

Sentiment Analysis Trading Indicators

Exemplified by Natural Language Processing and Bitcoin

KEVIN SINGPURWALA

UNIVERSITY OF TWENTE

P.O. Box 217, 7500AE ENSCHEDE

THE NETHERLANDS

k.singpurwala@student.utwente.nl

Permission to make digital or hard copies of all or part of this work for personal or classroom use is hereby granted without a fee, provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and a full citation. To copy otherwise, republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. The 35th Twente Student Conference on IT 2021, Enschede, The Netherlands. ©Copyright 2021, University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science.

ABSTRACT

Trading with strategies based on market indicators has become commonplace, especially in the domain of automated trading. A wide variety of different indicators and heuristics are available.

In this research, we use deep learning via natural language processing to create sentiment driven trading indicators. We figure out if there is a correlation between market sentiment and bitcoin price movements using Pearson's correlation coefficient and how these sentiment driven indicators can be used to create trading strategies. The aforementioned sentiment driven trading strategies are then compared to other trading strategies.

Surprisingly, it was found that market sentiment does not correlate with bitcoin price movements. On the contrary, bitcoin related tweet volume (how trending bitcoin is a topic) did correlate with bitcoin price movements. The sentiment driven trading strategy implemented based on tweet volume was the most profitable strategy created. It outperformed simply buying and holding bitcoin by 72.37%. Thus it can be inferred that sentiment driven trading indicators can manifest profitable trading strategies.

Keywords

Bitcoin, Market sentiment, Trading indicators, Natural language processors, Classifiers, Sentiment analysis

1 Introduction

“IF CRYPTO SUCCEEDS, IT’S NOT BECAUSE IT EMPOWERS BETTER PEOPLE, IT’S BECAUSE IT EMPOWERS BETTER INSTITUTIONS” – Vitalik Buterin (#LearnCrpto)

1.1 Overview

Bitcoin has emerged as a fascinating phenomenon in the financial markets. Without any central authority issuing the currency, bitcoin has been associated with controversy. Ever since its popularity, bitcoin has been accompanied by increased public interest. Here, we contribute to the discussion by examining the potential drivers of bitcoin price movements, from fundamental sources to speculative and technical ones (Ladislav Kristoufek, 2015).[1]

There are several sentiments describe in literature which can impact bitcoin trading. One such sentiment is investor sentiment, which could play an important role. For instance, Baig et al. (2019), shows that investor sentiment of bitcoin can affect its return. On the other hand, López-Cabarcos et al. (2019), indicates that investor sentiment of stocks can predict bitcoin volatility. Other factors which can impact the price of cryptocurrencies include supply and demand, the cost of mining, government regulations, power of social media and economical crises. As per Gerritsen (2020), there are seven families of trading rules which were publicised prior to the emergence of bitcoin trading on exchanges.[2] They are the moving average (MA), trading range breakout, moving average convergence divergence (MACD), rate of change (ROC), on-balance volume (OBV), relative strength index (RSI) and Bollinger band method (BB).

It was noted in the literature that all factors, sentiments and statistical trading rules give an individual an indication of whether the price of cryptocurrency is going to go higher or lower, but none of them give a specified percentage for price movement.

The author would like to develop an algorithm which delivers the strength of indicators, considering market sentiment and trading rules impacting the price of cryptocurrency using natural language processing. The author also aims to compare the price of bitcoin against the profitability of sentiment driven trading using the devised algorithm. Later the author would like to provide conclusions and future improvements based on the findings of the research.

1.2 Research Question

Under what circumstance is sentiment analysis a profitable trading indicator?

1.2.1 Sub Questions

- (a) Does bitcoin market sentiment correlate to bitcoin market price?
- (b) What kind of sentiment based indicators trigger a profitable trade?

2 Review of Literature

2.1 Understanding Cryptocurrency

According to Wolfgang Karl Härdle, Campbell R Harvey and Raphael C G Reule (2020), cryptocurrency refers to a type of digital asset that uses a distributed ledger or blockchain technology to enable a secure transaction.[3] Although this technology is widely misunderstood, many central banks are considering to launch their own national cryptocurrency.

Yukun Liu Aleh Tsyvinski (2018), states that cryptocurrency is a recent phenomenon that is receiving significant attention. On the one hand it is based on a fundamentally new technology. The potential of which is not fully understood. On the other hand, at least in the current form, it fulfils similar functions to other more traditional assets. Is cryptocurrency a form of currency, a commodity, a stake in a technological breakthrough, or a completely different instrument? Can the price of a cryptocurrency be extrapolated by factors available for other asset classes? Which industries may be affected by the development of blockchain technology?[4]

According to Wei-Ying Nie et al (2020), cryptocurrency is attracting more and more people who are ready to invest. Cryptocurrency could be seen as another safe haven asset, due to its popularity in the financial market.[5]

2.2 Understanding Bitcoin

Bitcoin has emerged as the leading digital currency and an investment destination that offers investors a novel opportunity (Dyhrberg, 2016; Bouri et al., 2017).[6, 7]

According to Jordi Herrera-Joancomartí (2014), Bitcoin has emerged as the most successful cryptocurrency, since its appearance back in 2009.[8] Besides bitcoin's security robustness, two main properties have probably been its key to success: anonymity and decentralization. Bitcoin is an online virtual currency based on public key cryptography, proposed in 2008 in a paper authored by someone behind the pseudonym, Satoshi Nakamoto. Bitcoin became fully functional in January, 2009. Its broad adoption, facilitated by the availability of exchange markets allowing easy conversion with traditional currencies (EUR or USD), has brought it to be the most successful virtual currency. Bitcoin avails of a public ledger, so all transactions on the network are readily viewable.

2.3 Understanding Different Market Sentiments for Cryptocurrency

Researchers have begun studying the determinants of cryptocurrency indicators. For example, factors such as a cryptocurrency's technical indicators (e.g., Gerritsen et al., 2019) and the impact of media attention (e.g. Philippas et al., 2019) might be related to the performance of cryptocurrency markets.[2][9]

Moreover, macroeconomic factors such as political uncertainty and global economic activity might be associated with the performance of the cryptocurrency bitcoin. A lack of intrinsic valuation methods for bitcoin has led many traders to explore the predictive power of exogenous variables such as bitcoin popularity and attention (Kristoufek, 2015; Dastgir et al., 2019)[1, 10], trading volume (Blacilar et al., 2017)[11], hashing difficulty (Hayes, 2017)[12], other cryptocurrencies (Bouri et al., 2019; Ji et al., 2019)[13, 14], geopolitical risks (Aysan et al., 2019)[15], economic uncertainty (Demir et al., 2018)[16], stock market uncertainty (Bouri et al., 2017)[7], and energy/commodity prices (Hayes, 2017; Bouri et al., 2018).[12, 17] Importantly, bitcoin prices seem to not take an unpredictable path but to exhibit inefficiency instead (Tiwari et al., 2018)[18], possibly emanating from its short history and the irrational behaviours of its market participants (Bouri et al., 2019).[13]

Kristoufek (2015), classifies potential pricing variables into three categories: economic drivers, technical drivers and attractiveness (popularity).[1] Examples of economic drivers measuring the demand for a cryptocurrency are the ratio between trade and exchange transaction volume, the number of transactions excluding exchange transactions and estimated output volume and the total number of bitcoins in circulation. Examples of technical drivers are the hash rate and the difficulty of mining. Variables capturing the popularity of bitcoin as an investment are the number of Wikipedia searches and metrics reported by Google Trends. Kristoufek's (2015) study also includes a financial stress index and gold prices to assess the relationship between cryptocurrency and variables considered as safe havens under financial turmoil. The main conclusion in Kristoufek's (2015) study is that the price of bitcoin is mainly driven by investors' interest in the cryptocurrency and not by macro economic fundamentals or financial ratios.[1]

2.4 Techniques Used to Measure Market Sentiments

Following Gerritsen (2021), we discuss seven families of trading rules which were publicised prior to the dawn of exchanges facilitating bitcoin trading.[2] The first family of rules we test are moving averages. Brock et al. (1992: 1733) states that moving average (MA) rules belong to the "most popular technical rules" applied in practice.[19] Academically, these rules have gained considerable attention (e.g., Neely et al.).[20] MA rules generally issue a buy/sell signal when the actual price, or a relatively recent average price, exceeds (is below) a longer-term average of the stock price. As recent prices, we use the 1-day, 2-day and 5-day averages, and for longer-term averages we use the 50-day, 150-day and 200-day averages.

A second rule that belongs to the most widely used indicators (Brock et al., 1992), is the so-called trading range breakout, also known as support and resistance levels. This indicator signals minimum and maximum prices, respectively, for which a security has traded over the past n days. Following Brock et al. (1992) we apply 50, 150 and 200 days for n . [19] A buy signal is issued when the price of bitcoin exceeds the recent maximum. A sell signal is issued when it is below a recent minimum.

The third set of rules concerns moving average convergence divergence (Murphy, 1999). This rule is associated with three trading indicators. One follows from the moving average convergence divergence (MACD) itself, the others from the MACD signal line and the MACD histogram. The

fourth rule is the rate of change (ROC) (Taylor and Allen, 1992). This rule relates the current price to the price n days ago. A common time period used is 10 trading days. Fifth, on-balance volume (OBV) is the best-known indicator based on trading volume (Granville, 1963). The OBV indicator stipulates that volume precedes price changes. The assumption of the OBV indicator is that rising prices reflect positive volume pressure, which in turn can lead to higher prices.

As a sixth rule, we use the relative strength index (RSI). Wong et al. (2003) suggests that the RSI is the most frequently used counter trend indicator. The RSI is an oscillator with a level between 0 and 100. According to the RSI, a level higher than 70 usually indicates that bitcoin has risen but is now overbought (i.e., one should sell the stock). A level lower than 30 indicates the converse.

Seventh, and last, the second counter trend indicator is the Bollinger band method (BB). This rule is related to MA trading rules because the BB method contains a moving average, around which two bands are plotted (Bollinger, 2001). According to Lento et al. (2007), using a contrarian's approach, the BB method outperforms the buy-and-hold strategy.[21, 22]

2.5 Observations and Literature Analysis

However, prior studies regarding the efficiency of bitcoin, assess predictability based on a departure from a random walk (e.g., Tiwari et al., 2018), without offering any practical inferences for the sake of bitcoin traders.[18] Furthermore, the extent to which technical trading rules perform when applied to bitcoin prices remains unclear. Although bitcoin has frequently been discussed on various financial blogs and even in mainstream financial media, the research community is still primarily focused on the currency's technical, safety and legal issues. Discussion about the economic and financial aspects remains relatively sparse.[23]

Trading stocks is related to human psychology. One tries to predict whether a stock is going up or down in value. One can profit off of this volatility if they have the right entry and exit points. To get the right entry and exist points one needs to have a trading signal. When this signal is triggered a buy/sell order is executed. This had me wondering whether one could use artificial intelligence to gather market sentiment as a means of creating a trading signal. To do this one needs to use natural language processing One tries to predict whether a stock is going up or down in value. A weakness that traditional trading strategies have is that they have no knowledge of current events which can greatly affect the price of a certain stock or asset. Here we will avail of market sentiment to try and better predict the future price of bitcoin. Many researches have been carried out on market sentiment prior to this body of work.

The current literature contains sentiment driven trading using a wide array of different natural language processors, market sentiment heuristics and data set retrieval methods. They analyse factors such as tweet count and tweet volume compared to the price of bitcoin. Once the author has the tweets, we also need some way to gauge market sentiment from them. For this we will avail of some deep learning techniques with natural language processing.

In the book, 2017 IEEE Symposium Series on Computational Intelligence, they were able to predict upcoming crashes and withdraw profit. This was done using sentiment analysis. They were able to outperform the market. [24] In the journal SMU Data Science Review, it was

found that tweet volume rather than tweet sentiment was a key indicator of price prediction. In tandem with google trends search volume, they were able to predict the price of bitcoin.[25] Both pieces of literature above used Twitter as a source of market sentiment. In this research paper we perform comparisons and observe what effects different variables have on the trading signal. Seeing whether tweet volume has any bearing on market sentiment and the price of bitcoin is something we will investigate in this research paper and it is then related back to the trading signal .

In this piece of literature a naive prediction model was used to create a signal. This model yielded an 83% accuracy rate, however, it outlines that further improvements to the lexicon and thus natural language processing would be needed to make more conclusive results.[26] In the journal, SMU Data Science Review , they use google trends search volume in combination with Twitter tweets as a way to collect market sentiment.They utilise a linear model that processes tweets and Google Trends data. As in the rest of the literature, Abraham et al. were able to predict the future price of bitcoin using market sentiment. [25]

In this piece of literature Twitter and StockTwit are availed of as sources of market sentiment. Kendall cross-correlations and non-parametric transfer entropy were used to attain a trading signal. It was concluded that market sentiment and the price of bitcoin are interconnected. [27] Research published in the The Journal of Risk Finance found that there was a correlation between media sentiment and the price of bitcoin. This piece of literature uses psycho-semantic dictionaries on Twitter tweets to generate a trading signal. [28]

Traditional trading bots such as freqtrade use non sentiment driven trading indicators to make trades. [29] One advantage that sentiment driven trading bots have is that they can react to changes in market sentiment due to recent news articles. Countless sentiment driven trading bots are on the market at the moment such as Whitebird which is the successor of Blackbird.[30] Furthermore, there is a sizeable amount of literature available on the topic of sentiment driven trading bots as seen in this section. 3.3.

Identifying trading indicators is no easy feat as one must account for many factors which may be confounding variables.

What makes this research unique is that it focuses on creating a profitable trading signal and analysis what factors trigger said trading indicator. This research will give a detailed trading signal by producing both entry and exit points. An individual using a trading bot to get a sense of the strength of a given indicator. This allows them to gauge the risk they take on. Different factors such as tweet, retweet, sentiment categories and the indicator trigger acceptance criterion were experimented with to see if they can manifest a good trading indicator. A quasi-experimental study was chosen because full control of the independent variable is not achievable and because the author has an exploratory hypothesis. Moreover, sentiment analysis indicators are then compared with more traditional indicators such as bitcoin mining prices and freqtrade by back testing. [29, 31]

3 Methodology

3.1 Introduction to Methodology

Our methodology contains both qualitative and quantitative metrics. To analyse the bitcoin market via sentiment

analysis, one needs to avail of concrete metrics such as the percentage of profit attained over one month using trading strategy X. While aggregating market sentiment at a given point in time, we also need qualitative data. Categorical data is also collected when it comes to deciphering where people gather market sentiment from and what sentiment categories are most influential. The combination of these two methods should yield complementary results.

In order to determine the sentiment of the market with regard to bitcoin we need a platform on which to evaluate sentiment. We will use twitter as they have an open source API.[25] This will make it easier for future work to build on top of this as there is no proprietary software.

We will make use of two twitter bitcoin tweet data-sets. Our First data-set¹ already has market sentiment marked in and will be used as training data. 1 being positive sentiment, 0 being neutral sentiment and -1 being negative sentiment. Our second data-set² is one which contains bitcoin tweets between February and June.

3.2 Quantitative Analysis

In quantitative analysis, I am planning to use a deep learning technique called transformer-based machine learning technique. This will be used to process tweets into market sentiments. It takes the tweets and evaluates whether the signals are positive or negative based on the natural language processing. ROBERTA (BERT) is a language neural network that takes text as input and can process text for a number of tasks, from sentiment analysis to speech prediction to language translation. The Transformer architecture of BERT is bidirectional and can learn context from both directions. Stacking two encoders we get BERT. BERT can be used for sentiment analysis.[32] BERT is pre-trained using Masked Language Modelling and does not require any labelling as in our old design. These “masked tokens” are simply missing words in a sentence, for which BERT fills in the gaps. Simultaneously, next sentence prediction (NSP) is used for BERT to understand the context of a sentence. After this BERT can be fine tuned for a specific task. In this case predicting whether a given tweet has positive, negative or neutral market sentiment.

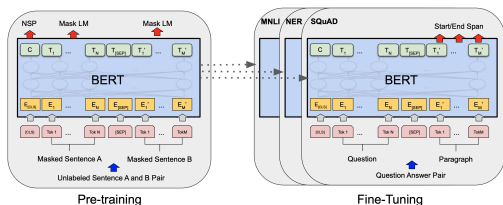


Figure 1: How BERT learns

In figure 1 one can see how BERT makes use of vector spaces to create a minimum spanning tree. The distance of each word from one another and the recovered syntax from transformed BERT vectors give this natural language processor the context of a sentence and allows it to process text to market sentiment effectively. RoBERTa is an

¹<https://data.world/mercal/btc-tweets-sentiment>
²https://www.kaggle.com/kaushikuresh147/bitcoin-tweets?select=Bitcoin_tweets.csv

optimized version of BERT which performs better than BERT.[33] RoBERTa gets rid of next sentence prediction (NSP) in its pre-training phase and has dynamic masking. Data masking is when certain words(tokens) are covered and RoBERTa has to predict what tokens to add in its place. Dynamic masking is when which tokens will be masked is not predetermined.³

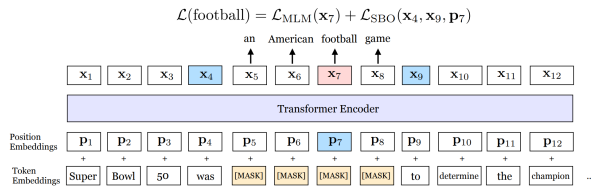


Figure 1: An illustration of SpanBERT. In this example, the span *an American football game* is masked. The span boundary objective then uses the boundary tokens *was* and *to* to predict each token in the masked span.

Figure 2: Masking

We will use RoBERTa for the purposes of tweet analysis because of the advantages in performance it has shown over BERT. We can use RoBERTa via the python package Ernie which avails of Keras and Tensorflow.[34, 35]

3.3 Quantitative Analysis Data Collection

An online survey consisting of 3 multiple choice questions and one ranking question was created⁴. The purpose of this survey was to collect qualitative data, from which some interesting findings may be drawn. The survey was designed to be short and easy for a respondent to fill in. The survey was created on Qualtrics because it offers many analysis tools and has a seamless user interface. One can preview there questions in mobile and tablet form.[36] The survey took place over a 10 day period in which respondents answered questions online. The multiple choice questions gathered data on whether one has traded before and which sources they gather market sentiment from. The ranking based question asked participants to order different types of market sentiment from most to least influential.

3.4 Algorithm

As aforementioned, the project uses market sentiment to deduce how to trade bitcoin. In a reinforcement learning system the agent usually chooses the best policy based on a given reward system, however trading based on sentiment automatically gives out a policy. Consequently, rather than focusing on what behaviour would help maximize profit, we needed to focus on how sentiment trading can be mimicked and what methods would be suitable. Therefore, RoBERTa: a robustly optimized Bidirectional Encoder Representations from Transformers pre-training approach was chosen. RoBERTa is a transformer-based machine learning technique that avails of deep learning. It is the

³<https://blog.inten.to/papers-roberta-a-robustly-optimized-bert-pretraining-approach-7449bc5423e7>
⁴<https://ql.tc/jr3fAx>

natural language processor(NLP) we use to process incoming tweet data and aggregate market sentiment from. The advantage of RoBERTa are mentioned in the literature section 2. Given this market sentiment, trades can then be made when a trading indicator is triggered. We firstly use data-set 1 to train our NLP. We train of 80% of the data and test on 20% of the data. Both the training set and the testing set are drawn from the same population. This means that the model expects the same distribution of features in the test set as in the training set. Thus, the data is first shuffled to avoid bias. Then the data is cleaned. The objective of cleaning is to limit white noise. Characters such as punctuation, special characters, numbers, and terms are removed because they do not affect sentiment and may confuse the classifier.⁵

```
>>> import preprocessor as p
>>> tweet = "@# Buy Bitcoin :) :) )"
>>> print(p.clean(tweet))
Buy Bitcoin
```

We then train the classifier on some NLP. We set epochs to 5, in line with the work of (Hao,Yaru and Dong et al. 2019).[37] In this case it was RoBERTa, had a low loss and an accuracy of just over 90%. We now have a model that can predict the sentiment of bitcoin tweets.

The second data set was sorted by date and every 50th tweet was chosen for processing. This was done because performing text analysis on 500,000 tweets was taking many hours and the machine did not have enough processing power in the given time span. An even distribution of tweets throughout the months was collected. All of which are available on Github, see footnote 5. Tweets from the second data-set were read by a csv reader and a sentiment score was appended to each tweet. This sentiment score was appended by RoBERTa using deep learning and natural language processing as depicted in figures 1 and 2. After this a bitcoin price is appended to each tweet for a given date.[38] We now have a tweet with date, bitcoin price and market sentiment. From here we group tweet data date and get the average market sentiment on a given day in section 4.2.1. The Author also gets the aggregate market sentiment in sections 4.2.2 and 4.2.3.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

Figure 3: Pearson’s correlation coefficient formula

To analyse all the graphs plotted. Pearson’s correlation coefficient is used and it tests if there is a linear relationship between two variables. 3

Once we get some sort of correlation, we can make use of a trading indicator to make a trading strategy. All averages for each strategy were calculated based on the data set given. In real trading, one can take the average of the last 100 days. One such strategy the author devised was coined “The Market Sentiment and Market Count Strategy”. This strategy triggers a buy order when market sentiment is less than the average sentiment per day and when the volume of tweets is greater than average number of tweets per day. It triggers a sell order when market sentiment is greater than

⁵<https://github.com/Singpurwala/SentimentTradingSignals>

average sentiment per day and when the volume of tweets is less than average number of tweets per day.

“The Count Strategy” is a similar strategy. This strategy triggers a buy order when the volume of tweets is greater than the average volume of tweets per day multiplied by 1.5. It triggers a sell order when the volume of tweets is less than half the average number of tweets per day.

A completed trade is one where a buy order is firstly issued and a subsequent sell order is issued. These are the only types of valid trades. In python this is expressed in sudo code as

```
if not alreadyBought and triggeredBuyOrder:
    buyOrderPlaced()
    alreadyBought = True

if alreadyBought and triggeredSellOrder:
    sellOrderPlaced()
    alreadyBought = False
```

Listing 1: How a buy/sell trigger is handled

This trading strategy is then compared to FreqTrade’s various default trading strategies. Freqtrade tests its strategies using backtesting and we do the same on our dataset.[39]

4 Findings & Analysis

4.1 Qualitative Results

Media articles and public opinion on social media can greatly affect the price of a certain asset and market sentiment. This can be seen throughout this literature review in section 2. The results of the ranking question from the survey can be seen in Table 1. The aim was to conduct a survey with over 50 participants and ascertain what platforms are best to gather market sentiment from. 71 out of the 77 respondents completed the survey. Analysis was conducted on these 71 participants.

Rank	Sentiment type			
	Political	Famous People	Mainstream Media	Other
Rank 1	23%	23%	28%	26%
Rank 2	38.5%	32%	22%	7.5%
Rank 3	26%	29%	32%	13%
Rank 4	12.5%	16%	18%	53.5%

Table 1: Ranked Political Sentiment

To find out which sentiment type the respondents thought were most important, this survey made use of positional voting on a ranking question. The question was: Rank these sentiments in order from most to least important. (1) being most and (4) being least. Political sentiment was the most influential type of sentiment, with Famous People’s sentiment and Mainstream Media sentiment closely following, in accordance with the Borda count method. [40]

Following the question, “Where do you get your trading news from?”, it was divulged that 42% of respondents get their trading news from social media, whilst 22% of respondents get their trading news from trading websites. The remaining respondents got their trading news from other

miscellaneous sources. software. Our qualitative results support the hypothesis that market sentiment may correlate with bitcoin price movements because the many of the respondents get there trading news from social media. It is a fair assumption that if they get their trading news from social media that social media has an effect on their trading habits. This serves as a segue in to section 4.2, quantitative results, which provides concrete findings.

4.2 Quantitative Results

4.2.1 Market Sentiment Against Bitcoin Price

To answer sub research question one, it was found that there is a relationship between the price of bitcoin and market sentiment. Surprisingly however, that relationship was not linear. A Pearson's correlation coefficient of 0.06 was present which means that there is no correlation.

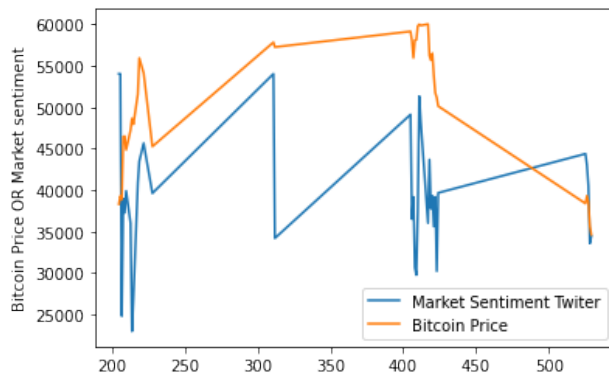


Figure 4: Sentiment Against Bitcoin Price

We can see in figure 4 that market sentiment fluctuates much more readily than the price of bitcoin. There may be some other type of non linear correlation between these two variables. For now though, it can be inferred that market sentiment and bitcoin price are not linear correlated.

4.2.2 Tweet Volume and Market Sentiment Against Bitcoin Price

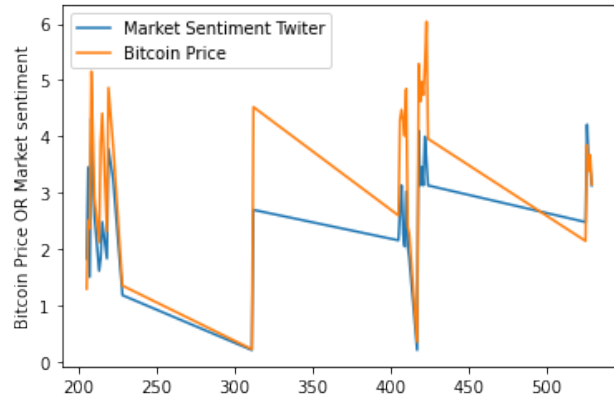


Figure 5: Sentiment & Tweet Volume Against Bitcoin Price

Here we aggregated market sentiment over a given day as in the previous section. What differed was that market sentiment was combined with tweet volume, to compare to bitcoin price movements.

Pearson's correlation coefficient was found to be 0.81 which shows a somewhat positive correlation. Depicted in figure 7

4.2.3 Famous People's Sentiment

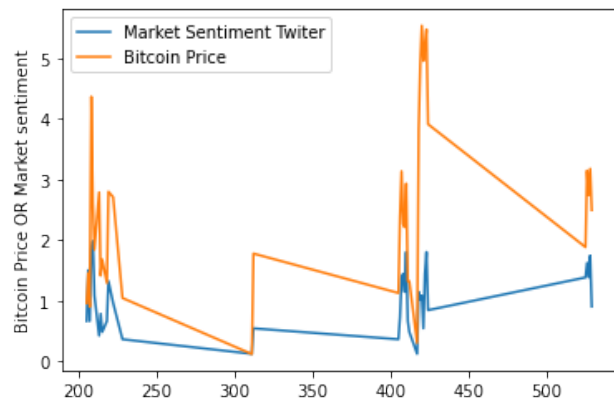


Figure 6: Famous Peoples' Aggregated Market Sentiment

Here only tweets posted by people with over 10,000 followers were considered for analysis. There was some correlation found but it was less than that in section 4.2.2

Pearson's correlation coefficient was found to be 0.54 which shows a somewhat positive correlation. Depicted in figure 8

4.3 Trading Strategies

The trading strategies were both profitable. We created a data-frame object of all buy triggers and all sell triggers. Only triggers in accordance with listing 1 in section 3.4. All figures are available in the appendices.

4.3.1 The Market Sentiment and Market Count Strategy

Firstly, let's look at "The Market Sentiment and Market Count Strategy". The buy triggers are seen in figure 9. The sell triggers are seen in figure 10 in appendix C. In our data set from the start date (10/02/2021) to the end date (25/05/2021), bitcoin's price went from \$44855.62 to \$38378.96. The buy and hold strategy is when one simply buys bitcoin and does not trade further. This would have yielded a cumulative loss of 16.87%. Comparatively, in accordance with our methodology algorithm, we get a cumulative profit of 35%, visible in figure 11. This answers sub research question two as it provides evidence that sentiment based indicators can create profitable trading strategies.

4.3.2 The Market Count Strategy

This strategy simply trades based on tweet volume on a given day. The sell triggers of the market count strategy are seen in figure 13 and the buy triggers are seen in figure 13. We get a cumulative profit of 55.5% as seen in figure 14. Bitcoin had a cumulative loss of 16.87% as discussed in the previous section. The "Market Count Strategy" outperforms buying and holding bitcoin. Further consolidating an answer to the second sub research question.

4.3.3 Sentiment Analysis Compared Against Other Trading Strategies

We have already compared our trading strategies to the strategy of simply buying and holding bitcoin. Now let's compare our strategy to those of a widely used trading bot called FreqTrade.[29] From appendix D, it is seen that Low BB is the best performing inbuilt Freqtrade trading strategy. This was measured via back testing.[39] Low Bollinger Bands (Low BB) had a cumulative profit of 475% in the same time frame as our trading strategy. This outperforms our trading strategies. Our trading strategies were able to outperform many other trading strategies created by Freqtrade such as the Scalp strategy. It is also interesting to note that the Market Count Strategy outperformed The Market Sentiment and Market Count Strategy. This provides further evidence that market sentiment does not correlate with bitcoin price movements and in actuality tweet volume (how much bitcoin is discussed online) correlates with bitcoin price movements.

5 Future Works

As a further continuation of this research, one could gather market sentiment on multiple cryptocurrencies at once and decide which one to trade based on which is predicted to be

most profitable. This can then be converted into a portfolio management tool.

The trading indicator in this research paper can also be branched into different versions running simultaneously. Tweaks can be made with each subsequent alteration. This will refine the trading indicators and trading strategies. The predicted outcome is that this will lead to better trading decisions. The aforementioned trading strategies will be tested using back testing. This testing depicts how a trading strategy would perform in the future under the assumption that past trends continue. Sentiment driven trading indicators can be combined with other trading indicators to find out what combination is most profitable.

6 Conclusion

It was observed that bitcoin market sentiment does not correlate with the price of bitcoin. It was bitcoin market sentiment in combination with tweet volume that correlated with the price of bitcoin. A Pearson's correlation coefficient of 0.8 was found.

Another finding was that a combination of bitcoin tweet sentiment and bitcoin related tweet volume did not produce the most profitable trading strategy. Instead it was bitcoin related tweet volume produced the most profitable trading. Thus, we can deduce that tweet volume correlates to bitcoin price movements better than a combination of tweet volume and tweet market sentiment. We were able to conclude that using a trading strategy based off of tweet volume ("the popularity of the topic of bitcoin on social media") are profitable.

One can conclude that sentimental analysis is never a profitable trading indicator but that trading based on "the popularity of the topic of bitcoin on social media" is a profitable trading indicator as long as there is variance in tweet volume, which is naturally expected to be there and should follow the normal distribution.

References

- [1] L. Kristoufek, "What are the main drivers of the bitcoin price? evidence from wavelet coherence analysis," *PLoS one*, vol. 10, no. 4, p. e0123923, 2015.
- [2] D. F. Gerritsen, E. Bouri, E. Ramezanifar, and D. Roubaud, "The profitability of technical trading rules in the bitcoin market," *Finance Research Letters*, vol. 34, p. 101263, 2020.
- [3] W. K. Härdle, C. R. Harvey, and R. C. Reule, "Understanding cryptocurrencies," 2020.
- [4] L. Yukun and A. Tsyvinski, "Risks and returns of cryptocurrency," in *Meeting Papers*, vol. 160, 2018.
- [5] W.-Y. Nie, H.-P. Cheng, and K.-C. Yen, "Investor sentiment and the cryptocurrency market," *The Empirical Economics Letters*, vol. 19, pp. 1254–1262, 11 2020.
- [6] A. H. Dyhrberg, "Bitcoin, gold and the dollar—a garch volatility analysis," *Finance Research Letters*, vol. 16, pp. 85–92, 2016.
- [7] E. Bouri, A. Jain, P. C. Biswal, and D. Roubaud, "Cointegration and nonlinear causality amongst gold, oil, and the indian stock market: Evidence from implied volatility indices," *Resources Policy*, vol. 52, pp. 201–206, 2017.

- [8] J. Herrera-Joancomartí, “Research and challenges on bitcoin anonymity,” in *Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance*, pp. 3–16, Springer, 2014.
- [9] D. Philippas, H. Rjiba, K. Guesmi, and S. Goutte, “Media attention and bitcoin prices,” *Finance Research Letters*, vol. 30, pp. 37–43, 2019.
- [10] S. Dastgir, E. Demir, G. Downing, G. Gozgor, and C. K. M. Lau, “The causal relationship between bitcoin attention and bitcoin returns: Evidence from the copula-based granger causality test,” *Finance Research Letters*, vol. 28, pp. 160–164, 2019.
- [11] M. Balcilar, E. Bouri, R. Gupta, and D. Roubaud, “Can volume predict bitcoin returns and volatility? a quantiles-based approach,” *Economic Modelling*, vol. 64, pp. 74–81, 2017.
- [12] A. S. Hayes, “Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin,” *Telematics and Informatics*, vol. 34, no. 7, pp. 1308–1321, 2017.
- [13] E. Bouri, C. K. M. Lau, B. Lucey, and D. Roubaud, “Trading volume and the predictability of return and volatility in the cryptocurrency market,” *Finance Research Letters*, vol. 29, pp. 340–346, 2019.
- [14] Q. Ji, E. Bouri, C. K. M. Lau, and D. Roubaud, “Dynamic connectedness and integration in cryptocurrency markets,” *International Review of Financial Analysis*, vol. 63, pp. 257–272, 2019.
- [15] A. F. Aysan, E. Demir, G. Gozgor, and C. K. M. Lau, “Effects of the geopolitical risks on bitcoin returns and volatility,” *Research in International Business and Finance*, vol. 47, pp. 511–518, 2019.
- [16] E. Demir, G. Gozgor, C. K. M. Lau, and S. A. Vigne, “Does economic policy uncertainty predict the bitcoin returns? an empirical investigation,” *Finance Research Letters*, vol. 26, pp. 145–149, 2018.
- [17] O. Alkhalzali, E. Bouri, and D. Roubaud, “The impact of positive and negative macroeconomic news surprises: Gold versus bitcoin,” *Economics Bulletin*, vol. 38, pp. 373–382, 01 2018.
- [18] A. K. Tiwari, R. K. Jana, D. Das, and D. Roubaud, “Informational efficiency of bitcoin—an extension,” *Economics Letters*, vol. 163, pp. 106–109, 2018.
- [19] W. BROCK, J. LAKONISHOK, and B. LeBARON, “Simple technical trading rules and the stochastic properties of stock returns,” *The Journal of Finance*, vol. 47, no. 5, pp. 1731–1764, 1992.
- [20] C. J. Neely, D. E. Rapach, J. Tu, and G. Zhou, “Forecasting the equity risk premium: the role of technical indicators,” *Management science*, vol. 60, no. 7, pp. 1772–1791, 2014.
- [21] J. Bollinger, “Bollinger on bollinger bands, 2001,” *McGraw Hill, New York*, 2001.
- [22] C. Lento, N. Gradojevic, and C. Wright, “Investment information content in bollinger bands?,” *Applied Financial Economics Letters*, vol. 3, no. 4, pp. 263–267, 2007.
- [23] S. Barber, X. Boyen, E. Shi, and E. Uzun, “Bitter to better — how to make bitcoin a better currency,” in *Financial Cryptography and Data Security* (A. D. Keromytis, ed.), (Berlin, Heidelberg), pp. 399–414, Springer Berlin Heidelberg, 2012.
- [24] R. C. Phillips and D. Gorse, “Predicting cryptocurrency price bubbles using social media data and epidemic modelling,” in *2017 IEEE symposium series on computational intelligence (SSCI)*, pp. 1–7, IEEE, 2017.
- [25] J. Abraham, D. Higdon, J. Nelson, and J. Ibarra, “Cryptocurrency price prediction using tweet volumes and sentiment analysis,” *SMU Data Science Review*, vol. 1, no. 3, p. 1, 2018.
- [26] E. Stenqvist and J. Lönnö, “Predicting bitcoin price fluctuation with twitter sentiment analysis,” 2017.
- [27] T. Aste, “Cryptocurrency market structure: connecting emotions and economics,” *Digital Finance*, vol. 1, no. 1, pp. 5–21, 2019.
- [28] V. Karalevicius, N. Degrande, and J. De Weerd, “Using sentiment analysis to predict interday bitcoin price movements,” *The Journal of Risk Finance*, 2018.
- [29] Freqtrade, “Freqtrade trading bot.” <https://www.freqtrade.io/en/stable/>, 2021.
- [30] T. Buchholz, “Whitebird.” <https://github.com/BuchholzTim/Whitebird>, 2021.
- [31] D. Sornette and K. Wu, “Backtesting of trading strategies for bitcoin,” *Eth zurich*, 2019.
- [32] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” *arXiv preprint arXiv:1810.04805*, 2018.
- [33] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, “Roberta: A robustly optimized bert pretraining approach,” *arXiv preprint arXiv:1907.11692*, 2019.
- [34] R. M. Castaño, “Ernie natural language processing.” Ernie, 2020.
- [35] R. M. Castaño, “Ernie natural language processing.” Ernie Secondary Git Project, 2020.
- [36] Qualtrics, “Qualtrics survey platform.” <https://www.qualtrics.com/uk/>, 2020.
- [37] Y. Hao, L. Dong, F. Wei, and K. Xu, “Visualizing and understanding the effectiveness of bert,” *arXiv preprint arXiv:1908.05620*, 2019.
- [38] CoinDesk, “Coindesk api to get bitcoin price.” <https://www.coindesk.com/coindesk-api>, 2012.
- [39] Freqtrade, “Freqtrade backtesting.” <https://www.freqtrade.io/en/2021.1/backtesting/>, 2021.
- [40] P. C. Fishburn and W. V. Gehrlein, “Borda’s rule, positional voting, and condorcet’s simple majority principle,” *Public Choice*, vol. 28, no. 1, pp. 79–88, 1976.

Appendices

A Market Sentiment and Tweet Volume Against Bitcoin Price



Figure 7: Market Sentiment and Tweet Volume VS Bitcoin Price

B Famous Peoples' Sentiment



Figure 8: Market Sentiment and Tweet Volume Vs Bitcoin Price

C Trading Strategies

C.1 The Market Sentiment and Market Count Strategy

sentiment	btcPrice	Count	SSMA_50
23	38378.9600	49	16.10
25	39269.3417	37	22.80
18	44855.6167	41	0.00
22	55923.1300	50	15.68
24	58087.1800	39	21.38
19	58119.5900	38	21.58
30	59791.8450	49	26.24

Figure 9: Potential Buys - “The Market Sentiment and Market Count Strategy”

sentiment	btcPrice	Count	SSMA_50
23	38378.9600	49	16.10
25	39269.3417	37	22.80
18	44855.6167	41	0.00
22	55923.1300	50	15.68
24	58087.1800	39	21.38
19	58119.5900	38	21.58
30	59791.8450	49	26.24

Figure 10: Potential Sells - “The Market Sentiment and Market Count Strategy”

```

Trading Strategy -- avgCount
buying $34613.67 worth of bitcoin
selling $38378.96 worth of bitcoin
buying $47211.6683 worth of bitcoin
selling $55923.1 worth of bitcoin
Overall Profit/Lose: 12476.721700000002
This is a 0.3496942930479437% cumulative profit/Lose

```

Figure 11: Market Sentiment and Market Count Strategy - Cumulative Profit

C.2 The Market Count Strategy

```

Trading Strategy -- avgCount
buying $35678.94 worth of bitcoin
selling $38306.2467 worth of bitcoin
buying $39299.3583 worth of bitcoin
selling $45256.4133 worth of bitcoin
buying $46436.09 worth of bitcoin
selling $51573.4067 worth of bitcoin
buying $51722.1333 worth of bitcoin
selling $57827.6133 worth of bitcoin
Overall Profit/Lose: 19827.1584
This is a 0.555710410679241% cumulative profit/Lose

```

Figure 14: Market Count Strategy - Cumulative Profit

sentiment	btcPrice	Count	SSMA_50
29	35678.9400	89	26.18
27	39299.3583	80	28.34
31	46436.0900	94	35.42
14	50123.7367	78	16.90
30	51161.1467	107	28.10
24	51722.1333	101	23.84
9	53812.0750	92	11.70
17	55668.2283	88	17.18
18	56499.2767	98	17.60

Figure 12: Potential Buys - “The Count Strategy”

sentiment	btcPrice	Count	SSMA_50
11	38306.2467	25	15.18
11	38862.3500	23	11.52
6	45256.4133	23	6.52
11	51573.4067	25	11.22
2	57827.6133	2	0.66
6	59140.6850	19	5.82
8	59856.7133	22	8.00
11	59990.8733	22	9.54
2	60023.5133	5	1.52

Figure 13: Potential Sells - “The Count Strategy”

D Freqtrade Trading Strategies

STRATEGY SUMMARY									
Strategy	Buys	Avg Profit %	Cum Profit %	Tot Profit BTC	Tot Profit %	Avg Duration	Wins	Draws	Losses
ADXmomentum	6178	-0.24	-1455.58	-0.26226642	-0.01	2:18:00	2777	0	3401
ASDTSRockwellTrading	124164	-0.23	-28974.09	-5.22055322	-0.29	0:43:00	26752	3	97409
AverageStrategy	103487	-0.26	-27086.47	-4.88044133	-0.27	1:26:00	17867	0	85620
AwesomeOsc	30006	-0.25	-7601.89	-1.36970941	-0.08	4:53:00	6313	0	23693
BbandRSI	12002	-0.04	-471.44	-0.08494441	-0.00	13:05:00	7136	0	4866
BinRV27	7815	0.01	74.11	0.01335333	0.00	4:35:00	4528	0	3287
BinRV45	1630	0.15	240.45	0.04332467	0.00	2:45:00	1351	0	279
CCIStrategy	1489	-0.02	-22.85	-0.00411762	-0.00	21:39:00	432	0	1057
ClucMay72018	2456	0.00	9.07	0.00163370	0.00	0:47:00	1760	0	696
CMCWinner	9655	-0.13	-1230.25	-0.22166741	-0.01	3:02:00	4205	4484	966
Cof1BitStrategy	21763	-0.14	-3099.60	-0.55848589	-0.03	0:30:00	8592	2	13169
CombineBinAndCluc	3082	0.15	472.89	0.08520351	0.00	1:13:00	1937	0	1145
DoesNothingStrategy	0	0.00	0.00	0.00000000	0.00	0:00:00	0	0	0
EMASkipPump	72009	-0.15	-10777.53	-1.94189538	-0.11	2:12:00	35969	5	36035
Freqtrade_backtest_validation_freqtrade1	81929	-0.25	-20138.99	-3.62864510	-0.20	1:50:00	18819	1	63109
InformativeSample	0	0.00	0.00	0.00000000	0.00	0:00:00	0	0	0
Low_BB	649	0.73	475.11	0.08560479	0.00	1 day, 14:05:00	70	0	579
MACDStrategy_crossed	3709	-0.32	-1190.00	-0.21441342	-0.01	1 day, 20:34:00	3272	0	437
MACDStrategy	16234	-0.17	-2817.29	-0.50761926	-0.03	11:24:00	11002	0	5232
MultiRSI	22028	-0.15	-3408.56	-0.61415518	-0.03	1:53:00	12071	0	9957
ReinforcedAverageStrategy	44239	-0.25	-11160.25	-2.01085445	-0.11	2 days, 1:31:00	7972	0	36267
ReinforcedQuickie	15064	-0.15	-2240.26	-0.40365019	-0.02	6:50:00	10099	0	4965
ReinforcedSmoothScalp	1895	-0.06	-105.98	-0.01909506	-0.00	2:21:00	1063	0	832
SampleStrategy	7657	-0.29	-2232.10	-0.40217967	-0.02	1 day, 0:40:00	6624	0	1033
Scalp	21962	-0.14	-2998.24	-0.54022260	-0.03	0:29:00	8953	2	13007
Simple	15499	-0.33	-5112.42	-0.92115609	-0.05	18:33:00	12286	0	3213
SmoothOperator	9494	-0.16	-1477.11	-0.26614670	-0.01	11:15:00	5431	1	4062
SmoothScalp	4551	-0.03	-125.88	-0.02268172	-0.00	2:37:00	2789	0	1762
Strategy001	9152	-0.39	-3527.98	-0.63567124	-0.04	1 day, 0:16:00	7682	0	1470
Strategy002	971	-0.10	-100.18	-0.01805061	-0.00	4:39:00	584	0	387
Strategy003	2595	-0.08	-217.06	-0.03910967	-0.00	4:49:00	1564	1	1030
Strategy004	3975	-0.02	-74.14	-0.01335863	-0.00	3:51:00	2407	0	1568
Strategy005	21098	-0.08	-1653.04	-0.29784564	-0.02	1:32:00	11192	0	9906
TDSquentialStrategy	20760	-0.19	-3977.26	-0.71662289	-0.04	6:25:00	11061	1	9698
TechnicalExampleStrategy	38422	-0.28	-10875.54	-1.95955428	-0.11	7:10:00	30455	0	7967
AdxSmas	44442	-0.23	-10269.59	-1.85037534	-0.10	2:56:00	14392	1	30049

Figure 15: Trading Strategy Back-testing Results