Using sentiment analysis in cryptocurrency price comparison

Mattia Carminati

Supervisor: Carlo Mamo

June - 2023

A dissertation submitted to the Institute of Information and Communication Technology in partial fulfillment of the requirements for the degree of BSc (Hons) in Software Development

Authorship Statement

This dissertation is based on the results of research carried out by myself, is my own composition, and has not been previously presented for any other certified or uncertified qualification.

The research was carried out under the supervision of Mr. Carlo Mamo.

01/06/2023

Date

Signature

Copyright Statement

In submitting this dissertation to the MCAST Institute of Information and Communication Technology I understand that I am giving permission for it to be made available for use in accordance with the regulations of MCAST and the Library and Learning Resource Centre. I accept that my dissertation may be made publicly available at MCAST's discretion.

01/06/2023

Date

Signature

Acknowledgements

I would like to thank Mr. Carlo Mamo who supported me during my dissertation work as my tutor.

Abstract

The area of research that will be investigated is related to the world of cryptocurrencies. Cryptocurrencies, such as Bitcoin and Ethereum, are not just stores of values but can also be used as a payment method. These virtual coins gained a lot of popularity in 2020, when they got the attention of people, newspapers, television, social media, and so on, thanks also to the pandemic, during which prices increased significantly. The aims of this research are to show if Bitcoin can influence the price of other cryptocurrencies and if positive tweets and announcements of people can influence the price of Bitcoin and other cryptocurrencies. Apart from Bitcoin (BTC), the other cryptocurrencies analyzed were Ethereum (ETH), Bitcoin Cash (BCH), Dogecoin (DOGE), Litecoin (LTC), Decentraland (MANA), and Solana (SOL). All the collected data were then stored in a NO-SQL database called Firestore, which was particularly suitable for these data since it was more flexible and perfect for unstructured data. Cryptocurrency prices were then exported into a .csv file and data were represented with line graphs, one for each cryptocurrency analyzed. After observing the graph of Bitcoin, it was possible to notice four periods of time where its price either increased or decreased drastically. During such periods, price variations of cryptocurrencies were calculated, and the percentages of Bitcoin were then compared with other cryptocurrencies. The results obtained have shown that when Bitcoin experienced a significant price change, even the other cryptocurrencies followed the same direction at the same time, even though with higher or lower percentages, which depended on the actual price of the cryptocurrency being analyzed. Since all prices were increasing or decreasing at the same time, it was not possible to confirm that Bitcoin had an influence on the price of other cryptocurrencies. To answer the second research question and determine if positive tweets and announcements of people can influence the prices of cryptocurrencies, primary research made use of the Twitter API and several Python libraries to fetch tweets from the web and collect the information required for this purpose. For each day between 1st January and 30 April 2023, and for each cryptocurrency analyzed, Tweepy library retrieved up to

50000 tweets per day containing their names. Tweets were then cleaned from special characters, emojis and hashtags, and were classified as positive, neutral or negative thanks to the TextBlob Python library. This library supports complex analysis and operations on textual data, and classified sentiments by its semantic orientation and intensity of each word in a tweet. TextBlob contained a pre-defined dictionary classifying positive and negative words, and after assigning a score to all the words, it calculated the final sentiment for each tweet. After the sentiment classification, the number of positive tweets was stored in a separate collection on Firestore, and data were extracted to .csv files. Afterward, positive tweets for each cryptocurrency were represented with line graphs including even the price of each cryptocurrency in a secondary axis. The results obtained have shown that for some cryptocurrencies, positive tweets might have influenced their prices, even though there are other elements not mentioned in this research that can influence prices more than positive tweets.

Contents

List of Figures

List of Tables

1 Introduction

When people wanted to try investing in cryptocurrencies, they just did a bit of research on the web to understand if it was a good investment or not, and most of the time, they only bought them because their prices were rising, and consequentially, they thought they would have kept going to that direction. Although sometimes this could be true, this strategy was extremely risky because, if after buying a cryptocurrency the market went down, people would have got a loss. One of the best ways to choose when to invest is to know what people are currently thinking about cryptocurrencies, and if people have a positive attitude that will make them buy or a negative which will make them sell their assets. For this purpose, the libraries used in this research will retrieve announcements and tweets of people who keep talking and investing in cryptocurrencies and will show the sentiment levels. By reading this research, people will get a better idea of the current market sentiment and will plan better their investments in cryptocurrencies.

1.1 Research Questions

To determine if major cryptocurrencies had an influence on others and if sentiment analysis could influence the trend of cryptocurrencies, a set of research questions was formulated to be answered. Table 1 shows the research questions and the motivations for their constitution.

Research question	Motivation
RQ1: Is Bitcoin influencing the prices of other cryptocurrencies?	The purpose of this question is to establish if there is a price correlation between Bitcoin and major cryptocurrencies.
RQ2: Do tweets and announcements of people influence the trend of cryptocurrencies?	The purpose of this question is to establish if positive tweets can influence the prices of cryptocurrencies.

Table 1: Research questions

1.2 Research Objective and Scope

The objective of this research is to determine if the price of Bitcoin can influence the prices of other cryptocurrencies and if positive tweets and announcements of people can influence the price of cryptocurrencies. To investigate and reach these objectives, Binance and Twitter APIs were used to retrieve both prices and positive tweets. The choice of using these APIs was done to retrieve real-time and accurate data, that are fundamental for comparing the price of cryptocurrencies and sentiment analysis. The gap found in the market was the lack of research that made use of sentiment analysis to compare cryptocurrency prices.

1.3 Cryptocurrencies

1.3.1 Bitcoin

Launched in 2009, Bitcoin is the world's first cryptocurrency built on distributed ledger (blockchain) with a proof-of-work mechanism that is not backed by any country's central bank or government. Bitcoin was created by Nakamoto [1], a pseudonym representing an individual or group of individuals, who published the whitepaper on October 31st, 2008. Bitcoin gained popularity because of its several features, that made it an asset different from the others, such as stocks and bonds. As mentioned by Singhal et al [2], since Bitcoin is a currency that is available only virtually, its transactions run on the blockchain, a distributed database where records are maintained across several computers that are linked in a peer-to-peer network. As described by Andolfatto [3], while a normal database usually structures its data into several tables, a blockchain structures data into blocks that are strung together. When implemented in a decentralized nature, this data structure inherently makes an irreversible timeline of data. Once a block is filled, it is set in stone and becomes a part of this timeline. All the blocks of the chain are given an exact timestamp when they are added to the chain. The fact that it works on the internet makes Bitcoin faster than banks or other types of transactions. In fact, sending money in the form of BTC from one side to another in the world takes a few minutes with low commissions, while if that money is sent through bank transactions, it can take a week or more, and higher commissions can be applied according to the banks that are involved in the transaction. All the transactions that run on the blockchain are public, meaning that everyone can see the date of the operation, the amount that was sent in Bitcoin and equivalent fiat currency, and the wallet addresses of the sender and receiver, but it is not possible to see the name and surname of both sender and receiver. While this is done to ensure total privacy of people, some have criticized this aspect, stating that it makes it very hard to identify scammers, when transactions are not genuine and are done with the purpose to steal or hide money. In fact, Bitcoin is totally decentralized, which means that it cannot be controlled or censored by any government or bank entity, allowing individuals to store capital with self-sovereignty. Another important feature of Bitcoin is its supply. While traditional currencies can be generated indefinitely by printing banknotes and minting new coins at any time, Bitcoin has a cap of 21 million, after that it is impossible to generate new Bitcoins. As described by Semret [4], since Bitcoin has the proof-to-work consensus, new Bitcoins are generated through a decentralized and competitive process called "mining", that uses computing power to process transactions. The computing power is provided by computers equipped with several types of hardware, such as ASICS, graphics cards and processors. When someone needs to send Bitcoins to another person, all the computers of all the miners connected to the internet start to solve an extremely complex computational math problem. The first computer to find the solution to that problem receives a small reward in Bitcoin, and then the process starts again. The reward that miners receive is an incentive that motivates them to monitor and legitimize Bitcoin transactions and ensure their validity. Because of its market cap, Bitcoin is being considered as the digital equivalent of gold, that is also limited to the quantities found in nature and cannot be created artificially.

1.3.2 Ethereum

As described by Singhal et al [5], Ethereum is a software platform that runs on its blockchain, and its cryptocurrency called ether is the second largest cryptocurrency by volume. Contrary to Bitcoin, the founder of Ethereum is public and called Vitalik Buterin, a programmer who invented this crypto in 2015. Ethereum is different from Bitcoin because it allows to do much more than simple payments on the blockchain. As mentioned by Buterin [6], Ethereum allows software developers to create new applications, that can go from lending apps to payment platforms and will make interactions by the use of smart contracts. A smart contract is the main feature of Ethereum, defined as a program that runs autonomously on the blockchain and performs several functions that are generically done by third parties. Smart contracts are self-executing pieces of code, written in Vyper and Solidity and tested using Ethereum Remix IDE, that will run once the conditions defined by the involved parties are met. Vyper is an experimental contract-based language inspired by Python, while Solidity is a high-level object-oriented language influenced not only by Python but also JavaScript and C++, designed to integrate with the Ethereum virtual machine (EVM). The EVM helps the Ethereum blockchain to maintain its state by defining the rules for computing new valid states from block to block. To represent it in a mathematical function, Y can be defined as a state transaction function. The function can be written as Y(S, T) = S' where:

- S is the old valid state.
- T is the set of new valid transitions to be added to the next block.
- S' is the new valid state.

This function takes the old valid state, and a set of new valid transactions to produce a new valid state as an output.

Although Ethereum currently uses the proof-of-work consensus like Bitcoin, its founder has shown his intention to change it to the proof-of-stake, a type of consensus where users can validate transactions according to the number of coins that they stake. In this way, users who stake more coins will have higher chances to be chosen to validate transactions on the network and get a reward. The advantages of using this type of consensus include energy efficiency because it does not need computer power to solve the algorithms, faster and inexpensive transaction processing because the consensus is established before blocks are constructed, improving its scalability.

1.3.3 Decentraland

As described by Ordano et al [7], Decentraland is a decentralized virtual reality platform running on the Ethereum blockchain, where it is possible to buy virtual lands on that to create, experience and monetize content and applications. Lands are non-fungible digital assets that can be bought by using Decentraland's cryptocurrency token called MANA. This token is the in-game currency that lets users purchase or trade LAND, and buy or sell goods and services on Decentraland's marketplace. Purchasable virtual assets include estates, wearables, avatars and unique names. MANA tokens use the proof-of-work consensus mechanism like Ethereum, which means that they can be mined using ASIC or graphics cards. Even though MANA uses this type of consensus, its decision-making tool called DAO uses an off-chain voting system, allowing MANA token holders to participate in the governance and vote to make changes to the smart contracts on which the platform is built. To start playing in Decentraland, users must create an avatar, that can be customized with different body types, accessories, shoes etc. In this world, users can use their avatars to do the same things that they do in real life, but in a virtual environment. For example, avatars can interact with each other, go to buy clothes in virtual shopping malls, attend events, enjoy parties and run businesses.

1.4 Crypto Exchange

By definition, a crypto exchange is a platform where it is possible to trade several cryptocurrencies. Therefore, this exchange allows people to buy crypto or digital currencies using standard currencies such as euros, United States dollars or pounds, or using other cryptocurrencies by simply doing a conversion. For each crypto available in a chosen exchange, there is a wallet associated with it, that works in the same way as our physical wallets, meaning that it is used to deposit cryptocurrencies that have been bought and withdraw them when it is time to sell them. According to Cong et al [8], at the same time crypto exchanges compete between them in many ways in the market, with the goal to attract utility tokens that have the potential to produce high trading volume, and by launching new tokens. According to Burilov [9], some of the existing exchanges such as Binance, Crypto.com, and KuCoin have already launched their own tokens, with the aim to incentive and fuel decentralization of token listing.

2 Literature Review

As mentioned by Othman et al [10], the world of FIAT currencies was old and not secure enough for some countries with bad economies, where inflation had reached negative records. Since 2009, cryptocurrencies have been emerging and changing the whole market, because they introduced new aspects such as the absence of government regulations, the almost total anonymity of users, and their security while managing payments. Bitcoin and Ethereum, apart from being the most capitalized cryptocurrencies in the world, are also the most used blockchains on which transactions of other cryptocurrencies also take place. This fact is due to several reasons. As explained by Nakamoto [1], Bitcoin has the most secure network because it is based on a proof-of-work mechanism, ensured by a volume of over 18 million miners that has an effect of a high level of decentralization of the network, ensuring flawless security of the Bitcoin payments system. As explained by Singhal et al [2], Bitcoin is also considered a reserve currency for cryptocurrencies, similar to the US dollar in the global stock markets. Most expert crypto traders never compare the prices of altcoins to fiat currencies and do the comparison with Bitcoin, as it is the market pair that matters the most to them. Regarding Ethereum, the dependency that other cryptocurrencies have on it is due to its protocol. As mentioned by Buterin [6], Ethereum enables the deployment of smart contracts and decentralized applications, giving it potential applications which are wide-ranging and powered by its native token. Users who want to use the Ethereum platform to build smart contracts must buy ethers, since those are the tokens that are used to make interactions and without them, transactions cannot be processed. During the last decade, the use of Bitcoin and Ethereum has increased a lot and governments have started to study and set up regulations regarding the exchange between cryptocurrencies and fiat currencies. These new systems have been adopted by individuals and businesses wishing to transact quickly and efficiently over the Internet without the need to supply Credit Cards or Banking information. But this is not the only reason why people and businesses got interested in cryptocurrencies. In fact, cryptocurrencies are extremely volatile, meaning that their values can change significantly and rapidly, at any hour of the day and, most of the time, without a valid reason. According to Kozak and Gajdek [11], many people started to invest in them, with the hope to invest a certain amount today and to get a huge profit by the end of the month, or even earlier. Obviously, prices did not increase every time, people had to find the right moment to invest their money, and to help them in this task, various trading bots have been created. As described by Andersson et al [12], trading bots are automated trading systems that conduct trades and execute transactions on behalf of human investors. Using these bots brings many advantages to people. For instance, when there are price variations, bots react quicker than investors, and act according to them, also because many investors don't have enough time to dedicate to always get the best trade. Most crypto trading bots offer various services, such as data analysis, risk prediction, and buying / selling crypto assets. Data analysis means that bots fetch from a variety of web sources, scan the market data found on them, interpret them, and decide whether to buy or sell. Also, many bots give the possibility to customize the types of data to provide refined results. Risk prediction is a must-have feature for a trading bot since it leverages market data to estimate the potential risk of assets. This information helps the bot decide how much to trade or invest.

2.1 Social Media Sentiment Analysis

As mentioned by Rahul et al [13], sentiment analysis can be defined as a Natural Language Processing which aims to find the emotions of public opinion by analyzing text generated by users. This analysis has been widely used on social media platforms, where people give their opinions and reviews on relevant topics. The data collected by social media, in the form of words contained in a message can be used by applications and bots to take intelligent decisions. The term "sentiment" in the finance context is a view or an opinion expressed about the condition of a market. Market sentiment is the collective attitude of traders and investors toward a financial or market asset, including cryptocurrencies. It conveys the crowd psychology of those involved in the trading and development of the cryptocurrency as reflected through social and trading metrics. As mentioned by Waltz et al [14] Even though it does have the power to influence market cycles, favorable market sentiment does not always lead to positive market conditions. In fact, strong positive sentiment may come before a market correction or even a bear market. An example of this situation can be represented by the research conducted by Tandon et al [15], which shows the case of the well-known meme crypto called Dogecoin, which took the attention of millions of investors, thanks to the announcements and continuous tweets posted by Elon Musk. A lot of its demand in its bull run came from social media hype, which led to positive market sentiment, and convinced traders and investors to buy it, without considering the project's white paper, but only because of the positive market sentiment. Market sentiment has been recently adopted by a huge number of applications, which transform human-generated information such as tweets, announcements, and news, into a sentiment signal, either positive or negative. That is why sentiment analysis is an essential part of trading strategies, which can help to investigate whether the hype is justified or simply a result of herd mentality. Common tools used for such analysis are sentiment indicators, which give an idea of whether a certain market asset is bullish or bearish and represent these feelings either with graphs or scale. The most known indicator of crypto market sentiment is the Bitcoin Crypto Fear and Greed Index, which shows Bitcoin market fear or greed on a scale of 0 to 100 by analyzing different information sources: volatility, social media, dominance, trends, and market volume.

2.2 Social Media and Cryptocurrency Pricing

When social media gained popularity, businesses started to invest on them, by advertising and moving their activities online. For example, newspapers stopped printing physical papers and published everything online, to make their articles accessible to a wider range of people. Nowadays, social media have a big influence on people, since most of them are connected and influenced by them. Even celebrities started to share their lives and make business by advertising products or services. In the same way, the world of cryptocurrencies has been impacted by social media and Twitter in particular, because of the advertisements and celebrities who started to talk about their investments, successes, and losses. Twitter is a social networking site and online news where several types of entities such as celebrities, advertisers and political figures share messages called tweets. This platform can count on around 300 million monthly active users. The research made by Tandon et al [15] has demonstrated that the popularity of Elon Musk, due to the rise of Tesla, has influenced the price of Dogecoin in a positive way. Elon Musk is the founder of many, important companies, two of which are Tesla and SpaceX. Starting in 2020, his tweets have made a huge impact on his companies by increasing the value of their assets, and on both share and stock market. In the study, liked tweets and price/volume data were collected about both Dogecoin and Bitcoin cryptocurrencies. After analyzing Elon Musk's tweets based on likes, tweet counts, retweeted tweets, and most used words, and collecting historical stock data of Bitcoin, it has been possible to represent by graph the price of Bitcoin and volume evolution over the years. The use of keywords such as Tesla, Bitcoin, cryptocurrency, Mars, starship, and the launch was accounted for segregating the tweets based on cryptocurrency. As a model, the study used the ARIMA model which predicts future values of a given cryptocurrency based on past behavior. The final result has shown that the perspective of the common man can be influenced by Twitter and social media in general, and get motivated to invest more money without paying attention to the volatility of the market. As a consequence, people can easily lose all their life savings and much more. The study of Rothman and Yakar [16] has used empirical analysis to analyze and evaluate tweets. The research used the Twitter API to collect millions of tweets related to several cryptocurrencies, containing the hashtag #Bitcoin, between November 2017 and August 2018. The social networks analyzed were Twitter, Telegram, and Reddit. This time frame includes the period when the price of Bitcoin hit 20000$ as well as the period of its drop. The empirical analysis has shown that the effect of the social media volume for the above-mentioned social media is positive, while only on Telegram and Reddit the social media volume have a significant effect on the price. Then, the study has shown that except for Bitcoin and ZenCash, all cryptocurrencies are correlated in price, which means that investing in them gives broad exposure to the crypto world. Activity on social media has a true relationship with cryptocurrency fluctuations in terms of volume/price. Another study made by Naeem et al [17] proposed a sentiment analysis combined with machine learning methods. Datasets, consisting in exchange rates and tweets were retrieved from Forex and Twitter respectively. The study has shown how the method to forecast the USD/PKR exchange rate was made up and its results could be used by the Pakistani business community to invest in both local market and forecast exchange rates for the future.

2.3 Sentiment Analysis

As discussed by Susrama et al [18], the sentiment is defined by its semantic orientation and intensity of the words present in a sentence or a message. Sentiment classification requires a pre-defined dictionary with positive and negative words already classified. To return the subjectivity and polarity found in a given sentence, **TextBlob** makes use of semantic labels to help with fine-grained analysis, which includes emojis, exclamation marks etc.

As discussed by Wang et al [19], in sentiment analysis subjectivity refers to the degree to which a person is involved in an object. This is measured by the personal connection and individual experiences which the person has with the object, which can differ from someone else's point of view. Subjectivity quantifies the amount of personal opinion and factual information contained in the text. The higher subjectivity means that the text contains personal opinions rather than factual information. Its value is of type float and it falls within the range 0.0 to 1.0, where 0.0 is very objective and 1.0 is very subjective.

On the other side, polarity refers to the strength of an opinion, which can be either positive or negative. When something has a strong emotion such as love, trust, admiration or positive feeling associated with it, this will have a specific orientation toward all the other aspects of that object's existence. The strength of both positive and negative polarities varies according to the situation, but they can still be considered strongly positive or negative. The important thing to consider here is the feeling of people toward something else through non-verbal communication. As for the subjectivity, the type of value for polarity is a float falling within the range -1.0 to 1.0, where 0 indicates neutral, 1.0 represents a very positive sentiment and -1 represents a very negative sentiment.

A good sentiment analysis tool must be able to distinguish between subjectivity and polarity to correctly analyze the opinions of people. These opinions can have a high degree of subjectivity if they are expressed as personal experiences, while low degrees might indicate someone else's viewpoint on something. In addition, sentiments can have several levels of polarity throughout several ways of communication, which can be tweets, emails, chats etc.

Moreover, people can express themselves differently when writing or speaking about specific things that might affect the outputs of the sentiment analysis significantly.

A lot can be said about how both polarity and subjectivity affect sentiment analyses, especially because they are closely related. While polarity refers to the overall positivity or negativity of a statement, subjectivity focuses on how much someone values an object and expresses his/her opinion about it. Sentiment analysis tools must be able to distinguish between both types in order to analyze users' feelings correctly.

3 Research Methodology

3.1 Search Strategy

The search strategy is essential to allow relevant studies to be included in search results to help researchers get as many as possible. The search strategy used was composed of the following elements: search terms and data sources. The type of approach used was quantitative because the research relied on numerical data, identified as the number of positive tweets, and prices of each crypto being analyzed over the period of time between January and April 2023. Secondary research was conducted by assessing appropriate academic journals, articles, and websites. This literature review was carried out both to determine if Bitcoin can influence the trend of other cryptocurrencies and also to determine if social media can influence the price of cryptocurrencies.

3.1.1 Search Terms

Search terms are used to match paper titles, keywords, and abstracts in electronic data sources during automatic searches. The following strategies were used to form the most relevant search terms for automatic search:

- Derive key terms from research questions and study topics;
- Identify synonyms, plurals, and related terms;
- Use the logical operator “OR” to incorporate synonyms;
- Use the logical operator “AND” to concatenate the parameters;
- Check terms in article titles, abstracts, and keywords.

The resulting search terms are composed of the following synonyms and terms: ALL("Sentiment analysis") AND ("Blockchain" OR "Cryptocurrency") AND (LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018)) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "cp")) AND (LIMIT-TO (SUBJAREA, "ECON") OR LIMIT-TO (SUBJAREA, "COMP") OR LIMIT-TO (SUBJAREA, "MATH") OR

LIMIT-TO (SUBJAREA, "BUSI") OR LIMIT-TO (SUBJAREA, "ENGI")) AND (LIMIT-TO (LANGUAGE, "English")) The additional terms “exchange” and “market” were also added since the results of the inclusion of these terms are also of interest to the research.

3.1.2 Data Sources

The electronic databases selected for the research are presented in Table 2 and were selected considering their popularity, ease of access, and the possibility of retrieving the full text of the articles. A small number of resources were selected from other electronic databases and institutional websites of foreign universities, since those resources were not included in the list of databases provided below, but were still considered helpful for this research. English was selected as the language considered to be the standard of most international journals, and the period of time was set between 2018 and 2022, in order to get the most updated articles.

Number of articles	Electronic Database	Search terms are matched with	Web address
200	Elsevier Scopus	Paper title, keywords, abstract	http://www.scopus.com
150	Research Gate	Paper title, keywords, abstract	http://www.researchgate.net
70	Google Scholar	Paper title, keywords, abstract	http://www.scholar.google.com

Table 2: Electronic databases

3.2 Inclusion and exclusion criteria

In order to enable only studies that met the keywords, inclusion and exclusion criteria were adopted. In particular, results of papers including the keyword "sentiment analysis" but not discussing cryptocurrencies were excluded, since the main area of this research is related to cryptocurrencies. Articles were selected firstly by the keywords, and those matching were then selected according to their abstracts.

Inclusion criteria
Articles related to social media and their influence on cryptocurrencies
Articles related to cryptocurrencies and using sentiment analysis to find correlations
Exclusion criteria
Non-scientific publications
Publications that do not meet the inclusion criteria

Table 3: Inclusion and exclusion criteria

3.3 Data Collection

3.3.1 Twitter Data Collection

The list of libraries used in this research includes **pandas**, **NumPy**, **Tweepy**, and **matplotlib**. The first one is a Python library started by Wes McKinney in 2008 and is used today for data analysis. It acts as a wrapper of **NumPy** and **matplotlib**, allowing to access many of their methods with less code. For example, the command **.plot()** combines several **matplotlib** methods into a single method, allowing to plot a chart in a few lines. **NumPy** is a library used for working with arrays and it is being used mainly to perform mathematical operations on them, by using its mathematical, algebraic, and transformation functions. As discussed by Govindasamy [20], the main advantage of using **NumPy** is its speed which provides an array object that is up to 50x faster than traditional Python lists. It manages to be faster than lists because its arrays are stored at one continuous place in memory unlike lists, so functions can manipulate them very efficiently. **Matplotlib** is a cross-platform library written by John D.Hunter in 2003 which is used for data visualization and graphical plotting. It offers a hierarchy of objects abstracting various elements of a plot. The particular feature of this library is the **pyplot** state machine that enables users to write concise procedural code. **Tweepy** is an open-source Python library that provides access to the Twitter API. Any request sent to this API requires OAuth for authenticating and implies the user to apply for a Twitter developer account, by following the official procedure available on the official website. To start using the API, 4 credentials have to be generated. These credentials include the Consumer key, Consumer secret, Access token, and Access secret. After generating them, these had to be saved in the local machine, inside the project folder.

Moreover, **Tweepy** provided an API class to access the Twitter RESTful API methods made for tweets, users, trends, and likes. Tweets are the elements that have been analyzed in this research because they can reveal the sentiment of people toward cryptocurrencies. These were messages containing up to 280 characters which included emojis, hashtags, and special characters. For this type of research, those elements were not needed and they were not reliable in terms of sentiment, meaning that a tweet could not be marked as positive by relying on them. The sentiment analysis process started by defining the search term, which was made up of the hashtag of the crypto being investigated followed by the filter. The variable named search_result was used to search the tweets by specifying also the language used by each tweet, the date from when tweets should be collected, and the tweet_mode, which in this case was "extended".

```
search_term = '#' + self.cryptoName + ' -filter:retweets'
search_result = tweepy.Cursor(api.search_tweets, q=search_term,
                    lang='en', tweet_mode= 'extended').items(self.
                    numberOfTweets)
```

Figure 1: Tweets research algorithm

The tweets retrieved by the application were then written in a Pandas DataFrame, a two-dimensional size-mutable tabular data structure with rows and columns. To check that the research has been performed successfully, the first 5 tweets were displayed.

The retrieved tweets were then cleaned from special characters because they were not reliable in terms of sentiment analysis. To do this, the function

Tweets
0 this is a modern and high potential project, t...
1 Jailbot found #bitcoin in a User vault at this...
2 @EverythingAjay Biggest risk for government is...
3 $APE #ETH #Bitcoin RIGHT NOW FULL SEND...
4 @EthaxCrypto #DeFi #ethax #ethaxcrypto #ethaxt...

Table 4: Uncleaned tweets retrieved by the sentiment class

called cleanTwt was written. This function took 1 parameter which represented the tweets and returned the same parameter after the cleaning process was completed. Tweets were cleaned from words containing hashtags, emojis, and hyperlinks, and the capital letters were written in small.

The results were stored in Table 5 showing 2 columns, one for the Tweets and the other for the Cleaned tweets.

Tweets	Cleaned_Tweets
0 this is a modern and high potential project, t...	this is a modern and high potential project, t...
1 Jailbot found #bitcoin in a User vault at this...	Jailbot found bitcoin in a User vault at this ...
2 @EverythingAjay Biggest risk for government is...	@EverythingAjay Biggest risk for government is...
3 $APE #ETH #Bitcoin RIGHT NOW FULL SEND...	$APE Bitcoin APE RIGHT NOW FULL SENDOOO OOO A...
4 @EthaxCrypto #DeFi #ethax #ethaxcrypto #ethaxt...	@EthaxCrypto cryptocurrency BitcoinTh...

Table 5: Tweets and cleaned tweets

3.3.2 Crypto Data Collection

Table 6 shows the cryptocurrencies that were taken into consideration, sorted by market cap.

Cryptocurrency	Symbol
Bitcoin	BTC
Ethereum	ETH
Solana	SOL
Dogecoin	DOGE
Litecoin	LTC
Bitcoin cash	BCH
Decentraland	MANA

Table 6: Cryptocurrencies

The choice of these 7 cryptocurrencies was determined by some inclusion criteria. From the cryptocurrencies obtained through the API, only those with a price greater or equal to 0.01$ were selected, meaning that gambling cryptocurrencies such as Shiba Inu, which had more than 5 decimals, were not included in the list. The last inclusion criteria, applied for the benefit of the sentiment analysis, was the full name of the crypto, which had to be not ambiguous, meaning that it didn't have to be related to any other topic outside of cryptocurrencies. With this criteria cryptocurrencies such as **Maker (MKR)**, **Internet Computer (ICP)**, **The Graph (GRT)**, and **Waves (WAVES)** were excluded. Data related to cryptocurrencies were collected by the use of Binance API. As described by Changpeng [21], Binance is a well-known exchange founded in 2017, which offers crypto-to-crypto trading in more than 500 virtual tokens and cryptocurrencies such as Bitcoin, Ether, Dogecoin, and its signature token Binance Coin. This exchange is particularly known for its low transaction fees and high liquidity. In fact, Binance offers several services around trading, fundraising, listing or cryptocurrency withdrawals. Its staking platform called Binance Earn allows people to stake and earn interests, depending on the coin and tenure, by depositing stablecoins with the exchange.

3.3.3 Firebase

Firebase is a platform developed by Google that is used to create apps for mobile and web devices that stores data on the cloud. All the data collected from

Twitter and the APIs mentioned previously were stored in a NO-SQL database called Firestore Database. This database was more flexible and dynamic to save data, and makes use of collections and documents instead of tables as shown in Figure 2. The main advantages of Firestore Database over the others include ACID transactions that are consistent, isolated, and durable across hundreds of documents and collections, the fact that there is no need to set up an intermediary server to manage access to data, and its horizontal scaling which is perfectly suitable to visualize prices and sentiment results in a specific date.

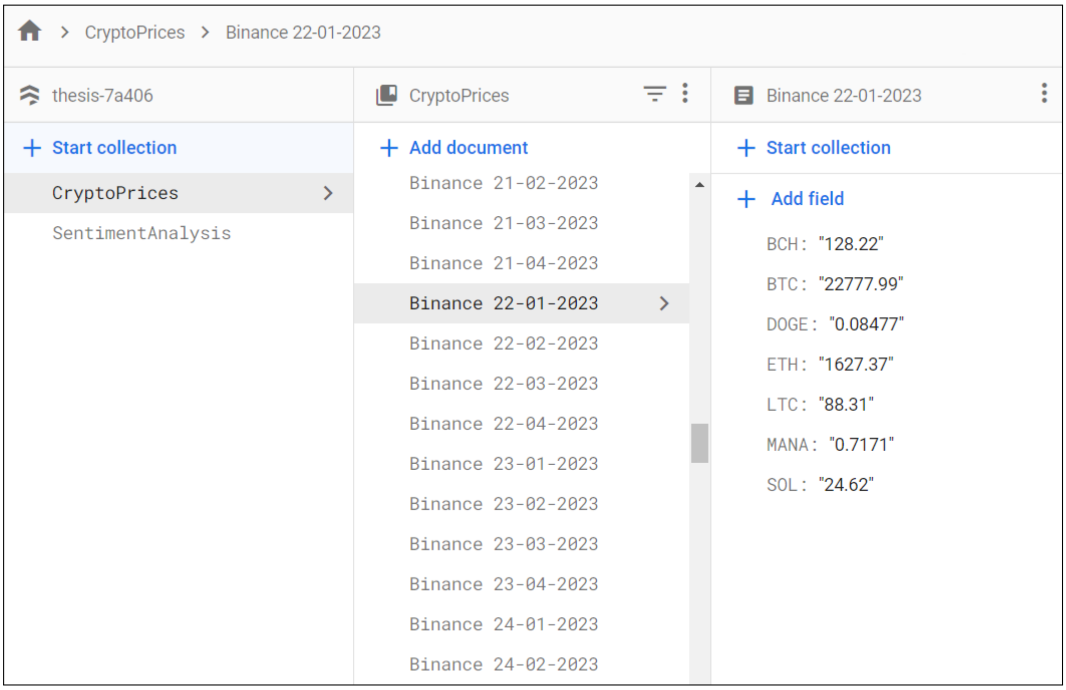


Figure 2: Firestore Database

4 Analysis of Results and Discussion

The period of time between January and April 2023 was chosen to determine if Bitcoin can influence the price of other cryptocurrencies, and also to analyze the market sentiment of cryptocurrencies by making use of sentiment analysis. The choice of a four-month period was made to increase the accuracy of the sentiment retrieved from tweets. For each cryptocurrency taken into consideration in this research, the code made use of Tweepy and TextBlob libraries to retrieve tweets, filter, and classify them as positive, neutral or negative. Since TextBlob had a dictionary where all the words were already classified as positive, neutral or negative, the process was relatively fast and easy to implement. Once the classification process was completed, the prices of cryptocurrencies and the number of positive tweets were inserted in the Firestore database and ready to be compared.

4.1 Crypto Comparison

This research was conducted with the aim to solve issues related to cryptocurrencies. The first aim was to determine if Bitcoin can have an influence on the price of other cryptocurrencies, which have a lower market cap. It was important to find this correlation because many people who want to start to buy crypto often do not know all the mechanisms behind them, ignoring those in which they do not have any interest and focusing exclusively on opinions from the web, newspapers, and/or social media posts. To determine the correlation between Bitcoin and other cryptocurrencies, the first step was to retrieve the price data from the API of Binance Exchange. As shown in Figure 3, Binance API returned a JSON string containing the pairs made up of crypto/USDT and their prices.

```
[
 {"symbol":"BTCUSDT","price":"21376.35000000"},
 {"symbol":"ETHUSDT","price":"1456.04000000"},
 {"symbol":"SOLUSDT","price":"21.73000000"},
 {"symbol":"DOGEUSDT","price":"0.08176000"},
 {"symbol":"LTCUSDT","price":"64.00000000"},
 {"symbol":"BCHUSDT","price":"182.80000000"},
 {"symbol":"MANAUSDT","price":"1.00080000"},
]
```


Figure 3: Cryptocurrency prices retrieved with the use of Binance API

The cryptocurrencies that were taken into consideration were Bitcoin (BTC), Ethereum (ETH), Bitcoin Cash (BCH), Dogecoin (DOGE), Litecoin (LTC), Decentraland (MANA) and Solana (SOL). Other cryptocurrencies retrieved from the API which had a market capitalization greater than the above-mentioned cryptocurrencies were discarded because of the few tweets retrieved during the period, which could not provide good accuracy when compared to their relative graph. All the prices retrieved from the API were then stored in Firebase in a collection called CryptoPrices, which contained all prices of each cryptocurrency for each day from January to April 2023. Afterward, all data were exported and merged into a single .xlsx file and ready to be compared. Bitcoin was the first cryptocurrency to be represented in a line graph, which included only 1 y-axis. Afterward, all the remaining cryptocurrencies were also represented in line graphs, but with 2 y-axis. The primary axis shown the price of Bitcoin and the secondary axis the price of the cryptocurrency being analyzed.

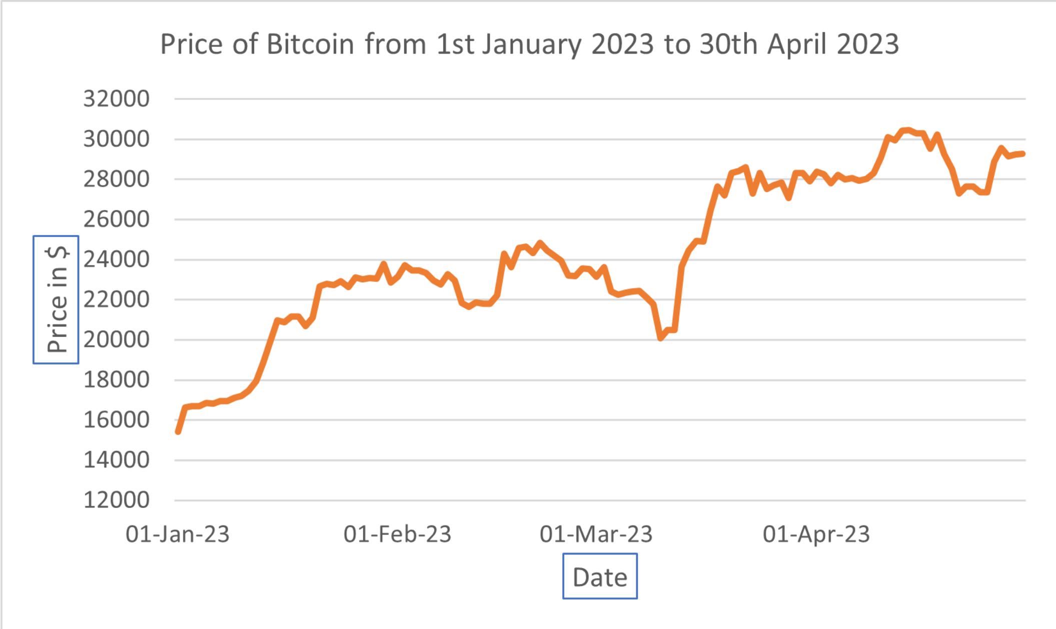
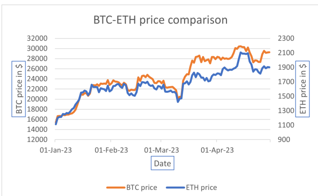
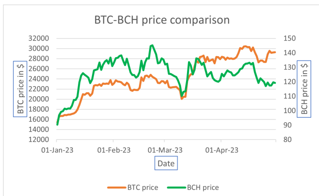


Figure 4: Bitcoin price

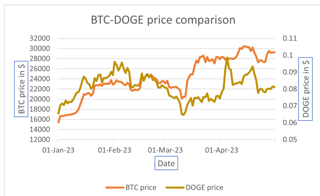
While comparing the graphs in Figure 5, it has been possible to notice some interesting aspects. Firstly, there was a rise made by the price of Bitcoin from $17000 to $23000 in January that affected even the other cryptocurrencies, making them increase as well and keeping the trend until the end of the month. Next, all prices were constant with highs and lows until the 10 of March, when there was a significant drop of Bitcoin, which eventually made another rise on the 21st of March and kept constant until the 12 of April when another rise occurred.



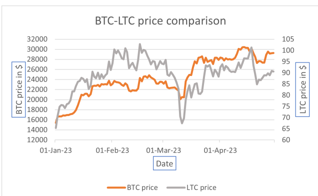
(a) Ethereum price



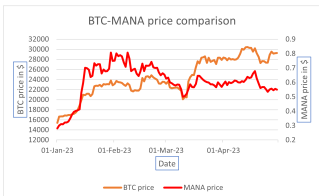
(b) Bitcoin Cash price



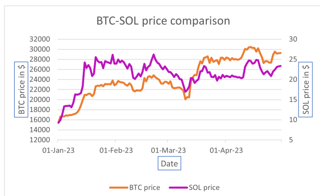
(c) Dogecoin price



(d) Litecoin price



(e) Decentraland price



(f) Solana price

Figure 5: Prices of other cryptocurrencies

When the graph of Bitcoin was compared with the others, the lines representing the price were following the same direction of the price of Bitcoin.

Crypto	01/01 - 22/01	22/02 - 10/03	10/03 - 22/03	18/04 - 21/04
BTC	+48%	-12%	+43%	-10%
ETH	+46%	-13%	+27%	-12%
BCH	+42%	-14%	+21%	-10%
DOGE	+30%	-23%	+16%	-17%
LTC	+35%	-20%	+16%	-16%
MANA	+156%	-30%	+22%	-17%
SOL	+165%	-31%	+28%	-14%

Table 7: Cryptocurrency price variations

To understand better the price movements, Table 7 shows the percentages of increase and decrease of prices. While looking at the period from 1st to 22 January and from 10 to 22 March, it was possible to notice that all other cryptocurrencies experienced an increase such as Bitcoin, with percentages that reached +165% for Solana in the month of January. On the other hand, between the periods 22 February - 10 March and 18 April - 21 April, Bitcoin price decreased and this also happened to the other cryptocurrencies. However, since the price of Bitcoin and other cryptocurrencies follow the same trend almost at the same time, it was not possible to establish that the price of Bitcoin had an influence on the prices of other cryptocurrencies. To confirm this, trading volumes of all cryptocurrencies being analyzed were retrieved from Binance.com. By looking at the trading volumes shown at Figure 6, it was possible to notice that there were no correlations between Bitcoin and the other cryptocurrencies, meaning that the increase/decrease of trading volumes were all occurring in different periods that were not correlated to each other.

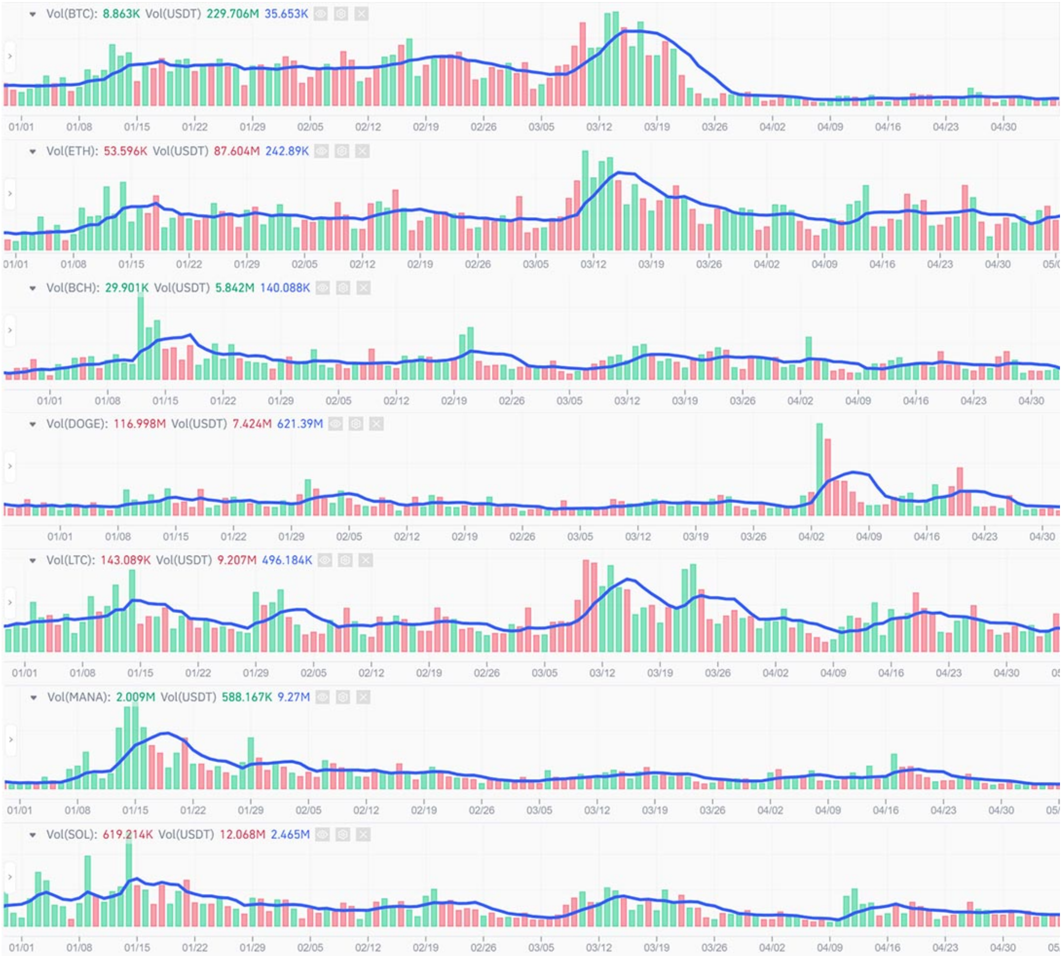


Figure 6: Cryptocurrency trading volumes

4.2 Twitter Sentiment Classification

The second aim of this research was to determine if positive tweets and announcements of people can influence the price of cryptocurrencies. The number of tweets that were collected by the Twitter API has been set to max 50000 per day for each crypto, which corresponds to the limit for the current type of API being used. The Python library called **TextBlob** was used to conduct the analysis. **TextBlob** is a library supporting complex analysis and operations on textual data, and makes use of **Natural Language ToolKit (NLTK)** to achieve its tasks. **NLTK** is a library that gives easy access to a lot of lexical resources and allows users to work with classification, categorization and many other tasks. To test the accuracy of TextBlob, a number of tweets retrieved with the Twitter API were trained with a classifier based on the Naive Bayes algorithm, as implemented in NLTK. Then, the classifier was passed into the constructor of TextBlob and its classify() method was called. The accuracy() method was used to calculate the accuracy of TextBlob for each tweet that was analyzed. The accuracy value could vary from 0 to 1, where 0 = not accurate and 1 = very accurate.

Tweet	Sentiment	Accuracy
Bitcoin price keeps increasing	Positive	1.0
Bitcoin is a scam	Negative	1.0
There are more holders of Bitcoin	Positive	1.0
It is a good moment to invest in BTC	Positive	1.0
More crypto whales have sold BTC assets since China has banned cryptocurrencies	Negative	0.83
Bitcoin has been adopted by more countries, even if some people are still afraid of Bitcoin.	Positive	0.86

Table 8: TextBlob accuracy test

After all the collected tweets were cleaned from the special characters, two new functions called **getSubjectivity** and **getPolarity** were used to classify the tweets and elaborate the sentiment analysis. This analysis can be defined as the process of analyzing and classifying data based on the need of the research, which will help to understand the emotions felt by the general public regarding a given topic. Figure 7 represented the Bitcoin scatter plot which used dots to represent the values given to each of the tweets retrieved previously. Polarity was set to x-axis and Subjectivity on y-axis.

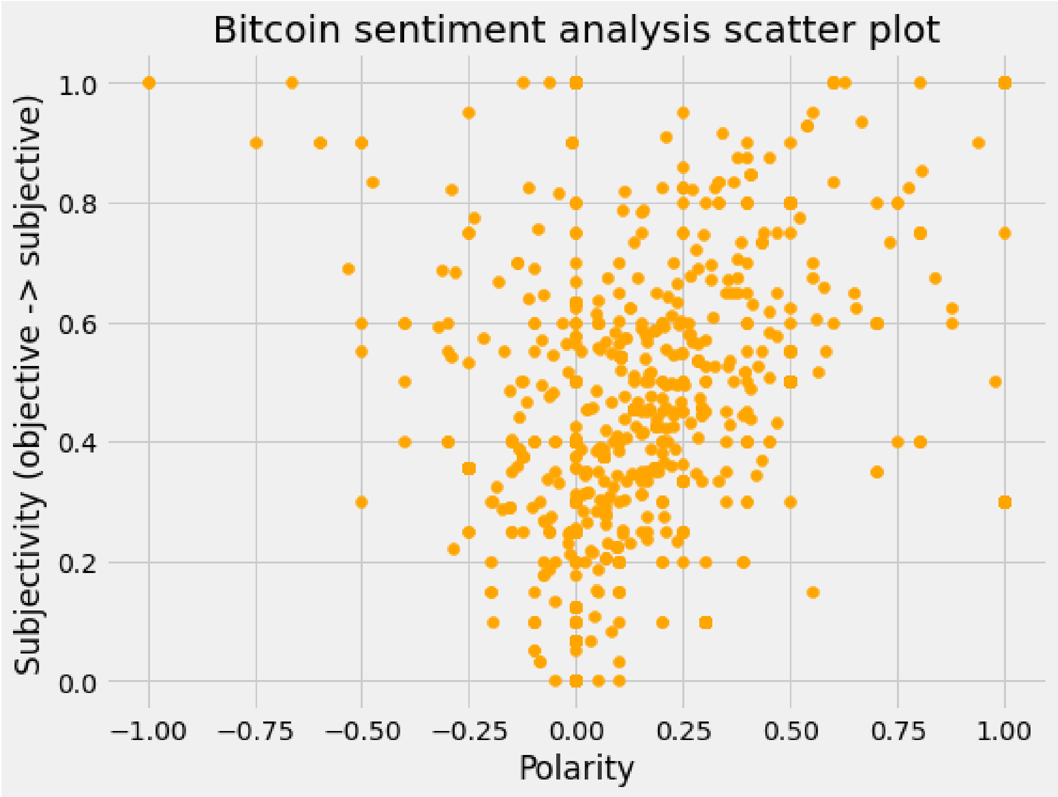


Figure 7: Bitcoin scatter plot

Once subjectivity and polarity were defined, the last function named **getSentiment** was used to calculate the number of positive tweets. After all tweets were displayed on the scatter plot graph, they were saved on Firebase like the crypto prices, but on a different collection and the positive ones were used to display the relative graphs. The graphs were structured to display the date on the x-axis, price of cryptocurrency on the primary y-axis and sentiment analysis on the secondary y-axis, so that it would have been easier to see and compare price vs sentiment analysis. While comparing the price of Bitcoin and its sentiment analysis at Figure 8, it has been possible to notice that the sentiment followed the price increases of January and March, while for the rest of the period had a movement quite similar to the one of the price. The drop of positive tweets that occurred in April was due to the low amount of positive tweets collected since the Twitter API did not retrieve many tweets on that particular day.

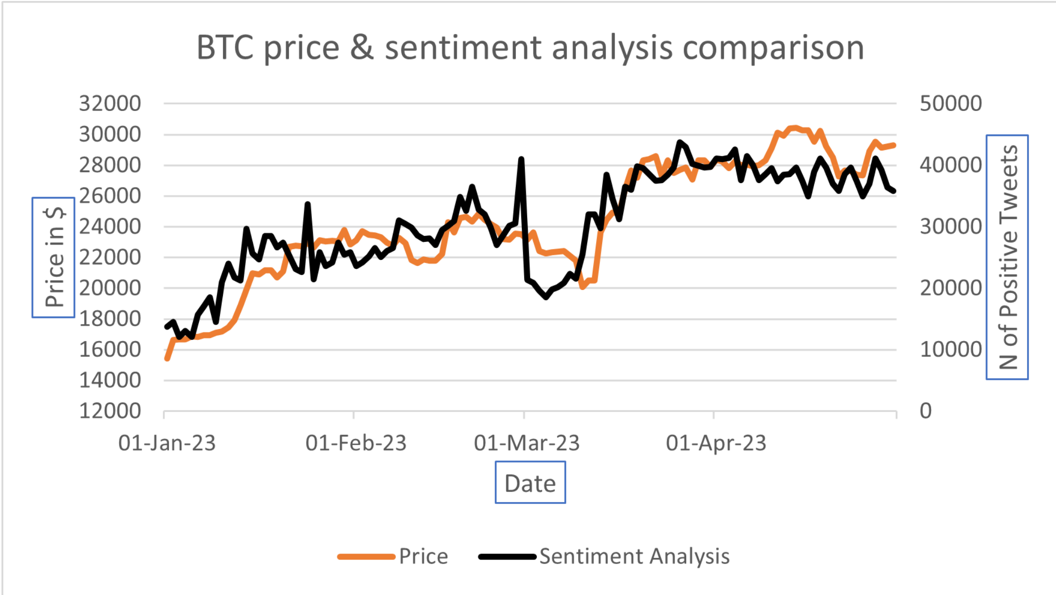
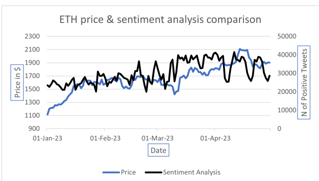
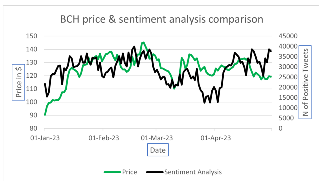


Figure 8: Bitcoin sentiment analysis

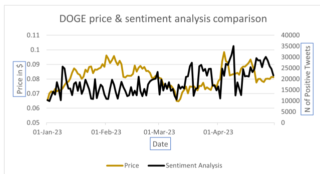
After comparing the other cryptocurrencies, the following results were obtained: the graphs shown at Figure 9 related to Bitcoin, Ethereum, Bitcoin cash, Litecoin and Solana have shown in many points that both price and positive sentiment went either up or down within a few days, while the other two cryptocurrencies Dogecoin and Decentraland have shown lines that were not going to the same direction and therefore, the correlation for these two cryptocurrencies was weaker.



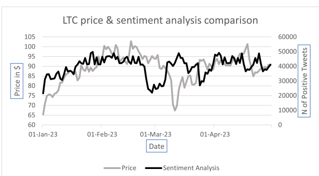
(a) Ethereum sentiment analysis



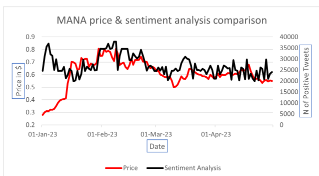
(b) Bitcoin Cash sentiment analysis



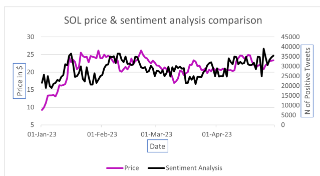
(c) Dogecoin sentiment analysis



(d) Litecoin sentiment analysis



(e) Decentraland sentiment analysis



(f) Solana sentiment analysis

Figure 9: Sentiment analysis of other cryptocurrencies

While comparing Table 9 and Table 10, positive tweets related to the period 1st - 22 January increased and the rise affected also the prices for all of them. During the 22nd February and 10th March, the number of positive tweets decreased and apart from Ethereum, even the prices of cryptocurrencies decreased. After the 10th March until the 22nd March positive tweets still increased and the prices of almost all of them followed the trend. During the last period, positive tweets increased for only 3 of them, but all the prices decreased.

Crypto	01/01 - 22/01	22/02 - 10/03	10/03 - 22/03	18/04 - 21/04
BTC	+48%	-12%	+43%	-10%
ETH	+46%	-13%	+27%	-12%
BCH	+42%	-14%	+21%	-10%
DOGE	+30%	-23%	+16%	-17%
LTC	+35%	-20%	+16%	-16%
MANA	+156%	-30%	+22%	-17%
SOL	+165%	-31%	+28%	-14%

Table 9: Cryptocurrency price variations

Crypto	01/01 - 22/01	22/02 - 10/03	10/03 - 22/03	18/04 - 21/04
BTC	+68%	-20%	+47%	-13%
ETH	+9%	+9%	+55%	-5%
BCH	+48%	-47%	+44%	+10%
DOGE	+68%	-15%	+143%	+16%
LTC	+93%	-2%	+2%	+8%
MANA	+20%	-22%	+1%	-5%
SOL	+9%	-5%	-10%	-7%

Table 10: Positive tweets variations

Hence, it could be that for some cryptocurrencies their prices were affected by positive tweets, but more research and techniques have to be used to determine it. Also, the thesis data is restricted as the posts and comments that were acquired from Twitter were from the general audience, and from the general review and consideration of individuals. Moreover, not all the users of social media are experts in cryptocurrencies. For instance, if Elon Musk mentions Bitcoin or other cryptocurrencies, his influence is stronger when compared to some other influential personalities. On the contrary, the effect of normal person perception and review has less effect on the cryptocurrency prices.

4.3 Discussion

This thesis was directed toward the investigation of the correlation between cryptocurrencies and sentiment analysis. The first research objective was to determine if Bitcoin can influence the price of the other cryptocurrencies. The data that were collected from the Binance exchange through the use of its API and their relative comparisons were represented with Figures 4 and 5. These graphs have shown that the price of Bitcoin did not influence the price of other cryptocurrencies. Crypto prices are typically measured in Bitcoin and since many of them cannot be purchased directly using traditional currencies, Bitcoin becomes their favorite substitute. If people who hold cryptocurrencies want to sell all of them in one go, it is most likely that they would sell their assets for Bitcoin at first, and then convert Bitcoin with the desired fiat currency. Despite this, if Bitcoin does experience a bull run on its own, it is not necessarily true that even the other cryptocurrencies will behave in the same way, and it can also be the exact opposite. The second research objective was to identify if tweets can influence the trend of cryptocurrencies. The results have shown that for some cryptocurrencies such as Bitcoin, Ethereum, and Bitcoin Cash, their prices might have been affected by positive tweets. However, the results that were achieved in this research and the used techniques were not enough to prove that there was always a correlation between the price of cryptocurrencies and tweets/announcements of people on Twitter. Furthermore, the work revealed that the volatility in both crypto prices and sentiment graphs is extremely high that restricted the study from acquiring the appropriate evaluation of the data correlation. The sentiment analysis has been able to collect and analyze up to 50000 tweets per day and classify them by the words that were composing them. However, the tweets collected were acquired from the general audience, meaning that some of them might have been written by people who were not experts in cryptocurrencies.

5 Conclusions and Recommendations

Conclusively, this thesis has directed toward the acquisition of the aim which is "Using sentiment analysis in cryptocurrency price comparison". The key purpose of this thesis was that there were some gaps left by prior studies for analyzing and comparing Bitcoin to other cryptocurrencies and determining if positive tweets can influence the price of cryptocurrencies. Past studies have utilized different techniques and models to analyze cryptocurrency prices and sentiments. Tandon et al [15] made use of the ARIMA model that predicts future values of a given cryptocurrency based on past behavior. This model has shown how the tweets of Elon Musk have significantly contributed to the rise of Dogecoin price. Although this research made use of the ARIMA model to predict the future value of cryptocurrencies, only Dogecoin was analyzed. Then, the study of Rothman and Yakar [16] made use of empirical analysis to evaluate the sentiment analysis of Bitcoin by analyzing tweets through the Twitter API. This research analyzed tweets from Twitter, Telegram and Reddit, and has shown that only on the last two the social media volume had a significant effect on the price of Bitcoin. Even though this research retrieved tweets from 3 different social media, it analyzed only one cryptocurrency. Another study made by Naeem et al [17] has proposed a research that combined sentiment analysis and machine learning methods. This research used datasets consisting of exchange rates and tweets retrieved from the Forex market and Twitter respectively, but instead of cryptocurrencies, the study has shown how the method to forecast the USD/PKR exchange rate was made up. Despite the above-mentioned ones, this research was conducted with the aim to determine if Bitcoin price can influence the prices of other cryptocurrencies and if positive tweets and announcements of people can influence the price of cryptocurrencies. To reach the first aim, this research started by retrieving the prices of Bitcoin and other cryptocurrencies through the use of the Binance API between the months of January and April 2023. Apart from Bitcoin (BTC), other cryptocurrencies that were analyzed were Ethereum (ETH), Bitcoin Cash (BCH), Dogecoin (DOGE), Litecoin (LTC), Decentraland (MANA), and Solana (SOL). Cryptocurrency prices were stored in a NO-SQL database called Firestore and afterward exported in an

Excel file ready to be compared. By looking at the graph of Bitcoin, it was possible to notice 4 periods of time in each month, when the price increased or decreased drastically, and during such periods price variations were calculated. The results obtained have shown that during the months of January and March, all cryptocurrency prices increased almost at the same time, while during the months of February their prices decreased. Therefore, it was not possible to confirm that the price of Bitcoin had an influence on the prices of other cryptocurrencies. Instead, it is possible that all of them have been influenced by external factors and / or events. To reach the second aim, this research made use of the Twitter API to fetch up to 50000 tweets per day containing all the cryptocurrency names mentioned previously. All the tweets collected were then cleaned from symbols, hashtags and emojis, and through the TextBlob python library, each word contained in the tweets was classified as positive, neutral or negative, and after assigning a score to each word, it calculated the final sentiment of each tweet. The research made use of TextBlob because it was built on the shoulders of NLTK and Pattern, making it simple and providing an intuitive interface to NLTK and because it provided functions such as language detection and translation, that were particularly helpful for all tweets written in foreign languages. Tests were carried out to check the accuracy of TextBlob and the results obtained have shown a high accuracy included between 0.83 and 1 where 0 = not accurate and 1 = very accurate. For each cryptocurrency, the number of positive tweets was exported, then represented in a graph and compared with the price of the respective cryptocurrency. By calculating the positive tweets variations in the same periods of time that were used to calculate price variations, it was possible to notice that for some cryptocurrencies, positive tweets can have an influence on their prices, but they are not certainly the only factors to do so. Except for Dogecoin, all other cryptocurrencies have shown that there might be a correlation between their prices and sentiment analysis. The reason of why Dogecoin had a weaker correlation might be because it is considered more as a meme coin instead of a serious asset like Bitcoin and Ethereum and therefore, its sentiment is much more speculative than the other cryptocurrencies being analyzed in the research. As a recommendation, this research can be improved in some aspects. For each cryptocurrency, price comparison can be evaluated more by extending the period of research up to a year or more. On the other hand, sentiment analysis can be improved by upgrading the API plan to a higher level. Having an API that retrieves more tweets would improve the results, making them more accurate. Then, more social media can be included, and their posts can be analyzed with different machine learning techniques that are more suitable to them. Also, since Twitter data is restricted to the posts and comments that were written by the general audience, implementing more filters that can fetch tweets posted only by celebrities and cryptocurrency investors would make the Twitter data more accurate when comparing them to the price of cryptocurrencies.

6 References

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Cryptography Mailing list at https://metzdowd.com*, 03 2009.

[2] B. Singhal, G. Dhameja, and P. Panda, *How Bitcoin Works*, pp. 149–217. 07 2018.

[3] D. Andolfatto, "Blockchain: What it is, what it does, and why you probably don't need one," *Review*, vol. 100, pp. 87–95, 01 2018.

[4] N. Semret, "Dynamics of bitcoin mining," 01 2022.

[5] B. Singhal, G. Dhameja, and P. Panda, *How Ethereum Works*, pp. 219–266. 07 2018.

[6] V. Buterin, "Ethereum: A next-generation smart contract and decentralized application platform.," 2014.

[7] E. Ordano, A. Meilich, Y. Jardi, and M. Araoz, "Decentraland, a blockchain-based virtual world,"

[8] L. Cong, X. Li, K. Tang, and Y. Yang, "Crypto wash trading," 08 2021.

[9] V. Burilov, "Utility token offerings and crypto exchange listings: how regulation can help?," 11 2018.

[10] A. Othman, N. Kawsar, A. Hasan, and F. Mahadi, "Determining the appropriate investment strategy and identify the leading monetary system before and during the covid-19 pandemic crisis: A case study of crypto-currency, gold standard, and fiat money," *Journal of Information Technology Management*, vol. 13, 03 2021.

[11] S. Kozak and S. Gajdek, "Risk of investment in cryptocurrencies," *Economic and Regional Studies / Studia Ekonomiczne i Regionalne*, vol. 14, pp. 294–304, 09 2021.

[12] S. Andersson, J. Häggström, U. Icimpaye, and H. Lindstedt, "Crypto proxima: An analysis of autonomous bitcoin trading," 2021.

[13] M. Rahul, R. R R, and S. Shine, "Social media sentiment analysis for malayalam," *International Journal of Computer Sciences and Engineering*, vol. 06, pp. 48–53, 07 2018.

[14] M. Waltz, A. Singh, and O. Okhrin, "Vulnerability-covar: investigating the crypto-market," *Quantitative Finance*, pp. 1–15, 05 2022.

[15] C. Tandon, S. Revankar, H. Palivela, and S. S. Parihar, "How can we predict the impact of the social media messages on the value of cryptocurrency? insights from big data analytics," *International Journal of Information Management Data Insights*, vol. 1, no. 2, p. 100035, 2021.

[16] T. Rothman and C. Yakar, "Empirical analysis owards the effect of social media on cryptocurrency price and volume," *European Scientific Journal ESJ*, vol. 15, 11 2019.

[17] S. Naeem, W. Mashwani, A. Ali, M. I. Uddin, M. Mahmoud, F. Jamal, and C. Chesneau, "Machine learning-based usd/pkr exchange rate forecasting using sentiment analysis of twitter data," *Computers, Materials and Continua*, vol. 67, pp. 3451–3461, 03 2021.

[18] G. Susrama, N. Mandenni, M. Fachrurrozi, S. Ilham Pradika, K. Manab, and N. Sasmita, "Twitter sentiment analysis as an evaluation and service base on python textblob," *IOP Conference Series: Materials Science and Engineering*, vol. 1125, p. 012034, 05 2021.

[19] W. Wang, B. Li, D. Feng, A. Zhang, and S. Wan, "The ol-dawe model: Tweet polarity sentiment analysis with data augmentation," *IEEE Access*, vol. PP, pp. 1–1, 02 2020.

[20] S. Govindasamy, "Twitter sentimental analysis," vol. 7, pp. 343–346, 11 2018.

[21] Z. Changpeng, "Binance exchange.," 2019.