Bitcoin Price-Predictive Model Based on Twitter News

by Rami El-Youssef

B.S. in Computer Engineering, State Polytechnic University of Kharkov, Ukraine (1995)
M.S. in Computer Engineering, State Polytechnic University of Kharkov, Ukraine (1996)
M.S. in Computer Science & Engineering, Southern Methodist University, Dallas (2007)

A Praxis submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Engineering

Aug 31, 2022

Praxis directed by

Shahryar Sarkani
Adjunct Professor of Engineering Management and Systems Engineering

John Fossaceca
Professorial Lecturer of Engineering Management and Systems Engineering

The School of Engineering and Applied Science of The George Washington University certifies that Rami El-Youssef has passed the Final Examination for the degree of Doctor of Engineering as of August 5, 2022. This is the final and approved form of the Praxis.

Bitcoin Price-Predictive Model Based on Twitter News

Rami El-Youssef

Praxis Research Committee:

Shahryar Sarkani, Adjunct Professor of Engineering Management and Systems Engineering, Praxis Co-Director

John Fossaceca, Professorial Lecturer of Engineering Management and Systems Engineering, Praxis Co-Director

Joseph P. Blackford, Professorial Lecturer of Engineering Management and Systems Engineering, Committee Member

© Copyright 2022 by Rami El-Youssef
All rights reserved

Dedication

I'd like to dedicate my work to my wife Noor, and my children Sara, Ahmad and Youssef who have motivated and encouraged me along the way. Without their encouragement and caring, this praxis would not have materialized.

I will also forever be indebted to my beloved parents Otra and Sobhi for instilling in me the love for knowledge and continuous education. They have provided me unfailing support and encouragement throughout my life.

Acknowledgments

First and foremost, I would like to praise and thank God, the most gracious, most merciful, who has bestowed his countless blessing and opportunity to start this journey and see it to completion.

The success of this journey depends on the guidance and support of many. I take this opportunity to express my gratitude to those who contributed to the successful completion of this Praxis.

At the top of the list, I wish to thank my doctorate advisors, Dr. Sarkani and Dr. Fossaceca at The George Washington University who have mentored and provided valuable input throughout the journey. They generously spent precious time giving guidance, and suggestions for this Praxis. Their continuous feedback and coaching have enabled me to enhance and better the results of the praxis and I cannot say thank you enough for their support and encouragement.

I would like to gratefully acknowledge my friends and colleagues who have helped along the way. It would be impossible to list all the names, but the following people deserve sincere and special thanks: Dr. Hammad Azzam, Dr. Ehap Sabri, Dr. Raed Sbeit, Dr. Mohammed Qaudeer, and Dr. Ayman Farran.

Abstract of Praxis

Bitcoin Price-Predictive Model Based on Twitter News

Bitcoin is a volatile investment and prices are hard to predict, with investors losses estimated at 70% of their portfolio from the market peak of Nov. 2021. This volatility in the price presents a great opportunity for investors. While there are various techniques to predict the Bitcoin price, sentiment analysis ranks highest in terms of prediction accuracy. Specifically, Twitter sentiment has shown to be valuable in predicting the increase or decrease in Bitcoin's price.

This praxis uses news tweeted on the Twitter social media platform to predict the direction of Bitcoin price for the next day. The model developed aims at improving the accuracy of the prediction of the direction of Bitcoin price by implementing advanced machine learning techniques as well as state of the art algorithms for sentiment analysis. In addition, to the sentiments, the model also incorporates reaction to the news by factoring in the number of likes and retweets to further improve the prediction accuracy.

The results obtained not only emphasize the importance of news in predicting the direction of Bitcoin price, but also, sets a new benchmark for predicting the direction of Bitcoin next day price achieving a 65.2% accuracy of prediction.

The solution presented has many practical and useful applications and can be expanded to other areas that can benefit from the application of social media sentiment analysis to predict future asset prices (ex. real estate, stocks, etc.)

Table of Contents

Page

Dedication ..... **iv**

Acknowledgments ..... **v**

Abstract of Praxis ..... **vi**

List of Figures ..... **xi**

List of Tables ..... **xiii**

List of Acronyms ..... **xiv**

CHAPTER 1: INTRODUCTION **1**

1.1. Background **1**

1.2. Research Motivation **2**

1.3. Problem Statement **3**

1.4. Thesis Statement **3**

1.5. Research Objectives **4**

1.6. Research Questions and Hypotheses **4**

1.6.1. Research Questions **4**

1.6.2. Research Hypotheses **5**

1.7. Scope of Research **5**

1.8. Research Limitations **6**

1.9. Organization of Praxis **7**

CHAPTER 2: LITERATURE REVIEW **8**

2.1. Introduction **8**

2.2. Review of Relevant Literature 8

2.3. The Rise of Bitcoin 15

2.4. Sentiment Analysis Techniques 17

2.5. BERT 20

2.6. ROBERTA 22

2.7. FinBERT 23

2.8. Social Media and Sentiment Analysis 25

2.9. News vs. Public Opinion 26

2.10. Twitter as a Source of Information about Bitcoin 26

2.11. Summary and Conclusion 27

CHAPTER 3: METHODOLOGY 31

3.1. Introduction 31

3.2. Collect Twitter and Bitcoin Data for 2020-2021 32

3.3. Classify Tweets Sentiments 36

3.4. Calculate Daily Sentiment Score 40

3.4.1. Polarity-Based Daily Sentiment Score 42

3.4.2. Strength-based Daily Sentiment Score 45

3.5. Construct Research Dataset 47

3.6. Identify Features 50

3.7. Develop ML Models for Price Prediction 52

3.8. Compare Results 53

CHAPTER 4: RESULTS 54

4.1. Introduction 54

4.2. Performance Metrics 55

4.2.1. Relative Variable Importance Chart 55

4.2.2. Accuracy 56

4.2.3. Confusion Matrix 57

4.3. Distribution of Bitcoin Next Day Price Direction 58

4.4. Validating Hypothesis 1 59

4.4.1. Binary Cart Classification 60

4.4.2. Discussion of the Results of Binary CART Classification 67

4.4.3. Correlation Attribute Evaluation 68

4.4.4. Discussion of the Results of Correlation Attribute Evaluation 73

4.4.5. PCA Analysis 75

4.4.6. Discussion of the Results of PCA Analysis 75

4.5. Preparing the Dataset for Machine Learning Runs 76

4.6. Validating Hypothesis 2 76

4.6.1. Naïve Bayes 77

4.6.2. Decision Stump 78

4.7. Validating Hypothesis 3 79

4.7.1. Ensemble Techniques 79

4.7.2. Stacking 80

4.7.3. AdaBoost 81

4.8. Summary 84

CHAPTER 5: DISCUSSION AND CONCLUSION 85

5.1. Discussion 85

5.2. Conclusion 86

5.3. Contribution to the Body of Knowledge 87

5.4. Recommendation for Future Work 88

References 90

Appendix A 113

APPENDIX DATA REPOSITORY 114

List of Figures

Page

1.1. Bitcoin Growth....2

2.1. Bitcoin Transaction Lifecycle....17

2.2. BERT Training and Fine-Tuning....21

2.3. BERT Pre-Trained Models, Data Sources, and Fine Tuning....24

2.4. From BERT to FinBERT....25

3.1. End to End High Level Block Diagram....32

3.2. Collect Twitter and Bitcoin Data Block Diagram....33

3.3. Bitcoin Historical Market Price....34

3.4. Classify Tweets Sentiments Block Diagram....37

3.5. Calculate Daily Sentiment Score Block Diagram....41

3.6. Construct Research Dataset Block Diagram....47

3.7. Identify Features Block Diagram....50

3.8. Develop ML Models for Price Prediction Block Diagram....52

3.9. Compare Results Block Diagram....53

4.1. Confusion Matrix....57

4.2. Bitcoin Daily Price Distribution....58

4.3. BERT- Polarity-Based Daily Sentiment Score Confusion Matrix....62

4.4. BERT- Polarity-Based Daily Sentiment Score Pareto Chart....62

4.5. BERT- Strength-Based Daily Sentiment Score Confusion Matrix....63

4.6. BERT- *Strength-Based Daily Sentiment Score Pareto Chart*....64

4.7. FinBERT- Polarity-Based Daily Sentiment Score Confusion Matrix (Minitab)...65

4.8. FinBERT- Polarity-Based Daily Sentiment Score Confusion Matrix 65

4.9. FinBERT- Polarity-Based Daily Sentiment Score Pareto Chart from Minitab 65

4.10. FinBERT- Strength-Based Daily Sentiment Score Confusion Matrix 66

4.11. FinBERT- Strength-Based Daily Sentiment Score Pareto Chart.....67

4.12. BERT- Polarity-Based Daily Sentiment Score Attribute Selection.....69

4.13. BERT- Strength-Based Daily Sentiment Score Attribute Selection.....70

4.14. FinBERT- Polarity-Based Daily Sentiment Score Attribute Selection 71

4.15. FinBERT- Strength-Based Daily Sentiment Score Attribute Selection 73

4.16. PCA Analysis for FinBERT-Polarity-Based Daily Sentiment Score 75

4.17. Naive Bayes ML Run Results.....77

4.18. Decision Stump ML Run Results 78

4.19. Stacking ML Run Results.....81

4.20. AdaBoost ML Run Results.....82

List of Tables

Page

2.1. Papers Reviewed, with Prediction Period and Prediction Method 10

2.2. Sentiment Analysis Categories and Algorithms 18

2.3. Bitcoin Price Predictions Accuracy Reported in Published Journals 30

3.1. Sample Price and Trading Volume from Blockchain.com 34

3.2. Sample Key Fields from Hashtag.com 36

3.3. Sample BERT Output 38

3.4. FinBERT Financial News Sample 39

3.5. Performance of Different BERT and FinBERT Models.....39

3.6. FinBERT Sample Output.....40

3.7. Example 1 of Calculating Polarity-Based Daily Sentiment Score.....43

3.8. Example 2 of Calculating Polarity-Based Daily Sentiment Score.....44

3.9. Example of Calculating Strength-Based Daily Sentiment Score.....46

3.10. Praxis Dataset Sample.....49

4.1. Praxis Dataset Sample.....54

4.2. Results of Binary CART Classification.....68

4.3. Results of Correlation Attribute Evaluation 74

List of Acronyms

BTC	Bitcoin
BERT	Bidirectional encoder representations from transformers
FinBERT	BERT for financial text mining
CART	Classification and regression trees
NLP	Natural language processing
ML	Machine learning
PCA	Principal component analysis
SVM	Support vector machines

CHAPTER 1:

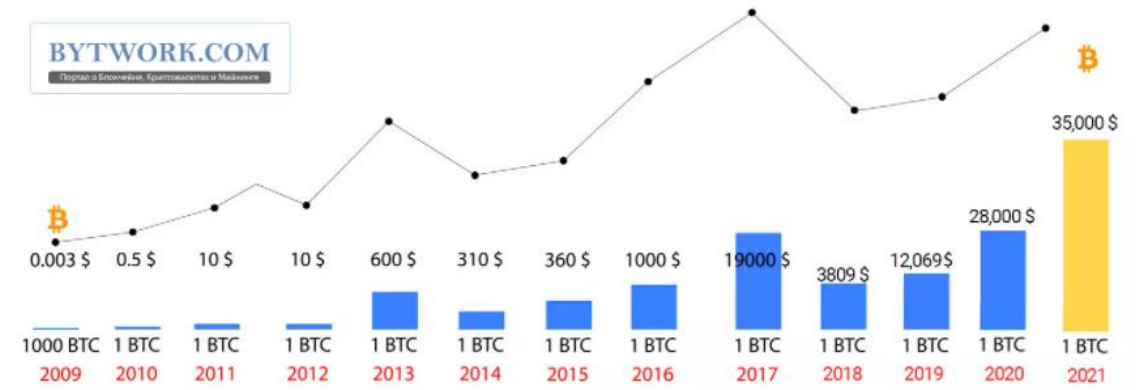
INTRODUCTION

1.1. Background

Cryptocurrency is money that is digital, secure, and anonymous. According to (Olena Bondarenko, 2019, p. 10), “Unlike electronic money, Cryptocurrency is created and produced via the Internet and does not need to be deposited into an account first”. Bitcoin is the leading cryptocurrency, accounting for more than 60% of the total digital currency market capitalization. Per (Ziaul Haque Munim, 2019, p. 2), “Due to the progressive price change and increase in the market capitalization, the popularity of investment in Bitcoin has been increasing steadily”. Furthermore, (P.Dwyer, 2015) explains that “The constant fluctuation in Bitcoin price has captured the attention of academic researchers”.

In the year 2020 alone, the market capitalization for the crypto market increased by 300%. According to (Smutny, Sulc, & Lansky, 2021, p. 2), “In January 2021, the capitalization of the cryptocurrency market exceeded USD 1 trillion, and in April 2021, it reached to USD 2 trillion”. (Umar, Su, Abbas Rizvi, & Shao, 2021) explains that “This increase is mainly driven by the economic crisis, accelerated by the global COVID-19 pandemic and increased interest from institutional investors”. Along with its capitalization growth, the Bitcoin price has also increased substantially for the years 2009-2021 as shown in Figure 1.1.

Figure 1.1
Bitcoin Growth



Note: From (BYTWORK.COM, 2022).

This research proposes a model to accurately predict the direction of Bitcoin next day price. It does so by applying sentiment analysis to news obtained from twitter financial and business news providers to calculate a daily sentiment score for predicting the Bitcoin next day price.

1.2. Research Motivation

With social media transforming the way that people communicate in the 21st century, there has been more interest in using social media platforms as a tool to approximate the sentiments of individuals, organizations, and society at large. Of the many social media platforms available today, Twitter has surfaced as the primary tool for performing sentiment analysis.

While there have been many attempts to predict bitcoin price, accurate prediction is still a challenge given the volatile nature of the cryptocurrency. Studies show that Bitcoin prices are driven by trust and sentiments and those in turn are largely shaped by opinions expressed in news. Using sentiment analysis is undoubtedly one of the best

techniques known today to predicting Bitcoin price. This praxis introduces a new predictive model for the direction of Bitcoin next day price using tweets from major financial and business news providers. The ultimate goal is to improve the accuracy of the prediction of the direction of Bitcoin next day price. The ability to capture the number of tweets' likes and retweets provides an advantage over existing research as it acts as a proxy for factoring in users' reaction to the news. In addition, the use of the latest technologies in language modeling (BERT and FinBERT) will further improve the accuracy of the analysis.

1.3. Problem Statement

Bitcoin is a volatile investment and prices are hard to predict, with investors losses estimated at 70% of their portfolio from the market peak of Nov. 2021.

The difficulty of predicting Bitcoin prices is understandable given the extreme volatility and fluctuations in the prices. As a result of this volatility, there is huge potential benefits to accurately predicting the Bitcoin prices. Worth noting that by Aug 2021, crypto market capitalization reached $2 Trillion according to Bloomberg and a $100 investment in Bitcoin in October 2010 would yield $48 million in 2021 (Reinicke, 2021).

1.4. Thesis Statement

A new predictive model using Machine Learning techniques with Twitter financial news sentiments improves the direction of the prediction of Bitcoin next day price.

While there are many social media platforms, Twitter was chosen as the platform of choice due to its succinctness. Per (CityFALCON, n.d.) ,"Due to the 280 characters

limit on Twitter posts, when a story is breaking, trusted and popular news sources can make fast, visible updates to the news story even if it isn't a fully developed story yet. In finance, the sooner important information is incorporated, the faster the decision-making can occur, and accordingly the more accurate the price targeted prediction".

1.5. Research Objectives

The primary objectives of this research are to improve the accuracy of prediction for the directional change in Bitcoin price for the next day by demonstrating a novel predictive model that uses sentiment analysis and incorporates Twitter financial and business news related to Bitcoin. Detailed objectives are as follows:

1. To identify the Twitter news features that are significant for Bitcoin next day price direction prediction.
2. To assess the ability of sentiment analysis applied to Twitter business and financial news in improving the prediction accuracy of the direction of Bitcoin price.
3. To develop a predictive model using ML models and sentiment analysis techniques applied to Twitter business and financial news to improve the prediction accuracy of the direction of Bitcoin price.

1.6. Research Questions and Hypotheses

The following research questions and hypotheses are the focus of this praxis.

1.6.1. Research Questions

RQ1: What Twitter features are most significant for accurately predicting the direction of Bitcoin next day price?

RQ2: Can a predictive model using Twitter features (tweets sentiments, # likes, #retweets) improve the prediction of the direction of Bitcoin next day price?

RQ3: Which ML techniques using Twitter features (tweets sentiments, # likes, #retweets) will improve the prediction of the direction of Bitcoin next day price?

1.6.2. Research Hypotheses

H1: Twitter features (tweets sentiments, # likes, #retweets) are significant for accurately predicting the direction of Bitcoin next day price.

H2: A predictive model using Twitter features (tweets sentiments, # likes, #retweets) will accurately predict the direction of Bitcoin next day price.

H3: An ensemble of machine learning techniques using Twitter features (tweets sentiments, # likes, #retweets) will improve the prediction of the direction of Bitcoin next day price.

1.7. Scope of Research

The scope of this research is to explore the feasibility of a predictive model based on Twitter sentiment analysis to predict the direction of Bitcoin next day price. This consists of capturing the financial and business news from Twitter and analyzing the predictive power of the daily sentiment score that is constructed using the tweets polarity.

The new proposed approach uses the latest language modeling techniques to classify the tweets and further builds a daily sentiment score by aggregating the tweets sentiment polarity daily.

The input data consists of 17761 tweets selected from major financial and business news providers on Twitter as well as the daily Bitcoin price and trading volume

collected from Blockchain.com. The data from Twitter is used to construct the daily sentiment score which is further merged with the Bitcoin price details to build the dataset that is used in this research.

1.8. Research Limitations

In this research, while sentiment analysis was used to construct the daily sentiment score which constitutes the primary predictor for the Bitcoin next day price direction, it is limited by the following factors:

- Not all financial and business news providers were selected. While the major providers were included in this research, this does not encompass all news providers in the market.
- The research is also limited to English news providers. Given the global nature of the Bitcoin market, the prediction could be improved by including non-English news providers as well.
- Twitter news sentiment analysis is used as a proxy for the larger public sentiment towards Bitcoin. While a reasonable assumption, one could argue that there might be some variation between sentiments expressed on Twitter from major business and financial news providers and the sentiments expressed by the public at large.
- Some of the news might be more of a reflection of the Bitcoin price after the fact as opposed to being a driver for the next day price. In other words, some component of the news is reporting on the current state of the price as opposed to the potential of a future change in the price.

1.9. Organization of Praxis

This Praxis consists of five chapters, as follows:

Chapter 1 begins with the background and the research motivation, then outlines the research objectives, research questions, and hypotheses. It ends with the scope and the limitations of this research.

Chapter 2 covers the literature review, starting with an introduction, a review of relevant literature, the rise of Bitcoin, sentiment analysis techniques that are available today, the importance of Twitter in sentiment analysis and the use of news as a source of information. The chapter ends with an explanation of BERT and FinBERT, the two sentiment analysis techniques used in this research.

Chapter 3 covers the methodology used in this research. It starts by providing a high-level overview of the methodology used, followed by a detailed description of each of the steps used in the methodology.

Chapter 4 presents the results of the praxis, followed by a discussion of those results and validation of the three hypotheses suggested in this praxis.

Chapter 5 summarizes the paper's main points and presents a discission and concluding thoughts. It highlights the potential contributions of this work to the body of knowledge and suggests recommendations for future work.

CHAPTER 2:

LITERATURE REVIEW

2.1. Introduction

Market price prediction is a challenging task because of erratic patterns and uncertainty of the market. This has motivated researchers to continuously think of better predictive methods. According to (Durgapal & Vimal, 2021), "Many studies have been conducted and are well documented in the literature, with many models to predict the market price accurately". When it comes to Bitcoin, the challenge is even bigger. Bitcoin prices are characterized by large number of fluctuations and sharp increases/decreases in price. This makes predicting Bitcoin price a fertile topic for many researchers and interested parties.

This chapter offers a comprehensive literature review of the topics related to Bitcoin cryptocurrency, Bitcoin price prediction models, sentiment analysis algorithms, and Twitter social media platform. The purpose of the review is to provide a summary of the research that has been published and to further examine the existing body of technical knowledge.

2.2. Review of Relevant Literature

In lieu of the high volatility and potential for huge amounts of gains for investors, it is no surprise that the topic of Bitcoin price prediction has been of great interest for many researchers especially in the last few years. To the best of my knowledge, this is the first research work that utilizes posts(tweets) from major financial and business news providers communicated through twitter to construct a sentiment score for the purpose of

predicting the direction of the Bitcoin next day price. In this review, we will analyze some of the articles and research work that has stood out and summarize the level of contribution to the body of knowledge in general and to the development of this praxis. Scanning and analyzing the different literature related to this topic, the overall literature can be categorized by:

1. Prediction period: next day price direction prediction vs. same day price direction prediction. Most researchers have focused on predicting Bitcoin next day price direction, while others tried to predict same day Bitcoin price direction and in some cases near real time price direction.
2. Prediction method: sentiment-based vs. non-sentiment-based. Some of the research work attempts to model the Bitcoin price prediction using the Blockchain network characteristics and historical price, while others use social media and sentiment analysis techniques to predict the Bitcoin price. It's worth noting that majority of the research work uses public sentiments while few rely on sentiments expressed in the news. For public sentiments, researchers mostly use social media platforms like Twitter and Reddit with Twitter being the preferred proxy for the overall public sentiment, while news sentiments are mostly gathered by directly crawling the news providers websites or using aggregator websites like google news.

Table 2.1 shows a summary of the papers reviewed categorized according to the definition above:

Table 2.1
Papers Reviewed, with Prediction Period and Prediction Method

Paper	Prediction Period	Prediction Method
(Greaves & Au, 2015)	Next day price direction	Blockchain network-based features
(Karalevičius, Degrande, & De Weerdt, 2017)	Same day price direction	Bitcoin-related news portals
(Lamon, Nielsen, & Redondo, 2017)	Next day price direction	Daily News Data
(McNall, Roch, & Caton, 2018)	Next day price direction	historical price data, mining difficulty and hash rate
(Abraham & Higdon, 2018)	Next day price direction	Twitter sentiment and google trends data
(Yao, Xu, & Li, 29 June-3 July 2019)	Next day price direction	News articles by crawling websites
(Luo, 2020)	Next day price direction	Covid-19 data and tweets sentiments
(Jeon, Oh, & Lee, 2020)	Next day price direction	Twitter posts
(Kumar, 2021)	Next day price direction	News content
(Gurrib & Kamalov , 2021)	Next day price direction	Bitcoin price information and Bitcoin news announcements headlines
(Vella Critien, Gatt2, & Ellul, 2022)	Bitcoin next day, next 3 days and next 7 days direction	Sentiment and volume of tweets

Below is a discussion of the research work reviewed in the writing of this praxis presented in a chronological order.

(Greaves & Au, 2015) examined the network-based features of Blockchain to predict the direction of Bitcoin price using support vector machines (SVM) and artificial neural networks (ANN). They achieve 55% accuracy of classification for the Bitcoin price up-down movement, concluding that Blockchain data has limited predictability. The low accuracy results are not surprising. While the network-based features might indicate a rise or fall of Bitcoin on the long run (more network activity implies more adoption), it's unlikely that these characteristics have an immediate impact on day-to-day price of Bitcoin.

(Karalevičius, Degrande, & De Weerdt, 2017) use a lexicon-based approach incorporating generic and finance-specific dictionaries and applying it to various Bitcoin-related news portals to predict best strategy for interday trading. They confirm that "interaction between media sentiment and the Bitcoin price exists and that there is a tendency for investors to overreact on news in a short period of time". This is one of the first papers to analyze the relationship between news and Bitcoin price and in a way inspired other researchers to consider news as a potential predictor for the Bitcoin price. Nevertheless, the paper focuses on interday trading which carries less value than predicting the direction of the next day price.

(Lamon, Nielsen, & Redondo, 2017) analyze the ability of news and social media data to predict fluctuations in price for three major cryptocurrencies including Bitcoin. This paper proposes an innovative approach to predicting the Bitcoin next day price direction. Instead of labeling daily news data as positive or negative sentiment, the authors use the actual price changes one day in the future for each coin. The model developed was able to predict fluctuations in price instead of having to predict sentiment first. Logistic regression achieved a prediction accuracy of 43.9% of price increases and 61.9% of price decreases. While an innovative approach to predicting the Bitcoin price, the accuracy achieved is still low compared to results documented in other research papers.

(McNall, Roch, & Caton, 2018) uses closing price, opening price, Blockchain mining difficulty and hash rate to predict the accuracy of the direction of Bitcoin price. After applying different deep learning models like RNN and LSTM, the highest accuracy of prediction achieved is 52.78%. This is a very low accuracy and is another indication

that using historical price and network-based characteristics are not ideal to predict the direction of Bitcoin price.

(Abraham & Higdon, 2018) use Twitter sentiment and google trends data to examine the correlation with Bitcoin and Ethereum price. The authors conclude that while there is a correlation between volume and price, sentiment analysis is less useful in its predictive capabilities of cryptocurrencies. The authors attribute the weak correlation between sentiments and Bitcoin price to the fact that people who tweet about cryptocurrencies view them more than just an investment opportunity thus making the tweets positively skewed. This is in-line with other research work published. While public sentiments could be a good predictor for other asset prices (ex. stock, real estate, etc.), public sentiments are not necessarily a good predictor for Bitcoin price. Most of the tweets about Bitcoin tend to be positive irrespective of the cryptocurrency performance.

(Yao, Xu, & Li, 29 June-3 July 2019) use machine learning to explore the influence of News articles on the same day fluctuation of Bitcoin price. News collected by crawling different websites and text mining techniques are applied to extract relevant information about Bitcoin. Different machine learning techniques are used: SVM, Random Forest, Multilayer Perceptron, Logistic Regression. Authors achieve a prediction accuracy of 59% and conclude that news articles do impact the price of bitcoin. This paper serves as a good basis to consider news sentiments as a predictor for the Bitcoin price, however, fails to explore using more advanced sentiment analysis techniques like BERT to further improve the accuracy of the prediction. Furthermore, paper is focused on same day price which carries less value than predicting next day price. Finally, adding

other predictors to approximate the users' reaction to the news might help improve the prediction accuracy.

(Luo, 2020) investigates the relationship between Covid-19 data, tweets sentiments and Bitcoin prices and analyzes the performance of different machine learning algorithms on predicting the return rate and price trend of Bitcoin. Paper concludes that "Twitter data does help in predicting Bitcoin price, but COVID-19 data does not improve the accuracy". Paper is beneficial for those who are exploring the importance of using social media(twitter) to predict Bitcoin prices and serves as a good basis to evaluating different ML techniques for predicting Bitcoin prices. However, the dataset used is limited only to 6 months, so there is not enough data to train the model especially for the purpose of using deep learning techniques.

(Jeon, Oh, & Lee, 2020) attempt to "find a partial correlation between the price fluctuation of Bitcoin and the fluctuation of the sentiment classes using a machine learning Algorithm". VADER technique is used to perform sentiment analysis on Twitter posts and the paper confirms the presence of a correlation between twitter sentiments and Bitcoin returns, observing a 62.48% accuracy of predictions. Unlike many of the research work that tries to predict Bitcoin next day price direction, this paper uses individual tweets predictive power for near real time price prediction. The results are based on just two months' worth of data, which is not representative. Nevertheless, the paper acts as a basis to explore other features that could be used to improve the prediction accuracy such as the number of tweets.

(Kumar, 2021) analyzes the ability of news content to predict cryptocurrency markets direction. Data is gathered from online Crypto blogs specifically coindesk.com

and cointelegraph.com. The author uses different ML algorithms (Logistic Regression, SVC, Random Forest) and achieves a 56% accuracy of prediction. While the research confirms the ability of news to predict the direction of Bitcoin price, the low accuracy of prediction can be attributed to failure in capturing negative sentiments because data is gathered from websites that are heavily vested in cryptocurrency (coindesk.com and cointelegraph.com). It's likely that collecting news from sources that are more objective in reflecting the true sentiments about Bitcoin in the marketplace would improve the prediction accuracy.

(Gurrib & Kamalov , 2021) uses the current Bitcoin price information and news headlines to predict the direction of Bitcoin next-day price. This is the first paper to explore using LDA for forecasting the Bitcoin next-day direction of Bitcoin prices. Nevertheless, using SVM achieves a higher prediction accuracy of 58.5%.
The authors achieve a relatively high accuracy as compared to other published research; an indication that using News headlines could be good predictor for Bitcoin price direction. The authors use VADER analysis to analyze sentiments, but potentially using more advanced sentiment analysis techniques like BERT and FinBERT and capturing users' reactions to the news could help improve the accuracy further.

(Vella Critien, Gatt2, & Ellul, 2022) use sentiment and volume of tweets to predict Bitcoin price direction and magnitude of the increase/decrease. This is one of the most interesting research papers as it explores predicting the direction of Bitcoin price using a lag interval. Unlike other research work that tries to predict same day or next day price direction, this paper explores predicting the price direction using a one day, 3 days and seven days intervals. It shows that highest accuracy of 60.1% is achieved for one day

lag interval using BiLSTM ML algorithm. The paper confirms that sentiments are most relevant when predicting the direction for the next day price of Bitcoin but are not adequate for predicting the direction of Bitcoin price over longer intervals. The authors achieve the highest accuracy reported from all the research papers examined. However, the authors use VADER analysis which is relatively outdated compared to more recent sentiment analysis techniques like BERT and FinBERT. In addition, the paper uses public sentiments which tend to be positively skewed.

2.3. The Rise of Bitcoin

According to (Manimuthu, Sreedharan, G., & Marwaha, 2019, p. 28), "The biggest thing that hit the global market since the invention of currency is the concept of "digital currency" "

(Shehhi, Oudah, & Aung, 2014) explain that "Money has taken different forms overtime, ranging from commodities to commodities-backed currency to fiat currency and most recently to the internet-based currency which is also referred to as cryptocurrency". Per (Simser, 2015), "Over the last two decades, the Internet has become an increasingly used platform for business-related transactions, which rely heavily on trust provided by financial institutions serving as third parties to process online payments". According to (Manimuthu, Sreedharan, G., & Marwaha, 2019, p. 30), "The disadvantages of the traditional financial institutions such as lack of transparency, high transaction fees and the potential risk of manipulation gave the rise to the idea of a transparent electronic payment system that is trustless and more efficient".

Bitcoin is a cryptocurrency that allows digital transfers of value between individuals and businesses. Per (Reid & Harrigan, 2011), "In simple terms, it is "an

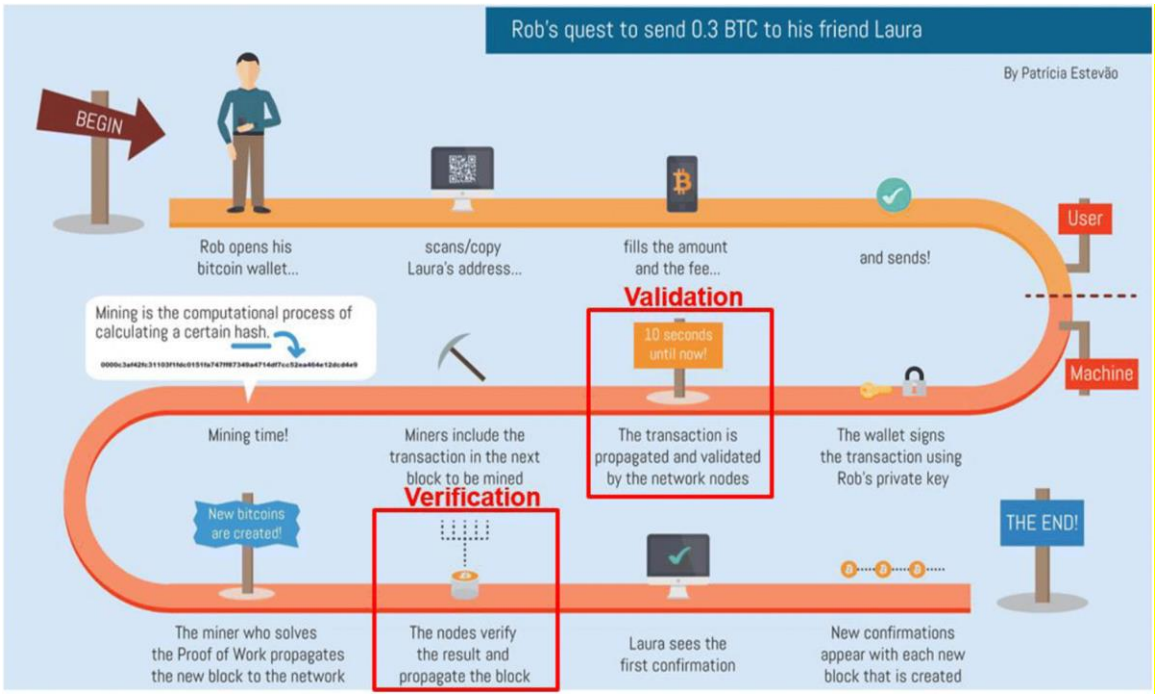
electronic equivalent of cash in the online world". It is decentralized which means that there is no central authority responsible for issuing Bitcoins and there is no need to involve a trusted third-party when making online transfers". Bitcoin was started in 2009 by Satoshi Nakamoto and is slowly moving to become a global digital value reserve (Partida, et al., 2022). Today, Bitcoin is the leading cryptocurrency by capitalization and as such "plays a unique and important role in the worldwide currency market" (Chen , Giudici, Misheva, & Trimborn , 2020). According to research published by Cambridge University, "As of 2017, there are 2.8 to 5.8 million unique Bitcoin users" (Williams, 2017) and as of December 2021, the number of accounts that hold Bitcoin worldwide has grown to an estimated 114 million according to crypto.com (Quiroz-GutIerrez, 2021). (Heston, 2018) estimates "2 billion USD worth of Bitcoin transactions happen every day and more than 300 million transactions have occurred till 2017". According to BitInfoCharts, "as of January 2021, there are 25 individual Bitcoin addresses holding more than a billion dollars" (Phillips, 2021), and an average of 150k bitcoin related tweets are posted on Twitter every day (Tweets Chart, 2022).

Bitcoin is also a growing topic in the literature, and with the amount of volatility it exhibits, there is a good amount of literature that tries to analyze fluctuations in Bitcoin prices and potential returns (Gangjin, Yanping, Chi, & Shou, September 2019). For example, (Dwyer, 2014) determined that "monthly volatility of Bitcoin returns is quite a bit higher than that of gold". (Weber, 2016) takes it a step further and explores "the potential for Bitcoin to replace the gold standard and become, in the future, the global digital reserve currency".

In addition, (Partida, et al., 2022) assert that "a growing number of players belonging to the mainstream financial system have started to accept Bitcoin and add it to their investment portfolio".

The architecture of the Bitcoin platform and the technology that empowers it (Blockchain) is beyond the scope of this research, but a high-level view of the Bitcoin transaction lifecycle is presented in Figure 2.1 (Ankalkoti, 2017).

Figure 2.1
Bitcoin Transaction Lifecycle



Note: From (Ankalkoti, 2017).

2.4. Sentiment Analysis Techniques

This praxis presents an argument that Bitcoin next day price direction is related to previous day news sentiments. To validate the argument, sentiment analysis is used to determine the sentiment of the news and use it as a predictor for the Bitcoin price.

Sentiment analysis, also known as opinion mining, is a technique used to approximate opinions and emotions for the purpose of gaining meaningful insights about products, services, or ideas (Huahu, Ahmad, Shabir, & Ahmad, 2019). Sentiment analysis seeks to assign a feeling or emotion to the text for the purpose of predicting whether the sentiment is positive, negative, or neutral (Goncharov, 2022). Twitter sentiment analysis is applying sentiment analysis techniques to classify the sentiments of tweets. The output of the sentiment classification determines the tweet's polarity which could be "positive", "negative" or "neutral".

According to (Nandwani & Verma, Aug 2021), "sentiment analysis techniques can be broadly classified into four categories: Lexicon-based, machine learning, deep learning, and transfer learning". Table 2.2 summarizes the different categories and algorithms that are used for sentiment analysis.

Table 2.2
Sentiment Analysis Categories and Algorithms

Category	Algorithms	
Lexicon-based	Dictionary-based	Corpus-based
Machine Learning	Naïve Bayes SVM	Random Forest Decision Trees
Deep Learning	CNN LSTM	BI-LSTM
Transfer Learning	BERT FinBERT ROBERTA	XLNet DistilBERT Etc.

Lexicon-based approach: This could be dictionary-based or corpus-based. According to (Nandwani & Verma, Aug 2021), "The dictionary-based approach maintains a dictionary of words in which each positive and negative word is assigned a

sentiment value. The overall sentiment of the document is then calculated as the sum or mean of the sentiment values. The corpus-based approach differs from the lexicon-based approach in that it includes domain-specific words where the polarity is assigned to the word according to its context or domain".

Machine Learning-based approach: With this approach, the dataset in its entirety is divided into a training dataset and a testing dataset. According to (Nandwani & Verma, Aug 2021, p. 10), "this uses supervised learning in which the training dataset is used to train the model while the testing dataset is used to validate the model built during training". The different types of algorithms used are decision trees SVM, naïve bayes, and others.

Deep Learning-based approach: Most recently, deep learning algorithms are becoming more popular in solving business challenges and performing sentiment analysis. According to (Ghorbani, Bahaghighat, Xin, & Özen, 2020, p. 5), "Deep learning (DL) can learn how to perform tasks using multilayer deep networks and enhance the power of learning of neural networks". The differentiating power of deep learning-based approach is its ability to detect text sentiments without performing engineering of the features. Among the many deep learning algorithms, recurrent neural networks (RNNs) and convolutional neural networks (CNNs) have surfaced as the key algorithms with an ability to produce more accurate results than other ML and deep learning models.

Transfer Learning-based approach: (Joshi, 2020, p. 2) defines transfer learning as "A technique where a deep learning model trained on a large dataset is used to perform similar tasks on another dataset". In other words, the learnings from one task can be

transferred to another task, allowing for saving time and producing more accurate results. As a result, according to (Zhuang, et al., 2021, p. 43), "transfer learning has become popular and promising and nowadays is considered the most sophisticated of all the machine learning/deep learning techniques". Transfer learning has been a great success in computer vision related classification, but it wasn't till 2018 when Google introduced the transformer model that this approach gained grounds in NLP and since then has become the most successful in solving many of the natural language processing tasks (Joshi, 2020).

The most popular algorithms for performing transfer learning on textual data are BERT, ROBERTA and FinBERT. These algorithms are described in the next few sections.

2.5. BERT

According to (Devlin, Chang, Lee, & Toutanova, 2019, p. 1), "BERT stands for Bidirectional Encoder Representations from Transformers and is a new language representation model that was first introduced by Google in 2018". BERT has shown tremendous improvement in language modeling for various NLP tasks (Won, Lee, Choi, & Jung, 2021). Before BERT, if we were to perform sentiment analysis, we would have needed large amounts of labeled data to build and train a custom model from scratch.

The challenge of producing labeled data is twofold:

1. Labeled data requires experts' domain knowledge that is not readily available
2. Quality labeled data requires substantial amount of effort to acquire.

But in 2018 the NLP world saw an innovative solution to the above challenges with the introduction of BERT. The idea was simple, use textual data that is available in abundance like Wikipedia and apply transfer learning efficiently for NLP problems. BERT is pretrained to understand language by training it on two unsupervised models:

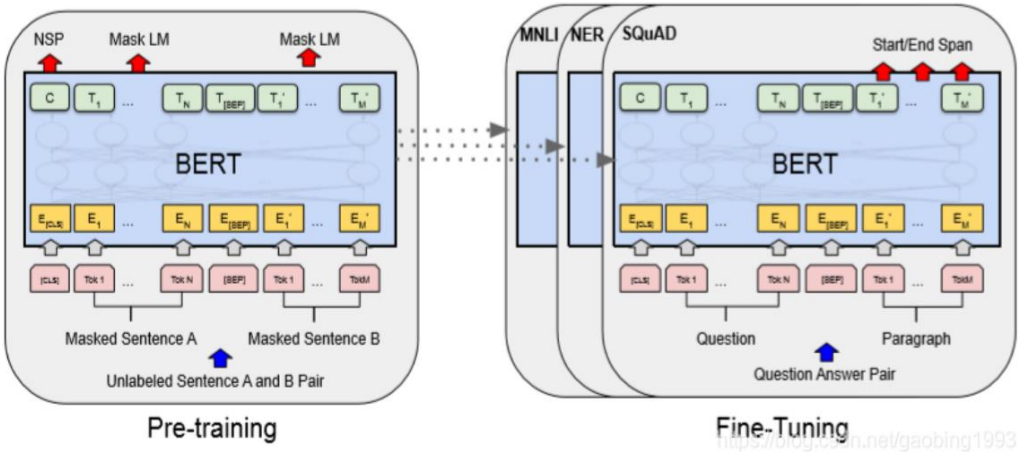
- Masked Language Model (MLM) – fill in the blanks. Context within a sentence.
- Next Sentence Prediction (NSP) – Does this sentence follow the first. Context across sentences (WANG, LEI, & MIAO, 2021).

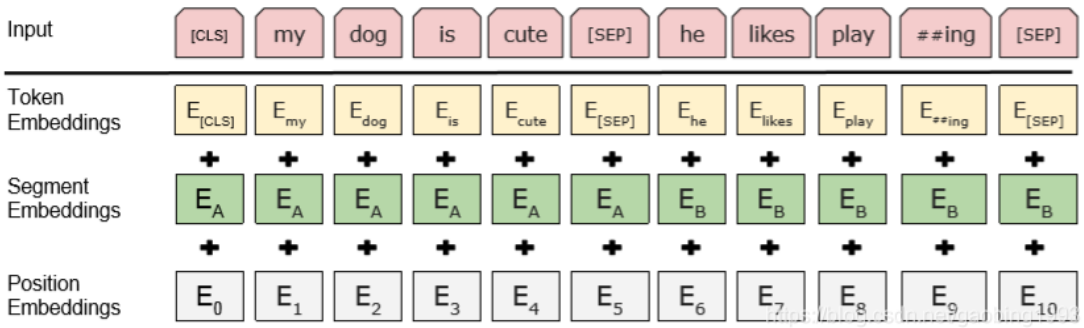
According to (Kana, 2019, p. 2), "The original BERT was trained on Wikipedia (2,500 million words) and Book Corpus, a dataset of more than10,000 books of different genres containing 800 million words".

Per (Guo, 2022), "There are two steps in BERT: pre-training and fine-tuning". According to (Devlin, Chang, Lee, & Toutanova, 2019, p. 3), "As part of pre-training, the model is trained on unlabeled data over different tasks. For fine-tuning, the BERT model is first initialized with the pre-trained parameters, after which they are fine-tuned using labeled data from the downstream tasks". Figure 2.2 (Matwin, Milios, Pralat, & Soares, 2021) shows the main structure used in pre-training and fine-tuning of BERT.

Figure 2.2

BERT Training and Fine-Tuning





Note: From (Matwin, Milios, Pralat, & Soares, 2021).

From the above figure, we can see that the pre-training and fine-tuning steps use the same main structure. (CodeTD, 2020, p. 1) explains that "Initially, each input is a sentence pair of two sentences, where some words are randomly masked, using CLS as the starting mark and SEP as the separation mark between sentences. Then these words pass through the embedding layer in Figure 2.2, comprised of three layers of embedding information: word embedding information, segment embedding information, and position embedding information. The final output is predicting the masked words and whether sentence B is the next sentence that follows sentence A".

The ability to perform transfer learning proved to be a game changer for language modeling and was applied to different aspects of the language: sentiment analysis, question answer, topic modeling etc. Furthermore, the ability to fine tune the model gave birth to thousands of BERT-based models that used the BERT original model as a base that was further trained using domain specific knowledge.

2.6. ROBERTA

RoBERTa is a method that is optimized for pretraining natural language processing (NLP) systems. According to (Indurthi, Syed, Gupta, & Varma, 2020, p.

4840), “RoBERTa is an improvement over BERT as it was trained with much more data – 160 GB of text instead of the 16 GB dataset originally used to train base BERT. Further, it is also trained for larger number of iterations. Finally, compared to BERT, it removes the next sequence prediction objective from the training procedure, and applies a dynamically changing masking pattern to the training data”.

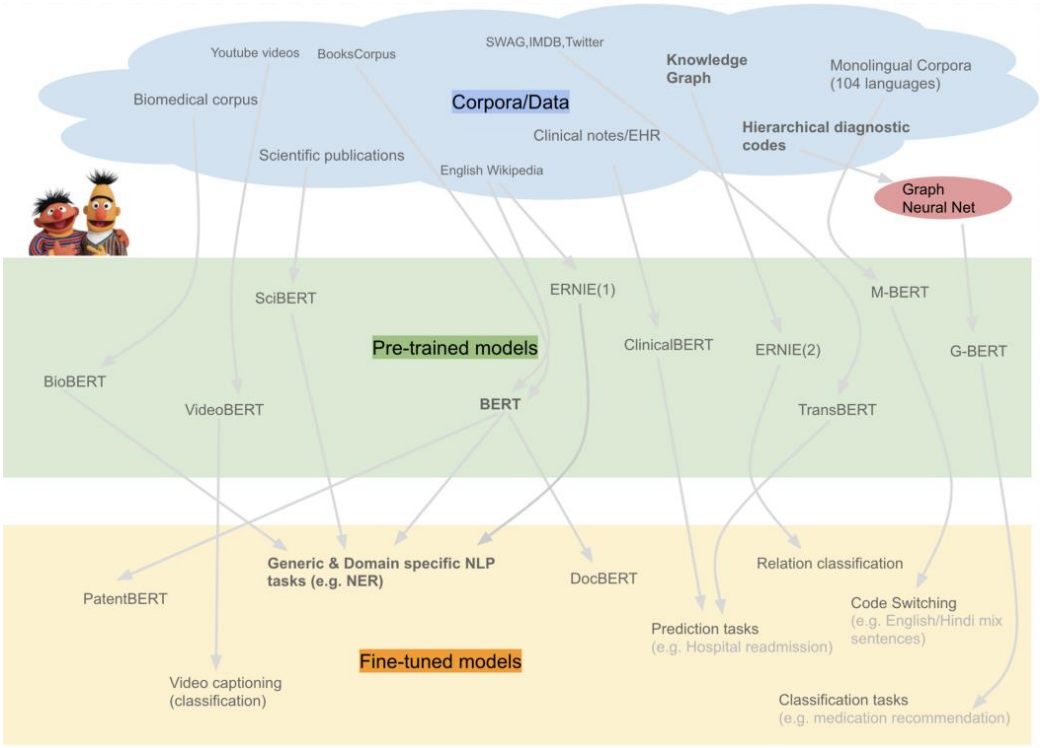
2.7. FinBERT

In many of the finance and economics domains, financial text is used heavily for analysis and prediction of future trends in the market. With the rapid increase in financial text data, manually analyzing these texts to gain actionable insights from them is not feasible anymore. That is especially true for analyst reports or official company announcements where textual analysis is used to gain insights about a company’s performance and future perspectives. With the most recent development in machine learning technology, mining financial text is not only gaining traction, but also becoming the de facto standard for performing textual analysis (Liu, Huang, Huang, Li, & Zhao, 2020).

One of the biggest advantages of BERT is that it can be fine-tuned to the task at hand. According to (Maiya, 2020), “Huggingface which is an immensely popular Python library, providing pretrained models that are extraordinarily useful for a variety of natural language processing (NLP), has more than 29000 models that are pretrained on domain/application specific corpus: example — BioBERT (biomedical text), SciBERT (scientific publications), FinBert(Finance)…”

Figure 2.3 (Rajasekharan, 2019) is a good depiction of the different pre-trained models, the data sources used to train those models and the types of tasks that can be fine-tuned for the different models. (Kana, 2019)

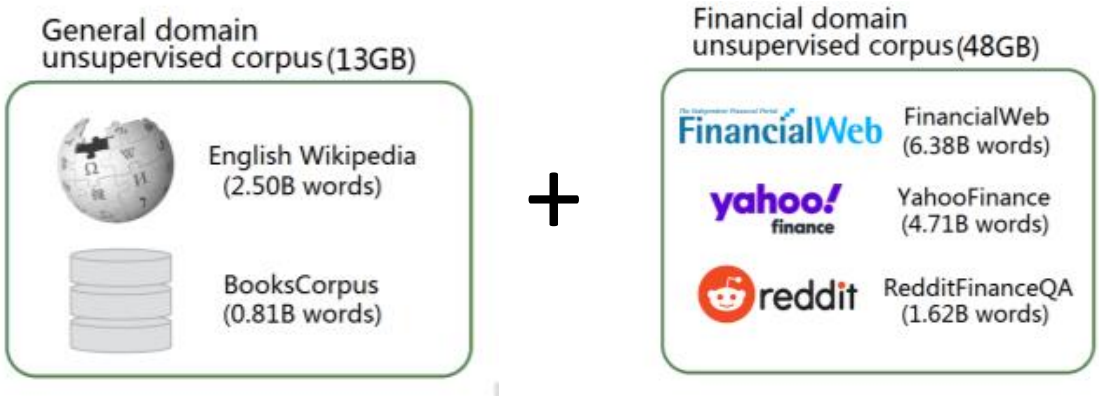
Figure 2.3
BERT Pre-Trained Models, Data Sources, and Fine Tuning



Note: From (Rajasekharan, 2019).

According to (Goncharov, 2022, p. 1), "FinBERT is a pre-trained NLP model to analyze financial text. It is built by further training the BERT language model in the finance domain, using a large financial corpus, and thereby fine-tuning it for financial sentiment classification". Figure 2.4 depicts a view of how FinBERT builds on the general BERT trained model and further fine-tunes it to satisfy the needs of the financial domain (Liu, Huang, Huang, Li, & Zhao, 2020).

Figure 2.4
From BERT to FinBERT



Note: From (Liu, Huang, Huang, Li, & Zhao, FinBERT: A Pre-trained Financial Language Representation Model for Financial Text Mining, 2020).

2.8. Social Media and Sentiment Analysis

In recent years, social media has become not only a platform for people to interact and share ideas, but a de facto proxy for approximating the sentiments towards a topic of interest. The advantage of social media over other classical methods for collating the public opinion is obvious. For one, it's very easy to gather data from social media without having to organize a special event to collect that data (ex. surveys) and secondly, the global and 24/7 presence of these platforms made it possible to collect large amounts of data that otherwise would have been impossible or unfeasible to do with traditional methods. In addition, social media platforms tend for the most part to reflect the raw opinion of users without additional filtering. This latter characteristic is one of the biggest advantages of using social media for sentiment analysis over classical methods like surveys that are difficult to shield from bias and are usually constrained by limited responses.

2.9. News vs. Public Opinion

Most of the research work to predict Bitcoin price relies on public sentiments expressed on social media platforms in general and Twitter in particular. Relying on public sentiments might have been driven by the success in using this approach to predict the prices of other financial assets (ex. stocks).

For Bitcoin, however, relying on public sentiments to predict the price does not yield the best accuracy results since Bitcoin posts tend to be skewed more positively. Bitcoin core supporters tend to post many more messages than those who do not support or necessarily care about it and accordingly would have a positive outlook of the currency irrespective of the reality on the ground. Example, (Qin & Livni, 2021) report that "in July 2021, when China intensified its crackdown on cryptocurrency, declaring all financial transactions involving cryptocurrencies illegal and closing the mining centers", most posts on social media platforms continued to be positive portraying a very different picture of the cryptocurrency market than what is should have been. The same trend was observed when U.S. officials voiced concerns about users having access to cryptocurrency exchanges that are not governed by the same rules of operation as the US.

In comparison, news provides a more objective perspective and hence a more accurate reflection of the Bitcoin sentiments.

2.10. Twitter as a Source of Information about Bitcoin

All social media platforms excel at delivering information to their users, but when it comes to sentiment analysis, Twitter surfaced out as the main platform of choice. While there were attempts to use news to collate sentiments by crawling news websites or using an aggregator like google news, the use of Twitter provides several advantages:

Per (CityFalcon, 2018, pp. 1-2), “Twitter imposes a limit of 280 characters on posts and that gives rise to the doctrine of succinctness in news. This means when a story is breaking, trusted and popular news sources can make fast, visible updates even if it isn’t a fully developed news story yet. In finance, more price action can be captured if important information is incorporated earlier. In addition, the 280 characters limit of the tweets drives Twitter users to post concise and distilled news and that gives more objectivity to the news. Twitter provides a public and suitable platform for this very concept”.

Furthermore, to get the most accurate representation of the Bitcoin related news, the following criteria was used in this research:

1. Publications that have Twitter presence and have large followings.
2. Publications that cover finance and business news.
3. Publications that are not considered biased towards Bitcoin. For example, there are many online publications that are almost always positive about Bitcoin (ex. coindesk.com and coinbase.com which were excluded from this research).
4. Publications that post about Bitcoin news more frequently.
5. Publications that post news in English.

2.11. Summary and Conclusion

This literature review section provided a comprehensive examination and analysis of journal papers, conferences, and articles on Bitcoin Price Prediction. The literature reviewed provides invaluable insights about the different techniques used for predicting the Bitcoin price and the progression over time in these techniques and accordingly improved reported accuracy.

The early studies focused on network-based features to predict the Bitcoin price. This aligns with the findings from (Jang & Lee, 2017) who points out that much of the research work in the early days of Bitcoin focused on the underlying blockchain characteristics such as hash rate, mining difficulty, transaction cost and block height to predict the bitcoin price. The findings from these papers are characterized by low accuracy of prediction compared to more recent techniques. This is understandable, given the limited amount of predictive power embedded in network features themselves. After all, network features could possibly be an indicator of the rise or fall of Bitcoin over longer periods of time, but not an accurate reflection of the state of Bitcoin price over a one-day period. Bitcoin price is largely dictated by exchanges whose behaviors lie for the most part outside the realm of the Bitcoin blockchain.

As social media developed and language modeling techniques evolved, applying sentiment analysis to social media provided a viable alternative to predicting the Bitcoin price using non-network features. The first attempts to predict Bitcoin price using social media were promising reporting on average a higher accuracy than that reported in the earlier studies that relied on network features. However, much of the work done during the time relied on public sentiments which are not necessarily a good reflection of the true sentiments about Bitcoin. Public sentiments about Bitcoin tend to be positively skewed and as such, the more recent papers seem to start moving in the direction of using news sentiments as a source for predicting the Bitcoin price. The accuracies reported using news sentiments on average exceed those reported in earlier works and further emphasize the importance of news in predicting the Bitcoin price.

While the literature review has identified a few methods of using sentiment analysis to predict Bitcoin price, this study introduces a creative approach to using Twitter as a source of information for news about Bitcoin. It does so, by collating news from major news providers and further constructing a sentiment score that is used to improve the accuracy of prediction of the direction of the Bitcoin price. The advantage of using Twitter as a source of news is two-fold: from a timing perspective Twitter news tend to be near real time and capture larger audience than traditional news outlets. In addition, the ability to capture users' reaction to news by factoring in the number of likes and retweets is another differentiating factor that sets this work apart from other related work. Finally, the use of the latest sentiment analysis techniques like BERT and FinBERT further contributes to the improved accuracy reported in this praxis.

Scanning the research body of work, the models to predict Bitcoin next day price direction have a varying degree of accuracy with the lowest accuracy reported as 52.78% and the highest being 60.1%. Worth noting that near real time or same day Bitcoin price predictions typically achieve higher accuracy with the best accuracy reported in the research community of 62.78%.

Table 2.3 summarizes some of the accuracies reported in key published papers.

Table 2.3

Bitcoin Price Predictions Accuracy Reported in Published Journals

Title	Scope	Accuracy	Comments
Predicting the price of bitcoin using machine learning	Uses historical price data, mining difficulty and hash rate to predict price.	52.78%	Using technical indicators might not be the best approach to predicting Bitcoin price.
Cryptocurrency Price Prediction Using News and Social Media Sentiment	analyzes news and social media ability to predict price fluctuations for three major cryptocurrencies including Bitcoin	43.9% of price increases and 61.9% of price decreases.	Does not share the overall accuracy
Using the Bitcoin Transaction Graph to Predict the Price of Bitcoin	Uses blockchain network-based features to predict the price of Bitcoin by applying support vector machines (SVM) and artificial neural networks (ANN).	55%	Concludes that Blockchain data has limited predictability
Analyzing crypto news sentiment to predict bitcoin prices	Data is gathered from online Crypto blogs specifically coindesk.com and cointelegraph.com.	56%	Using other sources of news and using more advanced NLP techniques like Bert might improve the prediction accuracy
Bitcoin price change and trend prediction through twitter sentiment and data volume	use sentiment and volume of tweets to predict Bitcoin price direction and magnitude of the increase/decrease		Using more recent sentiment analysis techniques like BERT and FinBERT and using news instead of public posts could improve the prediction accuracy

Based on the above, a success criterion for this research work is to build a model that can achieve a prediction accuracy for the direction of Bitcoin next day price that exceeds 60.1%.

CHAPTER 3:

METHODOLOGY

3.1. Introduction

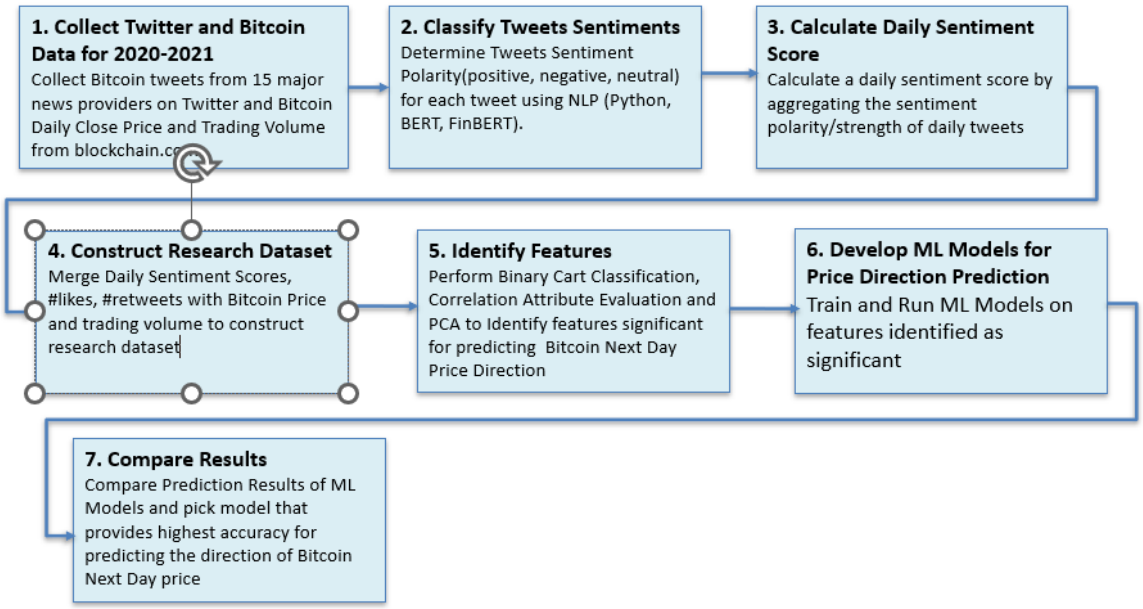
The goal of this praxis is to develop and evaluate how sentiment analysis performed on tweets from major business and news publications can yield better results compared to the research work that has been documented in professional journals and academic publications.

To achieve the goal of this praxis, various models are developed using a combination of tweets features as well as different sentiment analysis techniques. The chapter begins with a discussion of the raw data collection process which encompasses the data from Twitter obtained through trackmyhashtag.com as well as the Bitcoin price data downloaded from blockchain.com for the period 2020-2021. The chapter then proceeds to describe how tweets are classified using different sentiment analysis techniques. Multiple variations of a daily sentiment score are calculated by using different sentiment analysis techniques and factoring different variables in the calculation. Each of the daily sentiment scores calculated is then merged with the number of retweets, number of likes, Bitcoin daily trading volume and the daily Bitcoin price to construct the research dataset. Using the daily sentiment score, number of retweets, number of likes, Bitcoin daily trading volume as predictors and the Bitcoin next day price direction as a response, feature selection methods are applied to identify the features that are most relevant to predicting the direction of the Bitcoin next day price. Those features are then used to develop machine learning models and the results of the models are

compared against results published in research papers to confirm the prediction accuracy improvements offered by the developed model in this praxis.

A high-level block diagram depicting the end-to-end methodology steps followed in this praxis is shown in Figure 3.1.

Figure 3.1
End to End High Level Block Diagram



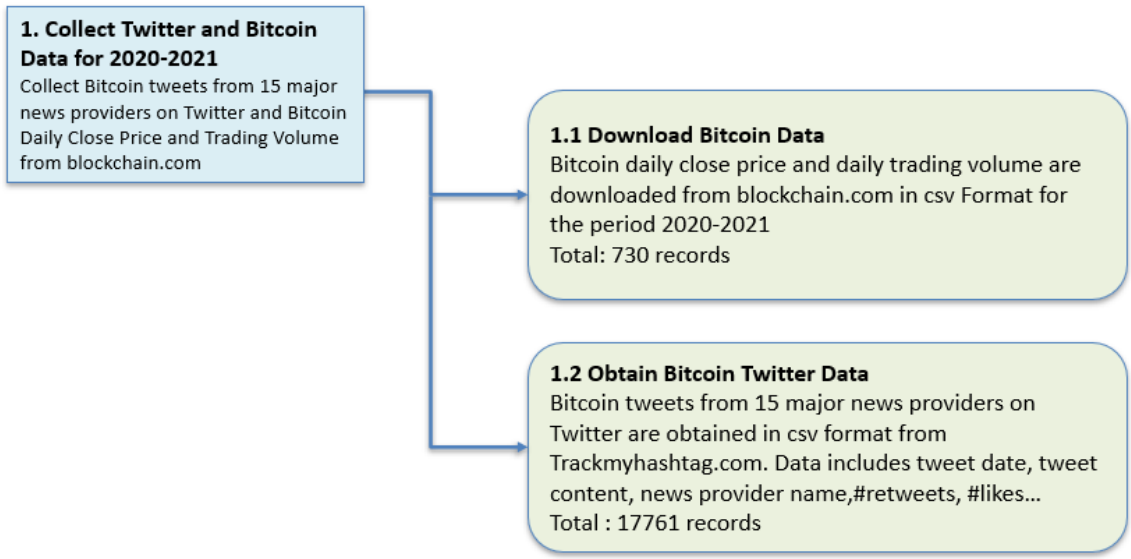
The subsequent sections below further break down each of the methodology steps and describes each of the steps in details.

3.2. Collect Twitter and Bitcoin Data for 2020-2021

The dataset used in this praxis has been obtained from Blockchain.com and Twitter. Data from blockchain.com was downloaded in csv format for the period 1/1/2020 to 12/31/2021. A two-year period was selected to ensure there is enough data to complete the research. Furthermore, years 2020 and 2021 were selected as opposed to earlier years

to make sure the model proposed, accurately reflects the latest development and maturity of the Bitcoin market. The data contains Bitcoin daily close price, Bitcoin trading volume and corresponding close date. Unlike the rest of the stock market which typically trades only on weekdays between 9 am est and 4 pm est, Bitcoin can be traded 24/7 throughout the year. For Bitcoin, the day starts at midnight 00:00:00 and ends at 23:59:59, UTC at which point the daily close price is established. The dataset downloaded from blockchain.com is comprised of 730 records. While many websites publish a daily close price for Bitcoin, blockchain.com is considered a reliable source given its' status as the world's most popular way to buy, sell and trade crypto. In addition, blockchain.com is one of the few websites where the historical Bitcoin data can be downloaded free of charge.

Figure 3.2
Collect Twitter and Bitcoin Data Block Diagram



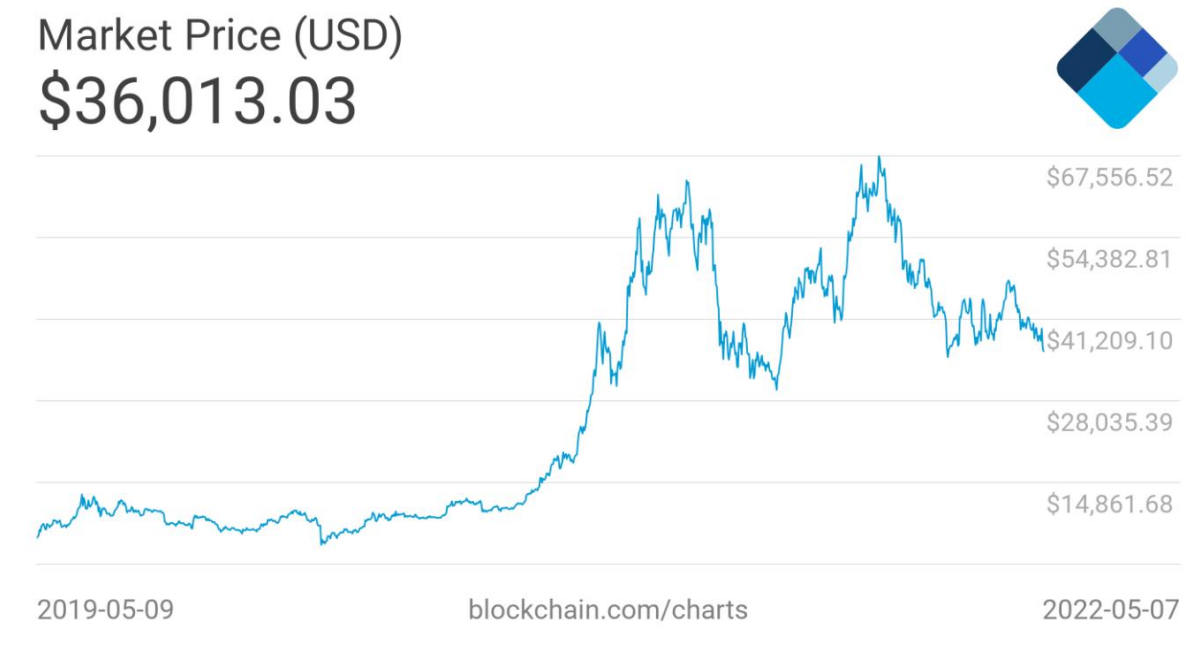
A sample of the data obtained from blockchain.com is shown in Table 3.1.

Table 3.1
Sample Price and Trading Volume from Blockchain.com

Date	Market Close Price	Bitcoin Trading Volume
3/3/2020	$8,912.82	111268001
3/4/2020	$8,754.34	85149526
3/5/2020	$8,758.90	70042033
3/6/2020	$9,067.39	122314467
3/7/2020	$9,155.89	83522226
3/8/2020	$8,898.63	59230883
3/9/2020	$8,039.38	149933311
3/10/2020	$7,931.94	223121982
3/11/2020	$7,885.46	155835851
3/12/2020	$7,936.65	130840913
3/13/2020	$4,830.21	574694715
3/14/2020	$5,609.03	684475989

A graphical view obtained from blockchain.com that shows the prices for the period 5/9/2019 till 5/7/2022 is shown in Figure 3.3.

Figure 3.3
Bitcoin Historical Market Price



Note: Data from blockchain.com.

Twitter provides an API to pull tweet data, however it has some major limitations as summarized below:

- The Twitter API does not allow retrieval of tweets older than 7 days.
- When trying to pull large amounts of data, Twitter will only provide a sample of all the available tweets for the selected period as opposed to the full dataset.
- The API can be queried only a few times within a specified time window.

Fortunately, Twitter data can be obtained from third party companies that provide the requested data for a service fee. For this research, the data was purchased from trackmyhashtag.com, a leader in providing a service to download the requested Twitter data in csv format.

Bitcoin Twitter data was obtained from trackmyhashtag.com for 15 major business and financial news providers for the period 1/1/2020 till 12/31/2021. The number of providers selected ensures there is enough variability in the news. Those providers also have very large following on Twitter which implies bigger influence in shaping up the sentiments of the investors and accordingly Bitcoin price. The total number of records in the dataset is 17761. Trackmyhashtag.com provides a list of more than 20 fields that constitute tweet attributes, but the most critical fields that are relevant to this praxis are: tweet date, tweet content, Retweets and Favorites. A sample of the key fields and records obtained from hastag.com is shown in Table 3.2.

Table 3.2
Sample Key Fields from Hashtag.com

Tweet Id	Tweet Date	Tweet Content	Retweets Received	Favourites Received	Name
"1358836766887862276"	2/8/2021	Bitcoin jumps 10% to record high on Tesla investment news https://t.co/mEZVFXahhc https://t.co/W2YEl5dENH	65	244	Reuters
"1358835437989744642"	2/8/2021	Tesla said that it has acquired $1.5 billion in bitcoins in January and that it could accept the worldâ€™s No. 1 digital asset for payment in the future. https://t.co/ZdFmjNYGkA	12	56	MarketWatch
"1358833925456551936"	2/8/2021	Watch now: ETF Edge on the rise of online betting, bitcoin ETF prospects and the Reddit trading rush https://t.co/MnNwsTW2hz	4	10	CNBC
"1358830488962666498"	2/8/2021	Bitcoin soared to a new all-time high after Tesla said it will soon accept bitcoin as payment for its vehicles. The electric car maker also said it has invested $1.5 billion in bitcoin. https://t.co/QB0TRDmweu	374	1255	CNN

The complete list of fields on Trackmyhashtag.com records are: Tweet Id, Tweet URL, Tweet Posted Time (UTC), Tweet Content, Tweet Type, Client, Retweets Received, Favourites Received, Tweet Location, Tweet Language, User Id, Name, Screen Name, User Bio, Verified or Non-Verified, Profile URL, Protected or Non-protected, User Followers, User Following, Favourites Count, Statuses Count, User Account Creation Date.

3.3. Classify Tweets Sentiments

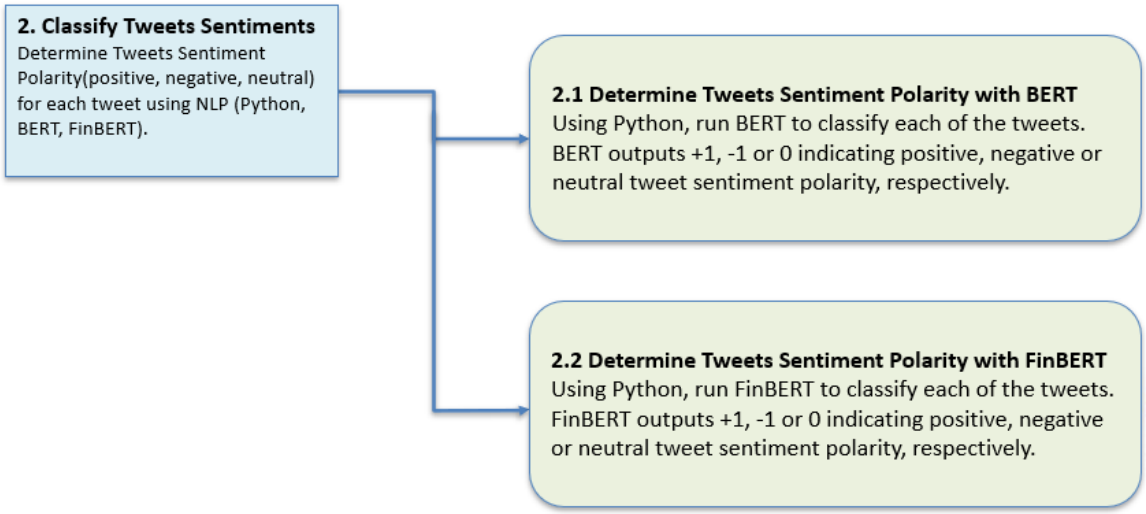
The next step in the methodology is to classify tweets sentiments. Classifying tweets sentiments is about determining the polarity of each of the sentiments in the text source of the tweet. The polarity of the tweet sentiment in this case would be positive if the tone expressed in the tweet is positive, negative, if the tone is negative and neutral if the tone is neutral.

Classifying tweets sentiments is one of the most important steps in the methodology and a building block for the steps to follow. Recall, the premise of this

study is that the sentiment score, which is a daily aggregate of the individual tweets' sentiments polarity, will improve the prediction of the direction of the Bitcoin next day price. Hence, the tweets sentiments polarity classification has a direct effect on the ability of the model to improve the accuracy of the predictions.

Due to the importance of this step, two different models to classify tweets sentiments are attempted as depicted in Figure 3.4.

Figure 3.4
Classify Tweets Sentiments Block Diagram



1) Determining Tweets Sentiment Polarity with BERT: As described in chapter 2 of this research, there are many algorithms for determining tweets sentiments polarity. In the last two years, BERT and its various derivations are surfaced as the best language model not only for generic NLP related problems, but also sentiment analysis. With that in mind, the first round of classifying tweets sentiments is conducted using the BERT model. The input to the model is the tweet text and the output is the polarity of the tweet sentiment. Worth noting that

the BERT model also provides as an output a probability for the tweets sentiment polarity classification which is something that potentially can be considered for future research work. A sample of a tweet text and the corresponding BERT output is shown in Table 3.3:

Table 3.3
Sample BERT Output

Tweet Date	Journal Name	Tweet Content	Sentiment	POS	NEG	NEU
2/23/2021	MarketWatch	Bitcoin drops further below key $50,000 level after Yellen calls it â€˜extremely inefficientâ€™ https://t.co/47nS6hA5oh	NEG	0.007	0.766	0.227

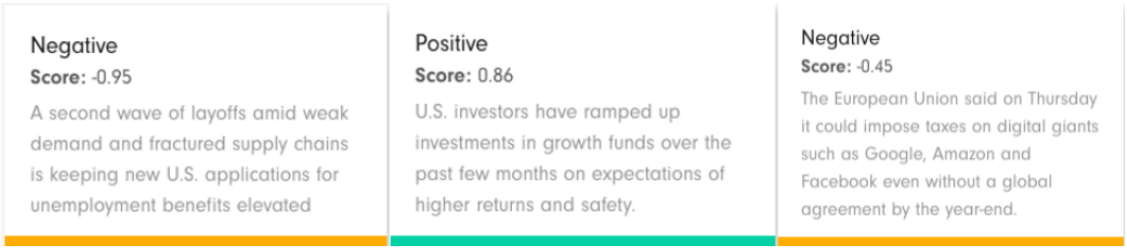
In the above sample, the BERT model classified the tweet sentiment as negative since the probability of the tweet being negative, 0.766(76.6%), is higher than it being positive or neutral. A manual examination of the tweet text would confirm the same. However, it's worth noting that sometimes the tweet classification from BERT does not align with the manual classification of the tweet sentiment polarity, which is understandable given the element of subjectivity inherent in the process of labeling the text for BERT training as well as the manual classification.

2) Determining Tweets Sentiment Polarity with FinBERT:

The second round of classifying tweets sentiments entails using a more advanced model that is built on top of the BERT base model. FinBERT, is expected to provide a better accuracy for the classification. That expectation stems from the fact that FinBERT has been trained specifically on the financial domain.

Table 3.4 shows an example of sentences from financial news classified using FinBERT (Genc, 2021).

Table 3.4
FinBERT Financial News Sample



It is worth noting that there are many variations of FinBERT trained on different datasets from the finance domain. This study uses FinBERT yang2020finbert, a model published by Yi Yang and others, that has been fine-tuned on analyst statements to predict whether the tone is positive, negative, or neutral.

Table 3.5 shows the accuracies reported by both BERT and FinBERT for different language models (Yi, Christopher Siy UY, & Huang, 2020).

Table 3.5
Performance of Different BERT and FinBERT Models

	BERT		FinBERT-BaseVocab		FinBERT-FinVocab	
	cased	uncased	cased	uncased	cased	uncased
PhraseBank	0.755	0.835	0.856	0.870	0.864	**0.872**
FiQA	0.653	0.730	0.767	0.796	0.814	**0.844**
AnalystTone	0.840	0.850	0.872	0.880	0.876	**0.887**

Like BERT, the input to the FinBERT model is the tweet text and the output is the polarity of the tweet sentiment, which could be positive, negative, or neutral. A sample of a tweet text and the corresponding FinBERT output is shown in Table 3.6.

Table 3.6
FinBERT Sample Output

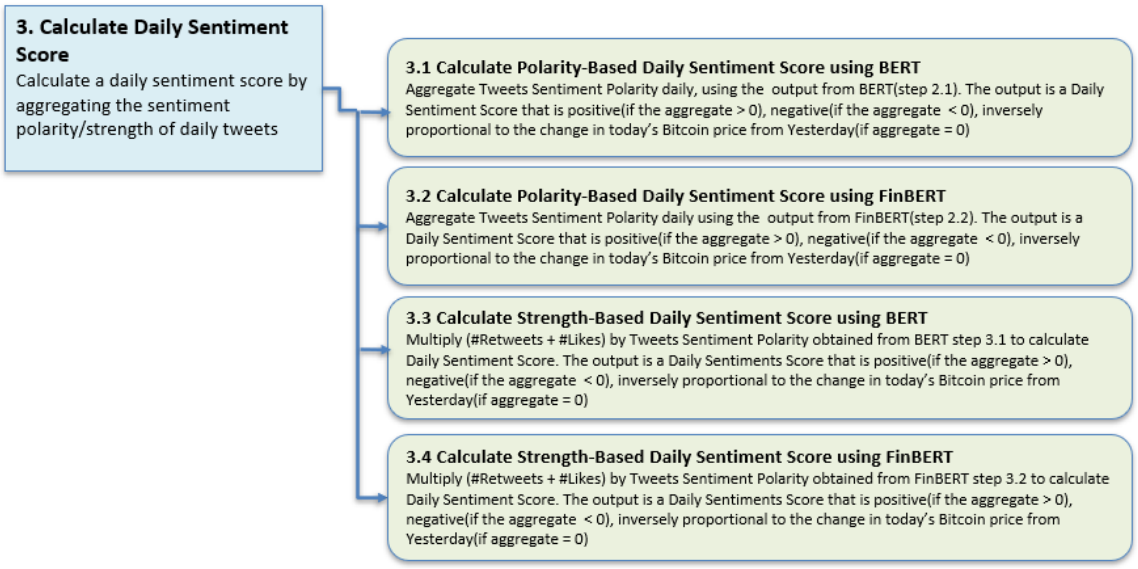
Tweet ID	Tweet Date	Journal Name	Tweet Content	Tweet Sentiment
"1242301983400304641"	3/24/2020	CNBC	Cryptocurrency value rises over $14 billion in 24 hours as bitcoin rallies 10% https://t.co/QB9VoLbTkS	positive

The code to run both the BERT and FinBERT models for this study is written in python using Google Colab (cloud-based IDE). One of the benefits of using Google Colab is to take advantage of the GPU cloud service which enables running classification algorithms at speeds 60 times faster than CPU speeds.

The polarity of the tweets as determined by BERT and FinBERT is used in the next step of the methodology to calculate a daily sentiment score. Each of the sentiment scores calculated will be subsequently used in the next step as depicted in Figure 3.5 to determine the model that provides the best results.

3.4. Calculate Daily Sentiment Score

Figure 3.5
Calculate Daily Sentiment Score Block Diagram



The next step in the methodology is to calculate the daily sentiment score. The daily sentiment score is used as the main predictor for the Bitcoin price. This step follows from and builds on the previous step of classifying each individual tweets sentiment. While the previous step assigned a polarity for each of the tweets' sentiments, this step is intended to assign an overall sentiment polarity for the day which again could be positive, negative, or neutral. As the goal of this study is to predict the Bitcoin next day price, it then becomes important to aggregate the tweets sentiments polarity daily to construct the daily sentiment score. For this study, the daily sentiment score is calculated in two different ways. The first method uses the tweet's polarity, and the second method uses tweet's strength which is calculated by multiplying the tweet's polarity by the associated number of likes and retweets. In the rest of this praxis, the two methods will be referred to as "Polarity-based Daily Sentiment Score" and "Strength-based Daily Sentiment Score" respectively. Furthermore, for each of the methods, two variations of the

sentiment score are calculated based on using the output from BERT or FinBERT tweets sentiment classification as determined by the previous step of the methodology. Below is a detailed explanation of the two methods:

3.4.1. Polarity-Based Daily Sentiment Score

The Polarity-based Daily Sentiment Score is calculated using the formula below:
Daily Sentiment score = 1 ("Positive Sentiment") if P – N > 0, where P and N is the number of positive and negative tweets for the day respectively. In other words, the daily sentiment score is positive if there are more positive Bitcoin tweets for the day than the negative tweets.
Daily Sentiment score = -1 ("Negative Sentiment") if P – N < 0, where P and N is the number of positive and negative tweets for the day respectively. In other words, the daily sentiment score is negative if there are more negative Bitcoin tweets for the day than the positive tweets.
If P – N = 0, then the number of positive sentiments is equal to the number of negative sentiments. In that case, the new sentiment polarity for the day is calculated by taking the reverse sign of the price change between yesterday and today per the rationale below: News could be reflecting the sentiments towards Bitcoin based on the current state of the market or could reflect the sentiments based on the market outlook. If the price change between yesterday and today was positive, then the expectation is that today's news is going to be mostly positive since a component of the news would be reporting on the positive price change examined throughout the day. However, as P-N = 0, then it follows that there is enough negativity in today's news, assumed to be driven by the outlook of the Bitcoin price, that lowers the positivity that was expected. Hence, even though the

number of positive and negative sentiments for the day is the same, the daily sentiment score for the day should be negative driven by the negative outlook. It follows that daily sentiment score moves in an opposite direction to the Bitcoin price movement between yesterday and today. Same logic applies for the case where P – N = 0, and the price change between today and yesterday is negative. Based on the logic described, below are the formulas used for calculating the daily sentiment score for the case where P-N = 0.

Daily Sentiment score = 1 (“Positive”) if Today’s Bitcoin price < Yesterday’s price.

Daily Sentiment score = -1 (“Negative”) if Today’s Bitcoin price > Yesterday’s price

The below examples demonstrate the logic using the available Twitter dataset. In the first example, on 6/25/2021 there are 4 positive tweets and 5 negative tweets as shown in Table 3.7.

Table 3.7

Example 1 of Calculating Polarity-Based Daily Sentiment Score

Tweet Id	Tweet Date	Tweet Sentiment Polarity	Digitized Sentiment
"1408522815108755461"	6/25/2021	negative	-1
"1408512821730549762"	6/25/2021	positive	1
"1408475130825687043"	6/25/2021	negative	-1
"1408461595076960264"	6/25/2021	positive	1
"1408417238042955777"	6/25/2021	positive	1
"1408365427680546816"	6/25/2021	positive	1
"1408323849427562498"	6/25/2021	negative	-1
"1408323588806123521"	6/25/2021	negative	-1
"1408239520638775301"	6/25/2021	negative	-1
Total			-1

Applying the daily sentiment score formula described earlier:

Daily Sentiment Score = P– N = 4 - 5 = -1.

Hence, the Daily Sentiment Score for 6/25/2021 is negative (-1).

In the next example, on 10/13/2021 three are 3 positive and 3 negative tweets as shown in Table 3.8.

Table 3.8

Example 2 of Calculating Polarity-Based Daily Sentiment Score

Tweet Id	Tweet Date	Tweet Sentiment Polarity	Digitized Sentiment
"1448357981805023232"	10/13/2021	positive	1
"1448336983256756226"	10/13/2021	positive	1
"1448329699634126849"	10/13/2021	negative	-1
"1448326223889715200"	10/13/2021	negative	-1
"1448321311168737280"	10/13/2021	positive	1
"1448310147001036803"	10/13/2021	negative	-1
Total			0

Applying the daily sentiment score formula described earlier:

Daily Sentiment Score = P– N = 3 - 3 = 0.

In this case we calculate the daily sentiment score using the change in price between today and Yesterday formula.

The Bitcoin close price on 10/13/2021 was 56242.94 and on 10/12/2021 was 57452.01.

Today's Bitcoin price – Yesterday's Bitcoin price = 56242.94 - 57452.01 = -1209.07

Since the change in price is negative, the daily sentiment score would be assigned a positive value.

Worth noting that the calculations above were performed on the output from FinBERT sentiment classification. The same logic applies if BERT sentiment classification was used instead.

3.4.2. Strength-based Daily Sentiment Score

In addition to the tweet's sentiment polarity, this method also incorporates the number of tweet's likes and retweets to calculate the tweet's strength which is further aggregated on daily basis to calculate the daily sentiment score.

For this method, a single tweet's strength is calculated using the following formula:

Tweet's strength = Polarity * (L + R) where:

Polarity is the polarity of the tweet (+1, -1,0) as determined by sentiment polarity classification performed by BERT or FinBERT.

L is number of tweet's likes

R is number of retweets.

From the above formula, it follows that tweet's strength could be positive, negative, or neutral. A positive strength is further designated as TP and a negative strength is designated as TN.

Daily Sentiment Score is then calculated using the formulas below:

Daily Sentiment score = 1 ("Positive Sentiment") if TP – TN > 0.

In other words, the daily sentiment score is positive if the sum of the positive tweets' strength for the day is greater than the sum of the negative tweets' strength.

Daily Sentiment score = -1 ("Negative Sentiment") if TP – TN < 0.

In other words, the daily sentiment score is positive if the sum of the positive tweets' strength for the day is less than the sum of the negative tweets' strength.

If P – N = 0, then:

- Daily Sentiment score = 1 (“Positive Sentiment”) if Today’s Bitcoin price < Yesterday’s Bitcoin price
- Daily Sentiment score = -1 (“Negative Sentiment”) if Today’s Bitcoin price > Yesterday’s Bitcoin price

Rationale for using the calculation above for the case where TP-TN = 0 is the same as the one used to explain the polarity based daily sentiment score calculation.

Below example uses the same sample that was used earlier in the polarity-based calculation but applies the strength-based calculation instead.

For example, on 6/25/2021 there are 4 positive tweets and 5 negative tweets which according to the polarity-based calculation yields a negative sentiment for the day as shown in Table 3.9.

Table 3.9
Example of Calculating Strength-Based Daily Sentiment Score

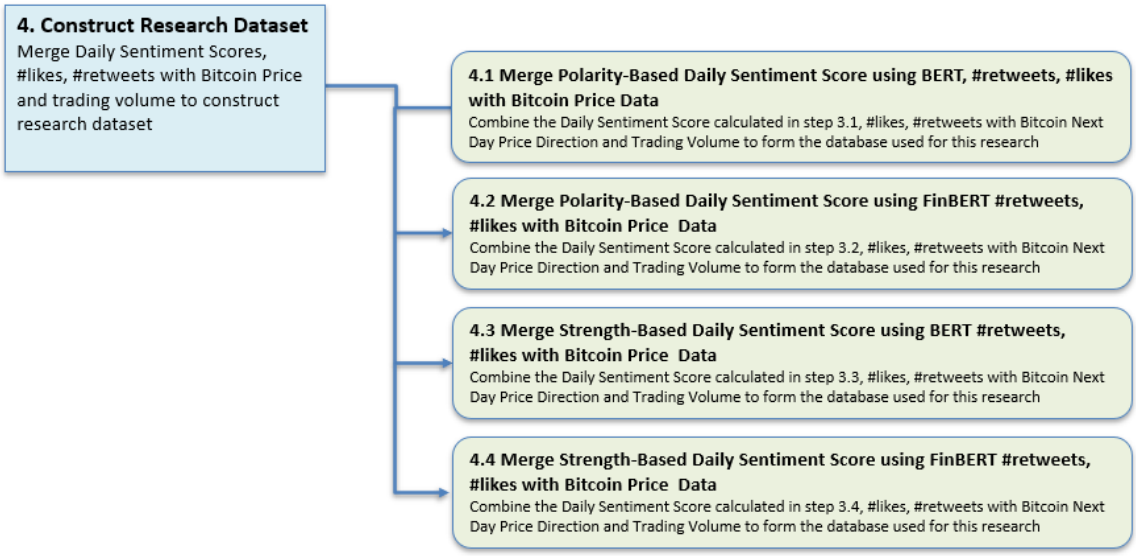
Tweet Id	Tweet Date	Tweet Sentiment Polarity	Number Retweets	Number of Likes	Tweet Strength
"1408522815108755461"	6/25/2021	negative	8	37	-45
"1408512821730549762"	6/25/2021	positive	17	34	51
"1408475130825687043"	6/25/2021	negative	10	23	-33
"1408461595076960264"	6/25/2021	positive	15	48	63
"1408417238042955777"	6/25/2021	positive	27	63	90
"1408365427680546816"	6/25/2021	positive	19	17	36
"1408323849427562498"	6/25/2021	negative	17	51	-68
"1408323588806123521"	6/25/2021	negative	10	30	-40
"1408239520638775301"	6/25/2021	negative	14	31	-45
Total					9

The aggregate of the tweets’ strength for the day is 9, which is positive. Hence the daily sentiment score for the day is positive.

Notice that for the same day, the polarity-based calculation yielded a negative daily sentiment, whereas the daily sentiment was positive when using the strength-based calculation.

3.5. Construct Research Dataset

Figure 3.6
Construct Research Dataset Block Diagram



Constructing the research dataset is a simple but important step in the methodology. The main attributes of the dataset are the daily sentiment scores, total number of tweets' likes for the day, total number of retweets for the day, trading volume for the day and Bitcoin next day price direction.

The Bitcoin next day price direction is calculated using the following formula:

Bitcoin next day price direction = Next day's Bitcoin close price – current day's Bitcoin close price.

If Bitcoin next day price direction > 0, then Bitcoin next day price direction = +1 ("Price Up")

If Bitcoin next day price direction < 0, then Bitcoin next day price direction = -1("Price Down")

If Bitcoin next day price direction =0, then Bitcoin next day price direction = 0 (“Price Unchanged”)

Since there are four different ways to calculate the daily sentiment score as explained earlier, it follows that there are four different variations of the dataset as outlined below:

1) Dataset1 is constructed by merging the Bert-polarity-based daily sentiment score with total number of tweets’ likes for the day, total number of retweets for the day, Bitcoin next day price direction and Bitcoin trading volume for the day, and today’s date.
2) Dataset2 is constructed by merging the polarity-based daily sentiment score from FinBERT with total number of tweets’ likes for the day, total number of retweets for the day, Bitcoin next day price direction and Bitcoin trading volume for the day, and today’s date.
3) Dataset3 is constructed by merging the strength-based daily sentiment score from BERT with total number of tweets’ likes for the day, total number of retweets for the day, Bitcoin next day price direction and Bitcoin trading volume for the day, and today’s date.
4) Dataset4 is constructed by merging the strength-based daily sentiment score from FinBERT with total number of tweets’ likes for the day, total number of retweets for the day, Bitcoin next day price direction and Bitcoin trading volume for the day, and today’s date.

Table 3.10 depicts a sample example of the dataset:

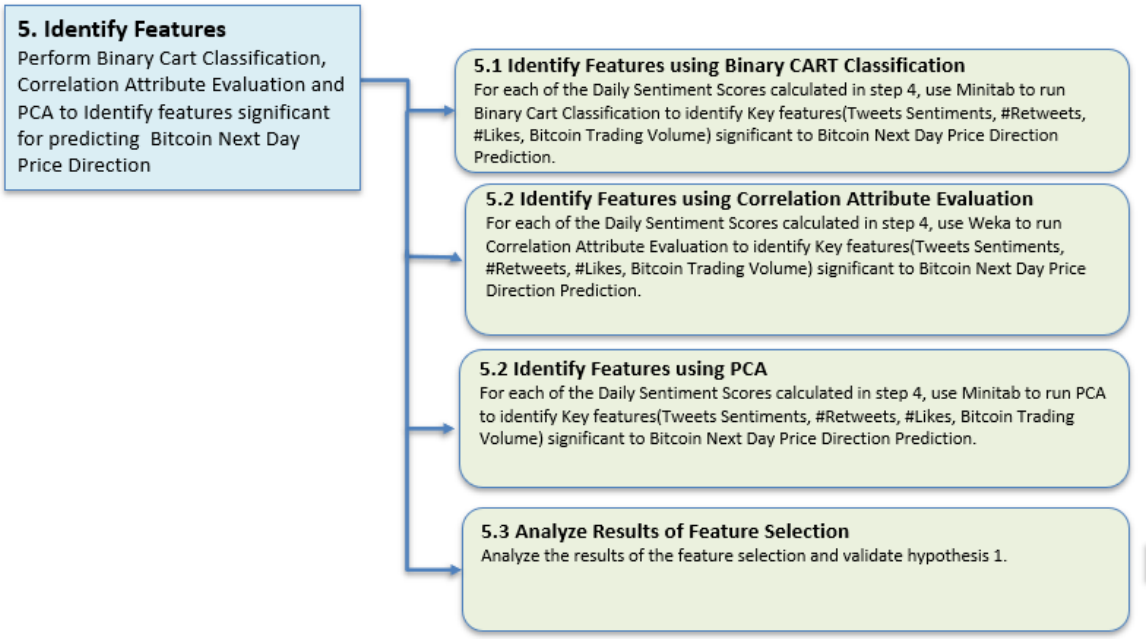
Table 3.10
Praxis Dataset Sample

Date	Daily sentiment score	Bitcoin trading volume	Number of likes	Number of retweets	Next day price direction
4/22/2021	Positive	629420905	1216	3720	Price down
4/23/2021	Negative	1187405699	2015	5788	Price down
4/24/2021	Negative	1413848471	1090	3776	Price down
4/25/2021	Positive	458274327	583	2018	Price down
4/26/2021	Positive	576635611	963	3334	Price up
4/27/2021	Positive	848207063	931	3607	Price up
4/28/2021	Positive	509014491	922	3413	Price down
4/29/2021	Positive	563311742	688	2701	Price down
4/30/2021	Positive	579652973	1138	3781	Price up
5/1/2021	Positive	683907405	1705	6245	Price up

The above table shows the dataset that was constructed using the Polarity-Based Daily sentiment score. The other variations of the dataset will have the same structure but might have different values for the daily sentiment score.

3.6. Identify Features

Figure 3.7
Identify Features Block Diagram



Having constructed the research dataset, the analysis phase of the study begins with identifying the features that are most significant to Bitcoin next day price direction prediction. This is the step where hypothesis 1 is validated.

Predicting Bitcoin price direction is considered a classification problem where the purpose is to forecast whether the price is going to move up or down. Classification is a type of supervised machine learning problem where the goal is to predict, for one or more observations, the category or class they belong to.

When using a classification/prediction model, it is important to evaluate the model performance. This is the process where the model determines how many of the predictions were right and how many were wrong.

In real-world classification problems, the most used performance metrics to evaluate the model is confusion matrix and accuracy.

For this study, feature selection is done in both Minitab and Weka. Minitab is one of the most sophisticated and easy to use statistical analysis tools in the marketplace. Binary CART classification algorithm was used to determine the features that are relevant to the prediction. According to (Ojha, 2020), "CART is a nonparametric technique, which means that it does not require any assumptions for underlying distributions. It is easy to use and can explain a continuous or categorical dependent variable in terms of multiple independent variables".

For the analysis, tweets sentiment polarity, is assumed to be the main predictor for the Bitcoin price. In addition, the number of retweets, number of likes and Bitcoin daily trading volume are potential predictors. The response variable used in the statistical analysis is the Bitcoin next day price direction.

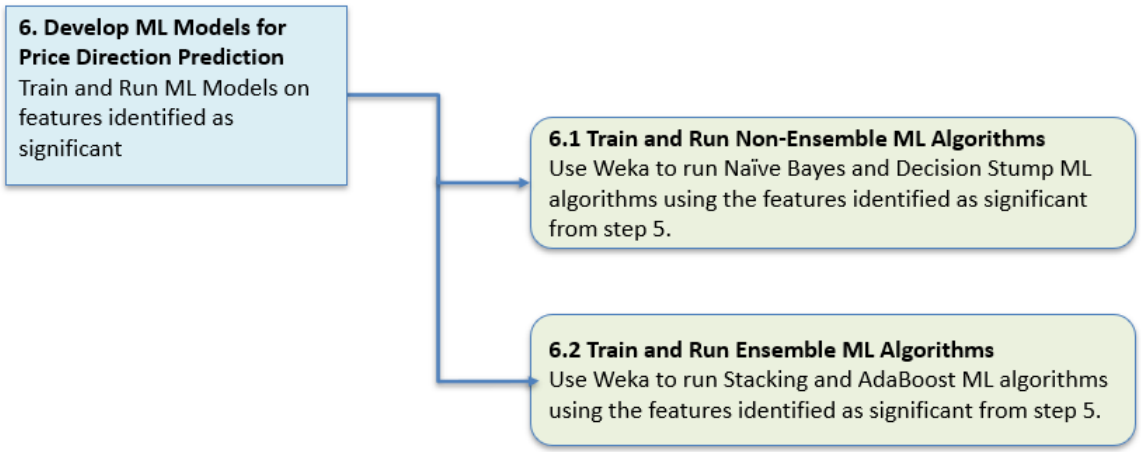
Feature selection was performed on four different variations of the daily sentiment score as described below:

- Feature selection using polarity-based daily sentiment score and tweet classification from BERT.
- Feature selection using polarity-based daily sentiment score and tweet classification from FinBERT.
- Feature selection using strength-based daily sentiment score and tweet classification from BERT.
- Feature selection using strength-based daily sentiment score and tweet classification from FinBERT.

A detailed description and analysis of the feature selection methods performed, the performance metrics and results will be covered in chapter 4 of this study, but the summary conclusion is that polarity-based daily sentiment score using output from FinBERT proved to have the highest relative variable importance and best accuracy as compared to the other tests. Daily Trading Volume, Number of Likes and Number of Retweets were also relevant to predicting the Bitcoin next day price direction.

3.7. Develop ML Models for Price Prediction

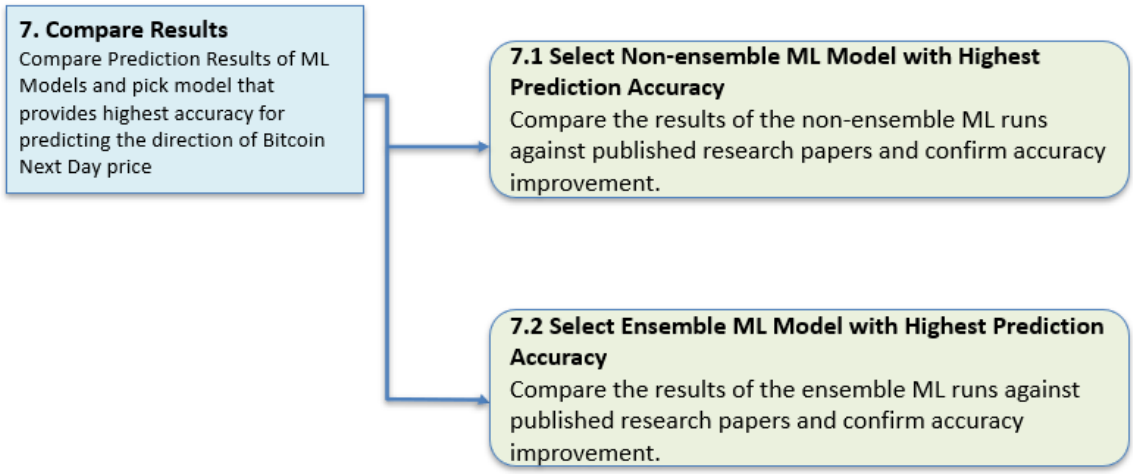
Figure 3.8
Develop ML Models for Price Prediction Block Diagram



In this step of the methodology, machine learning algorithms are trained and run to calculate the accuracy of the model. The previous step of the methodology identified the features that are most significant in predicting the Bitcoin next day price direction. In this step, the ML models are run in Weka using the significant features and the results are analyzed accordingly. The ML algorithms used as well as a discussion of the results obtained is discussed in detail in chapter 4 of this study.

3.8. Compare Results

Figure 3.9
Compare Results Block Diagram



In this step of the methodology, the results from the ML runs are compared. The primary performance metric used for the comparison is accuracy. Confusion matrix is also examined to provide additional insights about the breakdown between true positive, true negative, false positive and false negative predictions.

The model that provides the highest prediction accuracy is then selected and compared against the results published in research papers to confirm prediction accuracy improvements and accordingly success of this research.

Hypothesis 2 and hypothesis 3 are validated as part of this step to confirm the model built in this praxis provides an accuracy greater than 56%.

Analysis of the results, and the comparison are described in detail in chapter 4 of this praxis.

CHAPTER 4:

RESULTS

4.1. Introduction

This chapter provides comprehensive results of the different models constructed to improve the prediction accuracy of Bitcoin next day price direction. Four primary models to predict Bitcoin next day price direction are used and analyzed. These models are presented in Table 4.1.

Table 4.1
Praxis Dataset Sample

Model	Description
BERT-Polarity-based daily sentiment score	Predicting Bitcoin next day price direction by applying BERT classification of tweets and using polarity-based calculation of daily sentiment score.
FinBERT-Polarity-based daily sentiment score	Predicting Bitcoin next day price direction by applying FinBERT classification of tweets and using polarity-based calculation of daily sentiment score.
BERT-Strength-based daily sentiment score	Predicting Bitcoin next day price direction by applying BERT classification of tweets and using strength-based calculation of daily sentiment score.
FinBERT-Strength-based daily sentiment score	Predicting Bitcoin next day price direction by applying FinBERT classification of tweets and using strength-based calculation of daily sentiment score.

For each of the models, feature selection using Weka and Minitab is performed to select the features that are most significant to predicting the Bitcoin next day price direction. Multiple machine learning algorithms are then run using the features that are deemed significant based on feature selection run, and the results are analyzed to select the best performing model.

The next section of this chapter describes the performance metrics that are used to evaluate the different models. The sections that follow describe the feature selection

algorithms performed and the interpretation of the results. Based on the features that are determined to be significant in predicting the direction of the Bitcoin next day price, multiple machine learning algorithms are then trained and run. Selection of the best performing model as well as comparison between the best performing model and existing published predictions is presented at the end of the chapter.

4.2. Performance Metrics

As explained in Chapter 3, predicting Bitcoin price direction is considered a classification problem as the goal is to predict whether the price is going to be up or down. When using a prediction model, it is important to assess the model performance. This is the process where the model evaluates how many of the predictions were right and how many were wrong.

In real-world classification problems, the most used performance metrics to evaluate the model is confusion matrix and accuracy. In addition, to determine the features that are significant to the prediction model, different algorithms can be used. One algorithm that is used in this praxis is Binary CART classification which relies on "relative variable importance" as a metric to determine the relevance of each of the features in the model.

The section below provides a general understanding of relative variable importance as well as accuracy and confusion matrix.

4.2.1. Relative Variable Importance Chart

Relative variable importance is used to determine which predictors are most important to the classification tree. According to (Minitab, 2021), "The Relative variable

importance graph plots the predictors in order of their effect on model improvement. The variable with the highest improvement score is considered the most important variable, and the other variables follow in order of importance".

4.2.2. Accuracy

Accuracy can be defined as the model' ability to make correct predictions for the instances in question.

For accuracy, there are four measures that are important, namely, 'true positive' ,'true negative', 'false positive' and 'false negative'.

- **True positives (TP):** Represents the number of correctly predicted positive observations.
- **False-positive (FP):** Represents the number of incorrectly predicted negative observations.
- **True negative (TN):** Represents the number of correctly predicted negative observations.
- **False-negative (FN):** Represents the number incorrectly predicted positive observations.

According to (HASNAIN, et al., 2020), "Accuracy of a model is the number of correct predictions divided by the total number of predictions and can be calculated using the formula below:

Accuracy = Total number of correct predictions/Total number of predictions, or:

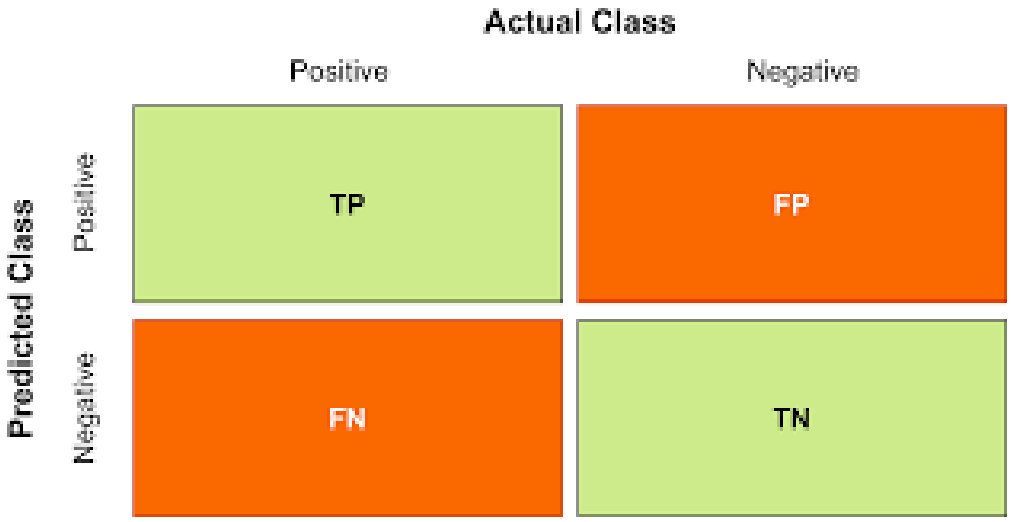
Accuracy=(TP+TN)/(TP+FP+FN+TN)

The accuracy score of a model is a value between 0 and 1, whereby a value of 1 represents a perfect model".

4.2.3. Confusion Matrix

A confusion matrix is a very useful tool that shows the number of predictions for each class that are correct and those that are incorrect. (Iqbal, Mumtaz, Shafi, & Zaidi, 2021) define a confusion matrix as "A visual representation of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN)" as shown in Figure 4.1.

Figure 4.1
Confusion Matrix

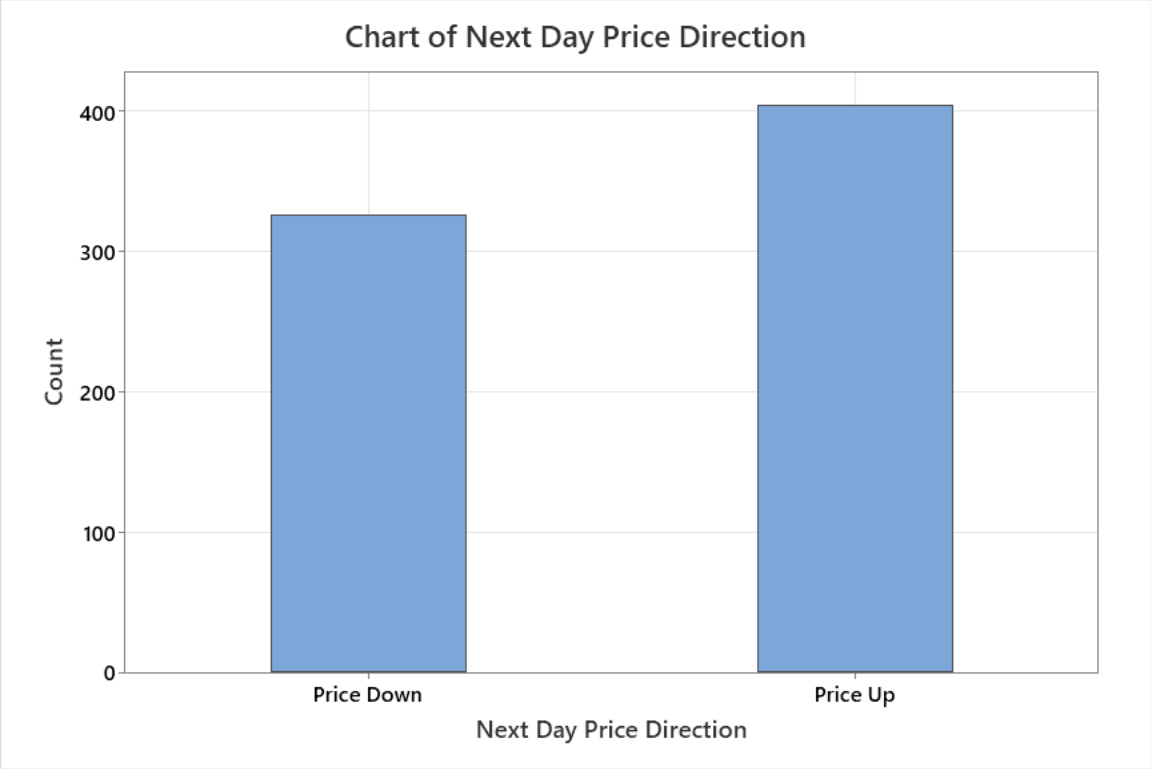


In the subsequent sections, each of the models will be evaluated in terms of accuracy. In addition, the confusion matrix will be presented to gain additional insights on the breakdown of the performance for each of the models discussed.

4.3. Distribution of Bitcoin Next Day Price Direction

Before selecting any of the features, it's important to establish a baseline against which we will measure whether the selected features could improve the prediction accuracy of Bitcoin or not. Figure 4.2 shows the distribution of the Bitcoin next day price direction for the 730 days included in this study.

Figure 4.2
Bitcoin Daily Price Distribution



The chart shows that for the period of this study (1/1/2020-12/31/2021), the Bitcoin price was up 404 days and down 326 days. That is the Bitcoin price was up on average 55.34% of the time. The implication of this, is that without using any prediction model, if we were to predict the Bitcoin next day price to be up every day of the year, we

would achieve a 55.34% prediction accuracy. To determine if the prediction model built in this praxis is adding any improvements, it is important to achieve an accuracy that is greater than this baseline of 55.34%.

In the next section, we perform a Binary CART Classification in Minitab to determine the features that are most relevant to predicting the Bitcoin next day price direction. We also use the feature selection tool in Weka to confirm the results of the feature selection.

4.4. Validating Hypothesis 1

Hypothesis 1 states that Twitter features (tweets sentiments, # likes, #retweets) are significant for accurately predicting the direction of Bitcoin next day price. To validate the hypothesis, three feature selection algorithms are used to determine the features that are significant in predicting the Bitcoin next day price direction. The first algorithm is Binary CART Classification, the second algorithm is Correlation attribute evaluation and the third is PCA (Principal Component Analysis). Minitab is used to perform the Binary CART classification as well as PCA while Weka is used to perform the Correlation attribute evaluation. The predictors and the response variable are described below and the results of each of the model runs are presented and discussed in the sections that follow accordingly.

1) Daily Sentiment Score: This categorical predictor is the main candidate predictor for the Bitcoin next day price direction. The value of this predictor is “Positive” if the sentiment for the day is positive and “Negative” if the sentiment for the day is negative.

As discussed earlier, there are four variations of this predictor as shown below. The method of calculating each of the daily sentiment scores has been explained in the previous section.

- BERT-Polarity-based daily sentiment score
- BERT-Strength-based daily sentiment score
- FinBERT-Polarity-based daily sentiment score
- FinBERT-Strength-based daily sentiment score

2) Daily Trading Volume: This continuous predictor is the Bitcoin trading volume for the day.
3) Number of likes: This predictor is calculated as the total sum of the number of likes for all Bitcoin tweets posted for the day.
4) Number of retweets: This predictor is calculated as the total sum of the number of retweets for all Bitcoin tweets posted for the day.
5) Bitcoin next day price direction: This binary response variable has a value of "Price Up" if the next day price is greater than the current day price and a value of "Price Down" if the next day price is smaller than the current day price.

4.4.1. Binary Cart Classification

Feature selection is an important technique in data analysis for determining the features that are most significant in preserving information.

Per (Dobilas, 2021), "Classification And Regression Tree (CART) is a predictive model, which explains how the values of an outcome variable can be predicted based on the values of input variables. As the name suggests, CART can be used for both classification and regression problems". Classification is used in areas where there is a

finite set of values for the output variable, e.g., whether the Bitcoin price will go up tomorrow or not.

Per (Questier, 2005), “The CART method allows to perform supervised feature selection by modeling one response variable (y) by some explanatory variables (x)”.

One of the advantages of using Binary CART classification for feature selection is its ability to show the relative variable importance of each of the features in a selected model, thus allowing the experimenter to pick the features that are most relevant to their study. Another advantage of using Binary CART classification for attribute selection is its ability to perform the attribute selection on both categorical and continuous variables, something other algorithms (ex. PCA) are not capable of doing. The next sections show the results of running Binary CART classification on the models that have been constructed for this praxis.

4.4.1.1. Binary Cart Classification using BERT- Polarity-based daily sentiment score

The test is performed in Minitab using the predictors and response variables described in section 4.3. The two important metrics to focus on to determine the important features for predicting the Bitcoin next day price direction are accuracy and relative variable importance. Figure 4.3 and Figure 4.4 show the results as obtained from Minitab, followed by a discussion of those results.

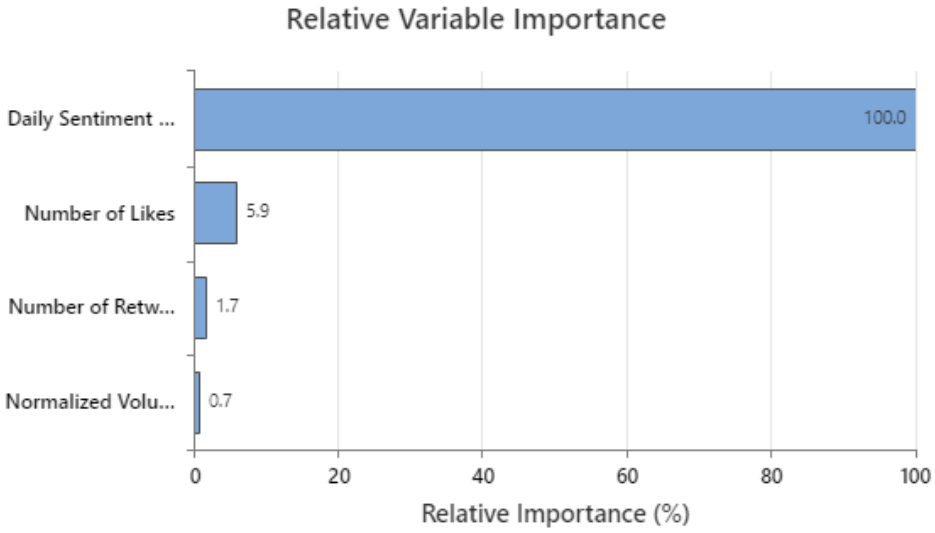
Figure 4.3
BERT- Polarity-Based Daily Sentiment Score Confusion Matrix

Confusion Matrix

		Predicted Class (Training)			Predicted Class (Test)		
Actual Class	Count	Price Up	Price Down	% Correct	Price Up	Price Down	% Correct
Price Up (Event)	404	256	148	63.4	256	148	63.4
Price Down	326	150	176	54.0	150	176	54.0
All	730	406	324	59.2	406	324	59.2

Statistics	Training (%)	Test (%)
True positive rate (sensitivity or power)	63.4	63.4
False positive rate (type I error)	46.0	46.0
False negative rate (type II error)	36.6	36.6
True negative rate (specificity)	54.0	54.0

Figure 4.4
BERT- Polarity-Based Daily Sentiment Score Pareto Chart



Variable importance measures model improvement when splits are made on a predictor. Relative importance is defined as % improvement with respect to the top predictor.

The prediction accuracy of the model is 59.2% as shown in the confusion matrix above. This accuracy is higher than the established baseline of 55.34% indicating that the features selected for this test are improving the prediction accuracy.

The pareto chart shows that the main predictor for the Bitcoin next day price direction is the daily sentiment score, which is a daily aggregate of each of the tweets' sentiments, followed by number of likes, number of retweets and daily trading volume.

4.4.4.2. Binary Cart Classification using BERT- Strength-based daily sentiment score

Test is performed in Minitab using the predictors and response variables described in section 4.3.

Figure 4.5 and Figure 4.6 show the results from Minitab, followed by a discussion of those results.

The daily sentiment score used is calculated as the daily aggregation of the tweets sentiment strength as determined by BERT. Each of the tweet's strength is calculated by multiplying the sentiment polarity of the tweet by the tweet's number of likes and retweets.

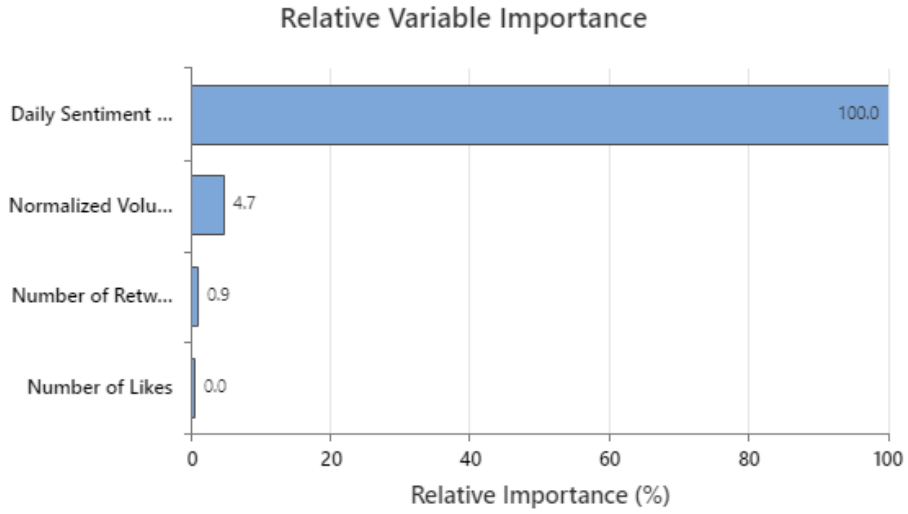
Figure 4.5
BERT- Strength-Based Daily Sentiment Score Confusion Matrix

Confusion Matrix

		Predicted Class (Training)			Predicted Class (Test)		
Actual Class	Count	Price Up	Price Down	% Correct	Price Up	Price Down	% Correct
Price Up (Event)	404	243	161	60.1	243	161	60.1
Price Down	326	142	184	56.4	142	184	56.4
All	730	385	345	58.5	385	345	58.5

Statistics	Training (%)	Test (%)
True positive rate (sensitivity or power)	60.1	60.1
False positive rate (type I error)	43.6	43.6
False negative rate (type II error)	39.9	39.9
True negative rate (specificity)	56.4	56.4

Figure 4.6
BERT- Strength-Based Daily Sentiment Score Pareto Chart



Variable importance measures model improvement when splits are made on a predictor. Relative importance is defined as % improvement with respect to the top predictor.

The prediction accuracy of the model is 58.5% as shown in the confusion matrix above. This accuracy is higher than the established baseline of 55.34% indicating that the features selected for this test are improving the prediction accuracy.

Like the previous run, the pareto chart shows that the main predictor for the Bitcoin next day price direction is the sentiment score which is a daily aggregate of each of the tweet's strengths calculated by multiplying the polarity of the tweet sentiment by the number of likes and number of retweets, followed by the daily trading volume and number of retweets.

4.4.4.3. Binary Cart Classification Using FinBERT- Polarity-Based Daily Sentiment Score

The test is performed in Minitab using the predictors and response variables described in section 4.3. Figure 4.7, Figure 4.8 and Figure 4.9 show the results, followed by a discussion of those results.

Figure 4.7
FinBERT- Polarity-Based Daily Sentiment Score Confusion Matrix (Minitab)

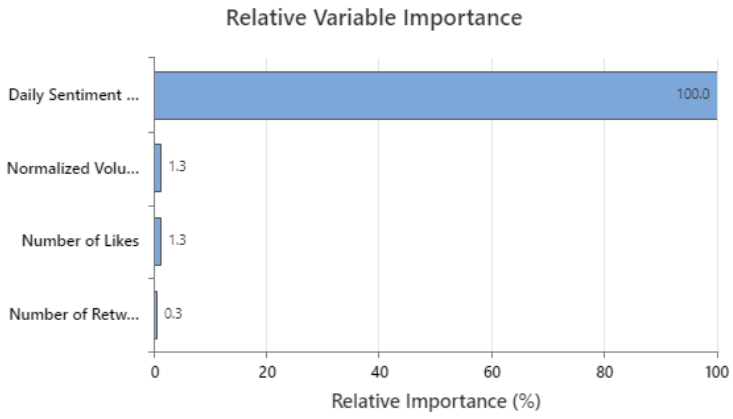
Confusion Matrix

		Predicted Class (Training)			Predicted Class (Test)		
Actual Class	Count	Price Up	Price Down	% Correct	Price Up	Price Down	% Correct
Price Up (Event)	404	286	118	70.8	286	118	70.8
Price Down	326	167	159	48.8	167	159	48.8
All	730	453	277	61.0	453	277	61.0

Statistics	Training (%)	Test (%)
True positive rate (sensitivity or power)	70.8	70.8
False positive rate (type I error)	51.2	51.2
False negative rate (type II error)	29.2	29.2
True negative rate (specificity)	48.8	48.8

Figure 4.8
FinBERT- Polarity-Based Daily Sentiment Score Confusion Matrix

Figure 4.9
FinBERT- Polarity-Based Daily Sentiment Score Pareto Chart from Minitab



Variable importance measures model improvement when splits are made on a predictor. Relative importance is defined as % improvement with respect to the top predictor.

The prediction accuracy of the model is 61% as shown in Figure 4.7. This accuracy is higher than the established baseline of 55.34% indicating that the features selected for this test are improving the prediction accuracy.

The pareto chart in Figure 4.9 shows that the main predictor for the Bitcoin next day price direction is the sentiment score, followed by the daily trading volume, number of likes and number of retweets.

4.4.4.4. Binary Cart Classification Using FinBERT- Strength-Based Daily Sentiment Score

The test is performed in Minitab using the predictors and response variables described in section 4.3. Figure 4.10 presents the results followed by a discussion of those results.

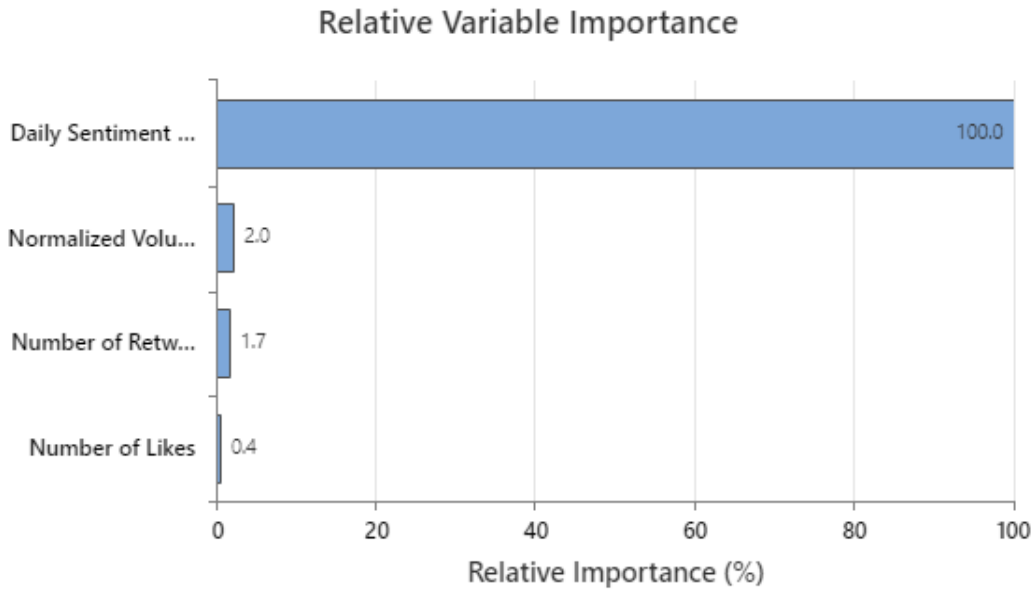
Figure 4.10
FinBERT- Strength-Based Daily Sentiment Score Confusion Matrix

Confusion Matrix

		Predicted Class (Training)			Predicted Class (Test)		
Actual Class	Count	Price Up	Price Down	% Correct	Price Up	Price Down	% Correct
Price Up (Event)	404	271	133	67.1	271	133	67.1
Price Down	326	168	158	48.5	168	158	48.5
All	730	439	291	58.8	439	291	58.8

Statistics	Training (%)	Test (%)
True positive rate (sensitivity or power)	67.1	67.1
False positive rate (type I error)	51.5	51.5
False negative rate (type II error)	32.9	32.9
True negative rate (specificity)	48.5	48.5

Figure 4.11
FinBERT- Strength-Based Daily Sentiment Score Pareto Chart



The prediction accuracy of the model is 58.8% as shown in Figure 4.11. This accuracy is higher than the established baseline of 55.34% indicating that the features selected for this test are improving the prediction accuracy. The pareto chart shows that the main predictor for the Bitcoin next day price direction is the sentiment score, followed by the daily trading volume, number of retweets and number of likes.

4.4.2. Discussion of the Results of Binary CART Classification

Table 4.2 summarizes the results of the Binary CART classification applied to the four variations of the daily sentiment score and the other predictors examined in this praxis.

Table 4.2
Results of Binary CART Classification

Predictors	Response	Relative Variable Importance	Accuracy
BERT-Polarity-based Daily Sentiment Score	Next Day Bitcoin Price Direction	Daily Sentiment Score: 100 Number of Likes: 5.9 Number of Retweets: 1.7 Trading Volume: 0.7	59.2%
BERT-Strength-based Daily Sentiment Score	Next Day Bitcoin Price Direction	Daily Sentiment Score: 100 Daily Volume: 4.7 Number of Retweets: 0.9 Number of Likes: 0.0	58.5%
FinBERT-Polarity-based Daily Sentiment Score	Next Day Bitcoin Price Direction	Daily Sentiment Score: 100 Trading Volume: 1.3 Number of Likes: 1.3 Number of Retweets: 0.3	61%
FinBERT-Strength-based Daily Sentiment Score	Next Day Bitcoin Price Direction	Daily Sentiment Score: 100 Trading Volume: 2.0 Number of Retweets: 1.7 Number of Likes: 0.4	58.8%

Analysis of the results shows that the FinBERT-polarity-based daily sentiment score achieves the highest accuracy in comparison to the other models. The relative variable importance for this model shows that daily sentiment score is the most relevant feature followed by daily trading volume, number of likes and number of retweets.

4.4.3. Correlation Attribute Evaluation

The correlation attribute evaluation method assesses the attribute worth by measuring Pearson's correlation between the attribute and the class. The search method selected is "Ranker" which ranks attributes by their individual evaluations, thus allowing us to select the number of attributes that we deem most important for the study at hand.

To perform correlation attribute evaluation, we used Weka and the results of running the method on the different models that have been constructed for this praxis are shown below followed by an analysis and discussion of the results.

4.4.3.1. Attribute Selection using BERT-Polarity-Based Daily Sentiment Score

The test is performed in Weka using the predictors and response variables described in section 4.3. The daily sentiment score used is calculated as the daily aggregation of the tweets sentiment polarity as determined by BERT.

Figure 4.12 shows a screen shot of the results followed by a discussion of those results.

Figure 4.12
BERT- Polarity-Based Daily Sentiment Score Attribute Selection

```
Attributes:   5
              Current Day Volume
              Number of Likes
              Number of Retweets
              Daily Sentiment Score
              Next Day Price Direction
Evaluation mode:    evaluate on all training data


=== Attribute Selection on all input data ===

Search Method:
        Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 5 Next Day Price Direction):
        Correlation Ranking Filter
Ranked attributes:
 0.1736   4 Daily Sentiment Score
 0.0184   3 Number of Retweets
 0.0148   1 Current Day Volume
 0.0144   2 Number of Likes
```

The results show that the main predictor for the Bitcoin next day price direction is the sentiment score, followed by the number of retweets, daily trading volume, and number of likes. Worth noting while all four predictors can add to the model accuracy,

the daily sentiment score seems to be the primary predictor for the Bitcoin next day price direction.

4.4.3.2. Attribute Selection Using BERT-Strength-Based Daily Sentiment Score

The test is performed in Weka using the predictors and response variables described in section 4.3. The daily sentiment score used is calculated as the daily aggregation of the tweets sentiment strength as determined by BERT. Each of the tweet's strength is calculated by multiplying the sentiment polarity of the tweet by the tweet's number of likes and retweets. Figure 4.13 shows a screen shot of the results followed by a discussion of those results.

Figure 4.13
BERT- Strength-Based Daily Sentiment Score Attribute Selection

```
Attributes:   5
              Current Day Volume
              Number of Likes
              Number of Retweets
              Daily Sentiment Score
              Next Day Price Direction
Evaluation mode:    evaluate on all training data


=== Attribute Selection on all input data ===

Search Method:
        Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 5 Next Day Price Direction):
        Correlation Ranking Filter
Ranked attributes:
 0.1652  4 Daily Sentiment Score
 0.0184  3 Number of Retweets
 0.0148  1 Current Day Volume
 0.0144  2 Number of Likes
```

The results show that the main predictor for the Bitcoin next day price direction is the sentiment score, followed by the number of retweets, daily trading volume, and number of likes. Worth noting while all four predictors can add to the model accuracy, the daily sentiment score seems to be the primary predictor for the Bitcoin next day price direction.

4.4.3.3. Attribute Selection Using FinBERT- Polarity-Based Daily Sentiment Score

The test is performed in Weka using the predictors and response variables described in section 4.3. The daily sentiment score used is calculated as the daily aggregation of the tweets sentiment polarity as determined by FinBERT. Figure 4.14 shows a screen shot of the results followed by a discussion of those results.

Figure 4.14
FinBERT- Polarity-Based Daily Sentiment Score Attribute Selection

```
Attributes:   5
              Daily Trading Volume
              Number of Likes
              Number of Retweets
              Daily Sentiment Score
              Next Day Price Direction
Evaluation mode:    evaluate on all training data



=== Attribute Selection on all input data ===

Search Method:
        Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 5 Next Day Price Direction):
        Correlation Ranking Filter
Ranked attributes:
 0.2004  4 Daily Sentiment Score
 0.0184  3 Number of Retweets
 0.0148  1 Daily Trading Volume
 0.0144  2 Number of Likes
```

The results show that the main predictor for the Bitcoin next day price direction is the sentiment score, followed by the number of retweets, daily trading volume, and number of likes. Worth noting while all four predictors can add to the model accuracy, the daily sentiment score seems to be the primary predictor for the Bitcoin next day price direction.

4.4.3.4. Attribute Selection Using FinBERT-Strength-Based Daily Sentiment Score

The test is performed in Weka using the predictors and response variables described in section 4.3. The daily sentiment score used is calculated as the daily aggregation of the tweets sentiment strength as determined by FinBERT. The tweets strength is calculated by multiplying the sentiment polarity of the tweet by the tweet's number of likes and retweets. Figure 4.15 shows a screen shot of the results followed by a discussion of those results.

Figure 4.15
FinBERT- Strength-Based Daily Sentiment Score Attribute Selection

```
Attributes:   5
              Daily Trading Volume
              Number of Likes
              Number of Retweets
              Daily Sentiment Score
              Next Day Price Direction
Evaluation mode:    evaluate on all training data



=== Attribute Selection on all input data ===

Search Method:
        Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 5 Next Day Price Direction):
        Correlation Ranking Filter
Ranked attributes:
 0.1578   4 Daily Sentiment Score
 0.0184   3 Number of Retweets
 0.0148   1 Daily Trading Volume
 0.0144   2 Number of Likes
```

The results show that the main predictor for the Bitcoin next day price direction is the sentiment score, followed by the number of retweets, daily trading volume, and number of likes. Worth noting while all four predictors can add to the model accuracy, the daily sentiment score seems to be the primary predictor for the Bitcoin next day price direction.

4.4.4. Discussion of the Results of Correlation Attribute Evaluation

Table 4.3 summarizes the results of running the Correlation attribute evaluation method to the four variations of the daily sentiment score and the other predictors examined in this praxis.

Table 4.3
Results of Correlation Attribute Evaluation

Predictors	Response	Variable Ranking	Degree of Correlation
BERT-Polarity-based Daily Sentiment Score	Next Day Bitcoin Price Direction	Daily Sentiment Score: 1	0.1736
		Number of Retweets: 2	0.0184
		Current Day Volume: 3	0.0148
		Number of Likes: 4	0.0144
BERT-Strength-based Daily Sentiment Score	Next Day Bitcoin Price Direction	Daily Sentiment Score	0.1652
		Number of Retweets	0.0184
		Current Day Volume	0.0148
		Number of Likes	0.0144
FinBERT-Polarity-based Daily Sentiment Score	Next Day Bitcoin Price Direction	Daily Sentiment Score	0.2004
		Number of Retweets	0.0184
		Daily Trading Volume	0.0148
		Number of Likes	0.0144
FinBERT-Strength-based Daily Sentiment Score	Next Day Bitcoin Price Direction	Daily Sentiment Score	0.1578
		Number of Retweets	0.0184
		Daily Trading Volume	0.0148
		Number of Likes	0.0144

Analysis of the results show that the FinBERT-polarity-based daily sentiment score has the highest degree of correlation with the Bitcoin next day price direction. The correlation factor shows that daily sentiment score is the most relevant feature followed by number of retweets, daily trading volume, and number of likes. This largely matches the results of the feature selection performed with Binary CART classification. Accordingly, the FinBERT-Polarity-based daily sentiment score, daily trading volume, number of likes and number of retweets will be used in the next round of machine learning runs to determine the potential improvements to the Bitcoin next day price direction prediction.

4.4.5. PCA Analysis

In addition to using CART and Attribute Evaluation Classification, PCA analysis is performed to identify the features that are relevant to Bitcoin Next day price direction. For this, PCA was executed using the following features: Daily Sentiment score, Bitcoin Trading volume, number of likes and number of retweets. The output from the PCA run is shown in Figure 4.16.

Figure 4.16
PCA Analysis for FinBERT-Polarity-Based Daily Sentiment Score

Eigenanalysis of the Correlation Matrix

Eigenvalue	2.2675	0.9974	0.7077	0.0274
Proportion	0.567	0.249	0.177	0.007
Cumulative	0.567	0.816	0.993	1.000

Eigenvectors

Variable	PC1	PC2	PC3	PC4
Current Day Volume	0.429	0.130	0.894	-0.000
Number of Retweets	0.635	-0.117	-0.288	0.707
Number of Likes	0.634	-0.127	-0.286	-0.707
Binary Daily Sentiment Score	0.102	0.976	-0.191	-0.008

4.4.6. Discussion of the Results of PCA Analysis

The results show that the Binary Daily sentiment score, which is a binary representation of the FinBERT-polarity-based daily sentiment score comprises a significant portion of the first two primary PCA components (PC1 and PC2). The eigenvalues show that at least three PCA components are needed to explain 99.3% of the variation. These results are in-line with the findings of the other two methods we used for attribute selection, that is the findings of both CART analysis as well as Correlation Attribute Evaluation.

4.5. Preparing the Dataset for Machine Learning Runs

Before running the machine learning algorithms, the praxis dataset is split into training and testing. The training dataset as the name implies is used to train the model and the testing dataset is used to validate the accuracy of the model. For this research, the dataset is split by a ratio of 70/30, which means 70% of the data is used for training the model and 30% is used to test the validity of the model. To avoid any bias in determining the instances that comprise the training vs. those that could be used for testing, the Weka feature "percentage split" is used. This randomly splits the dataset into a training and a testing partition.

With dataset split into training and testing, we are ready to run the machine learning algorithms and evaluate the results.

4.6. Validating Hypothesis 2

Hypothesis 2 states that a predictive model using Twitter features (tweets sentiments, # likes, #retweets) will accurately predict the direction of Bitcoin next day price.

To validate the hypothesis, several machine learning algorithms are run, and the results are compared accordingly. The input to the machine learning algorithms is the set of features that have been determined to be significant in predicting next day price direction. Based on the results discussed in the previous section, the features that are deemed to be significant are daily sentiment score, daily trading volume, daily number of likes and daily number of retweets. Accordingly, the machine learning algorithms will use all these features as an input to the model and the next day price direction as the expected output of the model.

Following machine learning algorithms were selected to validate hypothesis 2: Naïve Bayes and Decision Stump. The next sections describe the different machine learning algorithms used and discuss the results of the ML runs.

4.6.1. Naïve Bayes

(Ray, 2017, pp. 1-3) defines Naïve Bayes as, "A classification technique based on Bayes' Theorem where the predictors are assumed to be independent. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

Naive Bayes model is easy to build and is useful for very large data sets. Albeit the simplicity of the model, Naive Bayes is known to outperform some of the highly sophisticated and advanced classification methods".

Figure 4.17 shows the results of running Naïve Bayes algorithm in Weka:

Figure 4.17
Naive Bayes ML Run Results

```
Correctly Classified Instances         138               63.0137 %
Incorrectly Classified Instances        81               36.9863 %
Kappa statistic                          0.234
Mean absolute error                      0.4724
Root mean squared error                  0.4807
Relative absolute error                 95.7168 %
Root relative squared error             97.0002 %
Total Number of Instances              219

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.720    0.489    0.662      0.720    0.690      0.235    0.659     0.697     Price Up
                 0.511    0.280    0.578      0.511    0.542      0.235    0.659     0.615     Price Down
Weighted Avg.    0.630    0.399    0.626      0.630    0.626      0.235    0.659     0.662

=== Confusion Matrix ===

  a  b   <-- classified as
 90 35 |  a = Price Up
 46 48 |  b = Price Down
```

The results show that the accuracy of the model is ~ 63%, which is much higher than the established baseline of 55.34%.

4.6.2. Decision Stump

Decision stump is one of the most straightforward binary classification algorithms (Hsu, 2021). (Shaier, 2019) defines a decision stump as "A machine learning model consisting of a one-level decision tree. A decision stump makes a prediction based on the value of just a single input feature". The results of the feature selection runs performed in the previous section clearly show that the daily sentiment score has by far the highest predictive power; hence, it is expected that the decision stump algorithm will yield high accuracy results when the daily sentiment score is used as the single feature to predict the Bitcoin next day price direction.

Figure 4.18 shows the output of running decision stump in Weka, followed by analysis of the results.

Figure 4.18
Decision Stump ML Run Results

```
Correctly Classified Instances         142               64.8402 %
Incorrectly Classified Instances        77               35.1598 %
Kappa statistic                          0.2757
Mean absolute error                      0.4709
Root mean squared error                  0.4794
Relative absolute error                 95.4089 %
Root relative squared error             96.7302 %
Total Number of Instances              219

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
                 0.720    0.447    0.682      0.720    0.700      0.276    0.637     0.651
                 0.553    0.280    0.598      0.553    0.575      0.276    0.637     0.522
Weighted Avg.    0.648    0.375    0.646      0.648    0.646      0.276    0.637     0.596

=== Confusion Matrix ===

  a  b   <-- classified as
 90 35 |  a = Price Up
 42 52 |  b = Price Down
```

The results show that the accuracy of the model is ~ 64.8%. This is a tremendous improvement over Naïve Bayes. However, ensemble methods might improve the prediction accuracy, hence those models are explored in the next section of this praxis.

4.7. Validating Hypothesis 3

Hypothesis 3 states that an ensemble of machine learning techniques using Twitter features (tweets sentiments, # likes, #retweets) will improve the prediction of the direction of Bitcoin next day price.

To validate the hypothesis, ensemble machine learning algorithms will be trained and run, and the results will be analyzed to select the best model that will be compared against results published in academic and professional journals. The input to the machine learning algorithms is the set of features that have been determined to be significant in predicting next day price direction. Based on the results discussed in section 4.3, the features that are deemed to be significant are daily sentiment score, daily trading volume, daily number of likes and daily number of retweets. Accordingly, the machine learning algorithms use all these features as an input to the model and the next day price direction as the expected output of the model.

The next section describes the ensemble techniques followed by a discussion of the results of running ensemble ML algorithms to predict the Bitcoin next day price direction.

4.7.1. Ensemble Techniques

Per (Ayuya, 2020, p. 1), “Ensemble learning refers to a method where many base models are combined to carry out the same task. These base models are usually referred to as weak learners.

The main principle behind ensemble learning is that a weak learner predicts poorly when alone. But when combined with other weak learners, ensemble techniques create a strong learner". The premise is that the strong learner performs better than each of the weak learners.

Ensemble techniques are powerful and in general perform better than non-ensemble techniques. This is understandable given that the outcome of ensemble learning is a strong learner that benefits of the combined performance improvement of individual ML techniques.

Another advantage of using ensemble techniques is their ability to reduce prediction errors.

According to (Ayuya, 2020, p. 2), "Prediction errors of machine learning models are characterized by bias and variance. Bias refers to the difference between predicted and actual outcome. Variance describes the responsiveness of a model to small fluctuations in the training set. While ideally, it's best to have a model with low bias and variance, it's not an easy task to achieve.

Ensemble techniques can lower prediction variance through averaging of the performance of various models".

To validate hypothesis 3, two ensemble methods are used as described below.

4.7.2. Stacking

(Brownlee, 2020, p. 1) define Stacking as "An ensemble machine learning algorithm that tries to improve performance by learning how to best combine the predictions from multiple well-performing machine learning models".

Figure 4.19 shows the output of running the stacking algorithm in Weka, followed by an analysis and discussion of the results.

Figure 4.19
Stacking ML Run Results

```
Correctly Classified Instances         122               55.7078 %
Incorrectly Classified Instances        97               44.2922 %
Kappa statistic                          0.09
Mean absolute error                      0.4704
Root mean squared error                  0.5209
Relative absolute error                 95.3184 %
Root relative squared error            105.1075 %
Total Number of Instances              219

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.632    0.543    0.608      0.632    0.620      0.090    0.562     0.629     Price Up
                 0.457    0.368    0.483      0.457    0.470      0.090    0.562     0.494     Price Down
Weighted Avg.    0.557    0.468    0.554      0.557    0.555      0.090    0.562     0.571

=== Confusion Matrix ===

  a  b   <-- classified as
 79 46 |  a = Price Up
 51 43 |  b = Price Down
```

The performance of the model is not good compared to the models that were run earlier. According to (Brownlee, 2021), “This could be explained by the fact that, on classification problems, the stacking ensemble often performs better when base-models are configured to predict probabilities instead of defined class labels. The added uncertainty in the predictions would provide more context for the model when learning the best way to combine the predictions”.

4.7.3. AdaBoost

(Saini, 2021) defines Boosting as “An ensemble modelling technique that was first presented by Freund and Schapire in the year 1997. Since then, Boosting has been a prevalent technique for tackling binary classification problems. These algorithms improve the prediction power by converting several weak learners to strong learners”.

The principle behind boosting algorithms is simple and can be summarized in the steps below:

1) Using the training dataset, build a model.
2) Build another model to address the errors from the first model.
3) Continue the process until errors are minimal, and the output is correctly predicted.

AdaBoost is an ensemble Machine Learning technique. AdaBoost mainly uses Decision Stumps which are decision trees with only one split.

AdaBoost works by first building a model where all the data points have equal weights. Wrongly classified points are then given higher weights and are accordingly assigned more importance in the next model. AdaBoost keeps training the models until less errors are received.

Figure 4.20 shows the output of running the stacking algorithm in Weka, followed by an analysis and discussion of the results.

Figure 4.20
AdaBoost ML Run Results

```
Correctly Classified Instances         143               65.2968 %
Incorrectly Classified Instances        76               34.7032 %
Kappa statistic                          0.2823
Mean absolute error                      0.472
Root mean squared error                  0.4834
Relative absolute error                 95.6496 %
Root relative squared error             97.5355 %
Total Number of Instances              219

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
                 0.736    0.457    0.681      0.736    0.708      0.284    0.631     0.651
                 0.543    0.264    0.607      0.543    0.573      0.284    0.631     0.541
Weighted Avg.    0.653    0.374    0.650      0.653    0.650      0.284    0.631     0.604

=== Confusion Matrix ===

  a  b   <-- classified as
 92 33 |  a = Price Up
 43 51 |  b = Price Down
```

AdaBoost provides the highest accuracy of all the non-ensemble and ensemble techniques examined. The accuracy from running AdaBoost is 65.2% which is higher than

all the accuracies achieved from running the different ML algorithms in this praxis. Furthermore, the accuracy from running this model far exceeds the results obtained from the examined research papers, setting potentially a new benchmark for predicting the direction of the Bitcoin next day price. Worth noting that while the underlying algorithm used for sentiment analysis (FinBERT), in addition to using financial and business news from major providers on Twitter are important factors in predicting the direction of Bitcoin next day price, the selection of the ML algorithm is an equally important factor in improving the prediction accuracy.

4.8. Summary

Two sentiment analysis algorithms, BERT and FinBERT were used to classify the sentiment polarity for each tweet. Subsequently, four different daily sentiment scores were calculated using the output of BERT and FinBERT and a variation of polarity-based and strength-based calculations. Each of the sentiment scores was merged with other Bitcoin attributes (daily trading volume, number of likes, number of retweets and Bitcoin next day price direction) to form the praxis dataset. Feature selection algorithms using Weka and Minitab were run to determine the relevant features to predicting the Bitcoin next day price direction. Based on the analysis of the results, the following features were deemed relevant (daily sentiment score, trading volume, number of likes, number of retweets) with daily sentiment score being the most significant feature. Furthermore, the FinBERT-Polarity-based daily sentiment score yielded the highest level of accuracy and accordingly was selected to subsequently validate hypothesis 2 and hypothesis 3. Two machine learning algorithms were run to validate hypothesis 2(Naive Bayes and DecisionStump). The results were analyzed, and Decision Stump achieved the highest prediction accuracy. Finally, ensemble machine learning algorithms were run (Stacking and AdaBoost) to validate hypothesis 3. AdaBoost yielded the highest prediction accuracy of 65.2% which far exceeds the accuracy documented in the research body of work.

A detailed discussion of this result, contribution to the body of knowledge and potential future work is provided in chapter 5.

CHAPTER 5:

DISCUSSION AND CONCLUSION

5.1. Discussion

Predicting Bitcoin prices has been a critical research area in the academic literature in the last few years. This is due to the quick rise of cryptocurrency in general and Bitcoin in particular. Driven by the large and continuous fluctuations in price as well as the ease of trading, there has not been any other asset that had as much potential for returns as Bitcoin.

Today, much of the research work that is focused on predicting Bitcoin prices relies on social media platforms. In the last few years, social media platforms have evolved from a medium for information exchange, to a virtual gathering place, to becoming serious players in the retail and marketing world. Along the way, social media platforms became a good source of information about pretty much any topic of interest.

With the wealth of information available on social media platforms, applying sentiment analysis techniques to the text posted on social media platforms has become an increasingly popular way to determining the sentiments of the community towards a particular topic of interest.

Of all the social media platforms that are available today, Twitter has surfaced out as one of the most popular ones. The 280-character length limitation imposed by Twitter forces the posters to be succinct and concise. This means that not only important information gets conveyed, but also the information gets to be conveyed quickly.

Much of the research work on predicting Bitcoin prices has focused on applying sentiment analysis to all or a sample of the posts related to the topic. While there have

been some successes, most of the research has an inherent limitation of either using the public sentiments or a limited dataset. Unlike other topics of interest, where Twitter is a good proxy for the public opinion at large, that is not the case for Bitcoin. When analyzing a sample of Bitcoin tweets, one can quickly conclude that the Bitcoin posts are predominantly positive. For many, Bitcoin is more than just an asset. The success of Bitcoin fosters a philosophy that breaks away from the traditional models of centralized authority to a trustless system where the actors do not need to trust each other for the system to function. As such, it is no surprise that those who believe in Bitcoin are very active on social media and tend to portray a positive outlook of the crypto currency irrespective of the actual performance at a given point of time.

This praxis started with a simple hypothesis that applying advanced sentiment analysis techniques to Twitter news features, as opposed to Twitter public posts, is a determining factor in improving the accuracy of the Bitcoin price direction prediction. In the process of developing the praxis, this hypothesis, and the supporting hypotheses were validated along the way. The combination of business and financial news posted on Twitter, together with using some of the most advanced sentiment analysis techniques like BERT and FinBERT led to achieving a much higher accuracy than what has been published in research journals.

5.2. Conclusion

Many researchers have attempted to predict the Bitcoin next day price direction. While some relied on the blockchain network-based features to predict the price of Bitcoin, others used other indicators like historical price data, mining difficulty and hash rate. Some relied on applying sentiment to news by crawling the websites of news

providers while others relied on public opinion to predict the price of Bitcoin. Along the way, as the sentiment analysis techniques continued to advance and mature, different researchers experimented with different techniques to further improve the sentiment classification and accordingly the Bitcoin price prediction.

This paper examined and benefited from many of the published work and combined the best of the learnings from those works. The praxis developed a unique way to combine the ability of financial and business news to reflect the true sentiments towards Bitcoin with the conciseness and speed of delivering the news through the Twitter platform. Furthermore, by implementing some of the most advanced NLP techniques for sentiment analysis and experimenting with a variety of machine learning algorithms, we were able to achieve unprecedented accuracy in predicting Bitcoin next day price direction. The accuracy of 65.2% achieved compared to accuracy in published papers is an attestation and further confirmation to the uniqueness and validity of the hypotheses advanced in this paper.

5.3. Contribution to the Body of Knowledge

This praxis set a new benchmark for measuring the performance of Bitcoin next day price direction models, achieving a prediction accuracy of 65.2%. The key contributions to the body of knowledge are summarized below:

- Advancing the idea of relying on financial and business news as opposed to public opinion to determine the sentiments towards Bitcoin.
- Using Twitter as a source for business and financial news as opposed to news websites or news aggregators that have the inherent limitation of slow updates driven by the need for the story to be fully developed before it gets published.

- Combining the relevance of the news and Twitter ability to communicate the news quickly with the power of advanced sentiment analysis techniques and machine learning algorithms to achieve higher accuracy of the Bitcoin price direction prediction.

It's worth noting, that while the research work presented in this paper was focused on predicting Bitcoin next day price direction, the findings and learnings can be applied to many other areas that could benefit from applying advanced sentiment analysis and machine learning techniques to news expressed on the Twitter platform.

5.4. Recommendation for Future Work

The primary goal of this praxis is to improve the prediction accuracy of the Bitcoin next day price direction. To date, the accuracy results documented in the different research publications fall short of 60%. This praxis achieved an accuracy result of 65.2% which constitutes a major improvement to the earlier published results. While this is a great accomplishment, there are always opportunities for future work. Summarized below are some of the areas that could be further explored in the future to improve the Bitcoin next day price direction predictions.

- Expand the analysis to include business and financial news from non-English speaking providers. Bitcoin is being traded globally, so adding the non-English speaking major news providers to the analysis might result in a better representation of the sentiments overall and accordingly might lead to further improvement in the Bitcoin next day price direction prediction.

- Explore the use of more advanced sentiment analysis techniques like ROBERTA to improve the sentiment analysis accuracy and accordingly the Bitcoin next day price direction prediction.
- Consider using the tweet sentiment polarity probability from FinBERT instead of relying on sentiment polarity only. The sentiment polarity probability might be a better reflection of the tweet tone and accordingly is something that is worth exploring for future work.

I hope that the work presented here not only helps investors to maximize their profit returns, but more importantly inspires other researchers to take the learnings from this praxis and consider the recommendations presented to further advance the knowledge in this space.

REFERENCES

Ghorbani, M., Bahaghighat, M., Xin, Q., & Özen, F. (2020). ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing. *Journal of Cloud Computing:Advances, Systems and Applications*.

Abraham, J., & Higdon, D. W. (2018). *Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis*. Retrieved from SEMANTIC SCHOLAR: https://www.semanticscholar.org/paper/Cryptocurrency-Price-Prediction-Using-Tweet-Volumes-Abraham-Higdon/2a865d7bfcde586df5c142a5876c1e698726533a

Ainsworth, S., & Grant, D. (2012). Revitalizing scholarship in identity studies. *Scandinavian Journal of Management, 28*, 60—62.

Alvesson, M., & Sveningsson, S. (2003). Good Visions, Bad Micro-management and Ugly Ambiguity: Contradictions of (Non-)Leadership in a Knowledge-Intensive Organization. *Organization Studies, 24*(6), 961-988.

Alvesson, M., Ashcraft, K. L., & Thomas, R. (2008). Identity matters: Reflections on the construction of identity scholarship in organization studies. *Organization, 15*(1), 5-28.

Ankalkoti, P. (2017, May). *A Relative Study on Bitcoin Mining*. Retrieved from https://www.researchgate.net/: https://www.researchgate.net/publication/318850089_A_Relative_Study_on_Bitcoin_Mining

Antonioni, D. (1996, December). Two strategies for responding to stressors: Managing conflict and clarifying work expectations. . *J. Business and Psychology, 11*(2), 287-295.

Ashford, S. J., & Black, J. S. (1996). Proactivity during organizational entry: The role of desire for control. *j. Applied Psychology, 81*, 199-214.

Ashforth, B. E. (2001). *Role Transitions in Organizational Life: An Identity-Based Perspective.* New York: Routledge.

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review, 14*, 20-39.

Ayuya, C. (2020, Dec 13). *Ensemble Learning Techniques to Improve Machine Learning*. Retrieved from Section: https://www.section.io/engineering-education/ensemble-learning/#:~:text=Ensemble%20learning%20refers%20to%20a,learner%20predicts%20poorly%20when%20alone.

Baldwin, R. (2014, June). WTO 2.0: Governance of 21st century trade. *The Review of International Organizations, 9*(2), 261-283.

Balogun, J., & Johnson, G. (2005, November). From intended strategies to unintended outcomes: The impact of change recipient sensemaking. *Organization Studies, 26*(11), 1573-1601.

Bartunek, J. M., Rousseau, D. M., Rudolph, J. W., & DePalma, J. A. (2006, June). On the Receiving End: Sensemaking, Emotion, and Assessments of an Organizational Change Initiated by Others. *The Journal of Applied Behavioral Science, 42*(2), 182-206.

Battilana, J., Gilmartin, M., Sengul, M., Pache, A., & Alexander, J. (2010). Leadership competencies for implementing planned organizational change. *The Leadership Quarterly, 21*, 422-438.

Beck, T. E., & Plowman, D. A. (2009, October). Experiencing rare and unusual events richly: The role of middle managers in animating and guiding organizational interpretation. *Organization Science, 20*(5), 909-924.

Boyles, T. (2012). 21st century knowlledge, skills, and abilities and entrepreneurial competencies: A model for undergraduate entrepreneurship education. *Journal of Entrepreneurship Education, 15*, 41-55.

Brinkley, I. (2014). *Defining the knowledge economy; Knowledge economy programme report.* The Work Foundation. London: The Work Foundation.

Brownlee, J. (2020, April 10). *Stacking Ensemble Machine Learning With Python*. Retrieved from Machine Learning Mastery: https://machinelearningmastery.com/stacking-ensemble-machine-learning-with-python/

Brownlee, J. (2021, April 23). *Essence of Stacking Ensembles for Machine Learning*. Retrieved from Machine Learning Mastery: https://machinelearningmastery.com/essence-of-stacking-ensembles-for-machine-learning/

Burke, W. (2014). *Organization Change, 4th Ed.* thousand Oaks, CA, USA: Sage.

Burke, W., & Litwin, G. (1992). A causal model of organizational performance and change. *Journal of Management, 18*(3), 532-545.

Burrell, G., & Morgan, G. (1979). *Sociologoical Paradigms and Organisational Analysis.* Hants, England: Ashgate Publishing Ltd.

BYTWORK.COM. (2022). *Bitcoin price chart for the entire history from 2008 to 2022*. Retrieved from BYTWORK.COM: https://bytwork.com/en/articles/btc-chart-history

Caldwell, R. (2005, February). Things fall apart? Discourses on agency and change in organizations . *Human Relations, 58*(1), 83-114.

Cartwright, S., & Cooper, C. L. (1993, March). The psychological impact of merger and acquistion on the individual: A study of building society managers. *Human Relations, 46*(3), 327-340.

Cartwright, S., Teerikangas, S., Rouzies, A., & Wilson-Evere, E. (2012). Methods in M&A——A look at the past and the future to forge a path forward . *Scandinavian Journal of Management, 28*, 95-106.

Cerulo, K. (1997). Identity construction: New issues, new directions. *Annual Review of Sociology, 23*, 385-409.

Chen , Y., Giudici, P., Misheva, B., & Trimborn , S. (2020). Lead Behaviour in Bitcoin Markets. *Risks*, Volume 8 Issue 1.

CityFalcon. (2018). *Twitter, and The Importance of Financial Tweets and News to Investors and Traders*. Retrieved from CITYFALCON: https://cityfalcon.zohosites.com/twitter-tweets-important-for-financial-investors-traders

CityFALCON. (n.d.). *Twitter, and The Importance of Financial Tweets and News to Investors and Traders*. Retrieved from CityFALCON:

https://cityfalcon.zohosites.com/twitter-tweets-important-for-financial-investors-traders

CodeTD. (2020, October 14). *Natural language processing of BERT*. Retrieved from CodeTD: https://www.codetd.com/en/article/11892537

Conway, E., & Monks, K. (2011). Change from below: the role of middle managers in mediating paradoxical change. *Human Resource Management Journal , 21*(2), 190-203.

Crawford, L., & Nahmias, A. H. (2010). Competencies for managing change . *International Journal of Project Management , 28*, 405-412.

Creswell, J. (2014). *Research Design; Qualitative, Qantitative, and Mixed Methods Approaches* (4th ed.). thousand Oaks, CA: Sage Publications, Inc.

Crossley, N. (2008, April). Small-world networks, complex systems and sociology. *Sociology, 42*(2), 261-277.

Currie, G., & Procter, S. J. (2005). The antecedents of middle managers' strategic contribution: The case of a professional bureaucracy. *Journal of Management Studies, 42*(7), 1325-1356.

Curry, B. K. (2003). Organizational flux and its destabilizing influence on employee identity . *Management Decision, 41*(6), 558-569.

Da Palma, P., Lopes, M. P., & Soares, A. E. (2012). Relationship among identity, image and construed external image: A missing link on acquisitions. *International Business Research, 5*(10), 115-130.

Dawson, P., & Buchanan , D. (2005). The way it really happened: Competing narratives in the political process of technological change. *Human Relations , 58*(7), 845–865.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *SEMANTIC SCHOLAR*.

Dobilas, S. (2021, Jan 30). *CART: Classification and Regression Trees for Clean but Powerful Models*. Retrieved from https://towardsdatascience.com/: https://towardsdatascience.com/cart-classification-and-regression-trees-for-clean-but-powerful-models-cc89e60b7a85

Dube, L. (2014). Exploring how IT professionals experience role transitions at the end of successful projects. *Journal of Management Information Systems, 31*(1), 17-45.

Durgapal, A., & Vimal, V. (2021). Prediction of Stock Price Using Statistical and Ensemble learning Models: A Comparative Study. *2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON).* Dehradun, India: IEEE.

Dwyer, G. P. (2014). The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*.

Epstein, S. (1980). The self concept: A review and the proposal of an integrated theory of personality. In E. Staub, *Personality: Basic aspects and current research* (pp. 81-132). Englewood Cliffs, NJ, USA: Prentice-Hall.

Ericson, T. (2001). Sensemaking in organisations towards a conceptual framework for understanding strategic change. *Scandinavian Journal of Management, 17*, 109-131.

Forster, E. M. (1927/2010). *Aspects of the Novel.* New York: RosettaBooks LLC.

Gangjin, W., Yanping, T., Chi, X., & Shou, C. (September 2019). Is bitcoin a safe haven or a hedging asset? Evidence from China. *Journal of Management Science and Engineering*, Volume 4, Issue 3.

Genc, Z. (2021, July 31). *FinBERT: Financial Sentiment Analysis with BERT*. Retrieved from medium.com: https://medium.com/prosus-ai-tech-blog/finbert-financial-sentiment-analysis-with-bert-b277a3607101

Giddens, A. (1984). *The Constitution of Society: Outline of the Theory of Structuration.* Berkeley, CA: University of California Press.

Gilboa, S., Shirom, A., Fried, Y., & Cooper, C. (2008). A meta-analysis of work demand stressors and job performance: Examining main and moderating effects. *Personnel Psychology, 61*, 227–271.

Gioia, D. A., & Chittipeddi, K. (1991, September). Sensemaking and Sensegiving in Strategic Change Initiation. *Strategic Management Journal, 12*(6), 433-448.

Goncharov, I. (2022, Jan 18). *Financial Sentiment Analysis on Stock Market Headlines With FinBERT & Hugging Face*. Retrieved from wandb.ai: https://wandb.ai/ivangoncharov/FinBERT_Sentiment_Analysis_Project/reports/Financial-Sentiment-Analysis-on-Stock-Market-Headlines-With-FinBERT-Hugging-Face--VmlldzoxMDQ4NjM0

Graduate School of Education and Human Development, G. (2016). *Doctoral Handbook 2016-2017 Edition.* Washington, DC: The George Washington University, Graduate School of Education and Human Development.

Greaves, A., & Au, B. (2015). *Using the Bitcoin Transaction Graph to Predict the Price of Bitcoin*. Retrieved from SEMANTIC SCHOLAR: https://www.semanticscholar.org/paper/Using-the-Bitcoin-Transaction-Graph-to-Predict-the-Greaves-Au/a0ce864663c100582805ffa88918910da89add47

Greenhalgh, T., Stones, R., & Swinglehurst , D. (2014). Choose and Book: A sociological analysis of 'resistance' to an expert system. *Social Science & Medicine , 104*, 210-219.

Grunberg, L., Moore, S., Greenberg, E. S., & Sikora, P. (2008, June). The changing workplace and its effects: A longitudinal examination of employee responses at a large company. *Journal of Applied Behavioral Sciences, 44*(2), 215-236.

Guiot, J. (1977, October). Attribution and identity construction: Some comments. *American Sociological Review , 42*, 692-704.

Guo, T. (2022, March 24). *Learning From Human Correction For Data-Centric Deep Learning*. Retrieved from arxiv.org: https://arxiv.org/pdf/2102.00225.pdf#:~:text=There%20are%20two%20steps%20in,and%20next%20sentence%20prediction%20task.

Gurrib, I., & Kamalov , F. (2021, December 15). *Predicting bitcoin price movements using sentiment analysis: a machine learning approach*. Retrieved from emerald insight: https://www.emerald.com/insight/content/doi/10.1108/SEF-07-2021-0293/full/html

Hafeez-Baig, A. (2012, November 1). Does Information and Communication Technology (ICT) Facilitate Knowledge Management Activities in the 21st Century? *Journal of Software, 7*(11), 2437-2442.

HASNAIN, M., PASHA, M. F., GHANI, I., IMRAN, M., ALZAHRANI, M. Y., & BUDIARTO, R. (2020). Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking. *IEEEAccess*.

Heston, T. (2018). *Introductory Chapter: Making Health Care Smart.* IntechOpen.

Hogg, M. A., Terry, D. J., & White, K. M. (1995, December). A tale of two theories: A critical comparison of identity theory with social identity theory. . *Social Psychology Quarterly, 58*(4), 255-269.

Holden, L., & Roberts, I. (2004). The depowerment of European middle managers: Challenges and uncertainties. *Journal of Managerial Psychology, 19*(3), 269-287.

Holt, I. (2008). Role transition in primary care settings . *Quality in Primary Care, 16*, 117-126.

Hsu, R. (2021, Feb 25). *Decision Stump*. Retrieved from medium.com: https://medium.com/geekculture/decision-stump-b8e93c1f54d7

Huahu, X., Ahmad, W., Shabir, M. Y., & Ahmad, W. (2019). A Sentiment Polarity Categorization Technique for Online Product Reviews. *IEEE Access*, Volume: 8.

Huff, A. (2009). *Designing Research for Publication.* Thousand Oaks, CA: Sage Publications, Inc.

Ibarra, H., & Barbulescu, R. (2010, January). Identity as narrative: Prevalence, effectiveness, and consequences of narrative identity work in maco work role transitions. *Academy of Management Review, 35*(1), 135-154.

Ibarra, H., & Barbulescu, R. (2010). Identity as narrative: Prevalence, effectiveness, and consequences of narrative identity work in macro work role transitions. *Academy of Management Review, 35*(1), 135-154.

Ibarra, H., & Petriglieri, J. (2010). Identity work and play. *Journal of Organizational Change Management, 23*(1), 10-25.

Illeris, K. (2014). *Transformative Learning and Identity.* Oxon, UK: Routledge.

Indurthi, V., Syed, B., Gupta, M., & Varma, V. (2020). Predicting Clickbait Strength in Online Social Media. *28th International Conference on Computational Linguistics* (pp. 4835–4846). Barcelona. Spain: researchgate.net.

Insead, H. I., & Barbulescu, R. (2010). Identity as narrative: Prevalence, effectiveness, and consequences of narrative identity work in macro work role transitions. *Academy of Management Review, 35*(1), 135-154.

Iqbal, N., Mumtaz, R., Shafi, U., & Zaidi, S. (2021). Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. *PeerJ Computer Science*.

JAGANNATH, N., BARBULESCU1, T., SALLAM, K. M., ELGENDI1, I., OKON, A. A., MCGRATH3, B., . . . MUNASINGHE, K. (2021). A Self-Adaptive Deep Learning-Based Algorithm for Predictive Analysis of Bitcoin Price. *IEEE Access*, Volume 9.

Jang, H., & Lee, J. (2017). An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information. *IEEE*, Volume: 6.

Jeon, H. S., Oh, R., & Lee, J. D. (2020). Forecasting Bitcoin Price Fluctuation by Twitter Sentiment Analysis. *International Conference on Information Science and Communications Technologies (ICISCT).* ashkent, Uzbekistan: IEEE.

Joshi, P. (2020, July 21). *Transfer Learning for NLP: Fine-Tuning BERT for Text Classification*. Retrieved from Analytics Vidhya: https://www.analyticsvidhya.com/blog/2020/07/transfer-learning-for-nlp-fine-tuning-bert-for-text-classification/

Joshi, P. (2020, July 21). *Transfer Learning for NLP: Fine-Tuning BERT for Text Classification*. Retrieved from Analytics Vidhya: https://www.analyticsvidhya.com/blog/2020/07/transfer-learning-for-nlp-fine-tuning-bert-for-text-classification/

Kana, M. (2019, Sep 14). *BERT for dummies — Step by Step Tutorial*. Retrieved from towardsdatascience.com/: https://towardsdatascience.com/bert-for-dummies-step-by-step-tutorial-fb90890ffe03

Karalevičius, V., Degrande, N., & De Weerdt, J. (2017). Using sentiment analysis to predict interday Bitcoin price movements. *The Journal of Risk Finance*.

Kauppila, O.-P. (2014, July). So, what am I supposed to do? A multilevel examination of role clarity. *Journal of Management Studies, 51*(5), 737-763.

Kira, M., & Balkin, D. (2014). Interactions between work and identities: Thriving, withering, or redefining the self? *Human Resources Management Review, 24*, 131-143.

Kumar, A. (2021, August 31). *Analysing crypto news sentiment to predict bitcoin prices*. Retrieved from SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3913652

Lamon, C., Nielsen, E., & Redondo, E. (2017). *Cryptocurrency Price Prediction Using News and Social Media Sentiment*. Retrieved from SEMANTIC SCHOLAR:

https://www.semanticscholar.org/paper/Cryptocurrency-Price-Prediction-Using-News-and-Lamon-Nielsen/c3b80de058596cee95beb20a2d087dbcf8be01ea

Lewin, K. (1997[1947]). Frontiers in Group Dynamics. In K. Lewin, & G. W. Lewin (Ed.), *Resolving Social Conflicts: Selected Papers on Grouyp Dynamics* (pp. 301-336). Washington, DC: American Psychological Association.

Lichtenstein, B. (2010). Introduction: A New Science Of Leadership . In J. Goldstein, J. Hazy, & B. Lichtenstein, *Complexity and the Nexus of Leadership; Leveraging Nonlinear Science to Create Ecologies of Innovation* (pp. 1-14). New York: Palgrave MacMillan Press .

Liu, Z., Huang, D., Huang, K., Li, Z., & Zhao, J. (2020). FinBERT: A Pre-trained Financial Language Representation Model for Financial Text Mining. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20).* Yokohama: IJCAI.

Luo, J. (2020). Bitcoin price prediction in the time of COVID-19. *2020 Management Science Informatization and Economic Innovation Development Conference (MSIEID).* Los Angeles, CA.

Lutz, F. W. (1982). Tightening up loosely coupling in organizations of higher education. *Administrative Science Quarterly, 27*, 653-669.

Mai, F., & Bai, Q. (January 2015). From Bitcoin to Big Coin: The Impacts of Social Media on Bitcoin Performance. *SSRN Electronic Journal*.

Maiya, A. (2020, Jan 14). *Text Classification with Hugging Face Transformers in TensorFlow 2 (Without Tears)*. Retrieved from Towards Data Science:

https://towardsdatascience.com/text-classification-with-hugging-face-transformers-in-tensorflow-2-without-tears-ee50e4f3e7ed

Mallett, O., & Wapshott, R. (2012). Mediating ambiguity: Narrative identity and knowledge workers. *Scandinavian Journal of Management, 28*(1), 16-26.

Manimuthu, A., Sreedharan, R., G., R., & Marwaha, D. (2019). A Literature Review on Bitcoin: Transformation of Crypto Currency Into a Global Phenomenon. *IEEE ENGINEERING MANAGEMENT REVIEW,*, VOL. 47, NO. 1.

Matwin, S., Milios, A., Pralat, P., & Soares, A. (2021, December). *Survey of Generative Methods for Social Media Analysis*. Retrieved from researchgate.net: https://www.researchgate.net/publication/357046850_Survey_of_Generative_Methods_for_Social_Media_Analysis

McCann, L., Morris, J., & Hassard, J. (2008, March). Normalized intensity: The new labour process of middle management. *Journal of Management Studies, 45*(2), 343-371.

McConville, T. (2006). Devolved HRM responsibilities, middle-managers and role dissonance. *Personnel Review, 35*(6), 637-653.

McNall, S., Roch, J., & Caton, S. (2018). Predicting the Price of Bitcoin Using Machine Learning. *26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing.* Dublin: IEEE.

McNatt, D. B., & Judge, T. A. (2008). Self-efficacy intervention, job attitudes, and turnover: A field experiment with employees in role transition . *Human Relations, 6*(6), 783-810.

Mead, G. H. (1961a [1934]). The I and the Me. In T. Parsons (Ed.), *Thoeries of Society: Foundations of Modern Socialological Theory* (pp. 163-168). New York: The Free Press.

Mead, G. H. (1961b [1934]). Taking the role of the Other. In T. Parsons (Ed.), *Theories of Society: Foundations of Modern Sociolological Theory* (pp. 739-740). New York: The Free Press.

Mead, G. H. (1961c [1934]). Internalized Others and the Self. In T. Parsons (Ed.), *Theories of Society: Foundations of Modern Sociological Theory* (pp. 829-830). New York: The Free Press.

Merriam, S. (2009). *Qualitative research; a guide to design and implementation.* San Francisco: Jossey-Bass.

Minitab. (2021). *Relative variable importance chart for CART® Classification*. Retrieved from Minitab: https://support.minitab.com/en-us/minitab/20/help-and-how-to/statistical-modeling/predictive-analytics/how-to/cart-classification/interpret-the-results/all-statistics-and-graphs/relative-variable-importance-chart/

Moustakas, C. (1994). *Phenomenological Research Methods.* Thousand Oaks, CA, US: Sage.

Nakkula, M., & Ravitch, S. (1998). *Matters of Interpretation; Reciprocal Transformation in Theraputic and Developmental Relationships with Youth.* San Fransisco, CA: Jossey-Bass.

Nandwani, P., & Verma, R. (Aug 2021). A review on sentiment analysis and emotion detection from text. *Nature Public Health Emergency Collection*, 10.

Nicholson, N. (1984, June). A Theory of Work Role Transitions. *Administrative Science Quarterly, 29*(2), 172-191.

Niessen, C., Binnewies, C., & Rank, J. (2010). Disengagement in work-role transitions. *J. Occupational and Organizational Psychology , 83*, 695-715.

Ojha, A. K. (2020). *USE A CLASSIFICATION AND REGRESSION TREE (CART) FOR QUICK DATA INSIGHTS*. Retrieved from iSixSigma : https://www.isixsigma.com/methodology/lean-methodology/use-a-classification-and-regression-tree-cart-for-quick-data-insights/

Olena Bondarenko, O. K. (2019). *THE POSSIBILITIES OF USING INVESTMENT TOOLS BASED ON CRYPTOCURRENCY IN THE DEVELOPMENT OF THE NATIONAL ECONOMY*. Retrieved from https://pdfs.semanticscholar.org/680c/259b864721a6a4dda6a2be43a539042e1058.pdf

Olsen, T. H., & Stensaker, I. (2014, March). A change-recipient perspective on training during organizational change. *International Journal of Training and Development, 18*(1), 22-36.

P.Dwyer, G. (2015, April). *The economics of Bitcoin and similar private digital currencies*. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S1572308914001259?via%3Dihub.

Panchal, S., & Cartwright, S. (2001). Group differences in post-merger stress. *J. Managerial Psychology, 16*(6), 424-433.

Parris, M. A., Vickers, M. H., & Wilkes , L. (2008). Caught in the Middle: Organizational Impediments to Middle Managers' Work-life Balance. *Employee Resoinsibilities and Rights, 20*, 101-117.

Partida, A., Gerassis, S., Criado, R., Romance, M., Giráldez, E., & Taboada, J. (2022). Modeling Bitcoin plus Ethereum as an Open System of Systems of Public Blockchains to Improve Their Resilience against Intentional Risk. *electronics*.

Patel, A. (2021, July 21). *Support Vector Machines*. Retrieved from medium.com: https://medium.com/analytics-vidhya/support-vector-machines-805a24507f#:~:text=The%20goal%20of%20the%20SVM,as%20the%20Widest%20Street%20Approach.

Paul, G. W., & Berry, D. M. (2013). The importance of executive leadership in creating a post-merged organisational culture conducive to effective performance. *SA J. of Human Resource Management, 11*(1), 446-460.

Peltonen, T. (2007). In the middle of managers; Occupational communities, global ethnography and the multinationals . *Ethnography, 8*(3), 346-360.

Phillips, D. (2021, January 19). *How Many Bitcoin Billionaires Are There? (2021)*. Retrieved from Decrypt: https://decrypt.co/32694/how-many-bitcoin-billionaires-are-there

Qin, A., & Livni, E. (2021). China Cracks Down Harder on Cryptocurrency With New Ban. *The New York Times*, 1.

Questier, F. (2005). *The Use of CART and Multivariate Regression Trees for Supervised and Unsupervised Feature Selection*. Retrieved from researchgate.net: https://www.researchgate.net/publication/223724526_The_Use_of_CART_and_

Multivariate_Regression_Trees_for_Supervised_and_Unsupervised_Feature_Selection

Quiroz-GutIerrez, M. (2021). *A minuscule .01% of Bitcoin holders control nearly a third of the supply*. FORTUNE.

Rajasekharan, A. (2019, June 17). *A review of BERT based models*. Retrieved from https://towardsdatascience.com/: https://towardsdatascience.com/a-review-of-bert-based-models-4ffdc0f15d58

Ray, S. (2017, Sept 11). *6 Easy Steps to Learn Naive Bayes Algorithm with codes in Python and R*. Retrieved from Analytics Vidhya: https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/

Reid, F., & Harrigan, M. (2011, July 22). *An Analysis of Anonymity in the Bitcoin System*. Retrieved from arxiv: https://arxiv.org/abs/1107.4524

Reinicke, C. (2021, Feb 12). *Bitcoin hits another record. Here's how much you'd have if you invested $100 in 2009*. Retrieved from cnbc.com: https://www.cnbc.com/2021/02/12/how-much-youd-have-today-if-you-invested-100-in-bitcoin-in-2009.html

Ritchie, J., Lewis, J., McNaughton Nicolls, C., & Ormston, R. (Eds.). (2014). *Qualitative Research Practice, A Guide for Social Science Students and Researchers, 2nd Edition* (Second ed.). London, UK: Sage Publications.

Rizzo, J. R., House, R. J., & Lirtzman, S. I. (1970). Role conflict and ambiguity in complex organizations. *Administrative Science Quarterly, 15*, 150-163.

Romanelli, E., & Tushman, M. (1994, October). Organizational transformation as punctuated equilibrium: An empirical test. *Academy of Management Journal, 37*(5), 1147-1166.

Rothbaum, F., Weisz, J. R., & Snyder, S. S. (1982). Changing the world and changing the self: A two-process model of perceived control. *J. Personality and Social Psychology, 42*, 5-37.

Rouleau, L. (2005, November). Micro-practices of strategic sensemaking and sensegiving: How middle managers interpret and sell change every day. *Journal of Management Studies, 42*(7), 1413-1441.

Rouleau, L. (2010). Studying strategizing through narratives of practice. In D. Golsorkhi, L. Rouleau, D. Seidl, & E. Vaara (Eds.), *Cambridge Handbook of Strategy as Practice* (pp. 258-170). Cambridge , UK: Cambridge University Press.

Rouleau, L., & Balogun, J. (2011, July). Middle managers, strategic sensemaking, and discursive competence. *Journal of Management Studies, 48*(5), 953-983.

Saini, A. (2021, Sept 15). *AdaBoost Algorithm – A Complete Guide for Beginners*. Retrieved from Analytics Vidhya: https://www.analyticsvidhya.com/blog/2021/09/adaboost-algorithm-a-complete-guide-for-beginners/#:~:text=Boosting%20is%20an%20ensemble%20modelling,weak%20learners%20to%20strong%20learners.

Schein, E. H. (2010). *Organizational Culture and Leadership, 4th Ed.* . San Francisco, CA: Josey-Bass.

Schuman, D. (1982). *Policy Analysis, Education, and Everyday Life.* Lexington, MA: Heath.

Seidman, I. (2013). *Interviewing as Qualitative Research, Fourth Edition.* New York, NY: Teachers College Press.

Shaier, S. (2019, Feb 4). *ML Algorithms: One SD (σ)- Decision Trees Algorithms*. Retrieved from towardsdatascience.com: https://towardsdatascience.com/ml-algorithms-one-sd-%CF%83-decision-trees-algorithms-746e866ac3f

Shehhi, A. A., Oudah, M., & Aung, Z. (2014). Investigating Factors Behind Choosing a Cryptocurrency. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM).* Selangor Darul Ehsan, Malaysia: IEEE.

Silverman, D. (2010). *Interpreting Qualitative Data: Methods for Analysing Talk, Text, and Interaction, 4th Ed.* . London: Sage.

Simpson, B., & Carroll, B. (2008). Re-viewing ‘role’ in processes of identity construction. *Organization, 15*(1), 29-50.

Simpson, B., & Carroll, B. (2008). Re-viewing ‘role’ in processes of identity construction. *Organization, 15*(1), 29-50.

Simser, J. (2015). Bitcoin and modern alchemy: In code we trust. *Journal of Financial Crime*.

Smith, K., Collins, C. J., & Clark, K. D. (2005, April). Existing knowledge, knowledge creation capability, and the rate of new product introduction in high-technology firms. *Academy of Managment Journal, 48*(2), 346-357.

Smutny, Z., Sulc, Z., & Lansky, J. (2021). Motivations, Barriers and Risk-Taking When Investing in Cryptocurrencies. *MDPI*.

Sommers, M. (1994, October). The narrarive constitution of identity: A relational and network approach. *Theory and Society, 23*(5), 605-649.

Stones, R. (2005). *Structuration Theory*. New York: Palgrave Macmillan.

Swann, W. B. (1990). To be adored or to be known? the interplay of self-enhancement and self-verification. In E. T. Higgins, & R. M. Sorrentino, *Handbook of motivation and cognition* (Vol. 2, pp. 408-448). New York: Guilford.

Taylor, S. (1999). Making sense of revolutionary change: Differences in members' stories. *J. Organizational Change Management, 12*(6), 524-539.

Thomas, R., & Linstead, A. (2002). Losing the plot? Middle managers and identity. *Organization, 9*(1), 71-93.

Tocan, M. (2012, December). Knowledge based strategies for knowledge based organizations. *J. Knowledge Management, Economics and Information Technology*, *2*(6), 167-177.

Tubre, T. C., & Collins, J. M. (2000). Jackson and Schuler (1985) revisited: A meta-analysis of the relationships between role ambiguity, role conflict, and job performance. *Journal of Management, 26*, 155-169.

Tweets Chart. (2022). Retrieved from BitcoinInfoCharts: https://bitinfocharts.com/comparison/bitcoin-tweets.html#1y

Uhl-Bien, M., Marion, R., & McKelvey, W. (2007). Complexity Leadership Theory: Shifting leadership from the industrial age to the knowledge era. *The Leadership Quarterly, 18*, 298-318.

Umar, M., Su, C.-W., Abbas Rizvi, S., & Shao, X.-F. (2021, June). *Bitcoin: A safe haven asset and a winner amid political and economic uncertainties in the US?*

Retrieved from https://www.researchgate.net/publication/349572364_Bitcoin_A_safe_haven_asset_and_a_winner_amid_political_and_economic_uncertainties_in_the_US

Urry, J. (2004). Small worlds and the new social physics. *Global Networks, 4*(2), 109-130.

Van de Ven, A., & Poole, M. S. (1995, July). Explaining development nd change in organizations. *Academy of Management Review, 20*(3), 510-540.

Van Manen, M. (1990). *Researching Lived Experience.* London, Ontario, Canada: State University of New York Press.

Van Sell, M., Brief, A. P., & Schuler, R. S. (1981). Role conflict and role ambiguity: Integration of the literature and directions for future research. *Human Relations, 34*(1), 43-71.

Vella Critien, J., Gatt2, A., & Ellul, J. (2022). *Bitcoin price change and trend prediction through twitter sentiment and data volume*. Retrieved from Springer Open: https://jfin-swufe.springeropen.com/track/pdf/10.1186/s40854-022-00352-7.pdf

WANG, Y., LEI, C., & MIAO, D. (2021). S&I Reader: Multi-Granularity Gated Multi-Hop Skimming and Intensive Reading Model for Machine Reading Comprehension. *IEEEAccess*.

Watson, T. J. (2008). Managing identity: Identity work, personal predicaments and structural circumstances . *Organization, 15*(1), 121-143.

Weber, W. E. (2016). *A Bitcoin standard: Lessons from the gold standard.* Ottawa: Bank of Canada.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*, 1-19.

Weick, K. E. (1979). *The Social Psychology of Organizing, 2nd Edition.* Boston, MA: Addison-Wesley.

Weick, K. E. (1995). *Sensemaking in Organizations.* Thousand Oaks, CA: Sage Publications.

Wenger, E. (2000). Communities of Practice and Social Learning Systems. *Organization, 7*(2), 225-246.

Williams, D. (2017). *Cryptocurrency Compendium: A Reference for Digital Currencies.* Lulu.com.

Won, D., Lee, Y., Choi, H.-J., & Jung, Y. (2021). Contrastive Representations Pre-Training for Enhanced Discharge Summary BERT. *IEEE International Conference on Healthcare Informatics (ICHI).* Victoria, BC, Canada: IEEE.

Wright, C. (2009). Inside out? Organizational membership, ambiguity and the ambivalent identity of the internal consultant. *British Journal of Management, 20*, 309-322.

Yao, W., Xu, K., & Li, Q. (29 June-3 July 2019). Exploring the Influence of News Articles on Bitcoin Price with Machine Learning. *IEEE Symposium on Computers and Communications (ISCC).* Barcelona, Spain: IEEE.

Yi, Y., Christopher Siy UY, M., & Huang, A. (2020, July 9). *FinBERT: A Pretrained Language Model for Financial Communications*. Retrieved from arxiv: https://arxiv.org/abs/2006.08097

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., & Zhu, H. (2021). A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, Volume: 109, Issue: 1.

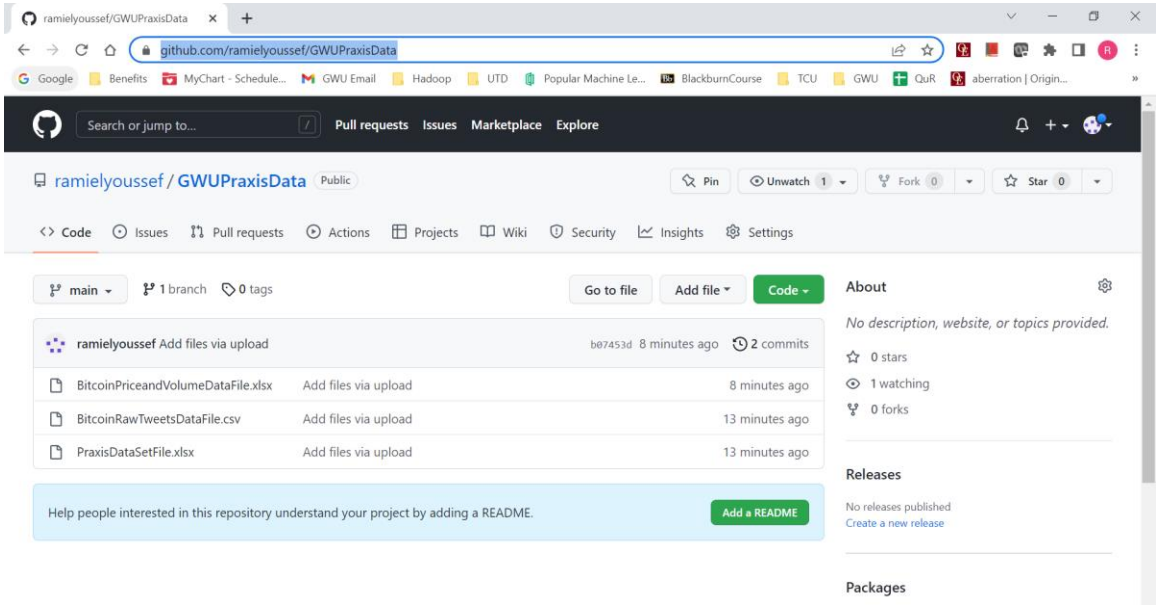
Ziaul Haque Munim, M. H. (2019). *Next-Day Bitcoin Price Forecast*. Retrieved from https://www.mdpi.com/1911-8074/12/2/103

APPENDIX A

Table A-1
Financial and Business News Twitter Accounts

Provider Name	Twitter Account
The Wall Street Journal	@wsj
Yahoo Finance	@yahoofinance
MarketWatch	@marketwatch
Reuters	@reuters
MSN	@msn
CNN	@cnn
Bloomberg	@business
CNBC	@cnbc
TheStreet	@thestreet
@NYT Business	@nytimesbusiness
Financial Times	@financialtimes
Guardian News	@guardiannews
BBC Business	@bbcbusiness
Guardian	@guardian
Forbes Tech	@forbestech

APPENDIX DATA REPOSITORY

The data used in this praxis is available at the following URL on GitHub:

https://github.com/ramielyoussef/GWUPraxisData