Sentiment Analysis for Arabic

تحليل المشاعر للعربية

By

Sanjeera Siddiqui

2013228044

Dissertation submitted in partial fulfillment of

MSc Informatics (Knowledge & Data Management)

Faculty of Engineering & Information Technology

Dissertation Supervisor

Professor Dr Khaled Shaalan

December-2015

DISSERTATION RELEASE FORM

Student Name	Student ID	Programme	Date
Sanjeera Siddiqui	**2013228044**	**MSc Informatics**	**15th December 2015**

Title
Sentiment Analysis for Arabic

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that the content of this dissertation is the direct result of my own work and that any use made in it of published or unpublished copyright material falls within the limits permitted by international copyright conventions.

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

I hereby agree that the material mentioned above for which I am 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 that copy available in digital format if appropriate.

I understand that I may apply to the University to retain the right to withhold or to restrict access to my dissertation 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

Abstract

Named Entity Recognition, Question Answering, Information Retrieval, Machine Translation, etc. fall under the tasks that follow Natural Language Processing approaches wherein Sentiment Analysis uses Natural Language Processing as one of the means to find the subjective text indicating negative, positive or neutral polarity. The united approach of text mining and natural language processing, termed to be as Sentiment Analysis has gained huge heights due to the increased use of social media websites like Facebook, Instagram, Twitter to name a few. Sentiment Analysis is a growing field and nevertheless a lot of research is done in English when compared to Arabic language. Analysis of Sentiments helps companies, government and other organization to improvise their products and service based on the reviews or comments. This dissertation not only depicts the various challenges faced by Arabic Natural Language processing in the Sentiment Analysis task, but this dissertation presents an Innovative approach that explores the role of lexicalization for Arabic sentiment analysis. Sentiment Analysis in Arabic is hindered due to lack of resources, language in use with sentiment lexicons, pre-processing of dataset as a must and major concern is repeatedly following same approaches. One of the key solution found to resolve these problems include applying the extension of lexicon to include more words not restricted to Modern Standard Arabic. Secondly, avoiding pre-processing of dataset. Third, and the most important one, is investigating the development of an Arabic Sentiment Analysis system using a novel rule-based approach. This approach uses heuristics rules that is triggered based on end-to-end mechanism of a particular word in a manner that accurately classifies the tweets as positive or negative. The manner in which a series of abstraction occurs resulting in an end to end rule-based chaining approach. For each lexicon this chain specifically follows a chaining of rules (i.e. rule A chains with rule B and if required rule C and so on), with appropriate positioning and prioritization of rules. Expensive rules in terms of time and effort thus resulted in outstanding results. Experiments were conducted on two dataset. They are chosen for a number of good reasons, including their availability and successfully used by other researches, richness and sufficient to come to a conclusion, and provision with electronic resources such as lexicon. Two set of

experiments were done. The first set of experiment was done only with two rules – “equal to” and “within the text”. The second set of experiment was done with rule chaining mechanism. The results thus achieved with end to end rule chaining approach achieved 93.9% accuracy when tested on one dataset, which is considered the baseline, and 85.6% accuracy on OCA, the second dataset. A further comparison with the baseline showed huge increase in accuracy by 23.85%.

Arabic Translation

التعرف علي الأسماء والإجابة علي الأسئلة واسترجاع المعلومات والترجمة الآلية الخ تندرج في إطار المهام التي تتبع نهج معالجة اللغة الطبيعية حيث يستخدم تحليل المشاعر معالجة اللغة الطبيعية باعتبارها واحدة من الوسائل للعثور على نص شخصي مشيرا إلى قطبية سلبية أوايجابية أو محايدة. النهج الناتج عن دمج التنقيب في النص ومعالجة اللغات الطبيعية تم وصفه علي أنه تحليل للمشاعر والذي اكتسب اهتمامات كبيرة بسبب الزيادة المطردة في استخدام مواقع التواصل الاجتماعي مثل فيسبوك وانستجرام وتويتر على سبيل المثال لا الحصر. تحليل المشاعر هو مجال متطور وعلى الرغم من ذلك فأن معظم البحوث تمت علي اللغة الإنجليزية بالمقارنة مع اللغة العربية. تحليل المشاعر يساعد الشركات والحكومات والمنظمات الأخرى إلى تحسين منتجاتها وخدماتها استنادا إلى الملاحظات أو التعليقات. هذه الأطروحة توضح ليس فقط مختلف التحديات التي واجهت معالجة اللغة العربية الطبيعية في مهمة تحليل المشاعر، ولكن تعرض نهج مبتكر يستكشف دور الكلمات في تحليل المشاعر العربية .تواجه نظم تحليل المشاعر في العربية معوقات نظرا لقلة الموارد، واللغة المستخدمة مع معجم المشاعر وضرورة المعالجة الأولية للبيانات وهناك تحفظ كبير من تكرار اتباع نفس النهج. واحد من الحلول الرئيسية المطروحة لحل هذه المشاكل يتضمن تطبيق تمديد المعجم لتشمل المزيد من الكلمات التي لا تقتصر فقط على اللغة العربية الفصحى الحديثة. ثانيا: تجنب المعالجة الأولية للبيانات. ثالثا: والأهم، التحقق من تطوير نظام تحليل المشاعر العربية باستخدام نهج قائم على القواعد. يستخدم هذا النهج قواعد الاستدلال التي يتم تشغيلها على أساس أسلوب النهاية-إلى-النهاية لكلمة معينة بطريقة تصنف بدقة التغريدات علي أنها إيجابية أو سلبية. الطريقة التي تعتمد علي اجراء سلسلة من التبسيطات والتي تؤدي إلى نهج تسلسل النهاية-إلى-النهاية القائمة على القاعدة. لكل معجم هذه التسلسل يتبع تحديدا تسلسل القواعد (أي القاعدة "ا" ترتبط بالقاعدة "ب" وإذا لزم الأمر القاعدة "ج" وهلم جرا)، مع تحديد الموقع وأولويات القواعد المناسبة. القواعد مكلفة من حيث الوقت والجهد أدت إلي نتائج باهرة.

وأجريت التجارب على مجموعتين من البيانات. تم اختيارهم حسب عدد من الأسباب الوجيهة، بما في ذلك توافرها واستخدامها بنجاح من قبل غيرها من الأبحاث والثراء وأنها كافية للتوصل إلى الخلاصة، وتوفير المصادر الإلكترونية مثل المعجم. تم القيام بتجربتين. وقد أجري الاختبار الأول علي المجموعة الأولي باستخدام قاعدتين فقط - "يساوي" و "داخل النص". وقد تم اجراء اختبار للمجموعة الثانية من التجربة باستخدام آلية النهج القاعدي المتسلسل. وقد حققت نتائج إلى نهج تسلسل النهاية-إلى-النهاية القائمة على القاعدة 93.9٪ من الدقة عند اختباره على المجموعة الأولي من البيانات ، والتي تعتبرالحد الأدني من الأداء، وبلغت الدقة 85.6٪ عند اختباره على المجموعة الثانية من البيانات. وأظهرت مقارنة أخرى مع الحد الأدني من الأداء زيادة كبيرة في الدقة بنسبة 23.85٪.

Dedication

This dissertation is dedicated to my family for their prayer, love and unlimited support during my lifetime.

Acknowledgement

Firstly and most importantly I would like to thank The Almighty - Allah for giving me strength to face new challenges and triumph over the current study.

I would like to express my sincere thanks to Prof. Khaled Shaalan, for supervising this thesis and offering his immense support throughout, without which I wouldn't have come up with a novel approach to sentiment analysis. His profitable comments and sagacious views have made it easier for me to exceed excellence in this research. Not to forget his patience in explaining and guiding me throughout this journey. He is a very good motivator as well as critical reflector, he motivated me when I did something right and at the same time provided me with his critical reviews on my work so as to enhance my research. He is the best advisor I could have ever dreamt off.

I would like to take this opportunity to thank Mrs. Ranju Anand, the principal of the Winchester School. She has been a role model for me and have learnt a lot from her in my tenure at the Winchester School for past 7 years which did helped me excel in the current study.

My husband, Syed Faizan supported me throughout, his constant motivation have really helped me achieve what I have achieved today, a big thank you to him for supporting me in my endeavor to achieve success in the study I conducted in this dissertation. Furthermore I would like to stretch out my thanks to few key people in my life including my parents, my in laws and my children in supporting me emotionally. Without their love, affection and motivation I wouldn't have been able to dedicate my time to this study.

Keywords

Sentiment analysis, Opinion mining, Rule-based approach, Arabic natural language processing

Contents

List of Tables

List of Figures

Chapter One

Introduction

Internet, also termed as World Wide Web, contains lots of information. Internet provides people with an open space to share their opinions or sentiments, their experiences, and their preferences on an entity. These opinion or sentiment shared online, forms a base for many organizations, companies and governments to take productive decisions to satisfy the needs of customers and to improve their businesses. Users do read these reviews and take further decisions to either opt for or reject a product or service. People in today's world are driven by online reviews and comments before they actually proceed with buying and selling anything. Hence there comes a need for building a system which can automatically perform the task of identifying negative and positive reviews.

Nevertheless this task of Sentiment Analysis has gained huge heights in English when compared to Arabic. Arabic is widely used language on Internet with more than 155 million users opting for it. Arabic Sentiment Analysis is hindered due to various constraints, some of the key major constraints include the lack of resources in terms of lexicons and standard dataset, pre-processing of dataset is a must step followed by most of the researchers, due to Arabic orthographic and morphologic challenges and most importantly absence of a novel approach to perform sentiment analysis which gives splendid results.

1.1 Overview of Sentiment Analysis

The mission of Sentiment Analysis is to recognize the text with opinions and arrange them in a manner conforming to the polarity, which includes: negative, positive or neutral. Opinions does take the businesses to huge heights [Taboada et al. 2013; Feldman 2013]. [Korayem 2012] stated that Subjectivity and Sentiment Analysis classification are processed in four dimensions: 1) subjectivity classification, to forecast on Subjective or Objective, 2) Sentiment Analysis, to predict on the polarity that could be negative, positive, or neutral, 3) the level based on document, sentence, word or phrase classification, and 4) the approach that is followed; it could be rule-based, machine learning, or hybrid. Sentiment Analysis could be either direct that is explicit or a kind of sentiment that is implicit. Direct sentiment illustrates a sentiment which is directly conveying feeling wherein a kind of sentiment illustrates something which looks like a sentiment. [Quirk et al. 1985] coined the term private state which illustrate an opinion or a state of conveying an opinion. [Pang and Lee 2008] coined various terms for sentiment analysis like sentiment polarity analysis or subjectivity analysis or opinion analysis, leaving the new bees in the field of sentiment analysis with an ambivalence.

Arabic Sentiment Analysis has been a huge focus for researchers [Abdul-Mageed et al. 2012]. The last decade has shown a growing interest in addressing challenges that underlie the development of a productive and robust Arabic Sentiment Analysis system. Sentiment Analysis is mostly classified as a domain specific problem [Aue and Gamon 2005].

1.2 Inspiration

1.2.1 Importance, Aims and Outcomes

Through Sentiment Analysis companies, organizations, businesses and so on collect reviews to know where they stand, what could be done to make things better? The key part of Sentiment Analysis distinguishes good and bad reviews, henceforth, giving a clear perspective to what extent a particular product or tool is liked or not liked by people. Sentiment Analysis and opinion mining are used reciprocally[Taboada et al. 2011], like [Nasukawa et al. 2003] used the term Sentiment Analysis and [Dave 2003] used the term Opinion Mining in their research conducted on identifying polarities.

Some other terms used for Sentiment Analysis includes review mining and appraisal extraction [Piao 2007]. [Wright 2009] compares currency with online reviews wherein the reviews can shape or smash a product in a market place. Other than this, Sentiment Analysis is a key area of interest for psychologist to study the mental make-up on a particular issue of interest [Pang and Lee 2008]. Business intelligence applications are also benefited by Sentiment Analysis [Glance et al. 2005].

Sentiment Analysis other application in politics includes finding out what voters think [Efron 2004; Goldberg et al. 2007; Hopkins and King 2007; Laver et al. 2003; Mullen and Malouf 2006]. Sentiment Analysis behaves like recommendation system by not recommending a product or item which have received negative reviews [Terveen 1997; Tatemura 2000; Glance et al. 2005]. Message filtering could be even handled using Sentiment Analysis [Mallouf and Mullen 2008].

Not to forget, Sentiment Analysis plays a vital role in measuring the impact on the news reviews and news channel reviews [Balahur et al. 2007]. Question Answering is another field where Sentiment Analysis is very useful [Somasundaran 2007; Stoyanov 2005; Lita et al. 2005]. Sentiment Analysis of Political reviews helps to some extent in foretelling the results. Hence, Sentiment Analysis has gained huge importance in every stream right from analysis of reviews of a branded tooth brush to analysis of an alien visit to earth.

The objective of this study is to investigate techniques that determine negative and positive polarity of the input text. The study aims particularly to derive rules that can axiomatically classify Tweets, or Movie or news reviews as positive or negative. One of the significant outcome would be to distinguish the proposed end-to-end rule chaining approach to lexicon-based approach on the selected dataset. Other than this it would be interesting to eliminate the pre-processing step and see if the built system respond to the entered tweets and gives an output polarity.

1.2.2 Sentiment Analysis in Natural Language Processing

Sentiment Analysis is one which involves fields like NLP, computational linguistics and machine learning. Languages spoken by people relates to their culture and what they speak, hence different languages are spoken or learnt in different places, which differ in features as well in characteristics. Hence, the approach or technique that was useful for one language may not be useful for another language. Natural language processing is violent as well as simple. Arabic Natural language processing is shifting most of the researcher’s mindset to Arabic, due to the increase use of Arabic language by individuals and the increased internet Arabic users. Arabic language holds one of the top ten position in the world wide used languages[1].

Arabic Natural language processing in sentiment analysis is taking huge attention due to the unavailability of resources, Arabic language complexity, and non-existence of appropriate tools with full fledge operators to perform various pre-processing tasks of Arabic datasets. This calls for a need to extend the work in Arabic Sentiment Analysis.

1.3 Knowledge Gap in Sentiment Analysis

In the literature, many key issues have been evident in the field of Arabic Sentiment Analysis. This chapter sheds the light on the gaps and issues found in the studies done so far. Handful of work is done so far in Arabic Sentiment Analysis with very little resources being built or shared openly.

Most of the research focussed on Modern Standard Arabic except for SAMAR which took an initiative to address dialect Arabic. SAMAR project, proposed by [Abdul-Mageed et al. 2012], is the one which is applied on four different genres; others hardly focussed only on newswires or user reviews. Most of the sentiment Analysis restricted to positive and negative polarity. Problems which needs a concrete addressing is to deal with different dialects. This problem still remains challenging due to the absence of resources and systems (tools).

Moreover, most of the work done so far involved Machine Learning approaches with very less work done on rule-based approach. The key underlying factor to be noticed is that none of the results exceeded 90% accuracy. This indeed calls for a novel approach for Sentiment Analysis. The key to be noticed is the very illustrative terminology behind Sentiment Analysis – it is must to do pre-processing add to the knowledge gap, where researchers missed to eliminate this step completely and measure the results without this step.

[1] http://www.internetworldstats.com/stats7.htm

Lexicon was only seen to find its significance in measuring the results [Nawaf et al. 2014] with repeated addition and deletion of words into the lexicons; no attempts to change the techniques or application of rules were seen. Another key area missed was lack of in-depth evaluation of the hick-ups with common words in both positive and negative lexicons. Whether these common words played a key role in the whole process. No comments on the exclusions or inclusions of these common words as part of rules creation were not evident to the literature to the best of our knowledge, which explores the role of lexicalized rules for Arabic sentiment analysis.

1.4 Problem statement and Research Questions

1.4.1 Problem Statement

Sentiment Analysis is one of the key area gaining huge heights due to the emerging use of online industry. Sentiment Analysis or Opinion Analysis or more precisely Sentiment Polarity Analysis focusses on revealing the sentiments of reviewers. The problem statement can be coined as problem of text classification into negative or positive, very well supported by [Pang and lee 2008]. English being highly investigated language with very little work done so far with Arabic gives a wakeup call to all the Arabic researches. One of the key reason for this deviation in Arabic and English research so far is the absence of resources for Arabic Sentiment Analysis. This lack of resources is due to time consumption and hard labor.

Reflecting on the availability and hinder in success in Arabic sentiment Analysis following are found to be the major problems.

- Pre-processing: Stemmers or stop word lists more than anything else is based on Modern Standard Arabic. Pre-processing techniques are used to clean the data [Aldayel and Azmi 2015].

- Machine learning approaches or Lexicon-based approaches are found to be biased with a clear miss on good feature set and classification algorithm on Arabic dialect, for example Jordanian dialect [Nawaf et al. 2013].

- Lexicons are mainly based on Modern Standard Arabic with a usual exclusion on highly weighted lexicons of Dialect Arabic [El-Baltagy and Ali 2013].

1.4.2 Research Questions

The research questions addressed in this study are as follows:

- RQ1: Is it possible to create a system for Sentiment Analysis without pre-processing the dataset?
- RQ2: Is it possible to improvise the performance of sentiment analysis system through an innovative rule chaining approach?
- RQ3: Does the use of end-to-end mechanism with rule chaining approach gives better results than the Lexicon-based approach?
- RQ4: Does the inclusion of common words based on the analysis of tweets in the negative and positive lexicon list enhance the results?
- RQ5: Does the positioning of rules make a huge difference to the rule-based approach?

The research hypothesis of this study is that the use of end-to-end mechanism in lexicalized rules to demonstrate the capability of the novel rule chaining approach in improving the performance of a Sentiment Analysis of positive and negative tweets.

1.5 Contribution

The end-to-end mechanism with rule chaining approach, accounts to address the above research questions and we carried out systematic enquiry against the hypothesis which results in the following contributions.

In order to develop an end-to-end rule chaining approach, first step taken was to improvise the [Nawaf et al. 2014]'s lexicon with the very addition of words falling at either ends of the review as well as one which were found to be mostly repeated.

The key weakness found in many research is to deal with pre-processing of dataset in Arabic. In this research, the pre-processing of dataset is completely eliminated which have not been done before. A novel approach targeted to rules chained in an orderly fashion is another key addition to research in Sentiment Analysis.

The inclusion of end-to-end rule chaining approach specific to a lexicon addresses the underlying need to cover the common sentiments found in both positive and negative lexicon. This rule based end-to-end chained approach is designed to not only overcome the hurdle of excluding common words found in both the list, but also to use these common words to enhance the overall results with appropriate positioning (order) of rules.

Two set of experiments were conducted. First set of experiment were done on system 1 which comprises of “equal to” and “search within the text” rule. Second set of experiment included a 360 degree rule coverage with an addition of “beginning with” and “ending with” rules to system 1 as well the inclusion of positioning and chaining of rules. Overall results obtained shows outperformance gained in the second sets of experiment performed for [Nawaf et al. 2014] and [Rushdi et al. 2011] dataset. 39.8% increase in accuracy in System 2 when compared to System 1 for OCA corpus [Rushdi et al. 2011] shows a high demand for rules covering all dimensions. System 2 experiment for [Nawaf et al. 2014] showed a 4.3% increase in accuracy when compared to the first experiment. Both the set of experiments further compared to [Nawaf et al. 2014] added to the glow of the rules proposed in this dissertation with a huge outperformance shown by both the experiments. This dissertation advances the state of the art of Arabic Sentiment Analysis by not just giving 93.9% accuracy for the [Nawaf et al. 2013] dataset used for building the system but also giving 86% accuracy for OCA dataset. Hence giving an open window to check this system on different datasets for depicting the polarity.

1.6 Chapter Summary

This chapter explicitly covers the crucial information of the topic of this dissertation. Online information in millions and billions resulted in the evolvement of sentiment analysis field, which is grabbing attention from researchers and had impacted the whole digital era. Review sharing is the most common thing found in the Internet Social Media World whether it is a brush or it is toothpaste or it is a service or anything. The key underlying parameter seen is the sharing of reviews on any context and how this impact the users to take decisions on many things right from buying a movie ticket to buying a property to many advance operations. It calls for a need for an appropriate tool which can do justice to identify sentiments polarity and help companies, organizations and not to forget the citizens of this planet earth to benefit from this classification. Arabic is the fourth used language on Internet according to the Internet World Stat (2015) highlighting the need for extensive work in Arabic especially in the field of Sentiment Analysis. This chapter comes to an end with inspiration, problem statement, research questions and hypothesis.

This dissertation is organized as follows. Chapter 2 covers a comprehensive depiction on the Arabic NLP. Chapter 3 covers the foundations for sentiment analysis, a close look at subjectivity analysis and sentiment analysis. Chapter 4 covers the approaches of Sentiment Analysis. Chapter 5 includes the data collection and the methodology followed. Chapter 6 covers a structured depiction of the system implemented. Chapter 7 covers the results, Chapter 8 presents the error analysis and lastly Chapter 9 presents our conclusion.

Chapter Two

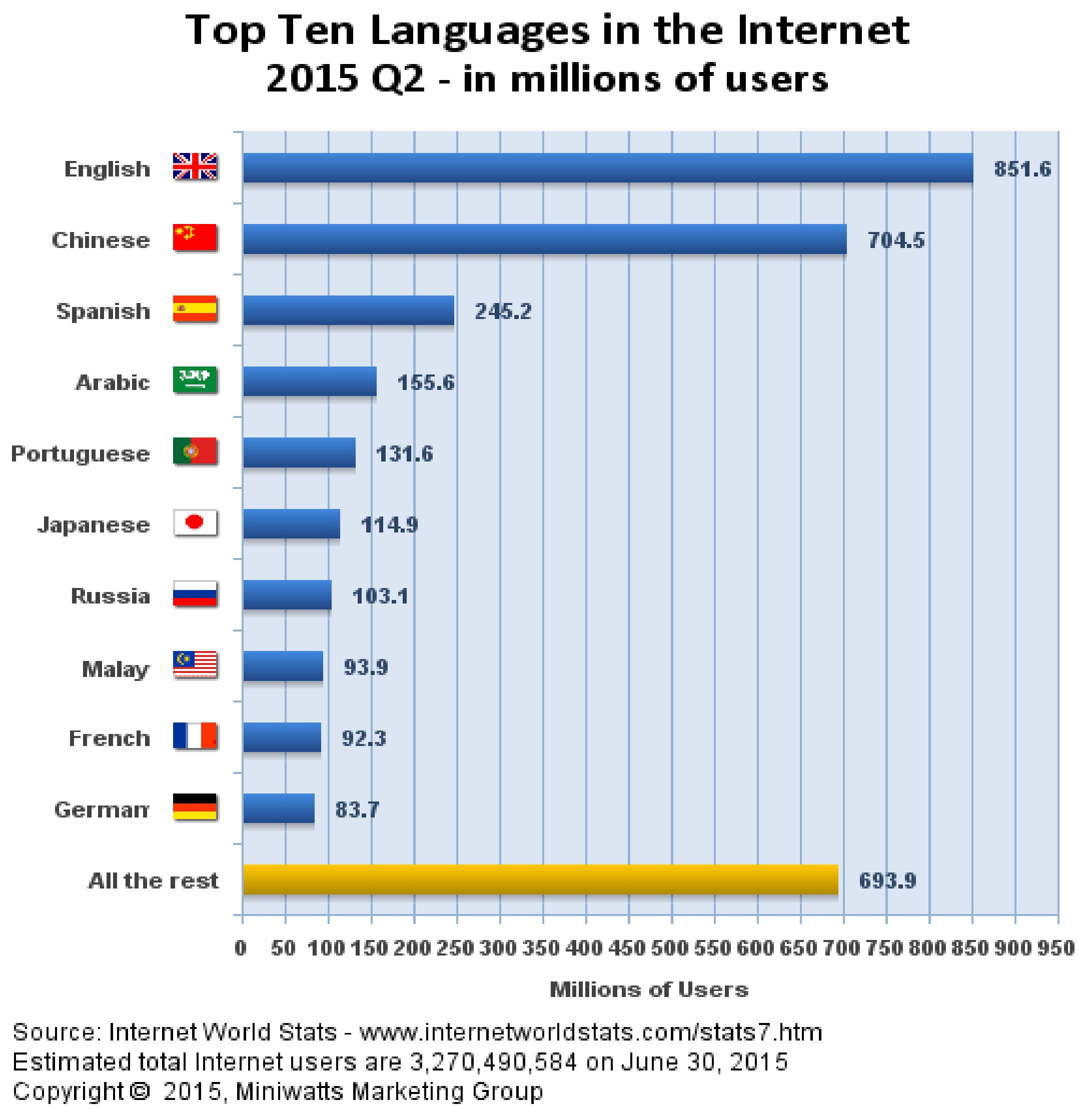
Arabic Natural Language Processing and its challenges

NLP has gained huge importance in machine translation and other applications like speech synthesis and recognition, localization multilingual information systems ... etc. Machine translation, Arabic named entity recognition and sentiment analysis are some of the Arabic tools which have shown considerable input in intelligence and security agencies. Arabic Sentiment Analysis has found to closely attract the researchers due to the growing Arabic Internet Users, the need to address to their online comments and reviews, the need to address various challenges that Arabic as a language brings along with it, and last but not the least the outrageous need to develop tools and resources. Sentiment Analysis field in Arabic Natural Language Processing is still sterile and requires fruit bearing efforts in order to achieve success and tremendous results. This chapter is an eye opener in the research field highlighting the key features, forms and of Arabic language, with an inclusive depiction on challenges encountered in Arabic Natural Language Processing with respect to Sentiment Analysis.

2.1 Arabic language

A language which is treated as challenging due to its complex linguistic structure is Arabic [Attia 2008]. [Farghaly and Shaalan 2009] stated the importance of Arabic language wherein Arabic is linked with Islam and more than 1.4 million Muslims performs their prayers five times a day using this language. Arabic is a Semitic language, which is different in terms of its history, structure, diglossic nature and complexity [Farghaly and Shaalan 2009]. Arabic is widely spoken by over 300 million people. Arabic Natural Language Processing (NLP) is challenging and Arabic Sentiment Analysis is not an exception. Arabic is highly inflectional [Bassam et al. 2002; Mostafa et al. 2006] due to the affixes which includes prepositions and pronouns. Arabic morphology is complex due to nouns and verbs resulting in 10,000 root [Darwish 2002]. Arabic morphology has 120 patterns. Beesley [1996] highlighted the significance of 5000 roots for Arabic morphology. No capitalization makes Arabic named entity recognition a difficult mission [Shaalan 2014]. Figure (1) Internet World Stat (2015) depicts the overview of the top ten languages in the internet, with a clear significance of 155.6 million users for Arabic positioning it at Fourth place.

Figure (1): Internet World Stats



2.2 Forms of Arabic Language

Arabic is classified intro three types [Elgibali 2005]. Classical Arabic, Modern standard Arabic and Dialect Arabic are the three main types of Arabic [Abdal 2008; Farghaly and Shaalan 2009; Habash 2010; Korayem et al. 2012]. Arabic language in use takes three these forms that is Quranic Arabic, Colloquial Arabic and Modern Standard Arabic based on three key parameters including Morphology, Syntax and lexical combinations [Rafaee and Rieser 2014].

2.2.1 Classical Arabic

Classical Arabic has its significance in Arab nation in old age. Quranic Arabic also termed as Classical Arabic found in Quran. Religious texts and many Arabic manuscripts are written in Classical Arabic [Shaalan 2014].

2.2.2 Colloquial Arabic

Colloquial Arabic which is informally used on daily basis by Arabs and is found be different in different countries as well as different in different regions within a country [Shaalan 2014], is classified into Mesopotamian Arabic, Arabian Peninsula Arabic, Syro Palestanian Arabic, Egyptian and Maghrebi Arabic[2]. A colloquial Arabic language found to be used mostly by Internet users [Al-Kabi et al. 2014], which is found to be used in writing rigorously in applications of social media [Shaalan 2014], differs from region to region is Dialect Arabic. In dialect Arabic, some of the words are obtained from Modern Standard Arabic. [Farghaly and Shaalan 2009] illustrated the importance of building native tools to work on both Modern Standard and Dialect Arabic.

2.2.3 Modern Standard Arabic

Modern Standard Arabic is commonly found to be in use in radio, newspapers, television … etc. Most of the research so far in Arabic Sentiment Analysis focussed individually on Modern Standard Arabic [Almas and Ahmad 2007; El-Halees 2011; Rushdi et al. 2012]. Media, education, news books, so on uses Modern Standard Arabic [Al-Maimani et al. 2011].

[Al-Kabi et al. 2014] introduced a tool for sentiment analysis accepting Modern Standard Arabic or Colloquial Arabic or both. Overall very limited work is done so far whether it is

[2] https://en.wikipedia.org/wiki/Varieties_of_Arabic

done on Modern Standard Arabic or Colloquial Arabic when compared to English language [Al-Kabi et al. 2014]

2.3 Arabic language Features

[Daya et al. 2007] discussed how the words are derived from a root comprising consonants termed as radicals in number of three of four. Arabic is distinctive in terms of its history, diglossic nature, internal structure, inseparable link with Islam, and the Arabic culture and identity. Arabic linguistics structure is inherited and is far beyond its social aspects. Consists of 28 letters with 25 consonants and 3 long vowels. Arabic is the topmost Semitic language spoken in the world, see Table (1).

Table (1): Arabic Position in Semitic Languages

Language	**Number of Speakers**
Arabic	300 million Native Speakers [Owens 2013]
Amharic	22 million[3]
Tigrinya	7 million[3]
Hebrew	5 million[3]

[3] https://en.wikipedia.org/wiki/Semitic_languages

Reference to Semitic languages, Arabic being one among them, the grammatical features looks like:

- Arabic language accept three numbers for nouns that is singular, dual and plural[3].
- Arabic language follows right to left writing orientation [Dave et al. 2003; Wiebe and Riloff 2005; Harrag 2009; Barney 2010; Ghosh et al. 2012].

[3] https://en.wikipedia.org/wiki/Semitic_languages

- Verb tense for Arabic is imperative, perfect and imperfect [Shaalan et al. 2015]. Imperative verb tense for Arabic is a result of an immediate action. Verb tense for Arabic is perfect for a fulfilled action and is imperfect for a not fulfilled action.
- Nominative, accusative and genitive are three cases of nouns and adjectives found in Arabic language.
- Arabic when compared to other languages shares few features in common like SVO and VSO order, morphology, and short and long vowels [Ryding 2005].
- Nominal and verbal are two types of sentences in Arabic. Based on the part of speech of the first word of a sentence, the sentence is illustrated as nominal or verbal. Nominal sentence comes into existence with a subject and predicate with no verb.
- Classical Arabic traditionally follows verb–Subject-object word order[4]. Exceptions exist but not used in tweets.

2.4 Arabic Natural Language Processing and its challenges

[David and Alan 2010] stated that to blossom, assess and reason theories a domain which is used fascinatedly on one side, to being hard on the other side, is NLP. This is human centric computing. NLP is important for wider area including people from industry (business), computer linguistics, and humanities computing from academia (they recognise NLP as computational linguistics) [Steven et al. 2009].

Natural language processing, computational linguistics, text analysis and machine learning are some of the topics found to be embedded in Sentiment Analysis era.

Natural language processing tasks like information extraction and answering systems are found to be benefited from Sentiment Analysis [Pang and lee 2008]. Information Extraction deals with extracting useful information from a given text wherein Sentiment Analysis deals with analysis the text and predicting the polarity as negative, positive or neutral. Hence bridging the gap between Information extraction and sentiment analysis could be easily framed as extracting positive reviews from the given dataset.

Arabic Natural Language Processing has grabbed lot of attention from researchers due to Arabic language complex nature and the deficiency of resources. Very well supported by [Farghaly and Shaalan 2009], Arabic Natural Language processing has started to come into existence.

Some of the key challenges hindering the Arabic Natural Language Processing when comes to Sentiment Analysis includes:

2.4.1 Pre-processing affected due to the changes in shape of a letter

Letters changes their shape based on their position in a word. Almost all Arabic letters differ in their shape based on their appearance in a word. Many researchers do pre-processing of data set. Figure (2) depicts the different shapes of Arabic letters based on their position in a word. Hence pre-processing of Arabic data set in terms of performing Tokenization brings in new challenges resulting in satisfying different shapes of letters found at the middle, end and beginning[4]. Different positions brings in different challenges for Arabic letter when it comes to Arabic Sentiment Analysis as it effects the overall formation of rules in terms of defining different rules keeping in mind different shapes of Arabic letters based on their positioning. Hence pre-processing is done so as to bridge this barrier.

[4] https://en.wikibooks.org/wiki/Arabic/Arabic_alphabet

Figure (2) – Arabic letters structure in initial, middle and final

Stand-alone	Initial	Medial	Final	Name	Transliteration	Phonetic Value (IPA)
ء	أ ؤ إ ئ ٔا ؤ ئ, etc.			hamza	' / '	[ʔ]
ا	—		ﺎ	'alif	', a, ae	various, including [æː]
ب	ﺑ	ﺒ	ﺐ	bā'	b	[b]
ت	ﺗ	ﺘ	ﺖ	tā'	t	[t]
ث	ﺛ	ﺜ	ﺚ	tħā'	tħ	[θ]
ج	ﺟ	ﺠ	ﺞ	jīm, Other Accents:(djīm, gīm)	j	[ʤ] / [g]
ح	ﺣ	ﺤ	ﺢ	ḥā'	hž	[ħ]
خ	ﺧ	ﺨ	ﺦ	ḫā'	kħ	[x]
د	—		ﺪ	dāl	d	[d]
ذ	—		ﺬ	ḏāl	dħ	[ð]
ر	—		ﺮ	rā'	r	[r]
ز	—		ﺰ	zāī	z	[z]
س	ﺳ	ﺴ	ﺲ	sīn	s	[s]

2.4.2 Named Entities addition to lexicon list is challenged due to lack of capitalization in Arabic

[Farghaly and Shaalan 2009] illustrated one of the key features of Arabic language includes lack of capitalization, which is one of the rich feature of Arabic language. These rich feature does brings in new challenges for Arabic Sentiment Analysis [Almutawa 2015]. Lexicons used as a major data source in the Sentiment Analysis and Named Entities plays a major role in the lexicon creation. Due to lack of capitalization in Arabic, the automatic retrieving of named entities for the lexicon becomes a big challenge this very missed feature in Arabic bounds a person to attempt pre-processing or manual inclusion of the named entities to form a lexicon. This indeed brings in new challenges in the formation of rules, wherein the rules creation for mapping the named entities involves a tedious job due to lack of capitalization. For Example: “Ahmed” in English contains “A” as capital in the beginning of the letter, which helps in rules formation, by automatically retrieving al the words which starts with a capital letter and place them in the lexicon list. “احمد” in Arabic contains no such discretion. Hence while attempting to satisfy this in the rule formation of Arabic brings is an additional constraint due to absence of capitalization in Arabic, hence leaving the researchers to either pre-process or manual jot down these words.

2.4.3 Orthographic and morphology are the main Arabic challenges for processing by a computational system

Orthographic nature of Arabic language as quoted by [Mesfar 2010], “We can cite the case of the letter "T" which can be written in two different orthographies with the same pronunciation and also the Hamza with six different orthographies ("أ" ,"إ" ,"ؤ" ,"ئ" ,"ئ " ,"ء ") for the same pronunciation”. Hence while considering the making of lexicon for positive or negative or neutral, one need to ensure the addition of words written in different orthographies though they convey the same meaning. Morphology is one of the key challenges due to the very fact of Arabic being highly derivational and inflectional [Beesley 2001; Farghaly 1987; McCarthy 1981; Soudy et al. 2007; Abdel Monem et al. 2008; Shoukry and Rafea 2012; Farra et al. 2010; Riyad and Obeidat 2008; Abd Al Salm 2009; Farra 2010; Shaalan et al. 2004]. Arabic being derivational involves the derivation of verbs from two or three root words. Likewise, all adjectives and nouns are derivational as well. Hence, lemma = root + pattern. Arabic is inflectional in the sense that an inflected word in Arabic can have zero or more affixes, i.e. Word = prefix (es) + lemma + suffix (es). The shape of Arabic letter changes based on its position in the word [Kareem & Walid, 2013].

Arabic proper nouns could be either original or Arabized (transliterated or romanized). [Habash et al. 2009] stated that for Arabic has nearly 333,000 morphological tag sets are possible when compared to English which is only 50 tags shrouding all morphological variations. A Morphological Analyzer was introduced in [Rafea et al. 1993] for inflected Arabic.

[Rafaee and Rieser 2014] stated the significance of Morphological Analyser mostly tackle Modern Standard Arabic and the Natural Language Processing challenge with Sentiment

Analysis in Arabic arises with the tweets or reviews written in dialect Arabic and mixed or spelled incorrectly, where the morphological interpretations fails.

Hence, Arabic containing many prepositions, pronouns, … etc. as affixes or clitics makes Arabic highly inflectional, apart from this it has 10,000 root and 120 pattern making its morphology highly derivational. Due to the rich morphology, data sparseness is the obstacle hence resulting in many researchers opting for NLP tasks dataset to be pre-processed. Two basic methods followed to work on pre-processing is through light-stemming or tokenization.

Light-stemming is to get rid of affixes and leaving behind the stems whereas Tokenization is to just separate the affixes and the stems through space. [Larkey et al. 2007] discussed various techniques for light-stemming.

2.4.4 Encoding another hurdle

With the availability of Windows CP-1256 and Unicode, individuals can read, write and save text written in Arabic. Problems do arise during the processing of these texts by any other programs, thus the use of transliteration comes into existence where the mapping of Arabic letters and Roman letters are possible. Hence, so as to resolve this problem, many researchers use transliteration as the pre-processing step [Abdul-Mageed and Diab 2012].

2.4.5 Parser suffering due to missed diacritics

[Green and Manning 2010] conducted a research on parsing through better grammar engineering wherein diacritics plays a major role in conveying the meaning of the word. Most of the researchers uses parsers as one of the key step in the Sentiment Analysis process. Due to deficiency or clear miss on complete diacritics in the Internet reviews, ambiguity arises and makes the job of parser very complex hindering the tokenization of the diacritics [Aldayel and Azmi 2015].

As stated by [Molijy, 2012], more than one form in writing Arabic letter brings in additional challenges like the letter — ع has four forms (عـ ,ع ,ـعـ,ـع) resulting in ambiguity if not explicitly defined in the parser.

2.4.6 Punctuation, a clear miss in Arabic hindered Natural Language Processing with respect to Sentiment Analysis

With the very precocious features shared by Arabic, it brings in new challenges. One of the key challenges found in Arabic NLP is the lack of strict and rigid rules with reference to punctuation, bring in difficulties for pin point sentence boundaries in Modern standard Arabic [Shaalan 2010]. Same constraint is found in Dialect Arabic, with no rules having control over the sentence boundaries.

No punctuations are followed by most of the people except for addition of a full stop at the end, which even is quite often neglected. With no punctuations, huge effect is seen in most of the NLP tasks including Arabic Sentiment Analysis making the selection of correct sentences difficult from a given text. With No capitalization. With less Arabic corpus available, it makes it more challenging.

2.4.7. Insufficient Resources

Annotated resources and tools, a complete miss for Arabic language [Diab et al. 2007] hindering the Natural Language Processing in Arabic. Resources are not sufficient in terms of lexicons which plays a major role in Sentiment Analysis.

Due to the lack of completeness in the lexicon for positive and negative hinders the analysis of Sentiment. [Farra 2010; Rania and Girju 2012; Omar and Chris 2013; Nawaf et al. 2014] stated no matter researchers are spending energy and time creating lexicons and collecting tweets and reviews forming dataset, but these are not made available open for others to use and explore, hindered the growth in Arabic Natural Language Processing with Sentiment Analysis in context.

Table (2) depicts the lexicon built for some of the domains which is very limited not covering many domains and are not freely available. Table (3) depicts the datasets built for multi domain but are not openly available as well are very limited.

Table (2): Lexicons constructed are not complete

Paper	**Method used**	**Lexicon build for domain**	**Own/ Translated**
[Elhawary and Elfeky 2010]	unsupervised learning	business reviews	Own
[El-halees 2011]	By translating English lexicons into Arabic	English lexicons for various reviews.	Translated
[Abdul-Mageed and Diab 2012	By labelling 4000 adjectives as negative, positive and neutral.	4000 adjectives words irrespective of any particular domain	Own

Table (3): Unavailable Corpus for different domains

Domain	**Paper**
Financial	[Ahmad 2006], [Almas and Ahmad 2007]
Multi Domain	[El-Halees 2011], [Elsahar and El-Beltagy 2015]
Web Reviews	[Rushdi et al. 2011]
News Corpus	[Elarnaoty et al. 2012], [Farra et al.2010]
Twitter Tweets	[Aldayel and Azmi 2015],

2.4.8. Sarcastic Tamper

Sarcastic way of representing Arabic opinions or sentiments results in big challenges for the system to discriminate a review as positive or negative. The sarcastic interference causes a huge damage to the judgement of overall polarity. Due to sarcasm a negative review could be predicted as positive and vice versa. An example sarcastic tweet: "أوه نعم، وأنا أتفق أنه في بعض الأحيان أنا لا أستمع لما تقوله، وأنا مجرد مشاهد لصعود وهبوط الفك - Oh yes, I agree that sometimes I don't listen what you say, I just watch your jaw go up and down". Table (4) depicts some of the sarcastic comments which creates the polarity interpretation more cunning task.

Table (4): Sarcastic Tamper

Sarcastic Comments	
Arabic	**English Translation**
أحيانا أنا لا أعرف إذا كانت اصرخ عليك أو أشفق عليك	Sometimes I don't know whether to shout at you or pity you
أنت مضحك بحيث انك تجعل الجميع يضحكون عليك وليس معك	You are so funny that you are making everyone laugh at you, not with you
الفستان الذي كنت ترتديه يبدو لطيفا ولكن ليس عليك	The dress that you are wearing looks nice, but not on you.
أحيانان أحتاج لشيء لا يمكن أن يعطيه أحد إلا أنت وهو أن تغلق فمك	Sometimes I need something that no one can give except you, which is your mouth shut.

2.4.9. One word represents two polarities

In Arabic, addition or deletion of one word turns the sentence into opposite polarity. In the example A: "أنا أحب المدرسة - I like school" and "أنا لا أحب المدرسة - I do not like school". Likewise, not necessary the same word always tends to give a negative meaning.

In example B: “لا أحد يصعد الشجرة - No one climbed the tree” contains the word “ لا- No” appears in this sentence but does not convey any negative sentiment. Hence this set of contradictions with one word representing two different sentiments in two different sentences results in major concerns.

While satisfying a negative sentiment, one can satisfy the word “ لا- No” as negative sentiment by placing the word in negative lexicon, but fails to satisfy example B by analyzing the statement as negative though it is not. Table (5) depicts the illustration with examples on the challenges that one negative word can bring to the Sentiment Analysis task in Arabic Natural language processing.

Table (5): One word changing polarity - examples with challenges faced

Positive sentiment word	**Positive review**	**Addition of one word**	**Negative review**	**Challenges to tackled both negative and positive polarities**
لذيذ – indicating positive polarity.	الطعام لذيذ - The food is delicious	ليس Converts the polarity to negative	الطعام ليس لذيذ - The food is not delicious	Brings in challenges due to the addition of لذيذ – delicious in positive lexicon, hence the negative sentiment containing the word لذيذ – delicious makes the system identifies the negative

				sentiment as positive. if "ليس – not" not tackled properly so as to satisfy negative polarity.
سعيد – happy	أنت سعيد في عملك – You are happy with your work	Addition of لست converts the polarity to negative	وأنت لست سعيدا في عملك – You are not happy with your work	Brings in challenges due to the addition of "سعيد – happy" in positive lexicon, hence the negative sentiment containing the word "سعيد – happy" makes the system identify the negative sentiment as positive, if "ليس – not" not tackled properly so as to satisfy negative polarity.

2.4.10. Indifferent Writing Style

Arabic reviews mostly found on different sites have a clear diversity in writing style with Egyptian and Gulf Arabic quite often used [Rania and Girju 2012]. Most of the Arabic reviews are not written in standard format and doesn’t follow Modern Standard Arabic [Omar and Chris 2013].

People write different dialect in Arabic which are informal and thus involves the use of personal pronunciation and vocabularies [Albalooshi et al. 2011]. As a result, the lexicon building process and rule based approached are critically challenged with this informal writing styles leaving researchers with thorough analysis of reviews written online so as to perform Sentiment Analysis.

Table (6) depicts various writing styles along with the challenges they bring in to the system to satisfy the negative and positive polarities.

Table (6): Different writing style resulting in inexactness

Review/ Comment	**Polarity**	**Negative /Positive words**	**Challenges**
انها ليست امرأة سيئة - She is not a bad woman	Positive	Negative words - Not and Bad	This type of writing style containing negative words but indicating positive polarity makes the job for the system a difficult task to handle. Not and bad due to the very nature of these words indicating their significance in negative reviews, joins the negative lexicon. When the system tries to tackle this positive sentiment, but due to the addition of negative words the polarity system gives as an output is negative but the correct polarity is positive.

عليك أن تكون كريما - you have to be generous	Negative	Positive words - generous	This type of writing style with addition of positive words resulting in negative polarity brings in challenges wherein the system matches this with both positive and negative lexicon and indicates the polarity as positive though this sentence is negative, if need to be is not tacked to resolve this challenge.
أنت لست انسانن جيدا - You are not a good human being	Negative	Positive word – good Negative - not	System struggles to tackle the existence of words like good and not belonging to both positive and negative lexicon. Based on the way the system is designed, the overall polarity goes in favour of either positive or negative. If system indicate this as positive polarity then the answer is wrong, indicating that the classification lacks in handling mixed words in a sentiment.
ليس هناك احد حزين في هذا البلد - No one is sad in this country.	Positive	Negative words - No Sad	This example though include two negative words but fails to satisfy the overall polarity if the system analyse this based on the negative lexicon list and identify this as negative review though the review is positive resulting in inexactness.

2.4.11. Free writing style

Another key issue found to have an impact on Arabic Natural Language Processing especially in Sentiment Analysis, wherein people do not follow grammar, punctuations, and spellings. People are very often seem to update their status and comment online without breaks, this might result in a tendency to miss spell or write in hurry missing the basic structure of writing a proper sentence. This makes the Sentiment Analysis process more complex. Spelling mistakes are high in online reviews and tweets [Al-kabi et al. 2014].

Table (7) depicts few examples of free writing style with the spelling mistakes as biggest challenge.

Table (7): Challenges faced due to spelling mistakes in free writing style

	Lexicon	**Review in Arabic**	**Review in English**	**Challenges due to the spelling mistakes**
Actual Sentence	أفضل	أنت أفضل شخص في المبيعات.	You're the best sales person	Actual Polarity is positive, due to the word misspelled as ضل for أفضل, turns the whole polarity of the review into negative.
Spelling mistake	ضل spelled incorrectly for أفضل	أنت ضل شخص في المبيعات.	You are lost sales person.	

2.4.12 POS TAGGER FAILURE

Part of speech tagging involves the annotation of tagging segmented words as nouns, pronouns, verbs, adverbs and adjectives [Maamouri et al. 2004]. [Lamya and Hend 2012; Shoukry and Rafea 2012] used POS tagger to make a dataset. POS Tagger focussed on adjectives and adverbs [Turney 2002] wherein Sentiment Analysis also depends nouns and verbs [El-Beltagy and Ali 2013] left [Turney 2002] with poor results on untrained dataset using POS tagger. [Farra 2010] highlighted the need for a part of speech tagger which is sophisticated and can tag all parts of speech.

2.4.13 Words written in short forms

Due to the limits on many social media sites example Twitter, with regards to the number of words that could be written, brings in new challenges for performing Natural Language Processing. The way the Arabic tweets are shortened by putting the short forms for the words hinders the System to find the match for the sentiment in the Lexicons. Like " غرامة – Fine" is often trimmed and written as "F9" [Abdayel et al. 2015].

2.4.14. Same words often used to reply to a different sentiment

Sometimes people reply to a comment or review or tweet through one word or phrases which can support both positive and negative review, hence falling in both positive and negative lexicons. Not only the first step that is the lexicon building process becomes challenging, but also the system building becomes immensely challenging. Huge analysis is needed in this situation to handle such words. Words like "okay, fine, alright, all right, very well – حسنا, " أنا أتفق - I agree", "انا لا اوافق- I do not agree" are quite often seen to support sentiments in negative as well as positive. Building system understanding on when to retaliate to these supporting or not supporting words or phrases as negative or positive is a major hiccup. Table (8) depicts a response to twitter tweets and the challenges these responses brings to the system to predict the polarity of the response.

Table (8): Person to Person Response Challenge

	Twitter Tweet	**Polarity**	**Challenges**
Person - 1	ساره دائما تتظاهربأنها جيدة جدا في الطبخ لكنها ليست جيدة في	Negative	This is the challenging phase wherein making the system understand when to give a positive polarity if the word "

	الطبخ. – Sara always pretend to be very good in cooking but she is not a good cook.		حق – right" is found in a tweet, as this word belongs to the positive lexicon, but in the current case it supports the negative polarity. Hence illustrating a negative sentiment by supporting a negative statement.
Person – 2 Responding to person -1	انت دائما على حق – You are always right.	Supporting a negative opinion – negative polarity	
Person -3	اعتقد انها لا تعرف الف باء الطبخ - I think she does' not know the abc of cooking	Negative comment	Person -3 is negative sentiment. This negative sentiment is not supported by person-4 instead gives a positive response by not agreeing to person-3. Due to the addition of " I do not agree - أنا لا أتفق معك ", brings in new challenges as the system will predict this as negative, due to " I do not agree - أنا لا أتفق معك " which falls under negative sentiment and the system will consider this to be a negative review but in actual this is a positive response to a negative review.
Person -4 responds to person -3	أنا لا أتفق معك - I do not agree with you	Positive response do not supporting a negative sentiment	

2.5 Chapter Summary

This chapter adds to the research in progress for Arabic NLP due to the rich morphology and highly derivational challenges. In this chapter, many challenges to Arabic Natural Language processing in Sentiment Analysis is presented. These challenges includes the rich features of Arabic language covering different structure for Arabic letters based on the position of the letter in initial, middle and final. As well features like lack of capitalization in Arabic results in challenges in the rules formation in the mapping of named entities.

Orthographic and morphology constraints for processing by a computational system, encoding Arabic language, parsing a big challenge due to missed diacritics, punctuation missed completely in Arabic writing, insufficient resources, sarcastic comments or reviews, one word representing two polarities, indifferent writing style, free writing style and POS tagging a big failure. Arabic linguistic resources are very limited when talked about lexicons and datasets. This chapter is going to enlighten the researchers as to what could be taken care off while attempting to perform Sentiment Analysis in Arabic. One of the research questions addressed in this dissertation is to measure the impact of elimination of the pre-processing of dataset which is a major step followed by many researchers to tackle the above challenges.

Chapter Three

Subjectivity and Sentiment Analysis

Sentiment Analysis is nevertheless a growing era founding its way out in today's digital era. Sentiments does play a major role in presenting ones emotions and feelings. These sentiments comes into existence from the discretion of a word belonging to the state of emotion or conveying an opinion or stimulating and expressing feelings on anything which matters to an individual or a group or a company or an organization or a business. This type of emotional or sentimental state belongs to subjective sentiment unlike objective which doesn't communicate anything emotional or any feeling. This chapter covers an overview of subjective and sentiment Classification along with the importance of sentiment analysis and the challenges it brings in. Different approaches to perform sentiment analysis and the state of art are also highlighted in this chapter.

3.1 Subjective Classification

As stated by [Pang and Lee 2008], Sentiment Analysis relied on two major phases, one of them being Subjective Classification and other one being Sentiment Classification. Subjectivity classification of a text includes the assignment of a text as objective and subjective. Mental and emotional states are represented as private states [Quirk et al. 1985], the significance of these private states in a sentence indicates that the sentence is Subjective. Sentence level subjective analysis includes the existence of private state in a sentence. A text is classified as objective if the text does not illustrate any feelings, opinions, or emotions wherein if the text depicts any of them then the text is subjective. [Roty 1997] presented a clear description on the term objective as the illustration of anything the way it is rather than the way they are presented. The text " سيارة جيس تم طرحها في السوق في يناير 2015" (CIAZ car was launched in January 2015) is objective. On the contrary, the text "السيارة جيس جيدة جدا" (CIAZ car is very good) is subjective. Once the text is classified as subjective, further steps are followed to classify the sentiments based on polarity.

[Muhammad et al. 2012] during the subjective phase used a binary classifier which successfully segregated objective from subjective occurrences. SAMAR system for subjectivity and sentiment analysis, proposed by [Abdul-Magged et al. 2012], depicted the complexity of Arabic language in subjectivity and sentiment analysis.

3.2 Sentiment Classification

The text which is classified as Subjective, is further worked on to classify as positive, negative, or neutral. The text "الطعام في مطعم سيزارس لذيذ جدا" - (CESSARS restaurant food is very delicious) and the text "والدتي هي أفضل طاهية" – (My mother is the best cook) is positive. On the contrary, the text "التعليم في مدرسة النجاح سيء جدا" - (The education in Al-Nagah School is very poor) and the text "وانا سأصبح مجنون مع هذا الغش" يا!" - (Oh! I am going to be mad with this cheating) is negative. The text " علي الرغم من أن الكلية توفر تعليم ممتاز لكن الأنشطة المتعلقة بالمناهج الإضافية جعلت بعض الطلاب يحيدون عن الدراسة" (Though the college provides excellent education but the extra-curricular activities led some students to deviate from studies) is mixed. The text "قد تغلق المدرسة" (The school may shut down) is neutral.

3.3 Importance of Sentiment Analysis

Sentiment analysis is self–extracting. With the very use of social media sites, reviews are shared on a large scale in different languages. Word, document and sentence level opinion analysis are the three levels at which opinions are studied [Xiowen et al. 2008]. Reviews are not confined to a particular entity, indeed reviews covers a huge area including products [Qui et al. 2011], business applications systems and recommender systems [Glance et al.

2005], services, news and media [kim et al. 2006], message filtering [Molouf and Mullen 2008] and very commonly [Somasundaran and Weibe 2008] in music, brands, videos, restaurants, colleges, universities, schools, and so on.

Hence putting forward, the need to extract the opinions, sentiments and emotions from the reviews, so as to deal with different issues or concerns. These reviews thus helps the businesses to keep track on their products or services as well helps them to strengthen their weak areas. Though Sentiment Analysis has been studied explicitly well in English, but Arabic still remains challenging with comparatively less work done in Arabic so far.

3.4 Sentiment Analysis Classification levels

Sentiment analysis operates at different classification level which includes phrase, sentence or document level. Different approaches have been deployed to different levels to find the polarity. There are two main approaches found to perform the Sentiment Analysis, one being supervised approach and other being unsupervised approach.

Supervised approach are also called as Machine Leaning Approach. Unsupervised approach is also termed as Semantic Orientation Approach. These two are discussed in detail in the following chapter along with the proposed approach in this dissertation. Other than this hybrid approach is also seen to be used with the combination of best features of semantic orientation approach and machine learning approach. Table (9) depicts the different levels worked upon in the literature.

Table (9): Summary of Different Classification levels in the Literature

Paper	**Classification Level**
[Ahmad, 2006]	Phrase
[Almas &Ahmad 2007]	
[Abdul-Mageed and Korayem 2010]	Sentence

[Aldayel & Azmi 2015]	
[Farra et al. 2010]	
[Abdul-Mageed et al. 2011]	
[Abdul-Mageed and Diab 2011]	
[Abdul-Mageed and Diab 2012]	
[Abbasi et al. 2008]	Document
[Farra et al. 2010]	
[Elhawary and Elfeky 2010]	
(El-Halees 2011]	
(Rushdi et al. 2011)	

3.5 Sentiment Analysis Challenges

To start with the challenges, the first one that has a direct influence is subjectivity classification. Due to the very fact of a trim variance in a sentence being subjective or objective brings in new challenges. Many other challenges are seen in Sentiment Analysis. If a negative sentiment contains an entry from the positive lexicon, it is overall graded as positive polarity wherein it is negative. For example, "شو يعني بدك ايانا نحكي من وراكي مش حلوة" (What is the matter? Do you want us to talk behind your back? Not cool!) is a negative sentiment but it contains the positive word "حلوة" meaning "sweet" (cool) in English.

A word found as an entry of both positive and negative lexicons brings in new challenges in the polarity assignment. For example, "با قرف" (it makes me sick or disgusting) was found in both negative and positive reviews; ultimately, joining positive and negative lexicons. Some people write sarcastic comments or different style of commenting using negated sentences in a positive review thus brings in new challenges. For example, " احسن من هيك رئيس وزراء ما في" (There is no Prime Minister better than this one).

As illustrated by [Pang et al. 2008] positive words in one domain might hold a different polarity in another domain. For example, a word like "بارد" (cold) were found in positive and negative reviews holding respective polarity. Other challenges includes, the tweet limit to 140 characters people ends up using short forms, grammatical mistakes in the tweets and most importantly use of slang language [Aldayel and Azmi 2015].

3. 6 Sentiment Analysis State of Art

Study done so far reveals lots of input from the researchers for English Sentiment Analysis. Arabic still remain center of attention. As the results and significance of state of art for Arabic Sentiment Analysis is not available we will be looking at some of the important aspects and work done so far.

Evaluation on results is done through different evaluation metrics. So as to work on Sentiment Analysis, the key parameter is the dataset. Recent efforts by [Farra et al. 2015] illustrated the importance of crowdsourcing as a very successful method for annotating dataset. Another important parameter being the lexicon, recently [Gilbert et al. 2015] introduced ArSenL lexicon achieved 67.3% accuracy on Sentiment Classification. Support Vector Machine classifier achieved 72.6% accuracy on twitter dataset of 1000 tweets [Shoukry and Rafea 2012].

[Alaa , 2011] worked on document level sentiment analysis using a combined approach consisting of a lexicon and Machine Learning approach with K-Nearest Neighbors and Maximum Entropy on a mixed domain corpus comprising of education, politics and sports reached an F-measure of 80.29%. A phrase-based sentiment analysis presented by [Muhammad and Mona 2012] achieved a precision of 60.32% using pattern matching approach on Penn Arabic Treebank, WIKI talk pages and web forums.

[Neura et al. 2010] worked on sentence and document level using grammatical and sematic approach. [Shoukry and Refea 2012] followed a corpus-based approach, achieved an accuracy of 72.6%. Positive, negative and neutral polarity classification done by [Ahmed et al. 2013] using n-gram feature selection and stemming though achieved some good results but the dataset collected was very less. [Nawaf et al. 2014] presented a lexicon and sentiment analysis tool with an accuracy of 70.05% on tweeter dataset and 63.75% on Yahoo Maktoob dataset.

[Abdayel and Nazmi 2015] using a hybrid approach that is Lexical and Support Vector Machine classifier produced 84.01% accuracy. [El-Halees 2011] achieved 79.90% accuracy with Hybrid approach comprising of lexical, entropy and K-nearest neighbor. [Shoukry and Rafea 2012] separately deployed two approaches, one being Support Vector Machine achieved an accuracy of 78.80% and other one being Lexical with an accuracy of 75.50%.

3.7 Chapter Summary

This chapter not only presents the difference between a subjective and an objective sentence but also depicts the importance of a subjective sentence forming a base in the analysis of a Sentiment. Challenges and Importance of Sentiment Analysis are also covered in detail. The significance of private state that is an opinion or feeling or emotions make the sentence subjective, adds to the analysis of this sentence to be either negative or positive. This chapter also highlights the importance of Sentiment Analysis in various fields including social media, government, online shopping, and online banking, politics etc., due to the rise in online comments and reviews on various products or service. People today rely on others opinions before buying a product or service. Classification levels including phase, document and sentence are also depicted in detail. Literature review in Arabic Sentiment Analysis covered in this chapter outputted the need for a novel approach to perform sentiment analysis in Arabic.

Chapter Four

Subjectivity and Sentiment Analysis

Supervised and Unsupervised approaches commonly used to perform Sentiment Analysis. Supervised approach also known as corpus based and unsupervised also known as lexicon based approach includes algorithms to perform Sentiment Analysis.

4.1 Supervised Approach

Supervised approach is also termed as Corpus-Based approach includes different classifiers. Decision Tree, K-Nearest Neighbor, Support Vector Machine … etc. are different Machine Learning classifiers. Machine Learning though includes many algorithms but the main challenge is to discover appropriate feature set.

Two types of approaches are commonly seen in supervised technique, they are the model approach and relationship-based approach, or merge of these approaches. A feature vector comprising of one or more features, is used to represent a text which will be undergoing analyses in a supervised technique. [Taboada 2011] stated that supervised approaches for sentiment analysis are uncommon case of text classification.

The process involved in supervised approach is based on division of dataset as training and test dataset. Machine Learning algorithms involve a selected set of features extracted from training dataset annotated with polarity in order to generate statistical models for sentiment classification prediction. Big dataset annotated by native speakers are the key success to supervised approach but nevertheless the annotation gets hindered due to sarcasm, time consumption and costly.

[Bo et al. 2002] putted in the tutoring Internet Movie data to Naïve Bayes, Support Vector Machine and Maximum Entropy so as find the polarity.

In the training documents, K is the number which represents the number of nearest neighbors which classifies an unannotated document with the help of k neighbors [Duwari and Islam 2014]. Here the calculation between similarities of unlabeled documents and the left out documents in the training dataset is done. As a result of this only the K similar documents are looked attentively at and most importantly majority voting or weighted average are taken into a new document. Japanese emotions were classified using K-nearest neighbor [Tokuhisa et al. 2008].

[McCallam and Nigam 1988; Rish 2001] used Naives Bayes classifier produced good results, which classifies based on the hypothesis that the prediction variable is autonomous.

[Fung and Mangasarian 2002] illustrated support vector machines as a mechanism to find a hyperplane this hyperplane is very well depicted as a vector that is found to more apart document vector belonging to one class from document vectors found in other classes. SVM^{light} was the classification algorithm found to be used by [Abdul-mageed et al. 2012], so as to perform Arabic Sentiment Analysis. [Wan 2009] produced good results with respect to Chinese sentiment analysis with the use of SVM classifier. Review classifier was used by [Elhawary and Elfeky 2010], in their attempt to perform Arabic Sentiment Analysis on business reviews.

4.2 Unsupervised Approach or Lexicon-Based Approach

Unlike supervised approach, unsupervised approach involves description of positive and negative lexicons based on the highest count. The word with highest count in a document is designated to be positive or negative.

A very famous tweaked way of unsupervised approached was seen through thumbs up and thumbs down presented by [Turney 2002]. In the unsupervised approach, positive is designated as polarity 1, negative (-1) and 0 for neutral. Even could be scaled on the strength polarity ranging from +1 to +5 for positive polarity, with +5 being the most positive one when compared to the +1 polarity. Automatic and non-automatic methods do give the lexicons (Dictionaries).

For automatic generation of the lexicons a root list is deployed on the similarity techniques to extend this list, ensure not to include the neutral words in the root list. The key method follows here with the very significance of polarity assignment to each word to the lexicon and then summing all the scores illustrating the sentiment of full document or text. [Korayem et al. 2012] pointed out that the lexicon-based approach is practical but does not fit well in all the domains.

4.3 Chapter Summary

K- Nearest neighbor, Naives Bayes, Maximum Entropy classifiers etc.. are used in supervised approaches. Some of the unsupervised approaches used by many researchers includes thumbs up and thumbs down, polarities like -1 indicating negative, +1 indicating positive and 0 indicating neutral.

Chapter Five

Data Collection

Data collection is the most critical and most challenging part in most of the sentiment analysis task. Sentiment Analysis includes two sets of data collection, one of which is the lexicon and the other one is the review or comment or twitter datasets. This chapter covers the lexicons and datasets overview. This chapter depicts the methodology followed for data collection and the method followed to improvise the lexicon.

5.1 Lexicon and Datasets

Lexicons mostly populated with adjectives are found to be very useful. [Benamara et al. 2007] stated that the adjectives are good signals to trace the polarity of a text. [Elhawary and Elfeky 2010] through unsupervised learning brought into existence lexicon for business reviews in Arabic using similarity graph method.

Wherein [El-halees 2011] proposed a lexicon for Arabic based on sentiment strength scale, termed as SentiStrength, and online dictionary, where English lexicons are translated into Arabic. Sentistrength is the strength of the review that is whether the review is highly positive or negative, average, or below average. [Abdul-Mageed and Diab 2012] created their own lexicon non-automatic, by labelling 4000 adjectives words as negative, positive and neutral. [Elarnaoty et al. 2012] put forward an Arabic news corpus, termed as MPQA. Annotation was done manually for the opinion holder corpus.

The pre-processing for fixing the morphological challenges and part of speech tags were done using RDI (Research and Development International) tool. One of the focus of this research is to exclude pre-processing step and see the significance of results with the addition of heuristic rules.

Table (10) depicts the overview of corpus in use based on our literature review.

Table (10): Overview of corpus genre in use

Paper	**Corpus Genre**
[Ahmad 2006]	Financial
[Almas and Ahmad 2007]	Financial
[Farra et al.2010]	News
[El-Halees 2011]	Multi Domain
[Rushdi et al. 2011]	Web Reviews
[Elarnaoty et al. 2012]	News Corpus
[Aldayel and Azmi 2015]	Twitter Tweets
[Elsahar and El-Beltagy 2015]	Multi Domain
[Nabil et al 2015]	Twitter Tweets

5.2 Data collection methodology

This chapter presents the methodology illustrating the actions performed to get the base in place so as to implement and test the proposed Arabic Sentiment Analysis System. Foremost the dataset is presented. Secondly, the lexicon construction is included.

5.2.1 Dataset in Hand

The dataset used in this paper is Twitter Tweets and movie reviews. Tweets are status messages created by humans on Twitter [Alec et al. 2009]. The dataset is taken from [Nawaf et al. 2014] and [Rushdi et al. 2011]. These datasets are used as these are available, rich and enough to come to a conclusion, as well electronic resources such as lexicon is provided. [Hazem et al. 2015] stated there is no freely available lexicon other than ESWN. Other researchers do not offer their data or it is not enough. Results are compared with [Nawaf et al. 2014]. As we noticed, that the results thus achieved by them are not the best. The system is tested on both OCA corpus [Rushdi et al. 2011] and [Nawaf et al. 2014]'s dataset.

[Nawaf et al. 2014]'s dataset contains 1000 positive tweets and 1000 negative tweets. 7189 words in positive tweets and 9769 words in negative tweets. The tweets collected were written in Modern standard Arabic and Jordanian Dialect, which covers Levantine language family. OCA corpus, termed as Opinion Corpus for Arabic, was presented by [Rushdi et al. 2011]. This corpus contains total 500 opinions, of which there are 250 positive opinions and 250 negative opinions. Table (11) depicts the statistics of the tweets taken from [Nawaf et al. 2014]. Table (12) depicts the overview of opinions collected by [Rushid et al. 2011]. OCA corpus is used in this paper during the testing phase.

Table (11): Statistics on the tweets dataset [Nawaf et al. 2014]

	Positive	**Negative**
Total tweets	1000	1000
Total words	7189	9769
Avg. words in each tweet	7.19	9.97
Avg. characters in each tweet	40.04	59.02

Table (12): Distribution of reviews crawled from different webpages [Rushid et al. 2011]

	Name	**Webpage**	**Rating system**	**Possible reviews**	**Negative reviews**
1	Cinema Al Rasid	http://cinema.al-rasid.com	10	36	1
2	Film Reader	http://filmreader.blogspot.com	5	0	92
3	Hot Movie Reviews	http://hotmoviews.blogspot.com	5	45	4
4	Elcinema	http://www.elcinema.com	10	0	56
5	Grind House	http://grindh.com	10	38	0
6	Mzyondubai	http://www.mzyondubai.com	10	0	15
7	Aflammee	http://aflamee.com	5	0	1
8	Grind Film	http://grindfilm.blogspot.com	10	0	8
9	Cinema Gate	http://www.cingate.com		0	1

10	Emad Ozery blog	http://emadozery.blogspot.com	10	0	1
11	Fil Fan	http://www.filfan.com	5	81	20
12	Sport4Ever	http://sport4ever maktoob.com	10	0	1
13	DVD4ArabPos	http://dvd4arab.maktoob.com	10	11	0
14	Gamraii	http://www.gamraii.com	10	39	0
15	Shadows and Phantoms	http://shadowsandphantoms.blogspot.com	10	0	50
Total				250	250

5.2.2 Lexicon Improvised

Lexicon does makes an explicit contribution to cover appropriate rules. The lexicon used in this paper contains the lexicon used by [Nawaf et al. 2014], which included sentiments, named entities and some randomly placed words. We further analyzed the dataset looking for evidence supporting our approach. [Nizar and Wassim 2013] built a lexicon with a united approach of combining two resources Arabic WordNet (AWN) [Black et al. 2006] and SAMA [Graff et al. 2009] lexicon. Based on the same grounds to combine different lexicons we improvised [Nawaf et al. 2014]'s lexicons with the addition of words from the dataset which were found to appear more than once. Based on the discretion of repetition of these words and their placement, they were included in both the list that is positive and negative lists. For example, “كاتب” (Writer) was repeated in both negative and positive reviews, positioned to be at the end in negative reviews and within the text in positive reviews. Hence, it was included in positive and negative lexicons.

But the major turnover was in the rules thus created. The words found to be common based on the analysis of the tweets were included in both positive and negative lexicons which were justified during the rule creation phase. One key addition to the lexicon used in this dissertation was the addition of words in the list which were single words that is the tweet which included only one word. Based on the polarity the words were thus placed in the appropriate list.

5.3 Chapter Summary

This chapter presented the methodology illustrating the actions performed to get the base in place so as to implement and test the proposed Arabic Sentiment Analysis System. Foremost the dataset was presented. Secondly, the lexicon construction was included. The data collection methodology included in this chapter covers the baseline dataset that is [Nawaf et al, 2014] dataset and OCA dataset. Lexicon building process is illustrated in detail. Adjectives are found to be very useful in building the lexicons for Sentiment Analysis, as they convey some kind of sentiment but there are other words like common words found in both positive and negative lexicon, which play a major role in uplifting the results to another level.

Chapter Six

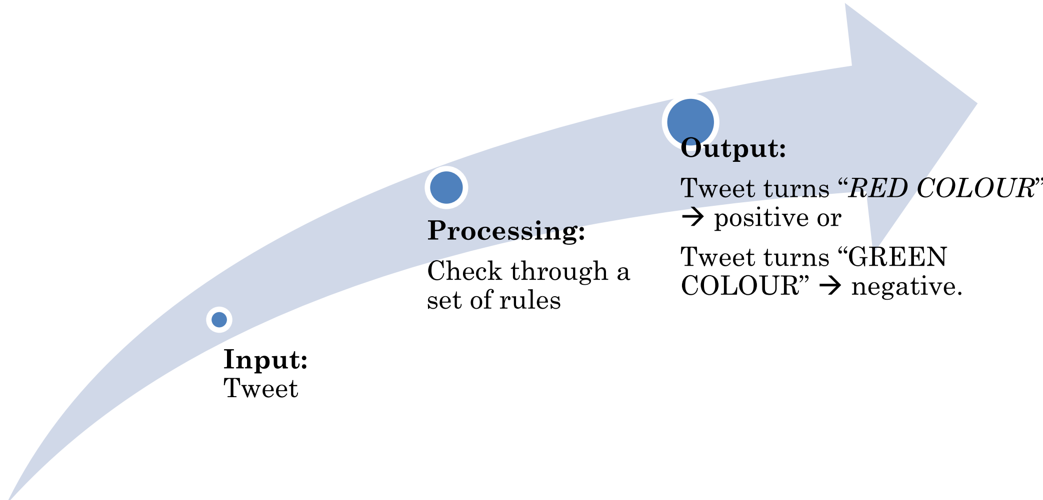
Implementation of Arabic Sentiment Analysis

This chapter presents the implementation of Arabic Sentiment Analysis. The rules formed and the significance of end-end rule chaining mechanism is illustrated in detail.

6.1 The system at a glance

The spreadsheet rule based system proposed in this dissertation included three key phases to display the results after processing. The first phase is to enter the tweet. The second phase includes rule based check wherein the entered text is passed through a set of rules checking for the tweet polarity either negative or positive. The third phase is to turn the text into “light red fill with Dark Red text” indicating text is positive or “Green fill with Dark green text” indicating text is negative. Figure (3) depicts the overview of the system proposed in this dissertation.

Figure (3): Overview of the system proposed



6.2 Rule-based approach

The approach followed in this dissertation is rule-based, we refer the reader for [Shaalan 2010] for a review about the significance of the rule-based approach and how it is used to apply different NLP tasks. Lot of research till date included pre-processing as the major steps. In this dissertation, with the very use of appropriate rules the pre-processing step was totally eliminated for the dataset, there were no amendments done to the dataset. The rule-based approach include patterns to look the entered tweet up in the lexicon.

6.3 Methodology for the formation of rules

The very existence of the methodology that we have proposed in this dissertation, has brought to life an innovative mechanism, so as to address the critical and undealt issues with the lexicon-based approach. Both the polarity lexicons and polarity reviews are analyzed. The methodology in this research covers many crucial areas found to be missing in the literature review that we conducted. Firstly, the rules are not confined to search within text. Secondly, rules are not excluding common words found in both positive and negative lexicon and lastly, the proposed rule-based system covers all the directions where the word has been found to have a huge impact. To best put the above three key areas into practice end-to-end mechanism with rule chaining approach is established.

End-to-End mechanism with rule chaining approach introduced in this dissertation, includes the chaining of rules based on the positioning of the polarities in the tweets. The rules includes a 360 degree coverage: 1) in the middle, we termed as “within the text", 2) at the boundary we termed as either "ending with the text" or " beginning with the text", and 3) full coverage, we termed as "equal to the text". The key underlying base ground factors which helped us formulate appropriate rules includes analysis of the tweets and the extension of positive and negative lexicons. The analysis of tweets resulted in identifying relations pertaining to words which were either disjoint, intersected or coexisted.

The words which were disjoint that is completely indicating either positive or negative polarity were included in their respective lexicons. The words which intersected that is the ones which were found to be common in both negative and positive reviews were included in positive as well as negative lexicons. The words which coexisted at the same place in the negative and positive reviews, that is the ones which appeared at the beginning or ending were placed in either positive or negative lexicon, based on the highest frequency of the word in the respective reviews.

6.3.1 Rules handling intersection with the end-to-end chaining mechanism

As an example consider the following positive tweet “اوقف القرار للحفاظ على الوطن” (the decision was suspended for protecting the motherland), the word “اوقف” (suspended) appeared in the beginning of this tweet and the same word was found in the negative tweets. Hence, this set of situations was handle with the very use of positioning and chaining of rules.

The steps involved to achieve the positioning and chaining of rules includes:

- Rules Formation: With the logical discretion of the word “اوقف” (suspended) being seen at the “beginning of” positive tweets and “within the text” for the negative tweets, the rules thus formed were “beginning with” for positive tweets and “within the text” for negative tweets.

- Positioning and Chaining of rules: In the current example “beginning with” rule needs to be positioned and chained in an orderly fashioned with the rule “within the text” so as to satisfy both positive and negative reviews. So, the rule “beginning with” was chained with the rule “within the text” for the word “اوقف” (suspended) by positioning the rule “beginning with” first followed by “within the text” rule. Hence, the rules are chained and positioned for the words which are found to repeat themselves in both positive and negative tweets.

6.3.2 Rules handling coexistence

The lexicons which are not seen to repeat themselves in either positive or negative tweets where handled with the rule – “within the text”. For example, “الحرامية” (Thieves) is set for search “within the text” “الحرامية” (Thieves), is majorly found within the text in negative review rather than in positive ones. Hence, based on the frequency of “الحرامية” (Thieves), the rule is set. Example tweet: “ انا مش مسؤول عن الحرامية يتفنجرو بمصاري البلد و انا اشحد” (I am not responsible for that thieves spend lavishly from the funds of the country and I beg.)

6.3.3 Rules handling disjoint

The cases wherein the words were not repeated in positive and negative cases and were found to have their significance at the end of the tweet, were handled using “ending with” rule. In the following positive tweet example تذكر من يكذب يكتب عند الله كذابا – اما من يتحرى الصدق يكتب عند الله”صادقا“ ” – (Remember who continues to speak falsehood he is recorded with Allah as a great liar – but who persists in speaking the truth he is recorded with Allah as an honest man). The word “صادقا” (Honest) appeared at the end of this tweet. Likewise, “صادقا” (Honest) was found to appear at the end in majority of positive reviews. Hence, the “end with” rule is set to search for “صادقا” indicating the system that if the review ends with the word “صادقا” (Honest) then it should be considered as a positive tweet.

6.3.4 Derivation of rules

Table (13) depicts the skeletal examples of derived rules. Conditional search includes two key phases, the first phase includes a condition which checks on the entered tweet mapping with the rules and deciding on whether the entered tweet/review is positive or negative. The second phase is the color coded output display of the entered tweet/review to indicate the polarity as positive or negative. The entered tweet changes its font color to – "Red for positive tweets and "Green" for negative tweets.

Table (13): Derivation of lexicalized Rule

Rule Type		**Derivation of Rules**
1) Search within the text	**A]**	**If** the Entered Tweet/Review *contains the word* " __ " **then** Fill the text with-"Red" for positive tweets
	B]	**If** the Entered Tweet/Review *contains the word* " __ " **then** Fill the text with- "Green" for negative tweets
2) Ending with the text	**A]**	**If** the Entered Tweet/Review *ends with the word* " __ " **then** Fill the text with- "Red" for positive tweets
	B]	**If** the Entered Tweet *ends with the word* " __ " **then** Fill the text with- "Green" for negative tweets
3) Beginning with the text	**A]**	**If** the Entered Tweet *begins with the word* " __ " **then** Fill the text with- "Red" for positive tweets
	B]	**If** the Entered Tweet *begins with the word* " __ " **Then** Fill the text with- "Green" for negative tweets
4) Equal to the text	**A]**	**If** the Entered Tweet *equals to the word* " __ " **Then** Fill the text with- "Red" for positive tweets

	B]	**If** the Entered Tweet *equals to the word* " __ " **Then** Fill the text with-"Green" for negative tweets

6.3.5 Description of rules

With reference to rule 1.A in Table (13) in the first row, if the word appeared within the text in a positive tweet. With reference to rule 2.B or 3.B in this table, if the same word appeared at the end or at the beginning in the negative tweet, then positioning and chaining of these two rules are done. For this case the positive rule was positioned and chained below the negative rule.

The rule "beginning with" or "ending with" unlike "within the text" only look for that word in the beginning or end which will tag the tweet as negative. If the placement of rule is reversed then "within the text" as positioned and chained before the rules "beginning" or "ending" with, when a negative tweet is passed through this, the tweet will be tagged positive. Hence positioning and chaining of rules places a key role. As this is a chained approach for a particular word the processing for the above case will be when a tweet is entered it undergoes the chain of rules for a specific word, which first checks its significance at the ending or beginning rule, if it does not belong to that then it undergoes the next rule which looks for the same word within the text.

6.4 End-to-end mechanism with rule chaining approach

After carefully executing the rules based on the satisfaction of disjoint word(s) which coexisted or common words in negative and positive reviews, the rule chaining was taken care of. End-to-end mechanism with rule chaining approach satisfies a word which belongs to a positive and negative review irrespective of its actual polarity. Hence, a negative word in a positive review and a positive word in a negative review is satisfied with appropriate chaining as discussed below with the aid of examples from negative and positive tweets.

Example 1 Rules in use – Rule A and Rule B

- Rule A in Table (13): In the negative tweet " ليست هناك حاجة لحماية وطننا من أيدي المنافقين الأردنيين" (There is no need to protect our motherland from hands of hypocrites Jordanian), contains the word "المنافقين" (hypocrites) within the text.

- Rule B in Table (13): In the positive tweet " اللهم لا تجعلنا من المنافقين"(O Allah! Place us not with the people who are hypocrites) contains the word "المنافقين" (hypocrites) at the last.

- Need for chaining: If Rule A is not chained with Rule B, then only one of the rule will be satisfied. To satisfy both the rules, that is to correctly identify negative and positive polarity, Rule A is chained with Rule B by positioning Rule A below Rule B, so as to allow the search to first visit the end rule first, then the within the text rule.

Example 2 with explanation on how the system works:

For example, the word **"**تشائم" (pessimism) was found to be part of both positive and negative reviews with a slight variation.

- In the positive reviews appears only in the middle whereas in the negative review "تشائم" (pessimism) was found to appear at the beginning. Hence the rules thus were created covering "within the text" (refer rule 1.A in Table (13)) and "beginning with" (refer rule 2.B in Table (13)), but the positioning was varied. By positioning the rule "beginning with" first and then the rule "within the text" helped in satisfying both positive and negative reviews.

- As the system checked for the word beginning with "تشائم" (pessimism) in the entered review and if found then the review was tagged as negative. Likewise, the system proceeded with the entered review containing the word "تشائم" (pessimism), if the word "تشائم" (pessimism) was not found at the beginning then the search proceeded further and identified it as positive.

6.5 Chapter Summary

This chapter presents the system implementation which involves three key phases, the first phase was the start of the process where the tweet/review is entered, second phase is the processing of the tweet/review which undergoes set of heuristics rules and last phase include the output which indicates whether the tweet is negative or positive. All the rules used in the formation of a system are covered in detail with the aid of examples. Along with the positioning, end-end chaining with prioritizing of rules, this chapter also presented the way the rules handled disjoint, co-existence and intersection relations.

Chapter Seven

Experiments and Results

This chapter presents the evaluation of the system developed. Precision, Recall and Accuracy are the evaluation metrics used to evaluate the results. Experiments done are analyzed and research questions answered in this chapter. Illustrative error analysis with the aid of examples are also included so as to provide researchers with a clear direction on what next.

7.1 Evaluation

To measure the improvement as well the quality of being trusted and believed in, evaluation plays a vital role. *Cross-Validation* and test data are very often used to evaluate the results in sentiment analysis. Cross-Validation is mostly found to be useful in the non-existence of standard gold dataset for testing or if there is not enough amount of data for testing and training (Pang & Lee, 2004).

7.2 Evaluation Metrics

The accuracy measures – Precision, Recall and Accuracy, which are commonly in use was deployed to measure the performance of the tools used in both the experiments. [Al-Kabi et al. 2013; Nawaf et al. 2014; Mohammad et al. 2013; Shoukry and Rafea 2012; Rushdi et al. 2013] have used Precision, Recall and Accuracy to compare their results. The equation are as follows:

Precision = TP/ (TP + FP)

Recall = TP/ (TP + FN)

Accuracy = (TP + TN)/ (TP + TN + FP + FN)

Where:

TP- True Positive, all the tweets which were classified correctly as positive

TN- True Negative, all the tweets which were correctly classified as negative

FP- False Positive, all the tweets which were incorrectly classified as positive

FN- False Negative, all the tweets which were incorrectly classified as negative.

7.3 Experiments Conducted

System 1 was built first. So as to improve the performance, *System 1* was improvised and *System 2* was brought into existence. Each system underwent two sets of experiments.

- *System 1*: It included only two types of key rules which were written based on "equal to" or "within the text" containing lexicon entries. For example, if the entered tweet contains the word "أصدق" (I believe), which is listed in the positive sentiment lexicon, then the polarity is positive. On the whole, if the tweet text contains the words or is equal to a word from the positive list then the text turned into "light red fill with Dark Red text" indicated the tweet is positive. If the tweet text "contain words" or "is equal to" from the negative list, then the text turned into "Green fill with Dark green text" indicating the tweet is negative.

- *System 2*: The amendment and additions of rules took place in the new version of *System 1* termed to be as *System 1*. The rules thus amended covered three major areas that is the entries of lexicon placement in the dataset, whether the sentiment word from the lexicon was found at the beginning or at the end. Based on this, the rules were thus amended or added. For example, "الفقراء" (The poor people) belongs to negative lexicon which was also a part of positive tweets. This rule was amended with the end-to-end approach, checking the appearance of "الفقراء" (The poor people) word in the dataset, whether the word appeared at the beginning or at the ending. For this example, "الفقراء" (The poor people) word appeared at the ending for the negative tweet and was within the text for the positive tweet, hence the rule was amended as – if inputted tweet ends with "الفقراء" (The poor people) then the tweet belongs to negative polarity. Creation of rules was the most challenging task, which included time and effort. The first set of experiments was conducted on *System 1*. The second set of experiments was conducted on *System 2*, which was improvised version of *System 1*.

7.4 Results

The results includes the comparison of all the experiments conducted in this dissertation. So as to do the comparison the accuracy of all the experiments are used. The *System 1* and *System 2* were tested on [Nawaf et al. 2014] and [Rushdi et al. 2011] dataset. Table (14) depicts the experiments results with regards to *System 1*.

Table (14): Results of applying System 1 on datasets

Dataset / **Acc. Measure**	**(Nawaf et al., 2014) dataset**	**(Rushdi et al., 2011) OCA dataset**
Precision	87.4	50.4
Recall	93.3	97.6

Accuracy	89.6	50.1

Table (14) clearly traces the outperformance of rules created in System 1 with huge accuracy for [Nawaf et al. 2014] when compared to the results on [Rushdi et al. 2011] OCA dataset. *System 1* provided to be successful with 39.5% more accuracy for [Nawaf et al. 2014] than [Rushdi et al. 2011]. The result variation in both the datasets calls for a system which could successfully improve the accuracy. So as to achieve this *System 2* was put in place, as discussed under Test Set 2 Chapter. Table (15) depicts the results achieved for Test Set 2 experiments on *System 2*.

Table (15): Results of applying System 2 on datasets

Dataset / **Acc. Measure**	**[Nawaf et al. 2014] dataset**	**[Rushdi et al. 2011] OCA dataset**
Precision	**92.5**	77.4
Recall	**95.8**	100
Accuracy	**93.9**	85.6

System 2 proved to be outstanding for both the datasets. 93.9% accuracy for [Nawaf et al. 2014] dataset wherein for OCA dataset 85.6% accuracy was measured. Still the accuracy for [Nawaf et al. 2014] was high with 8.3% more accuracy than OCA dataset.

7.4.1 System 1 Vs System 2 Results Comparison

While comparing *System 1* experiments with *System 2* few key conclusions were drawn. Figure (4) and Figure (5) depicts the comparison of *System 1* experiments with *System 2* with their respective datasets. Clear significance in increase in accuracy in *System 2* is noticed for [Nawaf et al. 2014] and OCA dataset.

Figure (4): *System 1* vs *System 2* [Nawaf et al. 2014] dataset

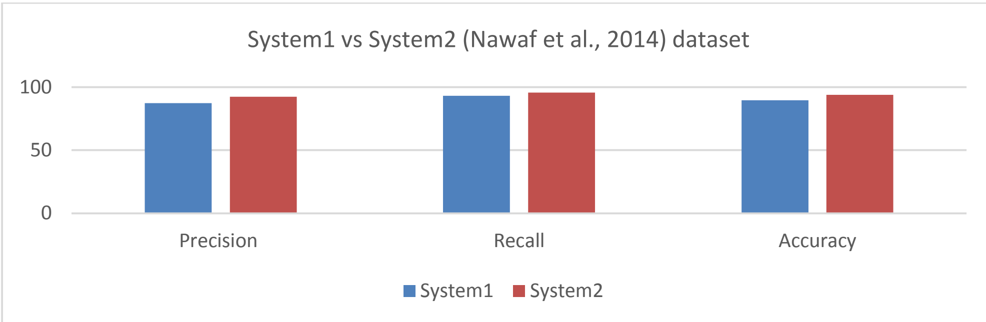
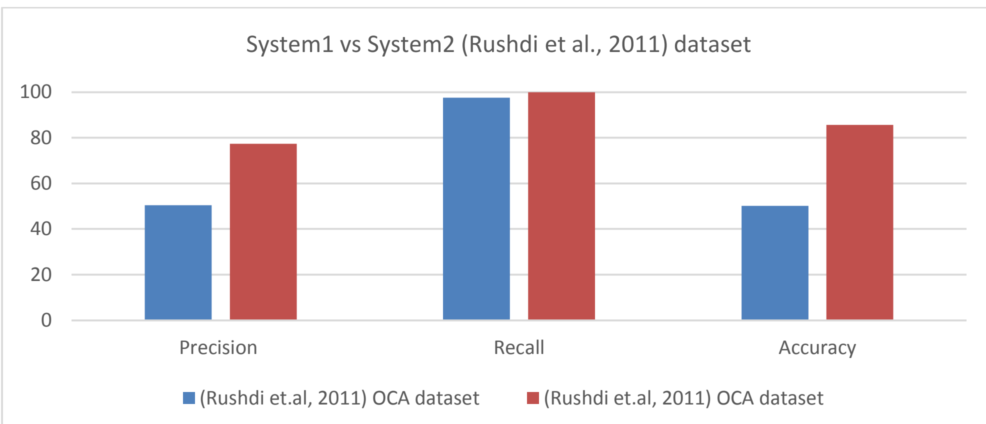


Figure (5): System 1 vs System 2 [Rushdi et al. 2011] dataset



7.5 Research Questions Answered

The rest of this chapter presents answers for the research questions.

RQ1: Is it possible to create a system for Sentiment Analysis without pre-processing the dataset?

Due to the very fact that the pre-processing of dataset was totally eliminated with no significance of any changes done to the dataset, resulted in *System 1* and *System 2* explained in Section 7.3. Hence clearly answering the first research question which was not addressed till date by any researcher that it is possible to create a system for Sentiment Analysis without pre-processing the dataset. One of the major step followed by [Nawaf et al. 2014] was the pre-processing of dataset which was completely excluded in the system we developed. Instead of pre-processing, time and efforts were invested on rule building.

RQ2: Is it possible to improvise the performance of sentiment analysis system through an innovative rule chaining approach?

The comparison of *System 1* and *System 2* clearly answers this question. *System 1* restricted to boundaries with two rules type with no chaining and System 2 an improvised version of *System 1* with chaining and appropriate positioning of the rules. Table (16) clearly depicts the outperformance of *System 2* when compared to *System 1*. Hence, the performance of sentiment analysis is improved through our innovative rule chaining approach.

Table (16): Comparison of System1 Vs System2

	System 1		System 2	
	[Nawaf et al., 2014] dataset	[Rushdi et al., 2011] OCA dataset	[Nawaf et al., 2014] dataset	[Rushdi et al., 2011] OCA dataset
Precision	87.4	50.4	92.5	77.4
Recall	93.3	97.6	95.8	100
Accuracy	89.6	50.1	93.9	85.6

RQ3: Does the use of end-to-end mechanism with rule chaining approach gives better results than the Lexicon-based approach?

This question is very well answered with a critical comparison of our end-to-end mechanism with rule chaining approach results with [Nawaf et al. 2014]'s lexicon based approach results. As the dataset used in this paper is taken from [Nawaf et al. 2014], we have compared our results this dataset. Our proposed system outperformed their results with 23.85% increase in accuracy when tested on *System 2* and 17.35% increase in accuracy when tested on *System 1*.

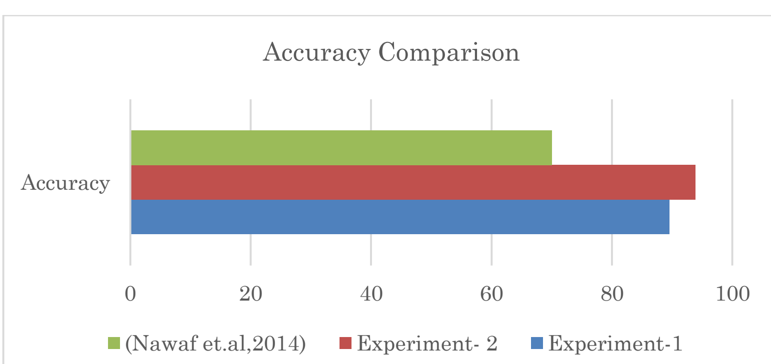
Hence, both the experiments outperformed when compared to the results reported in [Nawaf et al. 2014]. Table (17) depicts the comparison of the experiments conducted in this paper with [Nawaf et al. 2014].

Table (17): Results comparison with [Nawaf et al. 2014]

Rule-based Vs lexicon-based approaches	Accuracy
System1 Experiment	89.6
System2 Experiment	93.9
[Nawaf et al. 2014]	70.05%

Figure (6) shows graphical depiction of the results thus compared. In comparison of our achievements and our method with [Nawaf et al. 2014] shows an improvement which is very huge.

Figure (6): Graphical depiction of results



RQ4: Does the inclusion of common words based on the analysis of tweets in the negative and positive lexicon list enhances the results?

On comparing our method with [Nawaf et al. 2014] brought forward some key additions. [Nawaf et al. 2014] focused on extending the lexicons which was just addition of new words to the list even unrelated to the test set, has left them with no improvement in accuracy. [Nawaf et al. 2014] in his error analysis quoted:

> "Based on our reading of the related works in the literature, we expect such modest performance of the lexicon-based SA tool. The experimentation reveals more insights into why such tools perform badly. The first and most obvious issue is the modest lexicon size. Knowing the polarities of less than 5000 words is not enough for a complex language as the Arabic language. Hence, we had better seek for other techniques to build the lexicon along with the manual approach."

We have answered the above quote through our innovative lexicon improvising technique. In this technique, the addition of words which were found to be common in both positive and negative tweets/reviews, as well the addition of single word review, resulted in major success while drafting the rules.

RQ5: Does the positioning of rules make a huge difference to the rule-based approach?

The positioning of rules does makes a huge difference to the rule-based approach due to the very fact of placing the rules for a particular word in the chain at appropriate position so as to satisfy rules "within the text", "beginning with" and "ending with". If the positioning of these rules were not done then the results were reversed. Section 6.3, clearly depicts the importance of the positioning of rules with the aid of an example.

7.6 Chapter Summary

This chapter presented depicted *System 1* and *System 2*. *System1* included the rules which covered "within the text" and "equal to" rules. In *System* 2, the rules added to system were "beginning with" and "ending with" with positioning and prioritizing of rules in an end to end chaining mechanism. The research questions were also answered in this chapter. The first question was answered with the elimination of pre-processing step, indicating that it is possible to create a system for Sentiment Analysis without pre-processing the dataset. The second question was very well answered with the comparison of results in *System* 2 showcased the outperformance when compared to *System* 1, indicating that it is possible to improvise the performance of sentiment analysis through an innovative rule chaining approach, as *System* 2 was based on rule chaining approach. The third question addressed in this dissertation illustrated the outperformance of rule chaining approach that is *System 2* resulted in 23.85% in results when compared to [Nawaf et al. 2014]'s lexicon based approach. The inclusion of common words based on the analysis of tweets in the negative and positive lexicon list enhanced the overall result when compared to the baseline dataset, hence answering the fourth question. Last but not the least, the positioning of rules make a huge difference to the rule-based approach as appropriate positioning resulted in satisfying words which were found to be common in both negative and positive lexicons.

Chapter Eight

Error Analysis

The key error found was unhandled exceptions. This was significant in two cases.

8.1 Case 1:

In the first case, a negative review included a positive word within the text and the same positive word was mostly found within the text in positive reviews affecting positioning of rules. Table (18) illustrates few examples. Hence, the rule can either satisfy a positive review or negative review. Chain was not feasible to form in this case.

Table (18): Unhandled Exceptions – Case 1

Example	Word	Review	Comment
1	صحيح true Is a positive word which appears within the text in the negative review	اذا الخبر صحيح انا اول شخص بطالب بالهجرة لأي دولة خليجية او اوروبية او حتى اسرائيل If the news is true, I am the first person claiming to migration to any GCC country, European or even Israel	Due to the significance of the positive word – "صحيح" (true) within the text in both positive and negative reviews, restricted the rule to be created for either positive or negative. Chain was not able to form in this case.
	صحيح true Is a positive word which also appears within the text in most of the positive reviews	كلام صحيح من شان هيك الدول اللي ما فيها بطالة والمجتمعات المفتوحة بتقل فيها المشاكل النفسية Words properly stated, this is the reason that countries that do not have unemployment, and those with open societies, have less psychological problems	
		اي و الله صحيح أهم شيء الانسان كرامته	

		I swear god that this is true, the most important thing is the human dignity	
2	يسعد good Is a positive word which appears within the text in negative reviews/ sarcastic review	الله الله يا شيخ ابراهيم ..حكم والله يسعد مساااااااك يارب Oh my god! Oh my god! Sheikh Ibrahim … I swear god this is wisdom, I wish god to bless you with a good night على الأقل هناك رجل وفي امرأة أحبها وهذا يسعدني At least there is a man and a woman that I love and I am happy with this	Due to the significance of the positive word – يسعد within the text in both positive and negative reviews, restricted the rule to be only created for either positive or negative. Chain was not able to form in this case.
	يسعد happy Is a negative word which also appears within the	صبحكم الله يأهل الطيب بالخير ... و يسعد ضحاكم و الظهر و العشية God makes good morning to you good people … and makes	

	text in most of the positive reviews	you happy at morn, noonday, and eve.	

8.2 Case 2:

Likewise, in the second case, a positive review included a negative word within the text and the same negative word was mostly found within the text in the negative reviews affecting positioning of rules. Rule again restricted to either positive or negative review. Table (19) illustrates some examples.

Table (19): Unhandled Exceptions – Case 2

Word	**Review**	**Comment**
حرام Haram Is a negative word which appears within the text in the positive review	اللهم ازرقنا رزق حلالا تغننا به عن حرامك , سبحان القادر المقتدر بيده ملكوت كل شيء God grant us Muslim livelihood that enough for us to avoid Haram, glory to the capable Al-Muqtadir who has in hand the kingdom of everything	Due to the significance of the negative word – “حرام” (Haram) within the text in both positive and negative reviews, restricted the rule to be only created for either positive or negative. Chain was not able to form in this case.

حرام Haram Is a negative word which also appears within the text the negative reviews/sarcastic reviews	و الله حرام و الله موتوه لشعب الاردني من وين بدنا نجيب الكو من وين يا الله ارحموا من في الارض يرحمكم من في السماء و الله حرام Oh my god! This is Haram, Oh god! You are killing Jordanians, from where we can afford it, from where Oh god!, Be merciful to those who are on the earth, The one who is in the sky will be merciful to you all, Oh god! This is Haram.	
	و الله حرام الناس مش لاقيه توكل حتى ترفعو البنزين Oh my god! This is Haram. The people cannot afford food such that you raise the fees of benzene (fuel)	

8.3 Chapter Summary

This chapter depicted the error analysis due to the unhandled exceptions were two cases are seen. The first case in error analysis depicted the appearance of positive word in negative and positive review within the text, hindering the coverage of rule to satisfy both positive and negative sentiments. The second case in error analysis depicted the appearance of negative word in positive and negative review within the text, allowing the coverage of rule to satisfy either positive or negative sentiment.

Chapter Nine

Conclusion

This dissertation depicted how Sentiment Analysis has found its existence with the very advanced movements in online information. The key underlying parameter seen is the sharing of reviews on any context and how this impact the users to take decisions on many things right from buying a movie ticket to buying a property to many advanced operations. It calls for a need for an appropriate tool which can do justice to identify sentiments polarity and help companies, organizations and not to forget the citizens of this planet earth to benefit from this classification. According to recent published statistics Arabic is the fourth used language on Internet highlighting the need for extensive work in Arabic especially in the field of Sentiment Analysis.

This research sheds the light on the key features of Arabic Language and its peculiarities as well depicts the significance of Arabic language processing and its challenges for accomplishing the sentiment analysis task. Arabic found to have its base in Classic Arabic, Dialect Arabic and Modern standard. Arabic is being hindered to grow with very less work for dialect Arabic. The absence of rigid and strict rules accounts for the lack of success in Arabic NLP.

In this dissertation, we present the overview of subjectivity classification and sentiment classification. It depicts the importance of subjectivity classification, which forms the base for sentiment analysis. Challenges and Importance of Sentiment Analysis are also covered in detail. Shedding light on how a word in one sentiment accounts for positive polarity and how the same word could cause to make the sentence negative when used in a different context. Sentiment Analysis is found to be very useful in measuring the impact of a product or service, through the reviews that the people have shared on it. The key contribution to research in sentiment analysis through an end-to-end rule based chaining approach has proved to be a novel approach. This dissertation overcomes the lexicon building process through the appropriate placement of words as well not excluding the common words found in both the tweets for the lexicons. Nevertheless, the end-to-end rule chaining approach was the most challenging and expensive in terms of time and effort, but adds to the advancements in the state-of-the-art for Arabic Sentiment Analysis, through the orderly placed rules and through the correct use of different rules including “contains text”, “equal to”, “beginning with” and “ending with”.

Indeed, presented the newly developed sentiment analysis system, which outperformed in both sets of experiments when compared to [Nawaf et al. 2014]. *System 2* with rules extended to cover all areas was proved to increase the accuracy of OCA corpus by 39.8% and 4.3% accuracy for [Nawaf et al. 2014] when compared to *System 1*’s rules. Hence, initiating a wakeup call for all the researchers to divert their interest to rule-based approach. The clear significance in results thus obtained through the rules created makes the rule-based approach the most desirable approach.

Most of the researchers have focused only on text based output. Our output is new to the sentiment analysis era. The output was presented through the change in color of entered tweet to “Red” for positive tweets and “Green” for negative tweets.

Though this system was only focused on positive and negative tweets. Though this system was only tested on Tweeter tweets, left us with amazing results and hopes for future enhancements.

As a key task for further developing and enhancing this proposed system, we would be looking forward to improve the current proposed system to cover subjectivity classification and thus indicate neutral polarity as well. Other than this, the current system will be tested on Yahoo Maktoob and Facebook reviews. Interesting results could be drawn.

References

1) Almutawa, F., & Izwaini, S. Machine Translation in the Arab World: Saudi Arabia as a Case Study.

2) Ahmad, K. (2006). Multi-lingual sentiment analysis of financial news streams.PoS, 001.

3) Abd El Salam, A. L., HAJJAR, M., & ZREIK, K. Classification of Arabic Information Extraction methods.

4) Abbasi, A., Chen, H., & Salem, A. (2008). Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums. ACM Transactions on Information Systems (TOIS), 26(3), 12.

5) Abdel Monem, Azza, Shaalan, Khaled, Rafea, Ahmed, Baraka, Hoda, Generating Arabic Text in Multilingual Speech-to-Speech Machine Translation Framework, Machine Translation, Springer, Netherlands, 20(4): 205-258, December 2008

6) Abdul-Mageed, M., & Diab, M. T. (2012, May). AWATIF: A Multi-Genre Corpus for Modern Standard Arabic Subjectivity and Sentiment Analysis. In LREC (pp. 3907-3914).

7) Abdul-Mageed, Muhammad, Sandra Kübler, and Mona Diab. "Samar: A system for subjectivity and sentiment analysis of arabic social media." In Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis, pp. 19-28. Association for Computational Linguistics, 2012.

8) Abdul-Mageed, M., Diab, M. T., & Korayem, M. (2011, June). Subjectivity and sentiment analysis of modern standard arabic. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2 (pp. 587-591). Association for Computational Linguistics.

9) Abdul-Mageed, M., & Diab, M. (2014). Sana: A large scale multi-genre, multi-dialect lexicon for arabic subjectivity and sentiment analysis. In Proceedings of the Language Resources and Evaluation Conference (LREC).Abdulla, N. A., Ahmed, N. A., Shehab, M. A., Al-Ayyoub, M., Al-Kabi, M. N., & Al-rifai, S. (2014). Towards improving the lexicon-based approach for arabic sentiment analysis. International Journal of Information Technology and Web Engineering (IJITWE), 9(3), 55-71.

10) Abdul-Mageed, M., & Korayem, M. (2010). Automatic identification of subjectivity in morphologically rich languages: the case of Arabic. In Proceedings of the 1st workshop on computational approaches to subjectivity and sentiment analysis (WASSA), Lisbon (pp. 2-6).

11) Abdulla, N., Ahmed, N., Shehab, M., & Al-Ayyoub, M. (2013, December). Arabic sentiment analysis: Lexicon-based and corpus-based. In Applied Electrical Engineering and Computing Technologies (AEECT), 2013 IEEE Jordan Conference on (pp. 1-6). IEEE.

12) Albalooshi, N., Mohamed, N., & Al-Jaroodi, J. (2011, December). The challenges of Arabic language use on the Internet. In Internet Technology and Secured Transactions (ICITST), 2011 International Conference for (pp. 378-382). IEEE.

13) Aldayel, H. K., & Azmi, A. M. (2015). Arabic tweets sentiment analysis–a hybrid scheme. Journal of Information Science, 0165551515610513.

14) Almas, Y., & Ahmad, K. (2007, July). A note on extracting 'sentiments' in financial news in English, Arabic & Urdu. In The Second Workshop on Computational Approaches to Arabic Script-based Languages (pp. 1-12).

15) Albraheem, L., & Al-Khalifa, H. S. (2012, December). Exploring the problems of sentiment analysis in informal arabic. In Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services (pp. 415-418). ACM.

16) Al-Kabi, M. N., Alsmadi, I. M., Gigieh, A. H., Wahsheh, H. A., & Haidar, M. M. (2014). Opinion Mining and Analysis for Arabic Language. IJACSA) International Journal of Advanced Computer Science and Applications, 5(5), 181-195.

17) Al-Kabi, M., Gigieh, A., Alsmadi, I., Wahsheh, H., & Haidar, M. (2013, April). An opinion analysis tool for colloquial and standard Arabic. In The fourth International Conference on Information and Communication Systems (ICICS 2013)

18) Al-Kabi, M., Al-Qudah, N. M., Alsmadi, I., Dabour, M., & Wahsheh, H. (2013, April). Arabic/English sentiment analysis: an empirical study. In proceeding of: The fourth International Conference on Information and Communication Systems (ICICS 2013).

19) Almas, Y., & Ahmad, K. (2007, July). A note on extracting 'sentiments' in financial news in English, Arabic & Urdu. In The Second Workshop on Computational Approaches to Arabic Script-based Languages (pp. 1-12).

20) Al-Shalabi, R., & Obeidat, R. (2008, March). Improving KNN Arabic text classification with n-grams based document indexing. In Proceedings of the Sixth International Conference on Informatics and Systems, Cairo, Egypt (pp. 108-112).

21) Al-Maimani, M. R., Naamany, A. A., & Bakar, A. Z. A. (2011, February). Arabic information retrieval: techniques, tools and challenges. In GCC Conference and Exhibition (GCC), 2011 IEEE (pp. 541-544). IEEE.

22) Al-Sabbagh, R., & Girju, R. (2012, May). YADAC: Yet another Dialectal Arabic Corpus. In LREC (pp. 2882-2889).

23) Al-Subaihin, A. A., Al-Khalifa, H. S., & Al-Salman, A. S. (2011, December). A proposed sentiment analysis tool for modern arabic using human-based computing. In Proceedings of the 13th International Conference on Information Integration and Web-based Applications and Services (pp. 543-546). ACM.

24) Attia, M. A. (2008). Handling Arabic morphological and syntactic ambiguity within the LFG framework with a view to machine translation (Doctoral dissertation, University of Manchester).

25) Aue, A., & Gamon, M. (2005, September). Customizing sentiment classifiers to new domains: A case study. In Proceedings of recent advances in natural language processing (RANLP) (Vol. 1, No. 3.1, pp. 2-1).

26) Balahur, A., Steinberger, R., Kabadjov, M., Zavarella, V., Van Der Goot, E., Halkia, M., ... & Belyaeva, J. (2013). Sentiment analysis in the news. arXiv preprint arXiv:1309.6202.

27) Barney, B. (2010). Introduction to parallel computing. Lawrence Livermore National Laboratory, 6(13), 10.

28) Bassam Hammo, Hani Abu-Salem, Steven Lytinen, and Martha Evens, QARAB: A Question Answering System to Support the Arabic Language. In the proceedings of Workshop on Computational Approaches to Semitic Languages. ACL 2002, Philadelphia, PA, July. p 55-65 2002.

29) Benamara, F., Cesarano, C., Picariello, A., Recupero, D. R., & Subrahmanian, V. S. (2007, March). Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone. In ICWSM.

30) Beesley, K. R. (1996, August). Arabic finite-state morphological analysis and generation. In Proceedings of the 16th conference on Computational linguistics-Volume 1 (pp. 89-94). Association for Computational Linguistics.

31) Beesley, K. R. (2001, July). Finite-state morphological analysis and generation of Arabic at Xerox Research: Status and plans in 2001. In ACL Workshop on Arabic Language Processing: Status and Perspective (Vol. 1, pp. 1-8).

32) Dave, K., Lawrence, S., & Pennock, D. M. (2003, May). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. InProceedings of the 12th international conference on World Wide Web (pp. 519-528). ACM.

33) Daya, E., Roth, D., & Wintner, S. (2007). Learning to identify Semitic roots. InArabic Computational Morphology (pp. 143-158). Springer Netherlands.

34) Darwish, K. (2002, July). Building a shallow Arabic morphological analyzer in one day. In Proceedings of the ACL-02 workshop on Computational approaches to semitic languages (pp. 1-8). Association for Computational Linguistics.

35) Diab, M., Hacioglu, K., & Jurafsky, D. (2007). Automatic processing of modern standard Arabic text. In Arabic Computational Morphology (pp. 159-179). Springer Netherlands.

36) Duwairi et al (2014) worked on tweets/ comments equal to 2591 of which 1073 are positive and 1518 are negative. They found SVM to be the best method with a precision of 75.25.

37) Duwairi, R., & El-Orfali, M. (2014). A study of the effects of preprocessing strategies on sentiment analysis for Arabic text. Journal of Information Science, 40(4), 501-513.

38) Duwairi, R. M., & Qarqaz, I. (2014, August). Arabic Sentiment Analysis using Supervised Classification. In Future Internet of Things and Cloud (FiCloud), 2014 International Conference on (pp. 579-583). IEEE.

39) Efron, M. (2004). Cultural orientation: Classifying subjective documents by cocitation analysis. In AAAI Fall Symposium on Style and Meaning in Language, Art, and Music (pp. 41-48).

40) Elarnaoty, M., AbdelRahman, S., & Fahmy, A. (2012). A machine learning approach for opinion holder extraction in Arabic language. arXiv preprint arXiv:1206.1011.

41) Elgibali, A. (Ed.). (2005). Investigating Arabic: current parameters in analysis and learning (Vol. 42). Brill.

42) El-Beltagy, S. R., & Ali, A. (2013, March). Open issues in the sentiment analysis of Arabic social media: A case study. In Innovations in Information Technology (IIT), 2013 9th International Conference on (pp. 215-220). IEEE

43) Elhawary, M., & Elfeky, M. (2010, December). Mining Arabic business reviews. In Data Mining Workshops (ICDMW), 2010 IEEE International Conference on(pp. 1108-1113). IEEE.Farra, N., Challita, E., Assi, R. A., & Hajj, H. (2010, December). Sentence-level and document-level sentiment mining for Arabic texts. In Data Mining Workshops (ICDMW), 2010 IEEE International Conference on (pp. 1114-1119). IEEE.

44) ElSahar, H., & El-Beltagy, S. R. (2015). Building Large Arabic Multi-domain Resources for Sentiment Analysis. In Computational Linguistics and Intelligent Text Processing (pp. 23-34). Springer International Publishing.

45) El-Halees, A. (2011). Arabic opinion mining using combined classification approach.

46) El-Halees, Alaa. "Arabic opinion mining using combined classification approach." (2011).

47) Farghaly, Ali, Shaalan, Khaled. Arabic Natural Language Processing: Challenges and Solutions, ACM Transactions on Asian Language Information Processing (TALIP), the Association for Computing Machinery (ACM), 8(4):1-22, December 2009.

48) Farghaly, M. A. (1987). Study of severe slugging in real offshore pipeline riser-pipe system. Society of Petroleum Engineers.

49) Farra, N., McKeown, K., & Habash, N. (2015, July). Annotating Targets of Opinions in Arabic using Crowdsourcing. In ANLP Workshop 2015 (p. 89).

50) Farra, Noura, Elie Challita, Rawad Abou Assi, and Hazem Hajj. "Sentence-level and document-level sentiment mining for Arabic texts." In Data Mining Workshops (ICDMW), 2010 IEEE International Conference on, pp. 1114-1119. IEEE, 2010.

51) Farra, N., Challita, E., Assi, R. A., & Hajj, H. (2010, December). Sentence-level and document-level sentiment mining for Arabic texts. In Data Mining Workshops (ICDMW), 2010 IEEE International Conference on (pp. 111 Feldman, R. (2013). Techniques and applications for sentiment analysis.Communications of the ACM, 56(4), 82-89.

52) Feldman, R. (2013). Techniques and applications for sentiment analysis.Communications of the ACM, 56(4), 82-89.

53) Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, 1, 12.

54) Glance N., Hurst M., Nigam K., Siegler M., Stockton R., & Tomokiyo T.,"Deriving marketing intelligence from online discussion. In Proceedings of the 11th ACM SIGKDD international conference on knowledge discovery in data mining, pp. 419-428, 2005.

55) Green, S., & Manning, C. D. (2010, August). Better Arabic parsing: Baselines, evaluations, and analysis. In Proceedings of the 23rd International Conference on Computational Linguistics (pp. 394-402). Association for Computational Linguistics.

56) Goldberg, A. B., Zhu, X., & Wright, S. J. (2007). Dissimilarity in graph-based semi-supervised classification. In International Conference on Artificial Intelligence and Statistics (pp. 155-162).

57) Ghosh, S., Roy, S., & Bandyopadhyay, S. (2012). A tutorial review on Text Mining Algorithms. International Journal of Advanced Research in Computer and Communication Engineering, 1(4), 7.

58) Habash, N., Rambow, O., & Roth, R. (2009, April). MADA+ TOKAN: A toolkit for Arabic tokenization, diacritization, morphological disambiguation, POS tagging, stemming and lemmatization. In Proceedings of the 2nd International Conference on Arabic Language Resources and Tools (MEDAR), Cairo, Egypt (pp. 102-109).

59) Harrag, F., El-Qawasmeh, E., & Pichappan, P. (2009, July). Improving Arabic text categorization using decision trees. In Networked Digital Technologies, 2009. NDT'09. First International Conference on (pp. 110-115). IEEE.

60) Hopkins, D., & King, G. (2007). Extracting systematic social science meaning from text. Manuscript available at http://gking. harvard. edu/files/words. pdf, 20(07).

61) Habash, N. Y. (2010). Introduction to Arabic natural language processing.Synthesis Lectures on Human Language Technologies, 3(1), 1-187.

62) Kim, S. M., & Hovy, E. (2006, July). Extracting opinions, opinion holders, and topics expressed in online news media text. In Proceedings of the Workshop on Sentiment and Subjectivity in Text (pp. 1-8). Association for Computational Linguistics.

63) Korayem, M., Crandall, D., & Abdul-Mageed, M. (2012). Subjectivity and sentiment analysis of arabic: A survey. In Advanced Machine Learning Technologies and Applications (pp. 128-139). Springer Berlin Heidelberg.

64) Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. American Political Science Review,97(02), 311-331.

65) Larkey, L. S., Ballesteros, L., & Connell, M. E. (2007). Light stemming for Arabic information retrieval. In Arabic computational morphology (pp. 221-243). Springer Netherlands.

66) Lewis, M. P., Simons, G. F., & Fennig, C. D. (2015). Ethnologue: Languages of Ecuador.

67) Lita, L. V., Schlaikjer, A. H., Hong, W., & Nyberg, E. (2005, July). Qualitative dimensions in question answering: Extending the definitional QA task. In PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE (Vol. 20, No. 4, p. 1616). Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

68) Maamouri, M., Bies, A., Buckwalter, T., & Mekki, W. (2004, September). The penn arabic treebank: Building a large-scale annotated arabic corpus. In NEMLAR conference on Arabic language resources and tools (pp. 102-109).

69) Malouf R, and Mullen.T. "Taking sides: User classification for informal online political discourse." Internet Research 18.2: pp. 177-190, 2008.

70) McCarthy, T. (1981). The critical theory of Jurgen Habermas.

71) Mullen, T., & Malouf, R. (2006). A Preliminary Investigation into Sentiment Analysis of Informal Political Discourse. In AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs (pp. 159-162).

72) Mesfar, S. (2010). Towards a cascade of morpho-syntactic tools for arabic natural language processing. In Computational Linguistics and Intelligent Text Processing (pp. 150-162). Springer Berlin Heidelberg.

73) Mostafa Syiam, Zaki Fayed , and Mena Habib. An Intelligent System For Arabic Text Categorization. The International Journal of Intelligent Computing and Information Volume 6 Number 1 January 2006.

74) Mourad, A., & Darwish, K. (2013, June). Subjectivity and sentiment analysis of modern standard Arabic and Arabic microblogs. In Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (pp. 55-64).

75) Nabil, M., Aly, M., & Atiya, A. F. ASTD: Arabic Sentiment Tweets Dataset.

76) Owens, J. (Ed.). (2013). The Oxford handbook of Arabic linguistics. Oxford University Press.

77) Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis.Foundations and trends in information retrieval, 2(1-2), 1-135.

78) Pang, B., Lee, L., & Vaithyanathan, S. (2002, July). Thumbs up?: sentiment classification using machine learning techniques. In Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10(pp. 79-86). Association for Computational Linguistics.

79) Pang, B., & Lee, L. (2004, July). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In Proceedings of the 42nd annual meeting on Association for Computational Linguistics (p. 271). Association for Computational Linguistics.

80) Quirk et al. A comprehensive grammar of the English language. Vol. 397. London: Longman, 1985.

81) Rafea, A. A., & Shaalan, K. F. (1993). Lexical analysis of inflected Arabic words using exhaustive search of an augmented transition network.Software: Practice and Experience, 23(6), 567-588.

82) Refaee, E., & Rieser, V. (2014). An Arabic twitter corpus for subjectivity and sentiment analysis. In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, may. European Language Resources Association (ELRA).

83) Rish, I. (2001, August). An empirical study of the naive Bayes classifier. InIJCAI 2001 workshop on empirical methods in artificial intelligence (Vol. 3, No. 22, pp. 41-46). IBM New York.

84) Rorty, R. (1997). Justice as a Larger Loyalty1. Ethical Perspectives, 4(2), 139.

85) Rushdi-Saleh, M., Martín-Valdivia, M. T., Ureña-López, L. A., & Perea-Ortega, J. M. (2011). OCA: Opinion corpus for Arabic. Journal of the American Society for Information Science and Technology, 62(10), 2045-2054.

86) Ryding, K. C. (2005). A reference grammar of modern standard Arabic. Cambridge university press.

87) Shaalan K., Rafea, A., Abdel Monem, A., Baraka, H., Machine Translation of English Noun Phrases into Arabic, The International Journal of Computer Processing of Oriental Languages (IJCPOL), World Scientific Publishing Company, 17(2):121-134, 2004.

88) Shaalan, Khaled., Rule-based Approach in Arabic Natural Language Processing, Special Issue on Advances in Arabic Language Processing, the International Journal on Information and Communication Technologies (IJICT), Serial Publications, New Delhi, India, 3(3):11-19, June 2010.

89) Shaalan, Khaled, A Survey of Arabic Named Entity Recognition and Classification, Computational Linguistics, 40 (2): 469-510, MIT Press, USA, 2014.

90) Shaalan, Khaled, Magdy, Marwa, Fahmy, Aly. Analysis and Feedback of Erroneous Arabic Verbs, Journal of Natural Language Engineering (JNLE), 21(2):271-323, Cambridge University Press, UK, Sept. March 2015.

91) Shoukry, A., & Rafea, A. (2012, May). Sentence-level Arabic sentiment analysis. In Collaboration Technologies and Systems (CTS), 2012 International Conference on (pp. 546-550). IEEE.

92) Shoukry A and Rafea A. Preprocessing Egyptian dialect tweets for sentiment mining. In: The 4th workshop on computational approaches to Arabic script-based languages, 2012, pp. 47–56.

93) Somasundaran, S., & Wiebe, J. (2009, August). Recognizing stances in online debates. In Proceedings of the Joint Conference of the 47th Annual Meeting of the

ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1 (pp. 226-234). Association for Computational Linguistics.

94) Somasundaran, S., Wilson, T., Wiebe, J., & Stoyanov, V. (2007, March). QA with Attitude: Exploiting Opinion Type Analysis for Improving Question Answering in On-line Discussions and the News. In ICWSM.

95) Soudi, A., Neumann, G., & Van den Bosch, A. (2007). Arabic computational morphology: knowledge-based and empirical methods (pp. 3-14). Springer Netherlands.

96) Stoyanov, V., Cardie, C., & Wiebe, J. (2005, October). Multi-perspective question answering using the OpQA corpus. In Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (pp. 923-930). Association for Computational Linguistics.

97) Tatemura, J. (2000, January). Virtual reviewers for collaborative exploration of movie reviews. In Proceedings of the 5th international conference on Intelligent user interfaces (pp. 272-275). ACM.

98) Terveen, L., Hill, W., Amento, B., McDonald, D., & Creter, J. (1997). PHOAKS: A system for sharing recommendations. Communications of the ACM, 40(3), 59-62.

99) Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), 267-307.Qiu, G., Liu, B., Bu, J., & Chen, C. (2011). Opinion word expansion and target extraction through double propagation. Computational linguistics, 37(1), 9-27.

100) Tokuhisa, R., Inui, K., & Matsumoto, Y. (2008, August). Emotion classification using massive examples extracted from the web. InProceedings of the 22nd International Conference on Computational Linguistics-Volume 1 (pp. 881-888). Association for Computational Linguistics.

101) Turney, P. D. (2002, July). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In Proceedings of then association for computational linguistics (pp. 417-424). Association for Computational Linguistics.

102) Xiaowen Ding , Bing Liu and Philip S. Yu. A Holistic Lexicon-Based Approach to Opinion Mining. In Proceedings of WSDM 2008.

103) Wang, G., & Araki, K. (2007, April). Modifying SO-PMI for Japanese weblog opinion mining by using a balancing factor and detecting neutral expressions. InHuman Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers (pp. 189-192). Association for Computational Linguistics.

104) Wan, X. (2009, August). Co-training for cross-lingual sentiment classification. In Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1 (pp. 235-243). Association for Computational Linguistics.

105) Wiebe, J., & Riloff, E. (2005). Creating subjective and objective sentence classifiers from unannotated texts. In Computational Linguistics and Intelligent Text Processing (pp. 486-497). Springer Berlin Heidelberg.

106) Wright, A. (2009). Our sentiments, exactly. Commun. ACM, 52(4), 14-15

107) Yi, J., Nasukawa, T., Bunescu, R., & Niblack, W. (2003, November). Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In Data Mining, 2003. ICDM 2003. Third IEEE International Conference on (pp. 427-434). IEEE.

108) Zaidan, O. F., & Callison-Burch, C. (2014). Arabic dialect identification. Computational Linguistics, 40(1), 171-202.