons, youths, and families. Kutsar and Kurvet-Käosaar (2021) indicated youths and adolescents interact in various life domains, such as friendships, community through technology, school, and family, which were all impacted by the COVID-19 crisis. For example, Scott et al. (2021) conducted a qualitative study on adolescents and their experiences amid the coronavirus crisis. The researchers established fourteen thematic fields where the crisis influenced the group's way of life, and they include physical and mental health, friendships, family relations, attendance of key events, social relationships, and community, exposure to COVID risks, observance of COVID rules, routines, and the adoption of new technologies.

Similarly, Idoiaga et al. (2020) examined the emotional and social behaviour of children in Spain during the pandemic and analysed their coping capabilities. Literature on public health, child psychiatry, and psychology across the globe revealed that the COVID crisis increased mental health problems and child vulnerability, leading to a fall in subjective health. Disruptions to social life and schools were considered the main predictors of child vulnerability. However, other pandemic-related stressors such as anxiety and depression,

parental stress within families, and unprecedented parental problems also influenced young people's ability to cope with COVID-19.

The COVID crisis also posed a great threat to indigenous globally. This group already experiences limited access to healthcare, increased rates of both communicable and non-communicable diseases, and a lack of essential services. Even when this group can access healthcare services, they still experience stigma and discrimination. Their traditional lifestyles were a source of resiliency but were also a threat while attempting to contain the spread. This increased mortality rates globally, especially for such vulnerable groups.

2.4 The Psychological and Emotional Effects of COVID-19 Lockdowns

The pandemic of COVID 19 has impacted people all across the world. It spawned the creation of phrases like "lockdown," "emergency," and "curfew" and the introduction of preventative measures like physical separation in several countries (Xue et al., 2020). Not only has it economically and militarily impacted society, but it has also influenced its psyche. This unease in the human emotional quotient is caused by several interrelated elements, including the financial ramifications, the behaviour and support of family members, country-specific lockdown restrictions, media effects, and fear of the pandemic. The development of digital technologies and massive amounts of internet data on individuals has prompted researchers to investigate human brain functions and emotions to enhance the consumption of technology-based services. To increase the efficiency of pandemic management, Choudrie et al. (2021) noted the importance of understanding individual emotional differences to shed light on public perceptions of different government approaches to the problem. Increasing numbers of people have turned to Twitter and other microblogging services to share information and vent their concerns throughout the outbreak.

Xue et al. (2020) found that social media data and machine learning approaches may be used for infodemiology research, allowing researchers to evaluate how public discourse and

opinions shifted throughout the COVID-19 pandemic. While the situation deteriorated fast, several topics emerged as trending topics on Twitter, including confirmed cases and fatality rates, preventative measures, government policies and health authorities, the stigma associated with COVID-19, and adverse psychological reactions like fear. The public health sector gained invaluable information by real-time monitoring and analyzing Twitter comments and concerns. The public's trust will be affected by the dread, stigma, and mental health problems linked to the epidemic in case of a second wave.

Ning et al. (2020) aimed to assess Chinese citizens' adoption of protective behaviors during the COVID-19 pandemic. After controlling for potential confounding variables, the study found that individuals who perceived the pandemic as highly severe exhibited high levels of negative emotion, reported good health, paid close attention to and trusted the official media, and were more likely to take preventative measures. Also, it was assumed that women and older people were more prone to take precautionary measures.

Adaptation perceptions and strategies are crucial to the sustainability and resilience of society. Risk perception and enhancing individual and social resilience are essential elements of catastrophe risk management. The behavioural responses of the general public are crucial for minimizing the prevalence of developing diseases. The frequency of COVID-19, as it was at the time, was perfect for studying changes in behaviour. This is because the government did things it had never done before to stop the spread of the disease. For example, it closed educational and administrative centers, started a lot of public awareness programs, and gave different people different information. This was partly because scientists kept learning more about the new disease and when it would start spreading (Samadipour et al., 2022).

Addawood et al. (2020) stated that the coronavirus illness of 2019 significantly influenced global economies and communities. While governments were taking severe steps to prevent the spread of the virus, these new tactics began to harm the populace. With constraints such

as lockdown and social isolation, it became crucial to comprehend the public's emotional response to the pandemic. The study's findings indicated that people had a favourable attitude toward the pandemic. This positive response was at its peak during the start of the pandemic and has since begun to wane. The conclusion was that people supported one another during the pandemic.

2.5 Public Sentiments, Opinions, and Emergencies

Derivative public opinion and sentiments are often formed in the discussions and interactions of social network users on events resulting from emergencies. Such related events are considered derivative events, which are derived from original events. According to Li et al. (2010), the difference between the former and the latter is that the original events are consistent with regard to the cause and type of the occurrence, with linkage and continuity. In contrast, the derivative events are inconsistent in relation to type and cause of occurrence. Cai et al. (2021) argued that online public sentiments and opinions during emergencies encompass a synthesis of opinions, sentiments, and attitudes displayed on social platforms around the occurrence and distribution of emergencies in a given period and space while using social networks as communication tools.

Online users often express their views and evaluations on derivative emergencies through social networks to ensure dissemination and prediction of public responses to emergencies. Priors reveal that information mining approaches have been applied to understand the distribution mechanism of derived public sentiments and emotions in emergencies, and machine learning approaches have been applied to discover the prediction of public opinions and responses. For instance, Liu et al. (2019) proposed a digital model that simulates the propagation dynamics of a series of rumours in social platforms to establish special derivative public opinion and propagation aspects of network rumours in social networks. Additionally, Zhu and Wang (2020) proposed a control model. They elaborated on

the propagation mode of derived public sentiments based on the propagation approach to discover how to control derived public sentiments. Regarding the prediction of online public sentiments, Srijith et al. (2017) suggested using ML approaches such as HDP to analyse text semantics to establish the classification of derived public emotions and opinions. Further, deep learning algorithms, including neural networks, have been applied to discover the derived public sentiments through the analysis of social media data.

Yadav and Vishwakarma (2020) explored the application of sentiment analysis in network public sentiments and opinions. The authors noted that extracting users' tweets, posts, blogs, comments, and discussions on social networks necessitated mining and analyzing the attitudes and opinions of network users and then interpreting individual behavioural preferences and sentimental trends. Sentiment analysis of public opinion on emergencies chiefly focuses on two key aspects: factors influencing sentiments and the occurrence of the sentiments. Analysis of text information also allows comprehension of how sentiments change in different spaces and times and discovers the evaluation and forecasting of emergencies. For instance, Singla et al. (2017) divided the review text of Amazon mobile phone products into sadness, joy, anger, and trust. They understood the variables that affect consumers' emotions through review text analysis to provide feedback to the company.

Additionally, studies on factors that affect emotions employed sentiment analysis to explore the evolutionary exchange mechanism of users' opinions on social media to discover the prediction of emergencies. For instance, in accordance with sentiment transference and infectivity, Budiharto and Meiliana (2018) utilized social media to forecast national political elections. The authors found sentiment analysis approaches effective in generating good and accurate results. Sentiments are usually classified based on sentiment dictionaries and machine learning. In relation to the sentiment dictionary, words and phrases are used to mark the sentiment direction of the text via manual screening derived from the available dictionary

or approach of the corpus dictionary. Aslam et al. (2020), for example, suggested an NRC Sentiment Dictionary, which computes the sentimental performance of new headlines in multiple directions, usually eight, when emergencies occur. In addition, the point mutual derived available sentiment dictionary approach is applied to create emoji dictionaries that help compute document sentiments. On the other hand, the ML-based dictionary utilized unsupervised, supervised, and semi-supervised ML approaches, including Naïve Bayes, neural networks, and SVM, to classify sentiments based on the training data.

Emergencies like the COVID-19 pandemic significantly impacted public emotions, opinions, and sentiments. As such, analysis of COVID-19-related data depends on public emotion, sentiment, and opinion analysis in social media posts and blogs in English across multiple studies. For example, Kumar (2020) applied RNN models, including BiLSTM, SAB-LSTM, and LSTM, for training and testing datasets. The author employed the F1-score performance measure based on the specified score; the SAB-LSTM attained a higher F1-score compared to the other models for COVID-19 sentiment prediction. Sciandra (2020) found that social networks, including Reddit, enable healthcare providers to gather data on public opinions that can be applied to analyse human behaviour and discover knowledge. The researcher employed a generic method based on NLP, which helps extract vital aspects regarding comments associated with the COVID crisis. The sentiment classification system allows text classification only at the sentence level but not the document level. This method necessitates the extraction of characteristics associated with public sentiments and utilizes the WordNet lexical approach dataset to classify opinion words. The elements were organized under relevant opinion sentences, and the characteristics were scored based on the number of times they have been reviewed.

Jali et al. (2022) used Twitter data for sentiment analysis to discover public sentiments to examine the growing fear related to the COVID-19 pandemic. Twitter is among the most

often-used social networks, and it has recorded significant growth in the tweets associated with coronavirus, including negative, neutral, and positive tweets. Individuals have expressed their feelings about vaccination safety and efficiency on social platforms like Twitter. The proposed approach uses the COVIDSenti dataset to apply the Twitter Sentiment. It also evaluates DL and ML approaches evaluation metrics, including F1-score, precision, recall, and accuracy. The study established that the proposed approach provides improved accuracy of 93.88, 94.33, 95.22 and 96.66% of COVIDSenti_C, COVIDSenti_B, COVIDSenti_A, and COVIDSenti, respectively, compared to conventional DL and ML algorithms.

2.6 Precautionary Lockdowns on Public Sentiments and Emotions

Public sentiments are public opinions or attitudes concerning a circumstance, which can be positive, negative, or neutral. Due to precautionary measures such as lockdowns and working from home brought about by the widespread COVID-19, people took on social media platforms such as Twitter and Facebook to express their sentiments and emotions that supported or opposed the measures. Ideally, social media information, for example, tweets, have grown significantly to be used in health research associated with COVID-19 by analyzing public sentiments to make robust decisions regarding activities that support or inhibit the elimination of the virus. Studies conducted to examine the impact of precautionary measures on public sentiments have shown a variation between positive and negative sentiments about the control measures. For instance, to identify social media's insight on US mental health during the COVID-19 pandemic, Valdez et al. (2020) collected 86 million English tweets from January 22 to April 9, 2020. Using eight known topics to scrutinize the tweets, it was realized that the number of tweets increased and then dropped over time. Still, the general sentiment changed from negative to positive. However, some studies identified positive and negative sentiments concerning different topics related to the virus.

Chandrasekaran et al. (2020) conducted a study to examine subjects, trends, and views of tweets about the COVID-19 pandemic on 13.9 million tweets between January 1 and March 9, 2020. The results identified persistent negative opinions on subjects associated with the growth and spread of COVID-19, while opinions on preventive measures and treatment shifted from negative to positive. The above results were similar to the findings of Al-Rakhami and Al-Amri (2020), which established a high percentage of spotting reliable and non-reliable tweets involving COVID-19. Due to this, analysis of the Coronavirus pandemic and its consequential outcome on the people, especially lockdowns, have identified many resources channeled towards inspecting people's sentiments and reactions. Relatedly, a sentiment analysis study by Barkur et al. (2020) was carried out after the lockdown announcements in India, with Twitter being the main data source. The research used tweets posted between March 25th and 28th via two common hashtags (#IndiafightsCorona and #IndiaLokdown). A word cloud tool was adopted to aggregate the current sentiments and emotions conveyed in the tweets to establish that despite the fear of what the virus may spell for the Indians, their tweets were positive and hopeful about pulling down their curve.

Similarly, Agarwal et al. (2022) undertook a study to identify the consequence of lockdown measures on communal perceptions in some countries using temporal sentiment analysis on data collected from Twitter. The findings reveal that opinion in the United States of America ranged from neutral to negative, while in Italy and India, there was a mixed reaction and opinion regarding the virus. The study further examined the outcome of government control measures in the different stages of lockdown as per the public sentiments and emotions in India with the help of common hashtags, for example, #WorkFromHome. It was discovered that most people in India expressed positive thoughts during the lockdown period about working from home. Contrary to this were the findings of Medford et al. (2020), who examined the changes in Twitter events, sentiments, and contents concerning COVID-

19. The examination utilized tweets from January 14th to 28th, 2020, that were associated with infection prevention practices, such as lockdowns, while carrying out experiments to decipher views on the dominant and violent emotions. The study revealed that from January 21, 2020, the hourly mentions of keywords related to the virus significantly increased, with 50% of the tweets exhibiting fear while 30% showed surprise. Additionally, some of the tweets exposed cases of racial discrimination, with the political and economic impact of the virus being the popular topics of discussion.

COVID-19 precautionary measures posed a huge challenge in risk management regarding crisis communication by the mass media, impacting the emotions and sentiments of the public. Parallel to this is a study conducted in Spain to understand the influence of COVID-19 on emotion and digital ecosystems through sentiment analysis. Employing Web Scraping techniques and APIs, a sample size of 106,261 communications was obtained (de Las Heras-Pedrosa, 2020). It was realized that the four main emotions- fear, sadness, disgust, and anger were present. Further scrutiny was done on the emotions to determine themes that affected the reactions relating to crisis communication. Feelings of disgust, fear, sadness, and anger were realized through death forecasts, border closures, job losses, blaming the pandemic on people that can be grouped by sex, and the collapse of health bodies. The emotions were highlighted through the following themes; overwhelming social security system, loss of employment, delay in activating health alerts, and speedy development of the pandemic in Spain.

2.7 Sentiment Analysis and How it is Used

Sentiment analysis is also known as opinion mining. It is the procedure of locating and eliciting viewpoints, feelings, and emotions from the text. It is a sort of natural language processing (NLP) that analyzes text using machine learning algorithms to extract human information, including attitudes, emotions, and both positive and negative sentiments.

(Medhat et al., 2014). According to Medhat and partners, sentiment analysis is typically used in marketing, customer service, and public opinion research. It enables organizations to gain insights into public sentiment and reactions to their products and services and the news and events in their industry. Sentiment analysis typically involves classifying text into positive, neutral, or negative sentiment. This is generally done using a combination of natural language processing (NLP) methods and algorithms in machine learning (Medhat et al., 2014). For example, a sentiment analysis algorithm might use a bag-of-words approach to identify words that are associated with positive or negative sentiment, such as "happy" or "sad." It might also use machine learning algorithms to identify patterns in text and classify them accordingly.

According to Medhat et al. (2014), sentiment analysis is essential for discerning public opinion and gauging how people feel about specific topics. It is imperative in natural disasters, as it allows governments to assess the public response to their preventive measures and policies. This helps to inform decision-making, allowing governments to make better-informed decisions about responding to natural disasters. Additionally, sentiment analysis can provide valuable insights into public sentiment, allowing governments to understand better the public's reactions to their actions and policies. This can be used to create more effective future policies and measures, helping protect citizens from the effects of natural disasters. For example, sentiment analysis can be used to understand better public sentiment towards the government's response to a recent hurricane in the Gulf Coast region of the United States. This could inform the government's decision-making process, allowing them to better assess the public's reaction to their policies and measure the effectiveness of their response. Additionally, sentiment analysis can determine the public's reaction to the government's efforts to mitigate the effects of natural disasters, such as providing aid to affected

communities. This can help inform future policy decisions, allowing governments to better protect their citizens from the effects of natural disasters.

Sentiment analysis can also be used to analyze emojis and emoticons commonly used in social media. Emojis and emoticons are often used to convey feelings and emotions, and sentiment analysis algorithms can identify and extract these implicit sentiments from the text. For instance, a sentiment analysis algorithm might use a bag-of-emojis approach to identify emojis associated with positive or negative emotions, such as a "thumbs-up" or a "frowning face" (Ullah et al., 2020, pg.2). Additionally, sentiment analysis can be used to identify the sentiment of user-generated content, such as comments, reviews, and tweets, as well as to identify and extract subjective information such as opinions, emotions, and attitude. According to Uriah and partners, in this way, sentiment analysis can be used to gain insights into public views and reactions to events and news, as well as to understand the sentiments of customers and potential consumers.

Several methods are utilized or applied in sentiment analysis. These methods include lexicon-based, supervised machine learning, and unsupervised machine learning methods. Lexicon-based methods are the most straightforward form of sentiment analysis. They involve matching words or phrases in the text with a list of pre-defined words or phrases to calculate the sentiment of the text (Zhang et al., 2011, pg. 2). According to Zhang and his colleagues, supervised machine learning methods involve training a model on a labeled dataset, and then using the trained model to classify the sentiment of new text. On the other hand, unsupervised machine learning methods involve clustering text into groups with similar topics and opinions and then analyzing the clusters to infer the overall sentiment of the text.

2.8 Related Works

The use of sentiment analysis in social media to analyze public opinion during natural disasters has been studied by several researchers. Zhang et al. (2011) developed a sentiment

analysis system for Twitter that uses a machine learning algorithm and a lexicon-based approach to detect the public's reaction and opinion towards natural disasters. The study found that the system could accurately detect public sentiment and reaction towards natural disasters (pg. 4). The use of sentiment analysis in social media to examine the public's opinion during natural disasters has also been studied in the context of industries. For example, Obembe et al. (2019) used sentiment analysis to examine the public's opinion towards COVID-19 and the tourism industry. The study found that the public was generally hostile toward the conflict and that the sentiment analysis system could accurately detect the public's opinion. Finally, a study by Zvolensky et al. (2020) emphasized a data-driven approach for predicting public sentiment toward disaster management policies using Twitter data. The authors collected tweets related to the pandemic and used sentiment analysis and classification techniques to classify the tweets into negative, positive, or neutral categories. The authors then applied machine learning models, such as support vector machines (SVMs) and random forest (RF) algorithms, to predict public sentiment toward disaster management policies.

2.9 Impact of Sentiment Analysis among Various Societal Groups

The impact of sentiment analysis on various societal groups is complex. Sentiment analysis of various societal groups can provide valuable insights into how different communities may be affected by a natural disaster. For example, an analysis of the sentiment of low-income communities may reveal that they are more likely to feel neglected during a natural disaster due to a lack of resources or inadequate response from the government. This insight can help governments target resources more effectively and ensure everyone is treated fairly during a natural disaster. Sentiment analysis can measure public opinion on a range of topics, such as the effectiveness of government policies, the economic outlook, or public opinion on current events. This can benefit policymakers, as it can explain how different

groups are likely to respond to proposed policies or changes. However, it is essential to note that the impact of sentiment analysis on different societal groups may vary. Different groups may react differently to the same set of policies, depending on factors such as income, location, and social norms.

The impact of COVID-19 precautionary measures on public sentiment is common across all societal groups. A study by Zvolensky et al. (2020) showed that while the overall sentiment towards the lockdown policies was largely negative, there were distinct differences in sentiment among different groups (Pg. 3). For example, the study found that those with higher levels of income were more likely to be more accepting of the measures, while those with lower levels of income were more likely to be more resistant. Furthermore, the study found that those in rural communities were more likely to be less accepting of the measures than those in urban communities. This suggests that the impact of lockdown policies on sentiment is unique across all social groups. Therefore, different groups may have different reactions to the same set of guidelines. Undeniably, the confining measures set to mitigate the COVID-19 pandemic severely affected people's psychological, emotional, and behavioral activities. The impacts of the measures adopted to counter COVID-19 were distinct across almost all societies worldwide. One major societal group that was affected during the pandemic across the world is healthcare providers. Being on the frontline in serving their nations during the pandemic, these workers were forced to work in different environments and conditions. The conditions included a high workload, the agony of losing family members, patients, and colleagues, insufficient medical supplies, and personal protective equipment, mostly in evolving nations. Chew et al. (2020) research revealed that such situations and environments make health professionals more vulnerable to mental health burdens due to increased anxiety, stress, anger, depression, and insomnia. Another societal group that was greatly impacted across the globe is the elderly since they were more

vulnerable to the virus. The lockdown measure imposed on the elderly, instructing them to stay at home, has shown how the elderly are at a greater risk for depression and anxiety (Conejero et al., 2020). The social disconnection of the elderly caused and escalated depression, anxiety, neglect, and loneliness, further causing the elderly to incur long-term health problems (Bavel et al., 2020). Further studies showed that social discontinuation might escalate neurodegenerative disorders and, in addition, intensify suicidal cases among the elderly through the prevailing pandemic.

The adoption of lockdown measures following the ongoing pandemic has led to an unbearable economic crisis in every domain of work and business across societies worldwide, giving rise to unemployment. Based on Restubog et al. (2020), a survey conducted on US workers before and after economic outcomes revealed that pay cuts, challenging working conditions, and reduced work hours are also possible contingency plans for the COVID-19 pandemic. This has enhanced financial losses associated with psychological stress, further influencing mental health ailments. According to Conejero et al. (2020), the disturbing issues that impact economic conditions are one of the chief risks for suicidal actions identified in most countries during the pandemic. Data provided by the International Labor Organizations in March exposed that a study by Kawohl and Nordt (2020) approximated an increase of 2,135 suicides worldwide in a year in case the rise in unemployment prevails.

Additionally, measures such as working from home brought by the lockdown policies mandated that people stay at their homes, although to some, homes are not always safe, making them a place of abuse. The study by Bavel et al. (2020) showed how disaster-related instability, economic stress, and forced proximity influence domestic violence and aggression. This was distinct across nations like the United Kingdom, whose domestic abuse organization testified a 25% increase in calls to the domestic violence helpline after the announcement of lockdown measures by the government. This result corresponded with the

findings in Australia, where police sectors revealed a 5% upsurge in calls about domestic violence, and Google declared a 75% surge in internet explorations for domestic abuse support. In addition, there was a 21-35% increase in complaints about domestic violence in the US and a 32-36% increase in France after implementing precautionary lockdown measures.

2.10 Varied Impacts of Covid19 Containment Measures

According to Kausar et al. (2021), theprecedented nature of the COVID-19 Pandemic and the undertaken measures meant that communities across the globe were experiencing the same problem. As such, the researchers concluded that the COVID-19 lockdown measures undertaken in 11 nations across the globe elicited similar emotions that included uncertainty, fear, and anxiety. Kausar et al. (2021) findings suggest that COVID-19 lockdowns had similar impacts across different societies across the globe. These findings have been countered by other scholars who argue that the impacts of COVID-19 containment measures would have varying effects depending on individual communities' socioeconomic characteristics. For example, Rasul et al. (2021) explain that the Pandemic will affect all domains of society, including macroeconomic stability, economic growth, and income levels. Additionally, the pandemic countermeasures are expected to deepen the existing socioeconomic inequalities that deepen poverty, unemployment, and food insecurity for those in the working socioeconomic cluster. Shang et al. (2021) also point out that the effect of the pandemic will vary according to the individual nation's economic structure. Countries with LME economies (free markets) responded more to the undertaken containment measures than those with CME economies. In essence, the economic impact of the Pandemic was moderated by government involvement in nations with CME economies. For this reason, Shang et al. (2021) concluded that the impact of any pandemic could be moderated by increasing the overall involvement of the government authorities. Even though the entire world was facing a

common problem, individual countries' adoption of uniform containment strategies proved futile for several nations. A report released by Cambridge University Press demonstrates this point, which compared economic policy response models adopted by the UK and Germany, liberal market economy (LMC) and coordinated market economy (CME), respectively, in response to the pandemic. Despite having different economic structures, the two countries employed similar approaches to cushion their economies from the impact of the Pandemic. Since the systems differed, the results justified the view that policymakers should strive to customize policies to the individual nation rather than adopt a one-fits-all model (Hancke et al., 2021). Rocha (2021) commented on the general impact of the lockdown by noting that the Pandemic adversely affected all sectors, leading to widespread unemployment. The COVID-19 Pandemic is a multifaceted problem that impacts all aspects of life. According to Wildman (2021), there was significant heterogeneity in the number of cases and subsequent fatalities in the OECD nations, indicating that even though the Pandemic affected the global society, its effects are disproportionally higher for vulnerable members of society than others. The investigators concluded that income inequalities were a significant determinant of the efficacy of the employed containment measure. A positive correlation exists between the rate of socioeconomic inequalities and vulnerability to the Pandemic, with nations with high economic inequality recording higher pandemic burdens than those with low inequality levels. Wildman (2021) attributes this disparity to the fact that income inequality serves as a proxy for multiple socioeconomic disadvantages that can contribute to the contraction, spread, and deaths from COVID-19. Carreras et al. (2021) identify some socioeconomic disadvantage factors as obesity, pollution, and poor housing.

Like Wildman (2021), Carreras et al. (2021) argue that the impact of the Pandemic was not uniform, as poor communities were disproportionally affected. Whereas this fact is well-supported by scholars who have often supported their perspective by pointing to the overall

impact of poverty on health and well-being, little is known about the impact of poverty on compliance with government-imposed containment measures. Carreras et al. (2021) attempted to fill this knowledge gap by focusing on the effectiveness of Peru's containment measure. The study indicated that the economically vulnerable cohort was most likely to defy containment measures such as lockdowns despite the spiking number of positive cases and fatalities. Based on these findings, Carreras et al. (2021) concluded that factors such as poor housing and economic hardship make it impossible for the population belonging to the low socioeconomic class to adhere to lockdown policies.

Alaba et al. (2022) also note that socioeconomic inequalities increased following the COVID-19 lockdown. For example, in South Africa, cases of food insecurity, malnutrition, and obesity increased in slum communities following the COVID-19 lockdown. The authors attribute hunger to inequalities across several sectors, creating a need to adopt tailored pandemic control measures. This view is explored further by Nwosu & Oyenubi (2021), who argue that factors such as in Africa, such as per capita household income and experience of hunger, were significant predictors of income-related health inequalities brought forth by the Pandemic. The author notes that hunger, income, race, and employment determine socioeconomic inequalities that result in disparities in attempts to prevent and manage the virus. In other words, sub-Saharan African households with low to no regular income struggle to access basic necessities like food and healthcare.

McGowan & Bambara (2022) agree with Carreras et al. (2021) and Alaba et al. (2022) that the Pandemic is an unequal experience with high mortality rates being seen in the most deprived places in the world. This finding supports the view that a positive correlation exists between the Pandemic's mortality rate and poverty level. McGowan & Bambara (2022) reviewed multiple studies and uncovered patterns that indicated that even though individuals in the high-income cohort were at a higher risk of contracting the virus, the mortality rate

remained disproportionately high among those living in poverty. Similar patterns have been uncovered in Africa, South-East Asia, and Western Pacific region. The relationship between COVID-19 mortality and deprivation is multifaceted. In essence, the Pandemic has been termed a syndemic which describes a set of intertwined and mutually enhancing health issues that impact the overall health of a population within the context of a perpetuating social condition. Deprivation which encompasses multiple social determinants, including unemployment, poor housing, and lack of healthcare access, leads to multiple interacting adverse risk factors for Coronavirus mortality. These deprivation interactions could be summed up in four interrelated pathways: unequal exposure, vulnerability, transmission, and susceptibility.

Patel et al. (2020) explain that the vulnerable population faces high exposure to the virus compared to high-income individuals. Some of the factors that increase the exposure of the vulnerable population to the virus include overcrowded accommodations, poor working conditions, and the nature of employment for the majority of people belonging to low socioeconomic status who do not have the work-at-home option. Hence, containment policies like lockdowns mean a loss of income with immediate effects. McGowan & Bambara (2022) describes unequal exposure as a variation in one's ability to shield themselves from contracting the virus. Like Patel et al. (2020), McGowan & Bambara (2022) points out that the vulnerable population's poor living and working conditions increase the risk of contracting the virus.

Besides unequal exposure to the virus, poverty also increases the disproportionate risk of transmission. According to McGowan & Bambara (2022), unequal transmission could be described as the increased risk of the infection spreading within the household and larger community for individuals living in the most deprived area. As previously mentioned, vulnerable populations live and work in areas characterized by overcrowding and pollution.

Such conditions often make it impossible for those who have contracted the virus to isolate themselves, thus increasing the risk of transmission. Patel et al. (2020) Furthermore, about 20 percent of the UK's poorest households live in poor housing conditions that are often overcrowded and have limited outdoor space. All these factors make it impossible for such individuals to comply with the social distance protocol, increasing the overall risk of viral transmission. This point is further supported by an article published by the New York Times, which stated that people occupying the low socioeconomic strata are not only likely to catch the disease, but they are also more likely to die from the disease compared to those from high economic class (Fisher & Bubola, 2020).

The economically vulnerable cohort also faces unequal vulnerability to COVID-19, an increased mortality risk resulting from the high burden of non-communicable illnesses in the socioeconomically disadvantaged group. Shifa et al. (2021) agree with Fisher & Bubola (2020) on the view that economically disadvantaged groups are more likely to contract and spread the virus and are also likely to die from it compared to those who are economically stable. This variation could be attributed to poverty being associated with various COVID-19 risk factors such as respiratory complications, obesity, and heart disease. These risk factors could be traced to poor living and working conditions and lack of access to quality healthcare.

Focusing on the South African population, Shifa et al. (2021) illustrated that socioeconomic inequalities played a vital role in predicting the population's vulnerability to the virus. For instance, the researchers found that the vulnerability rate of the poorest quintile in South Africa stood at 30 percent while that in some of the wealthiest quantiles was only 2 percent. Similar patterns were uncovered by Rocha et al. (2021), who focused on Brazil. According to the researcher, the government adopted a uniform containment approach involving lockdowns and social distancing areas; a significant difference was noted between

areas with a high Socioeconomic Vulnerable Index (SVI) and those with a low SVI of mortality rate. This difference could be attributed to the fact that high SVI regions could enforce stringent prevention and management measures, which meant low transmission and subsequent death.

On the other hand, low SVI regions experienced significant challenges in enforcing these policies due to congestion and poor housing conditions. For these reasons, Rocha et al. (2021) concluded that the existing socioeconomic inequalities in Brazil rather than health status and age affected the Pandemic's course with a disproportionate adverse burden on municipalities and states with the highest socioeconomic vulnerability. Perry et al. (2020) support Rocha et al.'s., (2021) findings by noting that socioeconomic shocks of the initial stay-at-home orders by the US government disproportionately affected vulnerable groups.

The authors note that the observed variations on the impact of the Pandemic's lockdowns are not different from other major disasters like the Great Recession and Hurricane Katrina. In all these cases, the socioeconomically vulnerable groups face more hardships than those of higher economic status. In other words, the socioeconomically vulnerable groups' circumstances increase their risk of contracting the virus due to multiple poor-managed underlying conditions. At the same time, adhering to strict prevention measures could be impossible as such restrictions often mean a loss of the household's income. The congested living conditions might make it impossible to enforce social distancing, which means a high transmission rate. As such, scholars such as McCann & Matenga (2022) argue that the Pandemic has exacerbated existing socioeconomic inequalities, forcing billions into poverty. Even though lockdowns were necessary for combating the spread of the virus, the resulting economic crisis could culminate into serious socioeconomics and political issues in the future.

The impact of COVID-19 containment measures varied according to socioeconomic status and across racial and ethnic lines, with disproportionate suffering observed among minorities. This point is well-demonstrated by Ma (2021), who found that the Pandemic's burden was disproportionately high among racial minorities compared to the majority. For example, a survey targeting Long Island, US, found that 81 and 34 percent of the Hispanic and non-Hispanic population were infected and affected by the Pandemic once the herd immunity thresholds (HITs) were reached. This variation could be linked to racial minorities' contact level. Given the close association between racial minorities and poverty, it is pretty apparent that many ethnic and minority groups live and work in conditions that increase their overall exposure rate and do not allow social distancing and isolation. A report by the IZA Institute of Labor Economics revealed that racial minority groups like blacks suffered from high exposure to the virus yet remained exempted from the imposed containment measures (Bertocchi & Dimico, 2021). The scholars explain that blacks, especially black women, are overrepresented in low-paying jobs in critical sectors like transportation and healthcare. The report converges with Ma's (2021) view that the overall contact levels experienced by racial minority groups are high due to their work and living conditions. In essence, once one has contracted the virus, the congested living conditions make it impossible to self-isolate, thus increasing the risk of transmission to another family member.

Minorities are overrepresented in the low socioeconomic cluster, which means that a multitude of factors increases their risk of contracting the virus. At the same time, adopting containment measures might be impossible due to factors like overcrowding or the nature of their occupations (Bibbins-Domingo, 2020). Chu et al. (2020) state that racial minorities and poverty are closely intertwined. In most developed countries, systemic racism has subjected racial minority communities to poor living conditions where people have to live in crowded spaces, rendering containment policies like lockdowns and social distancing impossible.

Adebiyi al. (2021) point out that whereas containment measures like lockdowns were necessary, they disproportionately affected the marginalized communities occupying the low economic class. The lockdown meant the loss of income for unskilled and semi-skilled workers in non-essential sectors and increased contact with the virus for those working in the essential sector. In both cases, such minority families are bound to face both immediate and long-term effects of the Pandemic and lockdown. Those who have lost income are exposed to factors like food insecurity and unaffordability of health for other chronic conditions. On the other hand, the cohort that manages to retain their job is at a high risk of exposure and transmission of the virus.

Carreras et al. (2021) explain that minority groups are likely to oppose the lockdown requirement as staying at home means a loss of income. Even though the danger posed by the Pandemic is well understood, governments across the globe seem to favour a one-size-fits-all containment approach like lockdown without paying much attention to the implications of such policies to marginalized communities. Persistent racism in the labour market has often forced minorities to take up non-structural contracts. The situation is exacerbated by structural inequalities in housing and health services, forcing this cohort into poor living conditions with limited access to quality health care. For these reasons, policies like lockdowns place this population disadvantaged as they lack job security and access to quality healthcare. This situation is made worse by poor living conditions (Hooijer & King, 2021). Bhanot et al. (2021) argue that the lockdowns have inspired the idea of 'othering' as prejudice along racial, religious, and class lines is widening.

Romano et al. (2021) take a different perspective in examining the relationship between racial minorities and COVID-19 containment measures. The authors explain that governments across the globe followed scientists' advice that social distancing could significantly curb the spread of the virus. However, this finding could be ineffective when

aspects like air quality are not given consideration. As mentioned in most of the reviewed literature, minorities are marginalized in most societies, and thus, they live and work in poor conditions characterized by pollution and congestion. As such, enforcing policies that require people to remain indoors under such conditions could fail to lead to the intended results as prevailing conditions at home increase infection and transmission rates, implying the governments' skewed perspectives when formulating and implementing policies.

Additionally, studies conducted in different parts have confirmed resistance and low compliance with COVID-19 containment measures like lockdowns. For instance, in Brazil, Rocha et al. (2021) noted that the region's SVI primarily determined compliance rate as high SVI areas reported high compliance while low SVI areas had low overall compliance. Similar patterns were observed in the UK and Peru, where Patel et al. (2020) and Carreras et al. (2021) link poor neighbourhoods with low containment compliance rates. Since minority communities make up a good proportion of the population living in low socioeconomic regions, it can be argued that compliance is significantly low among minorities compared to the majority. Low compliance can be attributed to socioeconomic constraints such as loss of income linked to the stay-at-home order. This point is well-supported by Tan (2021), who argues that racial and ethnic minorities bear the highest burden of the Pandemic and the immediate and long-term impact of adopted containment measures.

2.10.1 Impact on Children and Teens

Significant psychological and societal impacts resulted from the COVID-19 epidemic. According to research, those most at risk of getting post-traumatic stress disorder, anxiety, depression, and other indicators of distress are children, college students, and healthcare professionals (Saladino et al., 2020). Social isolation and precautions affected people's relationships with one another and overall perceptions of the empathy of others around them. Telepsychology and other technical tools, as seen from this vantage point, were crucial in

mitigating the pandemic's impact. Online psychiatric treatment for patients may be enhanced by the availability of these tools, which allow for meetings in more convenient locations, such as the patient's home or the therapist's office, reduce expenses, and maintain the therapeutic alliance between the two parties. Students of all ages were affected by the quick emergence of the COVID-19 epidemic on such a massive scale. It was anticipated that the ongoing spread of the disease, travel restrictions, and the nationwide shutdown of educational institutions would substantially impact students' education, social life, and mental health (Saladino et al., 2020). Due to the COVID-19 pandemic, students from less wealthy households faced greater adverse effects, such as reduced family money, restricted access to digital tools, and expensive internet connectivity.

2.10.2 Workers

Despite their vital role in the global food supply chain, millions of agricultural workers endure extreme levels of poverty, hunger, ill health on the job, a lack of safety and labor protections, and various forms of abuse. Due to low and inconsistent compensation and a lack of social aid, many are obliged to work longer hours, frequently in risky circumstances, further endangering their families and themselves (Asmundson et al., 2020). Migrant agricultural workers are particularly vulnerable because of the dangers they experience while traveling, working, and living overseas and are not eligible for government-sponsored aid programs.

2.10.3 Adults

Women have been economically impacted more severely by the pandemic due to their disproportionate representation in sectors heavily affected by COVID-19, such as lodging and food services, and in frontline employment, such as the health and social services sectors. Due to the school cancellations, parents also had to assume extra caregiving tasks at home. Whether through the loss of jobs or the closing of schools, the pandemic posed a danger to

decades of progress in gender equality (Asmundson et al., 2020). While hospitals were filled with COVID patients, another illness was rife in many residences. Many were separated from their families as a result of lockdowns and quarantine, and the uncertainty of the economy caused many to struggle with mental health. Increased substance use, increased suicide ideation, and reduced mental health outcomes were reported by young adults, people of color, frontline workers, and unpaid adult caregivers (Asmundson et al., 2020).

2.10.4 *Elderly*

Global lockdown and preventative precautions have isolated individuals, impacted the global economy, and restricted access to physical and mental health care. Although these efforts may have been required to limit the virus's transmission, the detrimental bodily, psychological, and societal repercussions were clear. As a result, technology has been developed to offset these consequences by providing digital substitutes for many daily tasks that could no longer be performed naturally. Martins Van Jaarsveld (2020) discovered that the older population was the most adversely impacted by the virus and the lockdown precautions and reaped the fewest advantages from the digital solutions.

2.10.5 *Ethnic Minority Groups*

Ethnic minority communities were seen as especially susceptible to the consequences of these laws. Densely populated urban areas, more likely to have a wide variety of cultures and peoples, were the most likely to institute local lockdowns. Due to the disproportionate nature of the lockdown's economic and social effects, inequality may have worsened in certain places. Those who had to keep going to work or who lived in households with multiple adults and children saw no advantage from lockdowns (Kirby, 2020). Consequently, persons of color were less likely to benefit from lockdowns and more likely to be negatively affected by them. Instances of local lockdowns were associated with elevated rates of racial prejudice and stigma. Because of this, the infected and their social networks came to be held responsible for

the spread of COVID-19. As a result, members of ethnic minorities were at a higher risk of being found to be infected with COVID-19. Experts in the field noted that local lockdowns exacerbated racial tensions and posed a threat to social cohesiveness, especially when they were misinterpreted as targeting a certain group of people based on their race or religion (Kirby 2020).

2.11 Emoticons and Emojis in Social Media

No one can deny the power of graphical representations and expressions, from tiny expressive symbols that used to be inserted for fun to essential parts of any of nowadays' conversations or social media posts.

2.11.1 Emoticons

When compared to face-to-face type of communication, online communication is said to lack non-verbal cues that could provide additional contextual data on utterances, such as the emotional state or intent of the speaker (Pavalanathan and Eisenstein, 2015). To fill the void, the researchers claimed that different orthographic features such as emoticons, non-standard punctuations, emojis, and expressive lengthening are widely used on social media platforms such as Twitter, Facebook, and Instagram. Emoticons such as “:D” commonly contain typographical symbols and are a typical representation of facial expressions (Hogenboom et al., 2013). As stated by the researchers, emoticons can be read normally “(^_^)” or sideways “;-).” Emoticons were first introduced to individuals by Professor Fahlman, who used them in a message on September 19, 1982, and it did not take time before the emoticon phenomenon spread to the larger communities (Hogenboom et al., 2013). Researchers such as Pavalanathan and Eisenstein (2015) categorized the linguistic functions of these emoticons into three; to indicate an illocutionary force, to conventionally map onto facial expressions such as *joking* as a non-emotional indicator, and to directly map onto facial expressions such as *happy* for emotional meaning.

Researchers such as Childers and Houston (1984) and Shepard (1967) have demonstrated that humans are particularly influenced by the use of non-verbal cues in face-to-face communication. Non-verbal cues, as stated by Burgoon et al. (1989), dominate verbal cues when it comes to face-to-face communication in a case where both cues have equal strength. Seemingly, non-verbal cues are considered significant for individuals so as to understand the emotions and intentions of the person they are establishing a conversation. Translating the findings to a computer-mediated type of communication, therefore, does not seem far-fetched, were it not for the fact that plain-text mediated computer-mediated communication fails to leave room for non-verbal cues. However, users of the computer-mediated type of communication have discovered methods of overcoming the lack of personal contact by applying emoticons to their communication. In his message, Professor Fahlman proposed to use “:-)” and “:-)” to differentiate serious matters from jokes, respectively. Following its spread, individuals began sending kisses, hugs, and yells with the use of graphical symbols whose characters were found on a normal keyboard.

A decade later, individuals discovered ways of using emoticons in everyday computer-mediated communication, which had then translated into the Web’s paralanguage. At that time, Rezabek and Cochenour (1998) stated that 6.1% of the messages were on electronic mailing lists, and 13.2% were on UseNet newsgroup posts (Witmer and Katzman, 1997), which fully contained emoticons. Therefore, non-verbal cues are said to have emerged from computer-mediated communication. However, the cues are conceptually dissimilar from nonverbal cues in a face-to-face type of communication; often, cues such as weeping and laughing are regarded as an involuntary method of expression in face-to-face communication, whilst the use of their respective equivalents “:-)” and “:-)” in a computer-mediated type of communication is considered intentional (Kendon, 1987). By their very nature, emoticons allow individuals to express a subtle mood change, jokes, sarcasm, to signal irony, and to

disambiguate, stress, or show their (intended) sentiment, perhaps even more than communication in a face-to-face setting can through non-verbal cues. Due to this, harvesting data from emoticons appears to be a more than viable strategy to enhance state-of-the-art sentiment analysis. Yet, the question does not dwell much on whether rather than how researchers ought to account for emoticons when assessing a text for the sentiment.

2.11.2 Emojis

Often defined as successors of emoticons, emojis appear to be replacing emoticons on social media platforms such as Twitter (Pavalanathan and Eisenstein, 2015). Miller et al. (2016) describe emojis as pictographs or “picture characters” originating from mobile phones in Japan in the late 1990s that became recognized worldwide in the text-based type of communication following the introduction of smartphones that supposed the input and functioning of Emoji characters. A difference noted from emoticons that are built from ASCII character sequences, the most commonly-applied emojis are a representation of Unicode characters (Miller et al., 2016), and following the introduction of new characters in the new Unicode version, the number of new emojis is rising. Presently, there are 1791 emojis in the Unicode Standard (Unicode, n, d). For each emoji, a name and a code are provided by the Unicode Consortium. These emojis not only contain faces but incorporate ideas and concepts which, among others, are mainly used in social media sharing (For instance, emojis have been discovered to occur more in posted tweets), advertising, smartphone texting (Messages on the IOS of Apple have undergone a major “emojification.”

Kelly and Watts (2015) believe that in addition to the expression of emoticons, emojis can also be used for different purposes in mediated textual communication, such as creating a secret and shared uniqueness within a certain association, maintaining a conversational connection, and allowing a playful interaction. Felbo et al. (2017) and Hu et al. (2017) claim that the extensive use of emojis has also drawn increasing attention from researchers since

emojis convey sentimental and fruitful semantical data to visually complement the textual information that is important in comprehending the embedded emotional signal in various texts (Barbieri et al., 2016). For instance, the embedding of an emoji is proposed by Eisner et al. (2016) or Hu et al. (2017) to understand the semantics behind the emojis; the vectors used in the embedding can be useful in the visualization and prediction of the usage of emojis given their corresponding contexts. Previous researchers such as Felbo et al. (2017) showed that it was fundamental to pre-train a deep neural network on the task of predicting an emoji with pre-trained embeddings so as to learn the emotional signal of these emojis for other tasks such as sarcasm, emotion, and sentiment prediction. Nonetheless, previous works of literature have lacked consideration of the diversity and linguistic complexities of these emojis.

As a result, previous methods of embedding emojis have failed to handle the situation once the sentiments or semantics of the learned embedding of emoji depict contradiction with the information obtained from the corresponding contexts (Hu et al., 2017) or when the emojis display numerous senses of sentiments and semantics. In practice, emojis can either serve the purposes of emphasizing or summarizing the original tune of their contexts or as an expression of more complex semantics, such as sarcasm and irony, by being merged with contexts of contradictory sentiments or semantics.

2.11.3 Emojis/Emoticons and Sentiment

To assess emoticons and emojis in automated sentiment analysis, Hogenboom et al. (2013) suggested the need to first assess how emoticons and emojis are typically associated with the sentiment of any text in which they appear. Insights on the segments are impacted by emoticons and emojis in ways it appears important for progressing to the state-of-the-art sentiment analysis is significant in acquiring data from emojis and emoticons. Hogenboom et al. (2013) attempted to analyze the role of emoticons in relaying the sentiment of a text by

performing a qualitative analysis of 2,080 Dutch forum messages and tweets. First, the researchers hypothesized that emoticons possessed a rather local impact; that is, they impacted either a sentence or a paragraph. Typically, paragraphs address various perspectives for single or various topics, therefore, making it unlikely to apply an emoticon in a single paragraph or many others. In their sample collection and upon inspection, it was discovered that, in general, emoticons possessed a paragraph-level impact for paragraphs that only contained a single emoticon.

In an instance where a paragraph has various emoticons, the sample of the researchers generally showed that an emoticon was likely to impact a sentence in which it was present. Interestingly, 84% of the emoticons would be easily replaced at the last section of a paragraph, 9.0% were placed in the middle section of a paragraph, and 7.0% were included at the start of a paragraph (Hogenboom et al., 2013). From this, it can be concluded that the positioning of an emoticon is a clear indication that it cannot be categorized as a single word but rather as a segment of a text that impacts meaning.

Table 1

Instances of How Emoticons are used to Relay Sentiments (Hogenboom et al., 2013)

Sentence	How	Sentiment
<i>I love my work :-D</i>	Intensification	Positive
<i>The movie was bad :-D</i>	Negation	Positive
<i>:-D I got a promotion</i>	Only sentiment	Positive
<i>-- I love my work</i>	Negation	Negative
<i>The movie was bad --</i>	Intensification	Negative
<i>I got a promotion --</i>	Only sentiment	Negative

In addition to the findings, the researchers discovered that in cases where an emoticon is placed in the middle section of a paragraph with different emoticons, statistically, there is a likelihood for it to be linked to the previous segment of the text. In attempting to look into how an emoticon impacted a text, and discovered that, in general, an emoticon impacts a

sentence in three distinct ways; first, emoticons can be useful in disambiguating a sentiment, for example, in a case where the sentiment related to the sentiment-carrying word(s) need to be negated. Second, emoticons can be useful in expressing sentiment when sentiment is not relayed by a clear negative or positive word in the segment of a text, therefore, allowing the emoticon to carry the only sentiment in a sentence. Finally, an emoticon can stress sentiment by boosting sentiment that is already showing the sentiment-carrying word. Figure 3 indicates that the sentiment-related sentence can be distinct once an individual uses various emoticons, such as the “- -” emoticon to indicate extreme disagreement or boredom or the “:-D” emoticon to express happiness, regardless of the emoticon’s position.

According to Navigli (2009), the sentiment an emoticon carries is less dependent on its embedding text, thus making the word detect a disambiguation technique which is less significant for emoticons. In that respect, the sentiment of an emoticon in a sentence is likely to appear dominant with the sentiment through verbal cues, if any. In different cases, this might appear as a crucial property that can be exploited by approaches of automated sentiment analysis. For example, when an emoticon is the only cue in a sentiment-conveying sentence, typically, one is dealing with a phenomenon referred to as a factual sentiment. An instance is given by the sentence *I got a promotion* which only states the fact that an individual secured a promotion. However, securing a promotion is normally associated with positive emotions such as pride or happiness. Thus, a human interpreter is typically inclined to acknowledge the suggested sentiment and, as a result, take into consideration that the factual statement is positive. However, this requires a comprehension of context and mainly involves integrating knowledge of the real world into the entire process of sentiment analysis. For machines, Hogenboom et al. (2013) consider this a cumbersome task. In this regard, emoticons can be valuable cues for deriving the intended sentiment of an individual.

With the nonstop development of the Internet of Things (IoT), the growing popularity and accessibility of subjective content have offered new challenges and opportunities for sentiment analysis (Pavalanathan and Eisenstein, 2015). For instance, social media platforms such as Instagram and Twitter have been assessed for the massive content generated from user sentiment (Agarwal et al., 2011; Barbieri et al., 2016; Pak and Paroubek, 2010), where emojis are extensively utilized. Non-verbal cues of sentiment, such as emoticons which are defined as the previous generation of emojis, have been investigated for their sentiment impact prior to emojis taking over (Yuan et al., 2015). For instance, Hogenboom et al. (2013) predefined labels for emoticons and constructed an emoticon-sentiment dictionary. Liu et al. (2012) applied emoticons and emojis with the aim of soothing noisy sentiment labels. Similar works from Kralj et al. (2015) considered emojis as a component in the extraction of the lexical feature for more sentiment analysis. The researchers also constructed an emoji sentiment ranking that was based on emoji occurrences and sentiments annotated by humans of corresponding tweets where emojis were assigned with a sentiment score from negative to positive. However, the relatively intuitive usage of emojis by dictionary- and lexical-based approaches lacked an insightful comprehension of the complex semantics of an emoji.

Following the inspiration of success from the algorithms embedding semantic words such as Mikolov et al. (2013), the semantic embeddings of any emoji can be obtained by averaging the words from its descriptions and displays its effectiveness to take advantage of the embedding for Twitter Sentiment Analysis. Li et al. (2017) proposed a convolutional neural network for the prediction of emoji occurrences and jointly learned the embedding of emojis through a matching layer that is based on cosine similarities. Although the popularity of Twitter sentiment analysis continues to grow, there appears to be a limited number of emoji and emoticons datasets with sentiment labels made available since previous studies filter out emojis, URLs, and, at times, emoticons. Felbo et al. (2017), however, showed that it

remains effective to extract sentiment data from emojis for sarcasm detection and emotion classification tasks in the absence of learning vector-based emojis represented by pre-training a deep neural network for the prediction of emoji occurrence.

2.11.3.1 Emotions and Cognitive Reappraisal in Lockdown Compliance. According to Brooks et al. (2020), quarantine and lockdowns are stressful. In fact, Castellano-Tejedor et al. (2022) believe that such situations can be classified as fear- and anxiety-inducing experiences. Studies from previous health crises show that worry and anxiety that have been induced by diseases are predictors of greater behavioral compliance within measures of public health. Leung et al. (2005) discussed the outbreak of SARS in Hongkong in 2003 and discovered that anxiety was linked to a greater following of precautionary measures such as the possession of a facemask. Bults et al. (2011) stated that a few months after the Influenza A (H1N1) outbreak hit the Netherlands in 2009, compliance with suggested public health measures such as mask possession and wearing and social distancing, and anxiety about the virus declined in a parallel manner; a suggestion of a possible association between the factors. Harper et al. (2021) looked into the Coronavirus pandemic and discovered that in its early stages, adults in Britain with increased fear and anxiety of the virus reported even greater compliance with the regulations. In the U.S.A, adolescents presenting greater symptoms of anxiety took part in more COVID-19 pro-social acts, as claimed by Alvis et al. (2020). Conversely, adults presented more negative emotions as a result of the disease's related compliance, especially the mitigation measures (Van Rooji et al., 2020).

Contrastingly, Hania and Hongbo (2020) detected an even stronger association between behavioral compliance and COVID-19 fear in Pakistan. Taken as one, the results suggest that specific anxiety arising from the spread of COVID-19, worry, and emotions, including fear and anger, play a significant role in the motivation of compliance, behaviorally, with the health measures that have been created to curb the spread of the virus. Behavioral compliance

is also associated with the manner in which individuals choose to deal with their emotional experiences. According to Gross (1998), the regulation of emotions is the ability of any individual to shape when they experience it, what they experience, and how they experience emotion and express it. Particularly, Gross and John (2003) define cognitive appraisal as a strategy in the regulation of emotions involving the active re-analysis of emotional situations or stimuli with the aim of altering the emotional response in regard to one's goals. Research on emotional regulation, such as one conducted by Mauss et al. (2007), persistently discovered positive outcomes resulting from the use of reappraisal. As identified by the researchers, taking part in a successful reappraisal is linked with negative emotion, while the use of habitual reappraisal is linked to fewer psychopathological symptoms (Aldao and Nolen-Hoeksema, 2012). Nonetheless, present researchers such as Ford et al. (2019) highlight the significance of considering contexts where appraisal is applied and yields results.

In a context where negative emotions are considered key in motivating behaviors consistent with objectives, the use of appraisal to decrease any negative emotion is assumed to lead to a trade-off between the long-term instrumental objectives and short-term psychological well-being. Ford et al. (2019) issued an example; when Hillary Clinton lost the presidential elections in 2016, voters that used reappraisal to reduce their negative emotions displayed less likelihood of subsequently engaging in political action. This means that effective reappraisal was impactful in the short term as it restored the emotional well-being of the voters following a distressing event; however, the cost of attaining long-term objectives was subject to reducing the engagement in subsequent behaviors that were goal-consistent. An initial study by Wang et al. (2020) suggested that the trade-off was likely to extend to the mitigation measures of the pandemic. The researchers believed that individuals appraising the threat of the virus with less severity report lower psychological distress, less negative

emotions, and better mental health. Comparably, Jungmann and Witthöft (2020) discovered that appraisal use and the acceptance of strategies related to emotion regulation exhibited a correlation with low COVID-19-related anxiety. The appraisal associated with the virus appeared to affect the long-term engagement with mitigation measures. For instance, the view that the virus was less severe led Briki and Dagot (2020) to link behavioral compliance with hygiene measures and social distancing, whilst an appraisal of the disease as more severe exhibited a likelihood of engaging in disinfecting behaviors and social distancing (Oosterhoff and Palmer, 2020).

A study by Smith et al. (2021) of adults in the USA for over three months discovered that, subsequently, reappraisal predicted lower fear of the disease, which, as a result, predicted fewer efforts to mitigate the disease. When combined, the studies show that despite the engaging of individuals in appraisal and the associated decrease in negative emotions that surrounded the virus to restore emotional well-being, taking part in an appraisal increases the risks of a trade-off with the instrumental objective of being part of significant mitigation strategies. It is important for researchers to note that research exploring the effect of emotions and emotional regulation on behavioral compliance remains limited. Mitigation and response measures for COVID-19 have been reinforced inconsistently across the globe resulting in the limitation of the scope of mitigation behaviors that are based on less recommended behaviors by aid agencies, such as maintaining social distance, lockdowns, and quarantines, the wearing of face masks and practicing proper hygiene. The Arabic context offers an alternative for research where lockdown measures were enforced and followed by consistent and clear government messages that promoted engagement with mitigation strategies.

2.11.3.2 Arabic Sentiment Analysis during a Pandemic. Alayba et al. (2017) carried out a study with the aim of exploring the opinions of individuals on health services by gathering and assessing a dataset size of 2,026 Arabic Tweets. The researchers classified the

tweets into negative or positive classes and applied a collection of bigram and unigram approaches with the TF-IDF method to weigh the features in the corpus. The first 1000 maximum weighted features were put into the ML algorithm. For the classification process, exemplary results were achieved with an accuracy of 91% with the SVM applying the Linear Support Vector Classification and Stochastic Gradient Descent (SGD). With the same dataset, Aloqaily et al. (2020) implemented sentiment analysis with a distinct objective and classification model on 2000 Arabic tweets on the Syrian civil war and crisis. The tweets were gathered in 2014. The researchers categorized the tweets into either negative or positive classes, and the measure of Information Gain was applied to harvest and select features. Aloqaily et al. adopted a lexicon-based approach and ML algorithms and later compared their performance. For the lexicon-based methodology, they used the Arabic Emotion Lexicon to categorize the tweets and expressed them as BOW in two labels; negative and positive.

The above process was conducted through a lookup process where each word was identified in the Arabic Lexicon. When a word misses in the lexicon, it is considered neutral and awarded a zero. The total score of the tweet was then established, and based on such; the tweet was accorded a label that was either negative or positive. For the ML technique, the researchers applied different algorithms. The researchers then concluded that the approach based on algorithms on ML performed better than the approach based on lexicons in the prediction of the tweets' subjectivity. Regarding ML techniques, the Logistic Model Trees (LMT) algorithm registered an exemplary accuracy of 85.5% using the topmost 150 features. In a similar study, Aljarah et al. (2020) integrated ML approaches to acquire online hate speech in a dataset size of 1,633 Arabic Tweets. The tweets were labeled negative if they failed to register hate in the speech and were labeled positive if hate speech was detected. In extracting features, the researchers implemented a combination of TF-IDF and BOW. It can be noted that in the word features, Aljarah et al. considered profile-related and emotion

features such as ‘favorite-’ and retweet-count.’ Similarly, various ML algorithms were utilized for categorization, and the researchers achieved the conclusion that the highest accuracy is acquired using Random Forest (RF), with a 91% accuracy.

Al-Rubaiee et al. (2016) conducted a study to categorize Arabic tweets on whether they contained neutral, negative, and positive opinions on e-learning. The researchers began by gathering tweets and conducted a pre-processing that led to the collection of 1,121 tweets. The researchers later manually categorized the tweets as neutral, negative, or positive and applied Naïve Bayes and SVM to categorize the tweets using the TF-IDF weighting scheme together with N-grams for the extraction of features. Later, the researchers conducted two experiments with various classes. The first experiment had negative and positive classes, while the second one had neutral, negative, and positive categories. The researchers declared SVM as the best-performing model since it registered an accuracy score of 84.84% in the first experiment and 73.15% in the second one. Almouzzini and Alageel (2019) aimed to explore depressive emotions from the Gulf region using Arabic Twitter data. Initially, the researchers gathered tweets from 97 users, whereby 35 exhibited depressive behaviors. After pre-processing, Almouzzini and Alageel annotated the remaining data manually, which consisted of 6,122 tweets regarding the tools used in analyzing the depression scale.

Of the 6,122 tweets, 1,363 tweets displayed non-depressive thoughts while 1,359 tweets were categorized as depressed, and the remaining 3400 were categorized as ‘other’ and, therefore, eliminated. Subsequently, the applied BOW was used as an extraction feature with the handling of negation since negation was a significant factor in assessing the sentiment of the tweet. Therefore, the researchers considered a few modifications to the filter to improve compatibility in Arabic tweets by adding a list of negated Arabic words. Lastly, the researchers used Weka to conduct supervised ML using the following classifiers; Liblinear, AdaBoostMI, Naïve Bayes, and RF (Fan et al., 2008); an open-source packaging supporting

linear SVM and offered effective performance for the classification of tweets in large scale. Findings showed an 87.5% accuracy with the Liblinear classifier. Researchers such as Baker et al. (2020) proposed sentiment analysis on 54,065 Arabic posts that related to Influenza and labeled data as invalid or valid from their association with Influenza. The researcher used various ML techniques such as Naïve Bayes and SVM, Decision Tree (DT), and K-nearest neighbor (k-NN). The results from the three experiments showed that the NB algorithm presented the best accuracy of 83.20%. Among Arabic nations, Syria offered 89.06%, the highest value with the NB algorithm, while the Arabian Gulf region presented 86.43% using the k-NN. The SVM exhibited poor performance in all the experiments.

Alhajji et al. (2020) carried out research to investigate the attitude of the citizens of Saudi Arabia toward the Coronavirus pandemic preventive measures. The researchers applied Arabic sentiment analysis to posted tweets. The NB classifier was applied to 53,127 tweets that were gathered on seven health measures; 21-day curfew, suspension of Friday prayers, suspension of all sports activities, closure of restaurants, public parks, and malls, the closure of universities and schools, the lockdown of Qatif, and the closure of great mosques. Researchers concluded that the Twitter users of Saudi Arabia were willing to apply health measures and have positive opinions on announcements made by the government. Religious practices equally played a significant role in improving the overall positive sentiment.

2.11.4 Potentially Promising Techniques to Satisfy Hidden Annotation

Several techniques have been suggested to satisfy the hidden annotations on social media platforms and especially Twitter. Ritter et al. (2011) suggested the Named Entity Recognition (NER) technique, an approach used in the identification and categorization of entities such as locations, firms, and individuals. The researchers discovered and demonstrated that among the existing tools for tagging, NER performed poorly when applied to posts on Twitter. To address the named challenge, the researchers annotated tweets and

constructed tools that were trained on unlabeled, out-of- and in-domain data and showed significant improvement over the state-of-the-art counterparts with equal training. Similarly, the researchers identified Named Entity Classification (NEC) as a challenging task on social media platforms, especially Twitter. As a result of their terse nature, often, tweets and posts lack complete context for the identification of the various entities they contain. Additionally, numerous distinctive named entities were presented and necessitated large amounts of training data. According to the researchers, since capitalization on Twitter is considered less informative than local news, in-domain data is required for training models that depend less on capitalization and are also capable of using features provided by T-CAP.

Parts-of-Speech Tagging (POS) is a technique effective in identifying parts of a speech, such as an adjective, verb, and noun in a Tweet. The technique is considered applicable to a wider range of tasks in NLP, including the extraction of information and NES (Named Entity Segmentation). Prior researchers suggest that this tagging contains a strong baseline; the assigning of a phrase to where it is frequently tagged and assigning each Out of Vocabulary (OOV) word to the most commonly tagged in POS. According to Charniak et al. (1993), the obtained baseline was a 0.9 accuracy on the Brown-corpus. However, a similar application of the same baseline on tweets obtains a similar weakness that exposes the challenging nature of data obtained from Twitter. The most suitable explanation for the accuracy decrease is that Twitter contains more OOV words than any grammatical text. It was also observed that a large number of these OOV words came from spelling variations. Despite NNP being the most frequently tagged for OOV words, approximately a third were NNPs. A similar challenge identified on Twitter by the researchers is the performance of trained off-the-shelf news where POS taggers were identified. As discovered, the POS tagger on state-of-the-art Stanford enhances the baseline and therefore obtains a better accuracy. As such, the performance is considered impressive with its training nature.

Event extraction is another technique that can be useful in identifying and extracting events and associations from collected Twitter data. Ritter et al. (2012) presented an open-domain and scalable technique in the extraction and categorization of events from Tweets and status messages. The researchers manually assessed the quality of the events that exhibited a clear performance improvement. According to research, posts from Twitter are the most inclusive and up-to-date streams of commentary and data on current events; however, they have been discovered to be equally noisy and fragmented, and therefore, motivate the need for developing or adopting systems that can harvest, aggregate and classify significant events. Previous researchers on the extraction of structured representations of Twitter events have largely concentrated on newswire texts, that is, the unique characteristics presented on Twitter by new opportunities and challenges in the extracting of open domain events. The researchers demonstrated the accurate extraction of such a calendar of important events from Twitter posts and considered it feasible. Additionally, the researchers presented a novel approach for the discovery of significant event classification and categorized extracted events that were based on latent variable models. Through the leveraging of large unlabeled Twitter data volumes, the approach presented achieved a 14% increase over a supervised baseline. These techniques can also be combined and used in different ways to meet specific annotation requirements for Twitter data.

2.12 Applications of Sentiment Analysis in Different Contexts

Sentiment analysis has been heavily applied in many forms and shaped and utilized in various applications and fields. It is certainly not a novel application, but most definitely, an application that has not been fully discovered yet.

2.12.1 Sentiment Detection Approaches

Techniques suggested for sentiment analysis are classified into ones that employ algorithms of machine learning (ML) and are operating under supervised methodologies or

statistically-inspired methods under an unsupervised instance. Unsupervised methodologies deduce the sentiment of the document by exploiting the statistical properties of the word and/or leveraging the existing lexicons that possess emotional or polarized words, whereas supervised methods derive models that classify sentiment. The lexicons presented in the unsupervised methods can also be used to extract the representation features of a document considered useful in supervised methodologies of classification.

2.12.1.1 Supervised Machine Learning Approaches. Out of a textual resource, sentiment analysis is considered an issue that arises in the classification of documents and is aimed at separating a text or a document that expresses both negative and positive sentiments (at times, there is a consideration of neutral standings) through the exploitation of various linguistic and syntactic features. In recent times, relatively few more complex techniques have been suggested and mostly towards the identification of more refined emotions and not mere polarity (such as emotional analysis). According to Chatzakou et al. (2013), effective analysis or emotion analysis is considered a clarified form of sentiment analysis as it aims to uncover a detailed classification of either documents or texts centered on the emotions they show. The two versions of the presented challenges (emotion and sentiment analysis) follow similar methods with an existing suitably categorized collection of documents. This annotation can either be valued, for instance, 0.8 fear or 0.5 anger, or nominal, for instance, anger or positive. Pang et al. (2002) are acknowledged as the first pioneers of sentiment analysis that employed a machine learning approach with the aim of determining whether the written review is either negative or positive. The researchers experiment with the following three algorithms for the same purpose; Maximum Entropy, Naïve Bayes, and SVM, under

different parameters and features, such as bigrams vs. unigrams and feature presence vs. feature frequency.

Socher et al. (2011) suggested a framework of semi-supervised approaches based on the Recursive Auto-Encoders (RAE) that would be able to predict the distribution on a multi-dimensional level of any emotion at a sentence level. Zhang et al. (2011) also suggested a hybrid method, one where an improved lexicon-based methodology is first applied to data harvested from Twitter to conduct sentiment analysis on an entity level. After which, a binary classifier picks the results of the first steps and is formatted to allocate polarities of sentiments to opinionated tweets. In 2006, Cui et al. discovered the discriminative classifier's effectiveness, especially SVM, using high-order N-grams as features for the classification task of binary sentiments. Severyn and Moschitti (2015), initialized with the help of an unsupervised NL model, based their research on word embeddings and investigated a deep Convolutional Neural Network for the classification of sentiment. Ren et al. (2016) used Latent Dirichlet Allocation to acquire the topic distribution for sentences in the used dataset and later exploited the Recursive Auto-Encoders in learning word embeddings enhanced in a topic. For further improvement in performance, the researchers integrated the representation with conventional models such as logistic regression and SVM.

Approaches to ensemble these activities are proposed by Mesnil et al. (2015) and Wang et al. (2015). To get more detailed, the study by Mesnil et al. focused on blending discriminative and generative models (Recurrent Neural Networks and Naïve Bayes) for the prediction of sentiment. The log probabilities of the presented models are integrated through linear interpolation to harvest the final sentiments in the assignment. On the other hand, the work of Wang et al. suggested a new Random Subspace methodology named POS-RS for the classification of sentiments based on the analysis of the parts of speech that applies both function lexicon and content lexicon subspace rates to manage the complexity of the base

learners. As identified, supervised methodologies consider the classification of sentiment as a standard statistical classification containing a large number of labeled examples. There appear various widely used classification algorithms; the rule-based algorithm is utilized as a classification scheme that focuses on the use of IF-THEN rules to predict classes. Asghar et al. (2017) assessed the semantics of words by classifying the words into negative and positive classes with the aim of identifying and classifying domain-specific, general-purpose, modifiers words that were expressed by the public that soled the challenge of previous studies being less effective from the sparsity of data. Sentiment analysis that enhances sentiment and is based on rules is considered an effective alternative for bettering the classification of sentiment.

Deep learning is a section of the larger family of ML methodologies based on representations of data learning compared to algorithms based on tasks. Learning can either be unsupervised, supervised, or partially supervised. Additionally, deep learning (DL) has also been proven to play a significant role in the analysis of Big Data. Until now, various DL models have been integrated for the analysis of sentiments. Examples are; deep convolution neural networks, deep neural networks, and deep belief. Tang et al. (2014a) introduced a DL system named the Coooolll in the classification of sentiments based on the message level on Twitter. The system is constructed on a supervised learning framework by linking the features of sentiment-specific word embedding (SSWE), where the features are trained from Twitter on a large scale, gathered by both negative and positive emotions, without the application of manual annotation. The benefit of using this system is that it could effortlessly be re-integrated with the SSWE available in the public domain. As it is a fact, most existing algorithms are modeled to represent words continuously and, typically, only model the syntactic models; however, they ignore the text's sentiment.

Tang et al. (2014b) also suggested a methodology of learning on the classification of Twitter sentiment that addressed the presented issues by learning SSWE discovered to encode existing information on sentiment in a continuous representation of words. Additionally, the scholars also developed three neural networks to combine effective supervision from the polarities of the sentiment of the texts in loss functions.

2.12.1.2 Unsupervised Learning Methods. These approaches take a significantly different direction to the process of analyzing sentiment as they fail to make use of labeled documents. They, however, depend on the statistical properties of the documents, for instance, word cooccurrence existing lexicons and NLP processes with the words having a polarized or emotional orientation. Turney (2002) went ahead and presented an unsupervised sentiment analysis approach that categorized the social reviews as either positive or negative while establishing the semantic orientation of the phrases through the association of two words, poor or excellent. Lin and He (2009) suggested a methodology that utilized Latent Dirichlet Allocation for identifying the sentiment of a document and, simultaneously, its topic and achieved an accuracy in classification that was similar to the ones presented in supervised approaches. There also appear to be unsupervised methodologies that solely depend on lexicons to estimate the average emotion or sentiments exhibited in the presented document by the corresponding orientation of the words. Often, these approaches begin with a small set of opinionated words with an already identified orientation and, equally, attempt to broaden the set by using a well-identified thesaurus or corpus for synonyms. Similarly, these approaches are likely to exploit the syntactic patterns and structural elements existing in a text by integrating NLP processes such as POS tagging and lemmatization.

Heerschop et al. (2011) attempted to explore how knowledge harvested from the structural aspects of a document could be used to enhance the performance of the process. The researchers tested their hypotheses by discovering the most purposeful sections of a

document and scores based on documents and on word-level sentiment aggregation. Qiu et al. (2010) harvested sentences for opinions that were related to negative sentiments and identified sentence topics with the use of the rule-based methodology that integrated sentiment lexicons and syntactic parsing. The researchers tested their methodology in contextual advertising, that is, the issue of relating advertisements to web pages. Saif et al. (2016) introduced a lexicon-based technique named the *SentiCircles* that was capable of updating the orientation of the sentiments in words by acquiring the latent semantics from the patterns of their co-occurrences. Other researchers such as Carrillo-de-Alboronoz and Plaza (2013) leverage the NLP in identifying linguistic features such as modalities, intensifiers, and lexicons to discover the general sentiment or the emotion of a document under an approach that averages score.

Clustering, a common problem arising in unsupervised methods, is the process of categorizing data into various classes so that objects with similar clusters have an extremely high similarity, whereas ones with dissimilar clusters have an extremely high diversity. From a machine learning perspective, clusters correspond to a hidden pattern, while clustering remains the unsupervised technique of learning for searching clusters. Dissimilar to classification, these methods are independent of a presumed class or training cases presented with labels. Algorithms in clustering automatically determine latent patterns, which is considered the observation of learning. Additional data on clustering algorithms for the analysis of sentiments is discussed. With points in various spaces, the density-based algorithm classifies points that are closely packed, identifying outliers in low-density regions. According to Hassan et al. (2016), there is a chance to develop an analysis framework using distance- and density-based clustering. The suggested framework provided a higher accuracy rate and can be useful in performing various attacks prediction. To contextualize the

framework, the researchers also claimed that it provided an analyst with a feasible method of countering terrorism and related attacks.

The hierarchical algorithm classifies data over different scales by generating a dendrogram or a cluster tree. The tree is a multi-level hierarchy in which clusters are co-joined at a single level and transformed into multiple clusters at the next level. This permits individuals to make decisions on a scale or level of clustering that is considered most effective for the specific application. Wu et al. (2014) created a visual analysis system, OpinionFlow, which is capable of visually exploring and tracing opinion diffusions among the users of Twitter. Most importantly, the suggested system utilizes a hierarchical topic structure constructed by Bayesian Rose Tree (BRT) that identifies a topical multi-branch hierarchy from numerous tweets and presents them with the aid of a stacked topic tree. The tree is related to the flow of opinions to accelerate opinion diffusion analysis on various topics. Chen et al. (2016) suggested a method named the TCOL-Miner that incorporated leadership quality analysis and network structure methodologies structures to effectively discover opinion leaders in a larger social network. According to the researcher, the two-stage clustering methodologies are used to significantly decrease the effect of the overlapping issue in social networking.

Vaziripour et al. (2016) categorized the sentiments associated with single tweets to identify the different sentiments, over time, towards specific trending political topics. According to their findings, the vocabulary's hierarchical representation is created by following up the merging process of clusters that is later used as a feature.

2.12.1.3 Semi-Supervised Learning Methods. The semi-supervised learning methodologies present capabilities in using unlabeled data in the processes of training and improving application classification when the labeled data is scarce. From this point of view, these approaches are of significant practical value; this is because there is the discovery of a

balance between unsupervised and supervised learning. Fundamentally, there is the existence of various classical classifications of the semi-supervised techniques; topic-based, wrapper-based, and graph-based. On the graph-based methodology, semi-supervised learning is capable of managing either a document or a sentence. Lu et al. (2010) suggested a generic framework that incorporated social context data by incorporating the difficulties of regularization to the text-based predictor. Tan et al. (2011) utilized various social network structures to aid in the analysis of sentiments by adding linked information. The links were able to correspond to attention, for instance, when the user of a Twitter account wanted to concentrate more on the status update of another user. Pozzi et al. (2013) also suggested a generic framework that estimated the polarities of users on a particular topic by joining weighted approval relations and the contents of a post that were presumed to be a better representation of the contagion on social networks.

The wrapper-based methodology labels various unlabeled examples with the aid of a decision function that is then learned and iteratively added to the trained data. The well-established forms of wrapped-based techniques include co-training (Yu, 2014) and self-training (Zhao et al., 2014). Self-training begins with a supervised learner who has received training on available labeled data and is later iterated a number of times. Instead of utilizing one supervised learner, co-training utilizes two learners that perform several operations on various feature sets from independent views and, in turn, build more labeled information. Iosifidis and Ntoutsi (2017) showed a presentation on how to categorize collections on a large scale with labels that use co-training and self-learning, stream- and batch-processing. The piece of work solved the problem of the lack of large-scale labeled sets of data over a large period of time, which is extremely significant for streaming mining research. In the topic-based approach, data is acquired from various domains of analyzing sentiments. An example is Si et al. (2013), who used a continuous Dirichlet process mixture model to understand the

daily topic sets from Twitter to aid in the prediction of the markets such as the stock market. Xiang and Zhou (2014) suggested a sentiment mixture model from topics with particular information for the improvement of sentiment analysis on data harvested from Twitter that generated topic data through topic modeling based on LDA implementation.

2.12.2 Application Context of Sentiment Analysis

Applying the process of sentiment analysis has ranged from business and marketing, public action, and health to politics. The process remains unlimited to a single application as it offers a large field of applications in various domains to aid in the process of decision-making. According to Karamollaoğlu et al. (2018), the process of analyzing sentiments can be incorporated into events that have occurred worldwide, such as an activity, disaster, or a sport taking place in the world. A few of the instances are highlighted in a research study where the aim was to compare how individuals from Western and Eastern nations view ISIS. The findings by Mansour (2018) showed the two sides of the globe view the term or group (ISIS) in a similar manner; a terrorist. Similarly, the process permits the creation and raising of awareness on data security and the dangers associated with the breach of security. Also, the process assumes that the function of guidelines in firms is to respond to any breach in their security and, as a result, alter the shaping of perceptions (Hao and Dai, 2016). Moreover, sentiment analysis is also carried out on the rate of unemployment and the sentiment score of the same in social media (Shayaa et al., 2017). It remains evident that the application of the process in the healthcare sector and any area where the process is used as a service framework, spatiotemporal properties are used for the purposes of identifying locations of the outbreak of any disease (Ali et al., 2017).

Ragini et al. (2018) also stated that the analysis of sentiments could be useful in identifying the needs of individuals during a crisis and in preparing an appropriate and effective response to the rescue. Furthermore, according to Hassan et al. (2017), the process

facilitates the discovery of the levels of depression in individuals by overserving and assessing any emotions detected in a text. Joyce and Deng (2016) also claimed that sentiment analysis is useful in predicting political elections, where the researchers displayed that data assessed from Twitter offers more reliability with a 94% correlation in polling data. Also, researchers showed the potential of Twitter to become a platform that is capable of rivaling complex polling approaches. Lastly, feedback from a consumer also plays a significant role in supporting the application of sentiment analysis since it can help organizations and businesses in taking needed action to enhance the quality of their services and products and enhance their business strategy. This is evidenced in a study by Isah et al. (2014), where the researchers concluded on experiences and views of cosmetic and drug products among the users of social networks. According to Martin-Domingo et al. (2019), sentiment analysis equally facilitates detecting areas that require improvement in airport service quality and applying proper corrective measures, such as paying attention to feedback on social media.

Akter and Aziz (2016) also argue that the process enables the assessment of the characteristics and trends of food habits that are useful in business organizations when planning a marketing strategy for products. Sentiment analysis generates numerous advantages for business owners as it aids in the identification of their popularity among their customers and in how clients think about the services or products offered (El Rahman et al., 2019), and in evaluating the capability and effectiveness of the brand's communication (Poecze et al., 2018). Similarly, Suman et al. (2017) believe sentiment analysis assesses the flow of stock prices through social media platforms. Feedback from clients is significant as it aids in the recognition of a company's weakness; this is also highlighted by Ikoro et al. (2018), who compared sentiment data collected from tweets of consumers on Big Six. The findings indicated that sentiments from the perceptions and attitudes of the Big Six were more negative as a new entrant to the energy industry. Additionally, the analysis of

sentiments on social media platforms allows a firm to assess the success levels of any program, as established in a study by Yuliyanti et al. (2017), who established that positive sentiments are mostly acquired from tweets on community development program activities. The findings were declared useful as they could aid in improving the living standards of the community.

2.12.3 *Sentiment and Sentiment Analysis during Crises*

It is undeniably crucial to measure public sentiments during a pandemic or, in general, a crisis. Such a significant application provides authorities and decision-makers with the necessary data and information to build their actions and conclusions.

2.12.3.1 Sentiment during a Crisis. As a popular social media platform, Twitter plays host to numerous users enabling them to express their opinions. Its abundance of user meta- and textual data makes it useful in ML algorithms inputs. Trends and behavioral patterns can also be harvested from the platform to provide potential insight within societies and can also aid in the better understanding and predicting of public attitudes. The sections can aid government bodies in navigating crisis management and, in particular, in identifying new methodologies to follow up on the spread of the COVID-19 pandemic. Various instances have been explored, including contact-tracing phone applications through the use of social media data as a tool (Turiel et al., 2020) and the use of wastewater testing (Larsen and Wigginton, 2020). An unusual feature detected in the pandemic is its asymptomatic nature, which has resulted in numerous community transmissions with the virus remaining undetectable. This provides a motive to explore precautionary lockdowns and other restrictions and specifically through language processing. As discovered, this is likely to

predict public attitudes' evolution amidst the pandemic and remains of utmost significance in an era where there are resources.

As identified, social media platforms have become a significant alternative as an information channel to conventional platforms during disasters and emergencies. As stated by Lindsay (2011), these platforms have been useful in performing numerous tasks, including fundraising for disaster relief and warning individuals about unsafe areas. Social media also aids in keeping individuals informed, locating loved ones, and expressing support or notifying authorities during a disaster or an emergency. Of the diverse areas and the popularity of the discussions on social media, it is now also used as a tool by first responders for crisis management and disaster relief. Systems used for microblogging are used by a large population across the globe with the aim of broadcasting any information. Thus, these microblogs can equally be useful in managing disasters since they provide platforms where information can be relayed and, at the same time, ease accessibility. For instance, Twitter was used to distribute situation updates during major real-time disasters (Starbird et al., 2010; Sutton et al., 2008). Researchers, Qui et al. (2009) focused on the 2008 Sichuan Earthquake to display how much individuals used these platforms to disperse and gather information during crises. Similarly, Qu et al. (2011) conducted research on how microblogging systems would be of use in facilitating disaster response.

Involving social media platforms in managing disasters can roughly be classified into two; information sharing and situational awareness (Sakaki et al., 2013). Information sharing exhibits the behavior of individuals and their tendencies to share information on social platforms during disasters and can also be useful in directing resources to the public. Verma et al. (2011) define situational analysis as the identification, processing, and understanding of the critical elements of a situation or an incident to offer insights into safety- and time-critical situations. The process aids first responders in evaluating the degree of damage and locating

victims and their needs. The two processes aid in speeding responses and alleviating human and property losses in managing crises. The assessment and identification of posts relating to crises (Terpstra et al., 2012), the studying of automatic methods of information extraction (Imran et al., 2013), and the behavioral assessment during crises (Mendoza et al., 2010) include the numerous methods used in improving situational awareness. According to MacEachren et al. (2011) and Kumar et al. (2011), other analysis tools and applications also aid disaster relief organizations and humanitarians in tracking, organizing, and monitoring events related to posts and their sentiments.

It is evident that the impacts brought about by the spread of the Covid-10 pandemic were felt globally, and as a result, social media platforms became inundated with pandemic-related content. From the time of its initial identification in Wuhan, the spread of the virus to numerous counties across the continent has led to the rise of strategies to curb it. Among the strategies are imposed travel bans, quarantine, and lockdown measures restricting movement. Schools were also closed, and numerous employees found themselves unemployed. With the number of individuals that have been impacted by the diseases and the equivalent number of lives destroyed, social media platforms such as Twitter have transformed into a user's outlet that enables the expression of feelings, opinions, and concerns on the pandemic. According to Shannon and Kent (2020), social media platforms have also emerged as important conduits for health-related information since a majority of individuals around the globe use a certain form of social media. Silver et al. (2019) reference various Pew Research surveys and state that social media is a significant source of health information. Recently, the consumption and sharing of precautionary measures through social media platforms are becoming prevalent.

A study by Tri Sakti et al. (2021) discovered Twitter to be the most dominant platform used by the public in discussing large-scale social restrictions (LSSR) and the COVID-19 virus in general. Public sentiments towards social restrictions were mostly positive, with most

of the individuals expressing support for the intervention of the government in combating the disease. Nonetheless, negative sentiments were equally high and especially concerning the indecisiveness of the Indonesian government together with the defiance of the social restrictions. The emotion analysis showed that the citizens of Indonesia were mostly trustful; however, there were notable peaks in fear and sadness that corresponded to the increased death rates of the disease. A closer analysis of the tweets presenting the highest levels of engagement confirmed that most of the citizens were dissatisfied with the parties and the government and, as a result, failed to obey the regulations of the LSSR. Supporting this is the trends equally observed in the analysis. Criticisms of the relaxation of the LSSR and the indecisiveness of the government dominated the trends in discussions during the study. Despite the overall sentiment of the public being positive, both the neutral and the negative sentiments were considerably high, and the mere fact that almost half of the public failed to express a positive attitude towards the implementation of the LSSR and the government is worrisome and, in particular, in a global health crisis with such a magnitude.

Analyses of the overall sentiments surrounding the disease have presented similar findings in societies worldwide. The crisis in the health sector sparks concerns about the spread and containment of the pandemic (Xue et al., 2020), the survival of the economy (El Keshky et al., 2021), and the mental and physical well-being of citizens (Wang et al., 2021). The uncertainty that surrounds the pandemic has equally led to the public expressing mixed reactions. An assessment of the Twitter discussions in the early stages of the disease by Xue et al. (2020) discovered dominant emotions of anger, trust, and fear of death and overall cases surrounding the disease. Another researcher discovered a similar pattern of feelings of fear and anger, hope, solidarity, and trust in Italy (Stella et al., 2020). In Spain, de Las Heras-Pedrosa et al. (2020) claimed that the sentiments of the public were that of sadness, anger, fear, and disgust. It is also imperative to note that while emotions and sentiments exhibited

dominance at the onset of the pandemic, there has also been extreme change and polarity in the sentiments since that time (Wicke and Bolognesi, 2021). As such, researchers believe that the pandemic ought to be an awakening for governments to conclude the consequential purpose of strategic public policy and also acknowledge the significance of well-targeted, consistent, and clear public communication in curbing the crisis.

2.12.3.2 Sentiment Analysis during a Crisis. Analyzing the sentiments of a disaster post on social media platforms is one approach that could aid in improving post-detection for situational awareness. Particularly, it is fundamental to gain a better understanding of the dynamics associated with networking, including the concerns, panics, and feelings of the user, as it is useful in the identification of the sentiment's polarity expressed during crises to improve the process of decision-making. It aids authorities in findings answers to their queries and in making better decisions on event assistance without having to pay the cost of the conventional surveys issued to the public. Sentiment information could equally be useful in projecting the information on the recovery and devastation situation and donation requests to the population in different ways. According to Nagy and Stamberger (2012), using the findings acquired from the analysis of emotions and sentiments, authorities could discover areas they ought to search for specific information on the crisis, such as the types of emergency needs and the most affected areas. Numerous methodologies have been suggested to assess the role of sentiment analysis in crisis management. Most methodologies deployed variant ML approaches (Vo and Coolier, 2013; Varga et al., 2013) and other approaches, such as swarm intelligence (Chakraborty and Banerjee, 2013). The specific datasets that are used for assessments include posts from social media platforms associated with events such as the

Terrorist Attack in Mumbai, the California gas explosion, the earthquake in Haiti, the flooding of the Red River in 2009, and Hurricane Irene.

Among the earliest works in ML for managing disasters using sentiment analysis is one by Verma et al. (2011). Particularly, the researchers constructed a classifier that could automatically identify situational analysis tweets by classifying them into various classes, including impersonal or personal style, subjectivity, and formal or informal linguistic styles, using a collection of automatic and hand-annotated linguistic features. Features of subjectivity could be significant in evaluating the emotion a Twitter user expresses in the tweets for information extraction and situational awareness. The study used *OpinionFinder* (Wiebe and Rilof, 2005) to extract the tweet's subjectivity as a linguistic feature for the developed classifier. Comparably, researchers such as Nagy and Stamberger (2012) used Bayesian Networks with the aim of detecting the sentiments of a tweet in the California gas explosion from combined emoticons, sentiment ontology, a frequent list of sentiment words (Sebastiani and Esuli, 2006) and AFINN (Nielsen, 2011). Additionally, to categorize the sentiments of the tweets into negative and positive, Nagy et al.'s study (2012) could be useful in detecting information-based tweets to increase the credibility of a post. With the use of various features such as lexicon-based, a Bag of Words (BOW), and pruning features, researchers such as Mandel et al. (2012) train classifiers to classify messages based on different concern levels and leverage trained classifiers to understand the perception of the public towards crises by investigating the association between demographic features such as gender and sentiment.

Researchers Vo and Collier (2016) used the Japan earthquakes as a case study to track the emotions of the crowds in the microblog's posts during the earthquakes. The researchers considered six classes of emotions; relief, fear, anxiety, sadness, unpleasantness, and calm. They proceeded to deploy two ML classifiers to identify tweets relating to the earthquake and

categorized them based on the emotions expressed. With the tracking of the sentiments of the crowds, the researchers concluded that anxiety and fear were the emotions expressed immediately after the events, while unpleasantness and calm were unclearly expressed during smaller earthquakes and clearly during larger tremors. A similar case is the Westgate Mall attack in Kenya. Simon et al. (2014) applied various methodologies to analyze how emergency organizations utilize social media platforms to better their responses. The researchers attempted to discover the difference between the sentiments of the numerous posts made public by emergency managers and organizations and deployed document-level sentiment analysis of TwitterMate. They also mentioned the application of the statistical t-test and confirmed that the post of any manager has a more positive sentiment than that from an organization since the public considers them more approachable from the positive language of the posts.

Torkildson et al. (2014), while considering eight emotions such as no emotion, surprise, sadness, happiness, fear, disgust, anger, and accusation, categorized tweets that related to the Gulf Oil Spills using ALOE (Brooks et al., 2013) that deployed SVM. Buscaldi et al. (2015) used subjectivity, irony, and polarity detection tools to discover related tweets. The researchers believed that the subjective tweets emerged more from individuals at the event. Ironic tweets were posted following the disaster to blame or criticize, whereas event-related tweets presented a more negative sentiment instead of the day-to-day ones. Researchers have also expressed interest in the Genoa Flooding that took place in 2014, where researchers have assessed how sentiment analysis and NLP methods impact the extraction of useful information. Lu et al. (2015) sought to discover the effect of crises on the sentiment of social media streams. In more detail, the researchers investigated trends in both negative and positive sentiment regarding the crises and sentiments that related geographically. Particularly, the study proposed a measure of uncertainty to assess the disagreement in the

different classifiers using the vote entropy method and later used a committee vote scheme to establish the sentiment class of the Tweets.

From the vast amount of information that continues to be published on social media platforms during a disaster, corresponding messages about aids and reports on the problems were less likely to be successfully exchanged in almost all cases. Thus, numerous resources would be put to waste, and victims would be less likely to receive the necessary aid. Discovering matches between the aid agencies and the problem reports would improve communication between humanitarians and the victims. Varga et al. (2013) attempted to solve the issue of help with the reports and the recognition of their messages. Additionally, the researchers attempted to match the above with ML techniques by taking into consideration various features such as semantic, sentiment, and lexicon features. Moreover, the flow of information during events such as hurricanes is a key part of the management of disasters. Attempting to discover that, Kryvasheyeu et al. (2015) assessed the sentiments of more than 50 million tweets before Hurricane Sandy happened, during and after it to explore the behavior change in individuals on Twitter based on the number of times the hurricane reached various cities in terms of the expressed sentiments and the published posts. The researchers explored the Tweet's sentiments extracted from the normal ones and discovered that the assessment of sentiments, together with other metrics, facilitated the use of sentiment sensing for the detection and location of disasters.

Apart from the normal ML techniques, there exist other methodologies that are suggested for sentiment analysis in the management of disasters. For instance, a model by Chakraborty and Banerjee (2013) highlighted the evolution of individuals' views, sentiments, social dynamics, and opinions following a major event. The researchers used the swarm intelligence technique to explore the time-bound and dynamic properties of social disasters. In a sentiment analysis of 615 Nepal tweets under the hashtags Coronavirus and COVID-19,

the final results by Pokharel (2020) showed that approximately 58% of the tweets from the users displayed optimism, 27% displayed neutrality, and 15% displayed negativity.

Contrastingly, Barkur and Vibha (2020) showed a sentiment analysis of the tweets from India on lockdowns through the collection of Twitter data from two hashtags; *#IndiaFightsCorona* and *#IndiaLockdown*. Findings from the study showed 24,000 tweets were strictly dominated by positive sentiments. In a similar study by Manguri et al. (2020), who collected 530,000 tweets with the keywords COVID-19 and Coronavirus, discovered that more than 36% of the tweets posted by the individuals were positive, the negative ones totaled 14% with the neutral ones accounting for 40%.

Bhat et al. (2020) used two hashtags performed sentiment analysis related to the coronavirus pandemic. The hashtag *#Coronavirus* gathered 85 513, while *#COVID-19* put together 92,646. Sentiments that had been collected on the first hashtag displayed more neutrality (41.27%), while the negative ones resulted in 17.80%. Positive sentiments recorded 40.91%. Findings revealed that citizens were more optimistic about the second hashtag as most of the tweets were positive, 51.97%, tweets relaying negative sentiments 13.96%, and those displaying neutrality 34.05%. In general, the results showed that the opinions of a user during a crisis and especially a pandemic are mostly neutral or positive.

2.13 Techniques to Successfully Identify Emojis' Hidden Annotations

Sentiment analysis (SA) is a method of natural language processing (NLP) that tries to figure out how someone feels or what they think based on what they read. In sentiment analysis, one of the most important things to do is to classify the polarity of a given text and decide if the sentence expresses a positive, negative, or neutral emotion. With the rise of machine learning and deep learning, sentiment analysis jobs are using more than one way to group things (Liu et al., 2021).

2.13.1 Manual Annotation

Using Twitter as a reference, it is true that Tweets frequently include URL links, photos, and other forms of multimedia to increase the variety of emotional expression. As the most popular form of Twitter, emojis transmit numerous meanings in addition to graphic symbols. In past studies, however, complicated components such as emojis were frequently omitted (Ye, 2019). A WebApplication³ has been established to gather and manually annotate tweets submitted by our kind visitors to better comprehend the actual use of tweets. Visitors can pick between "Slightly positive" and "Slightly negative" as two additional tones in addition to the standard three-classification problem. This enables the annotators to make proper judgments when encountering emotionally ambiguous and difficult-to-determine tweets.

2.13.2 Automatic Annotation

Manual annotating has the disadvantages of being laborious and time-consuming. Thus, an automatic labeling approach can be utilized to successfully expand the dataset. When presented with a tweet, VADER-Sentiment-Analysis⁷ operates as an automatic annotator and returns a score between strongly negative and strongly positive. This score indicates the ordering of sentences calculated by adding the valence scores of each word. VADER can identify online slang properly and speculate on its sentiment based on its obscene meaning.

Extraction of emojis and emoticons can be a component of automatic annotation. Ye (2019) eliminated the location information of the remaining emojis because they carry only emotional information that is frequently independent of position. It was mentioned that the emotion mapping technique could be utilized to limit the number of emotions to evaluate. BERT's vocabulary is expanded to include processed emojis and emoticons. The word embedding associated with them can be learned in subsequent training to avoid them being incorrectly labeled as unknown or segmented.

Based on their research, Liu et al. (2021) concluded that emojis can be useful as extending features to boost the preciseness of sentiment analysis algorithms and that taking into consideration the various ways emojis are used can improve the performance of the algorithms. The CEmo-LSTM emoji embedding model, built around Bi-LSTM, was used in this study since it is the most accurate method currently available for interpreting Chinese characters in digital media. The sentiment analysis model used in this experiment was improved upon by including emojis and emoji use in a novel emoji-embedding algorithm.

2.14 Machine Learning Models for Sentiment Analysis

Studies (Taleby Ahvanooey et al., 2019; Yoo & Rayz, 2021; Liu et al., 2021) showed that technology has significantly evolved. However, the evolution has been accompanied by numerous challenges, such as the need to adopt data-driven policies and the need to predict unexpected events. Amidst such needs, there have been increased focuses on data encryption and the need to obtain meaningful information in sentiments passed across different social media platforms. This has been accompanied by the development of various machine-learning models for the analysis, classification, and production of human sentiments. According to Guibon et al. (2016), these methods range from general neural networks, support vector machines, and naïve Bayes algorithms. However, these models are considered highly theoretical (Guibon et al., 2016). Therefore, it is critical to examine their overall applicability in sentiment analysis, especially where emojis are concerned.

Eslami (n.d) recognized sentiment analysis as the key to monitoring and discerning user emotions and opinions. This importance has seen the model adopted to improve productivity and enhance key aspects of marketing; for this reason, Eslami (n.d) aimed to develop a sentiment analysis using a Convolutional neural network and long-short term memory. The analysis shows that convolutional networks are oriented towards processing structure arrays, thereby allowing image input, which is a key part of emoticons. As a result, it allows one to

embed emojis into the model for effective sentiment analysis. The model is further characterized by the presence of feed-forward networks that allow one to learn local features and account for temporal dependency. This aspect is accompanied by the use of a bidirectional long short-term memory that helps improve the accuracy of the model (Eslami, n.d). Therefore, Eslami (n.d) recommends the use of the CNN-Bi-LSTM approach to model sentiment analysis algorithms— the model's efficiency has also been emphasized in the works of Liu et al. (2021).

Surikov & Egorova (2020) stated that understanding a message requires one to recognize tonal variations across a given message. These elements allow one to examine the emotion indicated by any users within a social network. As a result, the study focused on developing a model that accommodates emotions, emojis, and punctuations within texts. This aspect of development was oriented towards classifying emotions as either positive or negative. The model was designed to ensure a total connection with vectors to ensure the classification of data. The classification elements were accompanied by a neural prediction network for testing and predicting emotions within a given text. The resulting model recorded a 91% accuracy in production (Surikov & Egorova, 2020). However, the prediction of positive emotional indicators was significantly higher than those associated with negative emotions.

Sentiment prediction and forecasting are critical tasks in natural language processing, and machine-learning approaches have become increasingly popular in recent years (Le, 2021). The most robust machine-learning approaches or hybrid techniques to attain significantly clear results in sentiment prediction and forecasting involve combining deep learning techniques with traditional learning techniques. The former model integrates supervised learning algorithms like Support Vector Machine (SVM) and Convolutional Neural Networks. These algorithms rely on labeled data to train the model and can enable

text classification into unique sentiment categories. Nonetheless, these techniques cannot handle the complexity and variability of natural language; hence they require a large amount of labeled data. The latter model consists of the combination of traditional machine learning methods with deep learning techniques. It also integrates a pre-trained deep learning model, such as a language model, and fine-tuning it on a specific task, such as sentiment analysis. Hybrid techniques enable significantly precise outcomes in sentiment prediction and forecasting. This is because the model leverages the strengths of different machine learning and deep learning methods. It links traditional machine learning methods like SVM with deep learning techniques like LSTM, handling the complexity and variability of natural language and achieving high accuracy in sentiment prediction and forecasting. Also, an effective sentiment prediction and forecasting approach involves ensemble methods, which combine the predictions of multiple models, helping improve the overall performance (Younis, 2022). The model can be deployed to link predictions of various machine/deep learning methods helping attain significantly precise outcomes.

Sentiment prediction is the process of using machine learning algorithms to predict the sentiment of a statement. This is typically done by training a model on a dataset of labeled sentiment-bearing texts and then using the model to predict the sentiment of a comment given as input. Machine learning techniques, for instance, support vector machines, Naive Bayes, and recurrent neural networks have been utilized successfully for sentiment prediction with high accuracy (Medhat et al., 2014). Support vector machines (SVMs) are a popular technique that can be used to group sentiment-bearing texts into negative, positive, and neutral categories. Naive Bayes is a probabilistic machine learning technique that can calculate the probability of a given text belonging to a particular sentiment category. Recurrent neural networks are a kind of artificial neural network that has been utilized

successfully to identify sentiment in the text (Singh et al., 2018, pg.2). These techniques have been used to achieve high accuracy in sentiment prediction.

On the other hand, sentiment forecasting uses machine learning algorithms to predict future sentiment. This is done by training a model on a dataset of labeled sentiment-bearing texts and then using the model to forecast sentiment given a set of parameters. Common machine-learning techniques used for sentiment forecasting include decision trees, linear regression, and logistic regression. Decision trees are a supervised learning technique that uses a tree-like model to predict a target value based on features. Linear regression is a supervised technique that uses linear models to predict a target value based on a set of features. Logistic regression is a supervised technique that uses a logistic function to predict a target class based on a set of features (Singh et al., 2018, pg.3). These techniques can be used to forecast sentiment accurately given a set of parameters. Sentiment prediction and sentiment forecasting can be used to analyze and predict public sentiment and reactions toward governments' preventive measures and policies during natural disasters. Sentiment prediction can be used to classify text-based reactions as positive, negative, or neutral. This can help to identify public sentiment towards a particular policy or measure. Sentiment forecasting could then be used to predict public sentiment toward a similar policy in the future, allowing decision-makers to anticipate potential reactions. Additionally, sentiment analysis in social media can be used to identify critical issues and topics that are important to the public, allowing policymakers to consider these when formulating their policies. The many types of sentiment analysis models are selected based on the anticipated volume of data and the required level of precision.

2.14.1 Rule- or Lexicon-Based Strategy

This technique for assessing attitude employs criteria for data classification that were developed by humans. This technique assigns a numerical value to the degree to which words

in a dictionary are positive or negative, so providing a measure of the polarity and sentiment strength of a given piece of text. When you use an expression, you can extend the capabilities of your program. By refining existing rules, rule-based sentiment analysis algorithms can be adapted to a wide variety of settings (Islam et al., 2020). It counts the amount of upbeat and downbeat words in the text. If there are more positives than disadvantages, the system will produce an optimistic assessment. If the two are equivalent, then the end result is not affected.

2.14.2 Methods based on Automation or Machine Learning

Instead of using predetermined criteria, the core of a statement is determined by machine learning in this method of sentiment analysis. This ensures that the analysis is more precise and that data can be processed based on various criteria without requiring excessive complexity. In this approach, supervised machine learning methods are used. An algorithm is trained with a large number of example- passages so that it can correctly predict the mood of the text (Grover, 2022). The classifier is then given a ton of data from which it attempts to draw conclusions about the author's likely positivity, pessimism, or apathy. Opinion mining and sentiment analysis using Machine Learning is reliable and effective. The two types of machine learning are:

2.14.2.1 Traditional Models. Gathering, analyzing, and training an algorithm using examples of positive, negative, and neutral classifications are all necessary steps for conventional models. Most of these methods are used to identify the polarity of written material. Naive Bayes, Logistic Regression, and Support Vector Machines (SVM) are all examples of classic machine learning techniques that are commonly employed for large-scale sentiment analysis due to their scalability. Nave Bayes is a common probabilistic classification technique since it is easy to implement and needs only a small quantity of data for training. In contrast, SVM uses a search for the optimal hyper-plane to classify data, and

it has been proven to work well with huge datasets (Grover, 2022). Finding the best kernel functions, however, is a laborious process.

2.14.2.2 Deep Learning Techniques. According to its definition, deep learning is "neural networks with numerous layers of perceptron inspired by the human brain," and it has been found to enhance text generation, word representation estimation, sentence categorization, and feature display (Alexandridis et al., 2020). In this context, "neural network models" refer to models that rely on neural networks to improve the accuracy of conventional models (Deep Neural Network). Inferring emotional states from text via NLP for sentiment analysis is a task well-suited to deep learning, which can find patterns in unstructured and unlabeled data over many layers. Regularized Neural Networks (RNNs) are a type of neural network optimized for modeling and predicting sequence data. As a word-based vector, it probes the interconnectedness of terms over the course of a textual corpus. An RNN's internal memory allows it to remember data it has already processed when the input is read in a sequential fashion (Alexandridis et al., 2020).

CNNs, on the other hand, are typically made up of three layers: convolution, pooling, and fully linked layers, and are designed to learn malleable spatial hierarchies of information (Alexandridis et al., 2020). The first two layers do feature extraction, while the final, fully-linked layer uses those features to generate the final result. Algorithms are frequently used to enhance parameter values since the recovered features may get increasingly complex over time and in a hierarchical structure. By encoding the semantics of both the immediate and remote past contexts, Bi-LSTM is an enhanced kind of RNN for analyzing input text (Alexandridis et al., 2020). It uses bidirectional LSTM units to do its tasks. Repetitive connections between memory blocks make up this structure. Each memory cell consists of an input gate that decides whether or not data is allowed in, a forget gate that determines how long data is stored, and an output gate that decides how data is released from memory. Recent

deep learning models have shown excellent performance in sentiment analysis tasks by addressing linguistic information (such as sentiment lexicon, negation, and intensification) (Alexandridis et al., 2020). In the realm of deep learning, LSTM is a type of recurrent neural network that is utilized for sequence prediction problems. The Twitter dataset, the Stanford Sentiment Treebank for movie reviews, and a mixed domains dataset were used to assess this neural network. Linguistically regularized LSTMs can improve sentiment categorization by utilizing linguistic resources such as negation, intensification, and vocabulary.

2.14.2.3 Hybrid Approach. The hybrid method of sentiment analysis combines statistical and knowledge-based approaches to identify polarity. High precision from machine learning's statistical methodologies and dependable operation from the lexicon-based approach are combined in this technique (Dang et al., 2021). Today's most popular method for analyzing user sentiment is a hybrid model that combines traditional methods with more recent advancements in AI. A well-designed hybrid system can reap the benefits of both automatic and rule-based systems. Hybrid models, which incorporate both the adaptability of personalization and the predictive ability of machine learning, can significantly improve productivity. Dang et al. (2021) tested the performance of hybrid deep learning models for sentiment analysis on eight textual datasets, such as tweets and reviews, by combining SVM, CNN, and LSTM with two-word embedding approaches, Word2Vec and BERT. A total of four hybrid vehicles were produced and then compared to base models. The paper states that experiments were conducted to determine if hybrid models are flexible and if hybrid techniques can handle a wide range of dataset types and sizes. According to the results, hybrid models outperformed other types of models in terms of reliability when it came to sentiment polarity analysis. When applied to sentiment analysis, a combination of deep learning models and the SVM technique yield better results than either method employed alone. In the majority of datasets studied, the dependability of hybrid models employing

SVM is larger than that of those without employing it; nevertheless, the computational time for those with SVM is much longer.

2.15 Summary

The rapid spreading of the Coronavirus disease threatened the global health system and resulted in serious challenges for governments and health systems globally. As a result, countries and regions adopted various prevention actions to curb the spread of the pandemic. Among the measures is the mandatory quarantine and lockdowns that spiked discussions on social media platforms, specifically Twitter, as it is considered the dominant platform for holding crises discussions, and despite efforts to curb negative sentiments that reduced compliance to the regulations, challenges continued to emerge. Further investigation into related literature revealed that, among others, supervised methodologies of ML are almost exclusively adopted by researchers as the most reliable technique for conducting sentiment analysis, without the lack of a few exemptions. Literature narrowed down to classification algorithms and exclusively established that they are dominant in this field of research. Despite classifiers, in general, demonstrating robust findings, other approaches have also warranted the attention of researchers in the field. For instance, the review established that unsupervised ML could generate promising results in sentiment analysis. Additionally, there appeared a limitation; the undeniable challenge of ambiguity in unpopular emojis and emoticons and their unidentified multiple or single annotations. Even in commonly used emoticons and emojis, the multiple meanings that could be derived arouse concern as it has significantly affected the accuracy of the different classifiers, producing a complex lexicon. Similarly, hidden emotions, sarcasm, or general implications are also listed as some of the challenging aspects when assessing a statement based on the emojis used. Multi-class or -vectors, together with classifiers, also displayed an incapability of interpreting complex human-related uncertainties.

In an Arabic context, the literature showed that an ML algorithm-based technique exhibited better performance than the one based on lexicons in the prediction of the tweet's subjectivity. Researchers concluded that Twitter users of such areas were more than willing to follow and spread health measures and displayed positive opinions upon government announcements. Religious practices equally played a significant role in the improvement of the overall positive sentiment. The empirical review provided data on sentiment analysis amidst the COVID-19 pandemic and, in the process, is assumed to make the following contributions to the study; firstly, in cementing the mentioned fact that Twitter is the most common social media platform used in extracting information -as the most of the reviewed studies used the platform in the social media context-. In cementing the above, it was discovered that the rising popularity of the platform is due to the richness and availability of content; millions of tweets are posted and viewed daily on any topic. This serves as an indication that social media platforms are now becoming precious sources of information. However, researchers also suggested the need to pay more attention to other platforms and sources, such as microblogs and YouTube, since the contexts might differ, and it is also important to explore other sources as they are likely to unravel new findings and knowledge.

Secondly, the review showed the method useful in assessing sentiment on social media platforms. There are numerous methodologies introduced by researchers; however, the most commonly used in Lexicon-based methods for ML are SVM and Naïve Bayes. It was also established that selecting an appropriate methodology for sentiment analysis solely depends on the selected dataset; the classifiers demonstrated relatively high accuracies. Researchers also highlighted the need to consider the amount of data, time, and structure of the text. If the data is large, opting for an ML-based method is recommended simply because it requires more data and time to train. Researchers identified a lexicon-based method for a messy structure as there is limited time and a relatively small amount of data to analyze. However,

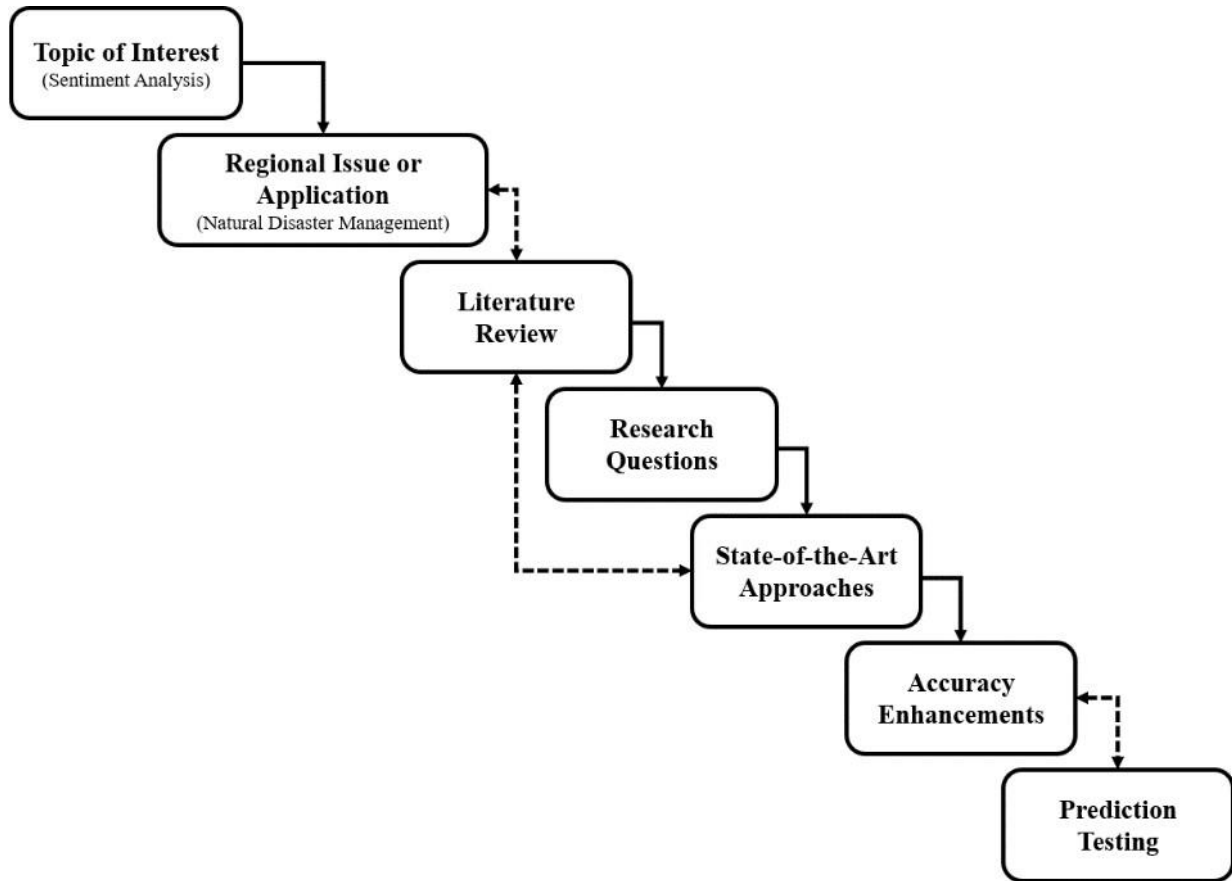
for researchers to enhance the quality and accuracy of their findings, suggestions were made to combine both ML and lexicon-based methods. Lastly, the review demonstrated the application of the sentiment analysis process on Twitter in various disasters and crises. Additionally, the review showed that the process could also be applied in a broader context and has been used in various areas, such as in the improvement of strategy and quality in businesses, the monitoring of an outbreak, in creating awareness of the significance of data security and in monitoring the perception of customers towards certain products, among others. As assumed, researchers in the review showed that sentiment analysis plays a significant role in understanding the perception of individuals and aids in decision-making processes. Generally, the review reiterates the need to develop a universal model for the analysis of sentiments that can be adopted when dealing with different data and identifies other platforms useful in obtaining the opinions of a user and developing the sentiment analysis application context.

Chapter 3: Research Methodology

3.1 Introduction

This chapter focuses on the research methodology that was adopted in this study and the rationale behind it. Like most application-based empirical research, this study pursued a purely quantitative data-driven approach by analyzing numerous datasets obtained from Twitter, using supervised learning classifiers and artificial recurrent neural network (RNN) algorithms. Concretely, the research will test the hypothesis that COVID-19 precautionary policies, measures, and restrictions considerably impact public sentiments in general, as well as having a noticeably diverse effect on various societal groups.

Based on the produced research questions and the examination of the previous related work, the most distinguished and promising data collection, data preprocessing, and data analysis models and algorithms will be employed to explore people's sentiments and feelings during COVID-19 announcements and amendments. Furthermore, the research will attempt to enhance the accuracy beyond what has been demonstrated in the literature by combining the leading techniques in every phase of the process, including the preprocessing pipeline, data initialization, feature extraction, and selection, as well as the training algorithm, as will be sufficiently elaborated in the following subsection. Ultimately, the final produced model's prediction accuracy will be validated against numerous datasets representing realistic scenarios and actual case studies. Depending on the results, additional enhancements might be required to achieve competitive outcomes. The model will be evaluated against the preceding state-of-the-art performances. Figure 5 summarizes the research methodology that will be pursued in this study.

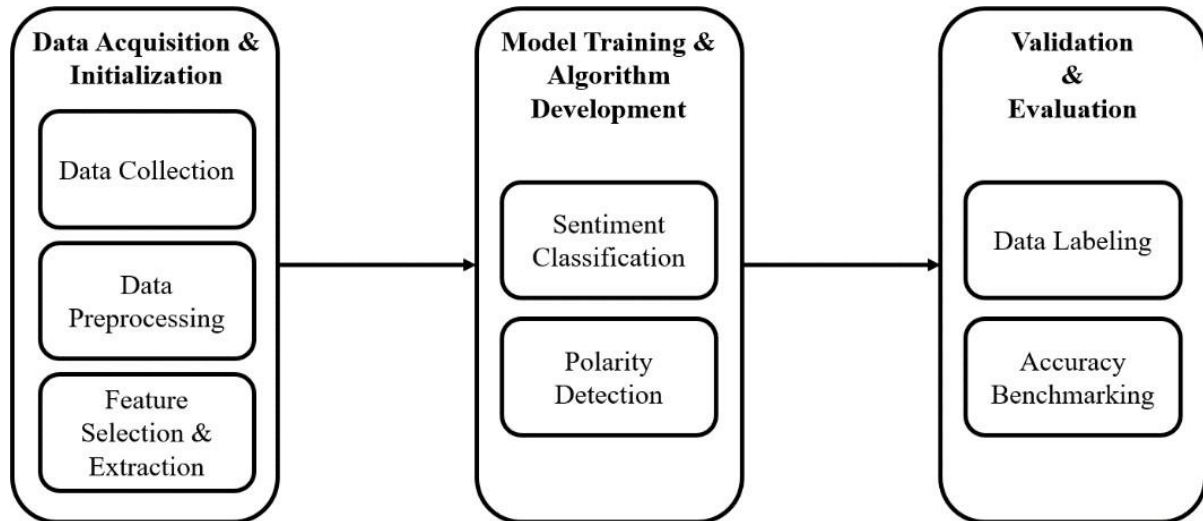
Figure 5*The Research Aim & Methodology***3.1.1 Research Approach & Justification**

This subsection will provide an in-depth exploration of the adopted conceptual framework or research paradigm. The study will implement the standard sentiment analysis scheme (Figure 6) that has been widely utilized in the literature and proven to be the most efficient strategy for executing and training a classifier (Angiani et al., 2016). Additionally, the conventional framework will be elegantly refined to enhance the accuracy beyond what has already been achieved or demonstrated. The paradigm consists of three primary stages:

- i. Data acquisition and initialization.
- ii. Model training and algorithm development.
- iii. Validation and evaluation.

Figure 6

The Research Paradigm (Angiani et al., 2016)



3.1.2 Data Acquisition & Initialization

This phase includes data collection, preprocessing and transformation, and feature selection and extraction. In short, this step prepares the collected data by converting it from human-readable text and emojis into machine-readable numerical attributes to perform any required mathematical operation, such as matrix arithmetic. For data collection, the study exclusively utilized Twitter as its primary data source for various reasons, including open accessibility, the nature and easiness of interactions, the Twitter application programming interface (API), the facilitated searching mechanism, the powerful hashtag society, and the tremendous variety of users in general (Ahmed et al., 2017). Twitter data collection and acquisition requires several steps, including creating a Twitter developer account and acquiring Twitter API keys by simply generating the access tokens. The ethical considerations section will comprehensively discuss the legal and ethical obligations concerning creating a Twitter developer account and importing data from Twitter. Furthermore, specific Python libraries are required to integrate Twitter API (including Twitter Trend API) with the Python

environment for the sake of convenience.

This study considered two of the most popular Python libraries for accessing Twitter API: *Tweepy* and *Twint* (Almatrafi et al., 2015). While *Tweepy* is a library or a tool to formally access Twitter through its official API, *Twint* is designed to overcome Twitter API limitations. Moreover, it is essential to note that Python is solely adopted in this study as the main programming language and machine learning environment for obvious reasons (Hugunin, 1997). *Google Collaboratory* is utilized as the primary Python environment due to its cloudstorage and computation capabilities, allowing the use of both Graphical Processing Units (GPUs) and Tensor Processing Units (TPUs). Data has been gathered considering various COVID-19 precautionary policies and restrictions phases, including compulsory leave-home permits, mall closing, grocery store working hours adjustments, border lockdowns, and complete internal lockdowns. The data was then imported based on specific announcement dates, particular timings, certain hashtags, and precise search terms. Additionally, the study only analyzes English tweets considering the massive number of English interactions compared to any other language (Hong et al., 2011). Lastly, Emojis and Emoticons are strongly considered due to their undeniable role in conveying emotions through compelling graphical implications (Shiha et al., 2017). In other words, Emojis' sentiment is a major and integral part of this study.

After integrating Twitter API into Python, the imported data has been loaded, stored, and structured using essential data science libraries such as *Pandas* and *NumPy*. Both libraries offer the ability to import data in multiple formats like *SQL* and *CSV*, as well as to provide a convenient and robust structure and analysis functions. Data preprocessing and transformation, on the other hand, have uncountable methods and techniques. It is almost impossible to plan every preprocessing and transformation step ahead of the actual implementation (Hemalatha et al., 2012). In general, data preprocessing and transformation

techniques can be categorized into three main categories: tokenization, normalization, and substitution (Mayo, 2017). Tokenization is an essential procedure in which text is converted into tokens (Webster et al., 1992). Tokenizing a specific tweet would basically convert it into sentences and then into individual words. This conversion will greatly simplify cleaning or filtering any unwanted words and characters known as noise. The cleaning process eliminated punctuations, irrelevant numbers, special characters (such as # or @), and stop words. Stop words are a set of frequently used words, such as pronouns and articles, which do not contribute to the meaning or the sentiment of a phrase (Saif et al., 2014). These words are usually discarded by search engines. Tokenization can be accomplished using *Natural Language ToolKit (NLTK)* or *SpaCy*. Both are among the most commonly used natural language processing packages in Python (Garg et al., 2016).

Normalization can be defined as retransforming the word back to its standard origin by eliminating any unnecessary additions, conversions, or even unintentional errors (Nayak et al., 2014). Normalization includes lowercase conversion, elimination, negation, lemmatization, and many more techniques depending on the specific requirement and implementation. However, some techniques are essential when dealing with sentiment analysis models as they can entirely reverse the outcome and significantly enhance accuracy. Negation is one of those essential techniques. Negation is as simple as removing apostrophes from negating words like “can’t” and replacing them with “not,” as in “cannot.” There are multiple open-source dictionaries and look-up tables available online that can be utilized to eliminate apostrophes or any other related shortcuts. Stemming or lemmatization are other essential normalization procedures (Korenius et al., 2004). Lemmatization reverts the word back to its original root or form. Both techniques exclude affixes such as “-ed,” “-ing,” and “-s.” However, unlike stemming, lemmatization takes the word’s context into consideration. Lemmatization can be implemented in Python using *NLTK* or *Gensim* packages.

Additional and more advanced normalization techniques such as dictionaries and lexicons, emoticons, splitting, slang, and grammar and spell correction have been only slightly deployed, as needed. Substitution is concerned with removing the noise from tweets in their raw format. Tweets naturally contain *HTML* markups as any text obtained from the web. In addition, Twitter allows its users to share *URLs* in their tweets. During the substitution stage, these undesired *HTML* markups and *URLs* must be entirely cleaned out. Furthermore, decoding could be the most crucial substitution technique. Decoding is involved in unifying the textual data encoding format. Textual data could be subject to multiple decoding formats, such as *ASCII*, Unicode, and *UTF-8*. It is highly recommended to consolidate the data into a single widely recognized format such as *UTF-8* due to its languages' compatibilities and memory efficiency (Zhang et al., 2017).

Feature selection and extraction is a crucial stage where the pre-defined attributes are converted from the raw text (human-readable) to numerical arrays or matrices (machine-readable). In other words, this step is sometimes called Word Embeddings, which simply implies the transformation process of text to numerical representations (Giatsoglou et al., 2017). Broadly, word embeddings can be classified into two main classes: frequency-based and prediction-based (Kutuzov et al., 2016). Frequency-based embedding techniques such as *Bag of Words (BOW)* and *Term Frequency-Inverse Document Frequency (TF-IDF)* are widely adopted in many NLP applications due to their fundamentally unsophisticated approach (Djuric et al., 2015). *TF-IDF*, for example, weighs and analyzes the significance of words based on the appearance of a specific word, not only in a single tweet but in the entire set of tweets. This specific feature enables *TF-IDF* to discard pronouns, prepositions, and other frequently used words that do not contribute to the actual meaning. In fact, according to Beel et al. (2016), more than 80% of all recommender systems utilize *TF-IDF* as their primary extraction technique. *TF-IDF* can be implemented in Python using the *Scikit-Learn* machine-

learning library.

Prediction-based embeddings, on the other hand, are more state-of-the-art in terms of analogies detection. Unlike frequency-based embeddings, prediction-based detects context, semantics, and relationships. In addition, prediction-based techniques display words in a vector space representation, where words with similar contexts are positioned in close vicinity to one another (Goldberg et al., 2014). Two of the most widely utilized prediction-based embeddings implementation are *Word2Vec* and *Global Vectors (GloVe)*. *Word2Vec* was first introduced by Google in 2013 (Yao et al., 2017). *Word2Vec* achieves word embeddings with two-layer shallow neural networks for context detection purposes. The algorithm itself consists of a combination of two different algorithms: *Continuous Bag-of-Words (CBoW)* and *Skip-Gram*. Each algorithm performs the opposite task of the other in terms of context-from-word prediction and vice versa. *Word2Vec* can be implemented in Python using the *Gensim* library. *Word2Vec* has been widely used in various sentiment analysis implementations due to its well-known accuracy (Goldberg et al., 2014). However, *TF-IDF* has also shown some promising results in various NLP applications, such as recommender systems (Beel et al., 2016). This particular study will consider and compare both techniques' abilities for accuracy validation purposes.

Emojis embeddings, on the other hand, are a fresh yet highly trending topic (Chen et al., 2018). Like *Word2Vec*, emojis can be vectorized and embedded based on context and meaning. Through the literature, two methods were distinctively noticed: *The skip-Gram* approach & *Word2Vec* method. The *Skip-Gram* algorithm utilizes frequently appearing nearby words from around 100 million English tweets to interpret an emoji (Barbieri et al., 2016). The *Word2Vec* method, however, exhibits a more scientific approach by utilizing the emojis' official Unicode definitions and keywords (Eisner et al., 2016). Afterward, the description words are managed as normal text using *Google's Word2Vec* model. Both

techniques are publicly available and can be imported as dictionaries; however, this study will consider the *Word2Vec* approach (*emoji2vec*) due to the adoption of the official emojis Unicode (Eisner et al., 2016). Finally, a more detailed explanation and in-depth representation of the exact approach that has been utilized to produce this study is discussed in the next chapter.

3.1.3 Model Training & Algorithm Development

Sentiment analysis algorithms, in general, consist of two broad stages: sentiment classification and polarity detection (Symeonidis, 2018). In terms of sentiment analysis approaches, there are three main techniques or schools of thought:

- i. Machine Learning-based method.
- ii. Lexicon-based method.
- iii. Hybrid.

The machine learning approach consists of supervised and unsupervised learning. Supervised Machine Learning includes many widely used methods, such as decision tree classifiers, rule-based classifiers, linear classifiers, and probabilistic classifiers (Symeonidis, 2018). Each method within the main supervised learning approach contains various algorithms, models, and techniques. However, based on the previous literature review, it can be noticed that linear and probabilistic classifiers have been popularly utilized and demonstrated some robustly accurate outcomes. In particular, Long Short-Term Memory Recurrent Neural Network (LSTM) as a linear classifier, and Multinomial Naïve Bayes as a probabilistic classifier. Other models are the transformers model, K-nearest neighbor (KNN), Random Forest classifier, and Logistic regression model. Similar to the traditional feed-forward neural network, recurrent networks are designed to deal with data sequentially as the previous step's output is fed to the current one (Mikolov et al., 2011).

However, RNNs, especially LSTM, have the ability to process multiple sequences of inputs using their long short-term memory. Thus, RNNs are utterly capable of capturing and

utilizing patterns for prediction purposes. Although all inputs and outputs are assumed to be independent, LSTM demonstrated tremendous results in sentiment analysis (Wang et al., 2016). Similarly, Naïve Bayes Classifiers naively assume independent features in order to anticipate the category of a training sample class. However, Naïve Bayes classification heavily relies on probability calculations using the Bayes theorem. Thus, the algorithm outputs the classification that achieves the highest probability (Murphy, 2006). Naïve Bayes model has been widely used in various NLP applications due to its simplicity, reliability, and accuracy, as well as being one of the fastest approaches in terms of performance (Islam et al., 2007). Compared to other tremendously more sophisticated linear classifiers such as Support Vector Machines (SVM), Naïve Bayes classifiers have demonstrated similar or even better outcomes with a much simpler approach. Multinomial Naïve Bayes implies the type of the probability distribution of each feature as multinomial rather than various combinations of distributions for the sake of simplicity (Stuart et al., 2003). KNN algorithm selects the k-nearest of data points and assigns them to the commonly established class among the adjacent data points. The Random Forest classifier, on the other hand, performs observation and feature selection randomly to generate several trees and computes the average results. The logistic regression model is used for binary classification by assigning 1 or 2 classes to a specified input. It utilizes the set of input features and uses the mathematical function known as the logistic function to change the linear combination of input features into a probability value between 0 and 1. Finally, the transformer model is a neural network that learns context and tracks relationships in sequential data. The model is based on the concept of self-attention and attends to input sequences to calculate how each token or word is represented. The model has an encoder and decoder and uses attention to enhance the speed of the model. It can learn large dependencies between tokens in the sequence, which is critical in model classification. These model algorithms can be implemented in Python using *TensorFlow* and

Scikit-Learn machine-learning libraries. Additionally, *Google Cloud Natural Language API* and other machine-learning services might be considered for the implementation.

This study exclusively employs machine learning-based methods and models to train and apply sentiment analysis. LSTM, Multinomial Naïve Bayes Classifiers, Transformers, K-nearest neighbor (KNN), Random Forest classifier, and Logistic regression model were also used. Similar to the traditional feed-forward neuralnetwork, recurrent networks will be explored, utilized, and compared in terms of accuracy and performance. It is worth mentioning that other machine learning-based approaches, such as decision trees, rule-based classifiers, and unsupervised learning, will not be considered due to the lack of literature that might imply ineffectiveness. In addition, lexicon-based approaches (dictionary or corpus-based) will not be considered either. In general, whether statistical or semantical, lexicon-based approaches heavily rely on the accuracy of the utilized lexicons or dictionaries (Hailong et al., 2014). Ultimately, a lexicon can incorporate only a few words, greatly limiting its capabilities and accuracy. Moreover, it can be clearly observed from the literature that machine learning-based approaches' performance exceedingly surpasses lexicon-based ones, although they require tremendously more training datasets.

3.1.4 Validation & Evaluation

Testing a machine learning model, in general, requires pre-verified testing datasets that have been sentimentally categorized by humans (Wei et al., 2013). Such datasets are publicly available in many open-data sources, such as *Kaggle*, *DrivenData*, *Codalab*, and *CrowdANALYTIX*. In addition, application-specific Twitter testing data for NLP, in general, and sentiment analysis, in particular, are widely available. Such data can be utilized for both training and testing purposes. Pre-categorized testing data are ideal for measuring the model analysis against actual human behavior and genuine annotations (Wei et al., 2013). Additionally, the acquired and preprocessed data through Twitter API can be used for both

training and testing.

Many practical studies have shown that the best outcomes are achieved when using 70-80% of the obtained data for training, whereas the other 20-30% for testing (Gholamy et al., 2018). Using regionally obtained Twitter data, the model will undergo multiple performance analyses and accuracy evaluations using global measurement metrics and benchmarks, including recall, precision, F-score calculation, and overall accuracy (Goutte et al., 2005). The accuracy benchmark provides rich insight into the model's performance regarding correct sentiment predictions. In comparison, recall measures the accuracy of neutrality detection. F-score, on the other hand, provides a powerful intuition combining both precision and recall. These widely adopted benchmarks will provide a comprehensive perspective on the overall performance of the model (Goutte et al., 2005). Finally, the model's performance, ratings, and benchmarks have been compared against state-of-the-art sentiment analysis algorithms and models such as *IBM Watson Tone Analyzer*, *Amazon Comprehend*, and *Microsoft Language Understanding Intelligent Service (LUIS)*.

Automated data labeling has been entirely adapted to initially tag or pre-annotate the training datasets. Whether internally or externally, manual data labeling will be completely excluded in this research due to its ineffectiveness, data size limitations, time consumption, and other financial expenses (Haas et al., 2015). In addition, external data labeling involving third parties, such as crowdsourcing, imposes potential security risks and certain privacy violations, as will be discussed in the next section. The study initially incorporated two automated labeling techniques: semi-supervised learning (SSL) and transfer learning. SSL integrates both supervised and unsupervised machine-learning techniques to label a relatively massive amount of data using a small pre-tagged dataset (Rei, 2017). On the other hand, transfer learning employs a machine learning algorithm that has been pre-trained using an equivalent training dataset that is decently comparable or adequately similar to the original

unlabeled dataset (Chen et al., 2019). Both methods have been closely examined and tested to overcome the limitations of the high dependency on the small pre-trained dataset, as well as the possible inaccuracy due to any incorrect predictions. *Keras (TensorFlow API)* and *Scikit-Learn* packages can both be employed to implement automated data labeling techniques.

3.2 Ethical Considerations

The revolutionary emergence of social media platforms has undoubtedly played a significant role in enabling the concept of big data, which in turn has led to the tremendous evolution of data analysis studies and research (Smith et al., 2012). However, such large amounts of scattered data impose great responsibilities, moral dilemmas, and ethical considerations (Smith et al., 2012). Hence, data analysts and researchers are highly responsible for paying extremely close attention to the various ethical and moral considerations surrounding such vast amounts of data in terms of data collection and acquisition, analysis and interpretation, and utilization (Schuller et al., 2016). In addition, various social platforms, academic institutions, public databases, and software library developers have their own privacy policies and ethical agreements. These strict policies and agreements incorporate and oblige certain ethical considerations to protect public users' data privacy. Corporations, websites, and institutions require prior agreements and consent. For instance, Twitter imposes stringent rules on specific users who want to use their developer tools to import other users' data. This set of rules is concerned with protecting sensitive information, off-Twitter matching, content redistribution, spam and auto-responses, and other privacy concerns (Talib et al., 2014).

This particular study will involve and comply with the ethical regulations of the British University in Dubai (BUiD), Twitter, public datasets websites such as *Kaggle*, libraries developers like *Scikit-Learn*, and machine learning model providers such as *Google*. The Research and Grant Office at BUiD requires researchers planning to conduct any type of

study that deals with potentially and ethically sensitive subjects under the universities' association to seek approval from the Research Ethics Committee, using the research ethics form self-assessment (Appendix A). The form was submitted along with the final research proposal prior to the official beginning of this study to obtain the board's authorization. Furthermore, online data science communities and databases dictate a precise Acceptable Use Policy (AUP). For example, *Kaggle's* AUP requires prior agreements against any type of misconduct, violations, illegal activities, or even personal promotions and advertisements. Similarly, software libraries such as TensorFlow or online machine learning models and algorithms providers like *Google Cloud Natural Language API* demand similar consents and agreements.

Sentiment analysis, as a specific application of NLP, carries various moral dilemmas due to its nature. The text and emojis annotation process can be significantly misleading if not handled with extreme caution and neutrality (Cambria et al., 2013). Whether it is the limited annotations within an embedding or a lexicon, or an insufficiently trained model, these unintentional technical issues can lead not only to undesired or impractical outcomes but also to unethical conclusions. Unethical outcomes concerning improper socio-biases have been previously reported due to the reasons discussed earlier (Mohammad, 2020). Moreover, as an empirical study, the results must be reported genuinely and transparently, including any negative or unwanted findings and conclusions. In addition, shortfalls, limitations, and caveats will be represented both at the proposal stage, in the next section, and at the final reporting stage. Finally, every source of information, no matter how significant it is, will be adequately cited and referred to, including prior related work and literature, software and tools, libraries and data sources, and online articles or tutorials.

Chapter 4: Experimental Results and Analysis

This section will dive deep into the technical details that were briefly overviewed in the previous section by explaining every single step toward implementing the backend of the study until generating and examining the produced outcomes.

4.1 Data Acquisition and Initialization

Under the data acquisition phase, the following processes are included; the process of data collection, data pre-processing, and the feature selection and extraction process. Simply stated, this step prepares the collection of data by transforming emojis and human-readable texts to machine-readable numerical attributes required to perform any necessary mathematical operations.

4.1.1 Data Collection

The study exclusively uses Twitter as its primary source of data. Following the conclusions by Ahmed et al. (2017), Twitter was deemed most appropriate due to its ease of interaction, open accessibility, the variety of users on the platform, powerful hashtag society, and its facilitated search mechanism. Additionally, and with reference to the extensive literature review discussed earlier, it can be evidently observed that Twitter is by far the most preferred resort for individuals to quickly express their opinions and feelings toward a trending topic or phenomenon in short textual messages. Twitter offers various Application Programming interfaces (API) that allow its users to acquire access to different types of data, such as the users' profiles, the tweets' content, and the number of retweets. Representational State Transfer (REST) is able to collect historical data, while Streaming API gathers real-time data on Twitter. According to Jiang et al. (2016), REST is the most optimal tool for gathering historical data on Twitter. However, there are several disadvantages of using REST for the process of collecting data, one being that APIs on Twitter have various restrictions on data amount and the downloading rate. It is also important to note that the process of data

collection might be declared inefficient once the time span is declared long, comparatively, for instance, a year. This is because of the limits set by REST API.

The collection of data requires multiple steps, with some being creating a Twitter developer account and acquiring API keys by the simple generation of access tokens. The ethical and legal obligations of the Twitter platform developer account and data importation are comprehensively explored in the ethical consideration section. Moreover, with the use of Python, there necessitates the need to integrate Twitter API, including trends, with the Python environment for convenience's sake. Definitively, the study tests the hypothesis that COVID-19 measures, precautionary policies, and restrictions considerably affect general public sentiments and have a noticeably diverse impact on different societal groups. Data is to be imported on the basis of precise search terms, particular hashtags, certain timings, and specific announcement dates. In addition, the study will assess English tweets taking into consideration the massive English interactions when likened to other languages, including Arabic.

Emoticons and Emojis are also strongly considered. Shiha and Ayvaz (2017) explain that this is because they play an undeniable role in relaying information through compelling graphical implications. Simply phrased, sentiment analysis from emojis is an integral and major section of the current study. Following the integration of Twitter API into Python, imported data is locked, stored, and formulated using critical data science libraries such as Pandas. Such libraries enable the importation of data into several formats and equally offer robust and convenient structure and analysis functions.

4.1.2 Dataset

For the research, the dataset utilized consisted of tweets associated with the COVID-19 infection and COVID-19-related tweets across the globe. The COVID-19 infection data were publicly provided, and the tweet data were collected using the Twitter API using a set of

predefined keywords, hashtags, and emojis whose facial expressions exhibited different kinds of emotions, such as sadness, happiness, anger, and fear. The dataset contained keywords based on the trending phrases associated with COVID-19 across the globe, including general posts about COVID-19, posts regarding the reporting of the number of infections, posts regarding vaccination, and posts about government policy –the emergency state declaration and the spread-prevention policy. These keywords included vaccination, COVID-19, coronavirus, lockdown, sanitize, lockdown, social distancing, quarantine, stay safe, stay home, and sanitization. Keywords were considered since they signify the meaning or main ideas of the dataset and are often used as the index to the data set contents. They allow the researcher to select the specific data they need and filter out irrelevant information. This was particularly important since the dataset involved was large and diverse. Keywords allowed collection of data that is relevant to the research question. Additionally, keywords enable the collection of data from multiple sources, for instance, from social networking sites such as Twitter, and remove irrelevant data. Hashtags allow the identification and collection of relevant data related to public reactions toward the COVID-19 pandemic. They help classify and label Twitter content, allowing for easy search and analysis. The utilization of hashtags in collecting data is useful in evaluating trends on social media and sentiment analysis. For instance, in this experiment, the researcher collected data on tweets related to public reactions toward the COVID-19 pandemic across the globe by searching for tweets with relevant hashtags. The above keywords were hash-tagged to allow easy collection of data. In the period from November 13th, 2020, to January 1st, 2021, 179109 tweets were collected. The period considered is when the COVID-19 pandemic had already spread throughout the world, and the outbreak led to unprecedented situations with long-term quarantines to minimize the spread of the virus.

4.1.3 Experimental Setup

The COVID-19 pandemic affected the entire world, and as such, this experiment utilized COVID-19-related English tweets across the globe. The dataset utilized contains key information regarding most of the tweets as their attributes. The attributes of our dataset include user_name [Name], user_location (Text), User_description (Text), user_created (DateTime), User_followers (Number), user_friends (Number), user_favourites (Number), user_verified (Text), date (DateTime), text (Text), hashtags (Text), source (Text), and is_retweet (Text). The tweets collected contained hash-tagged keywords such as #COVID-19, #covid, #coronavirus, #socialdistancing, #quarantine, #staysafe, #lockdown, #stayhome, etc. Table 2 (next page) represents an overview of the dataset. The study utilized the Keras machine learning platform in the experiment, which used Tensorflow for deep learning approach implementation as the backend. The data utilized in the experiment was both collected and scrapped using the libraries discussed in the previous chapter. As a quick recap, TWEETPY was used to access Twitter API; however, due to the many limitations in the API, TWINT was then used to scrap some additionally required data.

Table 2

COVID-19 Twitter Dataset Sample

	user_name	user_location	user_description	user_created	user_followers	user_friends	user_favourites
0	Viola	Astroworld	wednesday addams as a disney princess keepin i...	2017-05-26 05:46:42	624	950	18775
1	Tom Basile us	New York, NY	Husband, Father, Columnist &	2009-04-16 20:06:23	2253	1677	24

			Commentator.				
			Auth...				
2	Time4fisticuffs	Pewee Valley, KY	#Christian #Catholic #Conservative #Reagan #Re...	2009-02-28 18:57:41	9275	9525	7254
3	ethel mertz	Stuck in the Middle	#Browns #Indians #ClevelandProud #[] #Cavs ...	2019-03-07 01:45:06	197	987	1488
4	DIPR-J&K	Jammu and Kashmir	✍ Official Twitter handle of Department of Inf...	2017-02-12 06:45:15	101009	168	101

According to Drus & Khalid (2019), sentiment analysis, also called opinion mining, entails natural language processing to categorize the data as negative, positive, or neutral. Sentiment analysis relies on textual data, which upon analysis, helps researchers analyze and organize the public's perspectives as expressed in their social media platforms. Alsayat (2022) points out that sentiment analysis encompasses perceptions, opinions, thoughts, attitudes, and individual emotions shared on different social media platforms. Like Drus & Khalid (2019), Alsayat (2022) explain that the main objective of sentiment analysis is to classify the collected textual data into three categories; neutral, positive, or negative. Sentiment analysis reveals the public's opinion about a product, event, or occurrence, thus being explored across different fields. The first technique that could be applied in collecting sentiment analysis data on Twitter is the lexicon-based approach. This technique uses a valence dictionary to label and categorize opinions as positive, neutral, or negative. The

lexicon-based approach calculates sentiment orientations of a set of sentences or an entire document from a pre-established semantic orientation of lexicons.

Like in any other research situation, data collection is the initial step in sentiment analysis. Based on the nature of sentiment analysis, the text's findings can be combined with other types of data, such as graphics (emojis), audio, and video. The researcher could opt for the dynamic data source when focusing on a social media platform like Twitter. Alaoui et al. (2018) note that the lexicon-based model of sentiment analysis involves the collection of specified sentiment terms that are categorized into two groups; dictionary and corpus-based. The dictionary-based approach is designed to identify opinion words in a selected text before matching them to their semantic orientation in the dictionary. The most common dictionary used for sentiment analysis when using the lexicon-based approach is the SentiWordNet, but a manually-created dictionary can also be used. On the other hand, the corpus-based approach is designed to find opinionated words in a specific context orientation. This technique starts with a list of opinion words and moves to find other opinion words in a much larger corpus. When using this approach, researchers focus on verbs and adjectives as the text's indicators of semantic orientations.

Even though data for sentiment analysis could be extracted from different social media platforms, Twitter has been found to provide the most straightforward way to extract the users' views using the application program interface (API). According to Hota et al. (2020), Twitter provides simple, practical, and straightforward data for SA compared to other social media platforms. As such, SA researchers could extract tweets through the use of API referred to as the TWINT API, a Twitter intelligence tool with a feature that allows it to work as an advanced scrapping tool that is written in Python. For this reason, TWINT API enables a researcher to scrap tweets from Twitter without necessarily using the Twitter API, making TWINT API more reliable than API. Even though TWINT API is considered an essential

aspect of data scrapping, some researchers have attempted to scrap Twitter data without using the TWINT API (Krishnan et al., 2021). Note that TWINT has been described as an advanced Twitter scraping tool written in the Python program. TWINT uses search operators to allow the scrapping of specific users and tweets that relate to certain trends, topics, and hashtags. TWINT can replace the need for API in the scraping phase of data for sentiment analysis. Like Krishnan et al. (2021), Sulthana et al. (2018) use API instead of TWINT. According to Sulthana et al. (2018), there are two types of API; rest and search API. Whereas the rest API allows the researcher to access core Twitter data, including timelines, user information, and status data, search API allows one to interact with search and trend data on the platform. Another advantage of TWINT without Twitter API is that it reduces noise by scrapping off the users' followers and likes without using any logins, Selenium, API, or browser emulation. The TWINT option in sentiment analysis data collection offers several advantages over API. For instance, TWINT could retrieve unlimited tweets, whereas API can only retrieve about 3200 of the last tweets. Another significant advantage of TWINT is that it provides several formats for storing scraped data. These formats include SQLite, CSV, JSON database, and Elastic search database. Furthermore, TWINTs are also quick to have initial setups and use them anonymously without necessarily having a Twitter account. Lastly, TWINT offers the developer more freedom than API as it does not have any rate restrictions.

4.1.4 Data Pre-Processing

Data pre-processing has been declared a major step towards data mining, and in this study, it plays a rather significant role in reproducing data for further processes such as feature selection and extraction. A simple tweet has been discovered to include several parts; words, user mentions, and emoticons. Generally, the process of data pre-processing is classified into three categories; tokenization, normalization, and substitution.

Researchers have defined tokenization as a method used in creating tokens mostly split by characters. In this context, Tokenization Space is first performed. The tokenization of a certain tweet is to simply convert it into a sentence and then section it into individual words. The conversion greatly simplifies the filtering or cleaning of unwanted characters and words and classifies them as noise. The process eliminates stop words, special characters such as @ or #, irrelevant numbers, and punctuations. Saif et al. (2014) defined stop words as frequently used words such as articles and pronouns that fail to contribute to the sentiment or meaning of a phrase. Normally, such words are discarded by search engines. Tokenization is a process that can be completed using SpaCy or Natural Language ToolKit (NLTK). According to Garg et al. (2016), the two are Python's most commonly used NLP packages.

Normalization is the retransformation of words back to their standard origin by eliminating unintentional errors, conversions, and unnecessary additions. The process of normalization includes lemmatization, negation, elimination, lowercase conversion, and numerous other techniques that depend on the implementation and particular requirements. However, in the case of sentiment analysis models, some techniques become essential since they can completely reverse the outcome, thus significantly impacting accuracy. Negation falls under the category. Contextually, negation is the removal of apostrophes from tweets with words such as “can’t” and replacing them with “not” to form “cannot.” There are numerous look-up tables and open-source dictionaries online that can be used for the elimination of apostrophes and related shortcuts. Lemmatization, another significant procedure, reverts words back to their original form or root. Unlike Stemming, lemmatization considers the context of the word. Also, the process can be implemented in Python using NLTK packages. As needed, more advanced and additional normalization techniques such as spelling corrections, grammar, slang, emoticons, lexicons, and dictionaries might also be considered moving forward.

Substitution focuses on removing noise from tweets in their raw format. Naturally, tweets have HTML markups as it is with any text obtained from the web. Additionally, in their tweets, Twitter permits its users to share URLs. In the substitution stage, the URLs and undesired HTML markups ought to be cleaned out. Moreover, decoding might be the most crucial substitution technique. The decoding process involves unifying data into a textual data encoding format. As discovered, textual data might be subjected to several decoding formats, such as Unicode and ASCII. Researchers such as Zhang et al. (2017) recommend the consolidation of data into a single widely recognized format due to memory efficiency and the compatibility of languages.

Generally, in the process of data pre-processing, studies include hashtags as they could be significant indicators of the semantic meaning of various tweets. Additionally, stop words such as ‘the,’ ‘an,’ and ‘a’ are maintained as they contribute to NLP, especially in the sentiment analysis of the tweets. Similarly, it has also been deemed necessary to normalize words with similar meanings for further evaluation, whereby normalizing such components might appear difficult without complex algorithms. However, the above cannot be applied to the scope of the current study. Emoticons are significant in sentiment determination; therefore, they are replaced by their polarity by using the emoticon dictionary, as shown below.

Table 3

Replacement of Emojis by Polarity (Siddharth et al., 2019)

Emoticon	Polarity
:~) :) :o) :] :3 :c)	Positive
:D C:	Extremely-Positive
:-(:(:c :[	Negative
D8 D; D= DX v.v	Extremely-Negative
:	Neutral

Once the data has been arranged in a structured format, finding the correct information becomes easier. Thus, data pre-processing steps for this study include the following;

1. Elimination of stop words, special characters, irrelevant numbers, and punctuations.
2. Advanced normalization to check and eliminate spelling and grammar errors, slang and emoticons, and the transferring of all characters into lowercase.
3. The cleaning out of URLs and undesired HTML markups and decoding.

The study used the Adam optimization algorithm to train deep learning models incorporating two stochastic gradient descent extensions, including RMSProp and AdaGrad. Furthermore, the study used ReLU activation functions, `sparse_categorical_crossentropy` for the loss function, and the softmax activation function for ternary classification. In Text numerical conversions, the Tokenizers in the Keras NLP library were used. This method creates the vocabulary index based on word frequency. So, if it was given something like, "The cat sat on the mat." It will create a dictionary s.t. `word_index["the"] = 1`; `word_index["cat"] = 2`; it is a word-to-index dictionary, meaning that every word gets a unique integer value. 0 is reserved for padding. So lower integer means a more frequent word (often, the first few are stop words because they appear a lot).

Data processing is often utilized to clean raw data by observing particular steps to achieve improved results for further assessment. Data often contain missing values, noises, and unusable formats that cannot be directly utilized for ML models. This process is critical for data cleaning and making it suitable for the training model, which helps increase the accuracy and efficiency of the model. The steps in data pre-processing include data cleaning, data integration, data transformation, and data reduction. Data cleaning helps clean the data by smoothening noisy data, filling in missing values, and resolving the inconsistency. The dataset was cleaned by removing incorrect and redundant values and analyzing the tweets to establish tweet emotions as either positive, negative, or neutral. The main steps of data

cleaning are deleting duplicate data, removing irrelevant items, checking for outlier data, correcting typos and structural mistakes, checking for missing data, and validating our data. Pre-processing was performed on the collected dataset by developing a user pre-processing function based on NLTK (a Python library for NLP). A function `def clean_text(text):` was created to clean data by removing all URLs, punctuation marks, text in square brackets, and words containing numbers and making all text lowercase. The following table represents sentiments after cleaning:

Table 4

Cleaned Sentiments

```
0    if i smelled the scent of hand sanitizers toda...
1    hey yankees yankeespr and mlb  wouldnt it have...
2    wdunlap realdonaldtrump trump never once clai...
3    brookbanktv the one gift  has give me is an ap...
4    july  media bulletin on novel coronavirusupda...
Name: text, dtype: object
```

Following data processing, the study developed the Bag of Words (BOW) model with the help of frequently occurring words from lexicons and acquired a list of the most frequently used COVID-19 words. The dense word that represents the most used in the corpus is illustrated in Figure 7 below. The Word Cloud revealed the top 50 prevalent words in the tweets. There are words like 'sanitizers,' 'Yankees,' and 'coronavirus' that have a maximum frequency in our dataset, as shown in the following word cloud. Several #hashtags are visible in the tweets but are all nearly similar in sentiments, and as a result, they do not provide useful insights into the tweets' emotions.

Figure 7

Most Popular COVID-19 Related Words within the Corpus



Tokenization involves converting the text into tokens, usually separated with white space. In this process, the word boundaries are established, the special characters are removed, and numbers and abbreviations are processed. Depending on the morphological complexity of a language, the tokens might reach four due to the number of clitics and affixes. In our dataset, tokenization was applied to chop character sequences in suitable format pieces and throw some characters, like punctuations. A vocabulary size of 167824 was obtained, and afterward, the y feature was converted into a 1D array. This helped to split tweets into smaller units that can be easily assigned meaning. Stemming helps minimize the inflected words from their word base, root, or word. This was necessitated by the tokenization function, which split each of the sentences into smaller parts of the word.

Since sentiment analysis relies on user-generated texts, which are informal and based on habits, it is apparent that the collected data will have many nuances. Tweets analysis is complicated because they are unstructured, as they contain much irrelevant content. For this reason, cleaning the data to remove noisy information and normalize the language before analysis is necessary. Available literature indicates that there are several steps in Twitter data preprocessing. However, the first step of preprocessing would differ according to the

language of the tweets. In the case of Vietnamese, the researcher will have to undertake the tokenization step. Unlike English, Vietnamese words tend to be segmented as they are separated by punctuation and whitespaces. Failure to process tweets written in the Vietnamese language alters the sentence's meaning. Tweets in other languages could be translated into English before progressing to other preprocessing techniques. Alaoui et al. (2018) point out that the first step of sentiment analysis is data collection and storage. However, for this data to be processed and categorized into positive, negative, or neutral groups, it has to undergo a preprocessing process. Alaoui et al. (2018) summarize the preprocessing stage of Sentiment analysis as the cleansing step of turning the informal data into an analyzable format. This objective is accomplished by identifying and removing non-textual information and misspellings through various filters.

The first step of preprocessing that comes up in several works of literature is tokenization. The step involves identifying content words like nouns and verbs, standard emoticons, HTML tags, and URLs. Note that *tokenization* has also been defined as splitting text into a list of practical pieces referred to as tokens. The tokens consist of phrases, words, symbols, units, or even sentences to allow the researchers to work in the text effectively. Since tokenization serves as an input of the following process, it makes up an essential process of NLP. Tokenization can be done using punctuations, line breaks, and whitespace, but in most cases, it uses whitespaces. Several tokenizers exist in natural language processing. The first one is the Regexp tokenizer which splits a statement/sentence using regular expressions for matching tokens. The other natural language processing tokenizer is the TreebankWord tokenizer. Like Regexp, TreebankWord Tokenizer splits the sentence using the matching tokens' regular expression. However, it has an additional feature that treats punctuations as words, meaning that punctuation marks such as commas and apostrophes also undergo splitting. The two examples are the most commonly used but not

the only ones, as researchers could choose a preferred tokenizer that matches their needs. The idea of tokenization is based on the concept that words are the building blocks of natural language. However, tokens offer a way of representing words that software can interpret. As such, the tokens are converted into unique numbers because computers can interpret numbers much better than words. As such, tokenization converts the entire text into word vectors which are in turn converted into numbers (Dankhara, 2022).

Another important step in sentiment analysis data preprocessing is the removal of URLs, punctuations, and emoticons. The impact of emoticons on sentiment analysis is debatable because emoticons could positively affect sentiment scores. Regardless of this factor, removing emoticons, punctuations, and URLs is considered necessary as these features are often treated as an additional dimension in each word. For this reason, most researchers often opt to trade off emoticons' positive features of increased sentiment score with a reduction of dimensions to increase overall accuracy. While emoticons present some controversy that creates a dilemma of whether they should be removed or treated as part of the text, experts seem to agree that URLs should be removed from the data set. In essence, URLs does not have to provide any additional analyzable information that could be important in sentiment analysis. URLs could introduce an element of confusion whereby a negative comment is analyzed as positive and vice versa. As such, removing URLs is critical in avoiding such situations. But as Isik & Dag (2020) argue, URLs are often crucial for specific applications as they provide insights that might not be obtained from the provided context.

Preprocessing data for text analysis also involves expanding both abbreviations and acronyms. Both abbreviations and acronyms could be described as forms of contractions used to minimize repetition and save time and space. However, the two concepts are pretty different as abbreviations refer to shortened forms of words whereby only the first few letters are used. In contrast, acronyms use the first letters of each word in a phrase. Expanding both

acronyms and abbreviations is commonly used in a messaging application. As such, their expansion is crucial for understanding the context and information contained in a particular text. Abbreviation and acronym expansion is carried out alongside data cleaning, which entails removing data that is not helpful. In essence, the so-called unhelpful data does not contain any useful information about the text or sentiment; if they do, extracting the information might be difficult (Dankhara, 2022). Removing un-useful data significantly improves the decision-making process by reducing the noise affecting the algorithm. The algorithm is further improved by removing white spaces and stop words. White space is a term encompassing features like tabs, new lines, and spaces, while stop words are words that serve to make a sentence grammatically correct but do not have any meaning on their own. Stop words and white spaces increase data complexity which affects the algorithm's accuracy. Some of the most common techniques of stop-word elimination include the use of precompiled lists and the selection of rarely occurring words that are not related to the specific text.

Data preprocessing also involves lemmatization and stemmatization. Both approaches ensure that different versions of a word are not left out. Stemmatization involves the reduction of the word into its root forms, while lemmatization reduces a derivational word to its dictionary form before categorizing it in its part of speech and inflectional form. Stemmatization could be viewed as a generalized approach to word reduction as it does not consider the word's position in the sentence. On the other hand, lemmatization offers a progressive reduction and categorization of words by converting them into their dictionary and morphologically-related format. Lastly, the researcher has to undertake other measures such as lowercasing and removal of common words. Both steps are intended to minimize the algorithm's noise and thus increase analytical accuracy.

4.1.4.1 Bias Detection. Through previous research , the researcher acknowledged that there is the potential for bias in the data or algorithms used to make predictions in the sentiments, which can lead to inaccurate or inappropriate decisions. It was noted that it could be easy to bias results as it is easy to change the sentiment of a word by choosing which words to include and which to exclude. This is especially true when the sentiment of a word is subjective. Various studies as cited in this study also acknowledged that sentiment analysis by word is not very accurate because it is easy to mistake a positive sentiment for a positive word and a negative sentiment for a negative word.

As such, the researcher employed bias detection algorithms to flag potentially biased predictions, and one such algorithm was the lexical bias analysis, which seeks to examine the actual words and phrases that have been used in the text data. This algorithm was also able to identify terms or patterns that were most likely to skew the analysis. The study also employed statistical bias detection where the analysis of distribution of sentiment scores across different attributes were analyzed.

4.1.5 Feature Extraction

As of now, the process of pre-processing is completed for the further improvement of the features that can be extracted for the purposes of classification. Datasets from Twitter normally contain several variables with continuous-valued features that ought to be converted to discrete variables that will offer more responses to the work of classification. It is important to note that when input sizes are large in size, it becomes more complex to work with such data. Therefore, it would be better to reduce the dimensionality of the dataset and predict the specific variables that best suit the process of data extraction.

The feature selection and extraction process is crucial as this is where pre-defined attributes are converted from human-readable (raw text) to machine-readable (numerical matrices or arrays in simpler terms). The process, at times, might be named “Word

Embeddings,” a simple implication of the transformation process of a text to numerical representation. In broader terms, the process is categorized into two distinct classes; prediction- and frequency-based. According to Djuric et al. (2015), techniques that follow frequency-based embedding, for instance, the Term Frequency- Inverse Document Frequency (TF-IDF) and Bag of Words (BOW), appear to be the most widely used in numerous NLP applications, and this is attributed to their fundamentally unsophisticated approach. For instance, TF-IDF weighs and assesses the importance of words based on the appearance of a particular word and not in a single tweet but in a complete set of tweets. This specific feature allows the technique to discard prepositions, pronouns, and other frequently used terms that do not contribute to the actual meaning. In fact, researchers such as Beel et al. (2016) claimed that more than 80% of all recommender systems use TF-IDF as their primary extraction technique.

In Python, TF-IDF can be incorporated using the Scikit-Learn ML library. On the other hand, prediction-based embeddings are more state-of-the-art regarding their detection of analogies. Unlike frequency-based embeddings, prediction-based embeddings efficiently detect associations, semantics, and context. Additionally, prediction-based techniques display words in a vector space presentation, whereby terms bearing similar contexts are positioned in close vicinity to one another. Global Vectors (GloVe) and Word2Vec are among the most widely used prediction-based embeddings implementation. Introduced in 2013, Word2Vec, in this study, will be used to achieve word embeddings with a two-layered shallow neural network for the purposes of detecting context. The algorithm itself consists of a combination of two other distinct algorithms; Skip-Gram and Continuous Bag-of-Words (CBOW). The algorithms are predicted to perform opposite tasks regarding context from word prediction and vice versa.

Word2Vec can also be implemented in Python using the Gensim library. The technique has also been widely used in different implementations of sentiment analysis because of its renowned accuracy. However, Beel et al. (2016) claimed that TF-IDF equally shows some promising results in different NLP applications, such as recommender systems. The current study considers and compares the abilities of both techniques for the purposes of accuracy validity. On the other hand, emojis embeddings are considered a fresh yet highly trending topic (Chen et al., 2018). Similar to Word2Vec, emojis can be embedded and vectorized based on their meaning and contexts. Throughout literature, there have been two methods that have gained recognition; Word2Vec and Skip-Gram. The Word2ve method is expected to display a more scientific approach using official emojis' Unicode keywords and definitions. The Skip-Gram, however, used frequently appearing nearby words from numerous English tweets for the interpretation of an emoji. Afterward, the descriptive words in the current study are managed as normal texts using the Word2Vec model. The two techniques are publicly available and can be imported as dictionaries; the current study, however, considers the Word2Vec approach (emojis2vec) as a result of the adoption of the official emojis Unicode.

4.1.6 Summary of Methods of Feature Extraction

Many feature extraction approaches have been developed by researchers over the years. The variety of methods can simply be explained by the variety of applications and test cases.

4.1.6.1 Traditional Bag-of-words (BoW). Studies such as Asghar et al. (2014) and Ahuja et al. (2019) have reiterated what this study has already established in terms of methods of feature extraction in sentiment analysis in terms of the various methods that could be used, especially when Twitter data is involved. Some of the key methods mentioned include the Bag of Words (BoW) Model, the Term Frequency-Inverse Document Frequency (TF-IDF) Model, Word Embeddings, Part-of-Speech (POS) Tagging, and Named Entity Recognition (NER). The BoW model involves creating a vocabulary of all the unique words

in the dataset and representing each tweet as a vector of word frequencies. This is what the current study does using algorithms specified in the next sections of this chapter.

As far as this study is concerned, the following steps were undertaken in creating a BoW model for sentiment analysis of Twitter data:

1. The first step was to collect a dataset of tweets labeled with their corresponding sentiment (positive, negative, or neutral). This can be done through data labeling and polarity detection.
2. The next step is to preprocess the tweets, and this can be done in different ways, but in this study, it was done by removing stop words, punctuations, URLs, and some special characters.
3. Then, a vocabulary of all unique words in the dataset was created. The unique words are critical for differentiating the data, and in the context of this study, it was done by tokenizing the preprocessed tweets and storing the unique tokens in a dictionary.
4. Afterward, a vector representation of each tweet by counting the frequency of each word in the vocabulary of the tweet was created. This is done by iterating over each word in the tweet and incrementing the count of the corresponding word in the vocabulary. The iterative process is to ensure all the sentiments have been captured.
5. The following step was to normalize the frequency vector by dividing each element by the total number of words in the tweet. This helps to account for the varying lengths of tweets.
6. Now, it is time to use the frequency vectors as input to a machine learning algorithm, such as Naive Bayes, Random Forest Classifier, or Logistic Regression, to train a sentiment classifier.
7. Once the classifier is trained, it can be used to predict the sentiment of new tweets by converting them into frequency vectors using the same vocabulary and normalization

technique used during training. The classifier then predicts the sentiment based on the frequency vector. Later in this chapter, different algorithms and codes that perform these steps have been demonstrated and shown.

One of the fundamental reasons this model was used in this study is that the BoW model is a simple and effective method for sentiment analysis of Twitter data. However, the researcher was aware of the limitations of this model, including the inability to capture the semantic relationships between words and the lack of context awareness. The researcher, therefore, had to think of other methods or ways of overcoming these limitations, and this is because other feature extraction methods, such as Word Embeddings and Named Entity Recognition (NER), were considered in conjunction with the BoW model.

4.1.6.2 Term Frequency-Inverse Document Frequency (TF-IDF). Term Frequency-Inverse Document Frequency (TF-IDF) is a technique that is used to determine the importance of words in a given document or tweet (Asghar et al., 2014). This method assigns weights to the words in the BoW model based on their frequency in the tweet and their rarity in the dataset. This helps in giving more importance to words that are rare and less importance to words that are common. The procedure for using TF-IDF in sentiment analysis, especially when it comes to Twitter data in the current study, was done using the following steps. As cited by (Ahuja et al., 2019):

1. **The first step is Data Collection:** here, the dataset of tweets is collected, which is relevant to the topic that needs to be analyzed, and in this case, sentiments on the COVID-19 pandemic. This could be done by using Twitter's API or by scraping tweets from Twitter.
2. **Text Preprocessing:** the next step was to perform text preprocessing steps like tokenization, removing stop words, stemming, and lemmatization to transform the raw text of tweets into a clean and structured format.

3. **Calculation of Term Frequency:** the next step is the calculation of the term frequency (TF) for each word in the tweet. This is the number of times a word appears in the tweet divided by the total number of words in the tweet.
4. **Calculation of Inverse Document Frequency:** Calculate the inverse document frequency (IDF) for each word. This is a measure of how much information a word provides, calculated by taking the logarithm of the total number of tweets divided by the number of tweets containing the word.
5. **Calculation of TF-IDF:** Calculate the TF-IDF score for each word in the tweet by multiplying the TF score by the IDF score.
6. **Sentiment Analysis:** Assign a sentiment score to each tweet based on the sum of the TF-IDF scores of the words with positive or negative sentiment.
7. **Model Selection:** Choose a suitable machine learning model like Naive Bayes and Logistic Regression to classify tweets into positive, negative, or neutral sentiments.
8. **Evaluation:** Evaluate the performance of the model using metrics like accuracy, precision, recall, and F1 score. This will help you identify the model's strengths and weaknesses and fine-tune it for better performance.

In model classification, different machine-learning techniques are used. In this study, the following techniques are used:

1. **Naive Bayes:** Naive Bayes, as explained in more detail in this chapter, is a simple probabilistic classifier that is often used for text classification tasks like sentiment analysis and works by calculating the conditional probability of a given class (positive, negative, or neutral) given the words in the tweet.
2. **Logistic Regression:** Logistic regression is a statistical technique that can be used for binary classification problems like sentiment analysis, and it works by modeling the

probability of a tweet being positive or negative based on the TF-IDF scores of the words in the tweet.

3. **Support Vector Machines (SVM):** this is a powerful algorithm that is often used for text classification tasks like sentiment analysis. It works by finding a hyperplane that separates the positive and negative tweets in the high-dimensional space of TF-IDF features.
4. **Random Forest:** Random Forest is an ensemble learning technique that combines multiple decision trees to improve classification performance. It can be used for sentiment analysis by training a random forest classifier on the TF-IDF features of tweets.
5. **Gradient Boosting:** Gradient Boosting is another ensemble learning technique that can be used for sentiment analysis. It works by combining multiple weak learners (decision trees) to create a strong learner that can accurately classify tweets based on their TF-IDF features.

4.1.6.3 Word Embeddings. Word embeddings are a powerful technique in natural language processing that can be used in sentiment analysis to represent words as dense vectors in a high-dimensional space. Here are the steps involved in using word embeddings for sentiment analysis:

1. **Data Collection:** Collect a dataset of tweets that are relevant to the topic you want to analyze. This could be done by using Twitter's API or by scraping tweets from Twitter.
2. **Text Preprocessing:** Perform text preprocessing steps like tokenization, removing stop words, stemming, and lemmatization to transform the raw text of tweets into a clean and structured format.
3. **Word Embedding Training:** Train a word embedding model like Word2Vec, GloVe, or FastText on a large corpus of text data to generate dense vector representations of words. This is typically done by feeding the model a large amount of text data and using unsupervised learning to learn word embeddings.

4. **Embedding Extraction:** Extract the word embeddings for each word in the tweets. This involves looking up the embedding vectors for each word in the pre-trained word embedding model.
5. **Sentiment Analysis:** Assign a sentiment score to each tweet based on the embeddings of the words with positive or negative sentiment. This can be done by averaging the embeddings of the positive and negative words in the tweet and comparing them to a predefined threshold.
6. **Model Selection:** Choose a suitable machine learning model like Neural Networks, Support Vector Machines, or Random Forests to classify tweets into positive, negative, or neutral sentiments based on the embeddings of the words in the tweet.
7. **Evaluation:** Evaluate the model's performance using accuracy, precision, recall, and F1 score metrics. This will help identify the strengths and weaknesses of the model and fine-tune it for better performance.

4.1.6.4 Word to Vector (Word2Vec). This machine learning algorithm is used for natural language processing tasks such as sentiment analysis. As Asghar et al. (2014) define, it is a shallow neural network trained to predict the context in which a word appears in a sentence. The algorithm can learn the semantic meaning of words by analyzing the patterns of their co-occurrences in large datasets. The following is the summary of steps that are taken in executing this technique:

1. **Data Collection:** Collect a large dataset of tweets labeled with their sentiment polarity (positive, negative, or neutral). Several publicly available datasets can be used for this purpose, such as the Sentiment140 dataset.
2. **Data Preprocessing:** Preprocess the data by removing irrelevant information, such as URLs, hashtags, and mentions. Also, remove stop words and punctuations and perform stemming or lemmatization to reduce the dimensionality of the dataset.

3. **Training Word2Vec Model:** Train a Word2Vec model on the preprocessed dataset. The model should be trained on a large corpus of the text to learn the semantic meaning of words.
4. **Vector Representation:** Represent each tweet as a vector by averaging the Word2Vec vectors of all the words in the tweet.
5. **Model Training and Evaluation:** Train a classifier such as Logistic Regression, Support Vector Machine, or Naïve Bayes on the vector representation of the tweets. Split the dataset into training and testing sets and evaluate the model's performance on the testing set. Use standard evaluation metrics such as accuracy, precision, recall, and F1 score.
6. **Prediction:** Once the model is trained, it can be used to predict the sentiment polarity of new tweets.

4.1.6.5 Global Vectors for Word Representation (GloVe). Global Vectors for Word Representation (GloVe) is a word embedding technique that can be used for sentiment analysis, and the steps used in this technique highlighted by Ahuja et al. (2019) are as follows:

1. **Data Preparation:** The first step in sentiment analysis is to prepare the data. You need a dataset of labeled reviews with a sentiment label (positive or negative). You can use existing datasets like the IMDb movie review dataset or create your own dataset.
2. **Text Preprocessing:** Before training the GloVe model, you must preprocess the text data. This involves tokenizing the text, removing stop words, and performing other text normalization techniques like stemming and lemmatization.
3. **Build a Vocabulary:** Next, you need to build a vocabulary of words that will be used to train the GloVe model. You can use a pre-trained vocabulary or build your own based on the data.

4. **Train the GloVe Model:** Once you have preprocessed the text data and built a vocabulary, you can train the GloVe model. The GloVe model is trained on a co-occurrence matrix, representing the frequency of words appearing together in a given window size. The model learns to assign a vector representation to each word in the vocabulary based on its co-occurrence with other words in the corpus.
5. **Create Document Vectors:** After training the GloVe model, you can use it to create document vectors for each review in the dataset. This involves averaging the word vectors in each review to get a single vector representation for the entire review.
6. **Train a Classifier:** Finally, you can use the document vectors to train a sentiment classifier. You can use many different types of classifiers, such as logistic regression, support vector machines, or neural networks.
7. **Evaluate the Model:** Once the classifier is trained, you can evaluate its performance on a test set of reviews. This involves calculating accuracy, precision, recall, and F1 score metrics.

4.1.6.6 The Techniques in Perspective. Following data processing, a Bag of Words (BOW) model was developed with the help of frequently occurring words from lexicons, and a list of the most frequently used COVID-19 words was acquired. The dense word that represents most used in the corpus. Several #hashtags were visible in the tweets but were all nearly similar in sentiments, and as a result, they did not provide useful insights into the tweets' emotions.

Emojis' hidden annotations and vague implications consist of additional/underlying meanings that are not immediately obvious when interpreting a smiley (Jackson, 2019). For instance, an emoji of a smiling face can appear to convey positive sentiment and imply sarcasm or insincerity. A crying face emoji can convey sadness, irony, or mock sadness. This is a major challenge exhibited by the deployment of emojis. Hidden and vague implications

are difficult to detect via text analysis alone (Lin, 2022). Sentiment lexicons and contextualized word embeddings are key strategies that can be deployed to identify and solve these challenges. The former approach uses tools like "SentiWordNet" and a sentiment analysis toolkit to identify hidden annotations and implications in emojis. The tool lists words and phrases associated with a specific sentiment, such as positive, negative, and neutral. These words and phrases can be used to identify the sentiment expressed in the text, which is useful in inferring the sentiment of the emojis deployed. For example, when a social media post contains the words "happy" and "excited" along with a smiling face emoji, it can be identified that the sentiment expressed in the post is positive. Nonetheless, when the same post contains the words "sad" and "disappointed" along with a crying face emoji, it is determined that the sentiment expressed is negative (Bai, 2019).

Combined machine learning with sentiment lexicons can be used to analyze the context in which emojis and sentiment lexicons are used, helping determine the hidden annotations and implications. Lin et al. (2022) acknowledge that electronic word-of-mouth on social platforms affects different sections, such as stock trading and market confidence. Thus, sentiment analysis of comments linked to stock markets becomes vital in predicting stock markets. Deep learning methods have been diversified as essential in coping with prediction issues. The model can be used to analyze the context in which emojis are deployed. This is done via emoji embedding –numerical representations of smileys used in machine learning models. The approach improves the accuracy of sentiment analysis, as it captures the meaning of emojis more interpretably by machine learning models. A study reveals that combining sentiment lexicons and emoji embedding enhances the accuracy of sentiment analysis by eight percent.

Contextualized word embeddings aim to capture word semantics in various contexts (Wiedemann, 2022). It is a powerful tool helpful in identifying hidden annotations and

implications in emojis. These embeddings offer numerical representations of words that consider the context in which they are used, assessing the sentiment expressed in the text. Several benefits are exhibited by integrating contextualized word embeddings to solve challenges posed by emojis' hidden annotations and vague implications. First, the tool easily pinpoints hidden annotations and implications in emojis by considering the context in which they are used. An excellent example involves a social media post containing words like sad and disappointment and a crying face emoji. It can be inferred that the sentiment expressed in the post is negative, and the crying face emoji depicts sarcasm. Contextualized word embeddings can offer a more concise understanding of the sentiment expressed in the text by analyzing the context in which the emojis are employed (Naseem, 2020). Another advantage of contextualized word embeddings includes easier and more effective combinations with other techniques, such as sentiment lexicons, to achieve better performance. For instance, sentiment lexicons can be utilized to identify the sentiment of words in a text, while contextualized word embeddings can be deployed to determine the sentiment of emojis. Naderalvojud et al. (2020) argue that combining sentiment lexicons and contextualized word embedding techniques offers a more accurate understanding of the sentiment expressed in the text.

4.1.6.7 Emoji Translation. Emojis represent visual symbols or ideograms that represent facial reactions, ideas, and concepts such as happiness and joy, laughter 😄, confused 😕, sadness 😞, or disappointment 😞. Emojis were a useful feature in the analysis of the tweets as they convey information and emphasize context for understanding the sentiments of the tweets in the dataset. The input sequence of the dataset was augmented with corresponding emojis, which were concatenated as additional tokens in the dataset. The model then learned to associate the emojis with the sentiments of tweets and utilized the additional features to classify the sentiments. The classifier learned the link between the existing emojis and the sentiment of tweets and utilized the acquired information to classify them based on emoji features. The first way to utilise the emojis was to transform them into tokens for the performance of the sentiment analysis. The emojis were translated in the format utilized as input features in the LSTM classification. The emoji-to-text translation library was used to translate the emojis into the text message. The 'demojize ()' function was used to replace emojis into text messages with their corresponding text representation, with the help of (':') to enclose the text representation of the emoji. In the experiment, the emojis were transformed based on predefined words such as:

- *transform the emotion found in smiley faces into the sentiment "positive."*
- *transform the emotions found in sad faces into sentiment "negative."*
- *transform the emotion found in love into the sentiment "positive."*
- *transform the emotion found in neutral faces into the sentiment "Neutral."*

Then, the tokens from the NLTK library were used as input features for the LSTM model by embedding them using the pre-trained embedding layers. The following output was obtained after translating the emojis in the tweets:

Table 5*Emoji Translation*

```

10    cases still no response from hrdministry dr...
11    order here logo graphicdesigner logodesign log...
12    waving_hand_light_skin_tone pattyhajdu navdee...
13    lets all protect ourselves from real and the ...
14    rajasthan government today started a plasma ba...
Name: text, dtype: object

```

on row 12 the emoji is replaced by text it represents in this case waving hand emoji

Emojis and emoticons represented the mere signs, which posed significant meaning in the collected tweets. The study accounted for the context of their appearance to understand the symbols or signs. In reference to the dataset, emoji usage depended on events following the precautionary measures set due to the COVID-19 pandemic. The principal functions into which emojis were used include sentiment expression, sentiment modification, and sentiment enhancement. In sentiment expression, the tweet might have a neutral emotional state. Thus, emojis add emotions and polarity to the text. Sentiment modification slightly modifies the intensity of the expressed emotion by lowering it. Sentiment enhancement enhances the emotions of the tweet to predict the polarities in the dataset.

4.1.7 Challenges of Data Collection in Sentiment Analysis and How to Solve Them

Sentiment analysis is based on the idea that words are the most powerful tool for expressing opinions, thoughts, intentions, and preferences. As such, accurate sentiment analysis is a source of important insights that can guide significant decision-making. However, sentiment analysis is complicated because human language is highly complex and thus presents a significant challenge for AI. One such challenge is context-dependent errors such as polarity and polysemy. In most conversations, sarcasm expresses negative sentiments, yet the words are often positive. In such a case, the sentiment analysis algorithms often

interpret and categorize the sentiments as positive despite negative. The solution for the sarcasm problem is to train the dataset for sarcasm, such as using a multi-head self-attention-based neural network architecture that identifies terms including sarcasm. This algorithm identifies and connects the sarcastic tone to the past to obtain a general score. The problem of polarity is experienced when the emotional tone does not allow clear inferences by the employed algorithms. The polarity problem can be resolved by giving the trained dataset polarity scores to enable the algorithms to classify the statement variations. Lastly, a polysemy challenge occurs when a word has more than a single meaning, making it difficult for an algorithm to differentiate the intended meaning. This challenge can be resolved by incorporating domain knowledge in text annotation/model training stages to help the algorithms differentiate the words' meanings in different contexts.

Besides the context-dependent errors, negation detection is another major problem experienced during sentiment analysis. In essence, a sentence having a negation does not necessarily mean that the overall statement is negation. The existing negation techniques have not yet advanced enough to classify sentiments sufficiently. As such, a view like 'the lockdown is not unpleasant' has a negation, and thus, an algorithm would classify it as a negative, yet the statement conveys a positive meaning. The negation detection issues could be resolved by training the algorithm using large datasets containing as many negation words as possible. Multilingual data and emojis are other challenges related to sentiment analysis data collection. Even though English is a common language worldwide, society is highly multilingual. Thus, people use different languages to express their views about a common topic. However, since algorithms are trained to analyze sentiments in a single language, researchers must rely on translation when dealing with reviews from other languages. Translation poses the risk of losing particular sentiments in translation. The solution to multilingual data is designing a system that can be trained to handle multilingual content.

Lastly, the issue of emojis could be resolved by implementing them in the analysis algorithm to improve the overall accuracy of the analysis.

4.2 Model Training & Algorithm Development

This subsection will describe and explain the adopted approach used to develop the various models and algorithms and how these models have been trained and improved to achieve the most optimum possible result.

4.2.1 Algorithm Development

Different sentiment analysis models were developed to establish the accuracy of the built approaches in predicting sentiments with the COVID-19-related Twitter dataset. The data was trained with several ML classifiers that have proved higher accuracy in predicting sentiments in similar cases. The machine learning approaches include K-Nearest Neighbour, Random Forest classifier, Logistic Regression, and Long Short-Term Memory. The experiment examined the most accurate and efficient approach to classify the disaster tweets associated with the COVID-19 pandemic awareness and then utilized the accurate model to classify tweets to predict the sentiments about the pandemic.

4.2.1.1 K-Nearest Neighbour Model. KNN is utilized for sentiment analysis because it can support multiple labels. In classification, it is applied for efficiency and simplicity. It works by computing the distance target value and other values, establishing the most adjacent k-nearest values, and utilizing it to vote for labels. The algorithm selects the k-nearest data points for classification problems and assigns the commonly identified class among the neighbouring data points. It is based on supervised learning approaches and stores the available data, and based on the similarities, it classifies the new data point. This ensures the acquired data is easily classified with the help of the KNN approach.

The benefit of KNN is that the model is non-parametric; thus, it makes data distribution assumptions. Nonetheless, it is computationally expensive when large sets of data are

involved. As a brief overview in this subsection, KNN was applied to compare real and predicted values of the first 50 tweets in the cleaned dataset. After pre-processing, the data was split into training and testing datasets using the fit function *knn.fit(x_train, y_train_1d)*. The training dataset was utilized in training the KNN model, whereas the testing dataset helped evaluate the model's performance. The number of neighbours was selected as 5 (*n_neighbors = 5*). The number of neighbours is considered in the algorithm to make the prediction. Thus, the optimal value of the neighbour can be selected using approaches such as grid search or cross-validation. The performance of the trained KNN approach on the pre-processed testing dataset is based on metrics such as F1 score, recall, precision, and accuracy. The rationale behind using KNN is that it is a useful model for sentiment analysis in machine learning, especially when dealing with data that contains text and emojis. KNN is a simple and effective algorithm that can be used for both classification and regression tasks. One of the main advantages of KNN is its ability to handle non-linear data and complex relationships between features. In the case of sentiment analysis, this is important because the presence of emojis can significantly influence the sentiment of a message, and their meaning is not always directly related to the words used in the message.

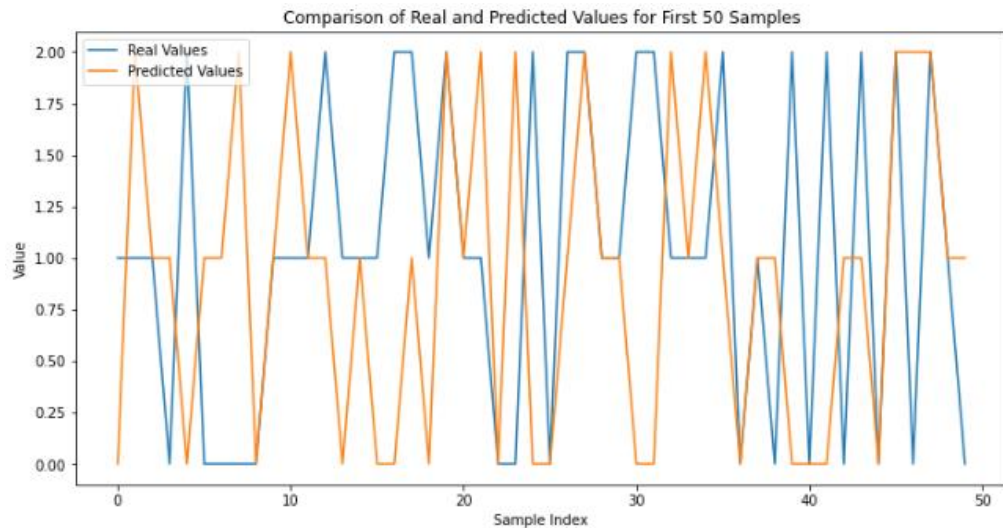
KNN works by finding the k-nearest neighbors of a new data point and using their labels to make a prediction, as mentioned. This approach is particularly useful in sentiment analysis because it can help capture the overall sentiment of a message rather than focus on individual words or phrases. Moreover, KNN is a flexible model that can be easily adapted to handle different types of input data, including text and numerical features. This makes it a good choice for sentiment analysis tasks involving a mixture of text and other data types, such as emojis.

The figure on the next page shows the application of the KNN classifier. It compares true values and predicted values in the dataset. Comparing these values is critical in

evaluating the performance and effectiveness of the model. It provides insights into model performance, feature importance, error analysis, and continuous monitoring. The comparison helps to enhance the accuracy, robustness, and generalization of the model, resulting in improved decision-making in different real-world applications.

Figure 8

KNN Classifier



The x-axis represents tweet sample indices ranging from 1 to 50, and the y-axis represents sentiment values, where 0 indicates negative values, 1 indicates positive values, and 2 indicates neutral values. The blue line represents the predicted values by the KNN model, and the orange line represents the values from the test data. The model's performance can be visually assessed by comparing to real value with the predicted values. The predicted values should closely align with the real values. If the lines are closely following each other, it implies that the model is making accurate predictions. However, if there are significant deviations between the lines, it shows that the model is not performing well. The above lines overlap, suggesting that the model might capture the dataset's underlying patterns and make predictions that match the true labels.

4.2.1.2 Random Forest Classifier. The Random Forest algorithm randomly performs observation and feature selection to create multiple trees. It then calculates the average

results. The model generates the subsets of features and creates smaller trees based on built subsets. It produces improved accuracy by handling missing values, maintaining the accuracy of large datasets, and cross-validation. Additionally, the classifier has a feature bagging aspect that helps maintain the accuracy of prediction when a certain portion of data is missing. Random Forest Classifier is highly accurate and suitable for overfitting. It can handle large datasets with high dimensionality and is utilized for regression and classification tasks. The model can be trained on sentiment categories and labelled datasets of text. It learns to predict patterns in the dataset that are related to several sentiment classes. During prediction, the model takes the text dataset and utilizes the learned patterns to make predictions on the text. The algorithm outputs the probability distribution over several classes of sentiments, and the class with the highest probability is chosen for the final prediction. A Random Forest classifier is an appropriate model in the analysis of sentiments to predict complex patterns in the dataset and handle variability in the dataset. Also, the model handles high-dimensional data and is less vulnerable to overfitting problems.

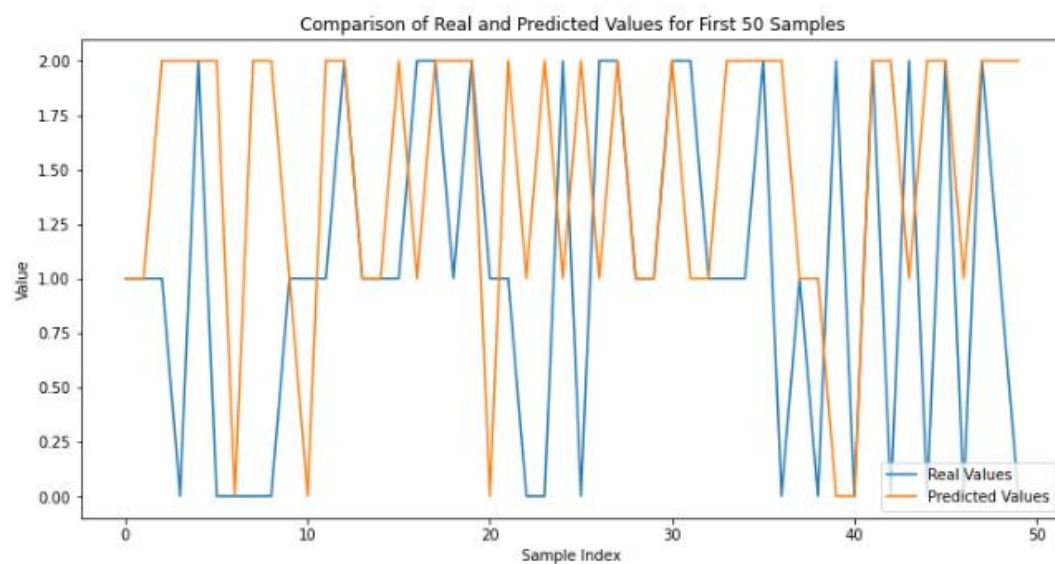
It combines multiple decision trees to produce a more accurate and robust model. The rationale for applying this method is that it can handle both numerical and categorical data, making it useful for sentiment analysis tasks that involve a mixture of text and other features. It can also handle missing values and noisy data often present in real-world datasets. Random Forest works by building multiple decision trees, where each tree is trained on a random subset of the features and data. This randomness helps to reduce overfitting and increase the model's generalization performance. During prediction, the output of each tree is combined to produce a final prediction. In the case of sentiment analysis, Random Forest can help to capture the complex relationships between words and their corresponding sentiment. It can also handle the presence of emojis in the data, as they can be treated as categorical variables. Another advantage of Random Forest is its ability to provide feature importance scores,

which can help identify the most important features contributing to sentiment classification. This information can be useful for further analysis and feature selection.

As a brief demonstration, the real values were compared to the predicted values of the first 50 samples of the COVID-19 tweet-related data. The following figure demonstrates a change in value in the selected sample index:

Figure 9

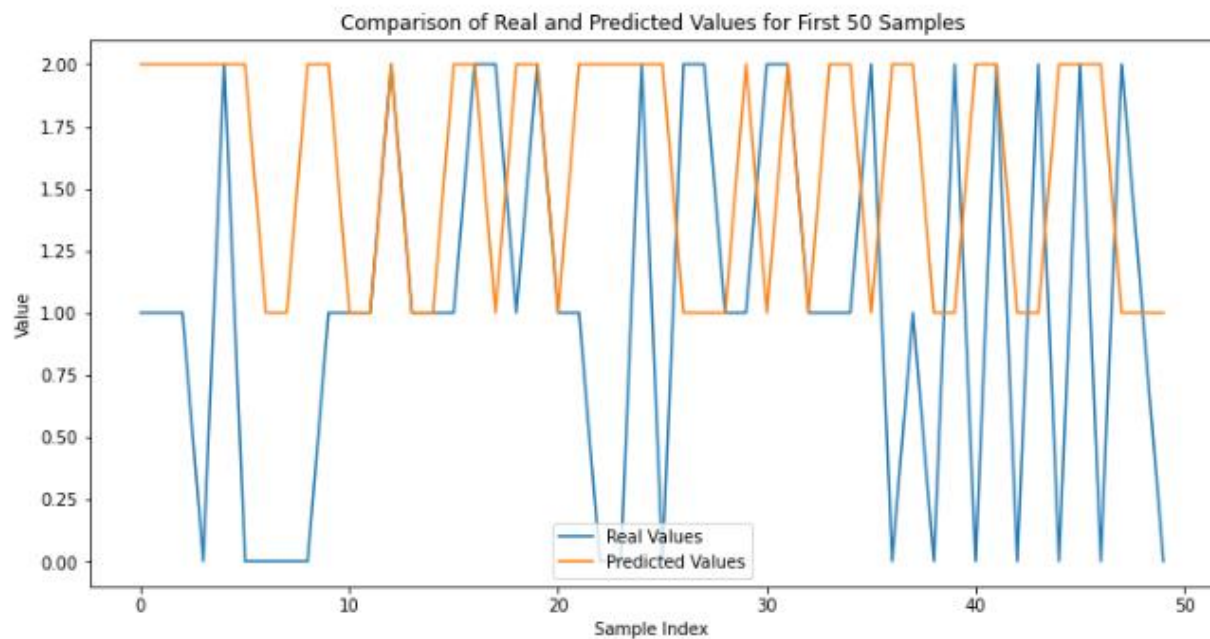
Random Forest Classifier



The x-axis represents tweet sample indices ranging from 1 to 50, and the y-axis represents sentiment values, where 0 indicates negative values, 1 indicates positive values, and 2 indicates neutral values. The blue line represents the predicted values by the Random Forest model, and the orange line represents the values from the test data. The model's performance can be visually assessed by comparing the real value with the predicted values. The predicted values should closely align with the real values. If the lines are closely following each other, it implies that the model is making accurate predictions. However, if there are significant deviations between the lines, it shows that the model is not performing well. From the above figure, the predicted and real lines are not closely aligned, indicating that the model might perform poorly.

4.2.1.3 Logistic Regression Model. A logistic regression classifier is a generalized linear approach where response records are a success or failure for a specified event. The model can be used to assess the corpus and perform multilabel classification. The model is usually for binary classification to predict whether a specified input belongs to a given class. It takes the set of input features and utilizes the mathematical function referred to as the logistic function to change the linear combination of input features into a probability value between 0 and 1. In training, the model uses the input features and computes the probability value for the existing classes. The predicted class is identified by choosing the class with the highest probability value. This algorithm is critical in binary classification, particularly when linearly separable and the link between input features and the outcome variable is linear. The model is first trained on the labeled dataset and sentiment classes to apply logistic regression in sentiment analysis. The input features are created with the help of multiple approaches like word embeddings and bag-of-words. The model is trained to learn the correlation between sentiment classes and input features. With regard to prediction, the logistic regression takes a new dataset as input and outputs the probability of sentiments as positive or negative. When the probability exceeds the predefined threshold (threshold = 0.5), the sentiment will be classified as positive; otherwise, will be negative. In sentiment analysis, the model's performance can be enhanced using different techniques such as regularization, tuning, feature engineering, and hyperparameter. Regularization approaches such as L1 and L2 are regarded as most appropriate as they help prevent model overfitting to the training data.

In the current study, the accuracy of the Logistic Regression Model was evaluated by comparing the real and predicted values of the selected samples. The model was trained using the *lr = LogisticRegression(solver='lbfgs', max_iter=1000)*, and the data was split into y and x sets to fit the model. The results are shown in the below figure.

Figure 10*Logistic Regression Classifier*

The x-axis represents tweet sample indices ranging from 1 to 50, and the y-axis represents sentiment values, where 0 indicates negative values, 1 indicates positive values, and 2 indicates neutral values. The blue line represents the predicted values by the Logistic Regression model, and the orange line represents the values from the test data. The model's performance can be visually assessed by comparing the real value with the predicted values. The predicted values should closely align with the real values. If the lines are closely following each other, it implies that the model is making accurate predictions. However, if there are significant deviations between the lines, it shows that the model is not performing well. As shown in Figure 10, there are inconsistencies between the predicted and real lines, implying that the Logistic Regression classifier might not capture the underlying features in the sample dataset correctly.

4.2.1.4 Multinomial Naïve Bayes. The text classification process can be performed with the help of the Naïve Bayes Classifier. The model is a supervised machine learning

algorithm based on Bayes' theorem. It is often used in NLP tasks like sentiment classification to identify the text corpus's emotional or opinion or sentiment tone. Naïve Bayes is a powerful classification model that incorporates datasets for high-dimensional training that helps generate fast machine learning classification models that make predictions with ease and quickly. This algorithm is mostly applied in text analysis, article classification, and sentiment analysis.

The Bayes theorem can be explained based on the following relation:

$$Prob(A/B) = (Prob(B/A) \times Prob(A)) / Prob(B)$$

Table 6

The Bayes Theorem

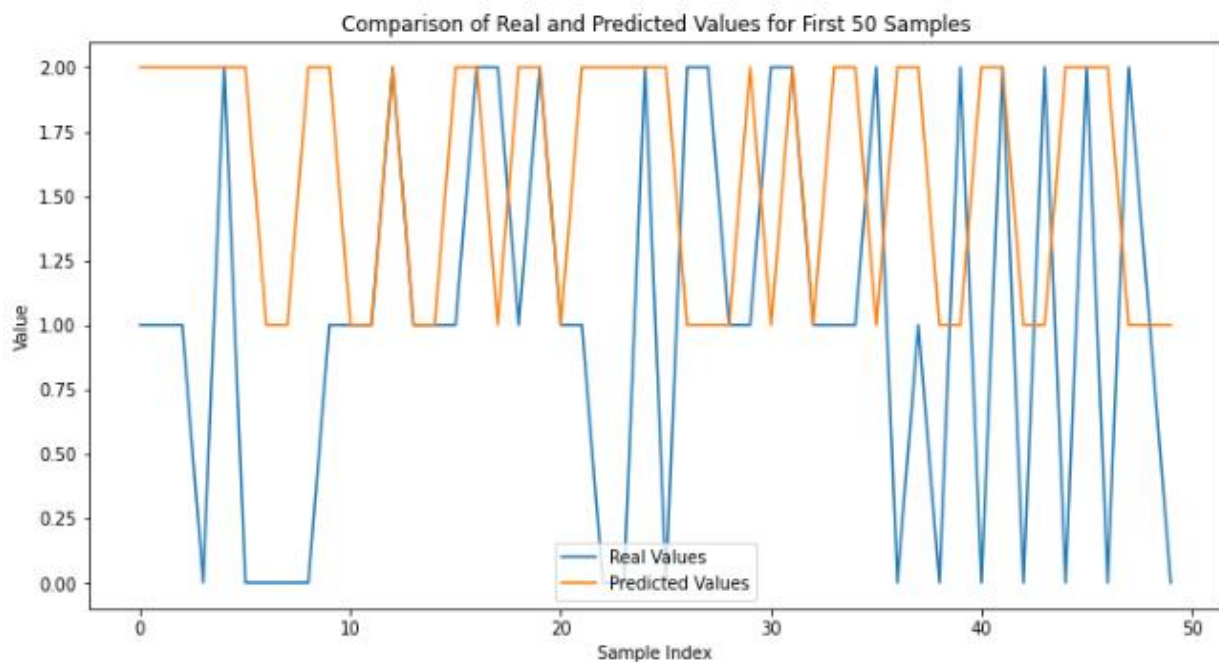
Element	Description
Prob (A/B)	Represents the posterior probability of hypothesis A, which occurs after a given condition is already met.
Prob B	It shows the likelihood probability. It represents the probability of evidence E when the given hypothesis is presumed to be true.
Prob A	The initial probability of A which does not have any condition.
Prob B	Prior known probability of B, which does not have any condition.

In mathematics, the Bayes theorem is applied to compute conditional probabilities. The theorem assumes that features are independent of each other conditionally, and all of them hold the same significance. This algorithm works by computing the probability of different classes in a set of input datasets, selecting the category with the highest probability, and utilizing the class for prediction. The Naïve Bayes model usually takes a short time and a

short prediction time. To apply the Naïve Bayes model, the model is first trained on the labeled dataset, where the dataset contains input features and their corresponding class label. The algorithm evaluates the probabilities of each input and given class and the prior probabilities of each class during training. The probabilities will be later applied in making predictions on new and cleaned datasets.

Even though the model is simple and efficient, the assumption about the independence of the inputs might not always be true, limiting its application in predicting the emotions and sentiments in the tweets. The sentiments in the dataset can be impacted by the context of texts, which translates to poor performance of the model on tasks critical in evaluating emoticons and emotions in text. The model lacks the context of the text in the dataset, which might have contributed to reduced performance in texts that required grammar of the text. In the study, the dataset exhibited imbalanced classes. Multinomial Naïve Bayes is usually inefficient in datasets that have imbalanced classes. This might negatively affect the overall performance of the model. Despite the limitations, the model is still effective and popular in sentiment classification.

To demonstrate the performance of the model, the real and predicted values of the first 50 samples of the dataset were assessed and compared, as shown in the figure below.

Figure 11*Multinomial Naive Bayes*

The x-axis represents tweet sample indices ranging from 1 to 50, and the y-axis represents sentiment values, where 0 indicates negative values, 1 indicates positive values, and 2 indicates neutral values. The blue line represents the predicted values by the model, and the orange line represents the values from the test data. The performance of the model can be visually assessed by comparing to real value with the predicted values. Ideally, the predicted values should closely align with the real values. If the lines are closely following each other, it implies that the model is making accurate predictions. On the other hand, if there are significant deviations between the lines, it shows that the model is not performing well. Figure 11 shows inconsistencies between the predicted and real lines, suggesting that the model is performing poorly.

4.2.2 Deep Learning Neural Networks

Many consider Deep Learning Neural Networks to be the bleeding edge of today's Machine Learning or Artificial Intelligence technologies. In simple terms, these work by

mimicking the human brain using a set of mathematical equations and algorithms. This subsection will explore the two most recent and promising models: LSTM and Transformers.

4.2.2.1 LSTM. In conventional sentiment analysis, LSTM has been established as more effective and efficient compared to similar neural models. The experiment exploits the sequential LSTM model to perform sentiment analysis of the COVID-19-related dataset. An embedding layer was utilized for the model architecture to generate the embedding vector for each input sequence. The embedding layer maps the words in the input sequence to the high-dimensional vector space, where the given dimensions represent different attributes or features of the text data. It can be initialized with pre-trained word embeddings, including GloVe or Word2Vec, or can learn the embedding of words from zero during training. Word2Vec model utilizes a neural network model to learn from a large corpus of text. It captures the context of words and simultaneously proposes an optimal way to preprocess raw text data. The words are preserved using the continuous bag-of-words and skip-gram model by considering neighboring words through word embeddings. The model is trained on the context of words in a corpus and aims to learn word vectors that capture semantic connections and store the syntactic features. It can effectively capture analogies and semantic connections easily, thus, is appropriate for sentiment classification. However, Word2Vec requires more training time and computational resources than GloVe, making it unsuitable, particularly for this study. In comparison, GloVe is easy to use and implement as it provides pre-trained resources and approaches that can be utilized for sentiment analysis tasks.

The approach is trained on co-occurrence statistics of words in a corpus, which allows easy learning of word vectors that have global tweet-tweet relationships. Moreover, it presents pre-trained word embeddings on a large corpus that can be directly applied to sentiment classifications. The provided embeddings can be accessed in multiple dimensions and trained on several datasets. This experiment utilized GloVe word embeddings, which is a

supervised algorithm that allows the acquisition of vector representation for words. It encodes the co-occurrence probability ratio between two words as vector differences.

The GloVe _EMB model was trained on non-zero entries of the global word-word co-occurrence matrix that records how frequently words co-occur with one another within a specified corpus. The tools provided in the Python environment automate the collection and preparation of co-occurrence statistics for input into the model. The training objective of GloVe is learning the word vectors so that their dot product is equivalent to the logarithm of the word's probability of co-occurrence. Upon execution, 400000-word vectors were found in the dataset. Vectorization through embedding helps represent words as semantically-useful dense real-valued values to overcome encoding-related issues. It enhances the generalization and performance of the LSTM model in training the dataset. Moreover, in training the model, the embedding layer optimizes word embeddings to lower the loss function and enhance the performance of the sentiment analysis task. It also includes contextual information like the order of words in the dataset or the presence of modifiers to enhance the strength of the embedding vector. Simply put, the embedding layer can be utilized in the preprocessing phase of sentiment analysis to generate vectors that are fed to the LSTM.

Additionally, the Conv1D Layer convolves data in smaller feature vectors and sequential data like text. In LSTM, the Conv1D layer is usually used in the algorithm to extract features from the text data of the dataset. The approach does not convolve data in small feature vectors but performs one-dimensional convolution on the dataset to extract key features. The Conv1D layer is fed into the LSTM model to capture dependencies in the text data. As a result, the LSTM model learns to categorize the sequence of words as either negative or positive. Using Conv1D, the model will extract meaningful insights from the dataset, which can enhance the accuracy of the sentiment analysis task. Finally, the Dense helped to connect layers in the dataset for classification. Dense layers are used as final layers

in performing the classification of the dataset. Once has been processed by the LSTM algorithm to extract useful insights, the model output represents the sequence of a single output vector.

The output vector is fed in the Dense layers to perform the nonlinear transformation of the input vector and output the vector of probabilities that represents the predicted class labels in the dataset. The possible classes that the model can predict are represented by the number of neurons in the Dense layer. In classifying the sentiments in the dataset as either positive negative, or neutral, the last Dense layer usually contains one neuron that outputs a value between 0 and 1 to represent the probability of the sentiments that belong to the positive category. Suppose the sentiment analysis task aims to categorize the sentiments into more than two classes (positive, negative, and neutral). In that case, the last dense layer will contain several neurons, where each neuron corresponds to one of the potential classes.

A new dataset was created that encompassed cleaned, preprocessed, and labeled tweets along with their corresponding negative (0) and positive (1) sentiments. Successively, two sets, x and y , were developed for the processed and classed tweets, along with their sentiment scores. A large number of COVID-19 exclusive words were generated by the LSTM model from the new dataset. Then, the words were converted into word vectors using Conv1D by setting the vector as 64 for collected N-grams within a tweet and building new x_{train} and x_{test} sets that consist of computed word vectors for data to be processed further. The TensorFlow framework and Keras library were utilized to add the LSTM model with Dense layers.

When configuring the LSTM, several key parameters were considered. The model was trained for ten epochs with two activation functions, an optimizer and a loss function. The model incorporated 300 embedding layers (EMBEDDING_DM) units, which help capture complex units in the model. The model expected the input sequences to have a certain shape.

The selection of the 300 embedding layers was due to the size of the dataset (the dataset involved was large) and the pre-trained embedding (GloVe), which can be leveraged by aligning with the dimensionality, high learning benefits, and improved model capacity. High embedding dimensionality helps generate denser and more continuous representations, which, as a result, improves the model's capacity to capture and generalize key patterns in the dataset. The input shape is defined as `batch_size`, `timesteps`, and `input_features`). The `batch_size` represents the number of samples in each batch. `Timesteps` represent the input sequence length, and `input_features` represent features recorded at each time step. The input shape had a `batch_size` of 10000, `timestep` of 32, and 24 input features (input features help capture and represent the information contained in the input data). The study applied the ReLU activation function for the initial dense layer with 512 units and the softmax activation function for the final Dense Layer with three units. The applied activation functions were integral to LSTM's input gate, forget gate, and output gate. During model training, the dropout approach was applied to the LSTM layer to prevent overfitting and improve generalization. The dropout for the x function was specified at 0.5. Similarly, the recurrent dropout (`recurrent_dropout = 0.2`), which functions as regularization, was applied to prevent the model from relying too heavily on certain memory cells. It was used together with regular dropout (`dropout = 0.2`) for bidirectional LSTM. Since the model involved multi-class classification, a sparse categorical cross-entropy loss function was used. These configurations contributed to building an optimal model for the prediction of disaster-related sentiments.

Computational constraints influenced the selection of the input features. The LSTM model is resource intensive; thus, using many input features would have increased the complexity of the model and memory requirements. Table 7 below summarizes the input features utilized in the experiment. The text represents the tweets used for analysis to predict public emotions and sentiments toward COVID-19. The `user_location` was also critical since

the collected tweets were filtered based on locations to show how individuals from different countries differed in their perception of the pandemic. User_followers, user_description, user_favourites, and user_verified informed the kind of tweets or information people shared on the platform. Other features were dropped since they were not important predictors of the desired output.

Table 7

Input Features

Input Features	
1	User_Name
2	User_Location
3	User_Description
4	User_Created
5	User_Followers
6	User_Friends
7	User_Friends
8	User_Favourites
9	User_Verified
10	Date
11	Text
12	Hashtags
13	Source
14	Is_Retweet
15	Keywords

17	Emojis
18	Sentiment Lexicon
19	Sentence Structure
20	Part-of-Speech (POS) Tags
21	N-grams
22	Average Word Length
23	Number of Stopwords
24	Number of Numerics

Figures 12 and 13 show the training accuracy vs. validation accuracy, and training loss vs. validation loss, respectively.

Figure 12

LSTM Training Accuracy vs. Validation Accuracy

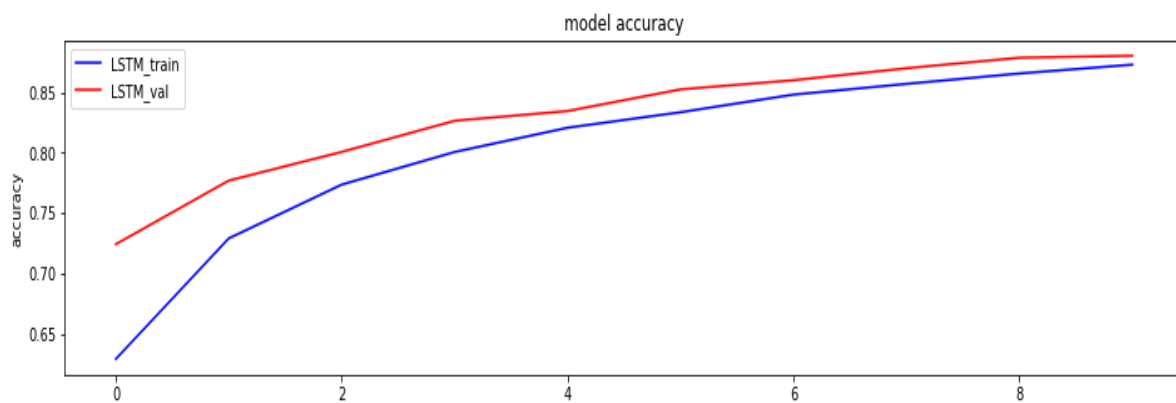
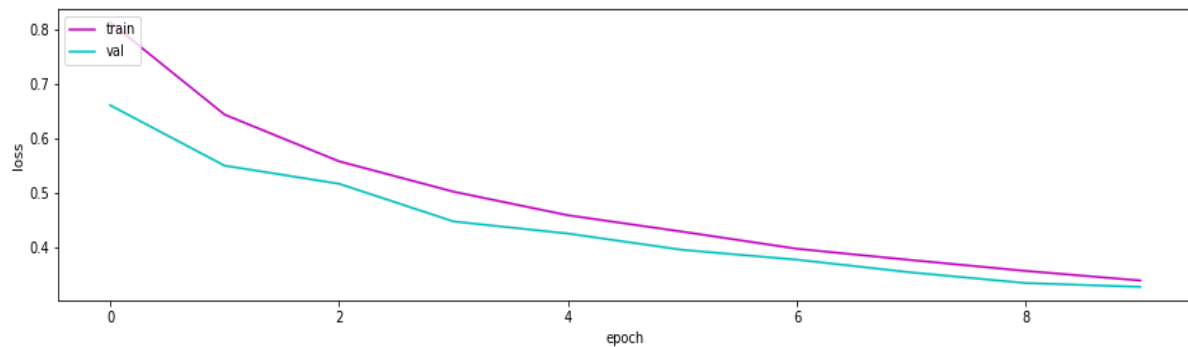


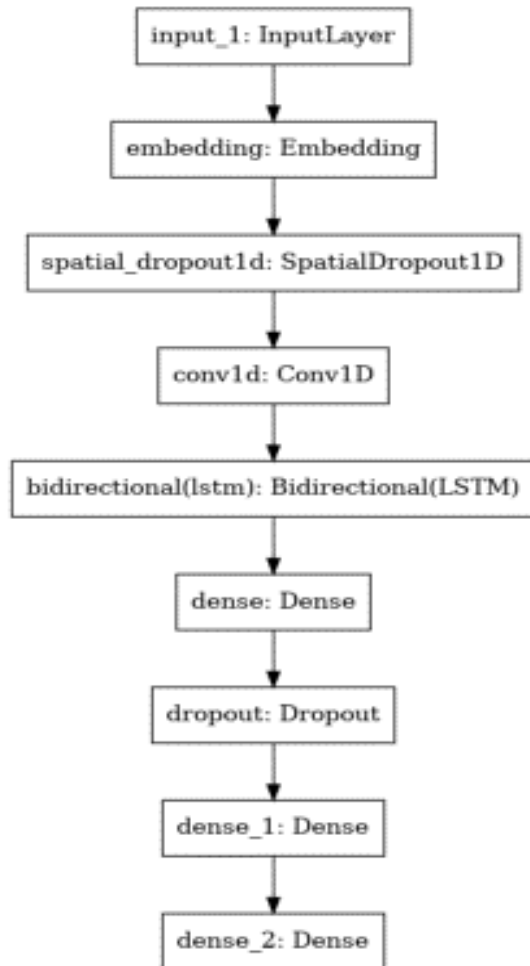
Figure 13*LSTM Training Loss vs. Validation Loss*

The values of training accuracy vs. training loss and validation accuracy vs. loss were compared as attained by the LSTM modeling during compilation. From the above graphs, it is clear that there is a significant loss difference between the training and validation epochs. In Figure 12, training and validation accuracy are increasing steadily. This shows that LSTM is a good model since both the validation accuracy and training accuracy. This model fits the training dataset and can be generalized to new datasets.

In the experiment, the Bidirectional LSTM architecture was utilized since the architecture can produce more meaningful output by integrating LSTM layers from both layers. Sentiment analysis can benefit from combining different features, such as word embeddings or syntactic information. The Bidirectional LSTM can easily combine these features and discover their relationships, leading to improved sentiment prediction. Throughout the process of training and executing the LSTM, 50,079,215 parameters were involved, where only 170,915 nodes were trainable, and 49,908,300 were non-trainable. Figure 14 (next page) and Table 8 (below) summarize the nodes used in modeling the LSTM algorithm:

Table 8*Nodes for LSTM Model*

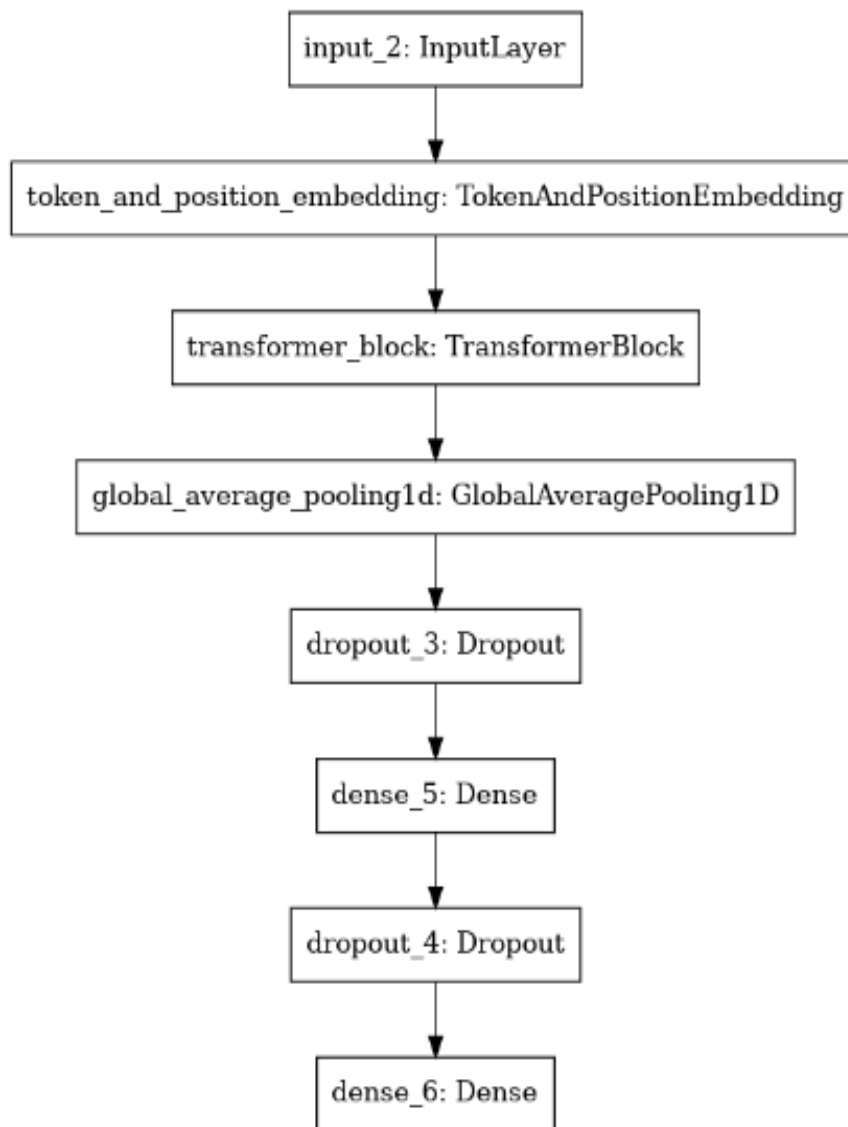
	Layer(type)	Output Shape	Param#
0	input_2 (InputLayer)	[(None, 50)]	0
1	embedding (Embedding)	(None, 50, 300)	49908300
2	spatial_dropout1d (SpatialDr)	(None, 50, 300)	0
3	conv1d (Conv1D)	(None, 46, 64)	96064
4	bidirectional (Bidirectional)	(None, 128)	66048
5	dense (Dense)	(None, 64)	8256
6	dropout (Dropout)	(None, 64)	0
7	dense_1 (Dense)	(None, 8)	520
8	dense_2 (Dense)	(None, 3)	27

Figure 14*LSTM Model*

4.2.2.2 Transformers. The transformer neural network is an architecture that aims to handle sequence-to-sequence tasks while solving long-range dependencies with ease. The model is effective in NLP tasks such as language modeling, text classification, and machine translation. It is based on the self-attention mechanism that enables the model to process variable-length input sequences and denotes huge dependencies between multiple sequence elements. This is attained through the computation of attention scores between different sequence elements and successively combining the input sequences with attention scores to generate a weighted representation of the sequence. Transformers consist of encoders and decoders, which have several self-attention layers and feedforward neural networks. Each

decoder in the model consists of three parts: the self-attention layer, decoder attention layer, and feed-forward neural net. Once the task in the first encoder is completed, the result is passed to the next encoder, and so on, until the last one. The encoder processes input sequences, whereas the decoder help generates the output based on input and the target sequence. The most common transformer model is Bidirectional Encoder Representations from Transformers (BERT), which has been established as effective in performing NLP tasks. Other models include Text-to-Text Transfer Transformer, Robustly Optimized Bert (RoBERT), and Generative Pre-trained Transformer (GPT). The aforementioned models are suitable for sentiment analysis to perform specific tasks on selected datasets. Transformers help perform complex sentiment analysis to identify the public emotions and reactions towards specific events.

The transformer model contains encoder-decoder attention layers, encoder self-attention layers, and decoder self-attention layers. In the encoder-decoder attention layer, queries are extracted from the preceding decoder layer, and the values and memory keys come from the decoder output. This allows different positions in the decoder to attend to all positions in the input sequence. In the self-attention encoder, all keys, queries, and values originate from the same place. Each position in the encoder can attend to all positions in the encoder's previous layers. Additionally, the self-attention layer in the decoder allows different positions in the decoder to attend to all positions in the decoder, including and up to that position. For every input, attention weighs the relevance of each other part and draws from the decoder to generate the output. The decoder and encoder layers are based on the feed-forward neural network to allow for further processing of the outputs and contain residual connections. The figure below shows the transformer model architecture:

Figure 15*The Transformer Model Architecture*

In this experiment, the multi-head attention layer available in TensorFlow /Keras deep learning framework was utilized. It performs several attention calculations in parallel with different heads or sets of weight. The study employed learned embeddings to convert input and output tokens to convert the decoder output to predicted next-token probabilities. Two embedding layers were created: the tokens and token index positions to manage positional embeddings. The experiment constituted a 32 embedding size for each token, two attention

heads, and 32 hidden layers in the feed-forward network inside the transformer. Each head learned to attend to different input sequences, which allowed the model to establish complex correlations between phrases and words in the tweets. While training the model, the dataset was split into the train set (x_{train}) and test set (x_{test}) based on an 80:20 ratio. The test size was 0.5, and the random state was 42. A ReLu activation function was applied for the initial activation layer of 20 units, and the softmax activation function was applied for the final three dense layer units. Additionally, Adam was used to optimizing the model (maintain a running estimate of variance and mean of the model gradient weight), the sparse categorical cross-entropy tested loss, and the metrics were applied to test model accuracy. A batch size of 64 and 10 epochs was involved during model training. Figures 16 and 17 indicate the training accuracy vs. validation accuracy and training loss vs. validation loss of the transformer model.

Figure 16

The Transformer Training Accuracy vs. Validation Accuracy

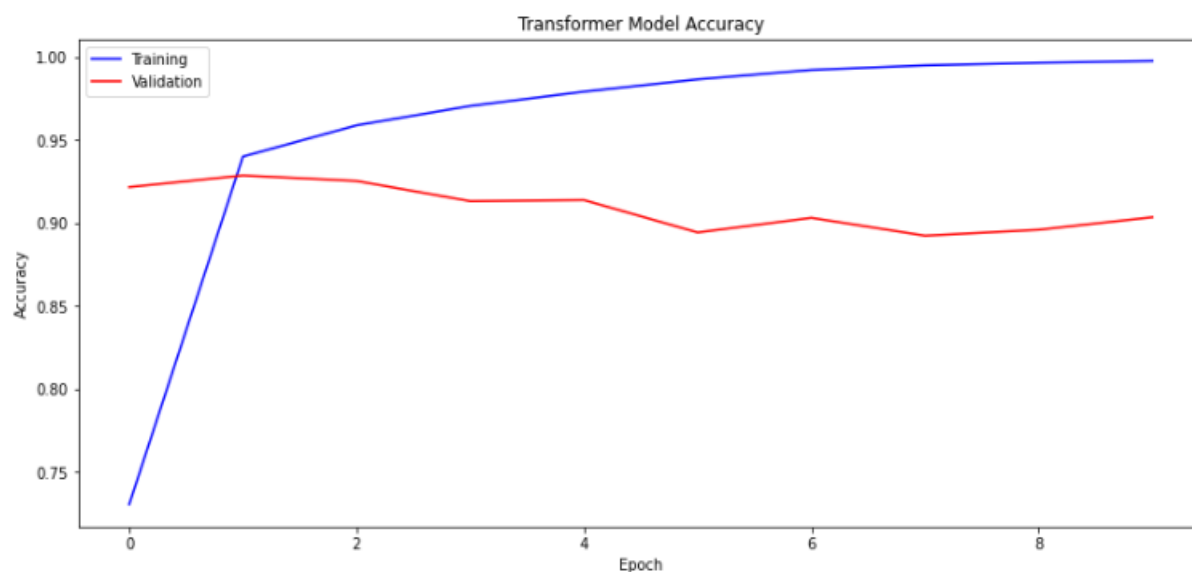
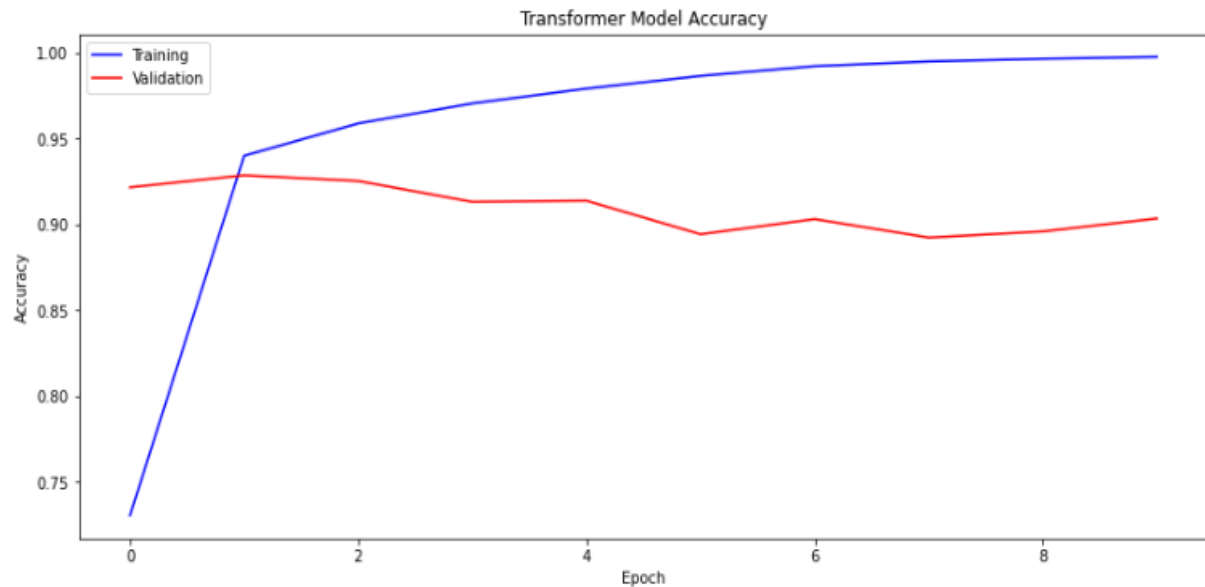


Figure 17

The Transformer Training Loss vs. Validation Loss



The values of training vs. validation accuracy and training vs. validation loss were compared based on the transformers model during execution. The above figures show a significant difference between the validation and training epochs. From Figure 16, the training loss is decreasing while the validation loss is increasing. This suggests that the model is overfitting; therefore, it is very specialized to the training dataset. Figure 15 shows that the training accuracy is slowly increasing while validation accuracy is plateaued. The results indicate that the transformer model is overfitting, revealing that the model is very specialized to the training data, thus, not suitable for generalizing the new datasets.

To address the overfitting problem, regularization and data augmentation approaches were applied. Weight regularization adds cost to the loss function of the model for huge parameter values or weights. L1 regularization adds cost based on the parameters 'absolute value (results in weights being equal to zero), and L2 regularization is based on the squared value of parameters, which leads to smaller weight. The approach forces the model to become more efficient and prevents it from depending much on certain neurons. Data preprocessing was also applied, which involved ensuring that preprocessing steps were appropriate and

consistent during model training. Numerical features were normalized, and necessary feature scaling was performed with the aim of improving the generalization of the model. The accuracy of the model improved but still exhibited the overfitting problem. This makes the model unsuitable for predictions and sentiment classification.

The experiment used 5,341,331 total nodes. These parameters were all trainable and constituted the sum of the input layer, token and position embedding, transformer block, global average pooling, dropout, and dense. The choice of these parameters was informed by the nature of the dataset, previous knowledge problem domain, and the architecture of the transformer model. The position embedding layer encodes the position of tokens in the sequence. The position embeddings were utilized to provide the transformer model with positional information. This helps differentiate several tokens and analyze the sequential order of inputs. The GlobalAveragePooling1D was used to reduce the spatial dimensions of the training set while preserving relevant information. It summarizes the sequence into a fixed length to capture critical details from different tokens and discard irrelevant information. The dropout layer was used to handle the overfitting problem. It serves as a regularization approach to ensure the model does not rely heavily on specific features or relationships. Lastly, the dense layer was used to determine the dimensionality of the output. It utilizes the linear transformation approach based on the activation function to generate predictions. Table 9 summarizes the nodes/parameters used in the transformers model:

Table 9*Nodes in the Transformer Model*

	Layer(type)	Output Shape	Param#
0	input_2 (InputLayer)	[(None, 200)]	0
1	token_and_position_embedding	(None, 200, 32)	5329952
2	transformer_block (Transform)	(None, 200, 32)	10656
3	global_average_pooling1d (Gl	(None, 32)	0
4	dropout_3 (Dropout)	(None, 32)	0
5	dense_5 (Dense)	(None, 20)	660
6	dropout_4 (Dropout)	(None, 20)	0
7	dense_6 (Dense)	(None, 3)	63

4.2.3 *Sentiment Classification*

Sentiment analysis is the type of NLP that utilizes statistical and machine learning approaches to establish and extract subjective information from the text. This analysis aims to identify the attitude, emotions, or opinions expressed in a given text, including news articles, reviews, or tweets. It accurately extracts people's opinions from huge sets of unstructured texts and classifies them into sentiment classes. Users engage in different activities on social media platforms like Twitter to read online news, write reviews or opinions about a certain

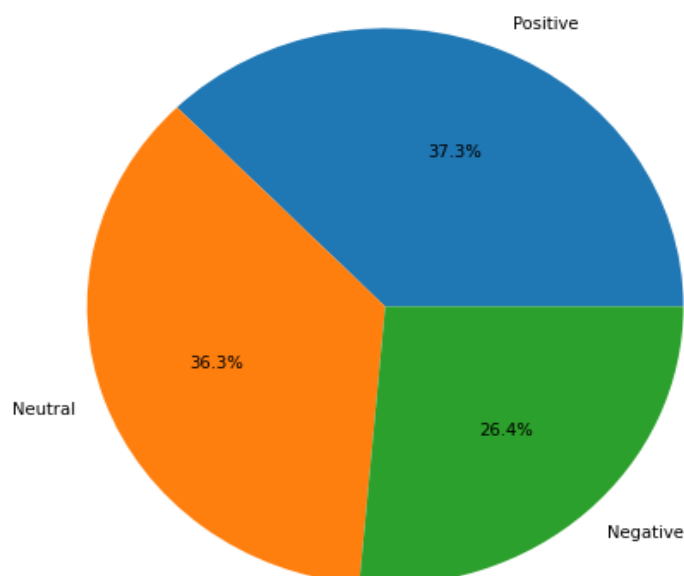
situation, and comment, like and share blogs, videos, and posts. Sentiment analysis algorithms usually analyze text for negative, neutral, or positive sentiments and can also be applied to establish specific emotions, like sadness, anger, or happiness. This analysis is critical to formulating policies following the disaster.

There are several sentiment analysis approaches, including machine learning algorithms, rule-based systems, and lexicon-based approaches. Lexicon-based approaches involve dictionaries of words annotated with sentiment strength and semantic polarity. The approaches attain good results with high accuracy but low recall. Large dictionaries that contain positives and negatives are utilized by lexicon-based models to evaluate polarity views. The pessimistic and optimistic words in the text were evaluated, and if the text consists of more positive words, the positive score will be assigned to the text, and if it contains more negatives, the negative score will be assigned. In this experiment, the VADER model was applied to assign labels to the tweets in our dataset. The VADER model uses a dictionary that comprises sentiments with over 9000 lexical features, such as idiomatic expressions and individual words, and rules for handling punctuations, modifiers, and other related aspects. The model analyses the text and assigns scores to the lexical features according to sentiment polarity (negative, neutral, or positive) and intensity. The general text sentiment is established by aggregating the scores of lexical features.

Machine learning models that receive words from a dataset and analyze it. After analysis, the sentiment intensity is measured, and a few inputs are trained with known outputs with the help of supervised and unsupervised algorithms. These techniques were applied to the Twitter dataset to provide valuable insights into public opinion and attitudes toward the COVID-19 pandemic. These models can learn to identify patterns in the dataset that represents specific sentiments. The experiment was conducted to classify or categorize tweets about COVID-19. The study explored people's sentiments towards COVID-19 measures by

looking at the tweets in order to devise critical policies. In this experiment, the dataset was annotated into three classes: positive, negative, and neutral. The classes were based on social media users' perception of the awareness of guidelines for containing the COVID-19 outbreak. The negative class incorporates tweets associated with denying the outbreak and carelessness, not wearing masks, visiting family and friends, or going out for no reason. The positive class is associated with increased awareness about precautionary measures such as washing hands, staying at home, social distancing, wearing masks, etc. The neutral category includes tweets that exhibited both negative and positive behaviours following the pandemic in the same tweet. The classification of these tweets was based on unique identifiers of keywords.

In Figure 18 below, the sentiment classification was represented with the overall value for each neutral, negative, and positive tweet within the dataset. It can be visualized that the sentiment classes are imbalanced since a significant number of social media users are either neutral or positive against the COVID-19 pandemic and the policies and guidelines associated with it. Very few individuals had extremely negative emotions or perceptions about the coronavirus pandemic.

Figure 18*Sentiment Distribution*

Analysis indicates that 37.3 % of social media users displayed positive emotions toward COVID-19 precautionary measures, 36.3% exhibited neutral emotions, and 26.4% had negative perceptions toward COVID-19 policies. The findings support H1 *‘COVID-19 precautionary policies, measures, and restrictions considerably impact the public sentiments in general and have a noticeably diverse effect on various societal groups*. Individuals from different societal groups were affected differently by the lockdowns and restrictions following the coronavirus pandemic. Individuals high in the economic spectrum displayed more positive sentiments and were less impacted by restrictions. They viewed the precautionary policies to positively impact their lives as they had increased awareness of the measures and their importance. However, minorities and low-income households were negatively impacted and exhibited increased negative sentiments. They perceived the COVID-19-related policies to lead to unequal exposure, transmission, and vulnerability among low-income groups than high-income groups. Restrictive policies, such as lockdowns,

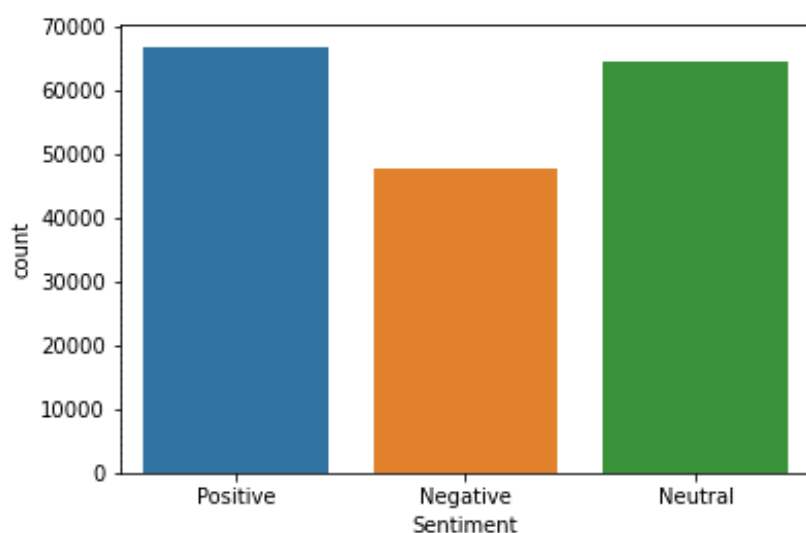
meant the loss of income with immediate effect on minorities and the highly vulnerable in society.

4.2.4 Polarity Detection

Polarity detection is critical in sentiment analysis as it involves establishing whether the sentiment polarity of the text dataset is positive, negative, or neutral. It classifies the tweets to have positive, negative, or neutral sentiments based on the words and phrases utilized in the text. Polarity can be expressed as a numerical rating called the sentiment score. For instance, the score can be a number between -100 and 100, with 0 representing a neutral sentiment. Sentiment scores can be fine-grained based on a specified use case, and in some instances, the categories can expand beyond just negative, neutral, and positive. The polarity shift remains a major issue in the sentiment classification problem since it is capable of reversing the sentiment polarity of the text. The classification performance of machine learning-based systems, rule-based approaches, and lexicon-based systems gets affected by the polarity shift. The figure below shows the distribution of polarities in the dataset.

Figure 19

Sentiment Distribution of Three Class Polarity



4.3 Validation & Evaluation

This subsection will examine and evaluate the obtained outcomes and results based on the standard performance benchmarks that will be explained in detail.

4.3.1 Data Splitting and Data Labelling

Data splitting and data labelling processes were applied during model training and sentiment labels to acquire the relevant polarities. Data splitting helps create the training data and test data because the training data and test data come from the same original dataset. To generate the model, the first step is starting with a single dataset and then splitting the dataset into two datasets: train and test. In this experiment, the models were trained after cleaning the dataset to find their accuracy and efficiency. The columns were separated into dependent and independent labels, and labels were split into train and test sets. The dataset was split into x_{train} and y_{train_1d} sets to fit the model. After splitting the dataset, the models were generated on the training set, and the predictions were performed on test set features using the $y_{pred_mnb} = lr.predict(x_{test})$ function. The evaluation included assessments in train and test percentages. The study utilized the 80% train/validation set and 20% test set configurations to measure the efficiency and evaluate the accuracy of the proposed models.

Data labelling involves identifying raw data, such as images or text files, and adding more meaningful and informative labels to create the context so that the training model can learn from it. The process begins by making a judgment about a specified piece of unlabelled data. The human labels provided are utilized to learn the patterns in model training, and the result can be utilized to make predictions on the cleaned dataset. I added the labels to the unlabelled dataset for sentiment analysis using the Vader Sentiment model, which is a lexicon and rule-based classification model designed for sentiments on social media platforms. The model is suitable for analyzing social media texts since it is designed to handle contextual

and linguistic nuances of such text types. It can handle emoticons and other non-standard features in social media-based data. VADER model is utilized for text sentiment analysis, which is sensitive to intensity (strength) and polarity (positive/negative). The model is dependent on the dictionary that integrates lexical features to emotion intensities referred to as sentiment scores. These scores can be acquired by adding the intensity of the word in the text.

VADER's `SentimentIntensityAnalyzer()` takes a string and returns a dictionary of scores, which are negative, neutral, positive, and compound. The labels –positive, negative, neutral, and compound were assigned, and the dataset was cleaned by removing unused columns. Columns such as `user_name`, `user_location`, `user_description`, `user_created`, `user_followers`, `user_friends`, `user_verified`, `user_favourites`, `date`, `hashtags`, `source`, and `is_retweet` were dropped. The code snippet in Appendix B was used to assign labels, scores, and intensity in the sentiments in the dataset. After cleaning the labelled sentiments in the dataset, the following table was obtained. The subjective information from Twitter was analysed and extracted to categorize text as either positive, negative, or neutral. The sentiment polarity of the cleaned and pre-processed tweets using NLTK-based Sentiment Analyzer revealed sentiment scores for positive, neutral, and negative classes to establish the compound sentiment score for every tweet. Table 10 shows the distribution of positive, negative, and neutral sentiments across the dataset.

Table 10*Distribution of Sentiments across the Dataset*

	Text	Positive	Negative	Neutral	Compound	Neutral
0	if I smelled the scent of hand sanitizers toda...	0.144	0.000	0.856	0.4939	Positive
1	hey Yankeespr and mlb wouldn't it have...	0.101	0.062	0.837	0.2263	Positive
2	wdunlap realdonaldtrump trump never once clai...	0.000	0.130	0.870	-0.2732	Negative
3	brookbanktv the one gift has give me is an ap...	0.267	0.000	0.733	0.7351	Positive
4	july media bulletin on novel coronavirusupda	0.223	0.000	0.777	0.3182	Positive

4.3.2 Accuracy Benchmarking

The selection of models has often been due to key performance issues. In this case, performance was based on the accuracy and precision of the analytical model. Accuracy is considered the most common measure of performance; this is due to its focus on the orientation of the prediction model in relation to the underlying variables. Precision, on the other hand, focuses on the positivity of the prediction score. While these measures are critical in evaluating the model, they are not the only applicable performance measures. Evidence shows the presence of other performance measures, such as the F-score and recall. An examination of the individual classifier algorithms shows varied levels of precision, recall, accuracy, and F-score. The transformer model recorded the highest precision, f-score, accuracy, and recall metrics. LSTM also recorded significantly high performance with regard

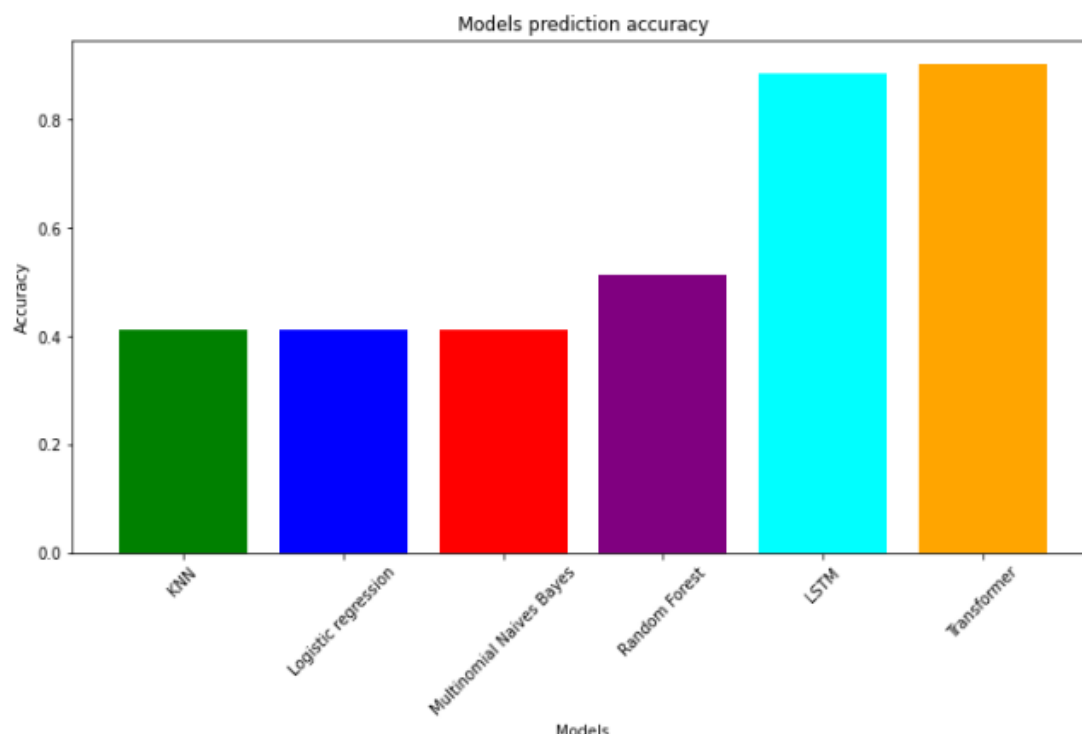
to accuracy and precision, but slightly lower than the Transformers model. The lowest performance was recorded by the KNN model— accuracy of 41.11% and precision of 41.23%. Where LSTM and Transformers are concerned, the models were considered highly accurate and most suitable models for sentiment analysis, with a precision of 88.45% and an accuracy of 88.44%, and 90.18% precision score, and 90.17% accuracy, respectively. The variation in accuracy used by each model is shown in Figure 20.

The figure below shows the model's accuracy, where LSTM was the only good model, with a higher prediction score of over 80%. The other models exhibited lower accuracy scores of below 50%, revealing that they are not efficient in the prediction of Twitter-related sentiments and reactions toward COVID-19-related precautionary measures and policies. These findings support the H2 '*LSTM is the most reliable and accurate approach to both measure and predict public sentiments for decision-makers while managing and administering the spread of COVID-19*'. The model exhibited an accuracy and precision of over 80% and fitted the selected dataset. The LSTM algorithm showed superior predictive precision with fewer errors when measured on the available real data sets. It recorded a significant accuracy score of 88.52%, which makes a good model for predicting the sentiment polarities in the provided dataset. Even though the transformers model (90.17% accuracy) outperformed LSTM in terms of accuracy and precision, LSTM remains a better model for analyzing tweet sentiments since it does not experience the issue of model overfitting and can be generalized to new datasets. On the other hand, the transformers model was found to be too specialized for the dataset and cannot be generalized to new datasets, making it unsuitable for prediction. Transformers can make inaccurate predictions on the Twitter dataset, resulting in poor performance in practice. Thus, LSTM is the most suitable model for evaluating and predicting public reactions and emotions toward precautionary measures and policies implemented following a global outbreak or disaster. This approach and predicted

consequences will help prevent COVID-19 pandemic infections by implementing appropriate measures to address the issue.

Figure 20

Models Prediction Accuracy



4.3.3 Adapting to the Use of Semantic Lexicons

Several solutions have been proposed to enhance word embedding by making use of semantic dictionaries. When training a skip-gram model, enhancing the context of a given word can help produce better embedding (Yu and Dredze, 2014). Each time an update is made, the context of a word is expanded using sampled terms drawn from a thesaurus or association data. Word similarity and relatedness tasks, as well as the TOEFL synonym and document categorization tests, make up the extrinsic evaluation. The word vectors perform better on all tasks after applying the augmentation approach. Several solutions have been proposed to enhance word embeddings by making use of semantic dictionaries. Multiple approaches are proposed by Yu and Dredze (2014) that use the CBOW architecture in tandem with a second objective function that seeks to optimize the relations present in some semantic lexicon. The models are put through their paces on language modeling and semantic

similarity tasks using the Paraphrase Database and WordNet as data sources. They state that using their technique results in better performance in both activities.

4.3.4 Supervised training

In most cases, a supervised classifier is used in tandem with pre-trained word embedding to perform sentiment analysis. The word embedding approach is used here as a feature extractor for the classification framework. Using a memory state to pick up on long-distance dependencies, recurrent neural networks like the Long Short-Term Memory network and the Gated Recurrent Units are a form of feed-forward network. Using GloVe vectors in conjunction with a recurrent neural network and training on the Stanford Sentiment Treebank is one method that has shown promise for text categorization tasks. They both train and test on these annotated phrases because the dataset includes sentiment annotations at each node of a parse tree. The use of Convolutional Neural Networks (CNN) for text classification has shown promising results. On seven datasets, including the Stanford Sentiment Treebank, skip-gram vectors have been employed as input to various Convolutional Neural Networks. A single-layer CNN that refines the initial skip-gram vectors during training performs well across datasets (Elfaik, 2021). The study by Barnes et al. (2017) shows that these methods now achieve state-of-the-art outcomes on numerous datasets, although they have not been compared to retrofitting or combined training methods.

4.3.5 Model Evaluation

The model focused on a specific dataset (data from Twitter) containing a wide range of sentiment-based information. An analysis of these sentiments shows the presence of numerous words that reflected different human emotions and attitudes towards government policies and the pandemic. While diversity was preferred and acceptable, it is critical to acknowledge the underlying classification challenges associated with the model. As a result, it was critical to classify the dataset into three major sentiments: positive, neutral, and

negative. The classification was accompanied by separating the dataset into training and validation. The retaining dataset accounted for 80% of the dataset, while validation accounted for 20%. The distribution of the dataset made it necessary to apply different classification models for further analysis. The introduction of machine learning has seen the development and use of various classification models. These models have unique strengths that allow them to be used for different analytical solutions. In this case, four machine learning models were adopted for the analysis of sentiment-based data; these models included (1) k-nearest neighbor, (2) random forest, (3) logistic regression, and (4) Multinomial Naïve Bayes. Each method had its unique advantages to the prediction process, creating varied results. This variation was primarily seen after evaluating the different models and comparing their performance metrics.

The model was evaluated based on accuracy, precision, recall, and f1-score metrics. Accuracy is a metric for evaluating model performance to classify tasks and accounts for all offline and online model performance. It measures the percentage of correctly classified instances from all instances. The metric is computed as the ratio of the number of correct predictions to the total number of performed predictions. Precision evaluates the percentage of the correctly classified positives based on all instances that were predicted as positive. It is the measure of the number of positive instances predicted by the model that is typically positive. Precision is computed as the ratio of the number of true positives (TP) to the total number of false positives (FP) and true positives. Recall is the percentage of true positive instances out of the total actual positive instances. It is a measure of the number of actual positive instances correctly established by the model. It is computed as the ratio of the number of TP to the false negatives (FN) and TP. Finally, F1-score is the weighted harmonic mean of recall and precision. It balances between recall and precision, and it is usually

appropriate when class distribution is imbalanced. It is computed as $2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$.

These metrics predicted and compared the performance and strength of the classification models used in the experiment. A high accuracy score suggests an overall good performance, whereas recall and precision scores imply that the model is suitable for the correct identification of positive instances. A high F1 score indicates a standard balance between recall and precision. The table on the next page presents the accuracy, precision, recall, and F1-score values utilized to assess the performance of all models. The accuracy of the LSTM model is 0.8844 (88.44%), which is significant. The precision score of the model is 0.8845 (88.45%), the recall value is 0.8844 (88.44%), and the F-score is 0.8842 (88.42%). For the Random Forest model, the precision score is 0.5243 (52.43%), the recall value is 0.5126 (51.26%), the F-score value is 0.4912 (49.12%), and the accuracy is 0.5126 (51.26%). For Logistic regression, the precision score is 0.4354 (43.54%), the recall value is 0.4123 (41.23%), the F-score value is 0.3478 (34.78%), and the accuracy is 0.4123 (41.23%). In the KNN model, the precision score is 0.4123 (41.23%), the recall value is 0.4111 (41.11%), the F-score value is 0.4091 (40.91%), and the accuracy is 0.4111 (41.11%). Lastly, the transformer model recorded a precision score of 0.901874 (90.19%), a recall value of 0.901742 (90.17%), an F-score of 0.901764 (90.18%), and an accuracy of 0.901742 (90.18%).

Table 11

Models Performance

	Model	Precision	Recall	F-score	Accuracy
0	KNN	0.412310	0.411121	0.409104	0.411121
1	Logistic regression	0.435435	0.412349	0.347821	0.412349

2	Multinomial Naive Bayes	0.435435	0.412349	0.347821	0.412349
3	Random Forest	0.526618	0.513176	0.494130	0.513176
4	LSTM	0.885083	0.885217	0.885084	0.885217
5	Transformers	0.901874	0.901742	0.901764	0.901742

The Transformers and LSTM performed better than other models with comparatively high precision, recall, F1-score, and accuracy scores. Even though both models were efficient, the Transformers model was superior to LSTM as it exhibited an accuracy of 90.17%, exceeding LSTM with 1.65%. Good accuracy in machine learning models is subjective; however, prediction above 70% is considered great and effective model performance. Simply put, the accuracy measure of predictions between 70%-90% is not only ideal but also realistic. In the current study, the code snippet below was written to plot the accuracy of the different models used in the experiment. Regardless of the performance, the transformers model was coupled with several issues, including data overfitting and overspecializing, making it unreliable for the current dataset. Therefore, LSTM remains the superior prediction model approach for sentiment classification expressed by the selected disaster-related dataset. At the end of the experiment, the model recorded 88.52% accuracy and fitted the model, which is good since it enables appropriate and accurate classification and generalization to the new dataset. As a result, it was utilized to analyze the sentiments of the provided dataset. The KNN model performed the least, with an accuracy of 41.11% and a precision score of 41.23%. The prediction accuracy of the models might have been influenced by the quality and quantity of the dataset that they were trained on, the context of the questions, and the complexity of the task at hand.

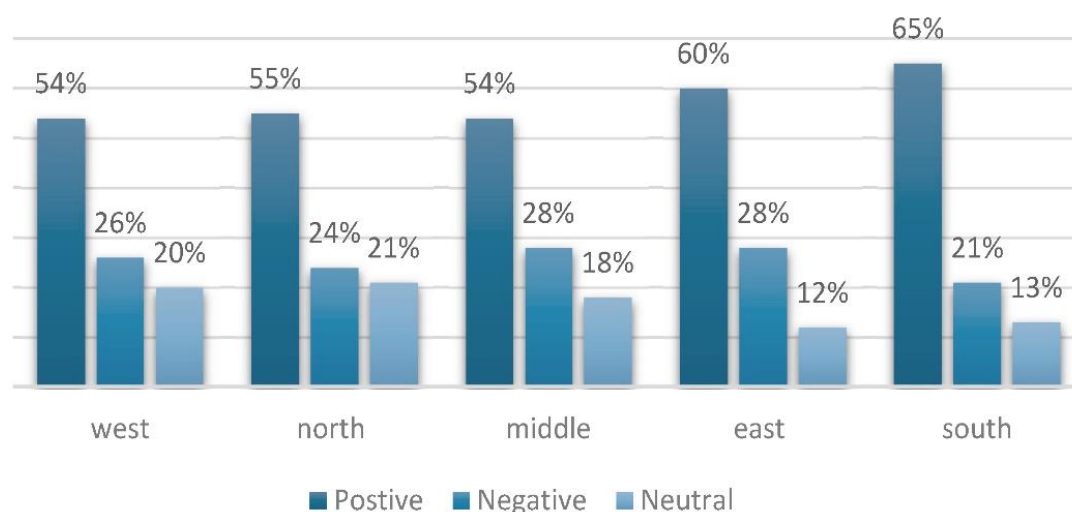
The variation in performance indicates the dynamism of the model across different fronts; this created the need to examine the models using actual data. As a result, tweets relating to global data were used as part of the preliminary testing. In the case of tweets from across the globe, the positive and negative sentiments were used as indications to show an increase or decline in awareness of precautionary policy measures imposed to deal with the COVID-19 pandemic. The results revealed that the southern region had the highest level of awareness, with more than 60% of tweets being classified as positive sentiments. The model and the western region had the lowest percentage of positive sentiments, about 55%. This indicates that such regions have a high level of awareness, which increases with time. Therefore, it is critical to acknowledge that policy acceptance and the nature of responses heavily rely on awareness and understanding. This implies that the number of sentiments associated with a given policy tends to increase as more people become aware of and understand its dynamics. This explains the variation in data during the early phases of lockdown in relation to the later stages of lockdown.

While focusing on the positive aspect of human sentiment, it is critical to acknowledge the diversity in the Twitter population. This diversity was best shown by varying attitudes and opinions towards the set precautionary policies. As a result, the negative sentiments were characterized by carelessness toward the provided measures. Each region had a varied number of negative tweets, showing a poor attitude towards various government actions. The eastern region recorded the highest number of negative sentiments (accounting for 28%) of the total tweets obtained from that region, while the southern region recorded the lowest proportion of negative sentiments. The diversity in the population was further highlighted by the presence of neutral tweets. These were individual sentiments that had both positive and negative elements. Such sentiments were predominant in the northern region, where they

accounted for more than 20% of the total tweets. The variation in the classification of tweets as positive, negative, or neutral sentiments is highlighted in Figure 21.

Figure 21

Tweet Classification



4.3.6 Users' Reactions During the Pandemic

The variation in individual perception is common in different events. According to Aljameel et al. (2021), the observed differences stem from the overall impacts of an event on their activities or lifestyle. During the pandemic, there was a notable change in lifestyle, ranging from social distancing to past-time activities. The cessation of social events and reduction of working hours was also a critical part of the new policies implemented during the pandemic. Amidst these regulations, social media platforms such as Twitter, Instagram, and Facebook have become commonplace for interaction and engaging in different virtual activities. These were accompanied by sharing of different opinions regarding the changes and implemented policies. Such opinions could be classified as different types of sentiments to determine the overall perception or attitude of individuals toward safety regulations or policies. The initial model focused on classifying sentiments using different techniques, with the main focus on the dynamics of the techniques. The adopted techniques allowed a

generalized approach to the classification of data, resulting in the need to focus on the general trend. While these elements were essential, they did not acknowledge the time series variation in the data. Specifically, individual opinions tend to vary from region to region and from time to time. As a result, it was critical to examine the individual sentiments across different periods during the pandemic, especially during the lockdown.

The shift in focus shows that the provided sentiments could be easily classified into three major categories: negative, neutral, and positive. The period between January and June 2020 was mainly characterized by variations in individual sentiments from various social media sites or platforms. The information provided ranged from very negative to positive. In January, the number of negative sentiments on different safety policies was significantly low compared to other periods. These observations were attributed to a lack of restrictive policy, with a high level of normalcy— individuals were not forced into a lifestyle change due to the pandemic. A significant increase was observed in the subsequent month, with higher negative sentiment and extremely low positive sentiments. This classification has been characterized by accompanied by the gradual emergence of more restrictive policies, creating significant challenges during the pandemic. In the months of March, April, and June, the highest number of negative sentiments were observed. Despite such observations, it is critical to acknowledge that the majority of social media users acknowledge the positives and the negatives of the policies implemented during the pandemic. Therefore, neutral sentiments were common during the majority of the period.

The recognition of human sentiment towards different policies is central to any data modeling approach. In this case, it was critical to acknowledge that the variation in sentiment was not one-dimensional. Instead, individuals across cyberspace tend to engage freely in different topics as a way of sharing ideas and communicating their opinions. This created the need to explore the provided dataset based on the topics. In this case, a topic modeling

function was applied to highlight the variation in topics and the underlying sentiments. The analysis shows that individuals engaged in specific topics in topic across different periods. The variation in spikes was mainly due to differences in data collected during the different periods. Nonetheless, the corresponding topics ranged from health issues to economic and social issues. The number of topic spikes per sentiment is best highlighted in the next table.

Table 12

Spikes per Sentiment

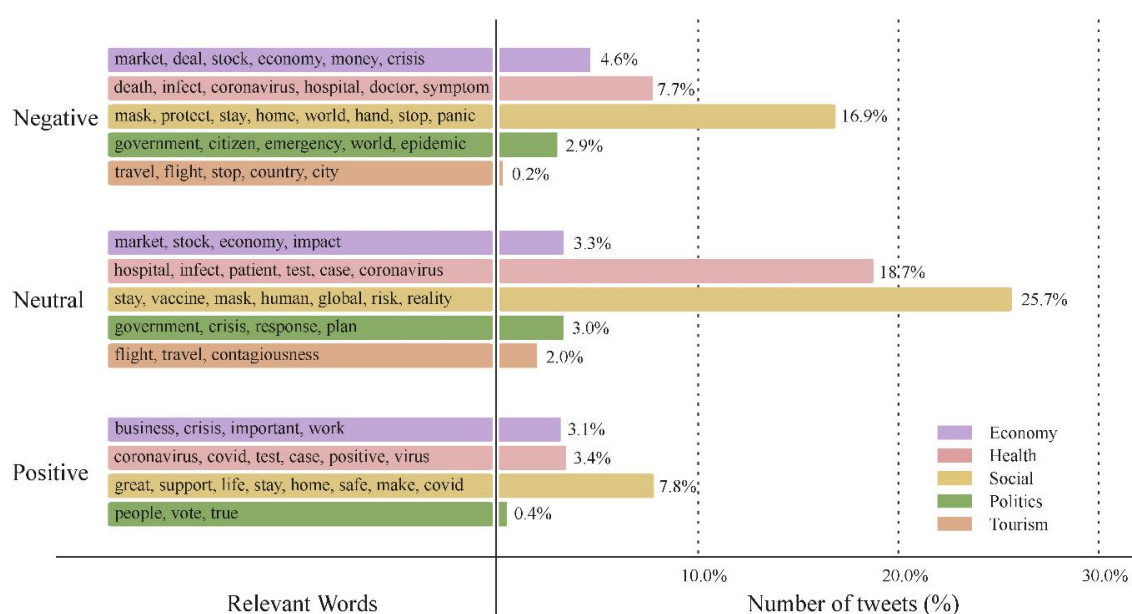
General Topics	Negative	Neutral	Positive
Economy	1	1	2
Health	4	2	1
Social	5	3	6
Politics	1	1	1
Tourism	1	1	0
Fashion	0	0	0
Topics Count	12	8	10
Content Score	0.4754	0.4132	0.4533

As the number of topic species can be related to topics in Table 10 above, they do not fully account for other key elements within the dataset. An examination of a data modeling component shows that the provided tweets had specific keywords there were observed in the different number of tweets. For instance, the number of the negative was highest when

dealing with social issues; in this case, the main keywords associated with the social issues included stay at home, panic, mask, protect, and hand— they accounted for more than 17% of the tweets. Similarly, the same obviation was observed where neutral and positive sentiments were concerned. Specifically, relevant words associated with healthcare sentiments included vaccine, masks, risk, reality, and human; these accounted for about 26% of the total number of tweets. Where positive sentiments are concerned, the keywords included support, safety, make, and grate, which accounted for approximately 8% of the tweets. Full variations of sentiments in relation to keywords and topics are shown in the next figure.

Figure 22

Variation of Sentiments in Relation to Keywords



The variation in the general trend was not the only factor that needed to be accounted for; evidence shows that social media platforms such as Twitter & Instagram have an international presence. This presence is often accompanied by differences in lifestyle and understanding of local and global policies. As a result, it was critical to integrate a geo-temporal element into the analysis. This aspect of the analysis was best achieved by identifying country mentions on the tweet and other temporal patterns within data collection.

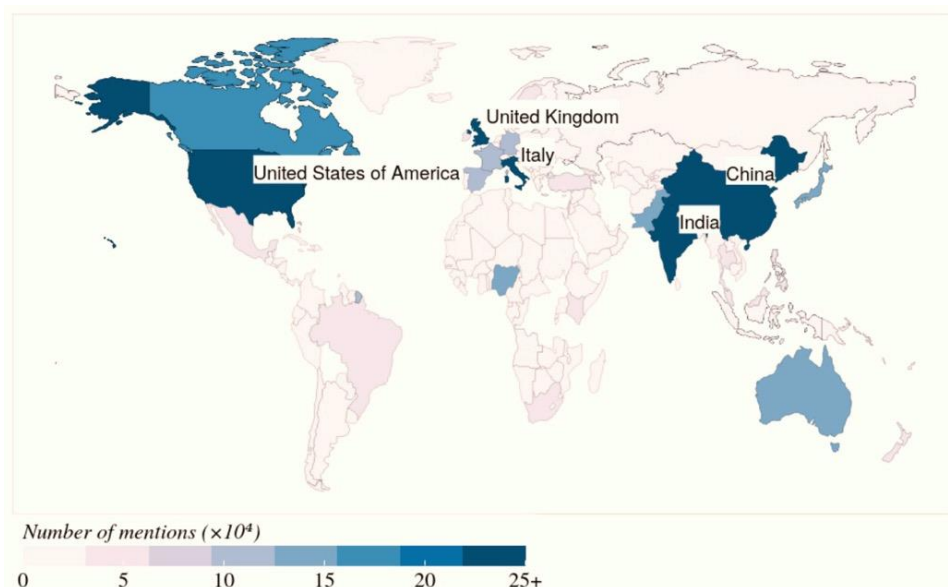
This aspect allowed the model to identify the various countries and what was tackled in relation to pandemic posts as people attempted to cope with a change in their lifestyle.

Therefore, a NER (Named Entry Recognition) tagger was incarnated as part of the language processing capabilities to determine country mentions, places, and dates during the testing and validation processes.

The inclusion of the NER technique allowed the model to identify country names, cities, and other accesses provided in the input. Despite this capability, the main focus was the name of the cities. Based on the analysis, the tweets provided mentioned more than 180 countries. China was the country with the most mentions, accounting for more than 40% of the total tweets. Other countries such as the United States, India, and the UK had mentions of more than 500,000, as can be seen in the figure next page. Given the presence of such spatial data, it was critical to acknowledge variation in sentiments from the provided tweets.

Figure 23

Spatial Distribution



The sentiment analysis shows that the number of types of sentiments varied significantly across countries. The variation was primarily observed by acknowledging Italy, China, the United States, and India as the top five countries discussing different topics associated with

virus. There is no clear indication of negative or positive emotions from the word cloud as people are largely focused on the origin of the virus and the virus itself. In Italy, the popular keywords related to the pandemic were case, quarantine, death, and people. Overall, the topics are negative, which denotes negative emotions among the public.

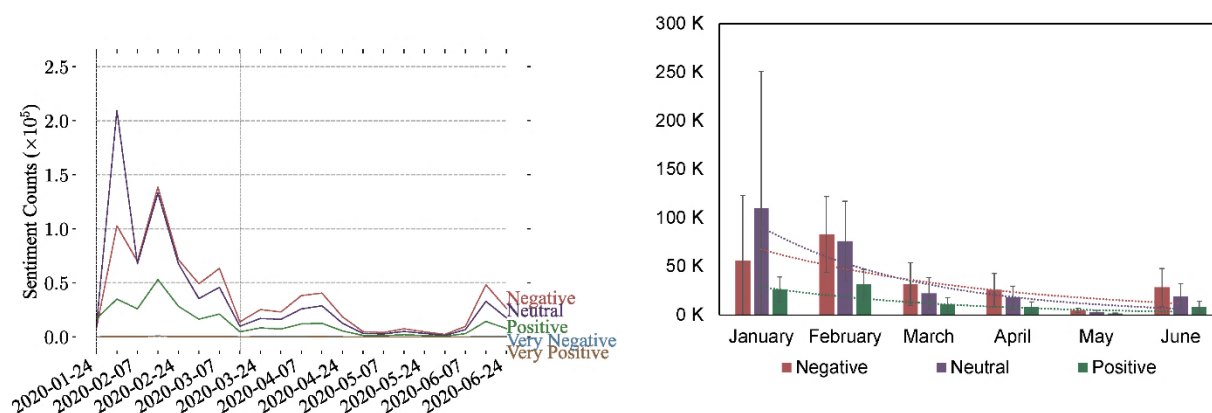
In India, the most common words were coronavirus, lockdown, people, infect, hospital, and vaccine. The dataset used in this study covered the periods between November 2020 and January 2021, and this was at a time when infections were soaring in India, and the country was imposing lockdowns in an effort to combat the spread of the virus. It was at this time as well that the first vaccines were available. In China, the most common words included Wuhan, China Virus, world, and Wuhan Virus. These words reflect a period when there was a lot of talk about the origin of the virus, and then former US President Trump often referred to COVID-19 as the “China Virus,” which originated in Wuhan province in China. In the US, the most common words were death, infection, fatality, coronavirus, hospitalized, and reported, and these words reflect a time when the number of deaths, infections, and hospitalizations in the US was soaring between 13th November 2020 and 1st January 2021. In Italy, within the period of the data (13th November 2020 to 1st January 2021), a famous journalist tested positive for COVID-19, and that is why the most common words include Fabrizio Romano, positive, and quarantine.

Previous research suggested that in China, January and February 2020 were primarily characterized by neutral sentiments, with positive sentiments being the lowest in the subsequent months; the number of negative sentiments significantly decreased as individuals had a negative perception of the policy measures put in place (Chen et al., 2020). Nonetheless, the subsequent months recorded a decline in social media interactions relating to the COVID-19 pandemic and respective government actions. A notable spike in sentiments was recognized at the end of June 2020— a positive-to-negative ratio of about 0.4 was

recorded during this period. The variation in sentiments across the six months period is shown in the figure below.

Figure 25

Variation of Sentiments Across a 6-Month Period in China (Chen et al., 2020)

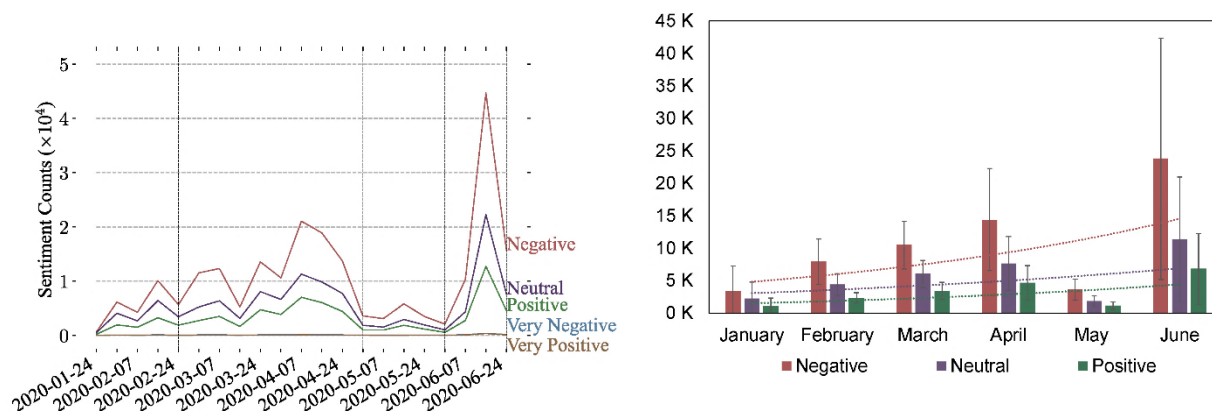


The observation in the United States, as Shekerdemian et al. (2020) explained, was notably different from those observed in China. Unlike China, the United States experienced a lower number of opinionated twitters during the initial months— less than 5000 tweets were attributed to varying sentiments in January 2020. Despite the low number of tweets to be accounted for, the main sentiments were predominately negative. Throughout the designated period, the total negative sentiments relating to pandemic safety measures fluctuated with the number of tweets. This fluctuation was primarily notable between April and May, as well as May and June. April recorded a significant rise in the number of positive sentiments, followed by a sharp decline during the month of May. A similar trend was observed in June, where the highest negative sentiments were recorded. An analysis of the variations shows their alignment with numerous economic and social policies; for instance, the extension of economic lockdown and social distancing-related policies created further negative perceptions across different regions of the country. In the process, the recorded negative sentiments were more than double the number of positive sentiments recorded in

June— a positive-to-negative ratio of about 0.2. This variation is best highlighted in the figure next page.

Figure 26

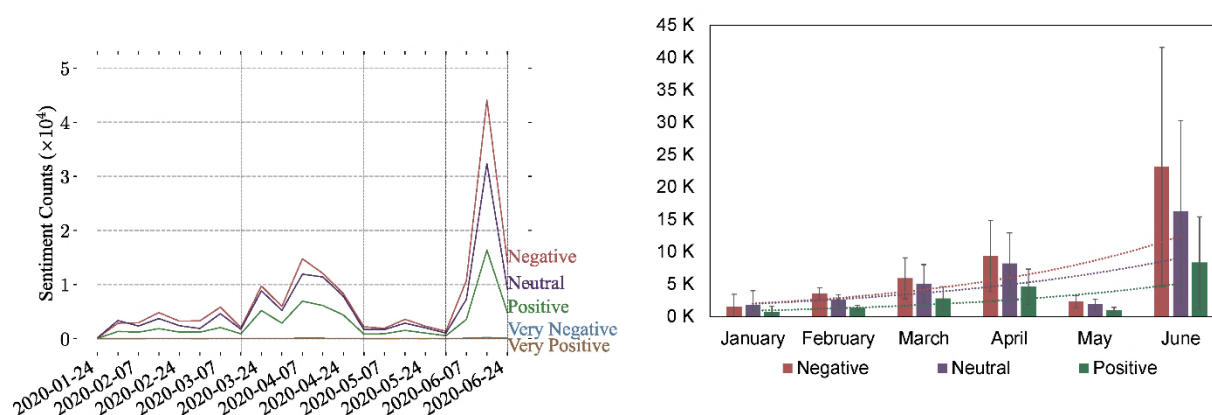
Variation of Sentiments across a 6-Month Period in the US (Shekerdemian et al., 2020)



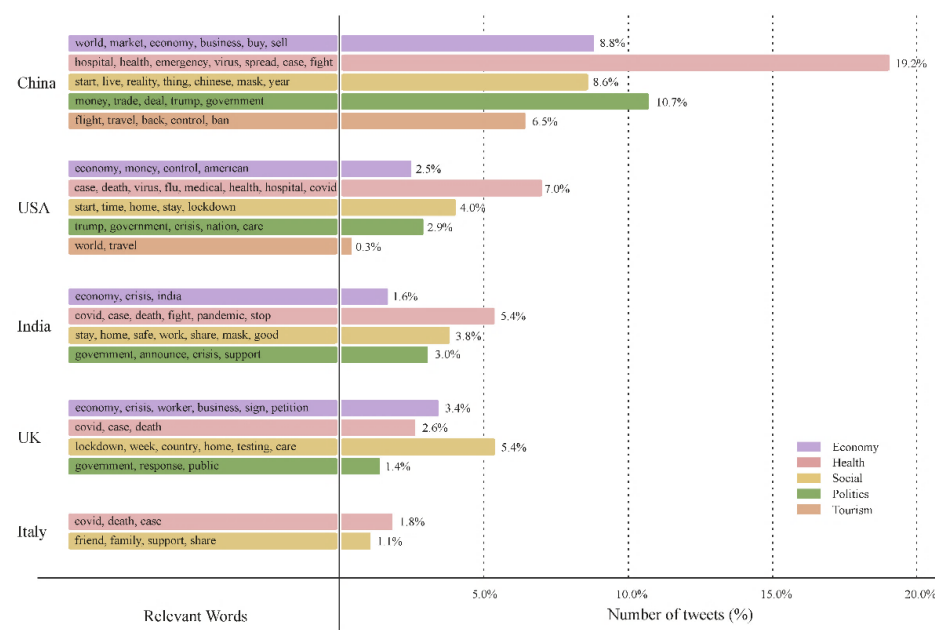
Where India is concerned, the trend is similar to that of the United States. The similarity is primarily observed where the lowest number of sentiment-based tweets are recorded during the initial months, with a significant fluctuation in the subsequent months. The spike in observation was also observed during the final period of June 2020 (Saigal et al., 2020). Despite the observed similarities, there was a notable difference in the dataset. These sentiments were mainly accompanied by the minimal fluctuation occurring between January and March 2020; during that period, each month had less than 5,000 sentiment base tweets, with the majority of sentiment being negative, except in January, where neutral and negative sentiment had the same value. Nonetheless, negative sentiments were prominent in the region. The prominence was primarily seen in June, where a maximum of 5,000 positive sentiments were recorded compared to 20,000 recorded negative sentiments (a positive-to-negative ratio of 0.25). These observations have been summarized in the figure below.

Figure 27

Variation of Sentiments across a 6-Month Period in India (Saigal et al., 2020)



While focusing on overall opinion, it is critical to acknowledge that the sentiments vary across different topics. This variation was observed across different countries. In the case of Italy, most tweets were related to health and social issues. The main keywords associated with a health-related tweet in Italy included death and COVID-19, while those relating to social issues were focused on support, family, and friendship— these accounted for 3% of the tweets. China had the most diverse tweet topics; they had tweets relating to each of the topic domains shown in Table 11. The analysis showed that keywords relating to political issues included government, trump, trade, and money. In the case of social issues, key elements included masks, reality, things, and start. The most dominant discussion was on health issues which constipated about 19% of the total tweets relating to China. The variation in sentiment-related keywords across different countries is shown in the figure next page.

Figure 28*Sentiment-Related Keywords across Different Countries***4.3.7 Techniques for Precise Outcomes**

The observed approach to sentiment analyses has focused on highlighting the polarity of human interaction in relation to external factors. The adopted model has been efficient in providing different outcomes due to its capacity to classify tweet-based data based on the underlying opinion. This approach was accompanied by the integration of various elements, such as natural language processing data mining and information retrieval techniques, to classify subjective data. While these elements have been designed to ensure a high level of accuracy, it is critical to acknowledge that prediction is reliant on precision. This acknowledgment creates the need to further explore the different dimensions of the sentiment analysis model to ensure improved precision across different datasets. Precision in a given model can be enhanced by the adoption of word embedding. In this case, the adopted dataset lacked contextual relationships. This created difficulty in specifying and obtaining relationships between different text-based variables. As a result, it was vital to advance word embedding modern that would define contextual relationships within datasets. This

integration allowed the model to understand different words and recognize suffices and prefixes during the training processes. In the process, the n-grams model was used to create boundary symbols for the case analysis. In the process, data will be classified and ready for sentiment analysis.

The adoption of n-grams was further enhanced by the use of a long short-term memory approach to analysis. The LSTM model allows the representation of words through corresponding features based on the embedded model. The features within the LSTM approach are usually represented in a network sequence. This attribute of LSTM aids the resulting model with the learning capability of the system, especially where a long-term dependency between input features is concerned. These advantages also focus on eliminating the singularity in analysis through the use of neural networks that all activities within a special domain. Despite the observed advantages, the company's approaches are designed to yield an accuracy of at least 80%; there is potential to further improve the analytical framework.

An analysis of the LSTM, especially the support vector machines, shows that accuracy can be improved through the use of N-gram models. Despite the high level of accuracy, the focus is on precision, which is central to any predictive analytic model. For this reason, a hybrid deep learning model is considered efficient for the analysis. The model integrates natural language processes, word embedding mechanisms, and systematic semantic labels to obtain connatural information. This is further enhanced by the introduction of multiple machine learning techniques, such as KNN, for a more detailed classification. Where prediction is concerned, the logistic regression technique will be included to enhance the overall prediction process. This integration accounts for the polarity of the datasets and variation in the relationship between context data and the emoticon used. Where optimization

is needed, a more dynamic boost algorithm will be adopted for a more efficient analytical system.

4.3.8 Approaches Used in Benchmarking

To better evaluate the developed models and generated results, state-of-the-art performance evaluation approaches were utilized, as will be explained in thorough detail in this subsection.

4.3.8.1 Binary and Multi-Class Benchmarking. Binary classification benchmarks are used to classify text into two categories, while multi-class classification benchmarks are used to classify text into more than two categories. Binary Classification Benchmarks in Sentiment Analysis refer to the process of classifying a piece of text as either positive or negative. This process is often used in sentiment analysis, which is the process of identifying the sentiment or emotion behind a piece of text. Sentiment analysis can be applied to a wide range of applications, including social media monitoring, product reviews, and customer feedback analysis. In order to evaluate the performance of sentiment analysis algorithms, researchers often use binary classification benchmarks.

These benchmarks involve training a machine learning algorithm on a dataset of labeled positive and negative text samples and then testing its accuracy on a separate set of labeled samples. The accuracy of the model can then be evaluated using metrics such as precision, recall, and F1 score. The use of binary classification benchmarks allows researchers to compare the performance of different sentiment analysis algorithms and to identify areas for improvement. The choice of the dataset for binary classification benchmarks can have a significant impact on the performance of sentiment analysis algorithms. Therefore, it is important to carefully select a representative dataset to ensure an accurate evaluation of the algorithm. Overall, binary classification benchmarks are a valuable tool for evaluating the

performance of sentiment analysis algorithms and improving their accuracy in real-world applications.

In order to evaluate the effectiveness of algorithms, researchers have developed multi-class classification benchmarks that measure the accuracy of classifiers across multiple categories of sentiment. The use of multi-class classification benchmarks is essential for evaluating the performance of sentiment analysis algorithms because it allows researchers to measure the accuracy of these algorithms across a range of different sentiment categories. One such benchmark is the Stanford Sentiment Treebank, which contains over 11,000 sentences that have been manually annotated with fine-grained sentiment labels. Another benchmark is the SemEval 2014 Task 4, which focuses on sentiment analysis of tweets and includes four distinct subtasks.

These benchmarks are important because they provide a standardized way of evaluating the performance of sentiment analysis algorithms, which can help to improve the accuracy and reliability of these algorithms over time. Additionally, the development of these benchmarks has led to increased interest in sentiment analysis research, as evidenced by the growing number of publications and conferences dedicated to this field. As researchers continue to develop more sophisticated algorithms and benchmarks, the field of sentiment analysis is likely to become even more important for a wide range of applications, including marketing, politics, and customer service.

One way to analyze public sentiment towards natural disasters is through the application of benchmarks. Benchmarks are a set of standardized measurements that allow for the comparison of different data sets. Benchmarks can be used to evaluate the effectiveness of disaster management strategies, as well as to assess the impact of natural disasters on communities. By comparing public sentiment towards natural disasters to established benchmarks, researchers can gain a better understanding of how individuals

perceive these events and what factors influence their reactions. For example, benchmarks can be used to analyze how different demographic groups respond to natural disasters, such as age, gender, and income. This information can then be used to develop targeted disaster preparedness campaigns that are tailored to specific groups. Overall, the application of benchmarks in analyzing public sentiment towards natural disasters is a valuable tool for understanding the impact of these events on communities and developing effective disaster management strategies.

Both binary classification benchmarks and multi-class classification benchmarks are crucial in sentiment analysis for analyzing and predicting public sentiments and reactions toward governments' precautionary measures and policies during natural disasters. The use of these benchmarks enables researchers to classify sentiments and reactions with a high degree of accuracy, which is essential for policy-makers to create effective policies and measures that address the public's concerns during natural disasters, in this case. Therefore, the research on sentiment analysis in social media using binary and multi-class classification benchmarks is essential for disaster management and response.

4.3.8.2 Aspect-Based, Fine-Grained, and Irony Detection Benchmarks. Sentiment analysis has become an essential tool for analyzing public reactions and sentiments toward specific topics. One such topic of interest is the public's response to governments' precautionary measures and policies during natural disasters. In recent years, the field of sentiment analysis has expanded to include aspect-based, fine-grained, and irony detection benchmarks, which have proven to be effective in studying public sentiments in disaster response. This section examined the use and importance of aspect-based sentiment analysis in disaster response, the effectiveness of fine-grained sentiment analysis as a tool for disaster

management, and the role of detecting irony in public sentiment on government policies during disasters.

Disaster response management involves collecting and analyzing vast amounts of data to make informed decisions. One critical aspect of disaster response management is understanding the sentiment of the affected individuals and communities. Aspect-based sentiment analysis (ABSA) is a powerful tool for disaster response management that enables the extraction of sentiment information from text data, providing a deeper understanding of public sentiment towards specific aspects of the disaster. ABSA can be used to monitor public sentiment toward relief efforts, track the effectiveness of response efforts, and identify areas where more support is needed. This method can also be used to analyze social media data and identify public sentiment towards various aspects of disaster response, such as government response, relief aid, and communication. By using ABSA, disaster response managers can make informed decisions, prioritize response efforts, and allocate resources more effectively.

The fine-grained benchmarking method has emerged as a promising tool. This method can be used to identify and classify tweets based on the emotions and sentiments expressed. This can help emergency responders quickly identify critical information, such as requests for help or reports of injuries, and respond accordingly. Additionally, sentiment analysis can also be used to identify areas of high stress or anxiety, which can help emergency responders prioritize their efforts. Overall, fine-grained sentiment analysis has the potential to improve the efficiency and effectiveness of disaster management efforts by enabling emergency responders to quickly identify and prioritize critical information.

During disasters, governments are expected to provide efficient and effective support to the affected population. However, public sentiment towards government policies during such times can be complex and reveal a range of emotions, including irony. Irony can be defined

as an expression of meaning using language that normally signifies the opposite, typically for humorous or emphatic effect. Detecting irony in public sentiment on government policies during disasters is crucial as it can provide insight into the level of dissatisfaction and frustration among the affected population. This can help government officials and policymakers to understand the areas where they need to improve their policies and communication strategies. Irony and sarcasm can be particularly potent sources of political criticism because they not only express disapproval but also imply a degree of sophistication and knowledge on the part of the speaker or writer. Therefore, it is important to pay attention to the use of irony in public sentiment on government policies during disasters, as it can reveal the underlying emotions and opinions of the affected population and help governments to improve their disaster management policies and practices.

Aspect-based sentiment analysis involves analyzing the sentiment of specific aspects or features of a product, service, or event. In the context of analyzing and predicting the public's sentiments and reactions towards government precautionary measures and policies during natural disasters, aspect-based sentiment analysis could involve identifying specific aspects of the government's response, such as communication strategies, evacuation procedures, and aid distribution. Benchmark datasets for aspect-based sentiment analysis include SemEval-2014 Task 4 and SemEval-2015 Task 12. On the other hand, fine-grained sentiment analysis involves analyzing the sentiment at a more granular level, such as identifying whether a sentiment is strongly positive or mildly positive. Finally, irony detection involves identifying and classifying ironic statements in the text. Incorporating these benchmarks into machine learning models can improve their ability to accurately predict public sentiment towards government policies during natural disasters. Machine learning models can provide more nuanced insights into public opinion by analyzing the sentiment of specific aspects and identifying the strength of the sentiment expressed. Additionally, incorporating irony

detection can help ensure that the sentiment analysis model is not fooled by sarcastic or ironic statements that express the opposite sentiment. Overall, by using these benchmarks, machine learning models can more effectively analyze and predict public sentiment towards government policies during natural disasters.

In conclusion, aspect-based, fine-grained, and irony detection benchmarks in sentiment analysis are important tools in studying and predicting public sentiments and reactions toward government precautionary measures and policies during natural disasters. Machine learning techniques play a crucial role in processing vast amounts of data and providing valuable insights into people's emotions and behaviors. These benchmarks can help policymakers make informed decisions and take appropriate actions to ensure the safety and well-being of citizens during natural disasters. As such, it is essential to continue exploring and developing these benchmarks to enhance the accuracy and effectiveness of sentiment analysis in predicting public reactions toward government policies and measures during natural disasters.

4.3.8.3 Domain, Time Series, and Emotion Detection Benchmarks. Domain benchmark refers to using a specific set of words or phrases commonly used in a particular domain or industry. In the context of COVID-19 policies, domain-specific words such as "lockdown," "vaccine," "social distancing," and "mask mandate" can be used to identify tweets related to COVID-19 policies. By using domain-specific benchmarks, sentiment analysis can be more accurate and relevant to the topic being analyzed.

Time series benchmark involves analyzing sentiment over time. This can be useful in understanding how public opinion toward COVID-19 policies has changed over time. For example, sentiment analysis can be performed on tweets from the beginning of the pandemic to the present day to see how public opinion has evolved. The emotion detection benchmark, on the other hand, involves identifying the emotions expressed in a tweet. This can be useful in understanding the specific emotions associated with COVID-19 policies. For example,

tweets expressing fear or anger towards COVID-19 policies may indicate a lack of trust in the government's response to the pandemic.

4.4 Chapter Summary

The impacts of the COVID-19 pandemic cannot be overstated. It influenced numerous aspects of human life, including the overall social structure and approach to the economy. Amidst these effects, there were a number of tweets that highlighted the impacts of the pandemic. At the same time, the tweets also informed the general public of different opinions or perceptions of the safety policies put in place across different regions. As such, this study used Twitter as the primary source of data. The Twitter dataset was related to COVID-19 infection and COVID-19-related tweets across the globe. Predefined keywords, emojis, and hashtags were employed in the data collection process using the Twitter API. The data was collected between 13th November 2020 and 1st January 2021. The study utilized the Keras machine learning platform in the experiment, which used Tensorflow for deep learning approach implementation as the backend. Since the dataset involved was large, it was trained in Kaggle. Data preprocessing was applied to clean raw data and was based on three major steps: tokenization, normalization, and substitution. The feature extraction process was attained using the bag-of-words (BOW) approach, TF-IDF model, and word embeddings such as GloVe. Since the study was based on emojis, the prediction algorithm (LSTM) learned to associate emojis with sentiments, transformed them into tokens for performance analysis, and translated them into the format used as input features in the classification. Integration of emojis in sentiment analysis and classification was found to improve the accuracy of the prediction model and yield better outcomes.

Several models were applied in this study, including KNN, LSTM, Transformers, Random Forest, and Logistic Regression. The capacity of these models to accurately predict and classify sentiments was assessed. Transformers model and LSTM emerged as the most

accurate models, with 90.17% and 88.44% accuracy, respectively. Despite transformers being the most accurate, LSTM was the preferred model for sentiment analysis since it did not exhibit overfitting and overspecialization issues that were present in transformers.

Transformers can result in inaccurate predictions on the Twitter dataset, leading to poor practice. Thus, LSTM is the most suitable model for the prediction and analysis of public reactions and emotions toward precautionary measures and policies implemented following a global outbreak or disaster.

The resulting sentiment analysis shows that most individuals have positive sentiments toward most of the lockdown policies. A comparatively large population also exhibited neutral sentiments. 37.3 % of social media users displayed positive emotions, 36.3% exhibited neutral emotions, and 26.4% had negative perceptions toward COVID-19 policies. This suggests that COVID-19 precautionary policies, measures, and restrictions considerably impact public sentiments in general and have a noticeably diverse effect on various societal groups.

Chapter 5: Discussion and Evaluation

In short, this chapter examines the findings of the current study and compares them with the most recent state-of-the-art studies.

5.1 Chapter Structure

The overarching research question in this study was as follows: What is the most reliable and accurate approach or technique to both measure and predict public sentiments for decision-makers while managing and administrating the spread of a pandemic or a natural disaster in general? The study utilized the COVID-19 outbreak-related data/sentiments from across the globe, and it investigated multiple subsidiary questions concerning the policies' impact on various societal groups, most promising accuracy enhancement procedures, emojis' sentiments practicality, and the abilities of state-of-the-art-based hybrid methods.

This chapter is organized into sections that correspond to the various research questions that the current study sought to answer. The research questions are as follows:

- i. In general, is there a notable and detectable pattern in public sentiments and emotions toward precautionary measures during natural disasters?
- ii. How distinct or diverse was the impact of these measures among various societal groups and categories?
- iii. How reliable and effective are emojis in enhancing the accuracy of sentiment analysis and predictions?
- iv. What is the most potentially promising technique to successfully identify emojis' hidden annotations and vague implications?
- v. Overall, what are the most robust machine-learning approaches or hybrid methods to achieve significantly precise outcomes in sentiment prediction and forecasting?
- vi. What is the performance of the developed model(s) compared with the state-of-

the-art benchmarks in the industry?

The study discusses the findings of each section based on the current study and compares it with various works of literature that have attempted to answer that research question. The chapter, however, starts with an ‘overview’ section that gives a brief discussion of the various machine learning models, briefly discusses sentiment analysis as a method of analyzing opinions on Twitter about the attitude or mood of people during the COVID-19 pandemic lockdowns and restrictions, and also briefly discusses Twitter as a source of the data being utilized in this study.

5.2 Overview

Machine learning is a rapidly evolving field of artificial intelligence that has the potential to revolutionize the way people interact with technology. It is a type of technology that enables computers to learn and make decisions without the need for explicit programming (Gopal, 2019). This section will provide an overview of machine learning and its applications, discuss methods, strategies, and techniques used in machine learning, and analyze the benefits and challenges involved. It will also compare the findings of this study and has these findings compared to previous research already reviewed in the current study.

In line with the overarching research question, which was to investigate the most reliable and accurate approach or technique to both measure and predict public sentiments for decision-makers while managing and administrating the spread of a pandemic or a natural disaster in general, the researcher proposed ML methods. The analysis evaluated different ML models to determine which of these methods can measure and predict public sentiments more accurately.

Machine learning (ML) is a sub-field of artificial intelligence that enables machines to learn and improve from experience without being explicitly programmed. ML algorithms use statistical techniques to approximate functions that can generate predictions and decisions

based on data. ML has applications in numerous fields, such as computer vision, natural language processing, robotics, and healthcare. According to research by Janiesch, Zschech, and Heinrich (2021), ML can be used to improve the performance of mechanical systems, including product design, manufacturing, and industrial control. For example, ML can be used to optimize product design by leveraging historical data to accurately predict the performance of a product under different conditions. In manufacturing, ML can be used to identify and control process variations, as well as to facilitate the development of advanced automation systems. In industrial control, ML can be used to predict and control system states, such as in the design of intelligent systems for predictive maintenance. In conclusion, ML has a wide range of applications in the field of mechanical systems and can be used to improve product design, manufacturing, and industrial control.

Machine learning is a powerful tool used to analyze data and recognize patterns in order to make better decisions. The study has discussed several methods, strategies, and techniques used in machine learning. These methods include supervised and unsupervised learning, deep learning, and reinforcement learning, each of which has its own strengths and weaknesses. Supervised learning involves the use of labeled data to train a model, while unsupervised learning does not require labeled data and instead uses algorithms to identify patterns in data. Deep learning is a type of artificial intelligence that uses multiple layers of algorithms to identify complex patterns and make decisions. Reinforcement learning is an iterative approach in which an agent interacts with its environment in order to maximize reward. By using a combination of these methods, strategies, and techniques, machine learning can make more accurate predictions and provide more advanced insights into data.

Machine learning has the potential to revolutionize the way people make decisions. This technology is increasingly being used to improve automation and efficiency in various sectors, such as healthcare, finance, and retail. Machine learning is used to analyze large

datasets quickly and accurately and can identify patterns, structure data, and make predictions. As a result, it can be used to inform decisions and reduce costs. For example, in finance, it can be used to detect fraud, which can save organizations millions of dollars each year. In healthcare, machine learning can help identify diseases and recommend treatments more quickly and accurately than human doctors.

However, there are some challenges associated with machine learning. For example, it requires large amounts of data to be effective and can be difficult to interpret and explain. Additionally, there is the potential for bias in the data or algorithms used, which can lead to inaccurate or inappropriate decisions. Finally, machine learning is often computationally intensive and requires powerful hardware and software, which can be expensive. Despite these challenges, machine learning is rapidly becoming an important tool for decision-making and is likely to be increasingly used in the future.

Machine learning has become an increasingly important tool for a range of applications. By understanding the methods, strategies, and techniques used in machine learning, such as supervised learning, unsupervised learning, and reinforcement learning, people can gain a better understanding of how machine learning has been and will continue to be applied to the development of improved data-driven applications. Furthermore, by understanding machine learning, individuals can learn more about how to utilize advanced and emerging technologies to solve complex problems in our daily lives. The challenges will be explored in detail in a later section.

The world has undergone a tremendous transformation recently, with the emergence of social networking platforms taking center stage. Social media has become a powerful tool for governments and individuals to interact, share information, and express their views and opinions. In this context, sentiment analysis of social media data has gained considerable attention and is rapidly becoming a popular research field. The importance of sentiment

analysis in social media has been widely recognized recently due to its potential applications in various areas. It enables organizations to gain insights into public opinion, identify trends, and monitor consumer sentiment. It also allows governments to understand better how the public perceives their policies and measures. Moreover, sentiment analysis in social media can be used to detect and predict public sentiment and reactions during natural disasters, such as the recent COVID-19 pandemic, which has caused unprecedented disruption and brought about drastic changes in how people live and interact. This paper reviews the literature on sentiment analysis in social media, focusing on analyzing and predicting the public's sentiment and reactions toward governments' preventive measures and policies during natural disasters, such as the recent COVID-19 pandemic.

Natural language processing (NLP), computational linguistics (CL), information retrieval (IR), machine learning (ML), and artificial intelligence (AI) are all components of sentiment analysis, which measures how people feel about a given product, service, person, business, organization, event, or subject (Birjali et al. 2021). Similar concepts to sentiment analysis include subjectivity analysis, opinion mining, sentiment classification, polarity classification, and affect the analysis.

Collecting and analyzing data on how people feel about your brand online is called social media sentiment analysis. Instead of simply tallying the number of times something was mentioned or commented on, sentiment analysis considers how people feel about those things. An approach to natural language processing, it analyzes written material to ascertain the author's intent. The three most common ways to categorize emotional tones are positive, neutral, or negative (Birjali et al., 2021). Understanding client satisfaction, keeping tabs on social media, and predicting stock prices all make considerable use of this method.

Sentiment analysis seeks, at its most fundamental level, to label a writer's attitude toward a set of topics as positive, negative, or neutral. Sentiment analysis is useful in many

fields, not just business intelligence, politics, and sociology (Birjali et al. 2021). Social media has become increasingly significant over the past five years, ranking as the fourth most popular medium for gaining access to critical information in the wake of disasters and crises. Sentiment analysis or opinion mining is a growing field of research that uses algorithms to analyze public reactions to messages in order to predict public opinion (Giachanou and Crestani, 2016). Applications in this field are also growing. In recent years, there has been a surge in interest in this area as digital platforms have become increasingly powerful tools for disseminating information and influencing public opinion.

One of the most commonly used methods for analyzing sentiment is the sentiment lexicon (Agarwal et al., 2016). This approach uses a large database of words and phrases to identify positive and negative sentiments in text. Using this information, researchers can then use the lexicon to identify patterns in public reactions. Another common approach is sentiment analysis through text mining. This involves using algorithms to extract information from the text in order to identify patterns in sentiment (Poria et al., 2014). This information can then be used to predict public opinion. Both of these methods have been used to analyze reactions to a range of topics, including political messages, marketing campaigns, and news articles. The field of sentiment analysis and opinion mining is still in its early stages, and there is still a lot of research to be done. However, the potential for this technology to be used to understand public opinion is very real.

Sentiment analysis is a method of analyzing the sentiment of a text to determine its mood or feeling using artificial intelligence to identify and analyze opinions in a given set of data (Ain et al., 2017). The benefits of sentiment analysis are numerous, as it can be used to determine the mood of a population, predict public reactions to events and more. Some potential applications include determining the mood of a population before an election or after the announcement of a certain policy to gauge the public reaction. Opinion mining can

be used to predict how people will react to new products or ideas as well. This is a powerful tool that can be used to improve understanding of the views. It has the potential to make lives easier and help to make better decisions.

Sentiment analysis is a critical tool used in a wide array of fields, including customer service, marketing, and social media. By systematically identifying and extracting affective states and subjective information, sentiment analysis can help organizations understand and improve the customer experience. Sentiment analysis has been used in customer service for years. For example, a company may use sentiment analysis to identify whether a customer is angry, frustrated, or happy. This information can then be used to improve the customer experience. Similarly, marketing teams can use sentiment analysis to understand the attitudes of their target audience. This information can then be used to create more effective marketing campaigns. Finally, social media users are constantly interacting with others online. Sentiment analysis can help users understand the emotions of those around them as well as help social media companies improve interactivity.

The COVID-19-induced lockdowns have stirred up conversations on all social media platforms, such as Facebook and Twitter, which were initially designed to offer a means of interaction between family and friends. However, the platforms gained wide recognition and acceptance for the sharing of emotions, feelings, ideas, and feedback, including among strangers, on almost everything and anything. Now, as noted by Latif et al. (2019), social media is viewed as an integral part of the life of an individual. Thus, during disease outbreaks, pandemics, and disasters, it is more than critical to develop an understanding of the social media shared content. Moreover, with little to no data on how the general public copes, responds, and reacts to lockdowns that have been influenced by infectious disease, the current study found it fit to explore the emotional dimension linked with the event. The

mentioned dimension is more likely to exhibit behavioural consequences, and as discovered, the COVID-19 pandemic increased anxiety and depression in numerous individuals.

Twitter appeared to be the most effective in sharing real-time information among all social media platforms. The current study discovered that emotions and sentiments expressed on Twitter provided insight into how individuals were coping with the lockdown.

Surprisingly, the more positive coping sentiment was the most recorded among social media users. Generally, this showed that during the first stages of the COVID-19 pandemic, individuals positively appraised the lockdown measures and felt confident in coping with the measures. Furthermore, the public appreciated the need for measures that facilitated social distancing and significantly reduced the chances of infection. The neutral coping sentiment was a reflection of the sentiment and emotion of the population towards the pandemic. The current study could also consider several potential factors behind why most positive citizens displayed a positive sentiment towards the lockdown measures. From the findings, lockdown measures were deemed a necessary precaution in the prevention of the further spreading of the virus, and the negative emotions meant that the general public believed that lockdown measures would not reduce the spread of the virus.

Szreter and Woolcock (2004) defined social capital as the sum of resource-rich networks, reciprocal, and trustworthy connections. From an individual view, literature revealed that social capital was a significant determinant of several health outcomes, with plausible pathways to mental health from social capital. From a similar perspective, social capital was also identified as an asset that mobilized and empowered society and its members. A specifically highlighted perspective is the significant role of social capital when individuals face catastrophes or disasters. For children, precautionary lockdown measures were discovered to make them come to the realization of an ongoing change across the globe. Literature has also shown that when compared to the mental health of adults, adolescents

exhibited fewer instances of depression and stress. A similar level was also identified between adults past the age of 50. Also, adolescent girls exhibited a likelihood of anxiety and depression symptoms.

The current study, featuring social capital, discovered that the effective functioning of social groups via relationships formed through the network is likely to impact mental health and might also be impacted by the COVID-19 pandemic. Despite the lifting of the lockdown measures, there were still recommendations on social distancing, with parents adopting advice to stop their youths from undertaking networking activities. The measures affected the social contact of youths, specifically with their community and friends, since online communication platforms were unable to compensate for the emotional demands identified in face-to-face community participation and communication. The elderly population was also identified as seeing the least benefits from social media platforms. The digital divide on the basis of age describes a longstanding inequality in access to, and various skills, to make use of new technology. While the issue remained common during the pandemic, it has led to the suffering of a large portion of the population from the negative impacts of the crisis, together with the inability to make use of the digital measures put in place to help.

5.2.1 Applications of Sentiment Analysis in Different Contexts

Sentiment analysis and opinion mining have gained prominence in recent decades as a tool for understanding public sentiment. The field of sentiment analysis has been used by businesses to parse and understand the sentiment of tweets and other online communications. The ability to understand sentiment not only assists in understanding public opinion but can also be used to improve the quality of online content. Sentiment analysis has been used by businesses for a number of purposes as well, from marketing campaigns to gauging customer reactions to products. By understanding the sentiment of communication, businesses can

better parse and understand the message being sent. This information can then be used to make decisions about the future course of the business.

Sentiment analysis has also been used by researchers in the field of opinion mining. Opinion mining is the process of understanding the sentiment of a dataset. By understanding the sentiment, researchers can better understand the ideas that are being conveyed. It can provide valuable information about the tone of a given piece of content. For example, positive sentiment can be indicative of a post that is likely to be liked or approved by readers, while negative sentiment can suggest that a piece of content is likely to generate a negative reaction. Secondly, sentiment analysis can help to identify patterns in opinions (Denecke and Deng, 2015). When a large number of positive responses are generated to a given piece of content, it could be inferred that the content is favorable. Conversely, if a large number of negative responses are generated, it could be inferred that the content is unfavorable. By tracking sentiment over time, it may be possible to identify shifts in opinion and monitor trends in a given topic. Overall, sentiment analysis is an important tool for understanding the tone of online discourse. While it can provide a wealth of information, it is important to be aware of the limitations of the technique. For example, sentiment analysis is not always accurate and cannot always identify all patterns in opinion (Cambria et al., 2017). Nevertheless, sentiment analysis is a valuable tool for monitoring trends and understanding the tone of online discourse.

A study such as Agarwal et al. (2016) presents a machine-learning approach to sentiment analysis, which they call "prominent feature extraction for sentiment analysis." The authors describe a two-step process for identifying important features in a text corpus. The first step is to identify the most prominent words and phrases, and the second is to identify the most important relationships between these words and phrases. The authors demonstrate the effectiveness of their approach by using it to analyze a news article about the Indian

Prime Minister, Narendra Modi. The article contains a mix of positive and negative sentiments, and the prominent features identified by the machine learning approach include words and phrases related to Modi's political goals (e.g., "victory," "proud," and "strong") and words and phrases related to his opponent (e.g., "weak," "opposition," "corruption"). Overall, the Agarwal et al. approach is promising and could be useful for sentiment analysis in a variety of contexts.

One of the most common ways to analyze sentiment is to look at the sentiment of individual words. This is called sentiment analysis by word. Sentiment analysis by word is a common method because it is easy to find specific words that have a particular sentiment. There are a few problems with sentiment analysis by word. First, it is easy to bias the results as it is easy to change the sentiment of a word by choosing which words to include and which to exclude. This is especially true when the sentiment of a word is subjective. Second, sentiment analysis by word is not very accurate because it is easy to mistake a positive sentiment for a positive word and a negative sentiment for a negative word. Finally, the use of sentiment analysis by word is not very sensitive to changes in sentiment. If the sentiment of a message changes, the results of the sentiment analysis by word will not change.

Instead of sentiment analysis by word, it is proposed to use sentiment analysis by context. This is a more sophisticated method that takes into account the surrounding text. Sentiment analysis by context is more accurate than sentiment analysis by word, but it is also more sensitive to changes in sentiment. There are several ways to do sentiment analysis by context. One way is to use a rule-based algorithm. This algorithm looks at the context of a word and decides whether the sentiment of the word is positive, negative, or neutral (Poria et al., 2014). Another way is to use a machine learning algorithm. This algorithm is trained on data sets of positive, negative, and neutral words. It can then be used to analyze new text.

Both of these methods are more accurate than sentiment analysis by word, but they are also more expensive and time-consuming.

While sentiment analysis of a single word is possible using lexicons, it is difficult to capture the sentiment of concepts due to the complexity of language and sentiment. To address this, the authors propose Sentic Patterns, which use a combination of semantic networks, concept-level sentiment databases, and dependency-based rules. The results show that their approach outperforms a number of existing sentiment analysis methods, implying that sentiment analysis based on machine learning is considered a superior method as opposed to analyzing individual words or contexts. In their work, the researchers conclude that their approach is a promising method for concept-level sentiment analysis and that it can be applied to other tasks, such as sentiment analysis of longer texts. They also suggest that additional research should be done to improve the accuracy of the approach.

There is a potential for sentiment analysis in providing valuable insight into the public's opinion on a variety of topics related to healthcare practices or any other practice in this field that continue to expand and become more complex. Sentiment analysis can be used to reveal public opinion on controversial topics, such as end-of-life decisions, genetic testing, and other areas of medical ethics. If it provides healthcare professionals with valuable information about how the public feels about specific issues, then it can be used to inform policy decisions.

While this is possible in healthcare, several challenges posed by the use of sentiment analysis in healthcare exist. These include the potential for bias in the data and the need for better methods for measuring sentiment. Furthermore, sentiment analysis has the potential to be used as a tool for predictive analytics to measure future events or reactions. It is clear from Denecke and Deng (2015) that sentiment analysis has the potential to offer valuable

insight to healthcare professionals who should take advantage of this technology to help inform their decisions, but there is a long way ahead before it becomes fully usable.

5.3 Q1 and 2: Public Sentiments Patterns During Natural Disasters

This subsection examines the findings of the current study in terms of how COVID-19-related restrictions affected different societal groups. The findings are then compared to existing research reviewed in this study. In this section, most of the information is relevant to the second hypothesis (hypothesis 2), and by the end of the section, the study provides a verdict on the extent to which the hypothesis has been supported. The hypothesis that COVID-19 precautionary policies, measures, and restrictions considerably impact public sentiments in general, as well as having a noticeably diverse effect on various societal groups, is mostly addressed in this section.

Different disasters, both natural and man-made, have caused immense damage to life and property, with the number of affected people and the extent of the damage increases over the years, with some people affected more than others depending on their unique circumstances. In recent times, the role of social media in natural disasters has been increasingly recognized. Social media serves as a platform for people to communicate and share information, which can be used to assess the impact of natural disasters in real-time. This has enabled governments to respond faster and to develop and implement more efficient strategies for disaster management. In recent years, the impact of these events has become more severe due to climate change and global warming. In response to such catastrophes, governments have taken various preventive measures and measures to protect their citizens. These measures can include evacuations, sheltering-in-place, and restrictions on movement and activity. However, the effectiveness of these measures and policies depends on the public's sentiment and reaction toward them. As such, there has been an increasing focus on sentiment analysis in social media to gauge the public's opinion.

The impact of COVID-19 precautionary measures on public sentiment and emotions has been widely studied in recent years. Studies have generally found that people largely support government measures such as lockdowns, travel restrictions, and social distancing. For instance, a survey conducted by Boon-Itt and Skunkan (2020) found that the government's decision to impose a national lockdown was supported by most, although the analysis of sentiments in this study revealed that in most parts of the world, there generally a negative reaction to these measures. Most of the literature reviewed in this study demonstrated that while there was some initial resistance to lockdowns, overall, the public was mainly in support of the measures. However, there is evidence that the impact of these measures is not uniform across different societies and groups. Studies reviewed in Chapter 2 have found that poorer and minority communities have been disproportionately affected by the pandemic and the associated restrictions.

Similarly, another study by Xue and partners found that people of lower socioeconomic status were more likely to report negative emotions such as fear, anxiety, and depression in response to the pandemic and associated restrictions (Xue et al., 2020, pg 4). These findings suggest that the impact of COVID-19 precautionary measures on public sentiment and emotions is not uniform and that poorer and minority communities are disproportionately affected by the pandemic and the associated restrictions. The impact of COVID-19 precautionary measures on public sentiment and emotions is also influenced by factors such as age, gender, and geographical location. This is to further suggest that different groups of people in society experienced the restrictions differently, again, depending on their unique circumstances. For instance, the study by Boon-Itt, and Skunkan (2020), found that depending on age, some young individuals between the age of 18 and 35 were more likely to report feeling stressed and anxious in response to the pandemic and the associated restrictions compared to older people aged for instance, 55 and above. These findings suggest that the

impact of COVID-19 precautionary measures on public sentiment and emotions is unique across different demographic groups.

The sentiments analyzed in this study showed that while the overall sentiment towards the lockdown policies was largely negative, there were distinct differences in sentiment among different groups. For example, the study found that those with higher levels of income were more likely to be more accepting of the measures, while those with lower levels of income were more likely to be more resistant. Furthermore, the study found that those in rural communities were more likely to be less accepting of the measures than those in urban communities. This suggests that the impact of lockdown policies on sentiment is unique across all social groups. Therefore, different groups may have different reactions to the same set of guidelines. Undeniably, the confining measures set to mitigate the COVID-19 pandemic severely affected people's psychological, emotional, and behavioral activities. The impacts of the measures adopted to counter COVID-19 were distinct across almost all societies worldwide. One major societal group that was affected during the pandemic across the world is healthcare providers. Being on the frontline in serving their nations during the pandemic, these workers were forced to work in different environments and conditions. The conditions included a high workload, the agony of losing family members, patients, and colleagues, and insufficient medical supplies and personal protective equipment, mostly in evolving nations. The current study also reviewed literature that demonstrated that such situations and environments make health professionals more vulnerable to mental health burdens due to increased anxiety, stress, anger, depression, and insomnia.

The elderly were more vulnerable to the virus, and as such, how they were affected by these precautionary measures differed significantly from how young people, for example, were affected. The lockdown measure imposed on the elderly, instructing them to stay at home, has shown how the elderly are at a greater risk for depression and anxiety. The social

disconnection of the elderly caused and escalated depression, anxiety, neglect, and loneliness that further make the elderly incur long-term health problems. As far as the findings of this study on these research questions were concerned, it can be concluded that while focusing on the positive aspect of human sentiment, it is critical to acknowledge the diversity in the Twitter population, who were the participants of this study. This diversity was best shown by varying attitudes and opinions towards the set precautionary policies in different parts of the globe since the data used in this study span the entire globe. Each region had a varied number of negative tweets about various government actions. As quoted in the previous chapter, the eastern region recorded the highest number of negative sentiments (accounting for 28%) of the total tweets obtained from that region, while the southern region recorded the lowest proportion of negative sentiments. The diversity in the population was further highlighted by the presence of neutral tweets. These were individual sentiments that had both positive and negative elements. Such sentiments were predominant in the northern region, where they accounted for more than 20% of the total tweets. These findings suggest that hypothesis 1 was sufficiently supported, and the conclusion made is that COVID-19 precautionary policies, measures, and restrictions considerably impact public sentiments in general as well as having a noticeably diverse effect on various societal groups.

5.4 Q3: Reliability of Emojis

The current study uses emojis as part of sentiment analysis, and one of the key research questions was whether or not emojis are reliable and effective in enhancing the accuracy of sentiment analysis and predictions. The findings of the current study have demonstrated that emojis in the Twitter data were used to express different emotions, from sadness to anger and happiness. The study demonstrated that emoticons were primarily used within a specific context and showed the relationship between the use of an emoji and the post topics. Analyzing the tweets revealed that the topic could be categorized into different clusters,

including Celebrity news, government reports, social events, and online games, as well as COVID-19. These topics generated more than 2 million tweets during the pandemic, and the use of different emoticons was closely related to different posts. For instance, tweets relating to the death toll and the number of infections were often accompanied by sad-related emojis, while those related to government actions that proposed further restrictions were accompanied by angry emojis. As explained in the previous chapter, these clues given by emojis could be used to reveal hidden messages during the pandemic. In the next paragraph, this section examines some of the key literature to determine how it compares to the findings of the current study.

The modern policy-making process trend is influenced by events within the mainstream and social media. According to Yoo & Rayz (2021), more than a billion social media posts are generated daily; with social media platforms growing, the number of daily posts is likely to increase. This increase has created the need to adopt sentiment analysis in different human-related applications. Sentiment analysis in this study has been defined as a natural language process mechanism used to analyze and identify attitudes and potential implications of information within a given text-based data. Users of social media have adopted unique systems of communication. One of these changes has been the use of emojis as a key part of the interaction. This raises the question of whether emojis can be reliable in improving the accuracy of sentiment analysis.

This study also focuses on investigating the use of emoji characters on social media and their effects on text mining and sentiment analysis. The authors recognized that the opinions, preferences, and emotions of individuals in social networks tend to vary depending on the topic of interaction. As a result, businesses have been able to adopt this approach in identifying some key elements of dissatisfaction within the market (Shiha & Ayvaz, 2017). This is the authors adopted text-based data that focused on positive and negative global

events. The results showed that the utilization of emoji characters in sentiment analysis tends to yield higher sentiment scores than plain texts. These observations were attributed to the ability of emojis to currently identify the emotions of an individual, thereby allowing analysis of intonations across text-based data. Further, Shiha & Ayvaz (2017) observed that emoji characters had higher impacts on the overall sentiment score of positive opinions than on negative opinions.

The insights provided by Shiha & Ayvaz (2017) highlight the importance of emojis in enhancing sentiment analysis. The analysis in the previous chapter indicates that emoji embedding is critical in improving different sentiment analysis methodologies. In the case of Liu et al. (2021), the improvements were attributed to the diversity of semantics and expressions associated with specific emotions. While recognizing these elements of sentiment analysis, it was vital to recognize that Liu et al. (2021) focused on the syntax and semantics of the Chinese language and lettering. At the same time, Liu et al.'s (2021) examination focus on adopting machine language models such as Bi-LSTM as a mechanism for emoji-embedding. The adopted model was considered higher in accuracy than other sentiment analysis models. Therefore, it is critical to acknowledge that emojis play a crucial role in improving the overall approach to sentiment analysis (Liu et al., 2021). This makes the approach highly suitable when dealing with policy issues surrounding unpredictable events such as natural disasters.

Sentiment analysis refers to a class of natural and computational language processing-based methods to extract, assess or characterize subjective data, like opinions, expressed in a specific text. The model's primary purpose entails classifying a person's attitude toward different topics into negative, positive, and neutral categories. Emojis are graphic symbols entailing specific meanings and are embedded within the micro-texts, helping express emotional meanings and offering novel information on an individual's sentiments. They are

integral to online communication, particularly social media. Data such as tweets and web postings express opinions on different events and topics. A study performed in Wuhan deploying quantitative analysis with the sample data deployed, totaling 38,183,194 posts taken from Weibo, shows that emojis can affect the accuracy of sentiment analysis (Liu, 2021). Approximately 40.88 percent (15,609,843) posted before the pandemic consisted of at most one emoji, while after the pandemic, 57.15 percent deployed emojis at least once. The findings revealed that the use of emojis could greatly enhance the accuracy of sentiment analysis. The results showed that emojis offer implicit information about the sentiment expressed in the text, which is more accurate than relying on the text alone. Research shows that the use of emojis improves the accuracy of sentiment analysis by 5-6 percent (Liu, 2021). They offer explicit information about the sentiment expressed, which can be more accurate than relying on the text alone.

As the author of this study established, emojis pose different challenges, including difficulty interpreting the meaning of smiley. Different emojis have unique meanings depending on the context and hidden annotations and implications that are not immediately obvious. A great demonstration of this entails a study by Aljameel Sumayh utilizing a sample dataset of Arabic COVID-19-related tweets as the primary data source. The process employed machine learning predictive models, encompassing, K-nearest neighbors, and Naive Bayes (NB), along with the N-gram feature extraction technique to process and analyze the data. The aim entailed investigating the effectiveness of the implemented models and feature extraction techniques in predicting individuals' awareness of the precautionary procedures during the COVID-19 pandemic in different regions. The finding showed that disasters increase containment measures by eighty-five percent. Li (2022) posit that containment measures have a direct link with the reliability and effectiveness of emojis in enhancing the accuracy of sentiment analysis and predictions. This is because the measures

affect the public's reactions toward the government's policies and actions during a natural disaster such as the COVID-19 pandemic. Using emojis in online communication offers more accurate information about people's sentiments concerning containment measures (Raamkumar, 2020). Their emotions and attitudes towards these measures can vary. Therefore, emojis offer explicit information about the sentiment, which is more informative than relying on the text alone.

Using emojis has been shown to be an effective way to enhance the accuracy of sentiment analysis. Emojis can provide additional context to a statement, allowing for a greater understanding of its sentiment. Emoji usage can also help to identify implicit or hidden annotations, which can be challenging to detect using traditional sentiment analysis methods (Shiha and Ayvaz, 2017, pg 2). For example, a sentence containing "I'm not happy about this decision" may be interpreted as a negative sentiment. However, including an emoji such as a smiley face may indicate that the sentiment behind the statement is much more positive. Thus, emojis can help identify hidden annotations that would otherwise be difficult to detect.

Emojis can help identify vague implications in a statement. This can be especially useful when analyzing texts in languages other than English, as the use of slang or idioms can often be challenging to interpret. For example, a statement written in Spanish may contain an expression such as "estar contento," which is often used to express satisfaction or pleasure. With the use of an emoji, it can be easier to interpret the sentiment of the statement. However, including an emoji, such as a smiley face, can indicate the sentiment behind the statement, allowing for a more accurate analysis.

Emojis are a popular means of expressing emotions on social media. They can be used to supplement sentiment analysis in natural disasters and government precautionary measures. Emojis can be used to convey sentiment to a greater degree than traditional

language, and they can be used to capture nuances in sentiment that might otherwise be overlooked. Moreover, emojis can identify the intensity of sentiment and capture both positive and negative sentiments (Shiha and Ayyaz, 2017, pg 6). Thus, they can enhance important insights into the public's perception of the government's actions.

However, the interpretation of emojis can be problematic. Emojis can often be interpreted differently, and their meaning can vary across cultures and contexts. Moreover, their usage can be open to interpretation, as multiple emojis can be combined to form complex messages that can be difficult to decode (Chuchu et al., 2021, pg 2). Thus, it is crucial to develop techniques to accurately identify emojis' hidden annotations and implications to enhance the accuracy of sentiment analysis and predictions.

With the ever-increasing popularity of social media networks like Facebook, Twitter, and Weibo, people are now openly using these platforms to express their opinions. These channels generate massive quantities of data every day, which, according to (Kavanaugh, 2012), is an important source of insightful information for various analysis reasons, including research purposes, election results forecasting, stock market analysis, and business analytics, among other data-driven tasks which is the main aim of social media data sentiment analysis. According to this author, sentiment analysis in the recent two decades has generated notable success from the analysis of user-generated data. However, the current analysis focuses broadly on text-based user data despite the fact that several occurrences of semiotic expressions such as Emojis, images, and memes are frequently used by users to express themselves creatively on these social media platforms. Kavanaugh (2012) highlighted that leaving out Emojis in sentiment analysis is inadvisable as they carry a tone of non-verbal information significant in people's expression of their opinions and feelings on social media.

As the author of this study established in the previous chapter, Emojis can be useful in enhancing the overall performance of sentiment analysis and predictions to some degree in

the case of social media. This study also highlights that Emojis could be used in conveying emotions and attitudes in text-based communication, and they can equally provide additional context that could not be captured by words alone. In this case, Emojis can help in disambiguating the sentiment of a message and distinguishing between similar words or phrases that could have different emotional connotations. The researchers were keen to note that the effectiveness of Emojis in sentiment analysis was closely related to several factors, including the language and culture of the users, the specific use case and domain, and the quality of the sentiment analysis model. An example is whereby some cultures and languages may use different Emojis to convey the same emotion, and some Emojis may have multiple meanings depending on the context. Additionally, poorly trained sentiment analysis models may not effectively process Emojis, which could limit their utility in these cases.

It was realized in this study that the effectiveness of Emojis in sentiment analysis is highly dependent on several factors, such as the users' language and culture, the specific use cases and domains, and the quality of the sentiment analysis model. His research on Twitter sentiment analysis (TSA) highlighted that sentiment analysis had been successfully implemented in the monitoring and forecasting of election results in multiple instances. This research, however, noted that many existing studies are mostly reliant on the extraction of sentiments from textual-based features, but a low number of studies have supported the inclusion of Emojis for sentiment analysis in election forecasting. The researchers hereby incorporate N-gram features in the prediction of 2017 Uttar Pradesh vote shares for the legislative elections. Tweets that contain Emojis exhibit a discernibly distinct sentiment distribution compared to those without Emojis. More than 300,000 tweets were collected with geo-tagging on search keywords to only extract exclusive election data. Seven lexicons were used in the labeling and classification of the tweets and the comparison of two methods to reduce prediction error – polarity of tweets and sentiment magnitude-based criteria. This

study, therefore, discovered that Emoji sentiments were more significant in the detection and prediction of vote shares. The researchers concluded by noting that the incorporation of Emoji sentiments and N-gram features greatly improves the overall prediction.

Government officials are constantly looking to leverage this by sentiment analysis of social media platforms which produce staggering amounts of information to ensure improved service delivery and communication for their respective citizens. According to Gao (2020), multiple potentials exist in the identification of real-world problems in real-time, which could be addressed quickly by emergency managers deployed to respond to these life-threatening calls. This study highlighted the need for high-quality medical care resources provisioning by many governments and how artificial intelligence (AI) integration into healthcare could help alleviate crises and act as a reliever to related healthcare workers. Using social media sentiment analysis to obtain general public opinions on this matter, the public perception of AI indicated that a high number of people held a positive attitude toward the inclusion of AI and machine learning in the healthcare sector. Therefore, based on sentiment analysis of social media content, governments can conduct analysis to find out what could be beneficial and widely acceptable by the general population before rolling out new technology or policies.

Sentiment analysis can be used to gain an understanding of public reaction toward precautionary measures by the government amidst the COVID-19 pandemic. This study utilized multiple machine learning algorithms to analyze a dataset of tweets related to COVID-19 that were collected during the pandemic. The aim was to determine which of the models performed best, as will be discussed in the subsequent sections. The best-performing model would be recommended for sentiment analysis. This model could be utilized by government agencies to determine the appropriate approach for taking related to each region based on their sentiments regarding certain occurrences. This is an indication that the

government could leverage the sentiment analysis approach to make informed decisions and tailor their responses to a pandemic in different regions.

Studies such as Li (2020) investigated the significance of noticing the credit of humor, slang, and pictograms for effective social media processing. In this study, the researchers introduced a Humor-Emoji-Slang-based (HEMOS) system approach geared toward sentiment categorization for China's population through machine learning. The researchers first collected 576 often-used social media slang expressions and used these to create a slang lexicon. Secondly, 108 Weibo Emojis were converted to their word equivalent, and these were used to generate an Emoji lexicon for the Chinese population. During the following step, two polarity annotations were created - "positive" and "negative." The two lexicons were applied on a machine learning model (AttBiLSTM) neural network, and their performances were tested on a labeled dataset. Experimental results indicated that the proposed approach could greatly help in sentiment analysis on Weibo, China's largest social media network.

The global acceptance of Emojis is an indication of multicultural and the normalization of Emojis usage. However, the author noted that Emojis used across different linguistic regions and diversity may bring about high differences in emotional expression. In this study, a comparison of Emoji usage in social media was carried out in relation to the context, frequency, and area of discussion across Asian countries (China and Japan) and Western countries (USA, Canada, and the UK). Analysis of this research indicated that Emoji usage both in Asian and Western countries follows a normative yet cultural-specific pattern. The researcher also went ahead to analyze the context of Emoji interpretation in combination with textual and non-textual cues, which was found to successfully improve the overall intended meaning of the Emoji use in the contexts of Asian and Western countries.

Derks (2008) investigated the role of Emojis in social media communications. This study consisted of an online questionnaire on the contexts of Emoji usage with an experimental part whereby 1,251 participants had to respond to some precise internet chats. The researchers contextualized the experiment into two broad chats – interaction between friend and stranger and the context either positive or negative. Results indicated that Emojis had a link with emotional expression in strengthening the message conveyed as well as in the expression of humor. This study, however, noted that the usage of Emojis was more evident in friends' communication as compared to communication with strangers. Similarly, Emojis were prevalent in positive contexts rather than ones that would be negative. The authors compared Emoji use in a similar way to face-to-face communication with regard to the interaction partner and/or social context.

Supervised machine learning has exceptionally done well in multiple tasks involving natural language processing, and its deployment for the identification of Emojis' meaning has borne substantial success in the training of models based on a dataset of tweets containing both texts and Emojis. This study noted that the development of the emotion-related collection is a difficult annotation task due to conflicting human opinions. One of the solutions to these problems includes the use of explicit emotional indications by the authors of the current social context in question as well as using self-annotated data. The researcher presents a study consisting of self-annotated Unicode Emojis based on a user's Twitter emotional situation. Emojis were found to be used far more extensively as compared to hashtags while expressing users' emotional positions. A set of tweets with indicative emotions were collected, and a sample annotation for emotions was created. The accuracy and efficacy of Emojis as annotations labels were evaluated using statistically trained models, and the results indicated that more caution is needed to optimize the message conveyed by Emoji usage.

Park (2018) presented a machine-learning model for solving Tweet effects with the main objective being on creating a model sentence and word-level interpretations of emotions conveyed by distinctly labeled data on Emojis and hashtags. Through the exploitation of the neural networks model, emotional knowledge coupled with collaborative filtering was extracted through traditional support vector and logistic regression models, whereby sentiment analysis results from these models proved superior in conveying an audience's emotions and expression socially. Sentiment analysis is a key task for government policymakers and researchers in social media analytics (Bertrand, 2013). In this study, the researchers analyzed data from Twitter's micro-blogging site to generate a sentiment classification of New York City. A classifier was developed based on Twitter messages to analyze certain keywords, phrases, and emoticons in aid of determining the mood of a given tweet. In combination with user geotagging information and collaborative filtering, the researchers were able to gauge public sentiments with far better classification results.

Semantic models could be used alongside other techniques to identify Emojis' hidden annotations and vague implications. The researchers highlighted that these models capture the underlying semantic relationships between words and concepts. By incorporating these models into Emoji analysis for social media sentiment analysis, the study found that it is possible to identify subtle differences in the meaning that might not be apparent from the literal interpretation of an Emoji.

Titov (2008) presented a novel model whereby ratable aspects were drawn from online user reviews. Extracted aspects were used in opinion mining and sentiment analysis, whereby the results were presented in a graph based-format. These approaches were effective in establishing the relationship between Emojis, words, and concepts. This allowed for more sophisticated modeling of Emoji meanings in any given context.

A user-based analysis approach could be useful in analyzing the behavior and demographics of social media users to gain an understanding of the meaning and implications of Emojis. This could be done more effectively by analysis of frequency and the context in which particular Emojis are used by certain groups of users. Doing so, according to this study, would reveal insights into such groups' attitudes and beliefs.

As previously mentioned, social media served as an essential platform for the public to air their opinions concerning COVID-19 containment measures like lockdowns. However, the interpretation of most of these tweets could have been better as the applied machine learning techniques often fail to categorize them accurately in the selected cohorts (Kandasamy et al., 2021). The subjectivity of the expressed opinions can make it challenging for researchers to categorize and analyze the social media sentiments on COVID-19 containment measures. As such, scholars like Yoo & Rayz (2021) uphold that the problem of interpretation of social media sentiments could be resolved using emojis which can be described as graphic symbols representing different concepts. According to the author, most people use emojis to clarify their emotions, increasing the overall efficiency of communication on social media platforms. The accuracy of sentiment analysis could be improved through proper emojis analysis. This view can be supported by the notion that, like other urban words, emojis tend to have closer meaning than word synonyms. Additionally, the meaning of emojis remains relatively stable over time. For this reason, researchers interested in sentiment analysis often replace emojis with a single word or a few words. The result has been a 2 percent increase in sentiment analysis accuracy. However, this argument is countered by Kandasamy et al. (2021), who point out that emojis tend to be additives with no tied meaning to the expressed opinion and thus should be deleted when analyzing sentiment. Therefore, expression emojis like crying and laughing were removed while a single term replaced others.

Emojis play a vital role in expressing sentiments related to COVID-19. However, the use of emojis varied demographically and geographically, with certain regions showing increased emoji usage than others and vice versa. Some of the most common emojis related to COVID-19 containment measures were thought of using masks, flight cancellation, and vaccination. According to the investigators, emojis could add another attribute to social media sentiment. Following this point, the current study suggested that emojis can serve as surrogacy for collecting emotional data. However, this idea still needs to be fully explored in the study. The idea that emojis help expresses deep emotions and are widely used in public sentiments concerning the COVID-19 pandemic is one of the central premises in the current study that guided further analysis and will be discussed in the next sections. As the Pandemic evolved, the public shifted towards emojis expressing sadness which is a logical reaction to the prevailing situation. Despite Stoikos & Izbicki's (2020) contribution in determining whether emojis help improve the reliability and effectiveness of sentiment analysis concerning COVID-19 lockdown measures. Chuchu et al. (2021) Attempt to clarify the essence of emojis in sentiment analysis by stating that the feature improves the accuracy of sentiment analysis. Additionally, the overall performance of the sentiment analysis algorithm could be increased by considering the usage of the different emoji. For instance, Chuchu et al. (2021) used the Cemo-LSTM emoji-embodying model to analyze the Chinese population's sentiments about the Pandemic. The researchers concluded that the Pandemic had a significant impact on Chinese public sentiment. As the Pandemic evolved and the population continued to live under stringent containment measures, the sentiments shifted to reflect passive emotions such as sadness and horror.

Emojis serve as effective vehicles for conveying ideas and sentiments and thus could offer essential insights into the public's discourse concerning the Pandemic. This argument stems from the view that people express themselves differently, and failure to consider certain

forms of expression alienates the views of specific population segments. However, the author does not elaborate on whether emojis could help increase the predictability and accuracy of the public's sentiments about COVID-19 containment measures. Feng & Zhou (2022) explain that understanding the public's perception concerning the imposed lockdown to minimize the spread of the virus should entail a close analysis of both polarized words and emoji. The ubiquitous nature of facial emojis makes them suitable for studying the population's expressed emotions and opinions as they convey mixed feelings that could be categorized into three classes; neutral, positive, and negative. Based on this categorization, the researchers detected 35.2, 18.1, and 46.6 percent positive, neutral, and negative emojis, respectively. These figures align with research findings that depicted the public's sentiment towards lockdown as largely negative, but it shifted to positive when governments started re-opening the economy. However, Feng & Zhou (2022) concluded that whereas facial emojis offer clues into the interpretation of emotions, they tend to be subjective and thus open to multiple interpretations. For example, a smiling emoji falls under the positive category, but it can also be used to express sarcasm which might subsequently affect sentiment detection accuracy. For this reason, the researcher suggests that emoji-based sentiment analysis should be fine-tuned to increase accuracy and predictability. (Chuchu et al., 2021) add that introducing emojis, especially when using an analytical technique that merges both words and emojis, could significantly improve the performance of sentiment analysis. Additionally, emoji tag words could be used directly when determining the features for classifier training. Once the emoji usage has been classified, the performance of each of the stipulated algorithms improved significantly.

Emojis, as this study has demonstrated, have gained popularity over the years and could offer important insights that would guide the interpretation of online sentiment. Despite this, the authors favored an approach that re-interpreted the emojis by replacing them with words

added to the main tweets as part of the data-cleaning process. The importance of emojis in supervised sentiment learning social media texts has also been mentioned in Choudrie et al.'s (2021) study as important add-ons that could help classify the text. This perspective counters Kandasamy et al. (2021), who point out that emojis tend to be additives with no tied meaning to the expressed opinion and thus should be deleted when analyzing sentiment. Like Abdulkader et al. (2022), Kumar (2022) notes the popularity of emojis in social media discourses. However, the researcher insists on the need to reconvert the emojis into descriptive texts or remove them alongside other features like extra space and special characters as part of the data cleaning process. On the other hand, Vitale et al. (2021) counter the treatment of emojis as nuisances during data analysis by noting that these features play a role beyond merely labeling emotional content, as they also help improve the accuracy of sentiment analysis. These objectives are well-attained if the meanings of the emojis are known. The authors justify their views by pointing out that their researcher employed an automatic translation technique designed for the visually impaired to analyze the public's sentiments concerning COVID-19 containment measures in Italy as the Pandemic evolves. The study concluded that public sentiment was generally positive concerning the stringent COVID-19 containment measures demonstrating that the public understood the government's rationale in undertaking the measures. Emojis have been around long before the Pandemic, and thus their application in sentiment analysis has been explored in a different area.

5.4.1 *Sentiment Analysis on Social Media*

Sentiment analysis is a powerful tool that can be applied to social media analytics to understand how people feel about specific topics. It can be used to measure the overall sentiment of a conversation on social media, trends in sentiment over time, and sentiment across different platforms. By using sentiment analysis, businesses can determine which

topics resonate with their audience, identify key influencers and measure customer sentiment about their brand.

Sentiment analysis for spam detection has come a long way in recent years, thanks in large part to advances in natural language processing. With the help of powerful algorithms, it is now possible to quickly and easily identify negative and positive sentiments in a text message, tweet, or other social media post (Hussein, 2018). This is important not only for identifying malicious content but also for gauging user sentiment and determining whether a post is likely to be interesting or engaging. For example, a tweet containing a link to a malicious site might be marked as negative because the message is promoting a harmful product or service. On the other hand, a post about a positive event might be marked as positive because it is happy and original. It is worth noting, however, that sentiment analysis is not infallible. For example, a tweet about the moon landing might be marked as positive even if the sentiment is slightly tongue-in-cheek. As social media continues to grow in popularity, it is important that platforms provide users with tools to detect and avoid harmful content. Sentiment analysis is a valuable tool for accomplishing this goal, and it is expected to become even more important as social media platforms become more complex and interactive.

In one study, a Chinese microblogging platform, Sina Weibo, partnered with a firm called Emotion Analysis Platform that aimed to identify negative emotions in online content (Serrano et al., 2015). The Emotion Analysis Platform is a sophisticated tool that can analyze content posted on Sina Weibo in order to detect negative emotions such as jealousy, suspicion, and envy. The platform was developed to determine the effectiveness of ads and content and to help brands and individuals target their marketing efforts more accurately. This important innovation was a significant step forward for social media monitoring and will help to create a more comfortable and interactive online society.

Sentiment analysis on social media involves the analysis of text, images, video, and audio data (Yadav and Vishwakarma, 2020). This data can be collected from public posts, comments, and reviews, as well as private messages. Natural Language Processing (NLP) is used to extract meaning from text, while computer vision and audio processing are used to understand the sentiment in images and audio. Once the sentiment analysis is completed, the results can be used to make decisions about products, services, marketing campaigns, customer support, and more.

Sarcasm is a tool of language that can be used to express feelings, emphasize points, and even add humor to a conversation. However, it is often difficult to detect sarcasm in written communication. Liebrecht et al. (2013) look at this issue and discuss sentiment analysis as a practical solution. Sarcasm often relies on the use of nonverbal cues, such as facial expressions and tone of voice, to convey its meaning. As a result, it can be difficult to interpret when these cues are removed from the equation. Liebrecht et al. (2013) note that this can be especially problematic in the context of social media, especially when it comes to tweets, due to the brevity of the messages. The researchers note that manually reading and interpreting the messages can be time-consuming and prone to errors, which can be especially problematic when dealing with large volumes of data. As such, they propose the use of machine learning algorithms for sentiment analysis. They note that these algorithms can be used to learn from the data, interpret the messages, and identify any sarcasm.

Liebrecht et al. (2013) also explain machine learning algorithms can have varying levels of accuracy, depending on the data they are trained on. They propose a combination of different algorithms, such as natural language processing and semantic analysis, to increase the accuracy of the detection. This underscores the importance of having sufficient data in order to ensure the accuracy of the algorithms.

Twitter is a powerful communication tool that allows users to share short messages and updates with their followers. These messages can be entertaining, informative, and even funny, but they can also be used to express political views and promote causes. There is a growing interest in using Twitter sentiment analysis to obtain insights about the sentiment of messages and to determine which messages are being more favorably received. This interest is based on the belief that negative sentiment (i.e., messages that are unfavorable) can influence a user's decision-making and behavior.

Topic-based mixture modeling is a machine-learning technique that can be used to improve the performance of Twitter sentiment analysis. This method is based on the assumption that the sentiment of a message is a product of the sentiment of the topics discussed in the message. In recent years, semi-supervised training has become a popular method for training machine learning models. Semi-supervised training is a type of learning algorithm that uses partially labeled data. This means that the training data does not include information about the labels of the data objects. The semi-supervised training algorithm developed by Xiang and Zhou (2014) is an example of a method that can be used to improve the performance of Twitter sentiment analysis. The algorithm works by using a database of positive and negative tweets to train the machine-learning model. The algorithm is able to identify positive and negative sentiment by using text sentiment analysis features) and lexical features.

The findings by Xiang and Zhou (2014) showed that the algorithm was able to identify positive and negative sentiments in tweets with an accuracy rate of 79.4 percent. The algorithm also showed a high correlation with human judge ratings of sentiment. Thus, topic-based mixture modeling and semi-supervised training are useful and practical methods.

5.4.1.1 On Twitter. Sentiment analysis is a popular field of study in the social sciences, including linguistics, sociology, and psychology. It is used to identify and gauge the

attitude, opinions, and emotions of individuals or groups toward a particular topic or idea. In recent years, sentiment analysis has gained recognition as an effective tool for understanding public opinion on various topics, largely due to its ability to measure the sentiment of large amounts of data, such as text from social media platforms like Twitter. Twitter is a popular social media platform that can be used to glean public opinion on various topics. As a result, sentiment analysis for Twitter has become an increasingly popular area of research.

One technique that can be used to analyze data on Twitter is supervised classification. Supervised classification is a type of sentiment analysis that uses labeled training data to classify tweets into categories of positive and negative sentiment (Tyagi and Tripathi, 2019). This type of analysis relies on a labeled dataset that has been pre-labeled by humans and uses machine learning algorithms, such as support vector machines and random forests, to identify and categorize tweets according to sentiment. This process is often done in order to provide a sense of granularity and understanding to the user as to what topics are being discussed in a given thread, as well as to make predictions about the sentiment of a given tweet.

The use of pre-labeled datasets has been shown to be a reliable method for classifying tweets, as it reduces the number of assumptions that need to be made about the data (Stine, 2019). Additionally, the use of machine learning algorithms allows for the classification of tweets to be done in a fast and accurate way. This is important, as it allows for supervised classification to be used in a variety of applications, such as understanding the sentiment of a given conversation or predicting the outcomes of a certain event.

The sentiment of a tweet is typically determined by examining the words in the tweet and their relationships to one another. For example, a tweet that is positive may contain words that are typically associated with happiness, such as “enjoy,” “love,” and “happy.” A tweet that is negative may contain words that are typically associated with sadness, such as “sad,” “depressed,” and “heartbroken.” SVMs are a very effective supervised classification

algorithm because they can learn to generalize from data that has been labeled with positive and negative sentiments. For example, if the data has been labeled with the sentiment “happy,” SVMs will be able to predict the sentiment of future tweets that are labeled with the sentiment “happy” (Bhuta and Doshi, 2014).

Random forests are a more recent supervised classification algorithm that is also very effective. Random forests use a Monte Carlo algorithm to predict the sentiment of a tweet. This algorithm is similar to the one used by SVMs, but random forests are more flexible because they are able to learn to generalize from data that has been labeled with a variety of sentiments (Tang et al., 2015). Therefore, supervised classification analysis is an important tool that is used to analyze Twitter data. The algorithm is able to identify and categorize tweets according to sentiment, which can be useful for a variety of purposes, such as marketing and research.

Unsupervised classification, on the other hand, is a type of sentiment analysis that does not require a labeled dataset. This type of analysis relies on unstructured data, such as tweets, and uses natural language processing (NLP) algorithms, such as topic modeling, to identify and categorize tweets according to sentiment (Ortigosa, Martín, and Carro, 2014). This type of analysis can be helpful in understanding a given topic or in predicting future events. Topic modeling is an NLP algorithm that uses text data to group tweets into topics. The algorithm works by considering the co-occurrence of words in a tweet and then using that information to determine the topic of the tweet. This method is often used in marketing and social media analysis to determine the topics of interest to a given audience. By using topic modeling, analysts can quickly and easily identify the main topics of a given data set. This information can then be used to better understand the sentiment of the tweets and to make predictions about future events.

Topic modeling forms a valuable tool for all types of data to identify the topics of a data set and the sentiment of the tweets within it. Topic modeling is a particularly effective approach for analyzing unstructured data because it can automatically identify and group tweets according to their topic. This information can then be used to develop trends and patterns in the sentiment of the tweets. By understanding the sentiment of a group of tweets, it is possible to gain insights into the behavior of the population that is being analyzed.

Sentiment lexicons is a technique that uses a list of words and phrases associated with a sentiment. These lexicons are used to analyze the sentiment of a tweet by counting the number of words or phrases that are associated with positive or negative sentiment. These tools are commonly used in machine learning and artificial intelligence applications to improve the accuracy of predictions.

One of the most popular sentiment lexicons is the Word2Vec library (Bhuta and Doshi, 2014). It includes over 1,500 words and phrases that are associated with the sentiment, including happy, angry, sad, and disappointed. This library is often used in natural language processing applications, such as search engines and Chatbots, to classify, explore, predict, and analyze the tone of language, among other aspects. The Word2Vec library is incredibly accurate, but it is not the only sentiment lexicon out there. There are also sentiment dictionaries, which are collections of word and phrase pairs that are associated with a sentiment. These dictionaries are usually smaller, but they are more flexible because they can include words and phrases that are not found in the Word2Vec library.

Since the advent of neural networks, these artificial intelligence (AI) tools have been used extensively to analyze data, including tweets. Deep neural networks are used to analyze a tweet and classify it according to sentiment. Examples include convolutional neural networks and recurrent neural networks. These networks are composed of multiple layers of neurons and can be used to identify features such as words, phrases, and emotions. This

technology is still in its nascent stages but has the potential to become an even more important tool in the future. By scanning through millions of pieces of text, neural networks can automatically detect important patterns that indicate the emotions behind a tweet. This can be incredibly useful for gauging public opinion, discovering extremist groups, and even predicting terrorist actions. However, there are a few limitations to this technology. Neural networks are not perfect and may sometimes incorrectly attribute sentiments to tweets that do not actually reflect the author's true feelings. However, this remains one of the most accurate ways to analyze sentiment in tweets. So far, neural networks have proven to be a valuable tool for social media monitoring and analysis.

Supervised classification, unsupervised classification, sentiment lexicons, and neural networks are all techniques that can be used to identify and categorize tweets according to sentiment. By utilizing these techniques, researchers can gain a better understanding of public opinion on various topics on tweeter. However, each of these has its own usefulness and drawbacks and should be carefully analyzed before application for fitness.

5.4.1.2 Emoji-Based. The goal of sentiment analysis is to identify the sentiment of a piece of text. Recent advances in natural language processing (NLP) have enabled sentiment analysis to become more accurate and efficient. Traditional sentiment analysis techniques rely on manual rules and techniques, such as sentiment lexicons, to identify sentiment. However, these techniques are not always accurate or dependable.

Emojis are a pictorial representation of a facial expression textually represented using punctuation and letters or pictures and have a metacommunicative function (Guibon, Ochs, and Bellot, 2016). They express the user's mood. For example, the smiley face with happy eyes conveys happiness, the frown face with pouting lips indicates anger, and the cartoon character with a surprised face signals surprise. Emojis have become ubiquitous on social media platforms like Twitter. They are often used to convey emotion and make

communication easier. For example, when a user sends a tweet with an emoji, the recipient can interpret the message based on the emoji's corresponding facial expression. This allows users to express their emotions without having to use words (Priya, 2019). Emojis have been around for a few years, but their popularity has grown in recent years. This is likely due to the fact that emojis are easy to use and they convey a lot of information in a short amount of time.

Emojis are widely used in communication and are increasingly being used to convey sentiment. Emojis have become so ubiquitous that they are now being used in sentiment analysis, with various studies suggesting that they can be used to accurately gauge user sentiment (Li et al., 2016). One of the key advantages of using emojis in sentiment analysis is that they can convey sentiment in a more nuanced and sophisticated way than text-based sentiment analysis. For example, a single emoji can convey a range of complex emotions, such as joy, sadness, anger, surprise, and confusion. This allows sentiment analysis using emojis to be more precise and accurate than traditional text-based sentiment analysis.

Emojis also offer an advantage over text-based sentiment analysis in that they are less prone to misinterpretation. Whereas text-based sentiment analysis relies on the understanding of the language being used, emojis are universal and are understood by all users regardless of language. This makes it easier to accurately identify the sentiment being expressed by a user. Overall, it is clear that emojis are closely related to sentiments and that they can be used effectively in sentiment analysis. Emojis offer many advantages over traditional text-based sentiment analysis, such as accuracy and clarity. As such, they are becoming an increasingly popular tool for sentiment analysis.

This idea about the usefulness and applicability of emojis has been echoed by nearly all researchers who have designed to investigate their role in sentiment analysis. One of the researchers (Stine, 2019) explains as well that emojis offer a unique way to present more

ideas or concepts when being used in sentiment analysis. By including emojis in sentiment analysis, analysts can access a wider range of emotional states and reactions. This can help to improve sentiment analysis accuracy and provide a more holistic view of the sentiment surrounding a given topic. For example, by including an emoji of a happy person in a sentiment analysis of a post about a happy event, the sentiment around the event would be seen as more positive. By including an emoji of a frowning person in a sentiment analysis of a post about a sad event, the sentiment around the event would be seen as more negative. Including emojis in sentiment analysis can help to broaden the scope of sentiment analysis and provide a more holistic view of the sentiment surrounding a given topic. Emojis lend themselves well to this type of analysis because they are versatile and can be used to express a wide variety of emotions.

The provided holistic view can be applied to reflect the underlying perceptions of specific policies. Specifically, the positive sentiments observed in the works of Stine et al. (2019) show the potential of individuals to accept a certain policy when implemented. During the COVID-19 pandemic, policies relating to safety were widely met with positive sentiments. However, the sentiments were not uniform; this is due to the connection between policies and various social restrictions. An example of such sentiment was observed during the initial phases of the lockdown. The goal was to control further spread, resulting in restriction of movement. Actions related to restrictions on social interaction were predominantly met by negative sentiments, while those focusing on individual safety were met with positive sentiments. Such variations, when applied to the government decision-making process, could be vital in improving the effectiveness of various disaster-related policies.

Emoji-based sentiment analysis is a new approach to sentiment analysis that uses emoji symbols to identify sentiment from text. In one study, Lou et al. (2020) proposed an

attention-based neural network for performing emoji-based sentiment analysis. The network takes a sentence as input and uses an attention mechanism to identify the sentiment of each word. The network then uses a softmax function to map the words to a set of emoji symbols. The authors evaluated their model on two datasets: Emoji Sentiment Ranking (ESR) and EmojiNet. The experiments showed that the model was able to outperform traditional sentiment analysis methods. On average, the model achieved accuracy scores of 70.9% and 81.3% on the ESR and EmojiNet datasets, respectively.

Overall, emoji-based sentiment analysis is a promising approach to sentiment analysis that has shown promising results. The development of effective models for this task may lead to improved sentiment analysis accuracy and more efficient sentiment analysis systems, especially on social media sites such as Twitter, where emojis are a common language for conveying emotions and sentiments.

Emoji-based sentiment analysis has seen a surge in popularity in recent years, as the widespread use of emojis has made the task of gauging user sentiment more accurate than ever before. However, while this method of analysis has been popularized and discussed in the media, there is still a lack of concrete evidence in the form of scholarly literature that demonstrates the reliability of this method with absolute efficiency.

In one of the studies by Priya (2019), the authors created a corpus of tweets which were then manually classified according to the sentiment expressed in them. Then, in order to examine the accuracy of emoji-based sentiment analysis, the authors built a machine-learning classifier that used the emoji-based sentiment as its input. The results of the study showed that the classifier was able to accurately classify the sentiment of tweets with over 80% accuracy. Another study conducted by Sun et al. (2019) examined the use of emoji-based sentiment analysis in the context of facial recognition. The authors used a dataset of over 4,000 images of facial expressions to test the accuracy of emoji-based sentiment analysis.

The results of the study showed that the method was able to accurately identify the sentiment of facial expressions with 85.7% accuracy.

Overall, the evidence provided by scholarly literature suggests that emoji-based sentiment analysis is a reliable and accurate method for gauging user sentiment. However, as with all methods of sentiment analysis, there is a need for further research to better understand the complexities of this analysis and to find ways to improve its accuracy.

The rapid adoption of digital platforms has led to an increasing number of emoticons, recently known as emojis. Emojis, which are icons that represent facial expressions and other body language, are commonly used in text messages, social media posts, and other online communication. However, emoji sentiment analysis became challenging for the researchers who wanted to include it in sentiment analysis because there is no standard format for storing and representing sentiment data. This has led to various researchers using different methods, such as tokenization rather than the usual lexicon-based or NLP methods (Li et al., 2016). Ultimately, a standard format for storing and representing emoji sentiment data will help make sentiment analysis more reliable and accurate.

As mentioned, traditional methods of sentiment analysis, such as analyzing text based on the use of words and phrases, are not always reliable or accurate in applying emoji analysis. One way to improve the accuracy of emoji sentiment analysis is to tokenize the emoji data. Tokenization is a process of breaking a text or data set down into individual tokens or units of meaning (Shiha and Ayvaz, 2017). By doing this, it becomes easier to analyze and track the specific emotions and sentiments conveyed by the emoji data. For example, the token "smiley face with eyes closed" could be analyzed to determine whether it is expressing happiness, sadness, or anger. This type of analysis could help researchers to better understand the mood of a group of people and improve their predictions about future behavior.

Overall, these studies suggest that a combination of techniques, including machine learning, contextual analysis, graph-based approaches, and collaborative filtering, may be the most promising approach to successfully identifying Emojis' hidden annotations and vague implications. By considering both textual and non-textual cues and training on large datasets, these techniques may help to improve the accuracy of Emoji interpretation in natural language processing applications. The findings in previous research are largely consistent with what is established in the current study, especially with respect to the role that emojis play in predicting sentiments accurately in sentiment analysis.

5.4.2 *Efficacy of Emoji Sentiment Analysis and Prediction*

Emojis have become a common element of modern communication. In the provided datasets, there are numerous instances where emojis have been used as part of text-based data in the communication and opinion-expressing process. However, the presence of these cases does not indicate that all individuals use emojis as part of their communication within Twitter space. An examination of the provided dataset shows a significant difference between male and female users in the Twitter space. The analysis shows that female users were the predominant users of emojis on the social media platform. The number of male use increased during the pandemic; however, their use primarily included formal language and very few emotional symbols. This variation created suitable conditions for dynamic and comprehensive sentiment analysis.

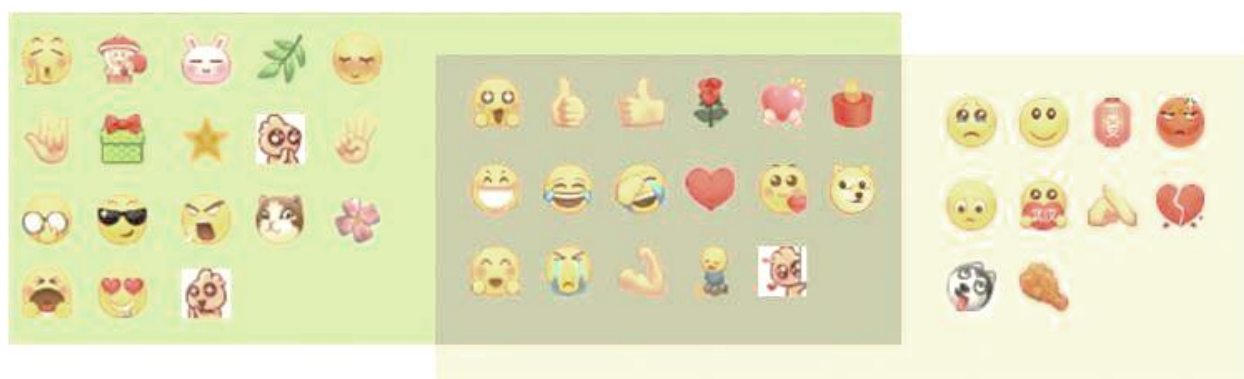
An analysis of initial models focused on text-based data in various languages and regions. While these were key components of the provided tweets, the underlying datasets were relatively more diverse than originally adopted. The presence of emojis best highlighted this diversity. According to Liu et al. (2022), emojis in modern communication occurs at a specific frequency which is critical in various interaction. Their presence usually conforms to Pareto's and Zipf's laws. Amidst conforming to these laws, they tend to create a sense of

heterogeneity within a given social platform, today more than 100 emojis are used in at least 96% of daily communication, with the top 10 of these emojis appearing in more than 50% of online interaction. This implies that the communication before and after the COVID-19 outbreak varied significantly. An analysis of the tweets prior to the pandemic shows that 154 unique emojis were used in communication across the platform. After the pandemic, more emojis were added. This raised the need to further explore the nature of emojis in relation to human interaction.

The provided datasets show that emojis tend to exhibit or convey different emotions. These new emojis were predominantly negative, which included key elements such as anger, sadness, and disappointment. However, they were not the only type of emojis introduced into the platform. Evidence shows the presence of a few positive emojis, such as the *peach lantern* and *cheers* for Wuhan. Each emoji was considered symbolic, with a specific message to different groups; for instance, the *peach lantern* and *cheer* for Wuhan emojis were created to express support for medical staff and show support for the anti-pandemic work. At the same time, the emojis were used to reflect the situation at hand. An instance of such reflection was the use of food-related emojis like the *chicken leg* when discussing the lives of quarantined communities.

Figure 29

Comparison of Emojis before and After the Pandemic (Liu et al., 2022)



An examination of communication patterns before, during, and after the pandemic show variation in the use of emojis within the digital space (Figure 32). The changes in the use of emojis spaces created the need to decode messages for effective communication. In this case, the model focused on classifying emojis into five major categories; these categories included *angry*, *sad*, *encouraging*, *happy*, and *questioning*. The classification allowed for the study to fully evaluate their use in communicating various sentiments throughout the pandemic. The analysis shows that there was a rapid decline in the use of happy-related emojis in China in January 2020. This development was accompanied by the introduction of mandatory quarantine and national lockdowns throughout the globe (Liu et al., 2022). This decline further explains the observed rise in negative sentiment in China. Furthermore, the rise in negative sentiments across the digital space in China further rose with its peak reaching February—the decline in the use of happy emojis also explains the increase in negative sentiments in regions of the United States, India, and Italy. A similar trend has been reflected in other countries, such as the United States and India, where a decline in the use of happy emojis has been accompanied by the rise of negative sentiments. Nonetheless, this variation cannot be fully explained by the decline in the usage of a single category of emojis.

5.5 Q4: Promising Technique to Identify Emojis' Annotations

There exist multiple machine-learning approaches useful in sentiment prediction and forecasting, with the most robust ones varying depending on the dataset, the features used, and the specific problem being addressed, and this has thus far been demonstrated in this study. Based on a review of multiple studies, as well as in light of what the current study has established, there are some of the most effective approaches for sentiment prediction and forecasting.

Sentiment analysis is the process of extracting subjective information from a text and using it to determine the attitude, opinions, or feelings of a speaker. Zhou (2016) agrees that

sentiment analysis is an essential area of machine learning that examines how natural language is used to express feelings, opinions, and emotions. In recent years, sentiment analysis has become increasingly popular due to its ability to detect and analyze sentiment in large volumes of structured and unstructured data. Sentiment analysis techniques using machine learning are divided into traditional and modern methods. Traditional machine learning techniques for sentiment analysis are used to classify documents according to their polarity (positive/negative) or subjectivity (objective/subjective). These traditional techniques can be further broadly categorized into supervised and unsupervised methods.

Supervised methods involve training data sets with labeled positive and negative classes which are then used by algorithms such as Naïve Bayes classifier, Support Vector Machines (SVM), Maximum Entropy Modeling (MaxEnt), Decision Trees, and others in order to learn how different features, affect the sentiment of a given document (Alpaydin, 2016).

Unsupervised approaches make use of clustering algorithms such as K-means Clustering in order to identify patterns within unlabeled data without relying on predetermined categories; this allows for more accurate detection than traditional classification models since it does not rely on predetermined labels.

Both supervised and unsupervised approaches have been widely used in sentiment analysis tasks due to their ability accurately detect sentiments expressed within documents while taking into account various nuances that may influence an individual's opinion about a particular topic; however, they each have their own strengths and weaknesses which must be taken into consideration when choosing between them depending upon one's specific application requirements.

5.5.1 *Unsupervised*

Unsupervised sentiment analysis techniques are a type of automated text processing that uses natural language processing (NLP) to identify, extract, and classify sentiment from

unstructured data sources such as social media posts or reviews. These techniques can be used to gain insights into the attitudes of customers towards products or services in order to inform decision-making processes within businesses. In the current study, techniques such as LSTM, Naïve Bayes, and Neural Networks were used to make predictions on the various sentiments on data extracted from Twitter.

Bibi (2022) focused on concept-based sentiment analysis (CBSA) methods within the field of natural language processing. CSBA methods take into account the underlying semantic meaning of the text in order to accurately identify sentiment labels. However, because the knowledge base utilized in CSBA methods is often limited, supervised learning techniques may be required to classify the entirety of the text. According to the researchers, the manual task of labeling large datasets could be tedious and time-consuming, so the article proposed an unsupervised learning mechanism for Twitter sentiment analysis. The proposed framework uses concept-based and hierarchical hybrid clustering methods along with two feature representation methods (Boolean and TF-IDF), and popular classifiers such as Naïve Bayes and Neural Network experimented for a balanced comparison. The accuracy measure was employed to assess the efficacy of lesser-explored techniques. The findings demonstrated that unsupervised learning techniques could perform as well as supervised learning techniques in sentiment analysis.

The most common unsupervised sentiment analysis technique is a lexicon-based classification which involves mapping words with predetermined positive or negative values based on their semantic meaning (Athey, 2018). This method relies heavily on dictionaries containing pre-defined keywords associated with certain sentiments; however, it cannot detect subtle nuances between different types of sentiments like sarcasm, irony, etc., making it less accurate than supervised methods, which use labeled training data sets for more fine-grained results.

Other popular unsupervised approaches include rule-based systems and topic modeling algorithms such as Latent Dirichlet Allocation (LDA). Rule-based systems rely on manually defined rules about how words should be classified, while LDA attempts to group documents by topics using statistical inference models trained over large datasets. However, both offer higher accuracy rates than lexicon-based methods but require more resources due to their complexity in implementation. Applications range from predicting stock market movements through analyzing news articles, detecting frauds in financial transactions, monitoring customer service interactions, identifying fake product reviews online, and others (Zhou, 2016).

In order to properly examine the use of unsupervised ML methods, they will be subdivided into three general areas of clustering, rule-based, and visualization & dimensionality reduction. *Clustering* is a powerful tool used in sentiment analysis to better understand customer feedback and opinions. Hierarchical cluster analysis (HCA) and k-means are two popular clustering algorithms that can be applied to sentiment analysis tasks (Burkov, 2019). Hierarchical cluster analysis works by creating clusters of similar items based on their proximity or similarity, with each item being assigned to the most suitable group. This method is especially useful for discovering patterns within large datasets as it allows for the efficient sorting of data points into meaningful groups without any prior knowledge about the dataset itself (Bell, 2022). Additionally, HCA does not require a predefined number of clusters; instead, it will automatically generate them depending on how close or distant different items are from one another in terms of their attributes, such as opinion scores or keywords associated with them.

K-means clustering works differently from HCA by requiring users to set an initial number K value which determines how many clusters should be generated before running the algorithm itself over all data points provided; this means that k-means requires more manual

work compared to hierarchical methods but at times may yield more accurate results due its ability tailor specific numbers of distinct groups according to user input parameters like the maximum distance between members inside the same cluster. In addition, unlike HCA, which uses dissimilarity measures between objects when forming clusters, k-means relies solely on distances calculated using Euclidean metric, making it easier interpretable than other techniques used in sentiment analyses such as natural language processing (NLP) (El Naqa and Murphy, 2015).

Based on overall use across literature, both hierarchical cluster analyses and K-Means of clustering unsupervised ML techniques offer unique advantages when applied to sentiment analysis tasks allowing researchers to gain valuable insights into customer preferences. At the same time, they offer the advantage of automating tedious parts related to grouping multiple pieces of information together quickly via machine learning algorithms, thus saving time and resources needed manually to process vast amounts of collected feedback data sets routinely encountered by today's businesses and social environments.

Association rule learning is another type of machine learning algorithm that is used to find complex relationships and correlations between data elements (Carleo et al., 2019). This type of learning has proven to be effective in a variety of tasks, such as marketing, financial forecasting, and sentiment analysis. In sentiment analysis, association rule learning is used to extract knowledge from large datasets and find relationships between different words, phrases, and sentiments. The goal is to identify patterns between words and sentiments that can be used to classify positive and negative sentiments in text documents. For example, a rule may be identified to determine that words such as “happy” and “delighted” are strongly associated with a positive sentiment. This type of rule is then used to classify sentiment automatically in new documents.

The association rule learning approach has become increasingly popular in sentiment analysis due to its ability to identify complex relationships between data elements. This approach is also useful because it can be used on both structured and unstructured data. Furthermore, it is relatively easy to implement and can be adapted to a variety of tasks. Typically, the association rule learning process starts by collecting a large dataset of labeled or unlabeled documents. This dataset is then analyzed to find relationships among words and sentiments. Machine learning algorithms are then used to look for patterns in the data and generate rules that can then be used to classify documents automatically. It is important to note, however, that feature selection plays an important role in this process (Mahesh, 2020). Feature selection is used to identify the type of features that are most relevant to the task and make the resulting rules more reliable.

Examples of specific techniques applied to association rule learning include Eclat and Apriori. Eclat and Apriori are two popular algorithms used in association rule mining (association rule learning). Association rule mining is a data mining technique used to uncover interesting relationships between variables in large datasets. Eclat (Equivalence Class Clustering and bottom-up Lattice Traversal) is an algorithm that uses a depth-first search approach to generate frequent item sets. It works by first generating a set of candidate items, then using a depth-first search to find the frequent item sets (Jordan and Mitchell, 2015). The algorithm is efficient in terms of time complexity, but it can generate a large number of candidate itemsets, which can lead to a large number of comparisons.

Apriori, on the other hand, is an algorithm used for mining frequent item sets. It works by first generating a set of candidate item sets, then using a bottom-up approach to find the frequent item sets. The algorithm is efficient in terms of memory usage, but it can generate a large number of candidate itemsets as well. This may have the advantage of leading to a large number of comparisons. The Apriori algorithm is based on the assumption that if an item set

is frequent, then all of its subsets must also be frequent. Both algorithms are used to identify relationships between items in large datasets. They are used to generate frequent item sets, which can then be used to generate association rules. These association rules can then be used to make predictions about future data.

Visualization and dimensionality reduction methods are used in sentiment analysis to reduce the complexity of data by reducing the number of variables while still preserving the important information contained in the data. Kernel PCA, t-distributed, and PCA is among the most commonly used methods in sentiment analysis. Kernel PCA is a non-linear dimensionality reduction technique that is used to reduce the dimensionality of data while preserving the important information contained in the data (Berry et al. 2019). It is used to reduce the data to a lower dimensional space while preserving the structure of the data. The technique works by mapping the data points to a higher dimensional space and then applying a kernel function to the data points. This allows the data points to be separated into distinct clusters.

T-distributed and PCA are other linear dimensionality reduction techniques that are used to reduce the dimensionality of data while preserving the important information contained in the data. T-distributed and PCA are used to reduce the data to a lower dimensional space while preserving the structure of the data (Jimenez et al., 2016). The technique works by projecting the data points onto a lower dimensional space and then applying a linear transformation to the data points. This allows the data points to be separated into distinct clusters. These methods are used in sentiment analysis to reduce the complexity of data while still preserving the important information contained in the data. By reducing the dimensionality of the data, sentiment analysis can be more accurate and efficient.

Lexicon-based methods are a type of sentiment analysis that uses a predefined set of words, or lexicon, to identify sentiment in text. This type of sentiment analysis is based on

the idea that certain words are associated with certain emotions or sentiments. For example, words like “happy” and “joyful” are typically associated with positive sentiment, while words like “sad” and “depressed” are typically associated with negative sentiment.

Lexicon-based methods are used to identify sentiment in the text by counting the number of words associated with positive or negative sentiment. This type of sentiment analysis is relatively simple and straightforward and can be used to quickly identify the overall sentiment of a text. However, lexicon-based methods are limited in their ability to accurately identify sentiment in the text (Zhou et al., 2017). This is because they rely on a predefined set of words, which may not accurately reflect the sentiment of a text. Furthermore, lexicon-based methods are unable to identify the context in which words are used, which can have a significant impact on the sentiment of a text. For example, the word “happy” may have a positive sentiment when used in one context but a negative sentiment when used in another context. One could conclude that lexicon-based methods are a useful tool for quickly identifying the overall sentiment of a text, but they are limited in their ability to accurately identify sentiment in text.

Unsupervised sentiment analysis, as examined, provides businesses with valuable insights into customers’ behaviors and preferences without requiring intensive manual labor. However, its accuracy depends largely upon the quality of the inputted dataset, so careful consideration must be taken when deciding which approach best suits specific business needs before implementing any solution.

5.5.2 Supervised ML

Supervised machine learning is an area of Artificial Intelligence and machine learning that has rapidly grown in popularity and importance in recent years due to its ability to produce more accurate and reliable results. This section will explore the various techniques and methods used in supervised machine learning, their usage and application, and the

benefits and implications of this technology. The discussion centers on how supervised machine learning can be used to improve the accuracy of predictions and provide insights into complex data.

Supervised machine learning techniques are used to help automate the process of extracting useful information from data. This is done by using the data to train a model, which is then used to make predictions about unseen data. According to (Denny and Spirling, 2018), supervised machine-learning techniques are used to build models from labeled data where the output is already known. Examples of supervised machine learning techniques include classification and regression, which can be used to classify data into different categories or predict continuous values, respectively (Wang and Gupta, 2015). These techniques are used to build effective models that can be used to accurately predict the outcomes of given data. Furthermore, supervised machine learning techniques can also be used to analyze and interpret complex data sets and identify patterns and trends. In conclusion, supervised machine learning techniques are a powerful tool for extracting useful information from data and can be used to build models that can accurately predict outcomes.

Supervised machine learning has been adopted for many applications due to its ability to learn from data and make predictions. Srivastava, Mansimov, and Salakhudinov (2015) established that supervised machine learning had been used for various tasks such as speech recognition, facial recognition, natural language processing, and medical diagnosis. These applications are able to leverage the power of supervised machine learning to identify patterns in data and make accurate predictions. Additionally, supervised learning methods are commonly used for predictive analysis, where a model is trained on a set of data to predict outcomes based on new input data (Celebi et al., 2016). This can be used to predict customer behavior or to forecast future trends. Furthermore, supervised machine learning is also used for image recognition, which can be used for applications such as autonomous driving and

object detection. In conclusion, supervised machine learning is a powerful tool that can be used for a wide range of applications, from prediction to image recognition.

Supervised ML has become a powerful tool for extracting knowledge from remote sensing data. As such, its benefits and implications should be explored. According to (Wang, 2016), the benefits of supervised ML include its ability to accurately identify objects and features in a scene, its ability to detect changes in land cover, and its ability to provide precise classification of land cover. All of these benefits are made possible through the use of ML algorithms that are trained on labeled data. The implications of supervised ML go beyond simply providing more accurate object and land cover classification. It can also provide valuable insights into environmental changes, provide a better understanding of land cover types and improve land use planning. Moreover, supervised ML can be used to monitor and assess the health of ecosystems, as well as to detect and predict changes in the environment. As such, supervised machine learning offers a range of benefits and implications for remote sensing applications.

5.5.3 *Convolutional Neural Network (CNN)*

According to Ouyang (2015), sentiment analysis is integral for most tasks involving natural language processing, particularly in the era of big data from social platforms. Deep learning models have shown promise in improving sentiment analysis performance. The researchers proposed a framework named Word2Vec + Convolutional Neural Network (CNN) in this study. This framework employed Word2Vec to generate vector representations of words, which were then used as input to the CNN. This initialization step helped to achieve better model performance. The CNN architecture consisted of three convolutional and pooling layers and utilized various techniques such as PReLU, normalization, and dropout to enhance both the accuracy and generalizability of the model. The framework was

tested on a public dataset of movie review excerpts with five labels, achieving a test accuracy of 45.4% which outperformed some other neural network models.

Zhang (2018) conducted a study to establish the most robust machine-learning algorithms for sentiment analysis sidelining both the strengths and weaknesses of each approach. The researchers were keen to highlight the importance of sentiment analysis in natural language processing. Similarly, they did not shy off from noting the main challenges of sentiment analysis, such as the presence of sarcasm, irony, and ambiguity in language. The researchers reviewed various machine learning algorithms, like convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid architectures that combine both CNNs and RNNs. The authors, after careful scrutiny of these deep learning architectures, came to the conclusion that the most robust approach presented for sentiment analysis using machine learning architectures was a hybrid combination of RNNs and CNNs. These researchers noted that this approach was successful in capturing both short and long-term dependencies inside the text and is effective in dealing with noisy and ambiguous language. They also noted that the use of pre-trained word embedding and attention mechanisms could further improve the performance of these architectures. However, the authors concluded by noting that more research is required to address the challenges of sentiment analysis, including the need for large amounts of tagged data and the toughness of manipulating rare and out-of-vocabulary words.

5.5.4 *ULMFiT Sentiment Analysis*

Howard (2018) described an approach that involved optimizing an already trained language algorithm on a target sentiment analysis task with an agenda of achieving state-of-the-art performance in various benchmark datasets. The researchers specifically used a Universal Language Model (ULMFiT) that was already trained on a vast corpus of text and then fine-tuned on a smaller, labeled dataset for a specific task in sentiment analysis. The

authors of this study introduced a novel fine-tuning process for pre-trained language models, which entailed gradually unfreezing layers of the model and then training them on the specific sentiment analysis task. This approach involved both discriminative fine-tuning and gradual unfreezing and was designed to optimize the performance of the model on the target task. This technique allows the model to learn task-specific features while retaining general language understanding from pre-training. Additionally, the authors introduced a novel approach to handling out-of-vocabulary words by using a combination of sub-word units and pre-trained embedding. The results of their experiments showed that optimizing a pre-trained language model using their proposed technique has a notable effect on performance compared to traditional approaches for sentiment analysis, achieving state-of-art results on multiple benchmark datasets. Overall, their approach demonstrated the advantages of leveraging pre-trained models for transfer learning in sentiment analysis.

As the need to understand the public's sentiments concerning major events and incidents increase, scholars have attempted to utilize various sentiment analytical technique. But only some of these techniques meet the required accuracy due to the complexity and dynamism of world languages that have made it difficult to understand texts. For this reason, most scholars are willing to explore alternative sentiment analysis techniques, such as emojis. The logic behind this shift is that emojis could express emotion that might not be captured in words. Thus, emojis have often been favored as accurate and direct expressions of human emotions that can provide crucial insights into the users' sentiments. Furthermore, emojis have gained popularity across all platforms and cultures, creating a sense of convergence of emotional expression (Chuchu et al., 2021). Similarly, Bai et al. (2019) note that emojis are add-ons that make up for the paralinguistic features in computer-mediated communication. Since their introduction in 1999, emojis have gained significant popularity across cultures so are their uses. Users use emojis to simplify conversations and make them convenient and conducive.

Despite having universal functions, emojis have diverse usage that differs at individual, cultural, and platform levels. Emojis improve online conversations making them immersive (Zou & Xiang, 2022).

Given emojis' ubiquitous nature, researchers favor sentiment analysis algorithms that combine emoji and sentiment analysis models. But this objective has yet to be achieved as most of the employed semantic analysis techniques tend to treat emotional emojis as noisy labels while others use emojis as classifiers. Both models have recorded poor performance due to excessive noise in the labels. At the same time, researchers like Chuchu et al. (2021) noted the complexity of interpreting Chinese texts due to the dynamism and diversity of meanings. For this reason, the authors suggested a technique referred to as the Chinese emoji-embedding LSTM model (CEmo-LSTM model) in studying the effect of emojis on sentiment analysis. The analysts use the bidirectional long short-term memory (Bi-LSTM) principle to introduce emoji usage in social media posts. Since the model is designed to study emojis as part of sentiment analysis in text, plain texts and those embedded with emojis make up the model's input fed in a trained corpus marked and filtered with the various emojis usages. The researchers minimize the data annotation workload by employing annotators that denote the emoji usage of every post during the labeling polarity phase of the experiment. Chuchu et al. (2021) CEmo-LSTM model's unique attribute is that emojis are not only used as one of the features but also all the training data is categorized according to their usage before being trained as classifiers. The researchers concluded that the model achieved the highest accuracy in sentiment analysis. When this model was applied in analyzing the sentiments of Wuhan, China residents' sentiments towards COVID-19, the researchers concluded that the pandemic led to passive emotions.

The Bi-LSTM principle has also been used by (Wang et al., 2020), who combined it with a multi-head attention mechanism to determine whether the accuracy of microblog

sentiment analysis could be improved. Integrating Bi-LSTM and multi-head attention mechanisms results in a model referred to as EBILSTM-MH used for sentiment analysis. Within the EBILSTM-MH model, the Bi-LSTM is used to learn the feature representation of the input texts given in the embedded words. At the same time, the attention mechanism assigns attentive weights to every word in the sentiment analysis exercise based on the effects of employed emojis. These attentive weights could combine with the hidden layer output to obtain the feature representations of the analyzed posts. Lastly, the researcher can obtain sentiment polarity through a dense connection layer. Even though the model proved feasible, the researchers needed to demonstrate its performance in analyzing the hidden annotations of emojis. Like Wang et al. (2020), earlier studies by Liu et al. (2012) proposed the combination of manually and noisy labeled data to train classifiers. As mentioned before, most sentiment analysis models treat emojis as a nuisance that should be cleaned before the data is analyzed for semantic analysis. Alternatively, some models favor replacing individual emojis with a single word, a technique that could be disadvantageous given the subjectivity and diversity of emoji usage. As such, Liu et al. (2012) suggest that noisy and manually labeled data should be combined for training to compensate for each model's weakness. If an investigator favors manually labeled data, the individual will face significant difficulty stemming from the limited size of the supervised model because the process is highly labor and time-consuming. On the other hand, noisy labels offer the advantage of having extensive data for the supervised training model. However, the noise in the labels significantly minimizes their performance. But the challenge of combining the two lies in how they should be integrated seamlessly into a common learning framework. (Chen et al., 2018) suggests that researchers interested in sentiment analysis on content characterized by sentences/words and emojis could use the bi-sense emoji embedding that uses short and long-term LSTM model. Using

the bi-sense emoji, the investigators could extract sentiment-aware embeddings of emojis to uncover the diverse meanings of emojis.

An alternative to the LSTM sentiment analysis approach is the convoluted neural network classification model (CNN). The main difference between the two models is that whereas the former uses a process that makes predictions in a data sequence, CNN can study spatial correlation in data and works well in image and speech. (Zou & Xiang, 2022). The CNN model is proposed by modifying it to analyze texts composed of words and emojis. The researchers recognize the popularity of short texts, which fail to convey sufficient information required for sentiment analysis despite their convenience and ease of use. These short texts are often accompanied by emojis that supply meaning beyond the accompanying words. Therefore, Zou & Xiang (2022) suggest a technique that blends short-text and emojis content which is then transformed into vectors that are in turn connected to the sentencing matrix. The sentence matrix I used as an input into a CNN classification model. Based on the results of this approach, the authors conclude that it recorded higher accuracy levels than other existing methods. The efficiency of CNN in sentiment analysis is well-supported by (Kumar & Vardhan, 2022), who notes that the uniqueness of the technique lies in its two-word embeddings and the classifier using frameworks like Kera and TensorFlow for training and classification purposes. Such a classifier has been proven effective in analyzing the public's sentiment on products, but the application still needs to be expanded to a statistic dataset. Additionally, through pooling, CNN could allow researchers to uncover important public sentiment details. But in most cases, this objective is rarely achieved due to challenges in obtaining contextual information. The performance of CNN could be improved by increasing the amount of neural data. This point is also supported by Weytjens & Weerd (2021), who point out that the predictive capacity of CNN is enhanced by increasing the dataset sizes. Additionally, changes in the data architecture by adding dropout layers

(Svensson, 2017). Social media platforms like Twitter use more emojis compared to others. This feature makes CNN the most effective sentiment analysis technique for tweets, as emojis can be classified with high accuracy, a feature that is further improved by training a larger dataset (Eslami, 2020).

Weytjens & Weerdt (2021) Compare CNN and LSTM and conclude that the two models record the same performance, disapproving like Wang et al. (2020), who argue that the attention mechanism adds significant value to LSTM. This view is countered by Weytjens & Weerdt (2021), who state that the attention mechanism does not give LSTM any additional features that can improve performance. In essence, both models have proven robust in terms of hyperparameters designed to attain maximum predictive power. Some researchers have bypassed this contradiction by adopting models that combine the CNN and LSTM models. One such researcher (Yang et al., 2022) proposes using a model called ALBERT-FAET. The investigators obtained text embedding through the ALBERT pre-training model and learned the inter-emoji through LSTM attention-based mechanism. Another feature added to the ALBERT_FAET model is a fine-grained attention mechanism designed to capture the relationship between emoji and plain text. The two features were combined and fed into the CNN classifier, which predicted the sentiment labels of the analyzed text. Yang et al. (2022) conducted comparison experiments to determine the precision, accuracy, and recall of the ALBERT-FAET model. They concluded that the model recorded improved performance which could be attributed to the fine-grained attention network. In essence, the investigators argued that the fine-grained attention network understands the diversified usage of emotions and thus could uncover the annoyed and hidden meanings often overgeneralized in other models. Alternatively, investigators could opt to go beyond Yang et al.'s (2022) ALBERT-FAET model by opting for the Emoji2vec model in calculating polarity scores. This mode combines features such as; CBOW, SG, CNN, LSTM, Bi-LSTM, and CNN-LSTM. CBOW

and SG are for feature extraction, while CNN, LSTM, Bi-LSTM, and CNN-LSTM were developed alongside their hyperparameters. The researchers concluded that CNN-LSTM demonstrated better performance, justifying Yang et al. (2022) findings. (Eslami, 2020) also suggests a model that merges CNN and Bi-LSTM with emoji-sense vector and embedded texts as input to predict the positivity and negativity of a comment or opinion.

5.5.5 Technique for Identifying Hidden Messages

The continuous advancement of technology is likely to advance sentiment analysis across all aspects of human life. According to Guibon et al. (2016), these aspects are designed to improve different aspects of human life. Based on this recognition, Guibon et al. (2016) aimed to examine the evolution of emojis to sentiment analysis. The analysis showed that understanding emojis and their uses is the key to applying them in sentiment analysis. This perspective is further presented by the need to recognize two types of emoticons used across different social media platforms: western and Eastern emoticons. Each group of emoticons tends to have a variety of representations that limit their overall interpretation. At the same time, it is critical to acknowledge that the underlying subject of interaction tends to influence their overall usage. Such aspects create a subjective element for the interpretation of emojis, especially when performing sentiment analysis (Guibon et al., 2016). Nonetheless, key methods such as keyword spotting and machine learning approaches are continuously advancing as critical aspects of sentiment analysis, especially where emojis are integrated into the natural language process.

The observations made by Guibon et al. (2016) indicate the polarity of emojis and underlying approaches for interpretation during sentiment analysis. The insights, while crucial to understanding the communication perspective of emojis and their uses, do not fully account for hidden messages within different platforms. This element has been accounted for by Iranmanesh et al. (2015), where the focus is shifted to analyzing the use of emoticons and

Lingoes for data handling during communication. In this case, Iranmanesh et al. (2015) recognize that emoticons have even been widely used as part of stenographic approaches, especially when hiding data in short service messages. This aspect enables individuals to build different networks. One advantage of this approach to hiding data is that it is less conspicuous due to its focus on short-messaging services (SMS) (Iranmanesh et al., 2015). While such insights are critical, it is critical to acknowledge that they do not account for hidden data within social media platforms. Nonetheless, it presents a unique gap that needs to fully explore when developing various sentiment analysis algorithms for different applications.

Taleby Ahvanooey et al. (2019) acknowledge the evolution of technology beyond short messaging services. According to Taleby Ahvanooey et al. (2019), data encryption has advanced beyond the use of the least significant bit, as shown by the migration to highly intelligent programming. An analysis of the modern method shows the increased use of secret watermarks to cover messages and files to protect information that individuals consider confidential. This implies that simple characteristics such as emojis may be embedded in the messages to be read by a specific audience. Such characteristics of modern data encryption allow messages to be passed through public platforms such as social networks (Taleby Ahvanooey et al., 2019). As a result, linguistic techniques and semantic and syntax analysis can be adopted as key approaches to detecting and decrypting such messages. These approaches, while critical, cannot be adopted haphazardly and in isolation; rather, intelligent programming approaches such as machine learning models can be adopted for efficient analysis and effective interpretation of the hidden messages (Taleby Ahvanooey et al., 2019).

Sentiment Analysis is defined as the extracting and quantifying of subjective information that includes the status of opinions, emoticons, and attitudes from different contents such as audio, images, and texts (Yuan et al., 2015). The process has attracted great

attention as a result of its wide applications in psychology and sociology, political science, and in business and government intelligence (Aldahawi, 2015; Elghazaly et al., 2016). From a technical point of view, the textual process is first investigated by researchers as an NLP (Natural Language Processing) task. According to Baccianella et al. (2010), the methods range from lexical-based techniques using features such as the keyboard where a word corresponds with a sentiment vector and entries representing the possibility of the word; phase-level features. As stated by Yuan et al. (2013), it was not until recently that scholars began focusing on multimodal sentiments and images and assessing how to take advantage of the resources in cross-modality (You et al., 2016). For multimodal sentiment analysis, an underlying assumption is that both of the modalities are an expression of similar sentiment, and such similarity is enforced to train a robust sentiment inference model. However, Felbo et al. (2017) claimed that a similar assumption does not stand in a modeling textual tweet since the complexities of natural language are in extensive existence, such as the use of sarcasm, jokes, and irony.

Although sentiment analysis recorded a recent surge in interest since its coining by Nasukawa and Yi (2003), the demand for data on opinion and sentiment in situations that require decision-making dates back to before the widespread use of the World Wide Web. Opinions are considered central to almost all activities involving humans since they are likely to influence their behaviors, especially when engaging in decision-making activities. For instance, there are numerous individuals that have asked their friends to recommend a dishwasher or to offer an explanation for whom they might vote for during the election period; some even request reference letters for job applications from their colleagues. During this era, Pang and Lee (2008) pointed out that the experiences and opinions of numerous individuals that are neither professional critics nor acquaintances are readily available, all thanks to the Web and the Internet. This, however, does not remain limited to individuals

only; the concept is also applied to companies and groups that display an eagerness to discover and interpret the opinions of their customers on their services and products.

When a business wanted to acquire the opinions of its consumers in the past, it carried out opinion polls and surveys. Currently, one is less limited to asking family and friends or carrying out surveys for opinions on their services and products; instead, a company uses volumes of user discussions and reviews in a public forum from the Web (Liu, 2012). Indeed, the internet has dramatically transformed the manner in which individuals express their opinions in social events, individuals, firms, services, and their products. Now, there are micro-blogs, blogs, discussion groups, and internet forums well suited for the user to freely make public their reviews on products and equally air their views on every activity performed online. Content that is user generated and exhibits word-of-mouth behavior are the common sources of data with numerous practical and immediate applications. Research on sentiment analysis is seen to have appeared earlier than 2003 according to the following researchers; Das and Chen (2001); Tong (2001); Wiebe (2000), while there some other researchers highlighted sentiment analysis even earlier on belief as later work or frontiers (Hearst (1992); Wiebe and Rapaport (1988) on interpreting viewpoints, affects, subjectivity, sentiment adjectives, metaphors and other related areas (Liu, 2012).

Contrasting the long history of NLP and linguistics, the area of assessing the sentiments and opinions of individuals has virtually been untraversable prior to 2000. However, after that, literature has witnessed numerous studies such as Kouloumpis et al. (2011) and Saif et al. (2012) as a result of different factors such as; 1) accessing datasets for the training of machine learning techniques prior to the introduction of the World Wide Web and in specific, the review-aggregation websites. 2). The rise of machine learning techniques in retrieved information and NLP and, lastly, realize a large number of applications in the industries that the area began offering (Pang and Lee, 2008). According to Liu (2012), the rapid growth of

the process (sentiment analysis) and, more significantly, the coincidence with the explosive popularity of social media platforms have transformed sentiment analysis into the chief point in research associated with social media platforms.

5.5.6 *Technique to Identify Emojis' Hidden Meaning*

The observed discrepancies in the correlation between emoji type and sentiments resulted in inaccuracies in prediction. Despite the inaccuracy, this does not necessarily imply poor model performance due to the integration of emoji in the analyses; rather, it highlights the failure to account for the appropriateness of the model during analysis. Therefore, it was critical to examine the variations using a maximum occurrence frequency model. The model was designed to focus on building a network of emojis using per nodes with the largest probability. During this analytical process, the edge of the node represented the occurrence probability between nodes hence resulting in 4 hub nodes in the network. These nodes were best highlighted by the four major emojis (, ,  and ) — these emoticons (hub nodes) represented a wide range of emotions that were paired within smaller sub-clusters. The maximum occurrence analyses show that after the COVID-19 outbreak, there were significant changes in the use of emoticons. The most notable change was the use of the longing emoticon (). The angry emoticon () was accompanied was connected to other gestures relating to sadness and anger. Additionally, different sub-clusters were included in the live node, while online interaction was a common theme during the pandemic. Therefore, the variance in emotion could be explained by changes observed in the maximum occurrence network, as demonstrated in this study.

While the concurrence may reflect the variation in sentiments, it was vital to acknowledge that emoticons were primarily used within a specific context. This impels a significant relationship between the use of an emoji and the post topics. This recurred, extracting all posts containing different hashtags or topics to demine the patterns in the

emojis. An analysis of tweets shows that the topic could be categorized into different clusters, including Celebrity news, government reports, social events, and online games, as well as COVID-19. These topics generated more than 2 million tweets during the pandemic (Liu et al., 2021). The analysis shows that the use of different emoticons was closely related to different posts. For instance, tweets relating to the death toll and the number of infections were often accompanied by sad-related emojis. Topics on government actions that proposed further restrictions were accompanied by angry emojis. Through this correlation, hidden messages during the pandemic could be easily unveiled—most of these messages expressed individual emotions and attitudes towards certain actions and overall attitudes towards their living situation.

As expression varied through the timeline, it is critical to acknowledge that the data accounted for the variation in sentiments after the pandemic. An analysis of the data shows that most tweet often focuses on the opportunities created by freedom. This variation saw a significant decline in negative sentiments across different regions. However, neutral and positive sentiments significantly increased. However, there were key sentiments related to the social distancing rules and guidelines on interpersonal interaction. This raised the question of the correlation between the changes in sentiment and the use of emojis during such changes. A correlational analysis across topics and the usability of various emojis indicate a notable rise in the use of happiness-related emojis. Therefore a few cases, the use of flowers as it gradually replaced the longing emoji. At the same time, the public was more focused on utilizing different opportunities hence a gradual decline in the use of anger-related emoticons. Such correlation indicated that emojis are likely to improve the performance of sentiment analysis models.

The likelihood of emojis to improve sentiment analysis was explored during the analytical phase of the project. In this case, the concept of rule-based sentiment primarily

stems from defining the underlying sentiment lexicon. Similar to the already established model, the approach classifies emojis as either positive, negative, or neutral (Figure 33). According to Liu et al. (2021), the sentiment lexicon requires the identification and definition of specific dictionaries that depict individual emotions and potential intentions. Based on this description, each word posted on the internet by individual tend to have strong ties to their emotional state and purpose in relation to interaction. Where emojis are concerned, their lexicon tends to depict a sense of heterogeneity. Therefore, each emoji tend to have a unique sentiment value that provides more insights into the observed social interaction in the digital space.

Figure 30

Classification of Emojis based on their Sentiments (Liu et al., 2021)



The definition of the underlying emoji lexicon created room for a new approach toward sentiment analysis. As a result, it was critical to examine its overall capabilities to improve the analytical model. An analysis of the adopted classification model shows an accuracy of 41.23% and 88.44% on the logistic regression and LSTM models, respectively. Where the multinomial Naïve Bayes Model was concerned, an accuracy of less than 45% was achieved (Table 13) without the inclusion of emojis. The introduction of emojis in the sentiment analysis model resulted in numerous changes. These changes were primarily reflected in the improvements in the accuracy model. A summary of improvements in the accuracy model is

highlighted in Table 13 below. The most notable improvement was on the KNN, Multinomial Naïve Bayes, and Logistic regression, whose accuracy increased to 52.14%, 50.22%, and 55.27%, respectively. Where the LSTM model was contended, the accuracy of prediction after emoji usage increased to 94.32%. These observations highlight the importance of emojis in enhancing the accuracy of sentiment analysis.

Table 13

Model Performance after the Inclusion of Emojis in the Sentiment Analysis

Model	Accuracy before emoji usage	Accuracy after emoji usage
KNN	0.411121	0.521433
Logistic Regression	0.412349	0.552721
LSTM	0.884435	0.943252
Multinomial Naïve Bayes	0.412349	0.502243

5.6 Q5: Most Robust Approaches to Achieve Precise Outcomes

Different studies have suggested different ML approaches as the most accurate in making predictions, and based on the findings of the current study, the best-performing ML techniques included LSTM, Random Forest, and Naïve Bayes. In the following paragraph, this section discusses top-performing ML algorithms based on existing literature.

Rahat (2019) carried out a study to establish the efficacy of Support Vector Machines (SVMs) and Naïve Bayes algorithm approaches for sentiment analysis using review datasets. Rahat (2019) also emphasized the importance of sentiment analysis and its application in various fields, such as marketing, product development, and customer service. They then described the review dataset used in the study, which contained reviews of movies, restaurants, and hotels. The accomplishment of Naïve Bayes and SVM algorithms was

compared for sentiment analysis using various evaluation metrics, including precision, accuracy, F1 score, and recall. Training and testing datasets were varied to evaluate the impact of dataset size on the performance of the algorithms.

The objective of Chen's (2020) study was to develop a framework and procedures for sentiment analysis on social media, using the MilitaryLife PTT board as the experimental dataset. This online forum is one of the largest in Taiwan and was utilized to test the proposed framework. In this study, the authors introduced a military sentiment dictionary which they developed themselves, with the aim of enhancing the accuracy of sentiment classification. They also evaluated the performance of different deep-learning models using various combinations of parameter calibration. The results of this study demonstrated that the model which utilized both the self-developed military sentiment dictionary and existing sentiment dictionaries achieved higher accuracy and F1-measure, as compared to using only the existing sentiment dictionaries. Additionally, the study found that implementing the Tanh activation function in combination with a Bi-LSTM network with two layers produced the best performance for sentiment classification. Therefore, the proposed approach of combining sentiment dictionaries and using specific deep-learning models with optimal parameters was found to be effective in sentiment analysis.

Huang (2021) notes that long short-term memory (LSTM deep-learning algorithms were being widely implemented for sentiment analysis, learning, and text detection. The researchers highlight that current existing sentiment analysis machine learning models could fail to include emoticons with these neural networks purposely based on words or text annotations. This, according to the author, could pose negative results on the overall performance in sentiment classification. This article, therefore, proposed a novel model referred to as AEC-LSTM, which aimed at improving the performance of text sentiment detection by incorporating attention and emotional intelligence (EI) mechanisms. This model

consisted of an emotion-enhanced LSTM (ELSTM) with an EI component designed to improve the classification capabilities of LSTM networks. Additionally, the ELSTM model also incorporated other operations, such as convolution, pooling, and concatenation which helped to better capture various structural patterns within text sequences. The topic-level attention mechanism of the proposed approach allowed for the adjustment of the weight of hidden text representation, thereby enabling the utilization of sentiment semantic information within the context of the text for more effective sentiment representation and classification. The experimental results on real-world datasets demonstrated that this approach significantly improved sentiment classification performance, outperforming state-of-the-art deep learning-based methods.

Agarwal (2019) used various machine learning models and compared their performance in sentiment analysis on the Yelp dataset. The researchers found that Random Forest (RF) and Support Vector Machine (SVM) models outperformed other traditional state-of-the-art benchmarks currently in use, such as Naïve Bayes and Decision Trees, and performed comparably to deep learning algorithms like Convolutional Neural Networks and Recurrent Neural Networks. In another study, Zhang (2018) conducted a study on sentiment analysis using machine learning models and found that these models, particularly the ones based on Recurrent Neural Networks (RNNs), had already achieved state-of-the-art performance on multiple sentiment analysis tasks, outperforming traditional machine learning models. However, the researchers noted that the performance of these machine-learning models is heavily dependent on the amount and quality of the training dataset.

In yet another study by Liu (2020), various machine learning models were compared for their performance in sentiment analysis based on datasets drawn from Twitter. The authors found that the Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models outperformed traditional deep learning models, including Naïve Bayes and

Logistic Regression, and also achieved competitive classification results with more complex models, such as BERT. Singh (2020) conducted a comprehensive study of the machine learning algorithms used for sentiment analysis, including various neural network architectures, including CNNs, RNNs, and LSTM networks, and compared their performance to state-of-the-art benchmarks. According to the article, state-of-the-art benchmarks for sentiment analysis typically cover more recent developments in the field, such as the use of deep learning models and the incorporation of contextual information. Additionally, these benchmarks often provide insights into the specific applications of sentiment analysis, such as in the domains of finance, politics, or customer service. The authors concluded by highlighting that state-of-the-art benchmarks provide a more comprehensive and rigorous analysis of the field, including recent developments and applications, as opposed to existing machine learning architectures.

Peng (2018) conducted research on sentiment analysis in the Chinese language. The authors discuss various machine learning approaches, including Naïve Bayes, SVM, and neural network models, and evaluate their performance on different datasets. This research offers a comprehensive review of the historical, current, and future developments in Chinese sentiment analysis, examining both monolingual and multilingual perspectives. The study investigates the construction of sentiment corpora and lexica and surveys the approaches to monolingual sentiment classification in Chinese. Additionally, the researchers introduce a sentiment classification method based on a multilingual approach. In comparison, state-of-the-art benchmarks for Chinese sentiment analysis are more technical and comprehensive and often involve the evaluation of multiple models on standardized datasets. These benchmarks typically cover a wider range of applications and domains, including social media, news, and reviews. Additionally, state-of-the-art benchmarks often incorporate recent developments in natural language processing, such as machine learning algorithm approaches and transfer

learning. State-of-the-art benchmarks also provide insights into the specific challenges of Chinese sentiment analysis, such as the complexity of Chinese Syntax and the lack of annotated data. These benchmarks often propose innovative solutions, such as leveraging external knowledge sources or applying domain adaptation techniques, to overcome these challenges. In conclusion, this research was harmonious with previous studies in that state-of-the-art benchmarks provide a more detailed and up-to-date analysis of the field, including recent developments, challenges, and solutions.

People all over the world took to social media during lockdown periods and beyond to share their thoughts and feelings about the pandemic and how it has affected their daily lives. The number of tweets on coronavirus, both good and negative, as well as neutral, on Twitter, one of the most widely used social media platforms, increased dramatically in a short time (Jalil et al., 2022). Due to the varied content of tweets, most academics shifted their focus to sentiment analysis, where they studied the public's complex emotions in response to COVID-19.

Sentiment analysis of COVID-19 was conducted by Jalil et al. (2022) using cutting-edge machine learning and deep learning methods. It was stated in the study that the COVID-19 epidemic had an effect on people's daily routines all around the world. Users of the social networking site Twitter who often voice their thoughts, were the focus of this research. Pre-processing, keyword trend analysis, word embeddings for feature extraction, and classification methods make up the study's recommended strategy. For the COVIDSenti dataset, which includes Twitter postings related to COVID-19, the suggested method provides improved accuracy and competence in its classification. To better understand and manage the pandemic, it is possible to detect early signs of COVID-19 emotions in collected tweets. Sentiment analysis of tweets reveals good, negative, and neutral categories.

Communicating using emoticons is not limited to just showing how you feel. As a result, they can be used to convey more than just the literal meaning of a message. It is a fast and effective way to convey emotion, often more so than words alone. When it comes to texting, emojis are the missing piece. This is most analogous to the nonverbal signs individuals pick up on when conversing with others face to face. The risk of miscommunication can be mitigated by the use of emojis and emoticons, much like it can be through the use of body language in face-to-face talks (Zhao et al., 2018). While emoticons can aid in communication and comprehension, they also create room for interpretational error, which can slow down the process.

However, as Liu et al. (2021) noted in their study on enhancing the precision of sentiment analysis through the use of emoji embedding, the difficulty of precisely identifying and distinguishing individual emotions from online texts is largely attributable to the wide variety of syntax and semantics used by different languages. To get over this constraint, the study included support for emojis, a new type of emoticon that contains hundreds of graphic symbols and is widely used in online interactions. The study concluded that emojis are useful as extending characteristics for enhancing the precision of sentiment analysis algorithms and that the performance of the algorithms can be improved by considering various emoji usages. When looking into how people's feelings changed over time online, researchers found that the COVID-19 epidemic had a profound effect on people's emotions, leading them to become more apathetic. In order to better manage emotion mining activities, an emoji embedding method was developed by combining emojis and emoji usage with the sentiment analysis model (Liu et al. 2021). In the process, a more dynamic approach to policy sentiment analysis and prediction of potential responses can be substantially improved.

According to the findings of Yoo and Rayz's (2021) study on emoji comprehension for sentiment analysis, the emoji's meaning can be slightly altered depending on the device it is

viewed on due to the variety of graphics used to render the emoji and the meaning of emojis can also vary from one culture to the next. In China, an emoji of a smiling face may be taken as a sign of happiness, whereas in Argentina, it may be taken as an indication of sadness. The survey confirms what many people already know: emojis are used increasingly frequently because they are a quick and easy way to convey meaning. According to the results, emoticons convey more emotion than words do. With this understanding, the researcher can improve the accuracy of sentiment analysis by a small margin by focusing on messages that contain multiple adjacent emojis, which demonstrate a stronger sentiment ratio. Since the emotions conveyed by emoji may be used to train other English words, using just one or two words in place of an emoji yields maximum accuracy. Therefore, emojis and emoticons are great tools for analyzing public opinion.

It is possible to think about emojis as a whole as a visual language for conveying feelings. They enhance the efficiency of computer-mediated communication by allowing for the expression of nuanced emotions beyond the confines of words. Modern emoji text messages can be automatically composed thanks to improvements in machine learning (Zhao et al., 2018). Since the volume of available data is constantly increasing thanks to social media users' contributions, understanding their actions is in great demand, especially in the midst of the current COVID-19 pandemic. With a standard set of emoticons and some modifications on the primary positive and negative polarities, Alsayat (2022) was able to determine the polarity of tweets. Whenever there is more than one emoticon in a text, it is observed that the first emoticon in the text is given the polarity of the entire text.

The current study identified LSTM and Transformers as the best-performing models, and although both had weaknesses and strengths, a comparison of these models, not just based on their performance in this study but also based on previous research, can provide a clearer picture of their respective strengths and weaknesses. Karita et al. (2019) conducted a

comparative study to establish the differences between LSTM and Transformer model architectures in sentiment analysis in the context of the speech application. The researchers conducted experiments using both Transformer and LSTM models on various speech tasks involving automatic speech recognition (ASR) and speech synthesis. Different datasets were used in the evaluation of the models to help out in the comparison of their performance in terms of their accuracy and computational efficiency. The study results indicated that Transformer models exhibited better performance results as compared to LSTM models. Transformer models achieved high accuracy results in ASR tasks, especially with inputs that had long sequence lengths.

In carrying out the analysis, some of the clear strengths of transformers included the fact that they have the ability to effectively capture long-range dependencies and are parallelizable in nature, which enables fast training of the model. The author also identified that the model faced several challenges involved in its training as a result of the large parameter sizes required. The researchers recommended several potential techniques that could be efficient in improving the model, including compression and knowledge distillation for sentiment analysis datasets.

According to Sudhir (2021), who provided a comprehensive overview of the different approaches, applications, and classifiers that are applied in sentiment analysis, extraction of subjective information from textual data plays a crucial role in decision-making processes. This study emphasized the importance of sentiment analysis in various domains, such as social media, customer feedback, and product reviews.

A close comparison of the Long Short-Term Memory (LSTM) and Transformer model indicated that some of the strengths of LSTM were its efficiency in the capture of sequential dependencies, thus making it suitable for long-term dependencies modeling as compared to the parallelizable architecture of the Transformers model, which makes it efficient for the

self-attention mechanism for capturing global dependencies. This study also highlighted the weaknesses of both neural network models highlighting that LSTM was prone to vanishing and exploding gradient problems making it limited to parallelism during model training. Compared, the Transformers model is less effective in modeling sequential information, requiring larger amounts of training data.

Consistent with the findings of the current study, Wagh et al. (2021) explored the performance and effectiveness of Long Short-Term Memory (LSTM) and Transformer models for sentiment analysis of tasks with long documents. A comparison of sentiment analysis conducted using both architectures based on accuracy, precision, recall, and F1-score indicated that both models achieved competitive performance in lengthy sentiment analysis tasks, with the Transformer model exhibiting slightly better in terms of accuracy and F1-score as compared to the LSTM model.

In terms of the strengths and the weaknesses of both models in the context of the current study, LSTM was superior in the capture of sequential dependencies and long-term dependencies, thus making it suitable for the processing of textual data that had temporal relationships. On the other hand, the Transformer model mechanism for self-attention enabled it to be effective in the capture of global dependencies and contextual information with long documents.

Lu et al. (2022) research focused on the comparison of LSTM and Transformer models for textual classification in sentiment analysis for unstructured medical notes with the aim of assessing the performance and efficiency of these two deep learning architectures in the handling of textual datasets with high levels of class imbalance. Analysis results of both models revealed that both of the models were effective in textual analysis of medical notes, with the Transformer model posting slightly better performance results. A focus on the strengths and weaknesses of both architectures indicated that LSTM models were capable of

capturing long-range dependencies and temporal information, thus making them suitable for the analysis of sequential datasets like medical notes. In comparison, Transformer models were effective in the capture of global context by leveraging their self-attention mechanisms.

5.7 Q6: Performance Comparison with the Start-of-the-Art Benchmarks

Many prior studies have applied different machine learning models to perform sentiment classification of opinions based on disaster-related data. The study findings were compared to other research works related to sentiment analysis. However, the comparison was not equitable since the other research works utilized different datasets, which provide information about the characteristics behind it and classification methods used with their outcomes in terms of accuracy, precision score, F-value, recall value, and accuracy. The models applied in the classification of COVID-19 global data performed differently with regard to accuracy and precision.

The transformer model recorded the highest precision, f-score, accuracy, and recall metrics. LSTM also recorded significantly high performance with regard to accuracy and precision, but slightly lower than the Transformers model. The lowest performance was recorded by the KNN model— accuracy of 41.11% and precision of 41.23%. Where LSTM and Transformers are concerned, the models were considered highly accurate and most suitable models for sentiment analysis, with a precision of 88.45% and an accuracy of 88.44%, and 90.18% precision score, and 90.17% accuracy, respectively. LSTM was the only good model, with a higher prediction score of over 80%. The other models exhibited lower accuracy scores of below 50%, revealing that they are not efficient in the prediction of Twitter-related sentiments and reactions toward COVID-19-related precautionary measures and policies. These findings support the H2 '*LSTM is the most reliable and accurate approach to both measure and predict public sentiments for decision-makers while managing and administrating the spread of COVID-19*'. The model exhibited an accuracy and precision

of over 80% and fitted the selected dataset. The LSTM algorithm showed superior predictive precision with fewer errors when measured on the available real data sets. It recorded a significant accuracy score of 88.52%, which makes a good model for predicting the sentiment polarities in the provided dataset. Even though the transformers model (90.17% accuracy) outperformed LSTM in terms of accuracy and precision, LSTM remains a better model for analyzing tweet sentiments since it does not experience the issue of model overfitting and can be generalized to new datasets. On the other hand, the transformers model was found to be too specialized for the dataset and cannot be generalized to new datasets, making it unsuitable for prediction. Transformers can make inaccurate on the Twitter dataset, resulting in poor performance in practice. This, therefore, means that LSTM is the most suitable model for evaluating and predicting public reactions and emotions toward precautionary measures and policies implemented following a global outbreak or disaster. This approach and predicted consequences will help prevent COVID-19 pandemic infections by implementing appropriate measures to address the issue.

In comparison with other models in past literature, the accuracy and precision in the current study showed significant differences in terms of performance. Yeasmin et al. (2022) applied several models, including logistic regression, support vector machine, Naïve Bayes, KNN, and Random Forest, to predict users' sentiments on the COVID-19 pandemic using tweets. In the study, the random forest classifier performed better than other machine learning models. The accuracy of random forest with TF-IDF was 97.11%, and the F-score was 96.41%, which was the highest. The other models performed well but recorded low performance. Logistic regression recorded 90.88% accuracy and 87.98%, Naïve Bayes had 90.83% accuracy and 87.70% F-score, and KNN recorded 90.23% accuracy and 87.68% F-score. Similar to the current study, Yeasmin et al. (2022) adopted the 80:20 approach of train-test-split after extracting key features to divide the dataset to assess the deep neural networks.

Our study employed LSTM, whereas Yeasmin et al. applied LSTM and CNN algorithms. The accuracy of the test data was established. The accuracy and F-score of the CNN model using Word2Vec embedding were 99.89% and 99.99%, respectively. The accuracy and F-score of LSTM using Word2Vec were 99.88% and 99.99%, respectively. CNN outperformed LSTM in the study and, as such, was considered more effective in predicting user sentiments about the COVID-19 pandemic. In both studies, LSTM recorded greater performance with an accuracy score of above 80%, which suggest its efficiency and effectiveness.

Similarly, Chintalapudi et al. (2021) applied deep learning models to classify sentiments of COVID-19-related tweets. Analysis of data was performed with the BERT model and compared to other models such as support vector machine, logistic regression, and long-short term memory (LSTM). The accuracy of each sentiment was separately evaluated. Accuracy was a primary parameter in evaluating the performance of the models. The study adopted the AdamW optimizer and applied the Adam algorithm by weight decay to fine-tune and duplicate the training features. The training accuracy was acquired since the model was trained with the epoch or individual stage. On the used dataset, the outcome test loss was 0.594, and the test accuracy was 0.890, which revealed that the model is more effective in generalizing tweets in an appropriate way. The model with the highest accuracy (89%) proved to be an appropriate model in the classification of tweets. The BERT algorithm produced an accuracy of 89%, the logistic regression produced 75% accuracy, the support vector machine recorded 74.75% accuracy, and LSTM produced 65% accuracy. Compared to the current study, where the LSTM model is the best model, LSTM in Chintalapudi et al. (2021) study underperformed against other models.

In comparison with the current study findings, Liu et al. (2021) utilized several machine-learning models to evaluate the accuracy of sentiment classification using emoji embedding. The experiment classified tweets in the datasets based on emoji usage and

evaluated the changes in the performance of models. Emojis were established as effective in improving the accuracy of sentiment analysis models, and the algorithm performance was further enhanced by accounting for different emojis. The accuracy of each model increased after evaluating the consistency between the sentiment of plain tweets and those with emojis. The study suggested that the use of emojis significantly enhanced the accuracy of all models, but CEmo-LSTM provided the best performance in the sentiment analysis tasks. In sentiment analysis of plain tweets, Logistic regression had 0.727 accuracies, support vector machine had 0.681, Naïve Bayes had 0.709, LSTM had 0.790, BERT had 0.762, and CEmo-LSTM had 0.8111. The models recorded improved accuracy with the usage of emoji-embeddings. The LSTM and Cemo-LSTM performed better than others, with 0.939 and 0.949 accuracies, respectively. The study established that the COVID-19 virus significantly affected individuals' opinions and sentiments, which resulted in more passive emotions like sadness and horror.

Different polarities and sentiments regarding the COVID-19 pandemic were extracted depending on the dataset and model applied. The analysis of tweets showed an imbalance in the sentiment distribution toward COVID-19 policies and precautionary measures, where 37.3% of tweets exhibited positive emotions, 36.3% displayed neutral emotions, and 26.3% displayed negative emotions. The higher number of positive perceptions suggested that individuals had confidence in how the government was handling the pandemic. The positive tweets in the experiment were attributed to a lack of fear among the public because, in lockdown, individuals were using social media more while sitting at home. The neutral and positive tweets more to ensure individuals support the government to prevent the spread of the coronavirus. At the end of the lockdown, public problems such as strict government guidelines kept increasing, which exacerbated difficulties for the public. These reasons changed the direction of people's sentiments to negative. Consistently, Yeasmin et al. (2022)

applied deep learning models to classify Twitter datasets. The study was conducted to explore people's emotions toward COVID-19 by looking at the tweet. Individuals who were mostly curious about COVID-19 fall under the neutral class. Based on the study, 61.2% of the emotions of individuals were neutral, 20.6% were positive, and only 17.6% were negative. The inclusion of emojis or emoticons in the sentiment analysis significantly improves the prediction model, hence the potential to improve the prediction of sentiment related to policies and actions during times of emergency.

Text Analytics, more particularly, sentiment analysis, is not a new concept by any means. Nonetheless, it has also undergone various iterations where related models have improved over time. First, researchers started with a bag of words approach to understanding whether specific words were capable of conveying certain emotions. Research then moved to LSTM, which used more sophisticated models to aid in understanding emotions through the required training; however, they lacked parallelization, and this made them resource intensive and slow. In 2017, researchers at Google suggested the concept of the Transformers Model, which appeared more efficient than its predecessors. The input embedding is multi-dimensional, and this means that it can process complete sentences and now a series of words one by one. Second, the model has a powerful multi-headed attention mechanism that allows sentences to maintain relationships and concepts between words in a sentence. This attention analysis is performed for each word multiple times to ensure adequate sampling. Lastly, it uses a feed-forward neural network to normalize results and offer sentiment prediction (polarity).

The issue with RNN is that they simply store previous data in their “short-term memory,” and once the memory there runs out, it also deletes the longest retained data and replaces it with new information. The LSTM model makes attempts to escape such issues through the retention of selected data in long-term memory (Ma et al., 2017). Long-term

memory is stored in the Cell State. Additionally, there is also a present hidden state. The state is the short-term model of the LSTM model, therefore, explaining the name Long Short-Term Networks. Performance-wise, Transformers are faster than LSTM since all inputs are ingested once. Similarly, training an LSTM model is harder than with Transformers networks because the number of parameters is more in LSTM networks. Furthermore, it is almost impossible to transfer learning in LSTM networks, and with that, it can be concluded that transformer networks provide the best accuracy and are equally accompanied by less computational cost and complexity.

One of the key benefits of LSTM, as established in this study, is that using the LSTM model provides better results and especially for unstructured data since it is able to acquire useful functionality for resources that contain long-term dependencies. At the same time, they appear better at dealing with long-term dependencies, and this can be attributed to their capability to remember information for an extended period of time. Also, Farzad et al. (2019) discovered that LSTMs are less susceptible to vanishing gradient issues, and this is due to the fact that they use a different activation function that aids in preserving information over long sequences. Lastly, the LSTM model appears efficient in the modeling of complex sequential data, and this is because it could learn high-level representations which capture data structure.

The use of the LSTM model also has its drawbacks, and some of them include; 1) they are more complicated than the conventional RNNs and, as a result, require more training data to learn effectively (Wang et al., 2018). Additionally, the model is less suited to perform tasks online since input data is not in a sequence. Lastly, the model could be slow in training on large datasets, and this can be attributed to the need for the model to learn the parameters of the LSTM cells, which could be computationally intensive.

According to literature, some models have evolved from the Transformers architecture, and with it were some identified weaknesses and strengths. Al-Rfou et al. (2019) claimed that

Transformers models were likely to capture better long-term dependencies than RNNs. Additionally, the model is able to capture short-term dependencies; at the same time, the model inferences for a more protracted context are faster when using a Transformers model such as Transformer-XL. The model is capable of achieving SOTA results for audio analysis and language modeling tasks. Notwithstanding its numerous strengths, a weakness emerges where the model is only useful in limited application areas such as machine translations and audio analysis.

Other strengths of the Transformers Model (s) include their ability to improve lexical robustness once applied, a strength specifically noted in the GPT model. The trained GTP model could be pre-trained and, later, fine-tuned to carry out other tasks without having to customize the model. Additionally, it also outperforms different other models on domain-specific datasets and produces SOTA results on a vast range of domain-specific language modeling tasks. Other models, such as GPT-3 and GPT-2, require no fine-tuning (Liu et al., 2021). However, such branches of the Transformers Model have a resource-intensive nature that makes it expensive to pre-train. Similarly, it is unable to model dependencies that are longer than the designated fixed lengths.

Other strengths of the Transformers Model include its ability to handle the extraction of contextual information because of its bi-directional ability (BERT). Training is faster and has been applied in numerous language modeling applications. However, the weakness of such a model includes its limitation to monolingual classification. The length of the input sentences equally limits the model. Similarly, the model suffers from pragmatic inferences (Ettinger, 2020); most importantly, using this model is likely to be computationally expensive. Ultimately, one of the key challenges of the transformers model experienced in the current study is the fact that the model can only deal with fixed-length text strings. What this means is that the text has to be split into a certain number of segments before being fed into the

system as input, and this segmenting or chunking of text causes context fragmentation. The findings of the current study, therefore, supported hypothesis 2, and the conclusion made is that LSTM is the most reliable and accurate approach to both measures and predict public sentiments for decision-makers while managing and administrating the spread of COVID-19.

5.8 Challenges of Sentiment Analysis in Natural Disasters

Despite its potential benefits, sentiment analysis in social media can be challenging in the context of natural disasters. This is since the data used for sentiment analysis is often noisy and unstructured, and the sentiment of the data can be difficult to identify accurately. Furthermore, sentiment analysis can be affected by the language used in social media and the context in which the data is collected (Wankhade et al., 2020). As such, it is essential to consider these challenges when designing sentiment analysis models for natural disasters. For example, the data used for sentiment analysis may contain misspellings, abbreviations, and slang words, making it challenging to identify the sentiment of the data accurately.

Wankhade and partners further note that the sentiment of the data may be affected by the language used in social media and the context in which the data is collected. For example, the sentiment of a tweet about a government's evacuation plan during a natural disaster may differ depending on whether the tweet is from a local resident or an outsider. As such, it is crucial to consider these challenges when designing sentiment analysis models for natural disasters. Sentiment analysis is also prone to bias because it relies on the sentiment associated with words. These sentiment scores can be inaccurate or incomplete, leading to erroneous results. Additionally, according to Wankhade and partners, sentiment analysis is limited by the data quality. Poorly written or incomplete text can lead to inaccurate results.

Chapter 6: Summary, Conclusion, and Recommendations

Lastly, this chapter will summarize the findings of the study, suggest and propose additional potential approaches to be explored in the future, and finally conclude.

6.1 Summary

Sentiment analysis in social media has become an increasingly important tool for researchers and policymakers to understand public opinion and reactions to various government precautions and policies related to natural disasters. The current study has examined the current state of research on sentiment analysis in social media during natural disasters and the effects of governments' precautionary measures and policies on the public's reactions. It has demonstrated the potential of social media data in understanding the impact of natural disasters and in analyzing and predicting the public's sentiments and reactions toward governments' precautionary measures and policies.

Available literature converges on several ideas. First, the public's sentiments concerning lockdowns to contain the COVID19 Pandemic were largely negative but shifted to positive as the government moved towards reopening the economy. This finding counters the initial idea that the prevailing situation will determine the public's perception concerning the undertaken measure. In essence, studies carried out earlier in the Pandemic suggested that as the number of cases surged, the public would be more willing to adhere to the strict containment measures adopted by most governments; but this assumption did not apply to all situations. For instance, in Italy, the public's sentiment towards the strict and prolonged containment measures was more positive than in the US and UK. In China, the containment measures elicited positive sentiments as the public felt that the government was responding to the situation, but over time the sentiments shifted toward negative. For these reasons, most researchers concluded that public sentiments toward lockdowns were dynamic, reflecting the uncertainty that characterized the period. But throughout this duration, a good proportion of

the global population was against the lockdowns and in favor of the reopening. Despite this simplification, multiple investigations indicated that most public views remained neutral as they were neither positive nor negative.

Based on the reviewed literature, the impact of the Pandemic's containment measures varied significantly. In other words, the enforced lockdowns did not have a uniform effect on the global society, as some groups were more affected than others. The effects of the Pandemic's containment measures varied according to factors such as socioeconomic status, race, and gender. Multiple studies have uncovered a pattern demonstrating a link between the individual community's socioeconomic class and compliance with the COVID-19 containment measure. A positive correlation exists between socioeconomic class and compliance rate, as high socioeconomic class leads to high compliance and vice versa. Another socioeconomic dimension is socioeconomic inequalities, whereby high socioeconomic inequalities adversely affect compliance rates and vice versa. Besides the economic factor, the impact of COVID-19 containment measures varied across races, with minority communities experiencing the most hardship. This point is supported by the fact that, unlike majority races, minorities are overrepresented in low-income labor with non-structural contracts, leaving them vulnerable to job insecurities once the lockdown policy is enforced. At the same time, the lockdown and social distance measures were inapplicable to an excellent proportion of minorities who work in essential sectors, which means they have high contact levels. Nevertheless, the lockdown measure only sometimes protects them. Additionally, COVID-19 lockdowns affected women and those living with disabilities. In summary, COVID-19 containment measures exaggerated the existing inequalities, with specific population segments experiencing more suffering than others.

Examining the literature resulted in the observation of several specific patterns within the field of emoji-based sentiment analysis and even textual opinion mining in general. There

are numerous notable similarities detected regardless of the recentness, origin, or target of the study. Among the most notable approaches or trends is the heavy reliance on lexicons, dictionaries, or simply emojis' annotations libraries to generate optimum accuracy. In fact, many studies attempted to create their own lexicons based on human annotations or by integrating multiple existing libraries in order to generate multiple interpretations for each emoji. Furthermore, considering the massive number of worldwide social platforms, Twitter is distinctly the most favored among researchers mainly due to its accessibility, limited number of characters, and versatile users. In general, social media is by far the most fertile source of sentimental and emotional data.

Further investigation into the literature, as well as based on what the current study establishes, it was revealed that supervised machine learning is almost exclusively adopted by authors as the most reliable approach for sentiment analysis, with some minor exceptions. In particular, classification algorithms are dominant in this field of study. Although classifiers, in general, have demonstrated some robust results, other techniques, and methods definitely do worth the attention of researchers in the field. For example, Wolny (2016) has demonstrated some potentially promising results using unsupervised machine learning algorithms. In addition, there is a certain limitation that was present in nearly all of the reviewed studies—the undeniable issue of unpopular emojis' ambiguity and their unidentified specific single or multiple annotations. Even with commonly used emojis, the multiple meanings concern has viciously impacted the classifiers' accuracy leading to the production of extremely sophisticated lexicons. Finally, sarcasm, hidden emotions, or implications, in general, are still some of the most challenging aspects when analyzing emoji-based statements or posts. Multi-vector or multi-class lexicons and classifiers are yet incapable of accurately interpreting such highly sophisticated human-like uncertainties.

Overall, as mentioned before, it has been clearly observed that disaster management has been somehow left out of the broad applications of sentiment analysis. While there are some minor attempts, the unquestionably significant subject of predicting the negative side effects of precautionary measures certainly deserves immensely deeper and more robust investigations. In summary, a combination of state-of-the-art approaches in every stage of the conventional sentiment analysis paradigm discussed later in section 4 has been oddly overlooked in the literature. Machine-Learning based approaches, recent neural network techniques, emojis' sentiments, multi-level classifications, automated data labeling, prediction-based embeddings, along proper benchmarking have never been observed in a single implementation. In theory, assembling all of these techniques possesses tremendous potential in achieving outstanding accuracy and significant possibilities that are certainly worth the attempt.

From the current research as well as previous research, it is pretty clear that sentiment analysis has received a lot of attention as the need to bridge the gap between online interactions and reality increases. However, researchers interested in this area face several hurdles in selecting the most appropriate techniques in terms of accuracy and precision. The situation is made worse by the fact that language is dynamic and ever-evolving, as evidenced by the rise of the popularity of emojis. Despite being used globally and across all social media platforms, emojis usage is diverse and often carries embedded meanings. Therefore, sentiment analysis experts are tasked with deciphering the meaning from short texts and interpreting the underlying meaning(s) of emojis. According to existing literature, it is apparent that multiple techniques in interpreting texts with emojis have been proposed, but the most outstanding are CNN and LTSM. Each model has its strengths and weakness; combining them could help make up for shortcomings and increase overall effectiveness and accuracy. As such, the most optimum technique to uncover emoji annotations and hidden

meanings could be CNN-LSTM. But this model has only been employed by three of the reviewed literature, and one attempted to make comparisons with other models. For this reason, one cannot conclude with certainty that the CNN-

In a social media platform, sentiment analysis is referred to as the process of using ML and NLP techniques for the extraction and classification of expressed opinions. The ultimate objective discovered in this review is to identify the sentiment of a particular text, whether it is neutral, negative, or positive, all on the basis of either the phrases or used words. Such information is useful in gaining insight into consumer sentiment, brand perception, and public opinion on the events, services, and products in a market. It is evident that assessing the sentiment expressed in a subjective sentence is among the most active fields in research. Researchers such as Thelwall et al. (2011) and Chatzakou et al. (2013) discovered that in all manners of emotional analysis, the classification of polarity is the most important task to execute in both emotional and sentiment analysis, that is, the assessment of both negative and positive feelings of emotions. In simpler terms, to assess the negative and positive meanings. Such categorization has also been discovered to possess a great public and commercial value that is likely to consume time and cost once such research is artificially permitted. As a branch of sentiment analysis, research on polarity classification has also been seen to perfectly meet the potentially increased demand for any consumer. However, researchers also show that studies on the simple dichotomy point of the process have ignored the diversity and richness of emotions of humans, their granularities, and their dimensions.

To a certain extent, multiple-level sentiment analysis was suggested to address the aforementioned issue. Furthermore, scale systems calculate the strength score of the negative or positive emotion in a text, and this translates to exposing the sentiment to a certain interval. The technique, as reviewed in the study's literature, has upgraded the comprehension of any textual emotion to a more comprehensible level. The identification of

objectivity or subjectivity is, in some cases, more sophisticated than the problem identified in the classification of polarity. An objective document, as identified by Wiebe and Riloff (2005), is likely to contain a subjective statement, while subjective phrases and words are based on the content. An aspect- or feature-based sentiment analysis is also likely to establish the sentiments or opinions on the different features or aspects of a certain entity that mostly involves the extraction of an opinion holder. Basically, the progressing development in existing application scales or an extreme demand surge for a specific service is likely to lead to an increase in the need for the process at various granularities. Researchers have been identified to broadly handle the challenge by categorizing the process of sentiment analysis into the following listed levels; word-, sentence- and document levels. This means that a multi-granularity sentiment analysis assumes a completely labeled corpus at a course- or fine-grained level, or even both. The review also highlighted both the disadvantages and advantages of the multi-granularity analysis with a focus on the merits and demerits of adopting a cross-domain sentiment analysis on a document level and also other levels.

Emojis are a means to enhance social media interactions using tiny expressive images that are capable of implying relatively rich content (Stark et al., 2015). These simple characters perform a substantial role in efficiently conveying ideas, expressing emotions, soothing a tense conversation, or merely injecting humor into any given context. Emojis, sometimes referred to as ideograms or pictograms, were formally introduced in the late 1990s. Since then, emojis have been widely involved in enhancing social media textual and non-textual content by implicitly infusing feelings (Wood et al., 2016), sarcasm (Joshi et al., 2017), and other sentiments (Novak et al., 2015). Although they started as an enhanced graphical representation of emoticons, and basic facial expressions that mainly utilize punctuations, emojis have rapidly developed into portraying flags, animals, common objects, races, genders, and many more genres. In fact, many technology giants, such as Apple,

Google, and Samsung, are heavily investing in improving their operating system emojis to further enhance their users' experience.

Emojis are distinguishingly superior to other forms of expression due to the ability of a single emoji to imitate multiple emotions and reactions depending on the context it is used for. In other words, emojis are being heavily used as highly effective shortcuts that are capable of abbreviating multiple sentences into a single icon. In fact, Twitter counts an emoji as a single textual letter as part of its 240 characters restriction. However, these potentials can commonly be very misleading, and emojis can occasionally be misinterpreted or misunderstood. In other words, people do not usually agree on the meaning or feeling that an emoji is meant to convey. For instance, the same emoji can be used in multiple contexts and yet convey completely different expressions or feelings depending on the nature of the context. In general, human beings can be very sophisticated with their emotions and expressions, which leads to very complicated and ambiguous emojis usage scenarios. In particular, emojis are mainly involved in posts containing complex content such as metaphors, rhetoric, mockery, and humor.

Recently, emojis started to gain the attention of both individuals and enterprises due to their powerful yet simple approach to conveying complicated sentiments and expressing intricate emotions. Hogenboom et al. (2013) pursued the notion of creating a dedicated emoji sentiment dictionary to achieve a state-of-the-art sentiment classification approach. The method was evaluated against 2,080 various and random social media posts that contain various emojis and have been humanly interpreted. The study discovered that emojis have a paragraph-level influence, particularly those having a single emoji. However, some paragraphs contain multiple emojis. In this case, emojis tend to have sentence-level effects. The research framework deployed two parallel data flows to compare emoji-based machine analysis with text-based human analysis. Furthermore, the study established that emoji-based

sentiment analysis substantially enhances the classification and interpretation accuracy with almost 70% improvement. Nevertheless, the model was executed exclusively on Dutch social posts. Thus, the ability to annotate multilingual tweets or messages remains indefinite and definitely requires further consideration.

The most tangible challenge concerning emoji-based analysis is the uncertain relationship between emojis and sentiment polarity. Nevertheless, it is highly crucial to take emojis into consideration while dealing with sentiment polarity. Otherwise, the algorithm will most likely achieve a misleading annotation. Wang and Castanon (2015) proved the current predominance of emojis in different social media platforms by examining the frequency of such graphical expressions in a large dataset obtained from Twitter. The study conducted four different approaches to explore the association between emojis and sentiment polarity in addition to the context in which emojis are usually used.

The first method examined a group of participants to analyze their perceived emotional polarity of the standard emojis using a qualitative survey. Secondly, an emoji-based lexicon was initiated to carefully investigate each emoji's annotations. The third approach correlated the sentiment polarity of social media posts prior to and after emojis were eliminated from the text. Moreover, the final step tested the assumption that eliminating emojis from a specific text significantly impairs sentiment classification. This has been achieved by training two separate machine learning algorithms, including and excluding emojis from text-based posts (Wang et al., 2015). The authors concluded that there is, in fact, a specific group of emojis that can be intensely reliable when it comes to sentiment polarity. These few emojis must be significantly considered and taken advantage of when dealing with sentiment analysis. In addition, the majority of emojis convey multiple and complicated emotions and thus must be handled with absolute caution.

Emojis, a new and more sophisticated class of emoticons, are being increasingly featured in electronic messaging and social networking. Studies show that nearly ten billion emojis have been posted on Twitter in the last two years (Novak et al., 2015). Emojis can be defined as Unicode visual icons used as shortcuts to convey thoughts and emotions. Compared to emoticons, emojis are significantly more powerful and highly versatile. Studies that have been cited in this study, for example, Novak and colleagues (2015) attempted to introduce the first emoji sentiment lexicon called the Emoji Sentiment Ranking. In addition to the lexicon, the study instructed and demonstrated a sentiment chart of the 751 most commonly used emojis. The emojis' sentiments have been derived from the sentiments of the particular social media posts in which they occur.

To achieve that, the authors have engaged around 80 human interpreters to identify emotion polarity (positive, neutral, or negative) using a dataset that consisted of 1.6 million tweets written in 13 different languages within the European region. However, only around 4% of the interpreted posts contained emojis. The emoji-based sentiment analysis enabled the researchers to achieve various interesting outcomes. The study has shown that most emojis, particularly the most prevalent ones, are indeed positive (Novak et al., 2015). Additionally, the distribution of emotions between tweets, including and excluding emojis, is substantially different. Emojis tend to frequently appear at the end of posts, with their sentiment polarity increasing proportionately according to the distance. Furthermore, no major variations were observed among the 13 examined languages and the Emoji Sentiment Ranking.

Another study by Shiha and Ayvaz (2017) explores the use of emojis in social platforms and their impact in the field of sentiment analysis and opinion mining. As in most studies in this particular area, Twitter was considered the primary data source due to its tremendous accessibility and numerous amounts of daily interactions. Text-based tweets engaged in global (positive or negative) incidents have been gathered to determine the potential influence

of such expressive characters in sentiment analysis. The article discovered that the sentiment scores accuracy is exceedingly higher in tweets with emojis against those with text-based expressions only (Shiha et al., 2017). Additionally, similar to previous studies, the authors discovered that the significant impact of emojis actually becomes very evident when analyzing positive opinions compared to negative ones.

6.2 Conclusion

The study sought to investigate and discover a reliably accurate framework to effectively measure and analyze public sentiments for decision-makers while managing and administrating the spread of a pandemic or a natural disaster. The analysis evaluated different ML models to determine which of these methods can measure and predict public sentiments on social media more accurately. Recently, users of social media platforms have become used to using various graphic symbols in their online conversations as a way of expressing their emotions; the characters are identified as emojis, with literature stating that they have become a global language across cultures, platforms, and applications. Generally, the current study discovered that the purpose of these emojis was to make available emotional cues that were otherwise lacking from typed conversations. When compared to words, emojis are more likely to express individual emotions more directly and accurately. In the assessment of short informal texts, including comments, blogs, or tweets, emojis were identified to offer significant information on the sentiment of the user. The prevalence of emojis was also discovered to offer a new opportunity to assess sentiment expressions in a textual context. Additionally, the introduction of emojis was identified to be beneficial in improving the performance of algorithms in sentiment analysis. Emoji tag words could also be useful in directly constructing features for classifier training. It is also worth understanding that, in the training set and following the classification of emoji use in posts, the performance of algorithms was significantly improved.

Singh et al. (2019) replaced emojis with text descriptions since it was considered usual to find embeddings of words in a text. The researchers experimented with a two-stack BiLSTM neural network and exhibited improvement in the categorization of sentiment and detection of irony. This gives reason to believe that emojis are more reliable and effective in the analysis of user sentiment. In the current study, from the collection of text data and from an assessment of the effect of emoji characters in sentiment analysis, the use of emojis showed an increase in sentiment scores. Moreover, there was an observation that the use of emojis in the analysis of sentiments had a higher effect on the overall sentiments of positive opinions when compared to the negative ones. Additionally, using emojis together with texts improved the performance of the classifier.

According to reviewed literature, researchers propose various techniques to satisfy the hidden annotations on social media, most especially Twitter. For instance, Ritter et al. (2011) suggested the NER technique, where they annotated tweets and constructed tools that were trained on unlabeled, out of, and in-domain data and identified significant improvement over the state-of-the-art counterparts that underwent equal training. Also, researchers focusing on the extraction of structured representation of Twitter events have largely been identified to concentrate on newswire text, meaning the unique characteristics that are presented on Twitter by new opportunities and challenges in the process of extracting open domain events. Here, researchers demonstrate an accurate extraction of a calendar of significant events from Twitter and consider it feasible. In addition, literature has also presented novel approaches for the simple discovery of significant event classification and categorized extracted events that were based on the latent variable models.

Like other studies, the current study aimed to identify the most appropriate model that would enable the identification of the hidden annotation associated with emojis and clear vague implications. The findings identified LSTM to be the most appropriate, with the main

idea of the technique being to present a gating technique that regulates the flow of information inside cell states. The technique was identified to compute three gates, namely, the forget, output, and input gates. LSTM recorded significantly high performance with regard to accuracy and precision, but slightly lower than the Transformers model. The lowest performance was recorded by the KNN model (41.11%) and a precision of 41.23%. Where LSTM and Transformers are concerned, the models were considered highly accurate and most suitable models for sentiment analysis, with a precision of 88.45% and an accuracy of 88.44%, respectively, and 90.18% precision score, and 90.17% accuracy, respectively. LSTM was the only good model, with a higher prediction score of over 80%. The other models exhibited lower accuracy scores of below 50%, revealing that they are not efficient in the prediction of Twitter-related sentiments and reactions toward COVID-19-related precautionary measures and policies. These findings support the H2 ‘LSTM is the most reliable and accurate approach to both measure and predict public sentiments for decision-makers while managing and administering the spread of COVID-19’.

6.3 Recommendations

According to reviewed literature, Pang et al. (2002) were acknowledged as the pioneers of sentiment analysis who employed ML approaches with the aim of determining whether a written review is positive or negative. The researchers conducted their experiment with three algorithms: Maximum Entropy, Naïve Bayes, and SVM under different parameters and features, such as bigrams vs. unigrams and feature presence vs. feature frequency. Other researchers, such as Severyn and Moschitti (2015), conducted their study on word embeddings with the aid of a deep Convolutional Neural Network NL. The researchers attempted to classify sentiments. Similarly, classified as an ML approach, Ren et al. (2016) used the Latent Dirichlet Allocation to gain the topic distribution for sentences in the used dataset and later exploited the Recursive Auto-Encoders in learning word embeddings that

were enhanced in the topic. To further improve performance, Ren et al. (2016) integrated representation with conventional models such as logistic regression and SVM. As identified in the literature, supervised ML methodologies considered the classification of sentiment as a standard statistical classification that mainly contains a large number of labeled examples.

Deep learning was also highlighted as a section of the larger members of ML methodologies based on various representations, such as data learning, when likened to algorithms that were based on tasks. There also appeared hybrid ML algorithms such as the SVM to aid in the selection of the best feature from the training data and then offer the features to the artificial neural network to perform a document-level sentiment classification. According to the literature review, the researcher used various ML and hybrid approaches for sentiment classification and forecasting, and there was no single classifier that was good for all kinds of datasets.

The findings of the current study proved otherwise, as from the analysis of various ML and hybrid approaches, the LSTM model performed better than the other models and proved to be the most suitable approach for sentiment classification and prediction, which was expressed by the selected disaster-related dataset. At the end of the experiment, the LSTM recorded an 89.47% accuracy which was declared “good” since it enabled accurate and appropriate classification. Based on these findings, therefore, this study recommends LSTM and the most accurate and reliable ML model in measuring and predicting public sentiments for decision-makers while managing and administrating the spread of a pandemic or a natural disaster in general.

The current study also reviewed the literature on state-of-the-art methods for sentiment analysis and forecasting with various comparisons of Machine Learning and Deep Learning techniques, methods such as RNN and LSTM. Moreover, to the best of the researcher’s knowledge, there was little to no information on the recent development in the domain of

probabilistic forecasting and, most especially, the contributions of deep learning techniques for the forecasting of sentiment analysis. Literature also highlighted the consideration of the Bayesian bi-directional (BLSTM) against other state-of-the-art approaches, together with Bayesian with standard ANN. From implemented findings, it can easily be inferred that the BLSTM outperforms deterministic approaches through the exhibition of the least errors for all variations of data together with the provision of prediction intervals for forecasting scenarios. Despite the feature of the Bayesian deep learning methods to improve the accuracy of forecasting, LSTM possesses a higher execution time, and a trade-off between the performance of the model and the cost of computation is highly required.

In a similar direction, the current study demonstrated an improved accuracy at aggregated levels in LSTM when it came to performing short-term forecasting. This was supported by Kong et al. (2017), who suggested RNN-based LSTM methods in performing short-term load forecasting for residential consumers. The findings showed improved accuracy at the granular level. Similarly, in the same study, the researchers mentioned that the inclusion of neuro-fuzzy logic together with LSTM outperformed other methods. A finding was identified in the current study. Forecasting performance was also mentioned to be improved by clustering individual sub-profiles on various patterns.

Additionally, future research can adopt joint training to help boost the performance of models in predicting sentiments. A sentiment term that makes use of logistic regression to predict the sentiment of a document is given to help with joint training. Semantic and emotional aims are combined to form the whole objective function. The work of Barnes et al. (2017) builds on prior research by acknowledging studies that have used a variety of sentiment and subjective standards to evaluate model performance, including the sentiment data during training results in lower performance.

Joint training in sentiment analysis refers to the process of training a single machine learning model to perform multiple related tasks simultaneously, such as sentiment analysis and named entity recognition. Combining two or more related tasks in a single model improves the accuracy of the model's predictions. For example, a joint sentiment analysis model might be trained to simultaneously classify text as positive, negative, or neutral while also identifying the named entities mentioned in the text, such as people, organizations, and locations. This helps to enhance the overall meaning and predictive outcomes to ensure that what is being studied is easily identifiable and can be improved with time. The advantage of joint training is that it can improve the performance of both tasks by allowing the model to learn from the correlations between them. For example, suppose the model learns that mentions of certain entities tend to be associated with positive or negative sentiment. In that case, it can use that information to improve its sentiment analysis predictions. In the example of a tweet, the combination of certain words, phrases, and even special inputs like emojis could create a more varied range of meanings. With joint model training, several aspects within the same dataset could be enhanced.

The topic of sentiment analysis in social media involves analyzing and predicting the public's sentiments and reactions towards government precautionary measures and policies during natural disasters, which means that the data collected would entail varying aspects. In this context, the model was trained in joint tasks of recognizing sentiment and emojis from Twitter data that incorporated both text and emojis. Some aspects of joint training in this sentiment analysis that made it possible to utilize the data included the obvious benefits of the process. Joint training can help improve the overall performance of the model by allowing it to learn from the correlations between different tasks. In the case of this research, recognizing emojis in addition to text did help the model better understand the sentiment expressed in a tweet. Emojis can often provide additional context or emotional cues that are not explicitly

expressed in the text. By incorporating both text and emojis, the joint model was able to improve the accuracy of its sentiment analysis predictions across all the training models utilized, as reported in the accuracy results.

Another aspect of joint training is the model architecture, which plays a crucial role. The model is thus designed to handle both text and emojis. This is because the model may consist of separate branches for text and emojis, with each branch, using different feature extraction techniques. The output from each branch can then be combined using concatenation or other fusion techniques to make the final prediction. Alternatively, the model may use a single shared feature extraction layer that can handle both text and emojis. This second aspect can be applied in this case in order to ensure that the meaning of the tweets is not lost due to separating the text and emojis.

6.4 Limitations and Suggestions for Future Research

The current study successfully compares different models (KNN, logistic regression, multinomial Naïve Bayes, random forest, transformers etc.) with LSTM and investigates their precision, accuracy, recall and F score. Future studies could benefit from a comprehensive comparison of LSTM with other models (such as variant of BERT and large language models), especially in the context of COVID-19 sentiments. Future studies could utilize various advanced techniques, including recent neural network methods, multi-level classifications, and prediction-based embeddings. Additionally, there is need for more insights into how to deal with ambiguity where same emoji can have different sentiments based on context can provide some insights. The uncertainty in interpreting emojis, especially in complex contexts such as metaphors or humor, remains a significant challenge. Future works could focus on refining approaches to handle ambiguous and sophisticated emoji usage. They should focus on analysis of combinations of emojis and words since an emoji sentiment is usually determined by the phrases and words in proximity, thus, understanding

the connection between different emojis and their surrounding textual content presents more robust and accurate sentiment analysis. Utilization of pre-trained models such as variance BERT or GPT can also help understand contextual data since they are trained in large sets of data, thus, able to capture the nuances associated with emoji usage.

The current study established that LSTM model yields the best results when uncovering emojis' hidden meanings and annotations because the performance variations between CNN-LSTM and other models have yet to be identified. This was, however, cleared by the current study, as it established that LSTM is, so far, the best model that could be used in making predictions in sentiment analysis. Future research could benefit from a more in-depth examination of emoji-based models, such as the CNN-LSTM model, including performance comparisons with other models to establish their efficacy in uncovering emoji meanings. This could also include qualitative analysis of the results to give in-depth understanding of emojis and how they are used to analyze sentiments. Some sample results, especially showing the cases where the model had some difficulties in correctly classifying the input can provide some insights. The model captures and recognizes patterns and features associated with different emojis compared to the existing LSTM model since it can learn many local features and increase the number of layers in model stack, which allows gathering more global features, hence increasing model accuracy and robustness. The combined CNN-LSTM model can provide qualitative information about how emojis contribute to the overall sentiment or meaning of a piece of text. The model will alleviate cases where the LSTM failed to correctly classify input as it helps understand the context in which emojis are used and how they influence the text interpretation. Analysis of model activations at different layers allows the researcher to gain insights into words or emoji sequences, which improves the overall prediction ability of the model.

Future studies could also be used to validate the findings of the current study by examining an array of datasets, as well as contexts that are beyond the COVID-19 sentiments. Future studies should also focus on dealing with ambiguity, especially where same emoji can have different sentiments based on context can provide some insights. The uncertainty in interpreting emojis, especially in complex contexts such as metaphors or humor, remains a significant challenge. Future works could focus on refining approaches to handle ambiguous and sophisticated emoji usage.

In terms of trade-offs between accuracy and computation cost, the current study noted that computational constraints can influence the selection of the input features. The LSTM model is resource intensive; thus, using many input features would have increased the complexity of the model and memory requirements. Performance-wise, the study established that Transformers are faster than LSTM since all inputs are ingested once. Similarly, training an LSTM model is harder than with Transformers networks because the number of parameters is more in LSTM networks. It was also noted that it is difficult to transfer learning in LSTM networks, and with that, it was concluded that transformer networks provide the best accuracy and are equally accompanied by less computational cost and complexity. Using the LSTM model provides better results and especially for unstructured data since it is able to acquire useful functionality for resources that contain long-term dependencies. At the same time, they appear better at dealing with long-term dependencies, and this can be attributed to their capability to remember information for an extended period of time. Future studies should focus more on understanding the cost implications of use these models in a more comprehensive manner.

6.4.1 Limitations

Sentiment analysis, in general, and this study in particular, entail certain limitations or constraints (Cieliebak et al., 2013). Some limitations are technically chronic and strongly

related to most NLP applications and their existing implementation approaches. Other constraints are more logistic and concerned with how a particular study is being executed. NLP applications are designed to deal with the complexity of human languages and expressions. However, humans are highly unpredictable and inconsistent, unlike machines (Hogenboom et al., 2015). In other words, people tend to engage widely in metaphors, implications, sarcasm, and humor in their daily interactions. Challenges such as name recognition, parsing, social abbreviations, slang, and rhetorical implications have long existed and impose an undeniable major limitation in the application (Hogenboom et al., 2015). Although this study attempted to address some of these challenges using online dictionaries, lexicons, or libraries, exact and perfect human-like interpretations are almost impossible. Naturally, it is impractical for any existing state-of-the-art word embeddings or lexicons to incorporate every existing word or emoji. Thus, no matter how sophisticated a particular machine-learning model is, it will still be incapable of handling some vocabulary and expressions (Liu et al., 2012).

Due to time limitations and constraints, the study overlooked many promising sentiment analysis approaches, such as Support Vector Machines (SVM), Maximum Entropy Probabilistic Classifiers, and Unsupervised Machine-Learning approaches. The latter has recently shown some potentially promising accuracy in the literature (Hu et al., 2013). Finally, depending on the time restriction and the archived results, the study might revert back and experimentally address some of the discussed limitations in order to obtain the most optimal and competitive accuracy possible.

References

- Abdukhamidov, E., Juraev, F., Abuhamad, M., El-Sappagh, S., & AbuHmed, T. (2022). Sentiment Analysis of Users' Reactions on Social Media during the Pandemic. *Electronics*, 11(10), 1648.
- Addawood, A., Alsuwailem, A., Alohal, A., Alajaji, D., Alturki, M., Alsuhaibani, J. and Aljabli, F. (2020). Tracking and understanding public reaction during COVID-19: Saudi Arabia as a use case. In *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*. <https://aclanthology.org/2020.nlpcovid19-2.24.pdf>
- Adebisi, B. O., Roman , N. V., Chinyakata, R. & Balogun, T. V. (2021). The Negative Impacts of COVID-19 Containment Measures on South African Families - Overview and Recommendations. *Open Public Health Journal*.
- Agarwal, A., Xie, B., Vovsha, I., Rambow, O. and Passonneau, R.J. (2011). Sentiment analysis of Twitter data. In *Proceedings of the workshop on language in social media (LSM 2011)* (pp. 30-38). <https://aclanthology.org/W11-0705.pdf>
- Agarwal, A., Yadav, A. and Vishwakarma, D.K. (2019). Multimodal sentiment analysis via RNN variants. In *2019 IEEE International Conference on Big Data, Cloud Computing, Data Science & Engineering (BCD)* (pp. 19-23). IEEE.
- Agarwal, B., Mittal, N., Agarwal, B. and Mittal, N. (2016). Machine learning approach for sentiment analysis. *Prominent feature extraction for sentiment analysis*, pp.21-45.
- Agarwal, B., Mittal, N., Bansal, P. and Garg, S. (2015). Sentiment analysis using common-sense and context information. *Computational intelligence and neuroscience*.
- Agarwal, B., Sharma, V., Harjule, P., Tiwari, V. and Sharma, A. (2022). The COVID-19 outbreak: social media sentiment analysis of public reactions with a multidimensional perspective. In *Cyber-Physical Systems* (pp. 117-138). Academic Press.

- Agusto, F. B. et al. (2021). Impact of public sentiments on the transmission of COVID-19 across a geographical gradient. *BMJ*.
- Ahmed, M. E., Rabin, R. I. & Chowdhury, F. N. (2020). COVID-19: Social Media Sentiment Analysis on Reopening. *arXiv*.
- Ahmed, W., Bath, P.A. and Demartini, G. (2017). Using Twitter as a data source: An overview of ethical, legal, and methodological challenges. *The ethics of online research*, 2, pp.79-107.
- Ahuja, R., Chug, A., Kohli, S., Gupta, S. and Ahuja, P. (2019). The impact of features extraction on the sentiment analysis. *Procedia Computer Science*, 152, pp.341-348.
- Ain, Q.T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B. and Rehman, A. (2017). Sentiment analysis using deep learning techniques: a review. *International Journal of Advanced Computer Science and Applications*, 8(6).
- Akter, S. and Aziz, M.T. (2016). Sentiment analysis on Facebook group using lexicon based approach. In 2016 3rd international conference on electrical engineering and information communication technology (ICEEICT) (pp. 1-4). IEEE.
<https://doi.org/10.1109/CEEICT.2016.7873080>
- Alaba, O. A., Hongoro, C., Thulare, A. & Lukwa, A. T. (2022). Drivers of socioeconomic inequalities of child hunger during COVID-19 in South Africa: evidence from NIDS-CRAM Waves 1–5. *BMC Public Health*.
- Alaoui, I. E. et al. (2018). A novel adaptable approach for sentiment analysis on big social data. *Journal of Big Data*, 12(5).
- Alayba, A.M., Palade, V., England, M. and Iqbal, R. (2017). Arabic language sentiment analysis on health services. In 2017 1st international workshop on Arabic script analysis and recognition (ASAR) (pp. 114-118). IEEE.
<https://arxiv.org/ftp/arxiv/papers/1702/1702.03197.pdf>

- Aldahawi, H. (2015). Mining and analysing social network in the oil business: Twitter sentiment analysis and prediction approaches (Doctoral dissertation, Cardiff University). <https://orca.cardiff.ac.uk/id/eprint/85006/1/2015aldahawihphd.pdf.pdf>
- Aldao, A. and Nolen-Hoeksema, S. (2012). When are adaptive strategies most predictive of psychopathology? *Journal of abnormal psychology*, 121(1), p.276.
<https://psycnet.apa.org/doi/10.1037/a0023598>
- Alhajji, M., Al Khalifah, A., Aljubran, M. and Alkhalifah, M. (2020). Sentiment analysis of tweets in Saudi Arabia regarding governmental preventive measures to contain COVID-19.
- Ali, K., Dong, H., Bouguettaya, A., Erradi, A. and Hadjidi, R. (2017). Sentiment analysis as a service: a social media based sentiment analysis framework. In 2017 IEEE international conference on web services (ICWS) (pp. 660-667). IEEE.
<https://doi.org/10.1109/ICWS.2017.79>
- Aljameel, S.S., Alabbad, D.A., Alzahrani, N.A., Alqarni, S.M., Alamoudi, F.A., Babili, L.M., Aljaafary, S.K. and Alshamrani, F.M. (2021). A sentiment analysis approach to predict an individual's awareness of the precautionary procedures to prevent COVID-19 outbreaks in Saudi Arabia. *International journal of environmental research and public health*, 18(1), p.218.
- Aljarah, I., Habib, M., Hijazi, N., Faris, H., Qaddoura, R., Hammo, B., Abushariah, M. and Alfawareh, M. (2021). Intelligent detection of hate speech in Arabic social network: A machine learning approach. *Journal of Information Science*, 47(4), pp.483-501.
<https://doi.org/10.1177/0165551520917651>
- Almouzini, S. and Alageel, A. (2019). Detecting arabic depressed users from twitter data. *Procedia Computer Science*, 163, pp.257-265.
<https://doi.org/10.1016/j.procs.2019.12.107>

- Alpaydin, E. (2016). Machine learning: the new AI. MIT press.
- Al-Rakhami, M.S. and Al-Amri, A.M. (2020). Lies kill, facts save: detecting COVID-19 misinformation in twitter. *Ieee Access*, 8, pp.155961-155970.
- Al-Rawi, A., Siddiqi, M., Morgan, R., Vandan, N., Smith, J., & Wenham, C. (2020). COVID-19 and the gendered use of emojis on Twitter: infodemiology study. *Journal of medical Internet research*, 22(11), e21646.
- Al-Rfou, R., Choe, D., Constant, N., Guo, M. and Jones, L. (2019). Character-level language modeling with deeper self-attention. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 33, No. 01, pp. 3159-3166).
- AL-Rubaiee, H.S., Qiu, R., Alomar, K. and Li, D. (2016). Sentiment analysis of Arabic tweets in e-learning.
- <https://uobrep.openrepository.com/bitstream/handle/10547/623004/jcssp.2016.553.563.pdf?sequence=2&isAllowed=y>
- Alsayat, A. (2022). Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model. *Arabian Journal for Science and Engineering*, 47(2), pp.2499-2511. <https://link.springer.com/article/10.1007/s13369-021-06227-w>
- Alvis, L., Shook, N. and Oosterhoff, B. (2020). Adolescents' prosocial experiences during the COVID-19 pandemic: Associations with mental health and community attachments. *PsyArXiv Preprints*.
- Asghar, M.Z., Khan, A., Ahmad, S. and Kundi, F.M. (2014). A review of feature extraction in sentiment analysis. *Journal of Basic and Applied Scientific Research*, 4(3), pp.181-186.

- Asghar, M.Z., Khan, A., Ahmad, S. and Kundi, F.M. (2014). A review of feature extraction in sentiment analysis. *Journal of Basic and Applied Scientific Research*, 4(3), pp.181-186.
- Asghar, M.Z., Khan, A., Ahmad, S., Qasim, M. and Khan, I.A. (2017). Lexicon-enhanced sentiment analysis framework using rule-based classification scheme. *PloS one*, 12(2), p.e0171649. <https://doi.org/10.1371/journal.pone.0171649>
- Aslam, F., Awan, T.M., Syed, J.H., Kashif, A. and Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanities and Social Sciences Communications*, 7(1), pp.1-9.
- Asmundson, G.J., Paluszek, M.M., Landry, C.A., Rachor, G.S., McKay, D. and Taylor, S. (2020). Do pre-existing anxiety-related and mood disorders differentially impact COVID-19 stress responses and coping?. *Journal of anxiety disorders*, 74, p.102271. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7342169/>
- Athey, S. (2018). The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda* (pp. 507-547). University of Chicago Press.
- Baccianella, S., Esuli, A. and Sebastiani, F. (2010). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Lrec* (Vol. 10, No. 2010, pp. 2200-2204). http://lrec-conf.org/proceedings/lrec2010/pdf/769_Paper.pdf
- Bai, Q. (2019). A Systematic Review of Emoji: Current Research and Future Perspectives. [Online] Available: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02221/full>
- Baker, Q.B., Shatnawi, F., Rawashdeh, S., Al-Smadi, M. and Jararweh, Y. (2020). Detecting epidemic diseases using sentiment analysis of arabic tweets. *J. Univers. Comput. Sci.*, 26(1), pp.50-70. <https://pdfs.semanticscholar.org/3514/560ef43da59e4af844362b2a11b498aa9632.pdf>

- Bansal, B. and Srivastava, S. (2019). Lexicon-based Twitter sentiment analysis for vote share prediction using Emoji and N-gram features. *International Journal of Web Based Communities*, 15(1), pp.85-99.
- Barbieri, F., Kruszewski, G., Ronzano, F. and Saggion, H. (2016). How cosmopolitan are emojis? Exploring emojis usage and meaning over different languages with distributional semantics. In *Proceedings of the 24th ACM international conference on Multimedia* (pp. 531-535).
https://repositori.upf.edu/bitstream/handle/10230/32660/barbieri_MM16_emoj.pdf?isAllowed=y&sequence=1
- Barbosa, L. and Feng, J. (2010). Robust sentiment detection on twitter from biased and noisy data. In *Coling 2010: Posters* (pp. 36-44). <https://aclanthology.org/C10-2005.pdf>
- Barkur, G. and Kamath, G.B. (2020). Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: Evidence from India. *Asian journal of psychiatry*, 51, p.102089.
<https://doi.org/10.1016%2Fj.ajp.2020.102089>
- Bavel, J.J.V., Baicker, K., Boggio, P.S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M.J., Crum, A.J., Douglas, K.M., Druckman, J.N. and Drury, J., Linden. S. *vd, Lange, P. v., Weeden, KA, Wohl, MJA, Zaki, J., Zion, SR, & Willer*, pp.460-471.
- Beel, J., Gipp, B., Langer, S. and Breitingner, C. (2016). paper recommender systems: a literature survey. *International Journal on Digital Libraries*, 17(4), pp.305-338.
- Bell, J. (2022). What is machine learning?. *Machine Learning and the City: Applications in Architecture and Urban Design*, pp.207-216.
- Ben-Ze'ev, A. (2000). *The subtlety of Emotions* (vol. 12).
- Berry, M.W., Mohamed, A. and Yap, B.W. eds. (2019). *Supervised and unsupervised learning for data science*. Springer Nature.

- Bertocchi, G. and Dimico, A. (2021). COVID-19, Race, and Gender, s.l.: IZA Institute of Labor Economics.
- Bertrand, K.Z., Bialik, M., Virdee, K., Gros, A. and Bar-Yam, Y. (2013). Sentiment in new york city: A high resolution spatial and temporal view. arXiv preprint arXiv:1308.5010.
- Bhanot, D., Singh, T., Verma, S. K. and Sharad, S. (2021). Stigma and Discrimination During COVID-19 Pandemic. *Frontiers in Public Health*, Volume 8.
- Bhat, M., Qadri, M., Kundroo, M., Ahanger, N. and Agarwal, B. (2020). Sentiment analysis of social media response on the Covid19 outbreak. *Brain, behavior, and immunity*, 87, p.136. <https://doi.org/10.1016%2Fj.bbi.2020.05.006>
- Bhuta, S. and Doshi, U. (2014). A review of techniques for sentiment analysis Of Twitter data. 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). [online] doi:10.1109/iciict.2014.6781346.
- Bibbins-Domingo, K. (2020). This Time Must Be Different: Disparities During the COVID-19 Pandemic. *Annals of Internal Medicine*.
- Bibi, M., Abbasi, W.A., Aziz, W., Khalil, S., Uddin, M., Iwendi, C. and Gadekallu, T.R., (2022). A novel unsupervised ensemble framework using concept-based linguistic methods and machine learning for twitter sentiment analysis. *Pattern Recognition Letters*, 158, pp.80-86.
- Birjali, M., Kasri, M. and Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, p.107134. <https://www.sciencedirect.com/science/article/abs/pii/S095070512100397X>
- Boon-Itt, S. and Skunkan, Y. (2020). Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. *JMIR Public Health and Surveillance*, 6(4), p.e21978. <https://publichealth.jmir.org/2020/4/e21978>

- Briki, W. and Dagot, L. (2020). “Liberate!”–Republicans Are More Willing to Get Back to Social Life Because They Are Less Scared of COVID-19.
- Brooks, S.K., Webster, R.K., Smith, L.E., Woodland, L., Wessely, S., Greenberg, N. and Rubin, G.J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The lancet*, 395(10227), pp.912-920.
[https://doi.org/10.1016%2FS0140-6736\(20\)30460-8](https://doi.org/10.1016%2FS0140-6736(20)30460-8)
- Brown, R., Coventry, L. and Pepper, G. (2021). COVID-19: the relationship between perceptions of risk and behaviours during lockdown. *Journal of Public Health*.
- Budiharto, W. and Meiliana, M. (2018). Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. *Journal of Big data*, 5(1), pp.1-10.
- Bults, M., Beaujean, D.J., de Zwart, O., Kok, G., van Empelen, P., van Steenbergen, J.E., Richardus, J.H. and Voeten, H.A. (2011). Perceived risk, anxiety, and behavioural responses of the general public during the early phase of the Influenza A (H1N1) pandemic in the Netherlands: results of three consecutive online surveys. *BMC public health*, 11(1), pp.1-13. <https://doi.org/10.1186/1471-2458-11-2>
- Burgoon, J.K., Buller, D.B. and Woodall, W.G. (1989). *Nonverbal communication: The unspoken dialogue*. Harpercollins College Division.
- Burgoon, J.K., Buller, D.B. and Woodall, W.G. (1989). *Nonverbal communication: The unspoken dialogue*. Harpercollins College Division.
- Burkov, A. (2019). *The hundred-page machine learning book* (Vol. 1, p. 32). Quebec City, QC, Canada: Andriy Burkov.
- Buscaldi, D. and Hernandez-Farias, I. (2015). Sentiment analysis on microblogs for natural disasters management: a study on the 2014 genoa floodings. In *Proceedings of the 24th international conference on world wide web* (pp. 1185-1188).
<https://doi.org/10.1145/2740908.2741727>

- Cai, M., Luo, H. and Cui, Y. (2021). A Study on the Topic-Sentiment Evolution and Diffusion in Time Series of Public Opinion Derived from Emergencies. *Complexity*, 2021.
- Cambria, E., Das, D., Bandyopadhyay, S. and Feraco, A. (2017). Affective computing and sentiment analysis. In *A practical guide to sentiment analysis* (pp. 1-10). Springer, Cham.
- Cambria, E., Schuller, B., Xia, Y. and Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent systems*, 28(2), pp.15-21.
<https://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/deliver/index/docId/72618/file/72618.pdf>
- Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L. and Zdeborová, L. (2019). Machine learning and the physical sciences. *Reviews of Modern Physics*, 91(4), p.045002.
- Carreras, M., Vera, S. & Visconti, G. (2021). A Tale of Two Pandemics: Economic Inequality and Support for Containment Measures in Peru. *Journal of Politics in Latin America*, pp. 358-375.
- Carrillo-de-Albornoz, J. and Plaza, L. (2013). An emotion-based model of negation, intensifiers, and modality for polarity and intensity classification. *Journal of the American Society for Information Science and Technology*, 64(8), pp.1618-1633.
<https://doi.org/10.1002/asi.22859>
- Castellano-Tejedor, C., Torres-Serrano, M. and Cencerrado, A. (2022). Psychological impact in the time of COVID-19: A cross-sectional population survey study during confinement. *Journal of health psychology*, 27(4), pp.974-989.
<https://doi.org/10.1177/1359105320985580>

- Cattell, R.B. (1940). Sentiment or attitude? The core of a terminology problem in personality research. *Character & Personality; A Quarterly for Psychodiagnostic & Allied Studies*.
- Celebi, M.E. and Aydin, K. eds. (2016). Unsupervised learning algorithms (Vol. 9, p. 103). *Cham: Springer*.
- Chakraborty, B. and Banerjee, S. (2013). Modeling the evolution of post disaster social awareness from social web sites. In 2013 IEEE International Conference on Cybernetics (CYBCO) (pp. 51-56). IEEE.
<https://doi.org/10.1109/CYBConf.2013.6617444>
- Chandrasekaran, R., Mehta, V., Valkunde, T. and Moustakas, E. (2020). Topics, trends, and sentiments of tweets about the COVID-19 pandemic: Temporal infoveillance study. *Journal of medical Internet research*, 22(10), p.e 22624.
- Charniak, E., Hendrickson, C., Jacobson, N. and Perkowski, M. (1993). Equations for part-of-speech tagging. In AAAI (Vol. 11, pp. 784-789).
<https://www.aaai.org/Papers/AAAI/1993/AAAI93-117.pdf>
- Charquero-Ballester, M., Walter, J. G. & Bechmann, A. (2021). Different types of COVID-19 misinformation have different emotional valence on Twitter. *SAGE Journals*.
- Chatzakou, D., Koutsonikola, V., Vakali, A. and Kafetsios, K. (2013). Micro-blogging content analysis via emotionally-driven clustering. In 2013 humane association conference on affective computing and intelligent interaction (pp. 375-380). IEEE.
<https://datalab-old.csd.auth.gr/wp-content/uploads/publications/ACII2013.pdf>
- Chen, J., Liu, X., Wang, D., Jin, Y., He, M., Ma, Y., Zhao, X., Song, S., Zhang, L., Xiang, X. and Yang, L. (2021). Risk factors for depression and anxiety in healthcare workers deployed during the COVID-19 outbreak in China. *Social psychiatry and psychiatric epidemiology*, 56, pp.47-55.

- Chen, J., Lu, H., Melino, G., Boccia, S., Piacentini, M., Ricciardi, W., and Zhu, T. (2020). COVID-19 infection: the China and Italy perspectives. *Cell death & disease*, 11(6), 438.
- Chen, L.C., Lee, C.M. and Chen, M.Y. (2020). Exploration of social media for sentiment analysis using deep learning. *Soft Computing*, 24, pp.8187-8197.
- Chen, Y., Yuan, J., You, Q. and Luo, J. (2018). Twitter sentiment analysis via bi-sense emoji embedding and attention-based LSTM. In *Proceedings of the 26th ACM international conference on Multimedia* (pp. 117-125).
- Chen, Y.C., Cheng, J.Y. and Hsu, H.H. (2016). A cluster-based opinion leader discovery in social network. In *2016 conference on technologies and applications of artificial intelligence (TAAI)* (pp. 78-83). IEEE.
- Chew, N.W., Lee, G.K., Tan, B.Y., Jing, M., Goh, Y., Ngiam, N.J., Yeo, L.L., Ahmad, A., Khan, F.A., Shanmugam, G.N. and Sharma, A.K. (2020). A multinational, multicentre study on the psychological outcomes and associated physical symptoms amongst healthcare workers during COVID-19 outbreak. *Brain, behavior, and immunity*, 88, pp.559-565.
- Childers, T.L. and Houston, M.J., 1(984). Conditions for a picture-superiority effect on consumer memory. *Journal of consumer research*, 11(2), pp.643-654.
<http://links.jstor.org/sici?sici=0093-5301%28198409%2911%3A2%3C643%3ACFAPEO%3E2.0.CO%3B2-W>
- Chintalapudi, N., Battineni, G. and Amenta, F. (2021). Sentimental analysis of COVID-19 tweets using deep learning models. *Infectious Disease Reports*, 13(2), pp.329-339.
- Choudrie, J., Patil, S., Kotecha, K., Matta, N. and Pappas, I. (2021). Applying and understanding an advanced, novel deep learning approach: A Covid 19, text based,

- emotions analysis study. *Information Systems Frontiers*, 23(6), pp.1431-1465.
<https://link.springer.com/article/10.1007/s10796-021-10152-6>
- Chu, Y.-H. I., Alam, P., Larson, H. J. & Lin, L. (2020). Social consequences of mass quarantine during epidemics: a systematic review with implications for the COVID-19 response. *Journal of Travel Medicine*, 27(7).
- Conejero, I., Berrouiguet, S., Ducasse, D., Leboyer, M., Jardon, V., Olié, E. and Courtet, P. (2020). Suicidal behavior in light of COVID-19 outbreak: Clinical challenges and treatment perspectives. *Encephale*, pp. S66-S72.
- Cui, H., Mittal, V. and Datar, M. (2006). Comparative experiments on sentiment classification for online product reviews. In *AAAI* (Vol. 6, No. 1265-1270, p. 30).
<https://www.aaai.org/Papers/AAAI/2006/AAAI06-198.pdf>
- Dang, C.N., Moreno-García, M.N. and De la Prieta, F. (2021). Hybrid deep learning models for sentiment analysis. *Complexity*, 2021.
<https://www.hindawi.com/journals/complexity/2021/9986920/>
- Dangi, D., Dixit, D. K. & Bhagat, A. (2022). Sentiment analysis of COVID-19 social media data through machine learning. *Springer Link*, p. 42261–42283.
- Dankhara, D. (2022). A Review of Sentiment Analysis of Tweets. *Research Gate*.
- Das, S. and Chen, M. (2001). Yahoo! for Amazon: Extracting market sentiment from stock message boards. In *Proceedings of the Asia Pacific finance association annual conference (APFA)* (Vol. 35, p. 43).
- Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A.Y., Gelbukh, A. and Zhou, Q. (2016). Multilingual sentiment analysis: state of the art and independent comparison of techniques. *Cognitive computation*, 8, pp.757-771. <https://doi.org/10.1007/s12559-016-9415-7>

- Davidson, R.J., Sherer, K.R. and Goldsmith, H.H. eds. (2009). Handbook of affective sciences. Oxford University Press.
- Davis, B. et al. (2020). Quantifying Depression-Related Language on Social Media During the COVID-19 Pandemic. *IJPDS*, 5(4).
- de Las Heras-Pedrosa, C., Sánchez-Núñez, P. and Peláez, J.I. (2020). Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems. *International journal of environmental research and public health*, 17(15), p.5542.
- Denecke, K. and Deng, Y. (2015). Sentiment analysis in medical settings: New opportunities and challenges. *Artificial intelligence in medicine*, 64(1), pp.17-27.
- Denny, M.J. and Spirling, A. (2018). Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it. *Political Analysis*, 26(2), pp.168-189.
- Derks, D., Bos, A.E. and Von Grumbkow, J. (2008). Emoticons in computer-mediated communication: Social motives and social context. *Cyberpsychology & behavior*, 11(1), pp.99-101.
- Djuric, N., Zhou, J., Morris, R., Grbovic, M., Radosavljevic, V. and Bhamidipati, N. (2015). Hate speech detection with comment embeddings. In *Proceedings of the 24th international conference on world wide web* (pp. 29-30).
- Drus, Z. & Khalid, H. (2019). Sentiment Analysis in Social Media and Its Application: Systematic Literature Review. *Science Direct*, Volume 161, pp. 707-714.
- Duyu, T., Bing, Q., Ting, L. and Qiuhui, S. (2014). Emotion analysis platform on Chinese microblog. *CoRR journal*.

Eisner, B., Rocktäschel, T., Augenstein, I., Bošnjak, M. and Riedel, S. (2016). emoji2vec:

Learning emoji representations from their description.

<https://arxiv.org/pdf/1609.08359.pdf>

El Keshky, M.E.S., Basyouni, S.S. and Al Sabban, A.M. (2021). Corrigendum: Getting

Through COVID-19: The Pandemic's Impact on the Psychology of Sustainability,

Quality of Life, and the Global Economy—A Systematic Review. *Frontiers in*

Psychology, 12, p.700815. <https://doi.org/10.3389/fpsyg.2021.700815>

El Naqa, I. and Murphy, M.J. (2015). What is machine learning? (pp. 3-11). Springer

International Publishing.

El Rahman, S.A., AlOtaibi, F.A. and AlShehri, W.A. (2019). Sentiment analysis of twitter

data. In 2019 international conference on computer and information sciences

(ICCIS) (pp. 1-4). IEEE. <https://doi.org/10.1109/ICCISci.2019.8716464>

Elfaik, H. (2021). Deep bidirectional lstm network learning-based sentiment analysis for

arabic text. *Journal of Intelligent Systems*, 30(1), pp.395-412.

<https://www.degruyter.com/document/doi/10.1515/jisys-2020-0021/html>

Elghazaly, T., Mahmoud, A. and Hefny, H.A. (2016). Political sentiment analysis using

twitter data. In *Proceedings of the International Conference on Internet of things and*

Cloud Computing (pp. 1-5). [https://www.researchgate.net/profile/Amal-](https://www.researchgate.net/profile/Amal-Mahmud/publication/301591444_Political_Sentiment_Analysis_Using_Twitter_Data/links/5cc62fefa6fdcc1d49b75f92/Political-Sentiment-Analysis-Using-Twitter-Data.pdf)

[Mahmud/publication/301591444_Political_Sentiment_Analysis_Using_Twitter_Data/li](https://www.researchgate.net/profile/Amal-Mahmud/publication/301591444_Political_Sentiment_Analysis_Using_Twitter_Data/links/5cc62fefa6fdcc1d49b75f92/Political-Sentiment-Analysis-Using-Twitter-Data.pdf)

[nks/5cc62fefa6fdcc1d49b75f92/Political-Sentiment-Analysis-Using-Twitter-Data.pdf](https://www.researchgate.net/profile/Amal-Mahmud/publication/301591444_Political_Sentiment_Analysis_Using_Twitter_Data/links/5cc62fefa6fdcc1d49b75f92/Political-Sentiment-Analysis-Using-Twitter-Data.pdf)

Elghazaly, T., Mahmoud, A. and Hefny, H.A. (2016). Political sentiment analysis using

twitter data. In *Proceedings of the International Conference on Internet of things and*

Cloud Computing (pp. 1-5). [https://www.researchgate.net/profile/Amal-](https://www.researchgate.net/profile/Amal-Mahmud/publication/301591444_Political_Sentiment_Analysis_Using_Twitter_Data/links/5cc62fefa6fdcc1d49b75f92/Political-Sentiment-Analysis-Using-Twitter-Data.pdf)

[Mahmud/publication/301591444_Political_Sentiment_Analysis_Using_Twitter_Data/li](https://www.researchgate.net/profile/Amal-Mahmud/publication/301591444_Political_Sentiment_Analysis_Using_Twitter_Data/links/5cc62fefa6fdcc1d49b75f92/Political-Sentiment-Analysis-Using-Twitter-Data.pdf)

[nks/5cc62fefa6fdcc1d49b75f92/Political-Sentiment-Analysis-Using-Twitter-Data.pdf](https://www.researchgate.net/profile/Amal-Mahmud/publication/301591444_Political_Sentiment_Analysis_Using_Twitter_Data/links/5cc62fefa6fdcc1d49b75f92/Political-Sentiment-Analysis-Using-Twitter-Data.pdf)

- Eslami, M.S. (n.d) Sentiment Analysis Using CNN-LSTM Based on Emoji-Sense.
- Ettinger, A. (2020). What BERT is not: Lessons from a new suite of psycholinguistic diagnostics for language models. *Transactions of the Association for Computational Linguistics*, 8, pp.34-48.
- Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R. and Lin, C.J. (2008). LIBLINEAR: A library for large linear classification. *the Journal of machine Learning research*, 9, pp.1871-1874. <https://www.jmlr.org/papers/volume9/fan08a/fan08a.pdf>
- Fang, X. and Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, 2(1), pp.1-14.
- Farzad, A., Mashayekhi, H. and Hassanpour, H. (2019). A comparative performance analysis of different activation functions in LSTM networks for classification. *Neural Computing and Applications*, 31, pp.2507-2521.
- Felbo, B., Mislove, A., Søgaard, A., Rahwan, I. and Lehmann, S. (2017). Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. <https://arxiv.org/pdf/1708.00524.pdf>
- Fisher, M. & Bubola, E. (2020). As Coronavirus Deepens Inequality, Inequality Worsens Its Spread. *New York Times* , 16 March.
- Ford, B.Q., Feinberg, M., Lam, P., Mauss, I.B. and John, O.P. (2019). Using reappraisal to regulate negative emotion after the 2016 US Presidential election: Does emotion regulation trump political action?. *Journal of personality and social psychology*, 117(5), p.998. https://eerlab.berkeley.edu/pdf/papers/Ford_et_al_2019.pdf
- Fu, X., Liu, W., Xu, Y., Yu, C. and Wang, T. (2016). Long short-term memory network over rhetorical structure theory for sentence-level sentiment analysis. In *Asian conference on machine learning* (pp. 17-32). PMLR. <http://proceedings.mlr.press/v63/Fu62.pdf>

- Gao, S., He, L., Chen, Y., Li, D. and Lai, K. (2020). Public perception of artificial intelligence in medical care: content analysis of social media. *Journal of Medical Internet Research*, 22(7), p.e16649.
- Giachanou, A. and Crestani, F. (2016). Like it or not: A survey of twitter sentiment analysis methods. *ACM Computing Surveys (CSUR)*, 49(2), pp.1-41.
- Gondhi, N.K., Sharma, E., Alharbi, A.H., Verma, R. and Shah, M.A. (2022). Efficient Long Short-Term Memory-Based Sentiment Analysis of E-Commerce Reviews. *Computational Intelligence and Neuroscience*, 2022.
- Gross, J.J. and John, O.P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology*, 85(2), p.348. <https://psycnet.apa.org/doi/10.1037/0022-3514.85.2.348>
- Gross, J.J. (1998). Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of personality and social psychology*, 74(1), p.224.
<https://www.wisebrain.org/media/Papers/EmotionsAntecedencesResponsesReg.pdf>
- Guibon, G., Ochs, M. and Bellot, P. (2016). From emojis to sentiment analysis. In *WACAI 2016*.
- Guntuku, S.C., Li, M., Tay, L. and Ungar, L.H. (2019). Studying cultural differences in Emoji usage across the east and the west. In *Proceedings of the international AAAI conference on web and social media (Vol. 13, pp. 226-235)*.
- Hancke, B., Overbeke, T. V. & Voss, D. (2021). *Crisis and Complementarities: A Comparative Political Economy of Economic Policies after COVID-19*, s.l.: Cambridge University Press.
- Hania, A. and Hongbo, L., 2020. Psychological Predictors Of Anxiety In Response To The Covid-19 Pandemic: Evidence From Pakistan. <https://doi.org/10.30773/pi.2020.0167>

- Hao, J. and Dai, H. (2016). Social media content and sentiment analysis on consumer security breaches. *Journal of Financial Crime*, 23(4), pp.855-869. <https://doi.org/10.1108/JFC-01-2016-0001>
- Harper, C.A., Satchell, L.P., Fido, D. and Latzman, R.D. (2021). Functional fear predicts public health compliance in the COVID-19 pandemic. *International journal of mental health and addiction*, 19, pp.1875-1888. <https://doi.org/10.1007/s11469-020-00281-5>
- Hassan, A.U., Hussain, J., Hussain, M., Sadiq, M. and Lee, S. (2017). Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression. In 2017 international conference on information and communication technology convergence (ICTC) (pp. 138-140). IEEE.
http://uclab.khu.ac.kr/LeanUX/wp-content/uploads/2017/12/ICTC_2017_aneesh.pdf
- Hassan, T., Bajwa, I.S. and Hassan, S. (2016). Prediction of terrorist activities by using unsupervised learning techniques. *Journal of Applied and Emerging Sciences*, 6(2), pp.pp56-60. <http://dx.doi.org/10.36785/jaes.62190>
- Hearst, M.A. (1992). Direction-based text interpretation as an information access refinement. *Text-based intelligent systems: current research and practice in information extraction and retrieval*, pp.257-274.
<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d3be556530c0cfdda24d569fe71f038e1bde63>
- Heerschop, B., Hogenboom, A. and Frasincar, F. (2011). Sentiment lexicon creation from lexical resources. In *Business Information Systems: 14th International Conference, BIS 2011, Poznań, Poland, June 15-17, 2011. Proceedings 14* (pp. 185-196). Springer Berlin Heidelberg.

- Henriksen, B.E. (2020). A Multimodal Approach to Researching Social Media Practices: A Case Study of Style Materialised in Online Texts (Doctoral dissertation, UCL (University College London)).
- Hogenboom, A., Bal, D., Frasincar, F., Bal, M., De Jong, F. and Kaymak, U. (2013). Exploiting emoticons in sentiment analysis. In Proceedings of the 28th annual ACM symposium on applied computing (pp. 703-710).
<https://personal.eur.nl/frasincar/papers/SAC2013b/sac2013b.pdf>
- Hogenboom, A., Frasincar, F., De Jong, F. and Kaymak, U. (2015). Using rhetorical structure in sentiment analysis. *Communications of the ACM*, 58(7), pp.69-77.
<https://personal.eur.nl/hogenboom/files/cacm15-sentistruct.pdf>
- Hogenboom, A., Hogenboom, F., Kaymak, U., Wouters, P. and De Jong, F. (2010). Mining economic sentiment using argumentation structures. In *Advances in Conceptual Modeling–Applications and Challenges: ER 2010 Workshops ACM-L, CMLSA, CMS, DE@ ER, FP-UML, SeCoGIS, WISM*, Vancouver, BC, Canada, November 1-4, 2010. Proceedings 29 (pp. 200-209). Springer Berlin Heidelberg.
- Hooijer, G. & King, D. (2021). The Racialized Pandemic: Wave One of COVID-19 and the Reproduction of Global North Inequalities. Cambridge Core, 11 August.
- Hosseinzadeh, P., Zareipour, M., Baljani, E. and Moradali, M.R. (2022). Social Consequences of the COVID-19 Pandemic. A Systematic Review. *Investigación y Educación en Enfermería*, 40(1).
- Hota, H. S., Sharma, D. K. & Verma, N. (2020). Lexicon-based sentiment analysis using Twitter data: a case of COVID-19 outbreak in India and abroad. Elsevier, pp. 275-295.
- Howard, J. and Ruder, S. (2018). Universal language model fine-tuning for text classification. arXiv preprint arXiv:1801.06146.

- Hu, T., Guo, H., Sun, H., Nguyen, T.V.T. and Luo, J. (2017). Spice up your chat: the intentions and sentiment effects of using emojis. In Eleventh international AAAI conference on web and social media.
<https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/viewFile/15568/14798>
- Hu, X., Tang, J., Gao, H. and Liu, H. (2014). Social spammer detection with sentiment information. In 2014 IEEE international conference on data mining (pp. 180-189). IEEE.
- Huang, F., Li, X., Yuan, C., Zhang, S., Zhang, J. and Qiao, S. (2021). Attention-emotion-enhanced convolutional LSTM for sentiment analysis. *IEEE transactions on neural networks and learning systems*, 33(9), pp.4332-4345.
- Idoiaga, N., Berasategi, N., Eiguren, A. and Picaza, M. (2020). Exploring children's social and emotional representations of the COVID-19 pandemic. *Frontiers in Psychology*, 11, p.1952.
- Ikoro, V., Sharmina, M., Malik, K. and Batista-Navarro, R. (2018). Analyzing sentiments expressed on Twitter by UK energy company consumers. In 2018 Fifth international conference on social networks analysis, management and security (SNAMS) (pp. 95-98). IEEE. <https://doi.org/10.1109/SNAMS.2018.8554619>
- Imran, M., Elbassuoni, S., Castillo, C., Diaz, F. and Meier, P. (2013). Practical extraction of disaster-relevant information from social media. In Proceedings of the 22nd international conference on world wide web (pp. 1021-1024).
https://chato.cl/papers/imran_elbassuoni_castillo_diaz_meier_2013_practical_extraction_disaster_crisis.pdf
- Iosifidis, V. and Ntoutsi, E. (2017). Large scale sentiment learning with limited labels. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge

- discovery and data mining (pp. 1823-1832). <https://kbs.uni-hannover.de/~ntoutsis/papers/17.KDD.pdf>
- Iranmanesh, V., Wei, H.J., Dao-Ming, S.L. and Arigbabu, O.A. (2015). On using emoticons and lingoes for hiding data in SMS. In 2015 International Symposium on Technology Management and Emerging Technologies (ISTMET) (pp. 103-107). IEEE.
- Isah, H., Trundle, P. and Neagu, D. (2014). Social media analysis for product safety using text mining and sentiment analysis. In 2014 14th UK workshop on computational intelligence (UKCI) (pp. 1-7). IEEE.
- <https://arxiv.org/ftp/arxiv/papers/1510/1510.05301.pdf>
- Isik, M. & Dag, H. (2020). The impact of text preprocessing on the prediction of review ratings. *Turkish Journal of Electrical Engineering and Computer Sciences* , 28(3), pp. 1405-1421.
- Islam, J., Xiao, L. and Mercer, R.E.(2020). A lexicon-based approach for detecting hedges in informal text. In *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 3109-3113). <https://aclanthology.org/2020.lrec-1.380/>
- Jackson, L.M. (2019). *White Negroes: When Cornrows Were in Vogue... and Other Thoughts on Cultural Appropriation*. Beacon Press.
- Jalil, Z., Abbasi, A., Javed, A.R., Badruddin Khan, M., Abul Hasanat, M.H., Malik, K.M. and Saudagar, A.K.J. (2022). COVID-19 related sentiment analysis using state-of-the-art machine learning and deep learning techniques. *Frontiers in Public Health*, 9, p.2276.
- Janiesch, C., Zschech, P. and Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), pp.685-695.
- Jiang, H., Qiang, M. and Lin, P. (2016). Assessment of online public opinions on large infrastructure projects: A case study of the Three Gorges Project in China. *Environmental Impact Assessment Review*, 61, pp.38-51.

- Jimenez Rezende, D., Eslami, S.M., Mohamed, S., Battaglia, P., Jaderberg, M. and Heess, N. (2016). Unsupervised learning of 3d structure from images. *Advances in neural information processing systems*, 29.
- Jordan, M.I. and Mitchell, T.M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), pp.255-260.
- Joyce, B. and Deng, J. (2017). Sentiment analysis of tweets for the 2016 US presidential election. In *2017 IEEE MIT Undergraduate Research Technology Conference (URTC)* (pp. 1-4). IEEE. http://uncg.edu/cmp/faculty/j_deng/papers/sentiment_urtc17.pdf
- Jungmann, S.M. and Witthöft, M. (2020). Health anxiety, cyberchondria, and coping in the current COVID-19 pandemic: Which factors are related to coronavirus anxiety?. *Journal of anxiety disorders*, 73, p.102239.
<https://doi.org/10.1016%2Fj.janxdis.2020.102239>
- Kandasamy, V. et al. (2021). Sentimental Analysis of COVID-19 Related Messages in Social Networks by Involving an N-Gram Stacked Autoencoder Integrated in an Ensemble Learning Scheme. *MDPI*, 21(22).
- Karamollaoğlu, H., Doğru, İ.A., Dörterler, M., Utku, A. and Yıldız, O. (2018). Sentiment analysis on Turkish social media shares through lexicon-based approach. In *2018 3rd International Conference on Computer Science and Engineering (UBMK)* (pp. 45-49). IEEE. <https://doi.org/10.1109/UBMK.2018.8566481>
- Karita, S., Chen, N., Hayashi, T., Hori, T., Inaguma, H., Jiang, Z., Someki, M., Soplin, N.E.Y., Yamamoto, R., Wang, X. and Watanabe, S. (2019). A comparative study on transformer vs rnn in speech applications. In *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)* (pp. 449-456). IEEE.

- Kausar, M. b., Soosaimanickam, A. & Nasar, M. (2021). Public Sentiment Analysis on Twitter Data during COVID-19 Outbreak. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 12(2), pp. 415-421.
- Kavanaugh, A., Fox, E.A., Sheetz, S., Yang, S., Li, L.T., Whalen, T., Shoemaker, D., Natsev, P. and Xie, L. (2011). Social media use by government: From the routine to the critical. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (pp. 121-130).
- Kawohl, W. and Nordt, C. (2020). COVID-19, unemployment, and suicide. *The Lancet Psychiatry*, 7(5), pp.389-390.
- Kelly, R. and Watts, L. (2015). Characterising the inventive appropriation of emoji as relationally meaningful in mediated close personal relationships. Experiences of technology appropriation: Unanticipated users, usage, circumstances, and design, 2. https://purehost.bath.ac.uk/ws/files/130966701/emoji_relational_value.pdf
- Kendon, A. (1987). On gesture: Its complementary relationship with speech. *Nonverbal behavior and communication*, pp.65-97.
- Kirby, T. (2020). Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. *The Lancet Respiratory Medicine*, 8(6), pp.547-548. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7211498/>
- Kong, W., Dong, Z.Y., Jia, Y., Hill, D.J., Xu, Y. and Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE transactions on smart grid*, 10(1), pp.841-851.
- Koto, F. and Adriani, M. (2015). A comparative study on twitter sentiment analysis: Which features are good? In *Natural Language Processing and Information Systems: 20th International Conference on Applications of Natural Language to Information Systems*,

- NLDB 2015, Passau, Germany, June 17-19, 2015, Proceedings 20 (pp. 453-457). Springer International Publishing. https://doi.org/10.1007/978-3-319-19581-0_46
- Kouloumpis, E., Wilson, T. and Moore, J. (2011). Twitter sentiment analysis: The good the bad and the omg! In Proceedings of the international AAAI conference on web and social media (Vol. 5, No. 1, pp. 538-541). <https://doi.org/10.1609/icwsm.v5i1.14185>
- Kralj Novak, P., Smailović, J., Sluban, B. and Mozetič, I. (2015). Sentiment of emojis. PLoS one, 10(12), p.e0144296. <https://doi.org/10.1371/journal.pone.0144296>
- Krishnan, V. G., Rao, P. V., Deepa, J. & Divya, V. (2021). Twitter Sentiment Analysis using Ensemble Classifiers on Tamil and Malayalam Languages. Journal of Tianjin University Science and Technology , Volume 54, pp. 433-451.
- Kryvasheyeu, Y., Chen, H., Moro, E., Van Hentenryck, P. and Cebrian, M. (2015). Performance of social network sensors during Hurricane Sandy. PLoS one, 10(2), p.e0117288. <https://doi.org/10.1371/journal.pone.0117288>
- Kumar, D.A. and Chinnalagu, A. (2020). Sentiment and emotion in social media COVID-19 conversations: SAB-LSTM approach. In 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART) (pp. 463-467). IEEE.
- Kumar, S., Barbier, G., Abbasi, M. and Liu, H. (2011). Tweettracker: An analysis tool for humanitarian and disaster relief. In Proceedings of the international aaai conference on web and social media (Vol. 5, No. 1, pp. 661-662). <https://doi.org/10.1609/icwsm.v5i1.14079>
- Kutsar, D. and Kurvet-Käosaar, L. (2021). The Impact of the COVID-19 Pandemic on Families: Young People's Experiences in Estonia. Frontiers in Sociology, p.163.
- Larsen, D.A. and Wigginton, K.R. (2020). Tracking COVID-19 with wastewater. Nature Biotechnology, 38(10), pp.1151-1153. <https://doi.org/10.1038/s41587-020-0690-1>

- Latif, M.Z., Hussain, I., Saeed, R., Qureshi, M.A. and Maqsood, U. (2019). Use of smart phones and social media in medical education: trends, advantages, challenges and barriers. *Acta informatica medica*, 27(2), p.133.
- Le Glaz, A., Haralambous, Y., Kim-Dufoir, D.H., Lenca, P., Billot, R., Ryan, T.C., Marsh, J., Devylder, J., Walter, M., Berrouguet, S. and Lemey, C. (2021). Machine learning and natural language processing in mental health: systematic review. *Journal of Medical Internet Research*, 23(5), p.e15708.
- Lee, Y. et al. (2022). Social media use and mental health during the COVID-19 pandemic in young adults: a meta-analysis of 14 cross-sectional studies. *BMC Public Health*.
- Leung, G.M., Ho, L.M., Chan, S.K., Ho, S.Y., Bacon-Shone, J., Choy, R.Y., Hedley, A.J., Lam, T.H. and Fielding, R. (2005). Longitudinal assessment of community psychobehavioral responses during and after the 2003 outbreak of severe acute respiratory syndrome in Hong Kong. *Clinical Infectious Diseases*, 40(12), pp.1713-1720.
- Li, D., Rzepka, R., Ptaszynski, M. and Araki, K. (2020). HEMOS: A novel deep learning-based fine-grained humor detecting method for sentiment analysis of social media. *Information Processing & Management*, 57(6), p.102290.
- Li, H., Chen, Q., Zhong, Z., Gong, R. and Han, G. (2022). E-word of mouth sentiment analysis for user behavior studies. *Information Processing & Management*, 59(1), p.102784.
- Li, M., Chen, J., Chen, T. and Yuan, H. (2010). Probability for disaster chains in emergencies. *Journal of Tsinghua University Science and Technology*, 50(8), pp.1173-1177.

- Li, M., Ch'ng, E., Chong, A. and See, S. (2016). The new eye of smart city: novel citizen sentiment analysis in twitter. In 2016 International Conference on Audio, Language and Image Processing (ICALIP) (pp. 557-562). IEEE.
- Li, X., Yan, R. and Zhang, M. (2017). Joint emoji classification and embedding learning. In Web and Big Data: First International Joint Conference, APWeb-WAIM 2017, Beijing, China, July 7–9, 2017, Proceedings, Part II 1 (pp. 48-63). Springer International Publishing.
- Liakos, K.G., Busato, P., Moshou, D., Pearson, S. and Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), p.2674.
- Lin, C. and He, Y. (2009). Joint sentiment/topic model for sentiment analysis. In Proceedings of the 18th ACM conference on Information and knowledge management (pp. 375-384). <https://doi.org/10.1145/1645953.1646003>
- Lin, S.Y., Kung, Y.C. and Leu, F.Y. (2022). Predictive intelligence in harmful news identification by BERT-based ensemble learning model with text sentiment analysis. *Information Processing & Management*, 59(2), p.102872.
- Lin, Y. (2022). Using Deep Learning Techniques in Forecasting Stock Markets by Hybrid Data with Multilingual Sentiment Analysis. [Online] Available: <https://www.mdpi.com/2079-9292/11/21/3513>
- Lindsay, B.R. (2011). Social media and disasters: Current uses, future options, and policy considerations. <https://nsi.org/ReferenceLibrary/983.pdf>
- Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies. Morgan and Claypool Publishers, San Rafael.
- Liu, B. (2020). Sentiment analysis: Mining opinions, sentiments, and emotions. Cambridge university press.

- Liu, C. (2021). Improving sentiment analysis accuracy with emoji embedding. [Online]
Available:
https://www.researchgate.net/publication/355432067_Improving_sentiment_analysis_accuracy_with_emoji_embedding
- Liu, C., Fang, F., Lin, X., Cai, T., Tan, X., Liu, J., and Lu, X. (2021). Improving sentiment analysis accuracy with emoji embedding. *Journal of Safety Science and Resilience*, 2(4), pp.246-252.
- Liu, C., Tan, X., Zhou, T., Zhang, W., Liu, J., & Lu, X. (2022). Emoji use in China: popularity patterns and changes due to COVID-19. *Applied Intelligence*, 1-11.
- Liu, J., Shen, D., Zhang, Y., Dolan, B., Carin, L. and Chen, W. (2021). What Makes Good In-Context Examples for GPT-3? *arXiv preprint arXiv:2101.06804*.
- Liu, K.L., Li, W.J. and Guo, M. (2012). Emoticon smoothed language models for twitter sentiment analysis. In the Twenty-sixth AAAI conference on artificial intelligence.
<https://www.aaai.org/ocs/index.php/AAAI/AAAI12/paper/viewFile/5083/5319>
- Liu, W., Wu, X., Yang, W., Zhu, X., and Zhong, S. (2019). Modeling cyber rumour spreading over mobile social networks: A compartment approach. *Applied Mathematics and Computation*, 343, pp.214-229.
- Lo, S.L., Cambria, E., Chiong, R. and Cornforth, D. (2017). Multilingual sentiment analysis: from formal to informal and scarce resource languages. *Artificial Intelligence Review*, 48, pp.499-527.
https://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=5878&context=sis_research
- Lou, Y., Zhang, Y., Li, F., Qian, T., and Ji, D. (2020). Emoji-based sentiment analysis using attention networks. *ACM Transactions on Asian and low-resource language information processing (TALLIP)*, 19(5), pp.1-13.

- Lu, H., Ehwerhemuepha, L. and Rakovski, C. (2022). A comparative study on deep learning models for text classification of unstructured medical notes with various levels of class imbalance. *BMC Medical Research Methodology*, 22(1), p.181.
- Lu, Y., Hu, X., Wang, F., Kumar, S., Liu, H. and Maciejewski, R. (2015). Visualizing social media sentiment in disaster scenarios. In Proceedings of the 24th international conference on world wide web (pp. 1211-1215).
<https://doi.org/10.1145/2740908.2741720>
- Lu, Y., Pan, J. & Xu, Y. (2021). Public Sentiment on Chinese Social Media during the Emergence of COVID-19. *Journal of Quantitative Description*, Volume 1.
- Lu, Y., Tsaparas, P., Ntoulas, A. and Polanyi, L. (2010). Exploiting social context for review quality prediction. In Proceedings of the 19th international conference on World wide web (pp. 691-700). <https://doi.org/10.1145/1772690.1772761>
- Lwin, M. O. et al. (2022). The Evolution of Public Sentiments During the COVID-19 Pandemic: Case Comparisons of India, Singapore, South Korea, the United Kingdom, and the United States. *JMIR Publications*, 2(1).
- Ma, K. C. (2021). Modeling the impact of racial and ethnic disparities on COVID-19 epidemic dynamics. *Epidemiology and Global Health*.
- Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y. and Wang, Y. (2017). Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction. *Sensors*, 17(4), p.818.
- MacEachren, A.M., Robinson, A.C., Jaiswal, A., Pezanowski, S., Savelyev, A., Blanford, J. and Mitra, P. (2011). Geo-twitter analytics: Applications in crisis management. In 25th international cartographic conference (pp. 3-8).
- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9, pp.381-386.

- Mandel, B., Culotta, A., Boulahanis, J., Stark, D., Lewis, B. and Rodrigue, J. (2012). A demographic analysis of online sentiment during hurricane irene. In Proceedings of the second workshop on language in social media (pp. 27-36).
<https://aclanthology.org/W12-2104.pdf>
- Manguri, K.H., Ramadhan, R.N. and Amin, P.R.M. (2020). Twitter sentiment analysis on worldwide COVID-19 outbreaks. *Kurdistan Journal of Applied Research*, pp.54-65.
<https://doi.org/10.24017/covid.8>
- Mansour, R., Hady, M.F.A., Hosam, E., Amr, H. and Ashour, A. (2015). Feature selection for twitter sentiment analysis: An experimental study. In *Computational Linguistics and Intelligent Text Processing: 16th International Conference, CICLing 2015, Cairo, Egypt, April 14-20, 2015, Proceedings, Part II 16* (pp. 92-103). Springer International Publishing. https://doi.org/10.1007/978-3-319-18117-2_7
- Mansour, S. (2018). Social media analysis of user's responses to terrorism using sentiment analysis and text mining. *Procedia Computer Science*, 140, pp.95-103.
<https://doi.org/10.1016/j.procs.2018.10.297>
- Martin-Domingo, L., Martín, J.C. and Mandsberg, G. (2019). Social media as a resource for sentiment analysis of Airport Service Quality (ASQ). *Journal of Air Transport Management*, 78, pp.106-115. <https://doi.org/10.1016/j.jairtraman.2019.01.004>
- Martins Van Jaarsveld, G. (2020). The effects of COVID-19 among the elderly population: a case for closing the digital divide. *Frontiers in psychiatry*, p.1211.
https://www.frontiersin.org/articles/10.3389/fpsyt.2020.577427/full?fbclid=IwAR1AsSQQr7RhGiisqVIZ9jdbMJueuTbf7BC_RSazIuStRSJpBbi6k7P6lH8
- Mauss, I.B., Cook, C.L., Cheng, J.Y. and Gross, J.J. (2007). Individual differences in cognitive reappraisal: Experiential and physiological responses to an anger provocation. *International Journal of Psychophysiology*, 66(2), pp.116-124.

https://eerlab.berkeley.edu/pdf/papers/2007_Mauss_Individual%20differences%20in%20cognitive%20reappraisal.pdf

- McCann, G. & Matenga, C. (2022). COVID-19 and Global Inequality. COVID-19 in the Global South, pp. 161-171.
- McGowan, V. J. & Bambara, C. (2022). COVID-19 mortality and deprivation: pandemic, syndemic, and endemic health inequalities. *The Lancet Public Health*, 7(11), pp. 966-975.
- Medford, R.J., Saleh, S.N., Sumarsono, A., Perl, T.M. and Lehmann, C.U. (2020). An "infodemic": leveraging high-volume Twitter data to understand early public sentiment for the coronavirus disease 2019 outbreak. In *Open forum infectious diseases* (Vol. 7, No. 7, p. ofaa258). US: Oxford University Press.
- Medhat, W., Hassan, A. and Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4), pp.1093-1113.
<https://doi.org/10.1016/j.asej.2014.04.011>
- Meier, K. et al. (2020). Public perspectives on protective measures during the COVID-19 pandemic in the Netherlands, Germany, and Italy: A survey study. *PLOS ONE*.
- Mendoza, M., Poblete, B. and Castillo, C. (2010). Twitter under crisis: Can we trust what we RT? In *Proceedings of the first workshop on social media analytics* (pp. 71-79).
http://snap.stanford.edu/soma2010/papers/soma2010_11.pdf
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.
<https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf>

- Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.
- <https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf>
- Miller, H.J., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L. and Hecht, B. (2016). “Blissfully happy” or “ready to fight”: Varying interpretations of emoji. In the Tenth international AAAI conference on web and social media.
- <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/viewFile/13167/12746>
- Mohammed, M., Khan, M.B. and Bashier, E.B.M. (2016). *Machine learning: algorithms and applications*. CRC Press.
- Mohd Hanafiah, K., Ng, C. and Wan, A.M. (2021). Effective communication at different phases of COVID-19 prevention: Roles, enablers and barriers. *Viruses*, 13(6), p.1058.
- Munezero, M., Montero, C.S., Sutinen, E. and Pajunen, J. (2014). Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE transactions on affective computing*, 5(2), pp.101-111. <https://doi.org/10.1109/TAFFC.2014.2317187>
- Naderalvojud, B. and Sezer, E.A. (2020). Sentiment-aware word embeddings using refinement and senti-contextualized learning approach. *Neurocomputing*, 405, pp.149-160.
- Nagy, A. and Stamberger, J.A. (2012). Crowd sentiment detection during disasters and crises. In *ISCRAM*.
- Naseem, U. (2020). *Hybrid Words Representation for the classification of low-quality text* (Doctoral dissertation).

- Nasukawa, T. and Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. In Proceedings of the 2nd international conference on Knowledge capture (pp. 70-77). <http://nactem.ac.uk/files/workshop06/TAKMIandSA2006.pdf>
- Navigli, R. (2009). Word sense disambiguation: A survey. *ACM computing surveys (CSUR)*, 41(2), pp.1-69.
https://www.cis.lmu.de/~fraser/wsd_mt_2014_seminar/ACM_Survey_2009_Navigli.pdf
- Nguyen, D. T., Tran, V. D. & Ghafoor, A. (2022). The Impact of COVID-19 lockdown on Intention to Follow Preventive Measures in Vietnam: Integrated Protection Motivation Theory and Theory Planned Behavior. *Cogent Business & Management* , 9(1).
- Nielsen, F.Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. <https://arxiv.org/pdf/1103.2903.pdf>
- Ning, L., Niu, J., Bi, X., Yang, C., Liu, Z., Wu, Q., Ning, N., Liang, L., Liu, A., Hao, Y. and Gao, L. (2020). The impacts of knowledge, risk perception, emotion and information on citizens' protective behaviours during the outbreak of COVID-19: a cross-sectional study in China. *BMC public health*, 20(1), pp.1-12.
<https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-020-09892-y>
- Nshakira-Rukundo, E. & Whitehead, A. (2021). Changing perceptions about COVID-19 risk and adherence to preventive strategies in Uganda: Evidence from an online mixed-methods survey. *Elsevier Public Health Emergency Collection*.
- Nwosu, C. O. & Oyenubi, A. (2021). Income-related health inequalities associated with the coronavirus pandemic in South Africa: A decomposition analysis. *International Journal for Equity in Health*.
- Obembe, D., Kolade, O., Obembe, F., Owoseni, A. and Mafimisebi, O. (2021). COVID-19 and the tourism industry: An early stage sentiment analysis of the impact of social

- media and stakeholder communication. *International Journal of Information Management Data Insights*, 1(2), p.100040.
<https://doi.org/10.1016/j.jjime.2021.100040>
- Oosterhoff, B. and Palmer, C.A. (2020). Psychological correlates of news monitoring, social distancing, disinfecting, and hoarding behaviours among US adolescents during the COVID-19 pandemic. <https://doi.org/10.31234/osf.io/rpcy4>
- Ortigosa, A., Martín, J.M. and Carro, R.M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in human behavior*, 31, pp.527-541.
- Ouyang, X., Zhou, P., Li, C.H. and Liu, L. (2015). Sentiment analysis using convolutional neural network. In *2015 IEEE international conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing* (pp. 2359-2364). IEEE.
- Pak, A. and Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (Vol. 10, No. 2010, pp. 1320-1326).
https://lexitron.nectec.or.th/public/LREC-2010_Malta/pdf/385_Paper.pdf
- Paltoglou, G. and Thelwall, M. (2010). A study of information retrieval weighting schemes for sentiment analysis. In *Proceedings of the 48th annual meeting of the association for computational linguistics* (pp. 1386-1395). <https://aclanthology.org/P10-1141.pdf>
- Pang, B. and Lee, L. (2005). Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. <https://arxiv.org/pdf/cs/0506075.pdf>
- Pang, B. and Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends®*, in information retrieval, 2(1–2), pp.1-135.
<http://dx.doi.org/10.1561/15000000011>
- Pang, B., Lee, L. and Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. <https://doi.org/10.3115/1118693.1118704>

- Parajuli, R. R. et al. (2020). Multidisciplinary approach to COVID-19 risk communication: a framework and tool for individual and regional risk assessment. NCBI.
- Park, J.H., Xu, P. and Fung, P. (2018). Plusemo2vec at semeval-2018 task 1: Exploiting emotion knowledge from Emoji an
- Patel, J. A. et al. (2020). Poverty, inequality, and COVID-19: the forgotten vulnerable. Elsevier Public Health Emergency Collection, pp. 110-111.
- Pavalanathan, U. and Eisenstein, J. (2015). Emoticons vs. emojis on Twitter: A causal inference approach. <https://arxiv.org/pdf/1510.08480.pdf>
- Peng, H., Cambria, E. and Hussain, A. (2017). A review of sentiment analysis research in the Chinese language. *Cognitive Computation*, 9, pp.423-435.
- Perry, B. L., Aronson, B. & Pescosolido, B. A. (2020). Pandemic precarity: COVID-19 is exposing and exacerbating inequalities in the American heartland. *PNAS*.
- Poecze, F., Ebster, C. and Strauss, C. (2018). Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts. *Procedia computer science*, 130, pp.660-666. <https://doi.org/10.1016/j.procs.2018.04.117>
- Pokharel, B.P. (2020). Twitter sentiment analysis during the COVID-19 outbreak in Nepal. Available at SSRN 3624719. <https://dx.doi.org/10.2139/ssrn.3624719>
- Poria, S., Cambria, E., Winterstein, G. and Huang, G.B. (2014). Sentic patterns: Dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems*, 69, pp.45-63.
- Pozzi, F.A., Maccagnola, D., Fersini, E. and Messina, E. (2013). Enhance user-level sentiment analysis on microblogs with approval relations. In *AI* IA 2013: Advances in Artificial Intelligence: XIIIth International Conference of the Italian Association for Artificial Intelligence*, Turin, Italy, December 4-6, 2013. *Proceedings 13* (pp. 133-144). Springer International Publishing. https://doi.org/10.1007/978-3-319-03524-6_12

- Prasad, A.G., Sanjana, S., Bhat, S.M. and Harish, B.S. (2017). Sentiment analysis for sarcasm detection on streaming short text data. In *2017, the 2nd International Conference on Knowledge Engineering and Applications (ICKEA)* (pp. 1-5). IEEE.
- Priya, B.G. (2019). Emoji-based sentiment analysis using KNN. *International Journal of Scientific Research and Review*, 7(4), pp.859-865.
- Qiu, G., He, X., Zhang, F., Shi, Y., Bu, J. and Chen, C. (2010). DASA: dissatisfaction-oriented advertising based on sentiment analysis. *Expert Systems with Applications*, 37(9), pp.6182-6191. <https://doi.org/10.1016/j.eswa.2010.02.109>
- Qu, Y., Huang, C., Zhang, P. and Zhang, J. (2011). Microblogging after a major disaster in China: a case study of the 2010 Yushu earthquake. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 25-34). <https://doi.org/10.1145/1958824.1958830>
- Qu, Y., Wu, P.F. and Wang, X. (2009). Online community response to major disaster: A study of Tianya forum in the 2008 Sichuan earthquake. In *2009, the 42nd Hawaii International Conference on System Sciences* (pp. 1-11). IEEE. <https://doi.org/10.1109/HICSS.2009.330>
- Raamkumar, A.S., Tan, S.G. and Wee, H.L. (2020). Measuring the outreach efforts of public health authorities and the public response on Facebook during the COVID-19 pandemic in early 2020: cross-country comparison. *Journal of medical Internet research*, 22(5), p.e19334.
- Ragini, J.R., Anand, P.R. and Bhaskar, V. (2018). Big data analytics for disaster response and recovery through sentiment analysis. *International Journal of Information Management*, 42, pp.13-24. <https://wiki.app.uib.no/info319/images/2/2f/P5-2019.pdf>
- Rahat, A.M., Kahir, A. and Masum, A.K.M. (2019). Comparison of Naive Bayes and SVM Algorithm based on sentiment analysis using review dataset. In *2019, the 8th*

- International Conference System Modeling and Advancement in Research Trends (SMART) (pp. 266-270). IEEE.
- Raschka, S. (2015). Python machine learning. Packt Publishing Ltd.
- Rasul, G. et al. (2021). Socio-Economic Implications of COVID-19 Pandemic in South Asia: Emerging Risks and Growing Challenges. *Frontiers in Sociology*.
- Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In *Proceedings of the ACL student research workshop* (pp. 43-48). <https://aclanthology.org/P05-2008.pdf>
- Ren, Y., Wang, R. and Ji, D. (2016). A topic-enhanced word embedding for Twitter sentiment classification. *Information Sciences*, 369, pp.188-198.
<https://doi.org/10.1016/j.ins.2016.06.040>
- Restubog, S.L.D., Ocampo, A.C.G. and Wang, L. (2020). Taking control amidst the chaos: Emotion regulation during the COVID-19 pandemic. *Journal of Vocational Behavior*, 119, p.103440.
- Rezabek, L. and Cochenour, J.(1998). Visual cues in computer-mediated communication: Supplementing text with emoticons. *Journal of visual literacy*, 18(2), pp.201-215.
- Ribeiro, F.N., Araújo, M., Gonçalves, P., André Gonçalves, M. and Benevenuto, F. (2016). Sentibench-a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science*, 5, pp.1-29.
- Ritter, A., Clark, S. and Etzioni, O. (2011). Named entity recognition in tweets: an experimental study. In *Proceedings of the 2011 conference on empirical methods in natural language processing* (pp. 1524-1534). <https://aclanthology.org/D11-1141.pdf>
- Ritter, A., Etzioni, O. and Clark, S. (2012). Open domain event extraction from Twitter. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge*

discovery and data mining (pp. 1104-1112).

<https://theory.utdallas.edu/seminar/G2S13/CC/rt080-ritter.pdf>

Rocha, M. d. L. R. (2021). COVID-19 Pandemic Effects on United Kingdom's Stock Market. MEGI.

Rocha, R. et al. (2021). Effect of socioeconomic inequalities and vulnerabilities on health-system preparedness and response to COVID-19 in Brazil: a comprehensive analysis. *The Lancet Global Health*, 9(6), pp. 782-792.

Romano, S. D. et al. (2021). Morbidity and Mortality Weekly Report. CDC, 70(15), p. 560–565.

Russell, J.A. and Barrett, L.F. (1999). Core affect prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology*, 76(5), p.805. http://www.ozrp-students.narod.ru/0910/0910_2/vo_4_spe/russell-barrett-1999.pdf

Saif, H., Fernandez, M., He, Y. and Alani, H. (2014). On stop words, filtering, and data sparsity for sentiment analysis of Twitter.

Saif, H., He, Y., and Alani, H. (2012). Semantic Sentiment Analysis of Twitter. In: *The Semantic Web–ISWC 2012*. 11th. Boston. November 11–15, 2012.

Saif, H., He, Y., Fernandez, M. and Alani, H. (2016). Contextual semantics for sentiment analysis of Twitter. *Information Processing & Management*, 52(1), pp.5-19. <https://doi.org/10.1016/j.ipm.2015.01.005>

Saigal, S., Gupta, S., Sudhindran, S., Goyal, N., Rastogi, A., Jacob, M., & Soin, A. S. (2020). Liver transplantation and COVID-19 (Coronavirus) infection: guidelines of the liver transplant Society of India (LTSI). *Hepatology international*, 14, 429-431.

Sakaki, T., Matsuo, Y., Kurihara, S., Toriumi, F., Shinoda, K., Noda, I., Uchiyama, K. and Kazama, K. (2013). The possibility of social media analysis for disaster management.

- In 2013 IEEE Region 10 Humanitarian Technology Conference (pp. 238-243). IEEE.
<https://doi.org/10.1109/R10-HTC.2013.6669048>
- Saladino, V., Algeri, D. and Auriemma, V. (2020). The psychological and social impact of COVID-19: new perspectives of well-being. *Frontiers in psychology*, p.2550.
<https://www.frontiersin.org/articles/10.3389/fpsyg.2020.577684/full?ltclid=>
- Samadipour, E., Ghardashi, F., Nazarikamal, M. and Rakhshani, M. (2022). Perception risk, preventive behaviors and assessing the relationship between their various dimensions: A cross-sectional study in the COVID-19 peak period. *International Journal of Disaster Risk Reduction*, 77, p.103093.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9174951/>
- Sciandra, A. (2020). COVID-19 outbreak through Tweeters' words: monitoring Italian social media communication about COVID-19 with Text Mining and Word embedding.
 In *2020 IEEE Symposium on Computers and Communications (ISCC)* (pp. 1-6). IEEE.
- Scott, S.R., Rivera, K.M., Rushing, E., Manczak, E.M., Rozek, C.S. and Doom, J.R (2021). "I Hate This": A Qualitative Analysis of Adolescents' Self-Reported Challenges During the COVID-19 Pandemic. *Journal of Adolescent Health*, 68(2), pp.262-269.
- Sebastiani, F. and Esuli, A. (2006). Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th International Conference on Language Resources and Evaluation* (pp. 417-422). European Language Resources Association (ELRA) Genoa, Italy.
- Serrano-Guerrero, J., Olivas, J.A., Romero, F.P. and Herrera-Viedma, E. (2015). Sentiment analysis: A review and comparative analysis of web services. *Information Sciences*, 311, pp.18-38.
- Severyn, A. and Moschitti, A. (2015). Twitter sentiment analysis with deep convolutional neural networks. In *Proceedings of the 38th international ACM SIGIR conference on*

- research and development in information retrieval (pp. 959-962).
<https://doi.org/10.1145/2766462.2767830>
- Shang, Y., Li, H. & Zhang, R. (2021). Effects of Pandemic Outbreak on Economies: Evidence From Business History Context. *Frontiers in Public Health*.
- Shannon S, Kent N. (2020). 8 charts on internet use around the world as countries grapple with COVID-19 Internet. Pew Research Center. URL: <https://www.pewresearch.org/fact-tank/2020/04/02/8-charts-on-internet-use-around-the-world-as-countries-grapple-with-COVID-19/>
- Shayaa, S., Wai, P.S., Chung, Y.W., Sulaiman, A., Jaafar, N.I. and Zakaria, S.B. (2017). Social media sentiment analysis on employment in Malaysia. In the Proceedings of 8th Global Business and Finance Research Conference, Taipei, Taiwan.
- Shekerdemian, L. S., Mahmood, N. R., Wolfe, K. K., Riggs, B. J., Ross, C. E., McKiernan, C. A., & Burns, J. P. (2020). Characteristics and outcomes of children with coronavirus disease 2019 (COVID-19) infection admitted to US and Canadian pediatric intensive care units. *JAMA pediatrics*, 174(9), 868-873.
- Shepard, R.N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and verbal Behavior*, 6(1), pp.156-163.
- Shifa, M., David, A. & Leibbrandt, M. (2021). Spatial inequality through the prism of a pandemic: COVID-19 in South Africa. NCBI.
- Shiha, M. and Ayvaz, S. (2017). The effects of emoji in sentiment analysis. *Int. J. Comput. Electr. Eng.(IJCEE.)*, 9(1), pp.360-369.
- Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H. and Deng, X. (2013). Exploiting topic-based Twitter sentiment for stock prediction. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 24-29). <https://aclanthology.org/P13-2005.pdf>

- Siddharth, S., Darsini, R. and Sujithra, M. (2018). Sentiment analysis on Twitter data using machine learning algorithms in Python. *Int. J. Eng. Res. Comput. Sci. Eng*, 5(2), pp.285-290.
- Siew, S. K., Chia, J. L., Mahendran, R. & Yu, J. (2022). Older adults' perceptions of government handling of COVID-19: Predictors of protective behaviours from lockdown to post-lockdown. *PLOS ONE*.
- Simon, T., Goldberg, A., Aharonson-Daniel, L., Leykin, D. and Adini, B. (2014). Twitter in the crossfire—the use of social media in the Westgate Mall terror attack in Kenya. *PloS one*, 9(8), p.e104136. <https://doi.org/10.1371/journal.pone.0104136>
- Singh, A., Blanco, E. and Jin, W. (2019). Incorporating emoji descriptions improves tweet classification. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 2096-2101).
- Singh, R., Singh, R. and Bhatia, A. (2018). Sentiment analysis using Machine Learning techniques to predict outbreaks and epidemics. *Int. J. Adv. Sci. Res*, 3(2), pp.19-24. <https://www.researchgate.net/profile/Rajeshwar-Singh-3/publication/332819639>
- Singh, S.K., Verma, P. and Kumar, P. (2020). Sentiment Analysis Using Machine Learning Techniques on Twitter: A Critical Review. *Advances in Mathematics: Scientific Journal*, 9(9), pp.7085-7092.
- Singla, Z., Randhawa, S. and Jain, S. (2017). Statistical and sentiment analysis of consumer product reviews. In *2017, the 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.
- Smith, A.M., Willroth, E.C., Gatchpazian, A., Shallcross, A.J., Feinberg, M. and Ford, B.Q. (2021). Coping with health threats: The costs and benefits of managing

emotions. *Psychological science*, 32(7), pp.1011-1023.

<https://journals.sagepub.com/doi/pdf/10.1177/09567976211024260>

Socher, R., Pennington, J., Huang, E.H., Ng, A.Y. and Manning, C.D. (2011). Semi-supervised recursive autoencoders for predicting sentiment distributions.

In *Proceedings of the 2011 conference on empirical methods in natural language processing* (pp. 151-161). <https://aclanthology.org/D11-1014.pdf>

Sokolova, M. and Lapalme, G. (2011). Learning opinions in user-generated web content. *Natural Language Engineering*, 17(4), pp.541-567.

<http://atour.iro.umontreal.ca/rali/sites/default/files/publis/SokolovaNLE2011.pdf>

Srijith, P.K., Hepple, M., Bontcheva, K. and Preotiuc-Pietro, D. (2017). Sub-story detection in Twitter with hierarchical Dirichlet processes. *Information Processing & Management*, 53(4), pp.989-1003.

Srivastava, N., Mansimov, E. and Salakhudinov, R. (2015). Unsupervised learning of video representations using LSTMS. In *International conference on machine learning* (pp. 843-852). PMLR.

Starbird, K., Palen, L., Hughes, A.L. and Vieweg, S. (2010). Chatter on the red: what hazards threat reveals about the social life of microblogged information. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 241-250).

<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=c9c23d05db7274582ca5c0fc6ca98ec5f98d2de9>

Stella, M., Restocchi, V. and De Deyne, S. (2020). # lockdown: network-enhanced emotional profiling in the time of COVID-19. *Big Data and Cognitive Computing*, 4(2), p.14.

<https://doi.org/10.3390/bdcc4020014>

Stine, R.A. (2019). Sentiment analysis. *Annual review of statistics and its application*, 6, pp.287-308.

- Sudhir, P. and Suresh, V.D. (2021). Comparative study of various approaches, applications, and classifiers for sentiment analysis. *Global Transitions Proceedings*, 2(2), pp.205-211.
- Sulthana, A. R., Jaithunbi, A. K. & Ramesh, L. S. (2018). Sentiment analysis in Twitter data using data analytic techniques for predictive modeling. *Journal of Physics: Conference Series*.
- Suman, N., Gupta, P.K. and Sharma, P. (2017). Analysis of stock price flow based on social media sentiments. In 2017 International Conference on Next Generation Computing and Information Systems (ICNGCIS) (pp. 54-57). IEEE.
<https://doi.org/10.1109/ICNGCIS.2017.34>
- Suratnoaji, C., Nurhadi & Arianto, I. D. (2020). Public Opinion on Lockdown (PSBB) Policy in Overcoming COVID-19 Pandemic in Indonesia: Analysis Based on Big Data Twitter. *Asian Journal for Public Opinion Research*, 8(3), pp. 393-406.
- Surdeanu, M., Hicks, T. and Valenzuela-Escarcega, M.A. (2015). Two practical rhetorical structure theory parsers. In Proceedings of the 2015 conference of the North American chapter of the Association for Computational Linguistics: Demonstrations (pp. 1-5).
<http://www.aclweb.org/anthology/N/N15/N15-3001.pdf>
- Surdeanu, M., Hicks, T. and Valenzuela-Escarcega, M.A. (2015). Two practical rhetorical structure theory parsers. In Proceedings of the 2015 conference of the North American chapter of the Association for Computational Linguistics: Demonstrations (pp. 1-5).
<http://www.aclweb.org/anthology/N/N15/N15-3001.pdf>
- Surikov, A. and Egorova, E. (2020). Alternative method sentiment analysis using emojis and emoticons. *Procedia Computer Science*, 178, pp.182-193.
- Sutton, J.N., Palen, L. and Shklovski, I. (2008). Backchannels on the front lines: Emergency uses of social media in the 2007 Southern California Wildfires.

- Szreter, S. and Woolcock, M. (2004). Health by association? Social capital, social theory, and the political economy of public health. *International journal of epidemiology*, 33(4), pp.650-667.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), pp.267-307.
https://doi.org/10.1162/COLI_a_00049
- Taboada, M., Voll, K. and Brooke, J. (2008). Extracting sentiment as a function of discourse structure and topicality. Simon Fraser University School of Computing Science Technical Report, pp.1-22.
- Taleby Ahvanooey, M., Li, Q., Hou, J., Rajput, A.R. and Chen, Y. (2019). Modern text hiding, text steganalysis, and applications: a comparative analysis. *Entropy*, 21(4), p.355.
- Tan, H. et al. (2021). Long-term Effects of the COVID-19 Pandemic on Public Sentiments in Mainland China: Sentiment Analysis of Social Media Posts. *Journal of Medical Internet Research*, 23(8).
- Tan, C., Lee, L., Tang, J., Jiang, L., Zhou, M. and Li, P. (2011). User-level sentiment analysis incorporating social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1397-1405).
<https://arxiv.org/pdf/1109.6018.pdf>
- Tan, M. K. (2021). COVID-19 in an inequitable world: the last, the lost, and the least. *International Health Journal*, 13(6), pp. 493-496.
- Tang, D., Wei, F., Qin, B., Liu, T., and Zhou, M. (2014a). Coooolll: A deep learning system for Twitter sentiment classification. In *Proceedings of the 8th International Workshop on semantic evaluation (SemEval 2014)* (pp. 208-212). <https://aclanthology.org/S14-2033.pdf>

- Tang, D., Wei, F., Qin, B., Yang, N., Liu, T. and Zhou, M. (2015). Sentiment embeddings with applications to sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, 28(2), pp.496-509.
- Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T. and Qin, B. (2014b). Learning sentiment-specific word embedding for Twitter sentiment classification. In *ACL* (1) (pp. 1555-1565). <https://doi.org/10.3115/v1/p14-1146>
- Terpstra, T., Stronkman, R., de Vries, A. and Paradies, G.L. (2012). Towards a real-time Twitter analysis during crises for operational crisis management. In *Is cram*. https://www.hkv.nl/wp-content/uploads/2020/07/Towards_a_realtime_twitter_analysis_ISCRAM2012_TT.pdf
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D. and Kappas, A. (2011). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 62(2), p.419.
- Titov, I. and McDonald, R. (2008). Modeling online reviews with multi-grain topic models. In *Proceedings of the 17th international conference on World Wide Web* (pp. 111-120).
- Tong, K. K., Chen, J. H., Yu, E. W.-y. & Wu, A. M. (2020). Adherence to COVID-19 Precautionary Measures: Applying the Health Belief Model and Generalised Social Beliefs to a Probability Community Sample. *NCBI*, 12(4), pp. 1205-1223.
- Tong, R.M. (2001). An operational system for detecting and tracking opinions in online discussions. In *Working notes of the ACM SIGIR 2001 workshop on operational text classification* (Vol. 1, No. 6).
- Tong, R.M. (2001). An operational system for detecting and tracking opinions in online discussions. In *Working notes of the ACM SIGIR 2001 workshop on operational text classification* (Vol. 1, No. 6).

- Torkildson, M.K., Starbird, K. and Aragon, C. (2014). Analysis and visualization of sentiment and emotion on crisis tweets. In Cooperative Design, Visualization, and Engineering: 11th International Conference, CDVE 2014, Seattle, WA, USA, September 14-17, 2014. Proceedings 11 (pp. 64-67). Springer International Publishing. <https://faculty.washington.edu/aragon/pubs/cdve2014-oilspill-torkildson.pdf>
- Tri Sakti, A.M., Mohamad, E. and Azlan, A.A. (2021). Mining of opinions on COVID-19 large-scale social restrictions in Indonesia: Public sentiment and emotion analysis on online media. *Journal of medical Internet research*, 23(8), p.e28249. https://preprints.jmir.org/preprint/28249?__hstc=102212634.372fc3266d2a014934de31dc6f765ea9.1675185950053.1675185950053.1675185950053.1&__hssc=102212634.1.1675185950054&__hsfp=577666595
- Tsao, S.-F. et al. (2022). Public Attitudes During the Second Lockdown: Sentiment and Topic Analyses Using Tweets From Ontario, Canada. *International Journal of Public Health*.
- Turiel, J., Fernandez-Reyes, D. and Aste, T. (2021). Wisdom of crowds detects COVID-19 severity ahead of officially available data. *Scientific Reports*, 11(1), pp.1-9. <https://link.springer.com/content/pdf/10.1038/s41598-021-93042-w.pdf>
- Turney, P.D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. <https://arxiv.org/ftp/cs/papers/0212/0212032.pdf>
- Tyagi, P. and Tripathi, R.C. (2019). A review towards the sentiment analysis techniques for the analysis of Twitter data. In Proceedings of 2nd international conference on advanced computing and software engineering (ICACSE).
- Ullah, M.A., Marium, S.M., Begum, S.A. and Dipa, N.S. (2020). An algorithm and method for sentiment analysis using the text and emoticon. *ICT Express*, 6(4), pp.357-360. <https://doi.org/10.1016/j.icte.2020.07.003>
- Unicode. (n, d). Smileys & Emoticons. <http://unicode.org/emoji/charts/full-emoji-list.html>

- Valdez, D., Ten Thij, M., Bathina, K., Rutter, L.A. and Bollen, J. (2020). Social media insights into US mental health during the COVID-19 pandemic: Longitudinal analysis of Twitter data. *Journal of medical Internet research*, 22(12), p.e21418.
- Van Rooij, B., de Bruijn, A.L., Reinders Folmer, C., Kooistra, E.B., Kuiper, M.E., Brownlee, M., Olthuis, E. and Fine, A. (2020). Compliance with COVID-19 mitigation measures in the United States. Amsterdam law school research paper, (2020-21).
<https://dx.doi.org/10.2139/ssrn.3582626>
- VandenBos, G. R. (Ed.). (2007). APA Dictionary of Psychology. American Psychological Association.
- Varga, I., Sano, M., Torisawa, K., Hashimoto, C., Ohtake, K., Kawai, T., Oh, J.H. and De Saeger, S. (2013). Aid is out there: Looking for help from tweets during a large-scale disaster. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 1619-1629).
<https://aclanthology.org/P13-1159.pdf>
- Vaziripour, E., Giraud-Carrier, C. and Zappala, D. (2016). Analyzing the political sentiment of tweets in Farsi. In Tenth International AAAI Conference on Web and Social Media.
<https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/viewFile/13089/12832>
- Verma, S., Vieweg, S., Corvey, W., Palen, L., Martin, J., Palmer, M., Schram, A. and Anderson, K. (2011). Natural language processing to the rescue? Extracting "situational awareness" tweets during a mass emergency. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 5, No. 1, pp. 385-392).
<https://doi.org/10.1609/icwsm.v5i1.14119>
- Vo, B.K.H. and Collier, N.I.G.E.L. (2013). Twitter emotion analysis in earthquake situations. *Int. J. Comput. Linguistics Appl.*, 4(1), pp.159-173.
<http://www.gelbukh.com/ijcla/2013-1/IJCLA-2013-1-pp-159-173-09-Twitter.pdf>

- Wagh, V., Khandve, S., Joshi, I., Wani, A., Kale, G. and Joshi, R. (2021). Comparative study of long document classification. In *TENCON 2021-2021 IEEE Region 10 Conference (TENCON)* (pp. 732-737). IEEE.
- Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., Ho, C.S. and Ho, R.C. (2020). Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. *International journal of environmental research and public health*, 17(5), p.1729. <https://doi.org/10.3390/ijerph17051729>
- Wang, C., Tee, M., Roy, A.E., Fardin, M.A., Srichokchatchawan, W., Habib, H.A., Tran, B.X., Hussain, S., Hoang, M.T., Le, X.T. and Ma, W. (2021). The impact of the COVID-19 pandemic on the physical and mental health of Asians: A study of seven middle-income countries in Asia. *PloS one*, 16(2), p.e0246824. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0246824>
- Wang, J.H., Liu, T.W., Luo, X. and Wang, L. (2018). An LSTM approach to short text sentiment classification with word embeddings. In *Proceedings of the 30th Conference on Computational Linguistics and Speech Processing (ROCLING 2018)* (pp. 214-223).
- Wang, L. (2016). Discovering phase transitions with unsupervised learning. *Physical Review B*, 94(19), p.195105.
- Wang, X. and Gupta, A. (2015). Unsupervised learning of visual representations using videos. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2794-2802).
- Wankhade, M., Rao, A.C.S. and Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, pp.1-50. <https://link.springer.com/article/10.1007/s10462-022-10144-1>

- Wicke, P. and Bolognesi, M.M. (2021). COVID-19 discourse on Twitter: How the topics, sentiments, subjectivity, and figurative frames changed over time. *Frontiers in Communication*, 6, p.651997. <https://doi.org/10.3389/fcomm.2021.651997>
- Wiebe, J. and Rapaport, W.J. (1988). A computational theory of perspective and reference in narrative. In 26th Annual Meeting of the Association for Computational Linguistics (pp. 131-138). <https://aclanthology.org/P88-1016.pdf>
- Wiebe, J. and Riloff, E. (2005). Creating subjective and objective sentence classifiers from unannotated texts. In *Computational Linguistics and Intelligent Text Processing: 6th International Conference, CICLing 2005, Mexico City, Mexico, February 13-19, 2005. Proceedings 6* (pp. 486-497). Springer Berlin Heidelberg. <https://people.cs.pitt.edu/~wiebe/pubs/papers/Temp/cicling05.pdf>
- Wiebe, J. (2000). Learning subjective adjectives from corpora. *Aaai/iaai*, 20(0), p.0. <https://www.aaai.org/Papers/AAAI/2000/AAAI00-113.pdf>
- Wiedemann, G. and Fedtke, C. (2022). From frequency counts to contextualized word embeddings: The Saussurean turn in automatic content analysis. In *Handbook of Computational Social Science, Volume 2*. Taylor & Francis.
- Wildman, J. (2021). COVID-19 and income inequality in OECD countries. *The European Journal of Health*, 22(3), pp. COVID-19 and income inequality in OECD countries.
- Witmer, D.F. and Katzman, S.L. (1997). On-line smiles: Does gender make a difference in the use of graphic accents? *Journal of Computer-mediated communication*, 2(4). <https://doi.org/10.1111/j.1083-6101.1997.tb00192.x>
- Wood, I. and Ruder, S. (2016). Emojis as emotion tags for tweets. In *Proceedings of the Emotion and Sentiment Analysis Workshop LREC2016, Portorož, Slovenia* (pp. 76-79). sn.

- Wu, Y., Liu, S., Yan, K., Liu, M. and Wu, F. (2014). Opinion flow: Visual analysis of opinion diffusion on social media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), pp.1763-1772.
<http://shixialiu.com/publications/opinionflow/paper.pdf>
- Xiang, B. and Zhou, L. (2014). Improving Twitter sentiment analysis with topic-based mixture modeling and semi-supervised training. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (pp. 434-439).
- Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y. and Zhu, T. (2020). Twitter discussions and emotions about the COVID-19 pandemic: Machine learning approach. *Journal of medical Internet research*, 22(11), p.e20550. <https://www.jmir.org/2020/11/e20550/>
- Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y. and Zhu, T. (2020). Twitter discussions and emotions about the COVID-19 pandemic: Machine learning approach. *Journal of medical Internet research*, 22(11), p.e20550. <https://doi.org/10.2196/20550>
- Yadav, A. and Vishwakarma, D.K. (2020). Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review*, 53(6), pp.4335-4385.
- Ye, Z. (2019). *Attention-Based Sentiment Analysis on Social Media* (Doctoral dissertation, Imperial College London).
https://zheyuye.github.io/files/TwitterSentimentAnalysis_report.pdf
- Yeasmin, N., Mahbub, N.I., Baowaly, M.K., Singh, B.C., Alom, Z., Aung, Z. and Azim, M.A. (2022). Analysis and prediction of user sentiment on the COVID-19 pandemic using tweets. *Big Data and Cognitive Computing*, 6(2), p.65.
- Yoo, B. and Rayz, J.T. (2021). Understanding Emojis for Sentiment Analysis. In *The International FLAIRS Conference Proceedings (Vol. 34)*.
https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Yoo%2C+B.+and+Rayz

- %2C+J.T.%2C+2021%2C+April.+Understanding+Emojis+for+Sentiment+Analysis.+I
n+The+International+FLAIRS+Conference+Proceedings+%28Vol.+34%29.&btnG=
- You, Q., Cao, L., Jin, H. and Luo, J. (2016). Robust visual-textual sentiment analysis: When attention meets tree-structured recursive neural networks. In Proceedings of the 24th ACM International Conference on Multimedia (pp. 1008-1017). <http://lcao.net/paper/t-lstm-acmmm16.pdf>
- Younis, E.M., Zaki, S.M., Kanjo, E. and Houssein, E.H. (2022). Evaluating Ensemble Learning Methods for Multi-Modal Emotion Recognition Using Sensor Data Fusion. *Sensors*, 22(15), p.5611.
- Yu, M. and Dredze, M. (2014). Improving lexical embeddings with semantic knowledge. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 545-550). <https://aclanthology.org/P14-2089.pdf>
- Yu, N. (2014). Exploring Co-training strategies for opinion detection. *Journal of the Association for Information Science and Technology*, 65(10), pp.2098-2110. <https://doi.org/10.1002/asi.23111>
- Yuan, J., McDonough, S., You, Q., and Luo, J. (2013). Sentribute: image sentiment analysis from a mid-level perspective. In Proceedings of the second international workshop on issues of sentiment discovery and opinion mining (pp. 1-8). <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=c5705bf5094ff7f760da74eff167cb695c0422f8>
- Yuan, J., You, Q., and Luo, J. (2015). Sentiment analysis using social multimedia. *Multimedia Data Mining and Analytics: Disruptive Innovation*, pp.31-59. <https://scholar.archive.org/work/lado6fe565d5vanjdbgatxdnzy/access/wayback/https://>

www.springer.com/cda/content/document/cda_downloadaddocument/9783319149974-c2.pdf?SGWID=0-0-45-1500780-p177220328

- Yue, L., Chen, W., Li, X., Zuo, W., and Yin, M. (2019). A survey of sentiment analysis in social media. *Knowledge and Information Systems*, 60, pp.617-663.
<https://cse.iitkgp.ac.in/~mainack/courses/2021-autumn/socialcomp/pdf/sentiment-analysis-survey-yue2019.pdf>
- Yuliyanti, S., Djatna, T. and Sukoco, H. (2017). Sentiment mining of community development program evaluation based on social media. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 15(4), pp.1858-1864.
<http://doi.org/10.12928/telkomnika.v15i4.4633>
- Zhang, J. et al. (2022). Stringency of COVID-19 Containment Response Policies and Air Quality Changes: A Global Analysis across 1851 Cities. *Environmental Science & Technology*, 56(17), p. 12086–12096.
- Zhang, L., Ghosh, R., Dekhil, M., Hsu, M. and Liu, B. (2011). Combining lexicon-based and learning-based methods for Twitter sentiment analysis. HP Laboratories, Technical Report HPL-2011, 89, pp.1-8. <https://www.hpl.hp.com/techreports/2011/HPL-2011-89.pdf>
- Zhang, L., Wang, S., and Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), p.e1253.
- Zhang, S., Zhang, X. and Chan, J. (2017). A word-character convolutional neural network for language-agnostic Twitter sentiment analysis. In *Proceedings of the 22nd Australasian Document Computing Symposium* (pp. 1-4).
- Zhao, J., Lan, M. and Zhu, T. (2014). ECNU: Expression-and Message-level Sentiment Orientation Classification in Twitter Using Multiple Effective Features. In *SemEval@*

COLING (pp. 259-264).

https://web.archive.org/web/20180601210629id_/http://aclweb.org/anthology/S14-2042

Zhao, P., Jia, J., An, Y., Liang, J., Xie, L. and Luo, J. (2018). Analyzing and predicting emoji usages in social media. In Companion Proceedings of the Web Conference 2018 (pp. 327-334).

https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Zhao%2C+P.%2C+Jia%2C+J.%2C+An%2C+Y.%2C+Liang%2C+J.%2C+Xie%2C+L.+and+Luo%2C+J.%2C+2018%2C+April.+Analyzing+and+predicting+emoji+usages+in+social+media.+In+Companion+Proceedings+of+the+The+Web+Conference+2018+%28pp.+327-334%29.&btnG=

Zhou, M., Tao, W., Pengxin, J., Shi, H. and Dongmei, Z. (2020). Table2analysis: Modeling and recommendation of common analysis patterns for multi-dimensional data.

In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 01, pp. 320-328).

Zhou, T., Brown, M., Snavely, N. and Lowe, D.G., (2017). Unsupervised learning of depth and ego-motion from video. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1851-1858).

Zhou, Z.H. (2016). Learnware: on the future of machine learning. *Frontiers Comput. Sci.*, 10(4), pp.589-590.

Zhou, Z.H., (2021). Machine learning. Springer Nature.

Zhu, L. and Wang, B. (2020). Stability analysis of a SAIR rumor spreading model with control strategies in online social networks. *Information Sciences*, 526, pp.1-19.

Zhu, X., Sobhani, P. and Guo, H. (2015). Long short-term memory over tree structures. <https://arxiv.org/pdf/1503.04881.pdf>

Zvolensky, M.J., Garey, L., Rogers, A.H., Schmidt, N.B., Vujanovic, A.A., Storch, E.A.,

Buckner, J.D., Paulus, D.J., Alfano, C., Smits, J.A. and O'Leirigh, C. (2020).

Psychological, addictive, and health behavior implications of the COVID-19 pandemic.

Behavior research and therapy, 134, p.103715. doi: 10.1016/j.brat.2020.103715.

Appendices

Appendix A - Data Splitting

```

train_data, test_data = train_test_split(
    clean_data, test_size=1-0.95, random_state=7
)

lb= LabelEncoder()
lb.fit(clean_data['Sentiment'])

y_train = lb.transform(train_data['Sentiment'].to_list())
y_test = lb.transform(test_data['Sentiment'].to_list())

y_train = y_train.reshape(-1,1)
y_test = y_test.reshape(-1,1)

y_train.shape
tokenizer = Tokenizer()
tokenizer.fit_on_texts(clean_data.text)

word_index = tokenizer.word_index
vocab_size = len(tokenizer.word_index) + 1
print("Vocabulary Size :", vocab_size)

x_train = pad_sequences(tokenizer.texts_to_sequences(train_data.text
), maxlen = MAX_SEQUENCE_LENGTH)
x_test = pad_sequences(tokenizer.texts_to_sequences(test_data.text),
maxlen = MAX_SEQUENCE_LENGTH)

```

Appendix B - Data Labelling

```

score = clean_data["Compound"].values
sentiment = []
for i in score:
    if i >= 0.05 :
        sentiment.append('Positive')
    elif i <= -0.05 :
        sentiment.append('Negative')
    else:
        sentiment.append('Neutral')
clean_data["Sentiment"] = sentiment
clean_data.head()

```

Appendix C - Emoji Translation

```
# Replace emoji with its description function
def get_emoji_regexp():
    emojis = sorted(emoji.EMOJI_DATA, key=len, reverse=True)
    pattern = u'(' + u'|'.join(re.escape(u) for u in emojis) + u')'
    return re.compile(pattern)

emoji_pattern = get_emoji_regexp()

def emoji_description(match):
    return emoji.demojize(match.group(), delimiters=(" ", " "))
```

Appendix D – KNN

```
knn = KNeighborsClassifier(n_neighbors=5)
knn.fit(x_train, y_train_1d)
y_pred_knn = knn.predict(x_test)
knn_report = classification_report(y_test_1d, y_pred_knn, output_dict=True)

prediction_reports.append({
    "Model": "KNN", "precision": knn_report['weighted avg']['precision'],
    "Recall": knn_report['weighted avg']['recall'], "F-score": knn_report['weighted avg']['f1-score'],
    "accuracy": knn_report['accuracy']
})
accuracyPlot(y_pred_knn, y_test, 50)
```


Appendix E - Logistic Regression

```
# Training the model
lr = LogisticRegression(solver='lbfgs', max_iter=1000)
lr.fit(x_train, y_train_1d)
#Predicting the labels for test data
y_pred_lr = lr.predict(x_test)

lr_report = classification_report(y_test_1d, y_pred_lr, output_dict=True)
prediction_reports.append({
    "Model": "Logistic regression", "precision": lr_report['weighted
avg']['precision'],
    "Recall": lr_report['weighted avg']['recall'], "F-score": lr_report['weighted avg']['f1-score'],
    "accuracy": lr_report['accuracy']
})
accuracyPlot(y_pred_lr, y_test, 50)
```

Appendix F - Multinomial Naïve Bayes

```
mnb = MultinomialNB()
mnb.fit(x_train, y_train_1d)
MultinomialNB()
#Predicting the labels for test data
y_pred_mnb = lr.predict(x_test)

mnb_report = classification_report(y_test_1d, y_pred_mnb, output_dict=True)
prediction_reports.append({
    "Model": "Multinomial Naives Bayes", "precision": mnb_report['weighted avg']['precision'],
    "Recall": mnb_report['weighted avg']['recall'], "F-score": mnb_report['weighted avg']['f1-score'],
    "accuracy": mnb_report['accuracy']
})
accuracyPlot(y_pred_mnb, y_test, 50)
```

Appendix G - Random Forest Classifier

```
# Create a Gaussian Classifier
rf = RandomForestClassifier()

rf.fit(x_train, y_train_1d)
y_pred_rf = rf.predict(x_test)
rf_report = classification_report(y_test_1d, y_pred_rf, output_dict
=True)

prediction_reports.append({
    "Model": "Random Forest", "precision": rf_report['weighted avg'
]['precision'],
    "Recall": rf_report['weighted avg']['recall'], "F-score": rf_re
port['weighted avg']['f1-score'],
    "accuracy": rf_report['accuracy']
})
accuracyPlot(y_pred_rf, y_test, 50)
```

Appendix H - LSTM

```
from tensorflow.keras.layers import Conv1D, Bidirectional, LSTM, Dense, Input,
Dropout, SpatialDropout1D

from tensorflow.keras.callbacks import ModelCheckpoint

from tensorflow.keras.optimizers import Adam

from tensorflow.keras.callbacks import ReduceLROnPlateau

!wget http://nlp.stanford.edu/data/GloVe.6B.zip

!unzip GloVe.6B.zip

GLOVE_EMB = '/kaggle/working/GloVe.6B.300d.txt'

EMBEDDING_DIM = 300

BATCH_SIZE = 1024

LR = 1e-3

EPOCHS = 10

MODEL_PATH = '../output/kaggle/working/best_model.hdf5'

embeddings_index = {}
```

```

f = open(GLOVE_EMB)
for line in f:
    values = line.split()
    word = value = values[0]
    coefs = np.asarray(values[1:], dtype='float32')
    embeddings_index[word] = coefs
f.close()

print('Found %s word vectors.' %len(embeddings_index))
embedding_matrix = np.zeros((vocab_size, EMBEDDING_DIM))
for word, i in word_index.items():
    embedding_vector = embeddings_index.get(word)
    if embedding_vector is not None:
        embedding_matrix[i] = embedding_vector

embedding_layer = tf.keras.layers.Embedding(
    vocab_size, EMBEDDING_DIM, weights=[embedding_matrix], input_length=MAX_SEQUENCE_LENGTH, trainable=False
)

#Training model
sequence_input = Input(shape=(MAX_SEQUENCE_LENGTH,), dtype='int32')
embedding_sequences = embedding_layer(sequence_input)
x = SpatialDropout1D(0.2)(embedding_sequences)
x = Conv1D(64, 5, activation='relu')(x)
x = Bidirectional(LSTM(64, dropout=0.2, recurrent_dropout=0.2))(x)
x = Dense(64, activation='relu')(x)
x = Dropout(0.5)(x)

```

```

x = Dense(8, activation='relu')(x)
outputs = Dense(3, activation='softmax')(x)
model = tf.keras.Model(sequence_input, outputs)
tf.keras.utils.plot_model(model)
# This LSTM uses Adam, optimization algorithm for Gradient Descent.
model.compile(
    optimizer=Adam(learning_rate=LR),
    loss='sparse_categorical_crossentropy',
    metrics=['accuracy']
)

ReduceLROnPlateau = ReduceLROnPlateau(factor=0.1, min_lr = 0.01, monitor = 'val_loss', verbose = 1)

history = model.fit(
    x_train, y_train, batch_size=BATCH_SIZE, epochs=EPOCHS,
    validation_data=(x_test, y_test), callbacks=[ReduceLROnPlateau]
)

scores = model.predict(x_test, verbose=1, batch_size=10000)
lstm_pred = np.argmax(scores, axis=1)
lstm_report = classification_report(y_test_1d, lstm_pred, output_dict=True)

prediction_reports.append({
    "Model": "LSTM", "precision": lstm_report['weighted avg']['precision'],
    "Recall": lstm_report['weighted avg']['recall'], "F-score": lstm_report['weighted avg']['f1-score'],
    "accuracy": lstm_report['accuracy']
})

```

```

}))

# Learning Curve of Loss and accuracy of the model on each epoch.
s, (at, al) = plt.subplots(2,1, figsize=[16, 8])
at.plot(history.history['accuracy'], c= 'b')
at.plot(history.history['val_accuracy'], c='r')
at.set_title('model accuracy')
at.set_ylabel('accuracy')
at.set_xlabel('epoch')
at.legend(['LSTM_train', 'LSTM_val'], loc='upper left')

```

Appendix I - Transformers

```

from tensorflow.keras.layers import MultiHeadAttention, LayerNormalization, Dropout, Layer

from tensorflow.keras.layers import Embedding, Input, GlobalAveragePooling1D, Dense

from tensorflow.keras.models import Sequential, Model

import warnings
warnings.filterwarnings("ignore", category=np.VisibleDeprecationWarning)

maxlen = 200 # Only consider the first 200 words of each movie review

X_train = pad_sequences(tokenizer.texts_to_sequences(train_data.text), maxlen = maxlen)

X_test = pad_sequences(tokenizer.texts_to_sequences(test_data.text), maxlen = maxlen)

X_train, x_val, Y_train, y_val = train_test_split(X_train, y_train, test_size=0.5, random_state=42)

class TransformerBlock(Layer):
    def __init__(self, embed_dim, num_heads, ff_dim, rate=0.1):
        super(TransformerBlock, self).__init__()
        self.att = MultiHeadAttention(num_heads=num_heads, key_dim=embed_dim)

```

```

        self.ffn = Sequential(
            [Dense(ff_dim, activation="relu"),
             Dense(embed_dim),]
        )
        self.layer_norm1 = LayerNormalization(epsilon=1e-6)
        self.layer_norm2 = LayerNormalization(epsilon=1e-6)
        self.dropout1 = Dropout(rate)
        self.dropout2 = Dropout(rate)

    def call(self, inputs, training):
        attn_output = self.att(inputs, inputs)
        attn_output = self.dropout1(attn_output, training=training)
        out1 = self.layer_norm1(inputs + attn_output)
        ffn_output = self.ffn(out1)
        ffn_output = self.dropout2(ffn_output, training=training)
        return self.layer_norm2(out1 + ffn_output)

class TokenAndPositionEmbedding(Layer):
    def __init__(self, maxlen, vocab_size, embed_dim):
        super(TokenAndPositionEmbedding, self).__init__()
        self.token_emb = Embedding(input_dim=vocab_size, output_dim=embed_dim)
        self.pos_emb = Embedding(input_dim=maxlen, output_dim=embed_dim)

    def call(self, x):
        maxlen = tf.shape(x)[-1]
        positions = tf.range(start=0, limit=maxlen, delta=1)
        positions = self.pos_emb(positions)
        x = self.token_emb(x)
        return x + positions

embed_dim = 32  # Embedding size for each token
num_heads = 2  # Number of attention heads
ff_dim = 32  # Hidden layer size in feed forward network inside transformer

inputs = Input(shape=(maxlen,))
embedding_layer = TokenAndPositionEmbedding(maxlen, vocab_size, embed_dim)
x = embedding_layer(inputs)
transformer_block = TransformerBlock(embed_dim, num_heads, ff_dim)

```

```

x = transformer_block(x)
x = GlobalAveragePooling1D()(x)
x = Dropout(0.1)(x)
x = Dense(20, activation="relu")(x)
x = Dropout(0.1)(x)
outputs = Dense(3, activation="softmax")(x)

model = Model(inputs=inputs, outputs=outputs)
model.compile(optimizer="adam", loss="sparse_categorical_crossentropy", metrics=["accuracy"])

ReduceLROnPlateau_ = ReduceLROnPlateau(factor=0.1, min_lr = 0.01, monitor = 'val_loss', verbose = 1)
history_transformer = model.fit(
    X_train, Y_train.ravel(), batch_size=64, epochs=10,
    validation_data=(x_val, y_val.ravel()), callbacks=[ReduceLROnPlateau_]
)

# Learning Curve of Loss and accuracy of the Transformer model on each epoch.
plt.figure(figsize=[14, 6])
plt.plot(history_transformer.history['accuracy'], c= 'b')
plt.plot(history_transformer.history['val_accuracy'], c='r')
plt.title('Transformer Model Accuracy')
plt.ylabel('Accuracy')
plt.xlabel('Epoch')
plt.legend(['Training', 'Validation'], loc='upper left')
plt.show()

# Plot the training and validation Loss
plt.figure(figsize=[14, 6])
plt.plot(history_transformer.history['loss'], c='m')
plt.plot(history_transformer.history['val_loss'], c='c')
plt.title('Transformer Model Loss')
plt.ylabel('Loss')
plt.xlabel('Epoch')
plt.legend(['Training', 'Validation'], loc = 'upper left')
plt.show()

```

Appendix J - Sentiment Distribution

```
#Plotting sentiments
sns.countplot(data=clean_data, x='Sentiment')

sentiment_data = clean_data['Sentiment']

value_counts = sentiment_data.value_counts()
percentages = value_counts / len(sentiment_data) * 100

plt.figure(figsize=(8, 8))
plt.pie(percentages, labels=value_counts.index, autopct='%1.1f%%')
plt.title("Distribution of sentiment")
plt.show()
```

Appendix K - Model Results/Performance

```
df_reports = pd.DataFrame(prediction_reports)

df_reports
```

Appendix L - Graphical Model Evaluation

```
# Plotting data
data_accuracy = df_reports['accuracy']

# Create the bar chart
fig, ax = plt.subplots(figsize=(12, 6))
ax.bar(range(len(data_accuracy)), data_accuracy)

ax.set_xticks(range(len(data_accuracy)))
ax.set_xticklabels(df_reports['Model'])
plt.xticks(rotation=45)

ax.set_xlabel('Models')
ax.set_ylabel('Accuracy')
ax.set_title('Models prediction accuracy')

# Display the chart
plt.show()
```