INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Influencers, are they responsible for Bitcoin's volatility? Transfer Entropy and Granger causality in prol of an answer

Jana Luisa Teixeira Lage

Master’s in Data Science

Supervisor:
Doctor Diana Elisabeta Aldea Mendes, Associate Professor
Iscte – Instituto Universitário de Lisboa

Co-Supervisor:
Doctor Fernando Manuel Marques Batista, Associate Professor
Iscte – Instituto Universitário de Lisboa

November, 2022

Department of Quantitative Methods for Management and Economics
Department of Information Science and Technology

Influencers, are they responsible for Bitcoin's volatility? Transfer Entropy and Granger causality in prol of an answer

Jana Luisa Teixeira Lage

Master’s in Data Science

Supervisor:
Doctor Diana Elisabeta Aldea Mendes, Associate Professor
Iscte – Instituto Universitário de Lisboa

Co-Supervisor:
Doctor Fernando Manuel Marques Batista, Associate Professor
Iscte – Instituto Universitário de Lisboa

November, 2022

Acknowledgements

To God, for allowing me to have health and determination to not get discouraged during the accomplishment of this study, even with all the obstacles between student, personal and professional life.

To my family, while even living in another country, supported me to make the move and were understanding even when I was too busy pick up the phone.

To friends who were understanding when I had to turn down invitations to go out.

To my professors, Diana and Fernando, who I chose to be my advisors due to the admiration that I have, both personally, as well professionally. Thank you for supporting me and for all the help whenever I needed it.

And last but not least, to my partner Camila, who was patient with me and my bad moods in times of stress. Thank you for all the support and comprehension of not being able to give due attention in several of these focus periods for the accomplishment of this study and my degree.

Abstract

Bitcoin, like any other cryptocurrency, is subject to fluctuations in price. The volatility of this market can be a reflection of several reasons, such as public opinion, social networks and news. Social networks, in particular Twitter, are increasingly used as an important source of value extraction because through this network, it is possible to find out about news in real-time, follow the repercussions, know what experts in the financial world are commenting or thinking and even decide based on influencer's opinion whether to invest or not. This study investigates the influence that a specific group of people exert on Bitcoin volatility. A selection of influencers from the “crypto world” was made, and through the Twitter API, it was possible to select the tweets of the object of study. To choose the classification model for sentiment analysis, two techniques were compared, one being very popular with a focus on the domain of social networks and the other recently created and focused on finance. From the selected technique, only positive and negative sentiments were considered, and then the daily series of the Sentiment Score was calculated. Next, the causal relationship between Bitcoin and sentiment was investigated using Granger causality and Transfer Entropy tests. Transfer Entropy showed encouraging results, suggesting that there is a transfer of information from Sentiment to Returns and that it is possible for an influencer to contribute to Bitcoin’s volatility.

Keywords: Bitcoin, BTC, Cryptocurrency, Sentiment Analysis, Granger causality, Transfer Entropy

JEL Classification: C32, G17

Resumo

O Bitcoin, assim como qualquer outra criptomoeda, está sujeito a flutuações no preço. A volatilidade desse mercado pode ser reflexo de vários motivos, tais como, opinião pública, redes sociais e notícias. As redes sociais, em particular o Twitter, cada vez mais é utilizado como uma fonte importante de extração de valor, isto porque através desta rede é possível saber das novidades em tempo real, acompanhar as repercussões, saber o que entendedores do mundo financeiro estão a comentar e decidir até mesmo com base na opinião de um influenciador se irá investir ou não. Este estudo investiga a influência que determinadas pessoas exercem sobre a volatilidade do Bitcoin. Foi feita uma seleção de influenciadores do "mundo crypto" e através da API do Twitter foi possível selecionar os tweets de objeto de estudo. Para a escolha do modelo de classificação para análise de sentimento foram comparadas duas técnicas, sendo uma muito popular com foco no domínio de redes sociais e a outra recém-criada e focada em finanças. A partir da técnica selecionada, apenas os sentimentos positivos e negativos foram considerados e então calculada a série diária do Sentiment Score. A seguir foi investigada a relação causal entre o Bitcoin e o sentiment utilizando os testes de causalidade de Granger e Entropia de Transferência. A Entropia de Transferência mostrou resultados animadores que sugerem existir transferência de informação de Sentiment para Returns e que, portanto, é possível que um influencer contribua para a volatilidade do Bitcoin.

Palavras-Chave: Bitcoin, BTC, Criptomoeda, Análise de Sentimento, Causalidade de Granger, Entropia de Transferência

Classificação JEL: C32, G17

Index

Index of Tables

Index of Figures

Glossary

ADF – Augmented Dickey-Fuller

AIC – Akaike Information Criterion

BERT – Bidirectional Encoder Representations from Transformers

BTC – Bitcoin

CRISP-DM – Cross Industry Standard Process for Data Mining

FinBERT – Financial Sentiment Analysis with BERT

Net TE – Net Transfer Entropy

NLP – Natural Language Processing

TE – Transfer Entropy

VADER – Valence Aware Dictionary and Sentiment Reasoner

1 Introduction

Between 2013 and 2016, economists felt that the use of cryptocurrency should be avoided as it was relatively new and the risk could not be accurately quantified (Folkinshteyn et al., 2015). However, it increasingly aroused the curiosity of investors and researchers, where they began to study cryptocurrencies such as Bitcoin and discovered that there were reasons for the cryptocurrency to become a complement to fiat currencies (Carrick, 2016).

Thus, the potential of Bitcoin led to an explosive growth of interest and development in cryptocurrencies throughout 2017 and early 2018, mainly caused by news reporting the unprecedented returns of cryptocurrencies, which subsequently attracted a type of gold rush (Kraaijeveld & De Smedt, 2020).

In an attempt to identify a possible influence on Bitcoin volatility and to help investors on when to invest or not in this cryptocurrency, the following question arose: “Are the comments made by influencers in the 'crypto world' capable of influencing the volatility of Bitcoin? ”

For Silva (2016), volatility is Bitcoin’s biggest problem. A special focus in literature is devoted to cryptocurrencies' risk and statistical properties, comparing them to stocks or exchange rates. For example, Hu et al. (2019) conducted research showing that large values of kurtosis and volatility characterize the returns of more than 200 cryptocurrencies and that the leading risk factor is Bitcoin itself, which is highly correlated with many altcoins. The result shown by Barbosa (2016) concluded that Bitcoin is four times more volatile than traditional currencies. Tress (2017) points out that volatility places cryptocurrency as a risky financial asset, just like a stock or commodity, rather than guaranteeing users a value reserve.

Neves (2018) cites three factors that influence the price of Bitcoin: i) macroeconomic and financial variables, such as the dollar quotation and the stock exchange index; ii) attractiveness, increased interest in the asset due to its appreciation over the years and iii) dynamics between supply and demand. He also highlights that events linked to the activity of cryptocurrencies and exogenous macroeconomic events significantly influence the price of Bitcoin.

In Bollen et al. (2011) it is pointed out that although news influences the stock price, the mood and sentiment of the public should also play an equally important role.

However, as shown in Kim et al. (2016), sentiment polarity contains changes that precede Bitcoin (BTC) price fluctuations. This result can vary its intensity according to the type of market regime.

There is a growing movement in the industry to analyse and make predictions based on social media data. With the accumulation of data and the development of new tools to analyse them, data mining and machine learning techniques are most frequently used to understand the relationship between human behaviour and financial market trends (Karppi & Crawford, 2016).

Twitter is currently one of the most used micro-blogging platforms among all existing social networks. Through Twitter, it is possible to check in real time what people are discussing, the volume of people commenting, and the repercussions of the news. With this type of data, it is possible to explore the behaviour and feelings of the users. For the investment market, for example, the possibility of verifying the relationship between people's common opinion and the movement, trading volume, and other financial data about a given asset is a valuable source to support decision-making (Kristoufek, 2013; Bukovina & Marticek, 2016).

Sentiment analysis, measured by natural language processing techniques applied to unstructured data, such as those available via the Twitter API, has been shown to have significant predictive effects on stock price movements (Bollen et al., 2011; Souza & Aste, 2016).

The time series predictive analysis is a technique capable of generating profitable trading strategies. Keskin (2018) carried out, through Transfer Entropy, a study to detect whether the social sentiment was causally related to price movements in four cryptocurrencies and observed that there was a greater transfer of information from price to sentiment.

In recent years, recurrent studies have been carried out on sentiment analysis (Barbosa & Feng, 2010; Liu, 2012; Hutto & Gilbert, 2014) associated with stock prices (Bollen et al., 2011; Li et al., 2014) and cryptocurrencies forecasting (Georgoula et al., 2015; Bukovina & Marticek, 2016; Kim et al., 2016; Steinert & Herff, 2018; Hamza, 2020 Kapar & Olmo, 2020), through various techniques, but few with a focus in the Transfer Entropy methodology. Therefore, it seems to be necessary to study this exchange of information and assess the direction of this flow of information.

The main purpose of this dissertation was to investigate the influence that cryptocurrency opinion-makers or influencers exert on Bitcoin volatility. To capture the influence, that is, whether the opinions of influencers had more positive or negative impact in Bitcoin, the technique of sentiment analysis of the FinBERT classifier was used. Moreover, in order to explain the causality between Bitcoin volatility and sentiment, Granger Causality (Granger, 1969) and Transfer Entropy (Schreiber, 2000) methods were used, with a primary focus on the latter.

With the main result obtained, the Transfer Entropy proved to be efficient in quantifying the flow of information between the considered time series and identifying the causality direction. It was identified that the sentiment towards Bitcoin transferred more information and that the influence of opinions made by financial market experts, specifically in cryptocurrency, can impact what the investor should do within three days of writing a tweet.

This thesis is structured as follows: the first chapter introduced the motivation, problem and the objective of this work. In the second chapter, a literature review was carried out for the theoretical and scientific basis. In the third chapter, an explanation on how Bitcoin, Twitter and the indicators created by investors were extracted, as well as the purpose of each one. The fourth chapter contains the methodologies used for extracting texts and test whether there is causality between the time series. The fifth chapter contains the results obtained and, finally, the conclusions and suggestions for future work.

2 Literature Review

This chapter analyses a number of scientific literature in order to identify gaps at a level of investigation as well as to obtain knowledge in relation to the proposed theme. For this, Bitcoin was researched, its history, what it is used for and what led to its appearance. Furthermore, the use of economic and social media indicators to assist in Bitcoin prediction. Finally, the use of sentiment analysis combined with causality to predict the volatility of the Bitcoin.

2.1 Bitcoin

For long periods, the evolution of monetary systems was significant. At every stage, from bartering to metallic money, paper, credit, plastic money, economic necessities, and the comparable efficiency of completing a transaction constituted the factors for the continuation of one form of money over another.

The banking system did not change profoundly with the financial crisis of 2008, despite the growing attention that institutions and the most varied banking corporations received with the scandals and alleged cases of corruption of the time, many more regulations were put in place, but attitudes of banks and even the attitudes of tax authorities have not been broadly modified (Viorica, 2013). Growing global resentment towards the current monetary system persists, especially after the financial turmoil that hit financial markets between 2007 and 2009, causing severe economic impacts worldwide (Almansour et al., 2021). There are also the “elephants in the room” about banking and fiat money: costs per transaction, transaction processing speed, and central power oversight over them.

In this context, alternatives arise, one of which is cryptocurrencies also called altcoins, because of their disruptive nature within the mindset of the modern banking system and established fiat money printing (Luther, 2015). Cryptocurrencies are mainly characterized by fluctuations in their price and number of transactions (Böhme et al., 2015).

The first cryptocurrency created, Bitcoin (Nakamoto, 2008), was launched in 2009 and aimed to create a radical change in the global monetary system. A few years later, the Bitcoin idea significantly took public interest into account. The world has witnessed rapid growth in the cryptocurrency market. The market capitalization of cryptocurrencies has hit record highs repeatedly. This phenomenon reveals the significant value of cryptocurrencies as an electronic payment system and financial asset (Bacilar et al., 2017). Therefore, which asset class cryptocurrencies belong to is difficult to define as they share characteristics of several existing asset classes. It could even be argued that cryptocurrencies form an entirely new asset class (Kraaijeveld & De Smedt, 2020).

Without delving into the nature of cryptocurrency, it is necessary to say that it is a type of currency that was born from decentralized blockchain technology that allows transactions over the network created as a consequence of the contact made between the holders of cryptocurrency used without any centralizing authority, or the amount of this currency issued (Catalini & Gans, 2020). This means that cryptocurrencies have incomparable advantages over the current traditional system, such as decentralization, strong security, and lower transaction costs. It can be stated that facilitating the creation of a borderless money transaction across the globe was one of the main purposes of crypto (Nakamoto, 2008).

2.2 Financial Indicators: A Prediction Aid

Vidotto et al. (2009) explain that technical indicators that are used to understand the price movement of traded assets, for example, on the stock exchange, are of crucial importance, as they can help identify trends and their reversal points.

According to Kim et al. (2021), the most representative cryptocurrencies (Bitcoin and Ethereum) cannot be predicted based on general economic indicators as is done for existing currencies or gold.

Based on the objective of the study by Pang et al. (2020) to create a consolidated analytical model to predict Bitcoin price movement and provide actionable signals to investors, the methods used for price prediction in the regular stock market may not work well for the cryptocurrency market, due to differences between them.

Furthermore, Ciaian et al. (2016) found that market-level indicators such as transaction volumes and user activity on Bitcoin community sites (e.g., number of posts, discussions and new members) are significantly associated with Bitcoin futures prices. As with other assets such as stocks and precious metals, the price of cryptocurrencies is influenced by several factors, for example, fake news, market manipulation, and government policies (Zhang et al., 2021). According to Lahmiri and Bekiros (2020) it becomes more difficult to assess during a pandemic, when there is a higher level of risk for the cryptocurrency market than for the stock market. In addition to the pandemic, there are other factors such as: war, natural disasters and other situations that cannot even be predicted or controlled, meaning that, digital currencies in these periods present greater instability, more significant irregularity and volatility, indicating a decline in the attractiveness of their investments in these periods.

Other studies have reported that macroeconomic factors such as stock indices, changes in crude oil and gold prices, global exchange rates, and Bitcoin blockchain information are critical factors in predicting the fluctuation of Bitcoin prices (Jang & Lee, 2017; Mallqui & Fernandes, 2019). It is also necessary to take into account the importance of studying the risk metrics of cryptocurrencies and estimating them since they are in a fast pace of development, and also because the volatility of cryptocurrencies usually is high, especially during high demand phase (Almansour et al., 2021).

Therefore, the usual technical drivers (bitcoin supply and demand), attractiveness indicators, and macroeconomic variables seem to have become lagged or no longer significant indicators to explain bitcoin price dynamics (Kapar & Olmo, 2020).

2.3 Social Media: A Prediction Aid

As the cryptocurrency market is relatively new, traditional media do not always report events on time, which makes social media the primary source of information for cryptocurrency investors, particularly Twitter (Kraaijeveld & De Smedt, 2020). Not only does Twitter provide real-time updates on the cryptocurrency world, it is also a rich source of emotional intelligence, as investors often express their sentiments, and with this, it can profoundly affect individual behaviour and decision-making (Kraaijeveld & De Smedt, 2020).

Due to the popularity of Twitter - a social network in the form of a micro-blog created in 2006, sentiment analysis in tweets has attracted more attention (Parikh & Movassate, 2009; Barbosa & Feng, 2010), followed by machine learning approaches for sentiment analysis of tweets. Over the past decade, there has been significant progress in sentiment tracking techniques that extract audience mood indicators directly from the social media content, such as blog content (Liu et al., 2007) and, in particular, large-scale Twitter feeds (Pak & Paroubek, 2010).

The work of Karalevicius et al. (2018), sought to measure the interaction between the price of Bitcoin and the sentiment expressed in the media daily through articles and news related to the topic. Sentiment analysis was based on the Harvard Psychosocial Dictionary and a dictionary focused on finance. This study proposed a strategy of trading based on sentiment measured during the day. Karalevicius et al. (2018) found a relationship between price and measured sentiment, where, after the publication of impacting news, the price initially goes in the direction of the sentiment expressed in the news, but the market tends to react with an impact toward correction movement. From a study on the factors that influence the excessive volatility of Bitcoin, Bukovina and Marticek (2016) showed that positive sentiment is more influential for the excessive volatility of Bitcoin.

Kaminski and Gloor (2014) analysed the present sentiment together of tweets and applied Pearson correlations to look for possible correlations between the price and daily volume of the currency. They found that negative tweets and the tweets related to a sentiment of uncertainty had a proportional relation towards the trading volume and an inverse relation towards the price. Georgoula et al. (2015) used time series analysis to study the relationship between the prices of Bitcoin together with sentiment analysis on Twitter and searches on

Wikipedia. The author used Support Vector Machines (SVMs) as a classifier, where a series of regressions showed that sentiment and public interest were positively correlated with currency prices.

Hamza (2020) also did a study to find out if Bitcoin prices could be predicted by using tweets from Twitter opinion leaders through a quantitative research method, utilizing a Vector autoregression model, and performing a time series analysis. The results indicated that the tweet sentiments impact the return and the volatility and that there is a Granger causality relation. Although sentiment does not Granger-cause volume, an impulse to sentiment impacts volume (Hamza, 2020). Nevertheless, Abraham et al. (2018) showed that sentiment analysis for cryptocurrencies is less effective in a bear market as overall Twitter data volume activity declines.

Results indicate that tweets by cryptocurrency influencers contain statistically significant information about the future value of Bitcoin, because these users can influence the decisions of other users (Hamza, 2020).

In addition to Twitter, the structure of Reddit forums has also been used to model audience sentiment and identify influencers (Wooley et al., 2019). Wooley et al. (2019) also showed a medium-term positive correlation between price and Reddit online activity and that such a relationship supports the validity of cryptocurrencies as speculative assets.

Previous studies have attempted to predict the price of Bitcoin and found that Google Trends searches for the keyword “Bitcoin” and Wikipedia views about Bitcoin are positively associated with Bitcoin price fluctuations. However, Kristoufek (2013) found that Google Trends predicted Bitcoin prices more accurately than Wikipedia views.

Later, according to research by Kim et al. (2021), it was seen that Google Trends search queries and Bitcoin tweet volumes are positively associated with Bitcoin price fluctuations, however, tweet volumes had a slightly higher positive correlation with Bitcoin price fluctuations than Google Trends. Furthermore, the association with the number of topics posted daily indicated that variation in community activities could influence price fluctuations, and unlike the price of cryptocurrencies, the number of transactions proved to be significantly associated with user responses rather than comments posted (Kim et al., 2021).

At a behavioural level, a study by Garcia et al. (2014) verified whether people's opinions and feelings contributed to the emergence of a bubble in the price of Bitcoin. For this study, collective behavioural traits and social phenomena were verified on Twitter, Wikipedia and Facebook. By analyzing user-to-user information shared on these three platforms, Garcia et al. (2014) found a social cycle and a user adoption cycle. In the social cycle, users exchanged information with each other, and in the adoption cycle, new users began to adopt the currency. As a result of this research, Garcia et al. (2014) observed that spikes in the search for information, presumably linked to external events, proceeded to drastic reductions in currency prices.

As analysed by Kim et al. (2021), the use of data on the Web (Cohen et al., 2013), the analysis of data from social networks (Bollen et al., 2011) and the reference to research volumes on Google lead to more accurate results. Furthermore, as the volatility of cryptocurrencies can be impacted by public opinion, influencers play a crucial role on social media as they have the ability to move the market with enough influence to guide the direction of public opinion (Hamza, 2020).

2.4 Sentiment Analysis

Sentiment Analysis, also known as Opinion Mining, is a field of research that studies and analyses people's opinions, feelings, evaluations, attitudes and emotions. (Liu, 2012). Sentiment analysis, or opinion mining, is considered one of the most active research areas in natural language processing. It is widely studied in data, web, and text mining (Liu, 2012).

In sentiment analysis, a sentiment lexicon must be defined initially, which is a set of words that express feelings. Defining this set of words is laborious, so there are already several lexicons of feelings that have been constructed. In addition, the task of automatic recognition of feelings in texts becomes more complex due to the so-called "noises", which is the mixture of objective and subjective information about a given topic. These noises, which are easily found in most tweets, range from simple expressions to complete sentences (stopwords, ironies, etc.), making the cleaning/modification of these words necessary with specific techniques (Anjaria et al., 2014).

To address the challenges of sentiment analysis on blog-like content and with the goal of developing an algorithm that is computationally fast for streaming data, without requiring a great deal of training and due to the fact that social media analysis is not yet generalizable, the Valence Aware Dictionary and Sentiment Reasoner (VADER) was developed (Hutto & Gilbert, 2014). This new lexicon proposed by Hutto and Gilbert (2014), was built with a focus on the domain of microblog-type social networks, such as Twitter.

For the evaluation of text-mining techniques, VADER was considered a suitable tool for sentiment mining and was used in previous research by Steinert and Herff (2018). VADER was originally a Python library used for dictionary-based sentiment analysis but has been ported to different languages and platforms. VADER (Hutto & Gilbert, 2014) has some characteristics such as: a focus on microblogging; the classification of slang language expressions, abbreviations and emoticons; valence score assignment for intensity feelings; and higher accuracy compared to human classifiers.

VADER has been successfully applied to tweets and online forum content with a F1-score (measure of classifier performance) rating accuracy of 96% compared to 84% of human raters (Hutto & Gilbert, 2014) and has been used for the purpose of predicting cryptocurrency prices (Kim et al., 2016; Kristoufek, 2013).

In the same year, in 2018, Google developed a machine learning framework for NLP (Natural Language Processing) called BERT (Bidirectional Encoder Representations from Transformers) that uses two tasks: masked language modeling and next sentence prediction. This algorithm was developed to perform complex context-based modeling tasks with little effort.

However, the use of general-purpose sentiment classification models is not sufficiently effective in the financial context due to the specialized language and scarcity of data labeled sentiment analysis, which makes the task challenging. This challenge occurs because financial texts have a unique vocabulary and tend to use vague expressions rather than easily identifiable negative/positive words (Araci, 2019).

Since this project deals specifically with texts about cryptocurrencies, that is, related to the financial sector, sentiment classification models were sought for this sector. This is because financial sentiment analysis differs from general sentiment analysis regarding domain and purpose (Araci, 2019).

Domain-specific models emerged using BERT as a basis and are used for NLP tasks: ClinicalBERT for clinical notes, BioBERT for Biomedicine and FinBERT for Finance.

Until the release of FinBERT, there were no pre-trained finance-specific language models available (Yang et al., 2020), this is because NLP models require large amounts of labeled training data and the application of deep learning for data mining. Financial text is generally unsuccessful due to the financial sector's lack of labeled data (Liu et al., 2021).

Choosing the most appropriate classifier for the analysis of sentiments in tweets is of paramount importance, as it must be able to automatically label new texts relating them to the trained content and guarantee the best possible assertiveness rate of the labels, according to your sentiment.

2.5 Granger Causality and Transfer Entropy

By implementing a specific cryptocurrency, lexicon-based sentiment analysis approach in combination with bivariate Granger (1969) causality tests, it was found that Twitter sentiment can be used to predict Bitcoin price returns (Kraaijeveld & De Smedt, 2020).

In Granger causality, despite the success in identifying couplings between interacting variables, the use of structural models restricts their performance (Gençağa, 2018). That means that in order to analyse Ganger causality correctly, some statistical requirements on time series are fundamental, namely, the stationarity of the stochastic process. Transfer Entropy is a quantity estimated directly from data and does not suffer from such restrictions, however, estimating Transfer Entropy from data is a numerically challenging problem (Gençağa, 2018). According to Schreiber (2000), Transfer Entropy can distinguish conduction and response elements and detect asymmetry in the coupling of subsystems. Furthermore, unlike Granger Causality, Transfer Entropy uses probability rather than system identification to determine causality (Schreiber, 2000).

Transfer Entropy can be thought of as a nonlinear generalisation of Granger causality (Barnett et al., 2009). The Transfer Entropy is a measure of the flow of information between the variables, that is, it is possible to measure the amount of information that a random variable contains about another. This information transfer quantifies the uncertainty that is reduced in one variable from the knowledge provided by the other variable (Schreiber, 2000).

Transfer Entropy has already been applied in several areas, such as in Medicine to quantify the interaction between maternal and fetal heartbeat rates (Marzbanrad et al., 2015), in Electrical Engineering in the diagnostic criterion of the occurrence of a possible blackout (Milligen et al., 2016), in Social Sciences, where social media data was analysed using Transfer Entropy to investigate the dynamics of collective social phenomena (Borge-Holthoefer et al., 2016), and in Economics for financial returns (Syczewska & Struzik, 2015) among others.

Recent studies have used Granger causality and Transfer Entropy in stocks and cryptocurrencies with social media data. Hamza (2020) used the Vector autoregression model, and performed a time series analysis, where the results indicated that the sentiments of tweets impacted the Bitcoin's return and volatility, which indicated that there was a Granger-cause relationship.

Another result of the study by Kraaijeveld and De Smedt (2020), was that the strongest predictors at the daily level are Twitter sentiment and message volume, while price returns are the strongest predictor variable at the intraday level, thus, one might suggest that Twitter causes, rather than follows, the cryptocurrency market. However, this difference is marginal as there are several cases in which price returns cause sentiment, mostly occurring at the intraday level (Kraaijeveld & De Smedt, 2020).

One of the experiments done by Keskin (2018) was to apply the Transfer Entropy methodology to find out the impact that social sentiment has on four popular cryptocurrencies, including Bitcoin. Using the non-linear Transfer Entropy calculation, a significant signal of information transfer was discovered between sentiment and Bitcoin returns with lags greater than one day and for both directions. However, Keskin (2018) noted that the significance of information transfer is greater than the market return to social sentiment, both in linear and non-linear information transfer tests.

3 Data Collection

For the development of this project, the daily closing values of Bitcoin were used, that is, the last price recorded on the day, Twitter data, such as tweets, date of tweet creation, number of likes, replies and retweets. In addition, financial indicators such as: Hash Rate, Crypto Fear & Greed Index, BTC Dominance and Mayer Multiple.

This thesis was based on the CRISP-DM (Cross Industry Standard Process for Data Mining) methodology, as explained in chapter 4. In this chapter, after obtaining the data, it was a matter of exploring the data, cleaning it and transforming it into suitable formats in order to prepare this data for the next steps.

3.1 Bitcoin Time Series

For the historical data of Bitcoin's daily closing values, the period from when the information was made available by the Glassnode [3] was considered, that is, from 2010-07-17 to 2022-06-15 (Figure 3.1), which corresponded to a time series with 4,351 days. However, due to Twitter data, this time period has been shortened, as explained in the following section.

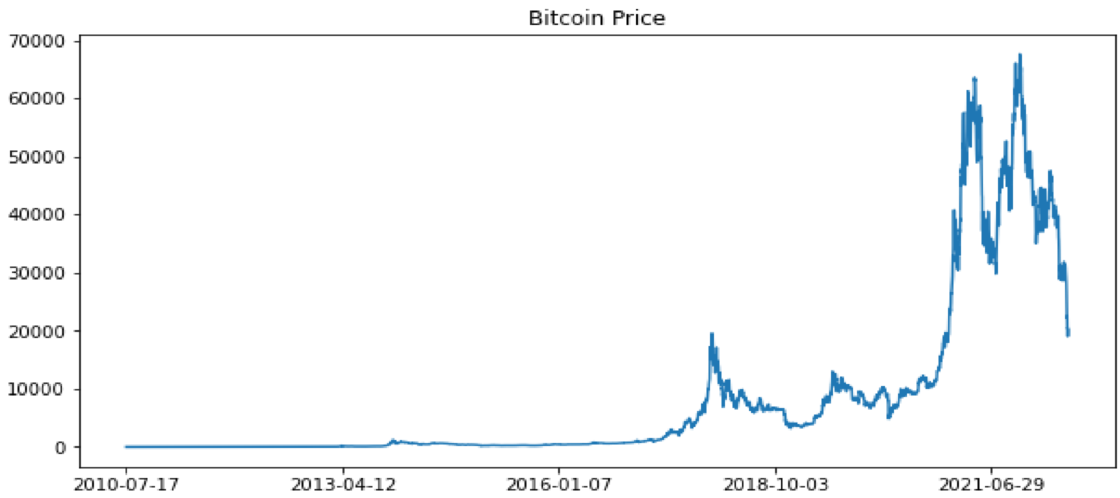

Figure 3.1 - Bitcoin price history provided by Glassnode

Bitcoin reached its maximum value of about US$65K on October 20, 2021. However, until the collection of this data on June 15, 2022, there was a decrease which made Bitcoin register the value of US$22K, which was also approximately recorded in November, 2020. The standard deviation in this period of Figure 3.1 was of US$15K, the median was US$672 and 75% of the prices registered were below US$8K. Such discrepant values are seen by the graph that registered the highest values at the end of the historical series.

3.2 Twitter

To collect the information from the tweets, a search was initially done for influencers in the “crypto world”, that is, which people have some kind of influence with their comments to be considered. For the choice of influencers, some considerations were made such as, which names are frequent in searched lists, if the content of the lesser-known ones (less followers) is mostly about crypto analysis and if they have engagement (likes and replies). Overall, 46 influencer accounts on Twitter were considered.

The Academic Research Twitter [4] account was used, where it is possible to extract up to 10M tweets per month, which is considered the best type of Twitter API account. However, when proceeding with extracting the Tweets from the selected accounts, some limitations were encountered. According to the documentation of the API reference [4], by default, the most recent tweets come back per request, but by using pagination, the most recent 3,200 tweets can be retrieved. Therefore, when searching for tweets for a specific account, it is only possible to retrieve the last 3,200 tweets and consider the replies. By excluding the replies, this number decreases to 800. That is, there are some situations: when considering the 3,200 tweets it may be possible to collect all the tweets from an account, but this means that in a long period of time, such as since 2010-11-06 (period for which the Twitter API makes the data available), there are not many tweets, and/or the account is recent. For most selected accounts, these have more than 5,000 tweets.

Another point is that accounts that have more than 3,200 tweets and that have many replies somehow fail to reach the main focus: the influencer's opinion about Bitcoin. This is because if a Twitter influencer wants to express their opinion, it will do so on its timeline so that it has a wide reach and not just with a specific reply to a person that not everyone will see. Furthermore, when considering the replies, the time period of the tweets is reduced and with the risk of not even capturing something related to the objective. So, without considering the replies, it is possible to get a larger data range window. With this, the conclusion was reached that the replies would not be considered when extracting the tweets. The total number of tweets collected was 35,960. After the tweets were collected, the tweets were filtered considering the following keywords: bitcoin and btc. This resulted in a final Tweet database of 12,246.

3.3 Hash Rate

The hash rate [3] is one of the most important concepts in the world of crypto investors. There is a relationship between the hash rate and cryptocurrency prices, although it is not always clear.

It is a measure of the total processing power used by computers that are mining a cryptocurrency and recording transactions on a blockchain network. The unit can also be used to scale the speed at which mining machines are completing complex mathematical calculations.

As presented by Tress (2017), miners are programmers who legitimize and build the Blockchain. Silva (2016) explains that blocks are added to the blockchain linearly and in chronological order. Tress (2017) says that for each successful mining, the miner receives a payment per operation, which is readjusted every 210,000 blocks mined or four years, which, according to Aragão (2016), is to stop the issuance of coins and avoid possible inflation.

Through the hash rate it is possible to know the security and integrity of a cryptocurrency's network. Because it is possible to identify, for example, how many computers are being used to maintain the blockchain and how many hashes are generated in a given time interval. Therefore, the higher the hash rate, the more secure the network.

In addition, this fee is one of the factors used to assess how many transactions are carried out, how many miners are active and also how profitable mining is. All of this impacts transaction fees, charged every time a cryptocurrency is moved, for example.

For the historical data of the hash rate, the Glassnode [3] API was used, in which the period considered was the same as the tweets.

3.4 BTC Dominance

Bitcoin Dominance, or BTC Dominance, is an important metric that measures Bitcoin's share of market capitalization relative to other cryptocurrencies, so it is calculated by dividing Bitcoin's market capitalization by the total market capitalization of all cryptocurrencies. Through this metric, it is known whether altcoins are performing better, worse or similar to Bitcoin, since Bitcoin has always had the highest dominance and thus tends to absorb market information faster.

Historically, Bitcoin Dominance has been a good indicator of where the market is and where it is going, which is why some cryptocurrency investors and traders use this indicator to adjust their trading strategies.

For the calculation of this metric, 9,816 active cryptocurrencies in the same period of extraction of the tweets were considered, through the CoinMarketCap [2].

3.5 Crypto Fear & Greed Index

The Crypto Fear and Greed Index, created by Alternative.me [1], calculates how the industry feels specifically towards Bitcoin. The index varies between 0 and 100, where the closer to 0 the greater the investors' concern about this asset in the future. This means extreme fear tends to lower prices, while extreme greed does the opposite. The Index is defined as follows in terms of the score range: between 0-24 points for extreme fear, 25-49 for fear, 50-74 for greed and above 75 points for extreme greed.

Alternative.me [1] (the API was used for this project), describes: “With our Fear and Greed Index, we try to save you from your own emotional overreactions” and their conclusions are based on two assumptions “People tend to get greedy when the market is rising which results in FOMO (Fear of missing out). Also, people often sell their coins in the irrational reaction of seeing red numbers”.

According to Alternative.me [1], the Fear and Greed Index is made up of five components:

• Volatility corresponds to 25% of the index: which measures the volatility of the current Bitcoin price and compares it to the 30-day and 90-day averages. They argue that an unusual increase in volatility is a sign of a fearful market.

• Market Momentum/Volume (represents 25% of the index): this measures Bitcoin's current trading volume and dynamics, and also compares it to the 30-day and 90-day averages, then combines the results. When you see high daily buying volumes in a positive market, you can conclude that the market is too greedy.

• Social Media (represents 15% of the index): it refers to the analysis of Twitter hashtags with a focus on Bitcoin, with a focus on speed and number of interactions. A higher-than-normal interaction rate represents a growing public interest in the coin which means greedy behaviour in the market.

• Dominance (10%): is associated with an increase in fear, for example, the increasing dominance means that funds are being withdrawn from riskier altcoins, on the assumption that Bitcoin is seen as “the safe haven of cryptocurrencies”. A decrease in the BTC dominance suggests the growth of greed represented by investments in risky currencies.

• Trends (10%): are based on Google Trend data analysis for various Bitcoin-related searches such as changing search volumes and other popular searches.

• Surveys (15%) currently paused: are weekly cryptocurrency surveys conducted by asking people how they view the market. Not much attention is currently given to these results, but it was quite helpful at the beginning of the studies.

The data extracted from the Alternative.me [1] was only available from February 2018, when this index was created.

3.6 Mayer Multiple

The Mayer Multiple is an indicator created in 2017 by investor Trace Mayer to help other investors analyse the price of Bitcoin. Some traders believe this indicator can identify speculative bubbles and low exhaustion moments.

The Mayer Multiple is the current price of bitcoin divided by its 200 daily MA (moving average), that is, the index compares the current price with the 200-day MA. This quantifies how close or far the current price is from the 200-day MA.

If the formula obtains a value lower than 1, it means that the current price is below 200 MA, and the Bitcoin market is in a low phase. A value greater than 1 means the current price is above 200-MA, so Bitcoin is in a high phase.

For the calculation of the Mayer Multiple, historical data previously extracted from Glassnode related to the closing prices were considered.

4 Methodology

This thesis is summarised in the following steps: the first step is the calculation of the daily sentiment through the technique provided by the FinBERT classifier which generates the sentiment time series. In the second step, Granger causality testing was carried out in order to identify the most expressive series between the indicators used by cryptocurrency investors and their sentiment, but considering the sentiment as primary (and not Bitcoin, since the indicators already consider Bitcoin data to compose them, and the focus of the thesis is on the influence of sentiment on the currency), and detect which lag to consider to predict cryptocurrency price fluctuations. In the last step, the Granger causality results are considered in Entropy Transfer application through the PyCausality package available in the Python library (Keskin, 2018).

According to Keskin and Aste (2020), there was no simple way in Python to detect causality using information theory techniques for predictive analysis. The closest approach requires a java function call, while PyCausality provides a Python-only interface, extending the functionality of familiar Pandas DataFrames to calculate transfer entropy between time series in a Pythonic manner. In this way, this package provides a simple and intuitive interface to explore coupled time series and to detect causality using autoregressive and model-free techniques to estimate linear and nonlinear transfer entropy (Keskin & Aste, 2020).

The CRISP-DM methodology was used in this thesis. This is a data mining methodology that defines the life cycle of a given project in order to direct the discovery of knowledge in decision making and is divided into the stages of Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation and Deployment (which is not applicable here). The Business Understanding stage corresponds in this thesis to the Literature Review (Chapter 2). Data Understanding and Data Preparation are discussed in the Data Collection (Chapter 3). Modeling and Evaluation in Results (Chapter 5) and Conclusion (Chapter 6).

4.1 Sentiment Score

After obtaining the set of textual documents, pre-processing techniques are applied to format and structure the texts, without removing their natural characteristics. The quality of the collected texts is a fundamental step in text mining processes.

In its original form, texts may contain noise, lack of language standards, inconsistency, and/or redundancy of information, which can affect the algorithms used in text mining processes, generating misleading and unrepresentative results (Tan, 2018). In most cases, pre-processing the text is a fundamental procedure to prevent unsatisfactory results in textual analysis tasks from being obtained (Alexander & Jetton, 1996). This is due to the fact that there are large amounts of words responsible solely for textual cohesion, such as: prepositions, conjunctions, articles, adverbs, numbers, pronouns, and punctuation. Pre-processing performs the treatment and cleaning of these words to try to avoid distortions in the results (Aggarwal & Zhai, 2012).

Nevertheless, VADER doesn't require much pre-processing because its algorithm englobes most of these steps. Furthermore, VADER also understands the valence of non-conventional text, including emojis, capitalization (i.e., “sad” vs “SAD”), extended punctuation (i.e., “!” vs “!!!!”), unlike some supervised methods of NLP (Hutto & Gilbert, 2014).

Unlike many other NLP algorithms, which require pre-processing inputs, the BERT algorithm takes entire raw sentences as inputs in fine-tuning, as well as in pre-training (Huang et al., 2022). Researchers who use non-BERT algorithms often need to decide the model structure, such as the number of layers and filters and size of word embedding, while FinBERT has its model structure fixed which makes it simpler (Huang et al., 2022).

VADER is more consolidated in literature with several publications compared to FinBERT, which is more recent and with fewer publications. Therefore, in order to test whether or not there would be an improvement in the classification of sentiments, it was decided to try to apply some pre-processing steps to FinBERT since this algorithm uses the texts in its original form.

Among the various text pre-processing techniques existent in literature, the following were applied: conversion of all characters to lowercase; expansion of contractions (example: “don’t” to “do not”); removal of URLs, numeric characters, punctuation, “RT”, @user, hashtags, and stopwords (these are important for understanding sentences, but alone do not bring meaning and at a computational level it takes up memory space and increases processing time). Additionally, the stemming technique was applied, which reduces inflected (or derived) words to their base or root word. Furthermore, lemmatization makes the grouping of inflected words to be analysed as a single word through its lemma.

After performing the previous steps, sentiments were calculated and classified using both methods.

In classifying sentiments with VADER, each tweet was categorized as “pos” (positive), “neg” (negative), “neu” (neutral) and “compound”. Scores categorized as positive, negative, and neutral are used when multidimensional sentiment indicators are needed. The “compound” measure is calculated by adding the valence scores of each word in the lexicon and normalized to be between -1 and +1, so this measure is more suitable for establishing a classification with standardized limits.

According to Hutto and Gilbert (2014), the typical values used in literature for the classification of sentences are:

positive sentiment: compound >= 0,05
neutral sentiment: (compound > −0,05) and (compound <0,05)
negative sentiment: compound <= −0,05

Differently to VADER, FinBERT does not provide a "compound" measure. FinBERT’s library returns three values in the following order: positive, negative, and neutral. The text classification corresponds to the one that obtained the highest value among the three possible classifications.

4.2 Granger Causality and Transfer Entropy

The causality analysis aims to trace a fault propagated through a process to its root cause. Transfer Entropy and Granger causality can be calculated between every pair of variables (Lindner et al., 2019).

The Granger causality test is used to determine whether a lagging variable can be introduced into an equation containing other variables, then if a variable is affected by the lag of other variables, the variables are considered to have Granger causality (Yao and Li, 2020). As mentioned by Lindner et al. (2019), the significance test is used to remove connections that fail the significance test and can be used to build an adjacency matrix.

The most common definitions of Granger causality rely on predicting a future value of the variable Y using the past values of X and Y. In this form, X is said to Granger-cause Y if the use of X improves the prediction of Y (Pierce & Haugh, 1977).

Consider a linear projection of y_t on past, present and future $x's$,

$$y_t = c + \sum_{j=0}^{\infty} b_j x_{t-j} + e_t \quad (1)$$

Where $E(e_t x_t)$ = 0 for all t. Then x fails to Granger-cause y iff b_j = 0 for j = 1,2,…

Therefore, in order to test whether the series in this thesis manage to temporarily precede each other, the null hypothesis of non-causality of Granger was tested between Bitcoin and Sentiment and between Sentiment and “crypto world” indicators (In Figure 4.1 shows the behavior of each "crypto world" indicator versus Bitcoin prices) at the 5% significance level. The objective with the Granger causality test was to identify possible lags for Transfer Entropy, and also to identify which series could be related to the Sentiment Score. Before proceeding with Granger causality, it was necessary to check whether the series were stationary. For this test, the Augmented Dickey-Fuller (ADF) test was used.

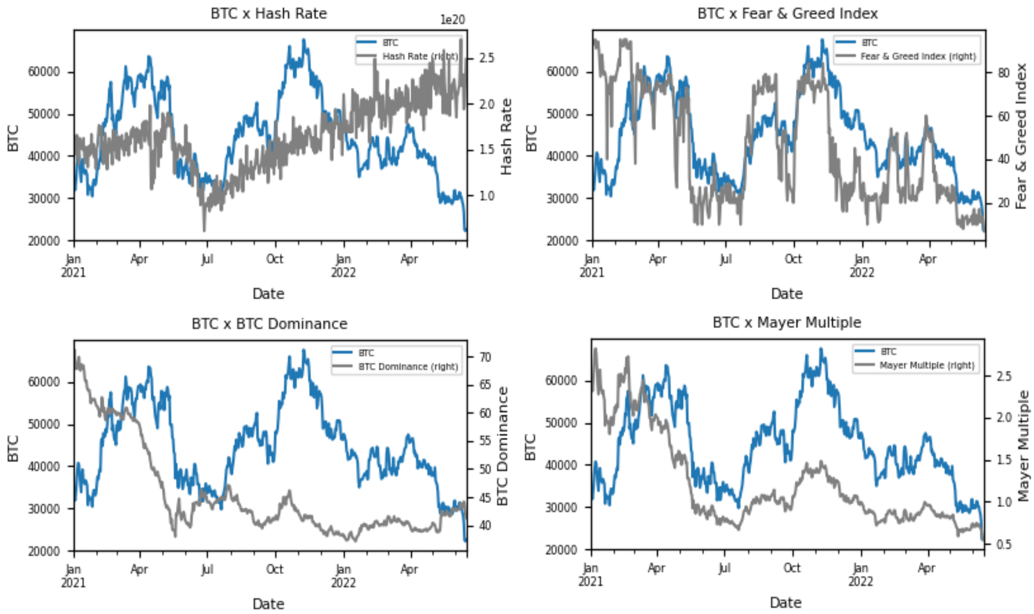

Figure 4.1 - Plot of the series between BTC and the BTC indicators used by investors

For Transfer Entropy, the PyCausality package in Python was used to calculate the series addressed, following the same parameters used by Keskin (2018), who was the main reference for this study, and who used the Entropy methodology defined by Shannon (1948).

Transfer Entropy is defined as

$$T_{X\rightarrow Y} = H(Y_t|Y_{t-l}^{(t-k)}) - H(Y_t|Y_{t-1}^{(t-k)}, X_{t-l}^{(t-l)}) \tag{2}$$

Where, the first term represents the uncertainty about Y_t given Y 's past. The second term represents the smaller uncertainty when X's past is also known.

Therefore, for Transfer Entropy it was necessary to define which parameters would be used, otherwise default values would be applied. Following parameters were used:

lag - describes the lagged term;

window size - represents the size of the window;

window stride - defines the step between consecutive windows;

pdf_estimator - histogram or KDE;

bins - equiprobable, sigma, MIC or knuth. If the parameter is none then the sigma binning is called.

The choice of estimator was important for the results in this thesis. As justified by Keskin (2018), due to susceptibility to bias and providing discontinuous and inconvenient probability distributions, the histogram becomes disposable and undervalued. However, this discard is more applicable to general density estimation problems, and of lesser importance in the estimation of transfer entropy, which is fundamentally a discrete measure, and in which systematic biases are expected, in general, to cancel out (Keskin, 2018). Finally, Keskin (2018) concludes that parameter selection is important and, as the histogram is significantly more efficient, it becomes an attractive option.

After defining the parameters, Transfer Entropy was calculated and it was verified the need to transform some series in stationary and then Transfer Entropy was calculated again. Through significance, it was possible to identify at which points the exchange of information actually took place. Also, Charts were generated for the amount of lags that could indicate which lag to use for a better Transfer Entropy. Finally, the Net Transfer Entropy measure was used to calculate which direction of information between the series was stronger.

Then, to summarise the directionality of the flow of information between time series, Net Transfer Entropy was used, defined as the difference between Transfer Entropy in both directions:

$$TE_{X\to Y} - TE_{Y\to X} \tag{3}$$

If $TE_{X\to Y} - TE_{Y\to X} > 0$, it means that the direction of transfer is predominant in the minuend and negative otherwise, that is, the flow of information from Y to X, otherwise, the flow of information from X to Y is considered (He & Shang, 2017).

5 Results

This chapter presents the results obtained by implementing the considered methodologies for the Bitcoin and sentiment time series. Python software was used for the data extraction and analysis, as well as for all tests performed and generated images.

With the main objective of knowing if the sentiment score is related to Bitcoin in order to be able to predict its volatility, classifiers were compared for sentiment analysis and from this result the relationships between sentiment score and Bitcoin were tested through Granger causality and Transfer Entropy.

5.1 Sentiment Analysis

Given that sentiment analysis is a domain-dependent task, firstly, a comparison of the two previously described sentiment classifiers was made in order to identify which one would be the best choice for the domain under study based on the analysed data.

For each tweet, VADER produces four different measures, where the “compound” measure combines all the other three measures and outputs a number between -1 and 1. Table 5.1 presents some examples of tweets and the corresponding VADER sentiment classification. In the scope of this study, only the “compound” measure was used.

Table 5.1 - Example of tweets and respective classification with VADER

User	Tweet	%Pos	%Neu	%Neg	Compound
LayahHeilpern	The UK is currently debating another lockdown while everyone in Miami is celebrating freedom, decentralisation and empowerment #bitcoin	0.331	0.669	0	0.836
danheld	Bitcoin is freedom money.	0.583	0.417	0	0.6369
saylor	Billions of people have not yet discovered #Bitcoin.	0	1	0	0
danheld	Buying US Dollars with your #Bitcoin is extremely risky in my opinion.	0	0.84	0.16	-0.2716

In order to analyse the accuracy of the two classifiers VADER and FinBERT, 162 tweets were randomly selected and manually labeled using three possible classes: Positive, Neutral and Negative. Table 5.2 shows automatic sentiment classifications using the three classes, revealing that VADER tends to classify Neutral tweets as Positive or Negative, and that FinBERT tend to classify Positive tweets as Neutral. FinBERT achieves a much better precision for all classes, while VADER achieves a better recall for the positive and negative class.

Table 5.2 - Sample comparison of sentiment ratings between VADER, FinBERT and manually sorted true labels

		VADER Prediction			Total	FinBERT Prediction			Total
		Pos	Neu	Neg		Pos	Neu	Neg	
True Sample Label	Positive	42	8	14	**64**	8	46	10	**64**
	Neutral	23	7	18	**48**	1	35	12	**48**
	Negative	8	12	30	**50**	0	9	41	**50**
	Total Prediction	**73**	**27**	**62**	**162**	**9**	**90**	**63**	**162**

From Table 5.2, it was observed that out of the 162 tweets, 79 were correctly classified, which corresponds to an accuracy of 49%. When classifying the same tweets with FinBERT, 84 correct classifications were obtained, that is, an accuracy of 52%.

It was observed that VADER classified 8 tweets as positive instead of negative, while FinBERT with the same tweets classified 0 tweets as positive instead of negative. Some of these examples can be found in Table 5.3. When making an analogy from an investment point of view, where a person's opinion is taken into account for decision making, it is considered more serious to read a comment as positive, when it’s in fact negative. This is dependent on the user's interpretation because the user who could choose to withdraw or transfer the amount would not do so and would be losing money. The opposite is not so serious, as the user would understand that he should redeem the money invested when in fact he could still profit.

Nevertheless, as the purpose of this thesis is to take into consideration the opinion of influencers for the prediction of Bitcoin volatility, the most interesting parts are both the positive and negative comments, and consequently the neutral classification can be excluded since it would not exert in any kind of upward (positive) or downward (negative) movements.

Table 5.3 - Examples of tweets with classifications obtained manually (true label) and from two classifiers

Twitter (@)	Text	VADER	FinBERT	True Label
APompliano	Bitcoin’s hashrate has officially hit a new all-time high. The network has never been more secure.	Neg	Pos	Pos
cz_binance	#bitcoin mining hash rate has recovered and hit an ATH.	Neu	Pos	Pos
nayibbukele	El Salvador just bought the dip! sv 500 coins at an average USD price of ~$30,744 #Bitcoin	Pos	Neu	Pos
saylor	In #Bitcoin We Trust.	Pos	Neu	Pos
danheld	Not buying #Bitcoin is an expensive mistake.	Neg	Neg	Pos
saylor	Happy Birthday #Bitcoin.	Pos	Neu	Neu
saylor	“He who works all day has no time to make money.” - John D. Rockefeller on #Bitcoin	Neg	Neu	Neu
Natbrunell	Probably going lower but I bought more #Bitcoin today.	Neg	Pos	Neg
BitBoy_Crypto	I said last year this would be the worst #bitcoin bear market in history and they laughed at me. Those laughs have quickly turned to tears.	Pos	Neg	Neg
danheld	BREAKING: another country bans themselves from Bitcoin	Neu	Neg	Neg

In an attempt to improve the FinBERT results, it was necessary to perform an additional pre-processing step. Table 5.4 shows the corresponding results for the same 162 tweets, considering all the three possible classes. The results show an overall decrease of the performance. Such result was not a surprise, given that FinBERT is prepared to process meaningfull sentences.

Table 5.4 - Comparison of samples of sentiment classifications between FinBERT + Pre-processing and manually sorted true sample labels

		FinBERT + Pre-processing Prediction			**Total**
		Pos	Neu	Neg	
True Sample Label	Positive	19	39	6	**64**
	Neutral	2	43	3	**48**
	Negative	7	29	14	**50**
	Total Prediction	**28**	**111**	**23**	**162**

After this experiment, it was concluded that FinBERT without any pre-processing seems to be the most promising for the context of this study and FinBERT in general presents better results for the context of Finance, as seen in recent literature.

Following the classifications made by FinBERT, the tweets in which the sentiment was classified as Neutral were excluded, which resulted in a total of 1699 tweets.

In FinBERT, each analysed text is classified with three labels, where all scores are positive, and the prevailing sentiment is the one with the highest score. Therefore, for calculation purposes, all tweets that were classified as negative were multiplied by -1. From these tweets and their respective classifications, sentiment calculations were made at daily intervals to generate the time series.

For the calculation of sentiment by FinBERT at each daily interval, the following formula was established:

$$index = \sum_{t=0}^{totaltweets} sentiment\_score(t) \; x \; likes(t) \; x \; RTs(t) \tag{4}$$

Where, “index” is the sum of positive and negative sentiment values multiplied by likes and retweets (RTs), which were considered as weights.

Therefore, each day the indicator represents a new range with a new line or bar, resulting in a daily time series.

Finally, in order to have some indication of whether the words most used by influencers were positive or negative, the frequency of words (tokens) was created and Figure 5.1 was generated. However, it was not possible to draw any conclusions.

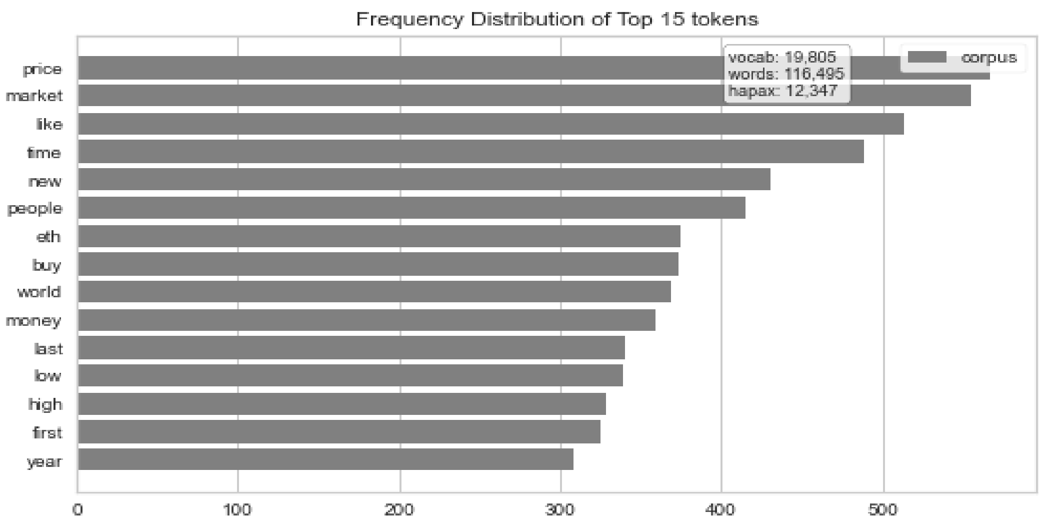

Figure 5.1 - Top 15 most frequent tweeted words in the dataset

The results obtained through the sample showed that neither the VADER classifier nor the FinBERT classifier with the addition of pre-processing were as good as the FinBERT without pre-processing. The accuracy obtained from the sample for FinBERT without pre-processing (52%) was slightly better than VADER (49%) and FinBERT with the addition of pre-processing (47%). Furthermore, when comparing VADER with FinBERT without pre-processing, it was seen that VADER misclassified more tweets than FinBERT. VADER ranked 22 tweets wrong (Table 5.2) out of 162 in total in the sample (8 as positive and 14 as negative), while FinBERT classified 10 wrong (with 10 as negative). When dealing with investments, the "value or weight" of classifying something as negative when it is actually positive is less bad than classifying something positive when it is negative, this is because the decision to buy or sell an investment, in this case Bitcoin, can be compromised if dependent on this misclassification.

These comparisons showed that the internal pre-processing already trained in FinBERT seems to be the most suitable for this dataset, and therefore, FinBERT without the addition of pre-processing is possibly able to obtain better classifications compared to the two previous ones tested.

It was also found to be better to consider the short period of tweets, since even though it is a smaller window, it contains fewer gaps between the data and more data proportionally to the number of days than the total period.

5.2 Granger causality and Transfer Entropy in prol of an answer

This section shows the causality between Bitcoin and Sentiment Score series that were graphically computed in the Figure 5.2, and, as illustrated, a long period between 2016 and 2021, with few tweets (positive and negative) from influencers, was observed.

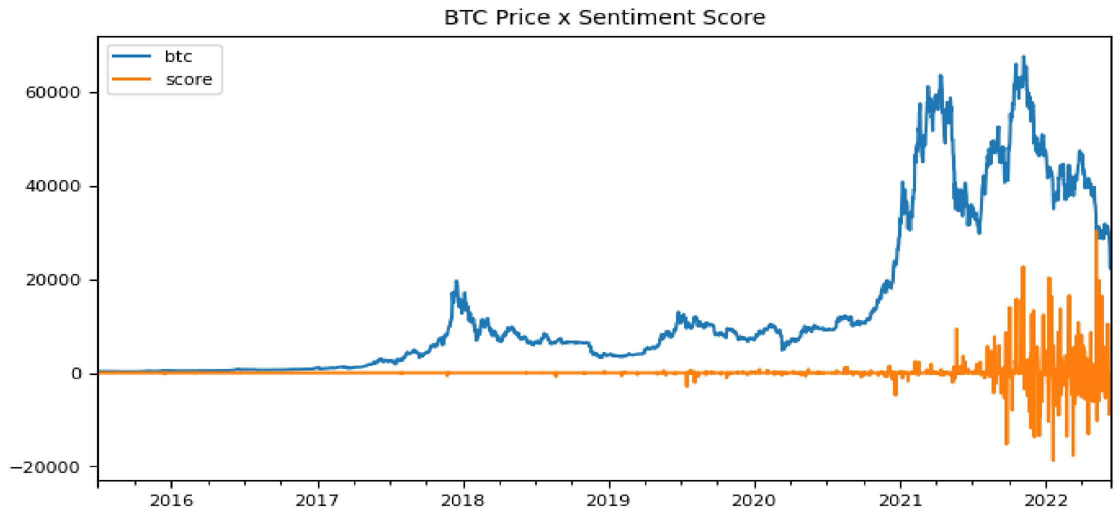

Figure 5.2 - Bitcoin and Sentiment Score series between 30th of June 2015 and 15th of June 2022

Although the first record of a "crypto influencer" talking about Bitcoin took place on the 30th of June 2015, after this there was not enough data for the entire period, that is, for many days and even weeks, no tweets were recorded. This way, only tweets data in the period beginning on 1st of January 2021 (Figure 5.3) was considered, that is, the period in which investors began to invest more in cryptocurrencies (as seen in the chart above) and talk about Bitcoin on Twitter more frequently.

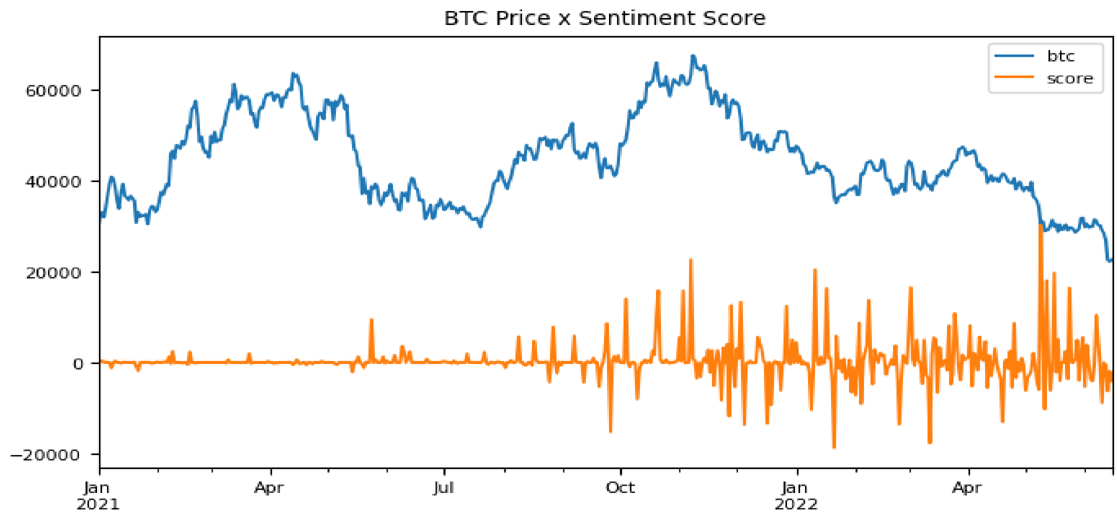

Figure 5.3 - Bitcoin and Sentiment Score series for the reduced period (between 1st of January 2021 and 15th of June 2022) of collected tweets

Table 5.5 shows that if the total time period were considered, the number of tweets per day would be one or none, regarding when an influencer spoke about Bitcoin (positively or negatively). Reducing this period to January 2021 - June 2022 would increase this amount to three tweets per day. When considering only the total number of days that had tweets in both periods, there would be an increase from one and three to seven and nine tweets per day. This way, the total number of tweets to be considered became 1,534.

Table 5.5 - Comparison of the number of tweets made by influencers in the two periods.

Period	Total days (with or without tweets)	Total tweets in Total days	Tweets/ day*	Total days with tweets	Tweets/ day*
Jun/30/2015 - Jun/15/2022	2,543	1,699	1	237	7
Jan/01/2021 - Jun/15/2022	531	1,534	3	180	9

Note: * Values were rounded up.

The reduction of tweets, that is, of data, until this step may seem precarious regarding the amount of information necessary to predict something. Nevertheless, according to the objective of this study, the quality of a tweet tends to exert more impact on volatility, since an influencer's tweet can impact an investor's decision. It does not matter if the influencer was writing 100 “neutral” tweets, but only ten had sentimental relevance. These ten tweets are likely to be reflected in volatility.

The choice of the reduced period corroborates with Keskin and Aste (2020), who say that the choice of time scale involves a trade-off: with a small timescale, there are not enough messages to estimate sentiment, but a long timescale represents a low-resolution sample which loses too much information about the underlying time series.

For the selected period, and in order to test whether one variable temporally precedes the other one, Granger causality was used.

As Granger causality (1969) can only be applied to stationary series, the ADF test (Augmented Dickey-Fuller) was applied to the six series (Bitcoin, Sentiment Score, Hash Rate, Fear & Greed Index, BTC Dominance and Mayer Multiple), to test the null hypothesis that the unit root is present in the time series. For this test, the Python statsmodels library was used and the autolag parameter (to determine the number of lags chosen to minimize the corresponding information criterion) was determined by using the Akaike Information Criterion (AIC). At a significance level of 5%, the results were: Bitcoin, Hash Rate and Fear & Greed Index are non-stationary series, while Sentiment Score, BTC Dominance and Mayer Multiple are stationary. Then the three non-stationary series were differentiated once, and the ADF test was applied again, and it was verified that they were stationary in the first difference or integrated of order one.

With all time series stationary (or stationarized), the Granger causality test was performed, and the adjacency matrix that was generated for lags (from 1 to 5) can be found in the Figure 5.4.

lag 1	score_x	returns_x	btc_dominance_x	mayer_multiple_x	diff_hash_rate_x	diff_fear_index_x
score_y	1	0.0814	0.4925	0.8141	1	0.6118
returns_y	0.9598	1	0.048	0.4832	0.2936	0.8652
btc_dominance_y	0.8833	**0.0002**	1	0.9047	1	0.3834
mayer_multiple_y	0.5362	0.7453	0.0086	1	1	0.9335
diff_hash_rate_y	1	1	1	1	1	1
diff_fear_index_y	0.921	**0**	0.8836	0.8812	1	1

lag 2	score_x	returns_x	btc_dominance_x	mayer_multiple_x	diff_hash_rate_x	diff_fear_index_x
score_y	1	**0.0439**	0.4395	0.1554	1	0.4771
returns_y	0.5868	1	0.048	0.3156	0.2936	0.8389
btc_dominance_y	0.4855	**0.0002**	1	0	1	0.3834
mayer_multiple_y	0.4379	0.7453	0.0086	1	1	0.9335
diff_hash_rate_y	1	1	1	1	1	1
diff_fear_index_y	0.921	**0**	0.0021	0	1	1

lag 3	score_x	returns_x	btc_dominance_x	mayer_multiple_x	diff_hash_rate_x	diff_fear_index_x
score_y	1	**0.0439**	0.4395	0.1554	1	0.4771
returns_y	0.5868	1	0.048	0.3156	0.2936	0.7517
btc_dominance_y	0.0671	**0.0002**	1	0	1	0.3834
mayer_multiple_y	0.4379	0.7453	0.0086	1	1	0.8564
diff_hash_rate_y	1	1	1	1	1	1
diff_fear_index_y	0.7312	**0**	0.0021	0	1	1

lag 4	score_x	returns_x	btc_dominance_x	mayer_multiple_x	diff_hash_rate_x	diff_fear_index_x
score_y	1	**0.0439**	0.4395	0.1554	1	0.4771
returns_y	0.4909	1	0.048	0.3156	0.2936	0.7517
btc_dominance_y	0.0671	**0.0002**	1	0	1	0.3834
mayer_multiple_y	0.4379	0.7453	0.0054	1	1	0.8564
diff_hash_rate_y	1	1	1	1	1	1
diff_fear_index_y	0.7312	**0**	0.0021	0	1	1

lag 5	score_x	returns_x	btc_dominance_x	mayer_multiple_x	diff_hash_rate_x	diff_fear_index_x
score_y	1	**0.0439**	0.4395	0.1554	1	0.4771
returns_y	0.4909	1	0.0253	0.3156	0.2936	0.7517
btc_dominance_y	0.0671	**0.0002**	1	0	1	0.3834
mayer_multiple_y	0.4379	0.337	0.0054	1	1	0.2412
diff_hash_rate_y	1	1	1	1	1	1
diff_fear_index_y	0.2243	**0**	0.0021	0	1	1

Figure 5.4 - Adjacency matrix for the Granger causality significance test for lags 1, 2, 3, 4 and 5

The results were: Returns does Granger-cause Sentiment Score, BTC Dominance and Fear & Greed Index to lag >= 2. Sentiment Score does not Granger-cause in any indicator, nor Returns. There are no indicators of “crypto world” does Granger-cause Sentiment Score. Returns does Granger-cause BTC Dominance and Fear & Greed Index to all lags. BTC Dominance does Granger-cause Returns, Mayer Multiple and Fear & Greed Index to lag >= 2. Mayer Multiple does Granger-cause Returns and Fear & Greed Index to lag >=2. Since the indicators use BTC information to compose themselves and this is not the focus of the study, the other "crypto world" series will not be the focus, but will be analysed just in order to identify further down the line if these indicators are impacted by the Sentiment Score or vice versa.

The result obtained by the Granger causality test was not what was expected for the Sentiment Score (due to the fact that the Sentiment Score does not temporarily precede Returns), so Transfer Entropy was resorted to, because its definition tends to capture better the exchange of data and therefore is more adequate than Granger for this dataset.

The next step was to calculate the Transfer Entropy (TE) to measure the directional information movement from Sentiment Score to Bitcoin Prices (Sentiment Score -> BTC) and from Bitcoin Prices to Sentiment Score (BTC -> Sentiment Score), as shown in Figure 5.5.

To calculate the TE, the parameters considered were: 1 lag, window_size = 1 month, window_stride = 1 day, histogram estimation and equiprobable_bins = 7.

TE from BTC to Sentiment Score dominates in almost the entire period, but in some periods (Mar/21 to Apr/21 and Mar/22 to Apr/22) of falling BTC price, the Sentiment Score dominates the transfer of information. Dominance during declines is not immediate, however, as soon as prices begin to drop, the news spreads and the Sentiment Score impacts on volatility. For high prices, the impact of the news is not as big. This effect for stock returns is known as the "leverage effect". The high significance for the entire period corroborates the result of the TE that, in fact, there is an intensity in the flow of information exchange throughout the period in both directions.

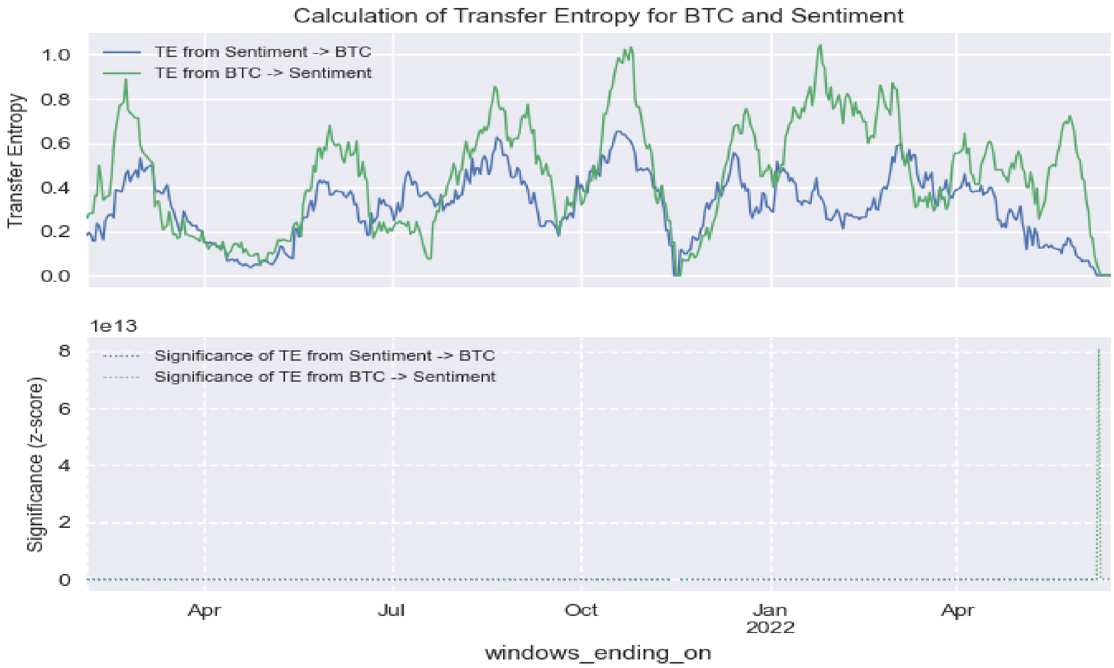

Figure 5.5 - Transfer Entropy between BTC and Sentiment for lag = 1

As this was done for Granger causality, it was also necessary to differentiate the Bitcoin series once to obtain the Returns (Figure 5.6).

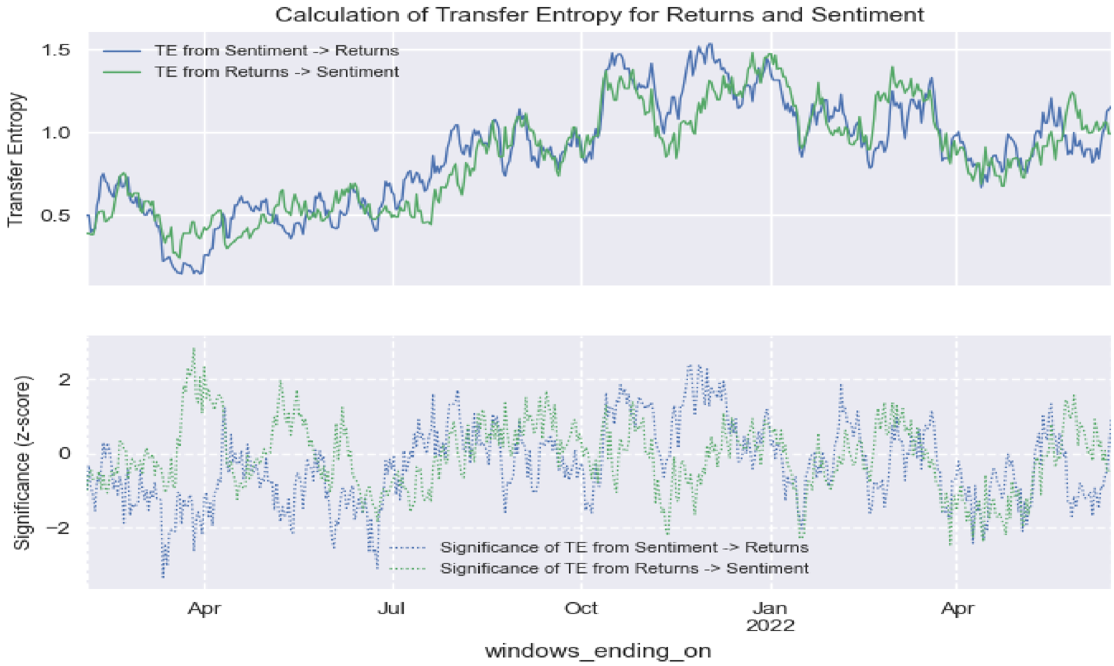

Figure 5.6 - Transfer Entropy between Returns and Sentiment series

For the calculation of the Transfer Entropy, the following parameters were considered: lag = 1, as it is understood that an investor will decide whether to invest or sell BTC based on what he reads from an influencer in the same minute or day. It is recalled that due to the limitation of tweets, that is, there is no volume of tweets per hour by influencers, the entire study was done in daily series. Size of the historical window for forecast (window_size): 1 month, because due to data limitation it was not possible to use a lower window. A window stride (window_stride) of one day since the objective is to know the forecast of the price fluctuation in up to one day. The histogram estimation and equiprobable bins (equiprobable_bins) with seven classes per dimension were used as suggested by Keskin (2018).

The result with the stationary BTC series (Returns) is higher than previous result (Figure 5.5). It is possible to observe that most of the information transfer flows for both directions have a TE greater than 0.5 and with peaks that reach a TE of 1.5, while most of the TE in Figure 5.5 is below of 0.6. Graphically, it is still not possible to identify which series have a greater flow of information.

Keskin (2018) recommends that the bins should be calculated using the AutoBins class, because if the parameter is “none”, or the bins dictionary is incompatible with the dimensions of the data, then the AutoBins sigma binning functionality will be called, passing the max bins parameter. For other applications where the high kurtosis of the distribution requires a thin partition, sigma binning is adopted (Keskin, 2018). Figure 5.7 contains the chart made for sigma bins to compare the results.

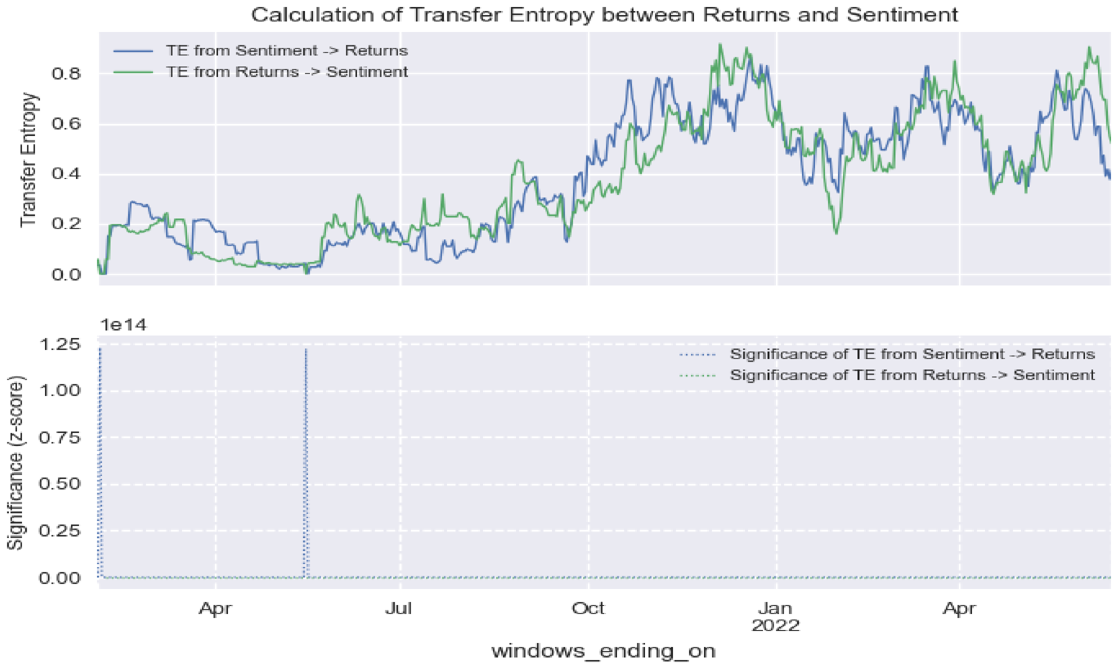

Figure 5.7 - Transfer Entropy between Returns and sentiment series considering sigma bins

To calculate the TE, the parameters considered were: 1 lag, window_size = 1 month, window_stride = 1 day, histogram estimation and sigma_bins = 7.

For the most period in Figure 5.7, the result obtained seems to show that the flow of information exchange is almost the same in both directions, with a few periods of dominance for Sentiment Score. TE values for sigma bins were lower than for equiprobable bins in Figure *5.7*, but with much higher significance. As in the study by Keskin (2018), this technique for this data was not helpful.

Following what Keskin (2018) did in one of his experiments to detect specific time-lags, the graph below was generated to help indicate which lag to consider. (Figure 5.8).

The most significant exchange of information occurred at lag 3 for the Sentiment -> Returns information transference.

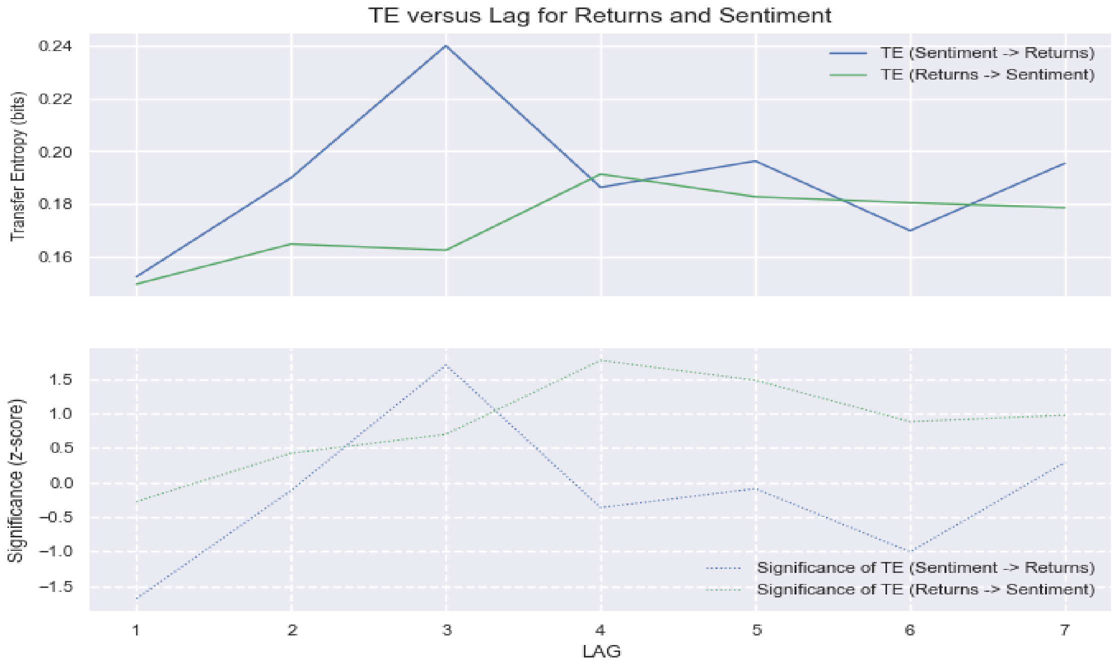

Figure 5.8 - Number of lags for Transfer Entropy between Returns and Sentiment

Therefore, the TE (Figure 5.9) was generated for lag 3, which showed the occurrence of information flow in both directions, and visually it seems to occur more TE for the Sentiment -> Returns flow, but it is still not possible to have sure about this.

In general, the behavior of both series for TE of lag 1 and 3 are very similar, but visually the TE of Sentiment -> Returns seems to have more intensity for lag 3, mainly in the 1st semester of 2021, December 2021 and June 2022, where the significance level mostly lies around -1 and 1.

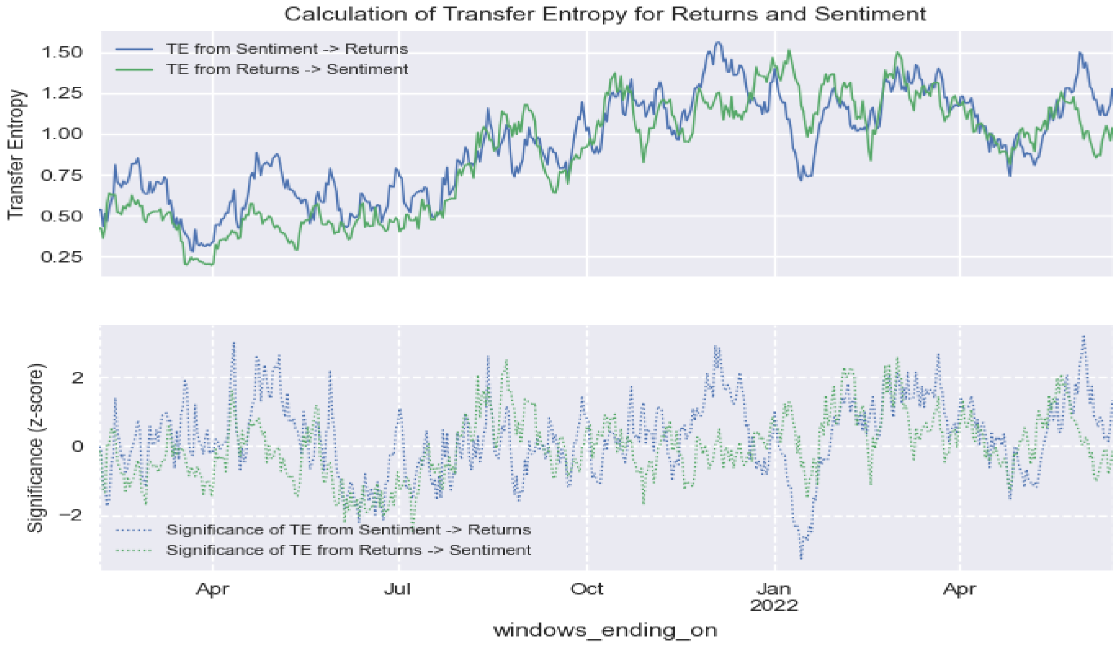

Figure 5.9 - Transfer Entropy between Returns and Sentiment series, considering 3 lags

Finally, to check the flow of information that must be considered and to quantify the intensity of the series, Net Transfer Entropy (see Equation 3) was used. In addition, Net TE was also calculated to compare the results observed in Figure 5.6 and Figure 5.8 (where lag 1 corresponds to the baseline, lag 2 corresponds to the result of the Granger causality test, lag 3 as the best result obtained and lag 4 to compare with the second highest peak). Thus, Table 5.6 and Figure 5.10 were generated. Below, the Net TE graphs between Sentiment -> Returns for four lags were plotted, in order to visually compare the information in the figure together with the table and analyse them.

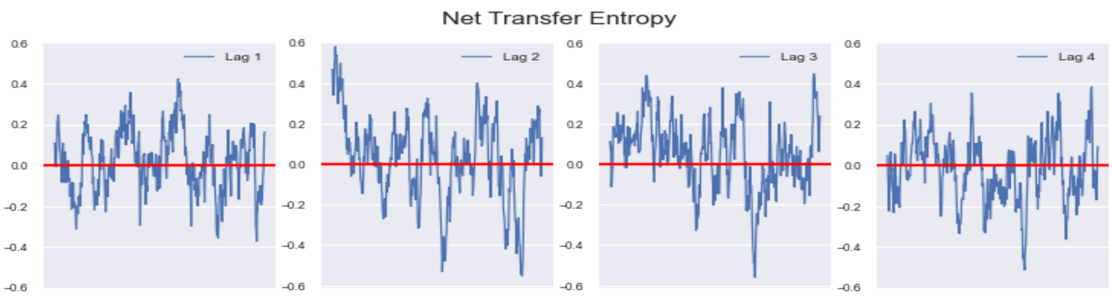

Figure 5.10 - Net Transfer Entropy graphs for lags 1, 2, 3 and 4 with the delimitation of positive and negative values

Table 5.6 - Net Transfer Entropy comparison for lags 1, 2, 3 and 4

Lags	Net TE	Net TE day count > 0	% Net TE day count > 0
1	9.23	266	53.63
2	12.89	299	60.28
3	28.15	324	65.32
4	-8.85	236	47.58

For Net TE it was considered $TE_{Sentiment \rightarrow Returns} - TE_{Returns \rightarrow Sentiment}$. Lags 1, 2 and 3 presented Net TE > 0, that is, the flow of information from Sentiment to Returns for any of the three lags is considered. For lag 3, it obtained the highest Net TE, recorded 324 positive daily points, corresponding to 65% of the total. The minimum values of $TE_{Sentiment \rightarrow Returns}$ for lags 1, 2 and 3 were 0.15, 0.21, and 0.28, respectively. The maximum values were also achieved for the mentioned lags, that is, 1.52, 1.53, and 1.56. Therefore, lag 3 conducts higher values for both the minimum and the maximum Net TE. Furthermore, it was observed that Sentiment transmits more information than it receives.

When verifying the relationship between the Sentiment Score and Bitcoin, the next step was to verify if the cryptocurrency indicators had any relationship with the Sentiment Score, since all indicators were based on Bitcoin information. To find out if there was a transfer of information between Sentiment and indicators, the same lag number was considered. Below in Table 5.7 is the Net Transfer Entropy, and the charts can be found in Figure 5.11.

Table 5.7 – Net Transfer Entropy between Indicators and Sentiment

	Hash Rate	BTC Dominance	Mayer Multiple	Fear & Greed Index
Net TE	43.8	30.1	24.2	-3.5

Thus, there was an exchange of Sentiment Score information for each of the three indicators (Table 5.7), with the highest intensity being Sentiment Score -> Hash Rate. This rate is directly linked, for example, to the number of transactions made and how many miners are active, so it was expected that the behavior would be similar to Returns, however with a higher TE intensity of Sentiment for Hash Rate (43.8 versus 28.15).

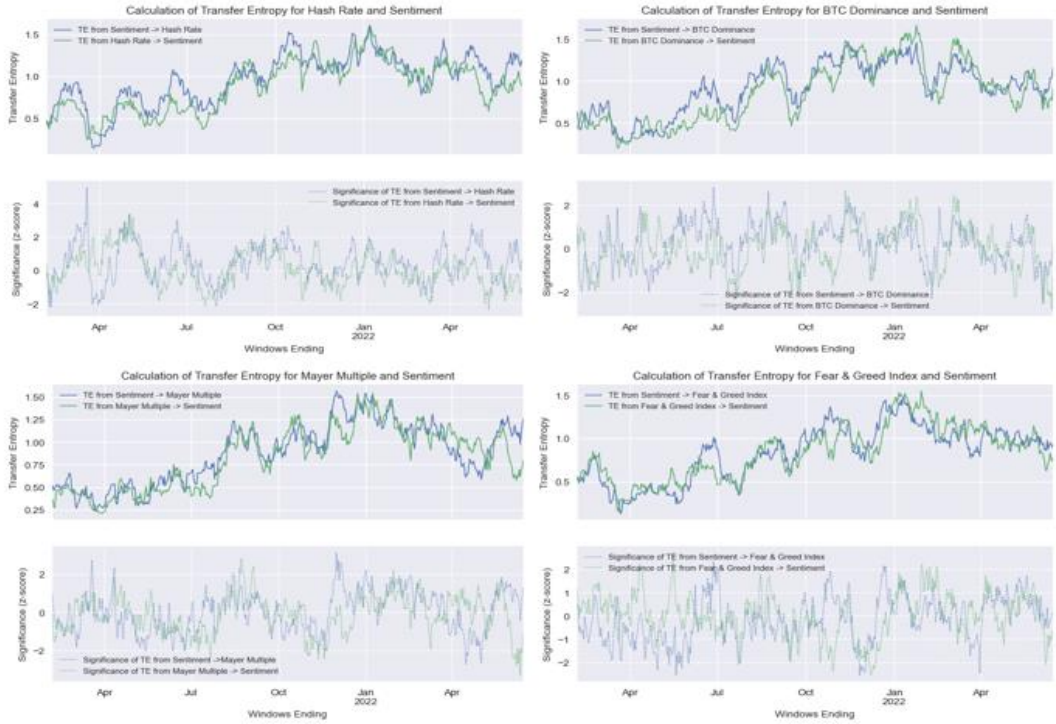

Figure 5.11 - Transfer Entropy between BTC indicators and Sentiment

The only indicator with the direction of information transferred in reverse was the Fear & Greed Index, which, when checking the number of possibly ideal lags, returned lag = 1 (Figure 5.12) and a new Net Transfer Entropy of 11.25. Therefore, regardless of the direction, Transfer Entropy was able to capture the non-linearity information flow that Granger linear causality could not.

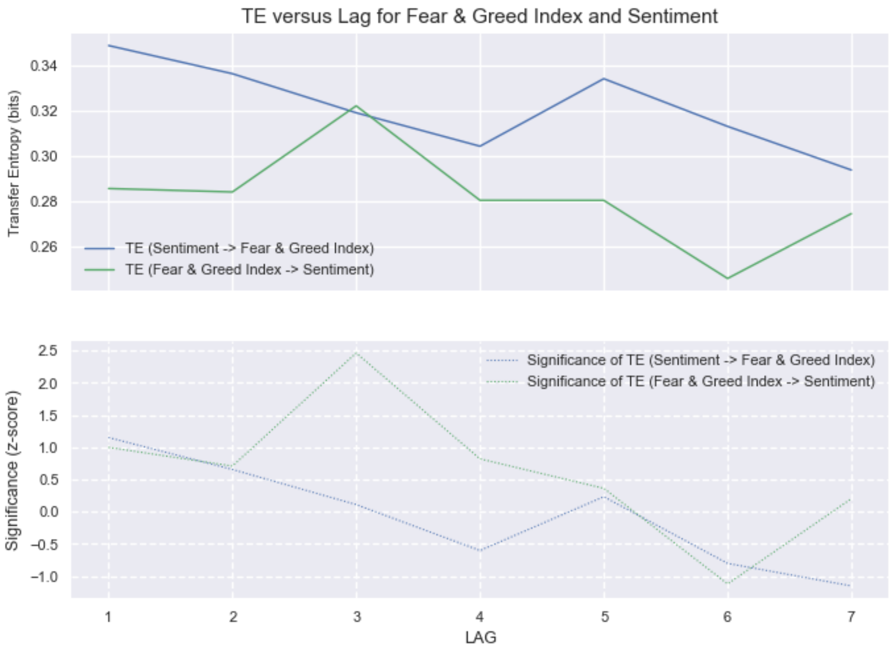

Figure 5.12 - Number of lags for Transfer Entropy between Fear & Greed Index and Sentiment Score

The results obtained from the classified data showed that the Granger causality test did not verify the usefulness of the Sentiment Score series in terms of forecasting BTC. The opposite was also tested, and it was found that Returns can be useful in the temporal precedence of the Sentiment Score for a lag greater than 1. The same test was applied to the sentiment score in relation to the indicators of the “crypto world”, but the result was not satisfactory neither for the Sentiment Score nor for the four indicators, that is, by the Granger causality test, none of the variables were able to temporally precede the other.

Other results were obtained about Transfer Entropy, which showed that there was an exchange of information between sentiment score and Returns for a lag greater than 2. The intensity of this flow was calculated by Net TE (Table 5.6), and it was obtained that from the sentiment score for Returns, this is where the strongest exchange of information took place.

When testing the flow of information between Sentiment Score and “crypto world” indicators for the same number of lags, the result obtained was that there was an exchange of Sentiment Score information for each of the three indicators (Table 5.7), however, only for the indicator (Fear & Greed Index) there was the most intense information exchange which occurred at lag 1 (Figure 5.12). Therefore, regardless of the direction, Transfer Entropy was able to capture the non-linearity information flow that Granger linear causality could not.

Through Transfer Entropy, it was possible to verify that the comments made by the influencers were able to influence the volatility of BTC in up to three days and that there was a causality between the sentiment score and the indicators of the “crypto world”.

6 Conclusion

Twitter has been considered a good source of data for some time due to its interactivity and speed of information, however, there are challenges in pre-processing this data and classifying sentiment. In summary, cleaning, transforming, and adjusting tweets to obtain a more accurate classification of sentiment requires knowledge of pre-processing techniques. For this study, two classifiers were tested: VADER, already well known for its ability to have good results in social network classifications in general, and FinBERT, which was created in 2020 with a focus on finance. Both techniques already have internal pre-processing aimed at each specificity, for example, VADER transforms emojis into words and FinBERT adjusts expressions from the financial market.

Having the correct classification of the data is important, since it is from these feelings that the up and down movement could be related to the Bitcoin series. Therefore, FinBERT showed slightly better rankings in a sample selection, with 52% accuracy versus VADER's 49% accuracy. However, VADER misclassified "the extremes" more often, that is, it classified positive as negative and vice versa more often, while FinBERT, when classified wrong, did it as neutral instead of positive/negative. This difference between ratings is very important when talking about investing, as getting reversed ratings can cause irreparable damage to an investor's business and finances. They may decide to keep the investment because it was based on wrong information, however, if the information was correct, they might have already taken another decision.

The Granger causality and Transfer Entropy information flow were analysed for the Sentiment Score series, Bitcoin daily closing prices and “crypto world” indicators. Based on Granger causality, it was identified that Returns does Granger-cause Sentiment Score for lag >= 2, but Sentiment Score does not Granger-cause Returns and no other indicator, as well as indicators from the “crypto world” does not Granger-cause Sentiment Score. By Granger causality, it is only possible to conclude that with the historical values of BTC it is possible to predict future sentiment values, however, that by adding what the influencer says does not help in these predictions.

The next step was to apply the Transfer Entropy methodology to find out if Sentiment Score has an impact on BTC.

When comparing the result of Keskin (2018) with this study, similar results were obtained, despite the difference between the selected period and the data set, since Keskin considered all public tweets for the sentiment series.

In this study, it was seen that non-linear causality occurs between both directions and with greater intensity from Sentiment Score to Bitcoin for a lag of three days. Therefore, it was concluded that it is possible to predict the volatility of Bitcoin based on the sentiment generated by the opinion of influencers.

As studied by Hamza (2020), the results of this thesis strengthen the reasoning of how the volatility of cryptocurrencies can be affected by public opinion, as influencers play a crucial role in social media, hence they have the ability to move the market with sufficient influence to guide the direction of public opinion.

In a future vision, but not so far, this study also agrees with Pang et al. (2020) when he says that social sentiment data will become the primary candidate for cryptocurrency trading, and also reinforces what was said by Jain et al. (2018), that user credibility, user popularity, and user network are other social factors that can be considered to measure the cryptocurrency price and increase the accuracy of price-prediction models.

Future work could consider specific sample periods to study whether the occurrences before and after specific events impacted the interaction of influencers. In addition, it could consider the occurrences of bots, as Kraaijeveld and De Smedt (2020) found that about 1 to 14% of tweets in the cryptocurrency datasets were posted by bots, which may take into account false information and somehow impair the analysis. For this study, as the influencer profiles were selected, none of them referred to bots. Nevertheless, there may be the incidence of bots in the total of likes and retweets, which are considered a weight in the sentiment calculation. This verification was not taken into account, because it is not simple.

References

Abraham, J., Higdon, D., Nelson, J., & Ibarra, J. (2018). Cryptocurrency price prediction using tweet volumes and sentiment analysis. SMU Data Science Review, 1(3), 1.

Aggarwal, C. C., & Zhai C. (2012). Mining text data. *Springer Science & Business Media*.

Alexander, P. A., & Jetton L. T. (1996). The role of importance and interest in the processing of text. Educational psychology review.

Almansour, M. (2021). Industrial Engineering & Management Systems 20 (2), 130-139.

Anjaria, M., & Guddeti, R. M. R. (2014). *Influence factor based opinion mining of Twitter data using supervised learning*. Sixth International Conference on Communication Systems and Networks (COMSNETS).

Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models.

Aragão, D. C. (2016). Bitcoin: Uma análise sobre a evolução das taxas de transação. Universidade Federal de Pernambuco.

Barbosa, L., & Feng, J. (2010). Robust sentiment detection on Twitter from biased and noisy data. Proceedings of the 23rd International Conference on Computational Linguistics, USA, 36-44.

Barbosa, P. C. L. (2016). Bitcoin e moedas fiat: um estudo de volatilidade comparada. Fundação Instituto de Pesquisas Econômicas–FIPE, São Paulo, 75.

Barnett L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for gaussian variables. *Physycal Review Letter*.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of computational Science, 2(1), 1-8.

Borge-Holthoefer, J., Perra, N., Gonçalves, B., González-Bailón, S., Arenas, A., Moreno, Y., & Vespignani, A. (2016). The dynamics of information-driven coordination phenomena: A transfer entropy analysis. *Science Advances*, *2*(4), 101-104.

Bukovina, J., & Marticek, M. (2016). Sentiment and Bitcoin volatility.

Carrick, J. (2016). Bitcoin as a complement to emerging market currencies. Emerging Markets Finance and Trade, 52(10), 2321-2334.

Catalini, C., Gans, J. S. (2020). Some simple economics of the blockchain. Communications of the ACM, 63(7), 80-90.

Ciaian, P., Rajcaniova, M., & Kancs, D. (2016). The economics of Bitcoin price formation. Applied Economics, 48(19), 1799-1815.

Folkinshteyn, D., Lennon, M. M., & Reilly, T. (2015). The Bitcoin mirage: An oasis of financial remittance. Journal of Strategic and International Studies, 10(2), 118–222.

Garcia, D., Tessone, C.J., Mavrodiev, P., & Perony, N. (2014). The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. *Journal of The Royal Society Interface*.

Gençağa, D. (2018). Transfer entropy. Entropy, 20(4), 288.

Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D., Sotiropoulos, D., & Giaglis, G. M. (2015). Using Time-series and sentiment analysis to detect the determinants of Bitcoin prices.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society, 424-438.

Hamza, S. (2020). The effect of tweets made by cryptocurrency opinion leaders on Bitcoin prices. *Saudi Journal of Economics and Finance*.

He, J., & Shang, P. (2017). Comparison of transfer entropy methods for financial time series. Physica A: Statistical Mechanics and its Applications, 772-785.

Hu, A. S., Parlour, C. A., & Rajan, U. (2019). Cryptocurrencies: Stylized facts on a new investible instrument. Financial Management, 48(4), 1049-1068.

Huang, A. H., Wang, H., & Yang, Y. (2022). FinBERT—A large language model approach to extracting information from financial text.

Hutto, C.J., & Gilbert, E. (2014). *VADER: A parsimonious rule-based model for sentiment analysis of social media text*. Proceedings of the 8th International Conference on Weblogs and Social Media.

Jain, A., Tripathi, S., Dhardwivedi, H., & Saxena, P. (2018). *Forecasting price of cryptocurrencies using tweets sentiment analysis*. 11th International Conference on Contemporary Computing.

Kaminski, J., & Gloor, P. A. (2014). Nowcasting the bitcoin market with twitter signals.

Kapar, B., & Olmo, J. (2020). Analysis of bitcoin prices using market and sentiment variables.

Karalevicius, V., Degrande, N., & Weerdt, J. D. (2018). Using sentiment analysis to predict interday bitcoin price movements. The Journal of Risk Finance, 19(1), 56–75.

Karppi, T., & Crawford, K. (2016). Social media, financial algorithms and the hack crash. Theory, Culture & Society, 33(1), 73–92.

Keskin, Z. (2018). PyCausality: A python package for effective measurements of information transfer in time series analysis.

Keskin, Z., & Aste, T. (2020). Information-theoretic measures for nonlinear causality detection: Application to social media sentiment and cryptocurrency prices: Non-linear causality detection. Royal Society Open Science, 7(9).

Kim, H.M., Bock, G.W., & Lee, G. (2021). Predicting Ethereum prices with machine learning based on Blockchain information. *Expert Systems with Applications*, 184.

Kim, Y.B., Kim, J.G., Kim, W., Im, J.H., Kim, T.H., Kang, S.J., & Kim, C.H. (2016). Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *Plos one*, 11 (8).

Kraaijeveld, O., & De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money, 65*.

Kristoufek, L. (2013). Bitcoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific Reports, 3*.

Lahmiri, S. & Bekiros, S. (2020). The impact of COVID-19 pandemic upon stability and sequential irregularity of equity and cryptocurrency markets. *Chaos, Solitons and Fractals, 138.*

Li X., Xie H., Chen J., & Deng X. (2014). News impact on stock price return via sentiment analysis*. Knowledge-Based Systems*, 69, 14-23.

Lindner, B., Auret, L., Bauer, M., & Groenewald, J.W.D. (2019). Comparative analysis of Granger causality and transfer entropy to present a decision flow for the application of oscillation diagnosis. *Journal of Process Control*, 79, 72-84.

Liu, B. (2012). Sentiment analysis and opinion mining. *Morgan & Claypool Publishers*.

Liu, Y., Huang, X., An, A., & Yu, X. (2007). *ARSA: A sentiment-aware model for predicting sales performance using blogs*. Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR'07, 607-614.

Liu, Z., D. Huang, K. Huang, Z. Li, & J. Zhao, J, (2021). *FinBERT: A pre-trained financial language representation model for financial text mining*. Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence.

Luther, William J. (2015). Cryptocurrencies, network effects, and switching costs. contemporary economic policy.

Marzbanrad, F., Kimura, Y., Palaniswami, M. & Khandoker, A.H. (2015). Quantifying the interactions between maternal and fetal heart rates by transfer entropy. *Plos one, Public Library of Science, 10*(12).

Milligen, B. P. V., Carreras, B., & Newman, D. (2016). Constructing criteria to diagnose the likelihood of extreme events in the case of the electric power grid. *Chaos: An Interdisciplinary Journal of Nonlinear Science, AIP Publishing, 26*(3).

Neves, R. H. D. (2018) Precificação do Bitcoin: Impacto de variáveis de atratividade. Fundação Getúlio Vargas.

Pang, Y., Sundararaj, G., & Ren, J. (2020). *Cryptocurrency price prediction using time series and social sentiment data*. Proceedings of the 6th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies, 35-42.

Parikh, R., & Movassate, M. (2009). Sentiment analysis of user-generated Twitter updates using various classification techniques, Stanford University.

Pierce, D.A. & Haugh, L.D. (1977). Causality in temporal systems. Characterization and a survey. *Journal of Econometrics*, *5*(3), 265-293.

Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, *85*(2), 461.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal 27*, 3, 379-423.

Silva, R. M. P. (2016). A evolução da moeda e a bitcoin: um estudo da validade da bitcoin como moeda. Revista da Graduação, *9*(2).

Souza, T. T., & Aste, T. (2016). A nonlinear impact: Evidences of causal effects of social media on market prices.

Steinert, L., & Herff, C. (2018). Predicting altcoin returns using social media. *Plos one*, *13*(12).

Syczewska, E. M. & Struzik, Z. R. (2015). Granger causality and transfer entropy for financial returns. Acta Physica Polonica A, Institute of Physics, Polish Academy of Science, 127(3A).

Tan, P. N. (2018). Introduction to data mining. Pearson Education India.

Tress, E. H. H. (2017). Uma revisão de literatura sobre o comportamento dos preços do Bitcoin: trata-se de uma bolha especulativa? (2017). Universidade Federal do Rio de Janeiro.

Vidotto, R. S., Migliato, A. L., & Zambon, A. C. (2009). O Moving Average Convergence-Divergence como ferramenta para a decisão de investimentos no mercado de ações. Revista de Administração Contemporânea, Associação Nacional dos Programas de Pós-graduação em Administração.

Viorica, I. (2013). Recent Evolution and Predictable Trends of the Banking Industry's Future Inside the European Space. *Procedia - Social and Behavioural Sciences*, *93*, 842–846.

Yang Y., & Uy M. C. S., Huang, A. (2020). FinBERT: A Pretrained Language Model for Financial Communications.

Yao, C.Z., & Li, H. Y. (2020). Effective transfer entropy approach to information flow among EPU, investor sentiment and stock market*. Frontiers in Physics, 8*.

Zhang Z., Dai H. N., Zhou J., Mondal S.K., García M.M., & Wang H. (2021). Forecasting cryptocurrency price using convolutional neural networks with weighted and attentive memory channels, 183.

Web References

[1] Alternative.me. (n.d). Fear and greed index. Retrieved June 15, 2022, from https://alternative.me/crypto/fear-and-greed-index/#api

[2] CoinMarketCap. (n.d). CoinMarketCap ID Map. Retrieved June 15, 2022, from https://coinmarketcap.com/api/documentation/v1/#operation/getV1CryptocurrencyMap

[3] Glassnode. (n.d.). Mining. Retrieved June 15, 2022, from https://docs.glassnode.com/basic-api/endpoints/mining#hash-rate

[4] Twitter. (n.d.). Developer Platform. Retrieved June 15, 2022, from https://developer.twitter.com/