Cryptocurrency Exchange Market Prediction and Analysis Using Data Mining and Artificial Intelligence

التنبؤ والتحليل لسوق العملات الرقمية بإستخدام طرق التنقيب عن البيانات والذكاء الإصطناعي

by

NASSER AL RAYHI

Dissertation submitted in fulfilment

of the requirements for the degree of

MSc INFORMATICS

at

The British University in Dubai

May 2020

DECLARATION

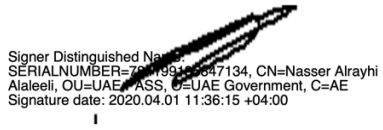
I warrant that the content of this research is the direct result of my own work and that any use made in it of published or unpublished copyright material falls within the limits permitted by international copyright conventions.

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

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

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

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



Signature of the student

COPYRIGHT AND INFORMATION TO USERS

Abstract

This thesis discusses the ability of using data mining and artificial intelligence techniques in order to predict crypto currency exchange market prices. The prices of crypto currency are virtual unpredictable and pose numerous liabilities and risks for its investors such as huge sums of money lost in unforeseen variations of crypto currency prices. Furthermore, there are some challenges when it comes to crypto currency prediction such as fake news. Cryptocurrency exchange market have a lot of fake news articles which can affect the market prices by fooling people either to buy or to sell. This research aims to tackle the unpredictability of fluctuating crypto currency prices by researching peer-reviewed articles and determine the applicability of AI algorithm and data mining and its implications for the cryptocurrency market. This thesis will also apply the algorithms in order to determine how applicable it is to predict crypto currency exchange market prices. Furthermore, this thesis discusses blockchain technology in general since this will be the building block for the student research.

This thesis is trying to answer the following questions:

1. To which extend can we predict Cryptocurrency exchange market prices using well known prediction algorithms?
2. How accurate is the prediction of the prices?

Crypto currency data has been downloaded from different sources and cleansing and normalization methods were used on them. The thesis has examined different algorithms that can be used for prediction such as Moving Average, Auto ARIMA, kNN, Linear Regression and Long Short Term Memory (LSTM). One of the best algorithms in terms of the result is the Long Short Term Memory (LSTM) since it is based on recurrent neural networks which uses loop as a method to learn from heuristics data. For the future improvements can be done to this work by integrating other related data such as social media news using sentimental analysis. Furthermore, more advanced algorithms can be used such as Support Vector machine (SVM) and XGBoost.

Abstract in Arabic

تناقش هذه الرسالة القدرة على التنبؤ والتحليل لسوق العملات الرقمية بإستخدام طرق التنقيب عن البيانات والذكاء الإصطناعي. أسعار العملات الرقمية لا يمكن التنبؤ بها افتراضيًا وتشكل العديد من المسؤوليات والمخاطر على مستثمريها مثل خسارة مبالغ ضخمة من الأموال في التغييرات غير المتوقعة في أسعار العملات الرقمية. علاوة على ذلك ، هناك بعض التحديات عندما يتعلق الأمر بالتنبؤ بالعملات المشفرة مثل الأخبار المزيفة التي تساهم في ارتباك السوق. يحتوي سوق تبادل العملات الرقمية على الكثير من المقالات الإخبارية المزيفة التي يمكن أن تؤثر على أسعار السوق عن طريق خداع الناس إما للشراء أو البيع. تهدف هذه الرسالة إلى معالجة عدم القدرة على التنبؤ بتقلب أسعار العملات المشفرة من خلال البحث في الأبحاث التي تم نشرها من قبل الزملاء وتحديد إمكانية تطبيق خوارزمية الذكاء الاصطناعي واستخراج البيانات وآثارها على سوق العملات الرقمية. ستطبق هذه الأطروحة أيضًا الخوارزميات لتحديد مدى قابلية تطبيقها للتنبؤ بأسعار سوق تبادل العملات الرقمية. علاوة على ذلك.

تحاول هذه الرسالة الإجابة على الأسئلة التالية:

١- إلى أي مدى يمكننا توقع أسعار سوق تبادل العملات المشفرة باستخدام خوارزميات تنبؤ معروفة؟

٢- ما مدى دقة التنبؤ بالأسعار؟

تم استيراد بيانات العملات الرقمية من مصادر مختلفة واستخدمت طرق التنقية والتوحيد عليها. قامت الأطروحة بتجربة خوارزميات مختلفة يمكن استخدامها للتنبؤ مثل Moving Average, Auto ARIMA, kNN, Linear Regression and Long Short Term Memory (LSTM). واحدة من أفضل الخوارزميات من حيث النتيجة Long Short Term Memory لأنها تعتمد على الشبكات العصبية المتكررة التي تستخدم الحلقة كطريقة للتعلم من بيانات الاستدلال. بالنسبة للتحسينات المستقبلية ، يمكن إجراء تحسينات على هذه الإطروحة مستقبلا من خلال دمج البيانات الأخرى ذات الصلة مثل أخبار وسائل التواصل الاجتماعي. علاوة على ذلك ، يمكن استخدام خوارزميات أكثر تقدمًا مثل Support Vector Machine (SVM)وXGBoost.

Dedication

I dedicate this work to my parents who always supported me in my education. I also dedicate this work to my current employer Telecommunication Regulatory Authority who encouraged and supported my decision to continue my education.

Table of Contents

1. Chapter One: Introduction

The technology of artificial intelligence (AI) has the potential of transforming many aspects of humanity's life, including the existing financial operations. The emergence of digital currencies, for instance, bitcoin in 2008 presented a question of analyzing this new financial market. This paper aims to review the application of AI in the context of blockchain finance by examining scholar articles to determine whether these digital currencies, best known for its volatile prices, can actually be precisely predicted.

1.1. Overview

The concept of Cryptocurrency refers to the conducting of financial transactions by using cryptographical functions. As opposed to other mediums of exchange, a cryptocurrency is based on the internet and the blockchain technology (Berentsen and Schar, 2). Cryptocurrency was introduced in 2009, the first digital currency to be used and released is Bitcoin which caught the interest of the mainstream media only in 2012. According to N Gandal, H Halaburda (2014) because of its unique characteristic such as decentralization and anonymity they picked up the pace of development rapidly. One of the important characteristics of blockchain and cryptocurrency is that it is not governed by any central authority. The blockchain is decentralized which makes cryptocurrencies hypothetically immune to the old ways of prices control and interference. The more the prices of Bitcoin get higher, the more we saw that other cryptocurrencies follows its path and increase in value, and vice versa which means that the prices of all the cryptocurrencies fluctuate in lockstep N Gandal, H Halaburda (2014). The cryptocurrency market value exceeded the limit of $100 billion market value in June 2017 ElBahrawy, Abeer, et al. (2017).

The utilization of the blockchain technology allows cryptocurrencies to be immutable, transparent and decentralized. A primary characteristic of the traditional payment mechanisms is the existence of a central authority which is the central banks of many countries which regulate

the flow of money. In the case of Cryptocurrency, national governments cannot restrict their functioning due to the decentralized operation of the blockchain. Internet-based transactions allowed by cryptocurrencies are characterized by reduced charges compared to the conventional banking fees. The lack of interference from centralized agencies creates the risk of double-spending within any given economy. However, the peer-to-peer nature of cryptocurrencies curbs the possibility of double-spending. While the concept of Cryptocurrency is relatively new, most people are embracing the idea which has witnessed the increase in the exchange value of cryptocurrencies such as the bitcoin.

Blockchain technology was initially developed and launched as the foundational mechanism and structure of bitcoin (Al-Saqaf and Nicolas, 339). Blockchain technology is established on the idea of an openly distributed ledger, openness, tamper-proof, and informational transparency, which allow for the completion of cryptocurrency transactions (Xu et al. 2). The technology was therefore developed primarily to facilitate internet-based mediums of exchange. The recent developments regarding the blockchain technology include the adoption of the concept in copyright protection, market monitoring, inventory chain management, healthcare, and finance. The foundation of blockchain technology is marked by consensus algorithms, smart contracts, cryptography, peer-to-peer networking, and distributed storage (Xu et al. 3). The idea of blockchain is associated with the financial crisis witnessed in the year 2008, and the reduced trust of the public on the normal banking systems. Contrary to the case where transacting rely on the bank as a reference point, whether a contract of money exchange was completed, Blockchain technology utilizes a digitally distributed ledger accessed by all members to record the event through cryptography (Al-Saqaf and Nicolas, 341). The functioning and design of blockchain technology prevent tampering by a single user (node), thus allowing the system to maintain data immutable at the absence of a centralized agency. The advantages and value of the blockchain technology are based on the transparency and credibility created by the system to the users (Leible et al. 1). The users of a blockchain form a community based on the needs of the involved people. The system is developed, allowing storage of information while benefiting from decentralization and immutability.

The cryptocurrency market is significantly big, primarily due to the continued interest of the system from the public, entrepreneurs, and investors (Giudici et al. 2). Since the introduction of Bitcoin in the year 2009, the current cryptocurrency market has increased to 5397 cryptocurrencies, according to CoinMarketCap.com. Based on market capitalization, several of the cryptocurrencies in the market have values running into millions of dollars.

A significant fact about the cryptocurrency market is the fact that the given digital currencies are not backed by physical assets (Giudici et al. 3). The huge growth of the Cryptocurrency and the lack of underlying security create the idea that digital currencies and the blockchain technology represent an economic 'bubble' calling for increased regulatory requirements from the market regulators such as governments and central banks (Giudici et al. 15). For Bitcoin, the Cryptocurrency recorded a 40% increase in price since the beginning of the year 2020 as of February 2020 (Bambrough, 1). The growth of the cryptocurrency market is proofed by the price of individual digital currencies. Taking the case of Ethereum, the Cryptocurrency recorded an increase in price to $4800 in October 2017 from $616 in October 2016 (Kyriazis et al. 1). From the initial response of the investors and entrepreneurs to cryptocurrencies to the current state, it is evident that the market size of the digital currencies is bound to increase into the future. The continued acceptance of cryptocurrencies is bound to spread beyond the developed economies and impact the operations of the global financial systems in the future.

The prices of cryptocurrencies are volatile and responsive to the market trends that affect investor decisions. The trading of the cryptocurrencies in the open market allows the intervention of market forces, including the future demand, to determine current prices. In the research by Kyriazis et al. (6), the price of cryptocurrencies is volatile to the whole outlook of the market, thus allowing the direction of the major digital currencies such as Ethereum and bitcoin to guide overall prices. In this case, the price of major currencies is directly related to those of the unpopular cryptocurrencies. Fluctuation in the price of a cryptocurrency is marked by exponentially descending or ascending line (Kyriazis et al. 6). The viewing of the cryptocurrencies as crypto assets allows the given instruments to be susceptible to the forces of demand and supply. The market system of cryptocurrencies is witnessed in the performances

between January and September 2018. Within the given period, the investors and entrepreneurs within the digital currency sector experienced a falling performance (Kyriazis et al. 7).

Cryptocurrencies have indicated the historical change in prices over time arising from specific events. On April 9, 2020, Toronto Stock Exchange made the initial list of public Bitcoin Fund, and the Bitcoin value declined to $7170.65 from $7314.95 within ten days (Beigei, 1). the event implied that the retail investors of Bitcoin would be subjected to the controlled stock exchange prices. The event seemed unfavorable for the cryptocurrency community, and the market response evidenced the volatility of prices to events. On April 1, 2020, CoinMarketCap was acquired by Binance, and the price of bitcoin rose to $6814.05 from $6671.95 over ten days (Beigei, 1). Binance is a big cryptocurrency exchange, and the retailers received the event as good news. Here, it is imminent that the market uses the state of facts to predict the future outlook of the cryptocurrency segment in making the sell or buy decisions. On March 23, 2020, the Federal QE Bazooka was announced, which implied an easing of the bitcoin federal reserve program. The event led to an increase in the price of the bitcoin to $6833.05 from $6492.94 (Beigei, 1). The examples provided allow for an in-depth understanding of the cryptocurrency market and the drivers of individual currency prices. While speculations on the safety of cryptocurrencies are bound to affect the long-term predictions and acceptance, the short-term rates are volatile to current trends. The operations of Cryptocurrency resonate with the situation in the case of the traditional currencies and other asset-backed instruments in the market.

Unlike traditional financial systems, blockchains and cryptocurrencies are not subject to centralized regulations. Here, the ability of the system to be used for criminal activities is imminent. According to Houben and Alexander (1), the cryptocurrency platform creates a ground for illegal activities, including tax evasion, terrorist funding, and money laundering. The non-regulation of the cryptocurrency trading and exchange prevents oversight and the government tracking of actions to prevent crime. While the users of the blockchain economy benefit from the immutability of data, the protection of the public from crime facilitated by Cryptocurrency is not guaranteed. The relatively new concept of information technology-based exchange mediums creates significant challenges for the legislators in most countries. As

opposed to the case of country-based currencies, the globalized nature of cryptocurrencies prevents in-depth scrutiny and assessment of risks and the vulnerable members of the society.

The advancement in technology is providing individuals and societies new avenues of creating efficiency and reducing business costs (Ghosh, 236). Cryptocurrency is enhancing social standards by eradicating the increased transactional expenses associated with traditional banking. Also, the control of the Cryptocurrency through cryptography guarantees individuals of transparency and decentralization. High reliance on centralized systems in information control, online surveillance, and censorship are explained as less productive as far as social wellbeing is concerned (Al-Saqaf and Nicolas, 338). Here, Cryptocurrency creates a platform for the realization of higher human rights as now as information sharing and access is concerned.

The widespread use of cryptocurrencies and the fact that they remained popular for over ten years facilitates the need for developing prediction models that will allow one to use these currencies as investments. AI technology has been developing for decades and has been a topic of discussion for many. It has implications for presenting an efficient analysis and presenting one with a practical evaluation of this financial market. However, the complexity of blockchain and the nature of this technology makes it impossible to analyze and predict using the standard financial analysis models. Therefore, behavioral factors and other elements should be accounted for in the successful algorithm. In regard to cryptocurrencies, a variety of researches is states that cryptocurrencies, such as bitcoin or altcoins, have a strong meaning for the developing economies because they mitigate the issue of inflation Darlington (2014). According to Salviotti, de Rossi, and Abbatemarco (2018), the current market of blockchain and cryptocurrencies is affected to bias and hype, which does not allow one to accurately apply AI or data mining for analysis.

1.2. Research Objectives

This research aims to examine peer-reviewed articles and determine the applicability of AI algorithm and data mining and its implications for the cryptocurrency market. The amount of data generated by exchange services and activities allow researchers to analyze behavior and trends and come up with valid prediction models.

This research aims to tackle the unpredictability of fluctuating crypto currency prices by researching peer-reviewed articles and determine the applicability of AI algorithm and data mining and its implications for the cryptocurrency market. This thesis will also apply the algorithms in order to determine how applicable it is to predict crypto currency exchange market prices. Furthermore, this thesis discusses blockchain technology in general since this will be the building block for the student research.
Additionally, the ability of AI to assess and evaluate this financial market in order to boost is one of the objectives. Understanding different models of AI and their results in regards to cryptocurrencies market predictions will be explored. Next, a test with the use of historical data will be conducted and results as well as their implications will be presented. In general, this paper should help develop an ethical tools and methodology for decision making in the context of AI facilitated cryptocurrency market analysis.

1.3. Research Questions

As mentioned previously, there is no sufficient research could be found on the concept of crypto currency exchange market prediction as comparison with the existing research done into fiscal and commodity currency exchange market predictions. Furthermore, as channels for cryptocurrency exchange became wide spread between end users, such as acting as a method of payment or acting as a method of investing, any research on this topic can generate data and findings that could be beneficial for all stakeholders.

In order to develop a structured research thesis, a set of research questions was compiled below which summarized what this thesis is trying to achieve by answering.

1. Q1: To which extend can we predict Cryptocurrency exchange market prices using well known prediction algorithms?
2. Q2: How accurate is the prediction of the prices?

1.4. Research Methodology

In order to locate answers for the questions that this paper aims to answer peer-reviewed articles from scholarly journals that focus on AI, finance, cryptocurrency, and blockchain were examined. The reports present a large number of data including background information and specific studied that use AI to make financial predictions. Thus, this paper will research the concept of blockchain, bitcoin, AI and investigate the connection between these elements and how that can be used for financial market purposes. The limitations of blockchain will be investigated as well as other aspects of this technology. A variety of models and strategies of AI application will be reviewed in order to determine a successful strategy for AI implementation. This information should provide a cohesive understanding of the cryptocurrency market and the implications of AI enabled forecasting.

1.5. Dissertation Structure

This dissertation is divided into six chapters. The first chapter gives an introduction about the overall done work in this thesis. The second chapter gives a detailed literature review about crypto currency in general and about prediction methods and previous existing work by other researchers. The third chapter would elaborate in detail about the research methodology that the student have followed and details about the used predicting algorithms. The fourth chapter would be used to conduct the experiments in order to test the algorithms and obtain the prediction results. The fifth chapter would give the conclusion and the suggested future work that the student or his colleges can do in order to improve the thesis. The final chapter would list down all references used by the student.

2. Chapter Two: Literature Review

The following paragraphs will focus on exploring the research works that examine on various aspects of blockchain and AI. It should be noted that the development of a framework that can be used to analyze the financial market of cryptocurrencies is critical due to significant impact of the blockchain technology on the economic and social life of people globally. Machine learning algorithms that will be discussed in this paper are an essential element of AI technology, and some examples include neural networks and deep learning. Due to the current popularity of cryptocurrencies and their impact on the economy on a global level, a need for forecasting the trends and specifics of value changes has arisen.

2.1. Background and Related Works

Firstly, it is crucial to define AI as well as blockchain technology and describe the application of the former in the financial markets. According to Zheng et al. (2018, p. 1), AI is “the core technology of new technological revolution and industrial transformation, is transcending the traditional means of simulating human intelligence by a computer.” The authors specifically focus on describing AI 2.0 and its application within the financial market because of the particular goals and suitability of this technology for the needs of the financial markets. Coeckelbergh and Reijers (2015, p. 172) state that "technologies have a temporal and narrative character: that they are made sense of by means of individual and collective narratives but also themselves co-constitute those narratives and inter-human and social relations; configuring events in a meaningful temporal whole." The third stage of Fintech development implies the integration of big data with other elements such as the Internet and blockchain for the achievement of better efficiency within the financial market. One example of an application that describes the efficiency of using AI in finance is the assessment of an individual’s credit score. Next, in order to understand blockchain one must have a sufficient knowledge of the information centralization on the Internet. Zheng et al. (2018, p. 2) state that various barriers obstructing individuals from free data sharing existed until blockchain was introduced, which is "a distributed, publicly available, and immutable ledger." Blockchain and bitcoin cryptocurrency are inseparable as they were simultaneously introduced in 2008 by Nakamoto (DeVries 2016). In his paper, DeVries (2016) argues that cryptocurrencies, for example, bitcoin, do not have a

potential of fully replacing the existing financial structures, however, they can transform the perception of market interactions. The conclusion made by the author suggests that due to the fact that cryptocurrencies imply fewer barriers or regulators, they can impact the existing approaches to currency exchange and national currencies perception.

Defining the concept of these currencies is crucial for understanding their application. DeVries (2016, p. 1) describes cryptocurrency as "an encrypted, peer-to-peer network for facilitating digital barter." The author states that the primary advantage of this exchange model is the lack of oversight from a third party, which allows internet users to exchange value more easily. According to Jani (2018, p. 1), "as of March 18, 2018, there are 1564 Cryptocurrencies available & traded in about 9422 exchanges." Therefore, the current market of cryptocurrencies is large and will continue to grow to provide investment opportunities for individuals. Miraz and Ali (2018) focus on the prospects of blockchain technology and its prospective applications beyond the cryptocurrencies described in this paper. According to the authors, "distributed storage systems, proof-of-location, healthcare, decentralized voting and so forth" are the prospective fields that can benefit from this technology in the future (Miraz & Ali 2018, p. 1). Therefore, blockchain is not only applied in peer to peer value enhance through cryptocurrencies but can be used in other fields as well.

A different perspective on the matter can help further improve the understanding of cryptocurrencies. Farell (2015, p. 130) provides the following description for the concept of cryptocurrencies – "virtual coinage system that functions much like a standard currency, enabling users to provide virtual payment for goods and services free of a central trusted authority." The author argues that although the successful implementation of blockchain occurred recently, the concept of cryptocurrencies was first explored in the 1980s.

The process of obtaining currencies is complex and requires one to solve an algorithm. Narayanan et al. (2016) describe mining puzzles that allow one to mine bitcoin and thus receive coins for the efforts. This is an essential element of the cryptocurrency market because as was previously mentioned the difficulty of mining is one of the aspects that affect the market. Thus, Narayanan et al. (2016) hypothesize that individuals will try to locate shortcuts for solving puzzles, which will allow them to achieve higher monetary rewards. This article provides a better

understanding of the forces impacting the cryptocurrency market as well as implications for developing AI and data mining algorithms for predictions regarding value fluctuations.
Other background information that can be useful for this research is the historic change of bitcoin evaluation. Gandal and Halaburda (2016) examined the price changes and factors impacting them in the early stages of cryptocurrency implementation. The findings suggest that despite bitcoin being the most well-known and the most initial cryptocurrency introduced to the market, the value of it is currently smaller than that of others. The authors applied the winner-take-all framework to analyze bitcoin against other cryptocurrencies. The analysis presented by the authors of this study suggests that further models and algorithms should not be based on the implications of the new price trends of cryptocurrencies because those show little impact on the future development of these coins.

The cryptocurrency market incorporates a variety of tokens or currencies. According to Lee, Guo, and Wang (2018), apart from the popular bitcoin, several other cryptocurrencies were developed based on the original blockchain technology, which is referred to as altcoins. Among the reasoning for these actions is the limited of 21 million bitcoins and high demand for electricity required to mine these coins. Additionally, these elements led to the increased interest towards cryptocurrencies that affected the financial market and subsequently the perceived value of these coins.

In regard to the prospects of bitcoin, blockchain, and cryptocurrencies, a variety of suggestions is offered by scholars. For instance, Darlington (2014) states that cryptocurrencies, such as bitcoin or altcoins, have a specific meaning for the developing economies because they mitigate the issue of inflation. However, limitations in regard to access and applicability still exist, as well as the possibility of uncovering problems with the mining algorithm in the future. Peters, Panayi, and Chapelle (2015) provide their insight on the topic of cryptocurrencies and prospective development trends while ElBahrawy et al. (2017) state that even though new cryptocurrencies emerge and leave the market continuously, the overall development trend has remained stable since 2013. This is consistent with the neutral evolution model that accounts for elements that facilitate the development of this financial market.

It is critical to understand the specifics of operations that distinguish bitcoin from other cryptocurrencies. Delmolino et al. (2015) provide an understanding of smart contracts that were applied by altcoins emerging after bitcoin. Those use specific user-generated rules for transactions resulting in the ability to change the approaches to mining and currency exchange introduced by bitcoin. Thus, this article provides valuable insight into the issue of smart contracts, which can be used to enhance one's knowledge of blockchain interactions.

2.2. Challenges and Limitations

Despite the overall benefits of cryptocurrencies, there are important issues that should be considered with their application and use. In his article, DeVries (2016) argues that contemporary in-state and international institutions are not tailored to the requirements of blockchain technology. Therefore, no clear regulations and legislation exist, which is both a challenge due to security and safety and an advantage due to the mitigation of unnecessary oversight. Azouvi, Maller, and Meiklejohn (2019, p. 127) state that "in a decentralized system, no one entity can act to censor transactions or prevent individuals from joining the network." Another issue that can significantly affect this financial market and AI-based predictions is the need for user acceptance. DeVries (2016) argues that this is the only force that determines the success and value of any cryptocurrency; thus, with a change in people's perception, the price of any cryptocurrency can change significantly.

The actual price of specific cryptocurrencies is unknown due to the current popularity of this technology. Salviotti, de Rossi, and Abbatemarco (2018) argue that due to the current demand and discussion surrounding blockchain and cryptocurrencies the current market is subjected to bias and hype, which does not allow one to adequately apply AI or data mining for analysis. Many companies and venture capitalists invested in enterprises in this market and the actual value of each cryptocurrency, including the most popular one, bitcoin will be seen over time. Lindman, Rossi, and Tuunainen (2017) point out the limitation of blockchain – is the network effect that impacts the cryptocurrencies and payment system limitations. More specifically, this technology is especially crucial for the financial markets because it eliminates a variety of risks associated with the industry. Additionally, blockchain can be applied in many governmental

operations including issuing of certificates, which will allow additional accuracy and transparency to this process.

Among the specific limitations that exist one should note that many countries, for instance, the US, introduced legislation aimed at regulating the cryptocurrency market. This can lead to a significant impact on the overall perception of the demand for coins. Hughes (2017) states that despite the current efforts the decentralized nature of the blockchain technology makes it impossible to enforce specific regulations. Delgado-Segura et al. (2018) state that P2P networks are the primary feature of cryptocurrencies that distinguishes them from other digital money. However, the issue is that no specific standard exists, which results in significant differenced within the functionality of these networks.

2.3. How AI, Cryptocurrency and Blockchain are Related

The following paragraph will explore the connection between the three technologies. Hassani, Huang, and Silva (2018) state that currently the cryptocurrency market is valued at trillions of US dollars, which showcases its significance in the context of global finance. The authors argue that the primary connection between AI and cryptocurrencies is the fact that the latter requires the application of Big Data analysis due to its complexity and incorporation of a large number of users. This corresponds with the five features of big data – volume, variety, velocity, veracity, and value.

The technology of blockchain is revolutionizing many elements of contemporary life, including the financial markets. Salviotti, de Rossi, and Abbatemarco (2018) state that within this model, users have a critical role because they need to continuously participate in the data exchange process to ensure the functioning of blockchain. This allows individuals to be a part of the peer to peer exchanges or transactions. In this regard, it should be noted that the connection between blockchain and bitcoin is the fact that the former was the first known successful application of blockchain technology. Therefore, cryptocurrencies emerged due to the development of the technology in question and their further development as well as the functioning of this financial market will depend on this concept.

In regard to the application of AI and data mining, prior research suggests that information gathered via the Internet can be used to make valid predictions regarding the changes in the financial markets. Colianni, Rosales, and Signorotti (2015) present an example of Twitter and the gathering of information from this social media platform that helped researchers make predictions regarding the movement of securities. Subsequently, the authors argue that social media can be used to gather information about cryptocurrencies that will enable one to develop adequate trading strategies. Thus, one can say that based on the findings of the study by Colianni, Rosales, and Signorotti (2015, p. 1), a conclusion regarding the successful use of machine learning within the financial market of cryptocurrencies predictions can be made, more specifically, the following algorithms were tested by the authors – "logistic regression, Naive Bayes, and support vector machines." The accuracy of such predictions was estimated at over ninety percent, which provides implications for further development of similar strategies. Cryptocurrencies are inevitably connected to blockchain as they are a result of a block exchange. Thus, tokens that are received as part of this process have the potential of transforming the economic system globally because they allow one to exchange value with others independently. Laskowski and Kim (2016) explore the technology of text mining in the context of cryptocurrencies. The authors created a framework that incorporates a number of information streams from social media and messaging applications which allowed them to develop a cohesive model representing the current financial market of cryptocurrencies. Therefore, unarguably, the nature of cryptocurrencies implies a need for a different analysis of information to make accurate predictions of market changes.

2.4. The Emergence of Cryptocurrency and Blockchain

As was previously mentioned, most researches reviewed in this paper agree that the first cryptocurrency introduced was bitcoin. It was created in 2008 and according to DeVries (2016) to ensure rarity, there is a specified number of this currency that can be generated by users. An unusual element of blockchain finance is explored by the author who points out that, unlike

traditional currencies, cryptocurrencies exist due to perceived value that individuals are accepting them as payments place on this technology. For instance, a vendor receiving bitcoins has to believe that this currency has value since no institution can provide support for it. Due to this reason, one can argue that the blockchain finance market is more complicated when compared to traditional ones and requires a more advanced technology such as AI for proper analysis and predictions.

2.5. Types of Blockchain

In order to understand the different kinds of blockchain one must have sufficient knowledge of the mechanisms that underline the blockchain technology. Price (2017, p. 2) states that "a blockchain database is a distributed ledger comprising transactions and blocks." This structure is critical for ensuring the safety of data because each block incorporates a hash from a previous block. Thus, it is highly unlikely that such data will be corrupted due to the integrity between different blocks. The cryptocurrencies that are created as a result of such exchanges or through the process called mining are intended to incentivize users to use the system and thus enable the exchange of information between different parties.

It should be noted that a variety of blockchain types exists that differ in accordance with the use of the algorithm. For instance, public blockchains use the proof of work model that can be seen in bitcoin. Users have to verify each transaction in order for this system to work. This model is open sourced and therefore can be used by anyone. Private blockchains, on the contrary, have limited access to data and thus can be used by a specified number of users.

2.6. Cryptocurrency Market Analysis

The formation of value within the cryptocurrency market is an essential feature that shapes the demand and price of each coin. Hayes (2015, p. 1308) argues that based on the empirical data and cross-sectional analysis it can be concluded that “the difficulty in 'mining 'for coins; the rate of unit production; and the cryptographic algorithm employed” are the critical elements. Kaplan, Aslan, and Bulbul (2018) focused on examining the world of mouth effect established through another social media platform Twitter and its prospective impact on the price of

cryptocurrencies. The regression analysis revealed that a correlation between rumors regarding altcoins and price value exists. Yilmaz and Hazar (2018a) state that investors choose cryptocurrencies based on five primary factors, which allows one to create a cryptocurrency that would correspond to all these elements and make predictions regarding the market success of a particular coin. DeVries (2016, p. 2) argues that "cryptocurrencies could possibly be the single most disruptive technology to global financial and economic systems." Thus, this technology is essential for the global economy and should be analyzed carefully.

Firstly, the decentralized nature of blockchain mitigates the traditional disadvantages of other online payment methods – including commissions, chargebacks, risks of doubles endings, or possible fraud. Heid (2014) provides an assessment of this financial market and argues that Bitcoin is a successful proof of concept for the blockchain technology. Despite the fact that the cryptocurrency market is more secure when compared to traditional financial markets, it has been subjected to various attackers. Heid (2014) provides examples of data breaches and attempts to target end users, and other malicious actions that were possible due to the fact that the protocol was experimental. The author presents the following explanation of the algorithm that allows blockchain transactions to function – the encrypted algorithm generates precomputed files each containing a pair of public and private keys and assigned to a specific owner. Next, individuals engaged in transactions can send data stored in the file walled.dat on their hard drive to other users. Dynamic wallet addresses contain information about the private keys while public keys contain information about the destination of a payment.

One crucial element of the cryptocurrency market is that a specified number of coins for each cryptocurrency exists. Therefore, the value of these elements is based on the supply and demand laws, which are significant forces within the cryptocurrency market. Zheng et al. (2018) conducted a study examining the existing articles that research the financial market and application of AI. The authors introduced the concept of financial intelligence, which is critical for this market. In addition, the level of difficulty in regard to mining a coin impacts the final price because it fluctuates depending on the conditions. The nature of this protocol implies no need for a third party because the open source protocol allows the transactions to be secure and reliable. In general, cryptocurrencies can be purchased through exchange marketplaces using fiat currencies. Thus, one can conclude that "the market is diverse and provides investors with many

different products" (Alessandretti et al. 2018, p. 1). Thus, forecasting within this domain will become increasingly important in the future.

To understand the specifics of the markets a study focusing on its functioning was explored. Hitam and Ismail (n.d.) compare the performance of different machine learning algorithms in regard to their ability for prediction changes in the cryptocurrencies market. The study uses technical analysis strategies for time series data forecasting which implies an assessment of information within the market such as price, the volume of sales, and future predictions. It disregards other elements applied in the fundamental analysis including outside forces that may impact the financial markets. In general, the algorithms proposed by the authors of the discussed study is reasonably successful in making accurate cryptocurrency market predictions.

An article by Farell (2015) offers a comprehensive analysis of the cryptocurrency market, which provides an understanding of the underlying forces guiding its development. Yamada and Nakajima (2016) introduce the concept of micro pricing in blockchain to develop a framework for understanding human behavior and its implications for the economy and financial markets. Gandal and Halaburda (2014) present an article that examines competition in the cryptocurrency marker using the network effect model. The findings suggest that currently cryptocurrencies are viewed as financial assets, which provides implications for developing AI for market predictions.

2.7. Algorithms and Methods of Predictions

Application of AI in the context of analysis and prediction of the price is a valid strategy that can be used to make investment recommendations in regards to cryptocurrencies. However, many components should be considered. Another study that applies AI in the form of neural networks for prediction of price fluctuation of cryptocurrencies was conducted by Gullapalli (2016), who used this framework to make predictions regarding high and closing prices of bitcoin on a daily basis. Both time delay and recurrent neural networks were used in this experiment to account for a variety of factors that may impact the market. In order to train these neural networks, historical data regarding the price of bitcoin as a reference. Components such as quarterly highest and the lowest value, closing costs, and volume of demand were taken into account by Gullapalli (2016).

In general, the algorithm developed by the author, more specifically the time-delay model was successful at predicting the price.

The ethical element of AI analysis is an essential focus of this paper. Wallach (2010) explores the question of ethics and decision-making the process that impacts the way humans choose to act. This research is relevant to the question of AI analysis applied to cryptocurrencies because as was previously mentioned, this market is based solely on the value perception that individuals have. Cognitive mechanisms, such as reasoning, have an impact on this process an with an appropriate model, AI can mirror this reasoning process, which will help make adequate predictions regarding the cryptocurrencies. In his study, Wallac (2010) mainly focuses on developing robots that can be applied in a variety of domains without posing a threat by applying algorithms that allow them to use the ethical values of humans.

The emergence of blockchain and cryptocurrencies is a result of a need for improvement in some aspects of financial operations. Jani (2018) explores the development and use of cryptocurrencies in India to provide an understanding of the enhancements that the introduction of this technology has as well as challenges that it poses. According to the author, approximately 21 countries responded to the widespread use of cryptocurrencies by introducing regulations aimed at protecting citizens from fraud. This affects the expectations of users and thus the ability to apply AI algorithm for analyzing this market.

One can argue that the application of theories from traditional finance can be used to improve the AI algorithm in regard to cryptocurrencies. Khuntia and Pattanayak (2018) claim that the adaptive market hypothesis can be applied to cryptocurrencies and evaluate their theory using the example of bitcoin. In this regard, new information that people obtain has an impact on the price of their assets, in this case – cryptocurrencies, in accordance with the martingale difference sequence.

However, Khuntia and Pattanayak (2018) state that over time, financial markets change and adapt to new conditions and are subjected to behavioral bias, which may be the primary limitation of the data mining and AI analysis. In order to prove this concept, several studies were examined that focus on behavior. Krafft, Penna, and Pentland (2018) conducted an experiment

that can provide important implications for the future development of AI and data mining algorithms tailored to the analysis of cryptocurrencies. The authors acknowledge the fluctuations of this market that are facilitated by bias and perception of the buyers and sellers and aim to account for it in their study. Krafft, Penna, and Pentland (2018) created bots that purchased small amounts of cryptocurrencies over a timeframe of six months to measure the impact of these actions on the overall market. The findings suggest that such actions have a significant short term effect, which can be used as an important element of algorithm design.

Several researchers examined the prospects of machine learning and plausibility of applying this strategy for the cryptocurrency market. Misnik et al. (n.d., p. 96) studied the effect of applying neural networks on the predictions made regarding bitcoin prices using “multilayer perceptron (MLP) and Long short-term memory (LSTM) neural networks." The authors added social and time elements to the standard model to improve precision. It was concluded that LSTM is more efficient due to the fact that it considers more factors. It can be found that the inclusion of a variety of factors that have a direct impact on the financial market, including behavioral elements, can significantly increase the accuracy of the AI algorithm. Napiah (2018) applied hybridization machine learning to facilitate the process of predicting changes in the cryptocurrency market. Other researchers focused on used a variety of AI frameworks to design a model that will accurately predict price fluctuations, For instance, Catania, Grassi, and Ravazzolo (2018), McNally (2016), and Jiang and Liang (2016) used AI while Yilmaz and Hazar (2018b) used conjoint analysis. This presents a variety of evidence suggesting that AI and data mining can be used to make accurate predictions for cryptocurrencies.

Most researchers focus on applying traditional financial models when analyzing the market in question, which leads to several difficulties. Alessandretti et al. (2018) conducted a study targeting the machine learning algorithms that can be used to determine the market changes of cryptocurrencies. The authors used gradient boosting decision trees as the primary strategy for predicting the value of cryptocurrencies. The findings are based on the analysis of 1,681 currencies and suggest that similar models that apply AI can be used to produce a profit from cryptocurrencies. Currently, Alessandretti et al. (2018, p. 2) state that the following algorithms were used to analyze cryptocurrencies, more specifically bitcoin – "random forests, Bayesian

neural network, long short-term memory neural network, and other algorithms." In general, from 2013 till 2017 the cryptocurrency market remained stable, showcasing its long-term properties. Li et al. (n.d.) focused on determining the prospects of using deep learning strategies in the prediction of cryptocurrency market fluctuations. The Python Library Keras was used for the development of this model, and two neural networks were created for the purpose of this experiment.

Different optimizers were used to adjust these models, and the findings suggest that it is plausible to develop a working neural network model that makes successful predictions regarding the value of cryptocurrencies, in this case, bitcoin, over short periods of time. Both classification and regression problems were resolved by Lee et al. (2018). This study provides implications for further development of deep learning networks that can be used to predict value changes accurately, as the authors of the described experiment aimed at determining changes within 3%. Pelletier (2018) uses Aylien API to make predictions regarding bitcoin, however, the study occurred during the plateau period experienced by the cryptocurrency which was on the peak of its popularity prior. The author argues that news regarding bitcoin can significantly affect its price and suggests using Natural Language Toolkit as an AI model for analysis. However, Pelletier (2018) states that according to the results of his studies such information impacts only individual users and has little effect on investors. Similar findings are introduced by Lamon, Nielsen, and Redondo (2017) who argue that both news and social media can be used to predict the price fluctuations of the cryptocurrency market. More specifically, the authors focus on the three popular currencies – bitcoin, litecoin, and etherum.

Phillips and Gorse (2017) improved the previously described algorithm that accounted for news and social media rumors. This was done by applying epidemic modeling that can be used to detect prospective cryptocurrency bubbles and thus avoid risky investments. Lee, Ulkuatam, Beling, and Scherer (2018) used inverse reinforcement learning together with agent-based modeling to design an algorithm that would allow one to make accurate predictions in the cryptocurrencies market.

The changes in the price that are prevalent for the market in question require additional attention. Conrad, Custovic, and Ghysels (2018) examined the volatility of bitcoin over both short and long timeframes using GARCH-MIDAS model. The findings suggest that in general changes in the price of bitcoin do not correspond with the standard volatility frameworks which can be observed in financial markets. However, the global economy and events affecting it have a direct impact on bitcoin, which provides implications for further analysis of occurrences affecting cryptocurrencies. Radityo, Munajat, and Budi (2017) used artificial neural networks, while Karasu, Hacioglu, and Atlan (2018) applied time series machine learning. The results provide an understanding of successful strategies that can be used to predicts volatility.

Accounting for social media impact is critical for accurate cryptocurrency forecasting. Islam et al. (2018) used text mining for financial analysis and understanding the implication of news in regard to market value. The study focuses on developing a variety of frameworks that can be used to classify significant information, which can be used for improving the AI algorithm. The authors suggest that "within the text mining techniques, stock predictive model (SPM) need to improve by the rank search method with the inclusion of gains and ratios along with forwarding selection methods by integrating dimensionality reduction techniques" (Islam et al. 2018, p. 770). Cocco, Concas, and Marchesi (2017) conducted an experiment in which the researchers used an artificial model of the cryptocurrency market. This model is helpful due to its recumbence with the actual cryptocurrency market and ability to showcase autocorrelation. This study is useful because the model can be used to test different strategies of investment in the cryptocurrencies market. Radityo, Munajat, and Budi (2017) conducted a study comparing several AI algorithms in their ability to predict the fluctuations of bitcoin accurately. The authors argue that AI was proven to have higher accuracy and efficiency than other strategies. Additionally, one should note that features included in the process of designing an AI network. Thierer and Castillo (2016) state that AI technology raises a variety of concerns from policymakers, however, the prospects of this technology can bring enormous economic and social benefits. When designing data mining and AI algorithms, some elements from the traditional financial analysis can be used to understand the basics of trading. Jegadeesh and Titman (1993) provide an assessment of the winners and losers framework that can be applied to

the modern day blockchain. This assessment offers implications for understanding the traditional financial market and buy-sell strategies that can be profitable.

Overall, this research paper explores the topic of AI and the application of the algorithm in the blockchain. In general, the findings suggest that the emergence of AI 2.0 and the development of the third Fintech stage can offer a variety of advantages for the financial industry. The studies examined in this paper suggest that AI and data mining can be applied to making an accurate prediction in regard to fluctuations in the financial market.

3. Chapter Three: Research Methodology

3.1. Data Choosing & Collection

For the calculations and predictions that will be conducted in this research the daily crypto currency data has been collected for Kraken crypto currency exchange market. Each crypto currency exchange market has its own set of unique crypto currency. From the well-known exchange cryptocurrency platform Kraken the following type of exchange values have been chosen: BTC/EUR, BTC/USD, ETH/BTC, ETH/EUR, ETH/USD, LTC/BTC, LTC/EUR, LTC/USD, XRP/USD.

All of the data were downloaded from CryptoDataDownload. CryptoDataDownload is a platform that provides free historical cryptocurrency for people to use in research and analysis. All of the data sets timespan are from the beginning of the first time that the cryptocurrency have been added for exchange in the crypto currency market.

3.2. Data Cleansing and Representation

All of the above-mentioned data sets have been structured into a unified format for easy processing. The format of the data consists of the following data sets:

Kraken_BTCEUR_d

Date	Symbol	Open	High	Low	Close	Volume BTC	Volume EUR
2019-10-08	BTCEUR	7484.2	7595	7423.3	7481.3	1053.79	7893982.15
2019-10-07	BTCEUR	7170.1	7574.8	7085	7484.2	5802.08	42800140.84
2019-10-06	BTCEUR	7420	7448	7108	7170.1	3335.28	24184058.52
2019-10-05	BTCEUR	7445	7470.7	7315.3	7420	2061.75	15227261.32
2019-10-04	BTCEUR	7515.2	7521.3	7325	7445	2698.83	20053234.43
2019-10-03	BTCEUR	7658.3	7688.5	7364.8	7515.2	3361.66	25183131.97
2019-10-02	BTCEUR	7618.5	7664.3	7470.6	7658.3	3341.02	25237082.75
2019-10-01	BTCEUR	7619	7820	7530.6	7618.5	4727.83	36347468.29

Figure 1: Crypto Currency Collected Data Format

The date column indicates for the data of the sample. The Symbol column referred to the exchange code that is being used in the market exchange, usually it indicates the two currencies that are being used to buy and sell the crypto currency. The Open column refers to the opening crypto currency value of the trade in that day. The High column refers to the highest crypto currency value of the trade in that day. The Close column refers to the closing crypto currency value of the trade in that day.

3.3. Processing Language

In order to be able to rapidly develop the algorithms the student have chosen to use Python as a supporting programming language to develop the algorithms.

Python's is an easy to learn easy to read language meant to be understandable similar to the human language Guttag, John V. (12 August 2016). Python have many modules and components that can be used in conjunction for rapid development. Python is an open source project which can be modified and developed for anyone, it also has two main version used by scientists, Version 2.7 and Version 3.

3.4. Data Storage

For sake of simplicity and performance matters, the student has kept all data in CSV format and did not integrate them in enterprise level databases such as MySQL Peter Zaitsev (10 April 2008).

3.5. Used Prediction Algorithms

Mainly six predicting algorithms has been used in this research, each one of the chosen algorithms will be used to predict the crypto currency market value and will be validated using another set of data. All of the prediction algorithms have been implemented using Python since python enables the researcher to apply the algorithms rapidly using available modules. For all of the below algorithm implementation the crypto currency exchange data set has been validated into training data set and validation data set. The training data set would be used to train the model and the validation data set would be used to calculate the accuracy of the algorithm. In

order to understand the accuracy of the algorithm the Root Mean Square Error (RMSE) would be used in order to verify the fit of the model to the data and to show how close the crypto currency data are to the model's predicted values Hyndman, Rob J.; Koehler, Anne B. (2006). Root mean square error is usually used in scientific researches and in regression analysis to validate experimental calculations.

3.5.1. Moving Average

Moving average is a statistical calculation method used to analyze data by creating a series of averages from different group of data. It has many other names such as rolling average, running average. There are different variations and types of the algorithm such as simple moving average, cumulative moving average and weighed forms moving average.

In the algorithm the data groups must be categorized in fixed subset size. The first group of the moving average is taken by the average of the initial subset of data series. After that, the subset of data is shifted forward by removing the first number from the calculation and including the next values. The shifting goes on and on until it reaches the end of the series.

Usually moving average is used with time series data such in our case with crypto currency exchange data where every day we have a record for the value of the crypto currency. Moving average helps in smoothing out short term fluctuations and gives more focus into long term predictions. Mainly Moving Average is used in analysis of financial data such as exchange predictions or trading volumes. Moving average is also considered to be a type of convolution thus it also can be seen used in signal and processing.

Simple moving average can be calculated is following:

$$\text{Simple moving average} = (P1 + P2 + P3 + P4 + ... + Pn) / n$$

In the student implementation data frames were used in order to conduct shifting and calculation for the Moving Average.

The student has developed the algorithm by using Python and open source available libraries that helped in rapid development. The source code of the algorithm can be found in Appendix 7.1.1

3.5.2. Linear Regression

The second used and discussed algorithm is Linear Regression. Linear Regression is used to model relationships between dependent and independent values. In case there is one dependent value the algorithm is called Simple Linear Regression. In case more than one it is called Multiple Linear Regression. It is considered to be one of the basic machine learning algorithms. Furthermore, Linear regression was the first type of regression analysis to be researched heavily.

The Linear Predictor functions are used in conjunction to model the relationship of unknown model parameters. The linear regression gives a model that decides the relationship between the independent variables and the dependent variable.

$$Y = \theta_1 X_1 + \theta_2 X_2 + \dots \theta_n X_n$$

Figure 3: Linear Regression Equation

In the student implementation the imported data has been sorted by date and features has been extracted from the dates in order to fit the model. Features were extracted using fastai library which extracts features. These set of features will be used in the predictive model for the Linear Regression.

The student has developed the algorithm by using Python and open source available libraries that helped in rapid development. The source code of the algorithm can be found in Appendix 7.1.2

3.5.3. Auto ARIMA

ARIMA is a time series forecasting method that stands for AutoRegressive Integrated Moving Average. In ARIMA the data series does not move thus the mean and variance should not have expanding gap with time. In ARIMA by using log transformation or differencing method the time series can be made stationary.

On the other hand, Auto ARIMA is an enhanced method for prediction which eliminates the need make the data frame stationary plus you do not need to specify autoregressive term and the moving average term.

The student has developed the algorithm by using Python and open source available libraries that helped in rapid development. The source code of the algorithm can be found in Appendix 7.1.3

3.5.4. k-Nearest Neighbors

k-nearest neighbors' algorithm (k-NN) is a classification and regression supervised machine learning algorithm used for prediction. Input is grouped into k closest training models since the Knn assumed that similar data points are close to each other. One of the main features of kNN algorithm is that its training is fast because it does not use the training data points to do any generalization. However, some of the cons of the algorithm that it requires a high memory space and it can be sensitive to irrelevant features.

The student has developed the algorithm by using Python and open source available libraries that helped in rapid development. The source code of the algorithm can be found in Appendix 7.1.4

3.5.5. Long Short Term Memory (LSTM)

Long Short Term Memory is a recurrent neural network algorithm used in deep learning. Recurrent neural networks are networks with loops in them, allowing data to sustain. With this feature it can not only process single data source (such as voice recording), but also entire sequences of data.

The student has developed the algorithm by using Python and open source available libraries that helped in rapid development. The source code of the algorithm can be found in Appendix 7.1.5

4. Chapter Four: Experimental Analysis

For all of the below testing scenarios 80% of the data has been used to train the model, and the remaining 20% were used to validate the result of the prediction. For all of below graphs that represents the prediction results, the blue line represents the known data that were used for the training model, the green line represents the predicted data and the orange line represents the known data used for validation.

4.1. Cryptocurrency BTC/USD Exchange Prediction

All of the above prediction algorithm was ran on the BTC/USD exchange data and the below is the results of the tests. Long Short Term Memory (LSTM) gave out the best results with the best RMS value.

4.1.1. Moving Average Result

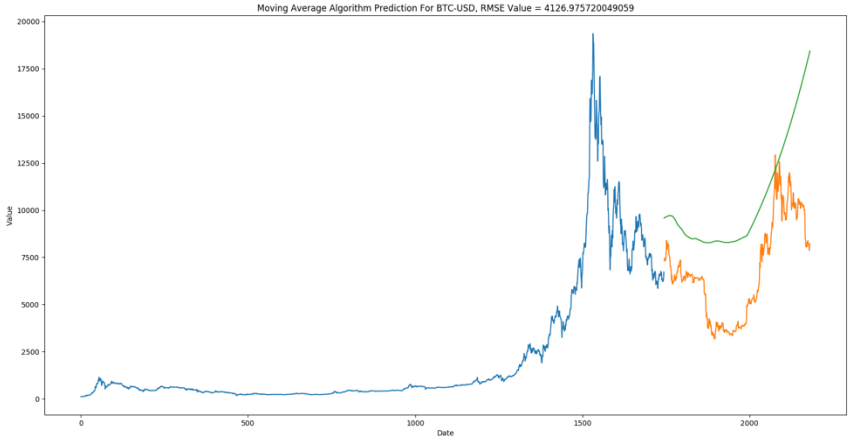


Figure 8: Moving Average Algorithm Prediction Results

RMSE Value is: 4126.97

4.1.2. Linear Regression Result

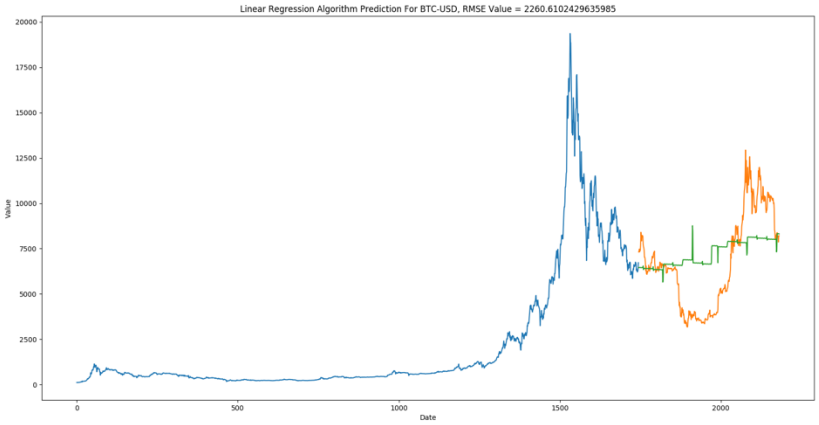


Figure 9: Linear Regression Algorithm Prediction Results

RMSE Value is: 2260.6

4.1.3. Auto ARIMA Result

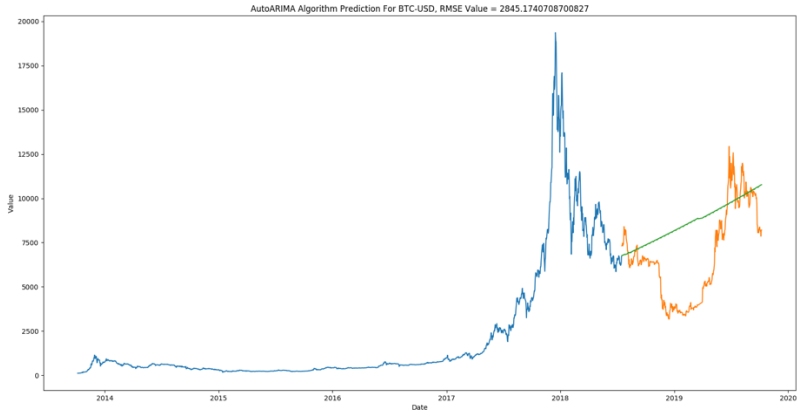


Figure 10: Auto ARIMA Algorithm Prediction Results

RMSE Value is: 2845.17

4.1.4. k-Nearest Neighbours Result

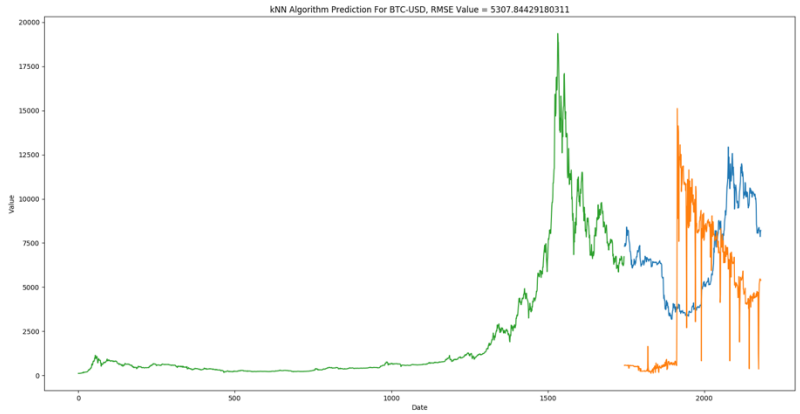


Figure 11: k-Nearest Neighbors Algorithm Prediction Results

RMSE Value is: 5307.84

4.1.5. Long Short Term Memory (LSTM) Result

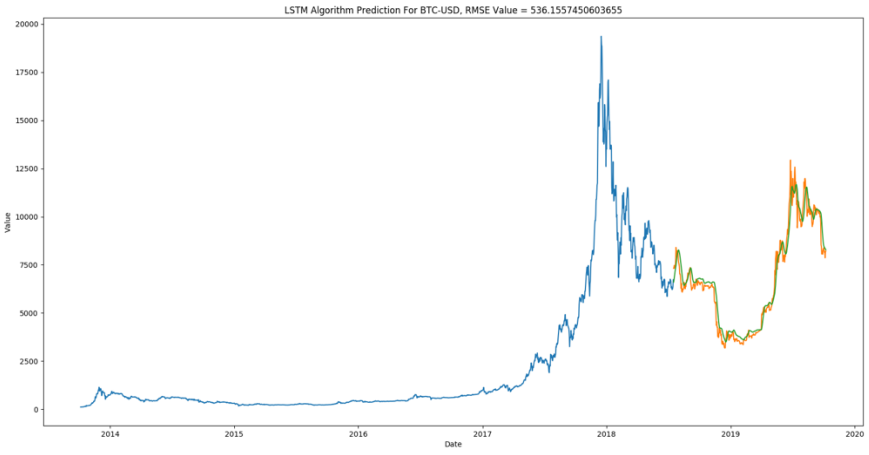


Figure 12: Long Short Term Memory (LSTM) Algorithm Prediction Results

RMSE Value is: 536.155

4.2. Cryptocurrency ETH/BTC Exchange Prediction

All of the above prediction algorithm were ran on the ETH/BTC exchange data and the below is the results of the tests. Long Short Term Memory (LSTM) gave out the best results with the best RMS value.

4.2.1. Moving Average Result

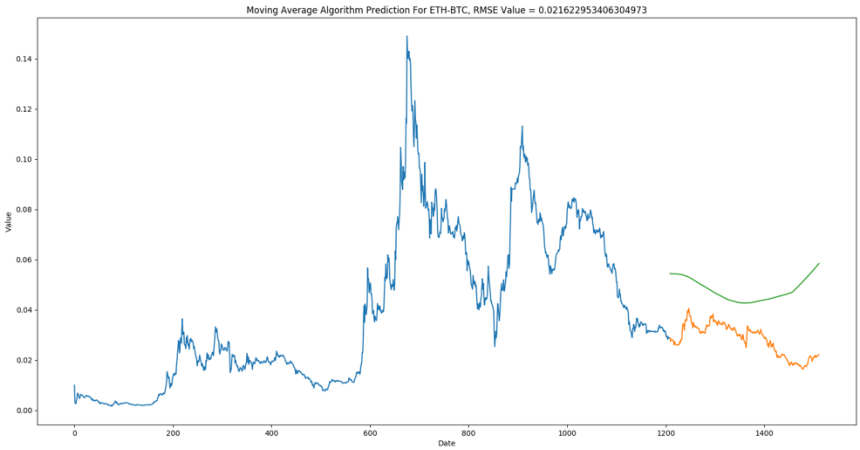


Figure 13: Moving Average Algorithm Prediction Results

RMSE Value is: 0.021

4.2.2. Linear Regression Result

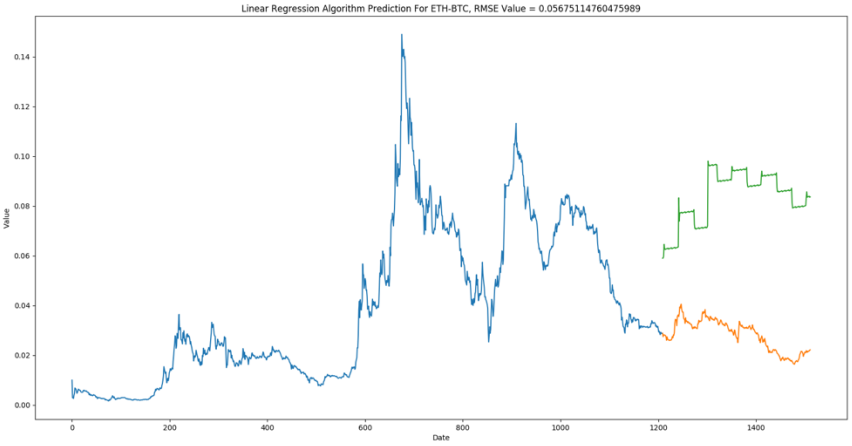


Figure 14: Linear Regression Algorithm Prediction Results

RMSE Value is: 0.056

4.2.3. Auto ARIMA Result

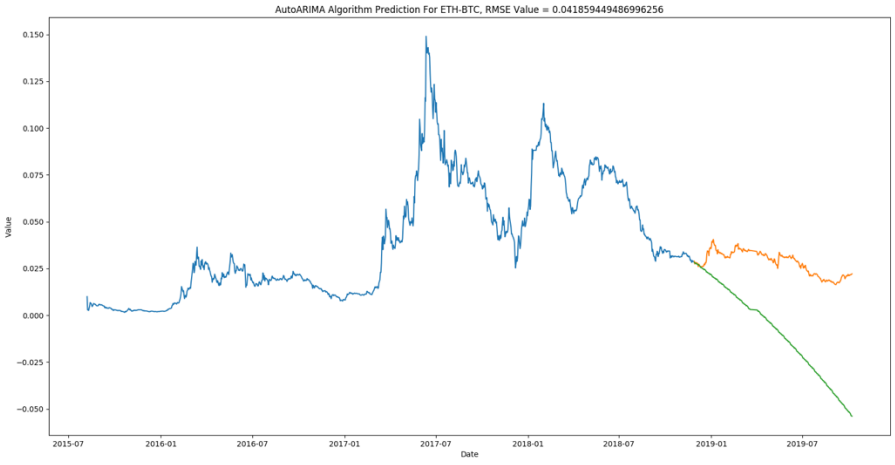


Figure 15: Auto ARIMA Algorithm Prediction Results

RMSE Value is: 0.041

4.2.4. k-Nearest Neighbors Result

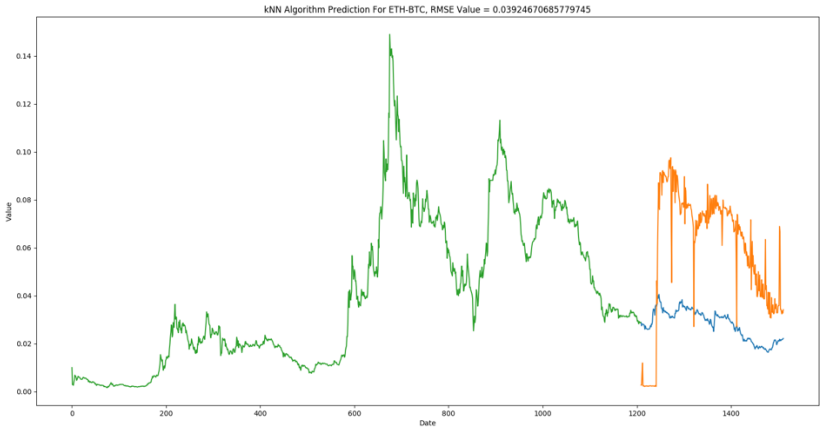


Figure 16: k-Nearest Neighbors Algorithm Prediction Results

RMSE Value is: 0.0392

4.2.5. Long Short Term Memory (LSTM) Result

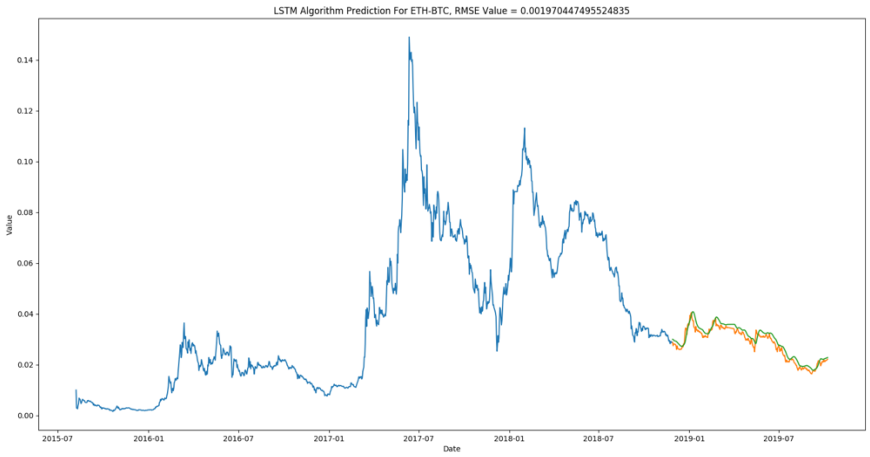


Figure 17: Long Short Term Memory (LSTM) Algorithm Prediction Results

RMSE Value is: 0.0019

4.3. Cryptocurrency ETH/USD Exchange Prediction

All of the above prediction algorithm was ran on the ETH/USD exchange data and the below is the results of the tests. Long Short Term Memory (LSTM) gave out the best results with the best RMS value.

4.3.1. Moving Average Result

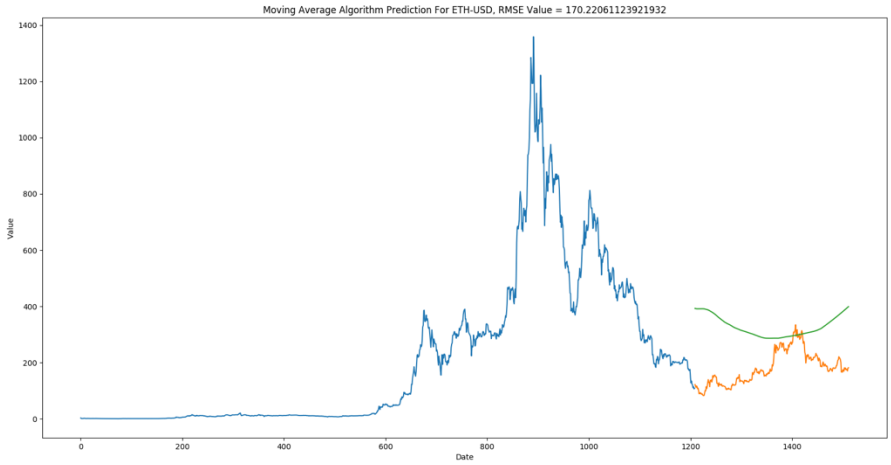


Figure 18: Moving Average Algorithm Prediction Results

RMSE Value is: 170.22

4.3.2. Linear Regression Result

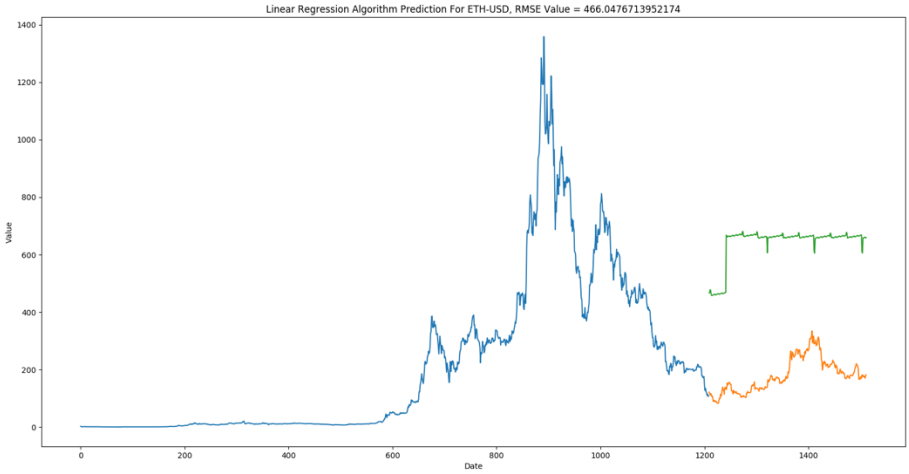


Figure 19: Linear Regression Algorithm Prediction Results

RMSE Value is: 466.0

4.3.3. Auto ARIMA Result

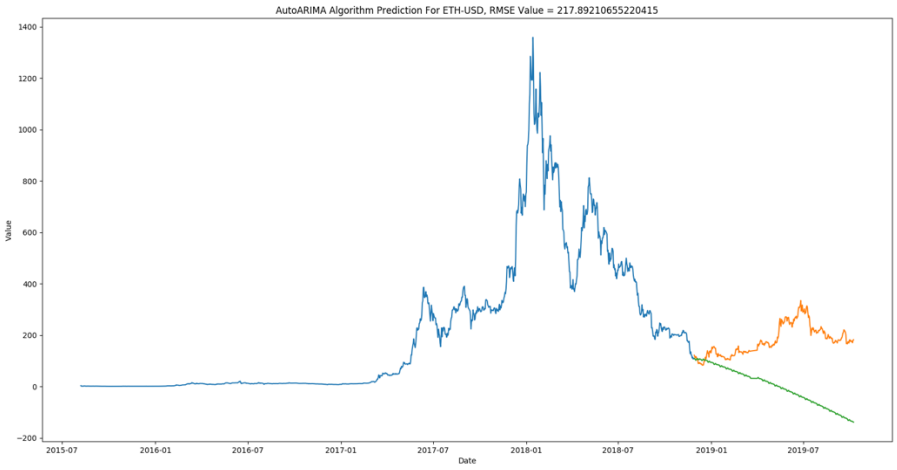


Figure 20: Auto ARIMA Algorithm Prediction Results

RMSE Value is:217.89

4.3.4. k-Nearest Neighbours Result

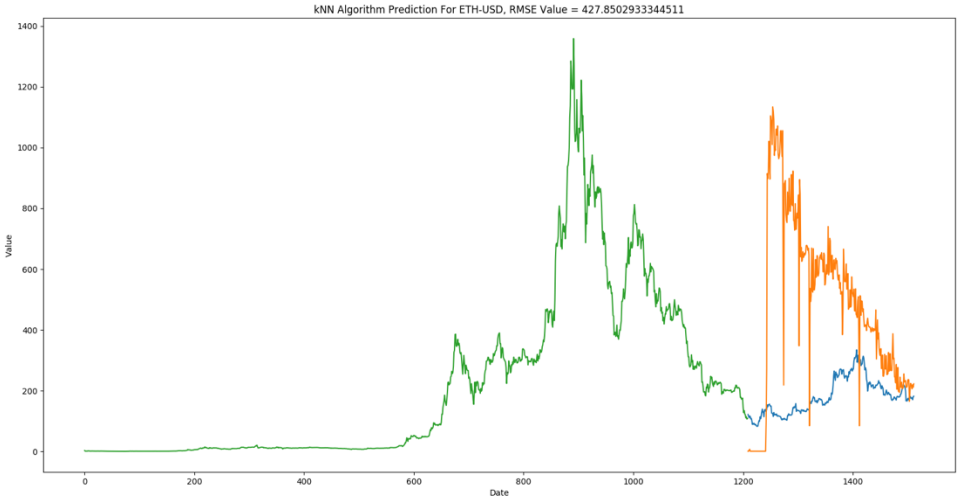


Figure 21: k-Nearest Neighbors Algorithm Prediction Results

RMSE Value is:427.8

4.3.5. Long Short Term Memory (LSTM) Result

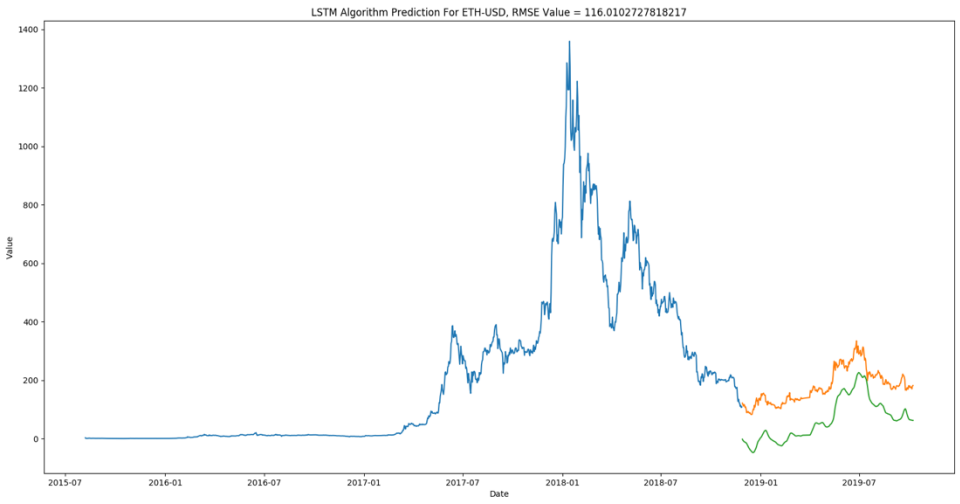


Figure 22: Long Short Term Memory (LSTM) Algorithm Prediction Results

RMSE Value is: 116.01

4.4. Cryptocurrency LTC/BTC Exchange Prediction

All of the above prediction algorithm was ran on the LTC/BTC exchange data and the below is the results of the tests. Long Short Term Memory (LSTM) gave out the best results with the best RMS value.

4.4.1. Moving Average Result

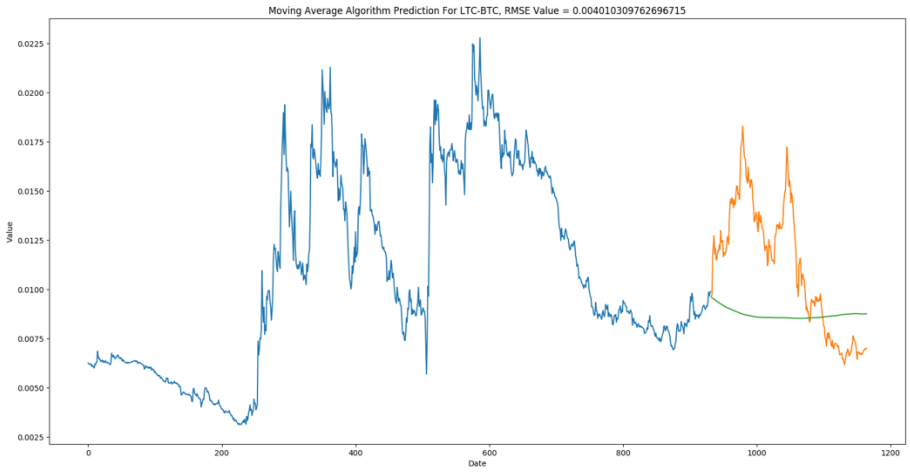


Figure 23: Moving Average Algorithm Prediction Results

RMSE Value is: 0.0040

4.4.2. Linear Regression Result

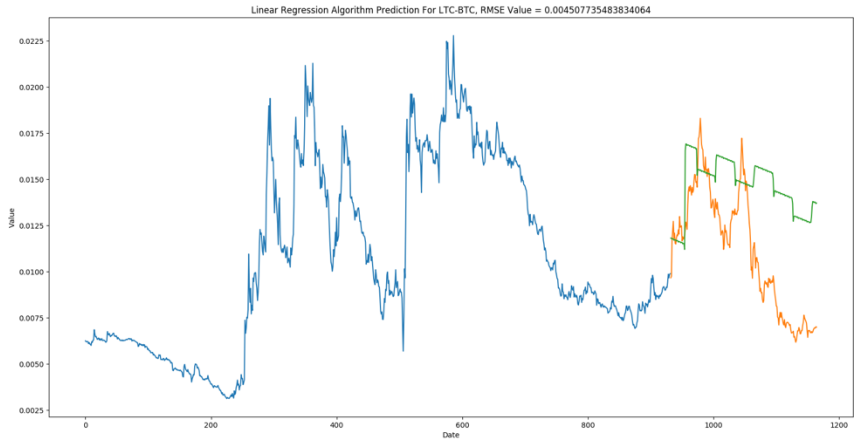


Figure 24: Linear Regression Algorithm Prediction Results

RMSE Value is: 0.0045

4.4.3. Auto ARIMA Result

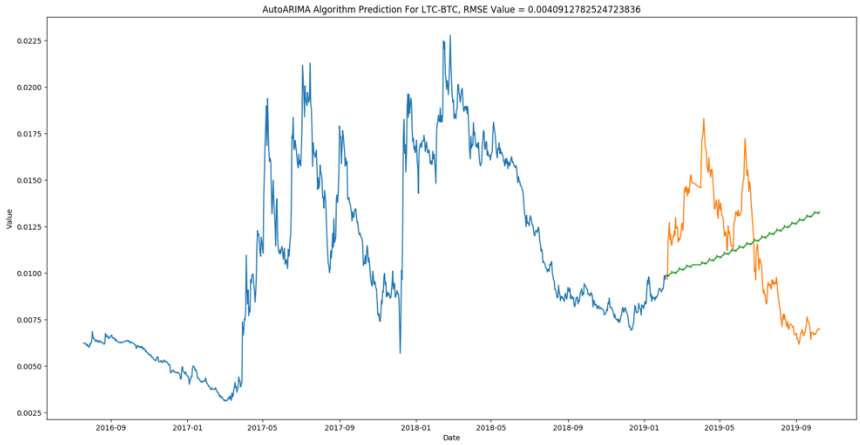


Figure 25: Auto ARIMA Algorithm Prediction Results

RMSE Value is: 0.0040

4.4.4. k-Nearest Neighbours Result

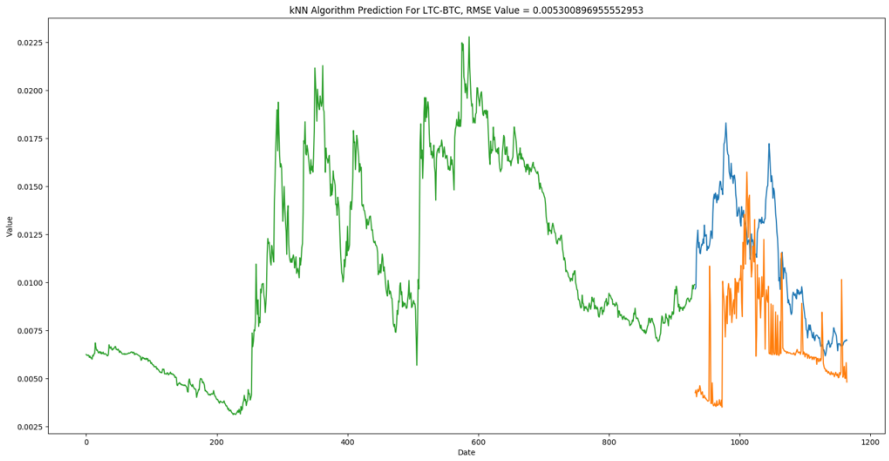


Figure 26: k-Nearest Neighbors Algorithm Prediction Results

RMSE Value is: 0.005

4.4.5. Long Short Term Memory (LSTM) Result

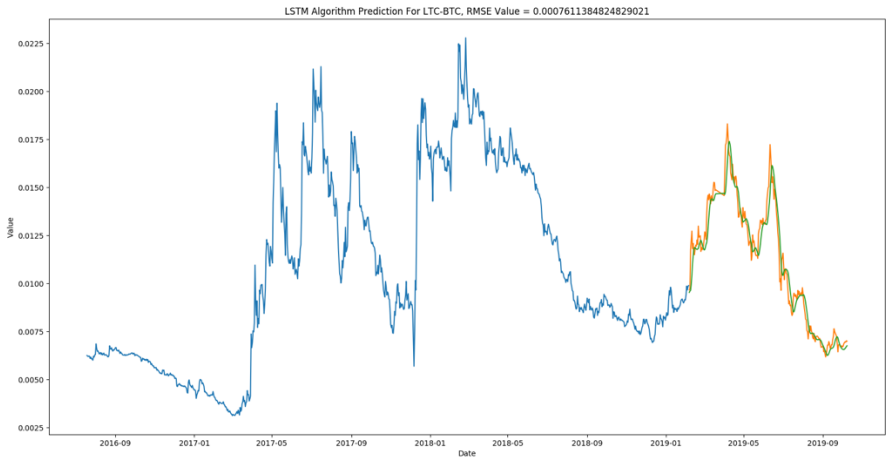


Figure 27: Long Short Term Memory (LSTM) Algorithm Prediction Results

RMSE Value is: 0.00076

4.5. Cryptocurrency LTC/USD Exchange Prediction

All of the above prediction algorithm was ran on the LTC/USD exchange data and the below is the results of the tests. Long Short Term Memory (LSTM) gave out the best results with the best RMS value.

4.5.1. Moving Average Result

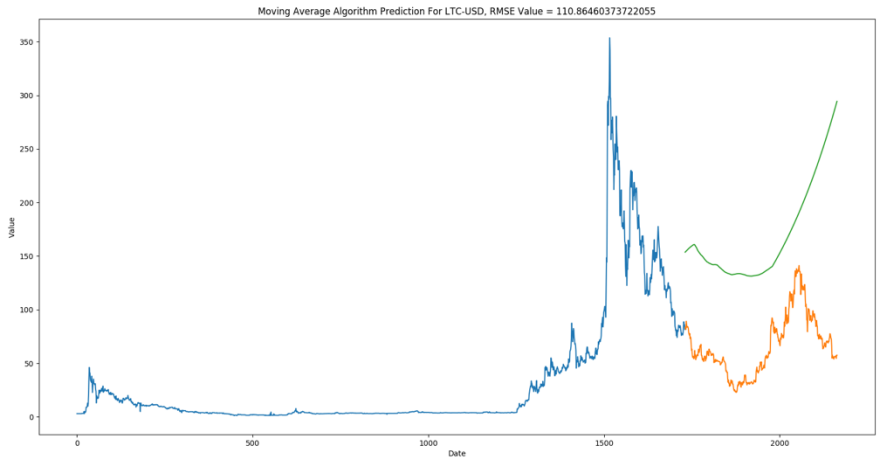


Figure 28: Moving Average Algorithm Prediction Results

RMSE Value is:110.86

4.5.2. Linear Regression Result

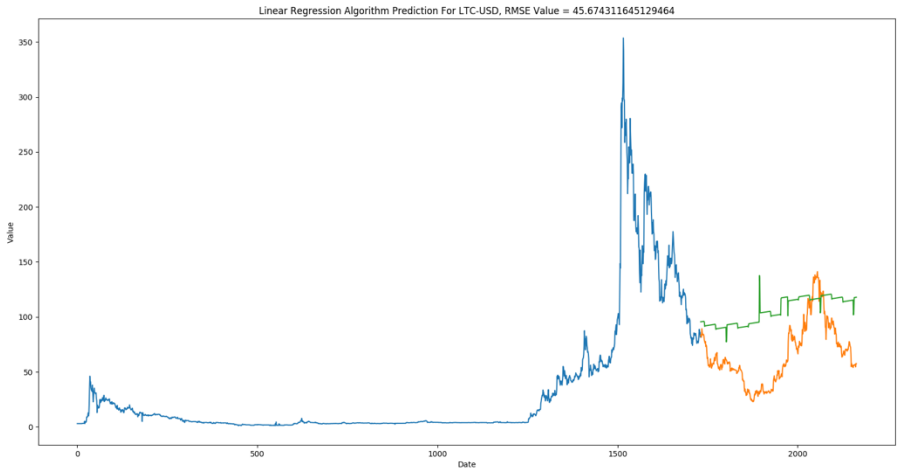


Figure 29: Linear Regression Algorithm Prediction Results

RMSE Value is: 45.67

4.5.3. Auto ARIMA Result

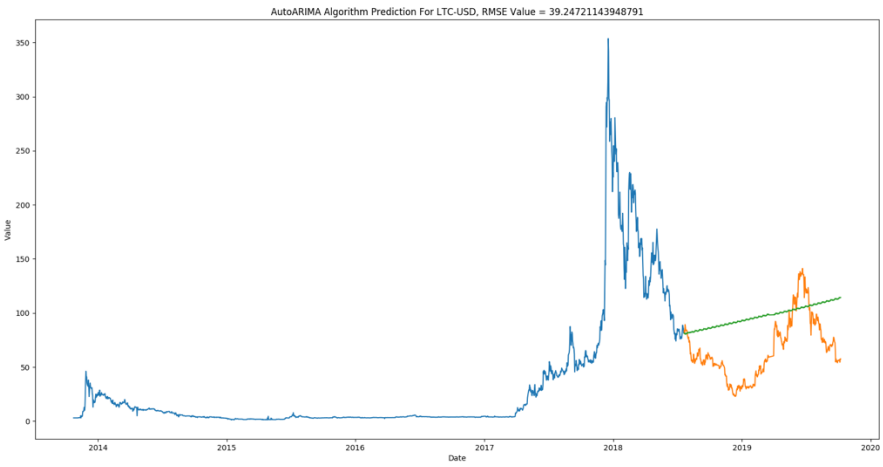


Figure 30: Auto ARIMA Algorithm Prediction Results

RMSE Value is: 39.24

4.5.4. k-Nearest Neighbors Result

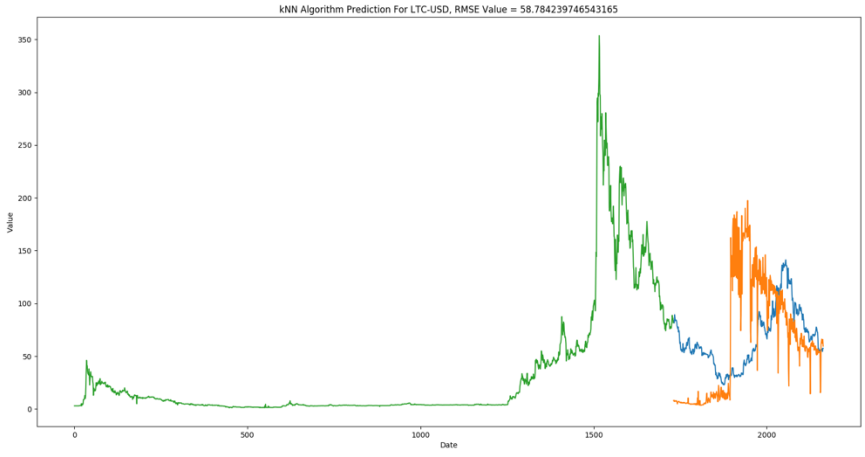


Figure 31: k-Nearest Neighbors Algorithm Prediction Results

RMSE Value is: 58.78

4.5.5. Long Short Term Memory (LSTM) Result

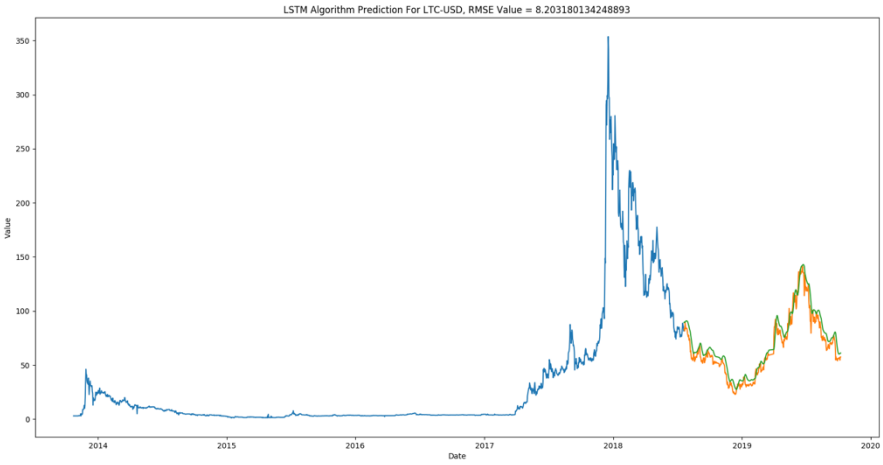


Figure 32: Long Short Term Memory (LSTM) Algorithm Prediction Results

RMSE Value is: 8.203

5. Chapter Five: Conclusion and Future Work

5.1. Conclusion

Cryptocurrency creates a globalized system for the realization of transparent and accountable exchange mediums. Prevention of double spending and ensuring distributed storage goes a long way in ensuring that individuals are protecting from fraud. However, there is a need for regulation of cryptocurrencies to prevent the susceptibility of the platform to criminal misuse. Ensuring acceptable transactions increase the social benefits associated with blockchain technology, which has allowed entities to create communities and control the storage and access of information. Prevention of unnecessary control, privacy intrusion, and censorship of information increase societal creativity and optimize the benefits of business transactions.

Throughout the thesis it was found out that it is possible to predict the stock market prices of crypto currency with some margin of error. There are different algorithms that can be used, in the student experiment the most suitable one and less error prone was the Long Short Term Memory (LSTM) algorithm. It showed less RMSE values since it takes into advantage its properties to use previous historical data.

Going back to the research questions was put down previously, the research in conclusion have answered each one of them by the following:

1. **Q1: To which extend can we predict Cryptocurrency exchange market prices using well known prediction algorithms?**
   A: It was found out that it is possible to predict the exchange market prices, however the farer the period of the predicted cryptocurrency value the less accurate it would be due to the other factors that the algorithms cannot predict such as events, news, hacktivism and any even that can affect the investors. Solutions for increasing the accuracy has been added to the Future Work section.
2. **Q2: How accurate is the prediction of the prices?**

A: The RMSE value of the Long Short Term Memory (LSTM) Algorithm and the k-Nearest Neighbors Algorithm showed very promising results. Their respective graphical results of the algorithm in comparison with the real results showed similarities in terms of the slope direction and in term of the prices. However, for both implementation there is a loss in the accuracy due to neglecting other facts such as events and news which can affect the investors decisions in the cryptocurrency community.

5.2. Future Work

For the future many things can be improved and added to the research. For a start the student can explore further techniques and algorithms such as Support Vector Machine Regression and XGBoost. Also, in order to enrich the set of features that are used in the algorithm the student could integrate social media and news feeds into the data and corelate it with the changes in the crypto currency stock market exchange. For example, a news that get released in the social media of a crypto currency exchange market getting hacked could lead into huge selling requests from the investors and lead to the price of the crypto currency to go down. Furthermore, the student can build comprehensive graphical user interface for the prediction components in order to make it easy to insert and predict crypto currency exchange data.

6. References

Alessandretti, L, ElBahrawy, A, Aiello, LM, & Baronchelli, A 2018, 'Anticipating cryptocurrencies prices using machine learning', Complexity, vol. 2018, pp. 1-16.

Azouvi, S, Maller, M & Meiklejohn, S 2019 'Egalitarian society or benevolent dictatorship: the state of cryptocurrency governance', International Conference on Financial Cryptography and Data Security, Nieuwpoort, Curaçao, pp. 127-143.

Catania, L, Grassi, S & Ravazzolo, F 2018, Predicting the volatility of cryptocurrency time–series, viewed 8 May 2019, <https://ideas.repec.org/p/bny/wpaper/0061.html>.

Cocco, L, Concas, G, & Marchesi, M 2017, 'Using an artificial financial market for studying a cryptocurrency market', Journal of Economic Interaction and Coordination, vol. 12, no. 2, pp. 345–365

Coeckelbergh, M & Reijers, W 2015, 'Cryptocurrencies as narrative technologies', ACM SIGCAS Computers and Society, vol. 45, no. 3, pp. 172-178.

Colianni, S, Rosales, S & Signorotti, M 2015, Algorithmic trading of cryptocurrency based on twitter sentiment analysis, viewed 8 May 2019, <http://cs229.stanford.edu/proj2015/029_report.pdf>.

Conrad, C, Custovic, A & Ghysels, E 2018 'Long- and short-term cryptocurrency volatility components: a GARCH-MIDAS analysis' Journal of Risk and Financial Management, vol. 11, no. 2, pp. 1-12.

Darlington, JK 2014, The future of bitcoin: mapping the global adoption of world's largest cryptocurrency through benefit analysis, viewed 8 May 2019, <https://trace.tennessee.edu/utk_chanhonoproj/1770>.

Delgado-Segura, S, Pérez-Solà, C, Herrera-Joancomartí, J, Navarro-Arribas, G & Borrell, J 2018, 'Cryptocurrency networks: a new P2P paradigm', Mobile Information Systems, vol. 3, pp. 1-16.

Delmolino, K, Arnett, M, Kosba, A, Miller, A, Shi, E 2016 'Step by step towards creating a safe smart contract: lessons and insights from a cryptocurrency lab', viewed 8 May 2019, <https://www.researchgate.net/publication/307507489_Step_by_Step_Towards_Creating_a_Safe_Smart_Contract_Lessons_and_Insights_from_a_Cryptocurrency_Lab>.

DeVries, PD 2016, ‘An analysis of cryptocurrency, bitcoin, and the future’, International Journal of Business Managnment and Commerce, vol. 1, no. 2, pp. 1-9.
ElBahrawy, A, Alessandretti, L, Kandler, A, Pastor-Satorras, R, & Baronchelli A 2017 ‘Evolutionary dynamics of the cryptocurrency market’, Royal Society Open Science, vol. 4, no. 170623, pp. 2-9.
Farell, R 2015 ‘An analysis of the cryptocurrency industry’, Wharton Research Scholars, p. 130-152.
Gandal, N & Halaburda, H 2014, Competition in the cryptocurrency market, viewed 8 May 2019, < https://www.bankofcanada.ca/wp-content/uploads/2014/08/wp2014-33.pdf>.
Gandal, N, & Halaburda, H 2016, ‘Can we predict the winner in a market with network effects? Competition in cryptocurrency market’, Games, vol. 16, no. 2, pp. 2-21.
Gullapalli, S 2004, Learning to predict cryptocurrency price using artificial neural network models of time series, viewed 8 May 2019, <https://krex.k-state.edu/dspace/handle/2097/38867>.
Hassani, H, Huang, X & Silva, E 2018, ‘Big-crypto: big data, blockchain and cryptocurrency’, Big Data and Cognitive Computing, vol. 2, no. 4, p.34-49.
Hayes, A, 2015, ‘Cryptocurrency value formation: an empirical analysis leading to a cost of production model for valuing bitcoin’, MCIS 2015 Proceedings, Stockholm, Sweden, p. 1308-1321.
Heid, A 2014, Analysis of the cryptocurrency marketplace, viewed 8 May 2019, <https://hackmiami.org/2014/04/20/whitepaper-analysis-of-the-cryptocurrency-marketplace/>.
Hitam, NA & Ismall, AR n.d. ‘Comparative performance of machine learning algorithms for cryptocurrency forecasting’, Indonesian Journal of Electrical Engineering and Computer Science, vol. 11, no. 3, pp. 1-11.
Hughes, Scott D. "Cryptocurrency Regulations and Enforcement in the US." W. St. UL Rev. 45 (2017): 1.
Islam, MR, Al-Shailkhli, IM, Nor, RB & Varadarajan, V 2018, ‘Technical approach in text mining for stock market prediction: a systematic review’, Indonesian Journal of Electrical Engineering and Computer Science, vol. 10, no. 2, pp. 770-777.
Jani, S 2018, The growth of cryptocurrency in India: its challenges & potential impacts on legislation, viewed 8 May 2019,

<https://www.researchgate.net/publication/324770908_The_Growth_of_Cryptocurrency_in_India_Its_Challenges_Potential_Impacts_on_Legislation>.
Jegadeesh, N & Titman, S 1993, 'Returns to buying winners and selling losers: implications for stock market efficiency', The Journal of Finance, vol. 48, no. 1. pp. 65-91.
Jiang, Z & Liang, J 2016 'Cryptocurrency portfolio management with deep reinforcement learning' Intelligent Systems Conference, London, UK, pp. 1-9.
Kaplan, C, Aslan, C & Bulbul, A 2018 'Cryptocurrency Word-of-Mouth Analysis viaTwitter', Conference: II. International Conference on Theoretical and Applied Computer Science and Engineering 2018-Summer, Istanbul, Turkey, pp. 1-10.
Karasu, S, Hacioglu, R & Atlan, A 2018, 'Prediction of Bitcoin prices with machine learning methods using time series data', 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, pp. 1-4.
Khuntia, S, Pattanayak JK 2018, 'Adaptive market hypothesis and evolving predictability of bitcoin', Economics Letters, vol. 167, pp. 26-28.
Krafft, P, Penna, N & Pentland, A 2018 'An experimental study of cryptocurrency market dynamics', Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18, New York, USA, pp. 1-13.
Lamon, C, Nielsen, E & Redondo, EF 2017, Cryptocurrency price prediction using news and social media sentiment, viewed 8 May 2019, <https://www.semanticscholar.org/paper/Cryptocurrency-Price-Prediction-Using-News-and-Lamon-Nielsen/c3b80de058596cee95beb20a2d087dbcf8be01ea>.
Laskowski, M & Kim, H 2016, 'Rapid prototyping of a text mining application for cryptocurrency market intelligence', 2016 IEEE 17th International Conference on Information Reuse and Integration (IRI), Pittsburg, USA, pp. 1-6.
Lee, DK, Guo, L & Wang, Y 2018 'Cryptocurrency: a new investment opportunity?', Journal of Alternative Investments, vol. 20, no. 3, pp. 16-40.
Lee, K, Ulkuatam, S, Beling, P & Scherer, W 2018 'Generating synthetic bitcoin transactions and predicting market price movement via inverse reinforcement learning and agent-based modeling', Journal of Artificial Societies and Social Simulation, vol. 21, no. 3, p. 5-16.

Li, B, Timaul, D, Lukanan, A, Lau, J, Steinmetz, E n.d. Applying deep learning to better predict cryptocurrency trends, viewed 8 May 2019, <https://pdfs.semanticscholar.org/fe77/d9c1285f3d07721a9cc92f85d972197a6932.pdf>.

Lindman, J, Rossi, M & Tuunainen, V 2017, 'Opportunities and risks of Blockchain Technologies in payments – a research agenda', Proceedings of the 50th Hawaii International Conference on System Sciences, Hawaii, USA, pp. 1533-1542.

McNally, S 2016, Predicting the price of Bitcoin using Machine Learning, viewed 8 May 2019, <http://trap.ncirl.ie/2496/1/seanmcnally.pdf>

Miraz, MH & Ali, M 2018, 'Applications of blockchain technology beyond cryptocurrency', Annals of Emerging Technologies in Computing (AETiC), vol. 2, no. 1, pp. 1-6.

Misnik, S, Krutalevich, S, Prakapenka, P, Borovykh, P & Vasiliev M n.d., Neural network approximation precision change analysis on cryptocurrency price prediction, viewed 8 May 2019, < http://ceur-ws.org/Vol-2258/paper13.pdf>.

Napiah, N 2018, Prediction cryptocurrencies using hybridization machine learning, viewed 8 May 2019, <https://www.researchgate.net/publication/327741155_Prediction_Cryptocurrencies_using_Hybridization_Machine_Learning>.

Narayanan, A, Bonneau, J, Felten, E, Miller, E & Goldfeder, S 2016, Bitcoin and cryptocurrency technologies: comprehensive introduction, Princeton University Press, Princeton.

Pelletier, C 2018, Computational approaches to predicting cryptocurrency prices, viewed 8 May 2019, <https://digital.kenyon.edu/dh_iphs_ai/8>.

Peters, G, Panayi, E & Chapelle, A 2015, 'Trends in crypto-currencies and blockchain technologies: A monetary theory and regulation perspective', Journal of Financial Perspectives, vol. 3, no. 3, pp. 1-25.

Phillips, RC Gorse, D 2017, 'Predicting cryptocurrency price bubbles using social media data and epidemic modelling', 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, USA, pp. 1-7.

Price, T 2017, Predictive cryptocurrency mining and staking, viewed 8 May 2019, <http://www.tdcommons.org/dpubs_series/889>.

Radityo, A, Munajat, O, Budi, I 2017, 'Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods', 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Bali, Indonesia, pp. 1-5.
Salviotti, G, de Rossi, LM & Abbatemarco, N 2018, 'A structured framework to assess the business application landscape of blockchain technologies', Distributed Ledger Technology, the Blockchain, pp. 3467-3476.
Thierer, A & Castillo, A 2016, Preparing for the future of artificial intelligence, viewed 8 May 2019, <https://www.mercatus.org/system/files/Thierer-Artificial-Intelligence-Permissionless-Innovation-PIC-v1.pdf>.
Wallach, W 2010 'Robot minds and human ethics: the need for a comprehensive model of moral decision making', Ethics and Information Technology, vol. 12, no. 3, pp. 243-250.
Yamada, Y & Nakajima, T 2016 'Blockchain-LI: a study on implementing activity-based micro-pricing using cryptocurrency technologies', MoMM '16 Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media, Singapore, Singapore pp. 203-207.
Yilmaz, NK & Hazar, HB 2018a, 'Determining the factors affecting investors decision-making process in cryptocurrency investments', Istanbul Finance Conference, Istanbul, Turkey, pp. 5-8.
Yilmaz, NK & Hazar, HB 2018b, 'Predicting future cryptocurrency investment trends by conjoint analysis', Finance and Accounting (JEFA), vol. 5, no. 4, pp.321-330
Zheng, X, Zhu, M, Li, Q, Chen, C & Tan, Y 2018 'FinBrain: when finance meets AI 2.0.', Frontiers of Information Technology & Electronic Engineering, pp. 1-11.
Guttag, John V. (12 August 2016). Introduction to Computation and Programming Using Python: With Application to Understanding Data. MIT Press. ISBN 978-0-262-52962-4.
Peter Zaitsev (10 April 2008). "TPC-H Run on MySQL 5.1 and 6.0". MySQL Performance Blog. Retrieved 8 June 2009.
Hyndman, Rob J.; Koehler, Anne B. (2006). "Another look at measures of forecast accuracy". International Journal of Forecasting. 22 (4): 679–688. CiteSeerX 10.1.1.154.9771. doi:10.1016/j.ijforecast.2006.03.001.
Statistical Analysis, Ya-lun Chou, Holt International, 1975, ISBN 0-03-089422-0, section 17.9.
Al-Saqaf, Walid, and Nicolas Seidler. "Blockchain Technology for Social Impact: Opportunities and Challenges Ahead." Journal of Cyber Policy. 2.3 (2017): 338-354.

Berentsen, A, and F Schar. "A Short Introduction to the World of Cryptocurrencies." Federal Reserve Bank of St. Louis Review. 100.1 (2018): 1-16.

Giudici, Giancarlo, Alistair Milne, and Dmitri Vinogradov. "Cryptocurrencies: Market Analysis and Perspectives." Journal of Industrial and Business Economics: Economia E Politica Industriale. 47.1 (2019): 1-18.

Ghosh, Jaideep. "The Blockchain: Opportunities for Research in Information Systems and Information Technology." Journal of Global Information Technology Management. 22.4 (2019): 235-242.

Houben, Robby, and Alexander Snyers. Cryptocurrencies and Blockchain: Legal Context and Implications for Financial Crime, Money Laundering, and Tax Evasion. Brussels: European Parliament, 2018.

Kyriazis, Nikolaos A, Kalliopi Daskalou, Marios Arampatzis, Paraskevi Prassa, and Evangelia Papaioannou. "Estimating the Volatility of Cryptocurrencies During Bearish Markets by Employing Garch Models." Heliyon. 5.8 (2019): 1-8

Leible, Stephan, Steffen Schlager, Moritz Schubotz, and Bela Gipp. "A Review on Blockchain Technology and Blockchain Projects Fostering Open Science." Frontiers in Blockchain. 2 (16): 1-28

Xu, Min, Xingtong Chen, and Gang Kou. "A Systematic Review of Blockchain." Financial Innovation. 5.1 (2019): 1-14.