COLEGIO UNIVERSITARIO DE ESTUDIOS FINANCIEROS

DOBLE GRADO EN DERECHO Y ADE

Trabajo Fin de GRADO - ADE

Sentiment Analysis of CBDCs and Cryptocurrencies

Institutional perception on the future of finance

Autor: Vargas Martino, Ramón

Tutor: Cuadros Solas, Pedro Jesús

Madrid, diciembre de 2022

TABLE OF CONTENS

1. Introduction

1.1. General context and definition of CBDCs and cryptocurrency

With the rise of a cashless economy, explained by increasing connectivity, and in part accelerated by the recent pandemic, Central Banks have taken the path of innovation to adapt their payment systems to the new standards. A CBDC can be defined as a digital representation of fiat money that can be stored and transferred via a network managed by a Central Bank.

On the other hand, the fundamental definition of cryptocurrency is: a digital stake within a decentralized network, to which the decentralization mechanisms (instead of a central authority) give value to. Some of these mechanisms are consensus algorithms, or nodes.

1.2. Bitcoin and DLT as positive catalyst for CBDC innovation

Central Bank Digital Currencies constitute the most recent effort from the main international financial institutions in coordination to create a more connected monetary system and make central banks more efficient in all aspects. To understand where all this innovation comes from, we should go back in time. In the rise of Cryptocurrencies, the underlying technology Blockchain, a type of DLT, became popular among business and professionals as it opened a world of possibilities for new solutions in many areas of business; mainly reducing or fully eliminating intermediaries of any kind within business processes. Overall, cryptocurrencies showed the practicality of using DLT to build payment systems that were transparent and had no intermediaries. Current national and international banking systems rely heavily on intermediary parties, making payments and any transaction settlement less efficient. Therefore, efforts to study the implementation of CBDC appeared shortly after the increase in popularity of DLT.

To see the rising importance of CBDC we can look at the Bank for International Settlements, which serves as a common platform for central banks. The BIS is the most significant contributor in this subject and its Innovation Hub has already tested 7 projects with central banks worldwide. It has studied and tested several prototypes for multi CBDC settlements as well as Distributed Ledger Technologies that could support these digital currencies. It is worth mentioning that still, some central banks are contrary to the use of DLT, but of all BIS projects the vast majority include DLT architecture.

In terms of practicality, working projects in the world are the e-Naira (Nigeria), Sand Dollar (Bahamas), and JAM-DEX (Jamaica), but many more are in pilot testing, namely: E-krona (Sweden), BoK (South Korea), Project Aber (Saudi Arabia and United Arab Emirate), Project Ubin (Singapore) and others.

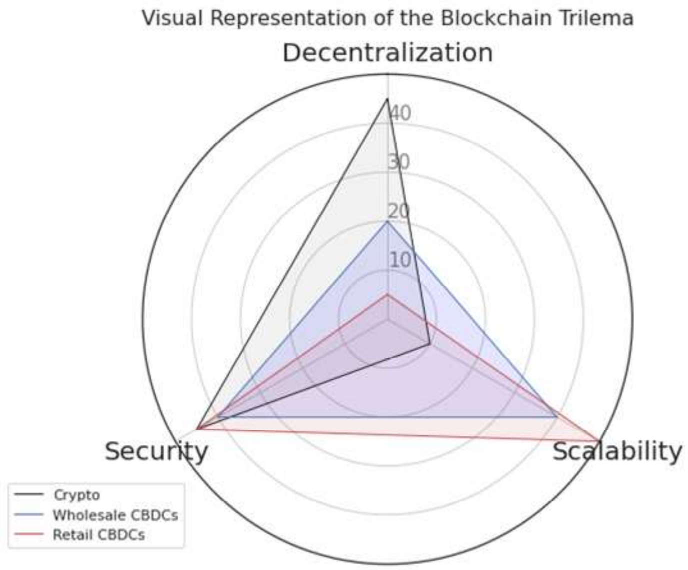
1.3. The idea behind it all.

It is no secret that CBDC innovation has been triggered by the idea of decentralizing intermonetary systems first introduced by Bitcoin. The first decentralized payment system was Bitcoin, and it was the successor of many years of research in computer science algorithms for the creation of decentralized networks. To understand where the concept of decentralization comes from, we must go back to 1982 when the Byzantine Generals Problem Research Paper was published (LESLIE LAMPORT, 1982). This research paper was funded by NASA, the BMDS Command, and the ARO and its core objective was to establish a way to handle conflicting information that malfunctioning components would give to different parts of a distributed system and thus making it reliable (I.e., missile systems). Whenever a distributed group of computers needs to achieve reliable communications, the network must be capable of solving the fundamental mathematical problem called the Byzantine Generals Problem. This problem can be solved through consensus algorithms which at the end are a set of rules. The main objective of these algorithms is to keep the system running while Byzantine faults or failures take place, i.e. one computer (node) stops working. Early studies on Byzantine Fault Tolerance date back to the 1950s in the aviation industry with solutions for flight computer systems. From all the research around Byzantine Faults in distributed computer systems, different algorithms and consensus protocols were proposed, from the early Micro-processor Flight Control System to the Paxos consensus protocol or the Practical Byzantine Fault Tolerance algorithm. In 1993 the idea of the Proof of Work (PoW) algorithm was first introduced by Moni Naor and Cynthia Dwork, and, some years later Bitcoin, an anonymous payment network appeared; with a Proof of Work consensus algorithm that used mining as the main mechanism for securing the network.

1.4. Centralized and Decentralized Payment Systems.

Now that we have a broad understanding of technical aspects of decentralization, for further understanding when we refer to centralized systems, we are speaking in terms of control over a payment system; for example, in a purely decentralized system, control is under a consensus algorithm while in a centralized system the banks assumes that role. The Blockchain Trilemma first described by Vytalik Buterin can help classify the level of decentralization and technical approach behind any project. This trilemma states that a given network (I.e., a CBDC payment system) is comprised of Decentralization, Security, and Scalability from which only two can be fully optimized to the detriment of a third one. For example, a central bank wanting to create a retail currency for wide national use, would maximize scalability and security instead of decentralization. We can visualize this in a radar chart that compares each project.

Chart 2: Blockchain Trilemma CBDCs vs Cryptocurrencies



2. Objective of the work

The aim of this work is to construct a framework of analysis that helps get a clear picture of how the latest innovations in the financial system are perceived and how far they are from being a reality. To do so, we will use Sentiment Analysis (Natural Language Processing technique) of text data from public and private banking institutions' statements on CBDCs, and cryptocurrency. With this analysis we will have a numerical measure of the state of each institutional party on the matter. This numerical approach will allow for a systematical analysis that can help navigate a vast amount of information about the issue at hand. After analyzing the main ideas and concerns from institutions we will build a model to see if sentiment data can predict the price trend of decentralized assets; in this case we will apply it to Bitcoin.

3. Exploring Sentiment on Cryptocurrencies and Central Bank Digital Currencies

By the end of this point, we will have a clear view of how central and commercial banks perceive CBDCs and cryptocurrencies. We will summarize the results of exploring the sentiment of central and commercial banks on CBDCs, and cryptocurrency. To achieve this, it is necessary to have the correct data at hand so that sentiment results are precise.

In the case of central banks, the data used to get the sentiment can be found in a database from the Bank for International Settlements. The database contains all speeches where central banks' representatives talk about CBDCs, cryptocurrencies and other related subjects.

On the other hand, data for commercial banks has been obtained through filtered Google search results. I have considered Google search as a primary source for commercial banks' posture as it contains news that generally reflect only the bank's opinion and rarely contains third party statements that could get mixed with public statements from the bank. For example, other well-known sources for sentiment analysis like twitter contain far less statements from commercial banks, being more useful for measuring the general public's opinion.

3.1. Sentiment Analysis of Central Banks' Speeches

In this stage we will go through an exploratory analysis of 284 speeches from 2016 until 2022 from many central banks on CBDCs and cryptocurrency to obtain a precise understanding of the posture of central banks. From this analysis we will be able to see which central banks are the most positive or negative towards both subjects.

3.1.1. Methodology

a. Obtaining the data.

First the data available at the BIS innovation Hub page is stored as a .CSV file with the following entries: "ISO2" (country code defined in the ISO 3166-1), "country_name", "Speech_title", "URL_text", "URL_link", "Date_m", and "speech_stance". After this, we read the .csv file programmatically with Python to obtain a Pandas Dataframe with the columns "Speech_title", "Date_m", "country_name" and "URL_text". This Dataframe will store all data in the form of a table as showed below.

Procedure 1: Obtaining CBDC and Cryptocurrency Central Bank Speeches data

```
df_url_list = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/NLP CBDC/BIS Speeches data analysis/Speeches_data.csv',delimiter=';')[['Speech_title','Date_m','country_name','URL_text']]

df_CBDC = df_url_list[~df_url_list['URL_text'].str.contains('pdf') & ~df_url_list['URL_text'].str.contains('central')]
```

⬇

	Speech_title	Date_m	country_name	URL_text
1	Yves Mersch: Distributed ledger technology - p...	01/04/2016	Euro Area	https://www.bis.org/review/r160426b.htm
2	Erkki Liikanen: Cash and the central bank	01/06/2016	Finland	https://www.bis.org/review/r160616e.htm
3	Mark Carney: Enabling the FinTech transformati...	01/06/2016	United Kingdom	https://www.bis.org/review/r160621e.htm
4	Carolyn Wilkins: Fintech and the financial eco...	01/06/2016	Canada	https://www.bis.org/review/r160622a.htm
6	Ravi Menon: Singapore's FinTech journey - wher...	01/11/2016	Singapore	https://www.bis.org/review/r161118a.htm

Additionally, a web scrapping method is created to automatically access each link from the table to then read and store the text from the speech web into the table. Now we have a table that contains what was said in each speech.

Procedure 2: Web Scrapping All .htm url to get the Transcript from every speech

Yves Mersch: Distributed ledger technology - panacea or flash in the pan?

Speech by Mr Yves Mersch, Member of the Executive Board of the European Central Bank, at the Deutsche Bank Transaction Bankers' Forum 2016, Frankfurt am Main, 25 April 2016.

Central bank speech | **26 April 2016**
by Yves Mersch
PDF version (99kb) | 4 pages

Is distributed ledger technology (DLT) a hype in need of demystifying? Is it a panacea that can heal the ailments of the financial industry, or it is a flash in the pan that will be forgotten in a few years? In my view, as it is often the case, the truth may lie somewhere in the middle.

```
for i in df_CBDC['URL_text']:
       response = rq.get(i)
       soup = bs(response.text,'html.parser')
       text = soup.find('body').find('div',attrs={'id':'container'})
       .find('div',attrs={'id':'cmsContent'}).text

       cleaned_text = (re.sub('\n',' ',text.strip()).split(' '))

       optimized_speech = ([i for i in cleaned_text if i not in stop
       _words])

       nlp_text.append(optimized_speech)
       all_text.append(cleaned_text)
       counter.append(coll.Counter(optimized_speech))
```

	text_str
	Is distributed ledger technology (DLT) a hype in need of demystifying? Is it a panacea that can heal the ailments of the financial industry, or it is a flash in the pan that will be forgotten in a few years? In my view, as it is often the case, the truth may lie somewhere in the middle. One of the main reasons why DLT has been capturing so much interest is because it could be a game

b. Data Cleaning.

This step is very important as the main goal is to get rid of meaningless words that in some sense add noise to the interpretation of polarity from the NLP module that we will use. To do so, we clean the words that do not contribute to a positive or negative read, for example: “hello” “ladies”,”and”, “gentlemen”, “Is”, “this”, etc. Additionally, regular expression statements are used to clean dates, numbers and other strings of text that follow the same pattern and do not add any positive or negative meaning.

Procedure 3: Cleaning Unmeaningful Words from Transcripts

text_str
Is distributed ledger technology (DLT) a hype ...

```
stop_words.extend(['Is','|','(',')','-
','.com','In','It','To','First','Second','Third','The','Ladies','ladies','gen-
tlemen' 'I', 'Dear', 'friends', 'Hello!','going','here','i.e.',r'[0-
9]{4}$','^(?=.{1,255}$)(?!-)[A-Za-z0-9\-]{1,63}(\.[A-Za-z0-9\-
]{1,63})*\.?(?<!-)$'])
optmized_speech = ([i for i in cleaned_text if i not in stop_words]
nlp_text.append(optimized_speech)
```

nlp_txt
distributed ledger technology (DLT) hype need ...

From the procedure above we achieve to pass to the NLP module a string that reads like this: “*distributed ledger technology (DLT) hype need demystifying? panacea heal ailments financial industry*” instead of: *“Is distributed ledger technology (DLT) a hype in need of demystifying? Is it a panacea that can heal the ailments of the financial industry, or it is a flash in the pan that will be forgotten in a few years?”*.

c. Analyzing the Speeches and Getting Sentiment and Most Frequent words Metrics.

At this point we will use NLP to create the sentiment metrics by analyzing the cleaned text in each row under the "speech_text" column. The Sentiment metric is obtained by using the Python TextBlob module which has a prebuilt Natural Language Processing tool called Sentiment. This tool analyses strings of text and returns a set of two values: subjectivity and polarity; but we are only interested in the polarity metric. The metric can have a value between –1 and +1, being +1 100% positive.

Procedure 4:

Applying Natural Processing Language Module to Speech Transcripts

Input

nlp_txt
distributed ledger technology (DLT) hype need ...

```
for i in df_CBDC['nlp_txt']:
    sent = tuple(TextBlob(i).sentiment)
    optimized_sentiment.append(sent)
```

Output

Polarity	Subjectivity
0.080962	0.380013

After this, some brief lines of code are implemented to find what are the most frequent words mentioned in each speech, and then they are stored in another column "freq_wrds".

Procedure 5:

Getting the Most Frequent Words in Each Speech

```
df_CBDC['cnt_txt'] = df_CBDC['cnt_txt'].ap-
ply(lambda x: {k: v for k, v in sorted(x.items(), key=lambda item
: item[-1],reverse = True) if v>2})
df_CBDC['freq_wrds'] = df_CBDC['cnt_txt'].ap-
```

cnt_txt	freq_wrds
{'market': 25, 'central': 25, 'DLT': 21, 'bank...	market central DLT bank financial would could ...

d. Filtering Speeches

At this stage we will filter the complete set of speeches with two main criteria to get the most precise speeches so that the polarity is more representative. The first filter is to get rid of those speeches where the subject is not crypto or CBDC. And the second one is to filter those that do not have a length of 10000 characters.

Procedure 6:

Filtering Speeches about CBDC and crypto

```
df_crypto_from_CB = df_CBDC[df_CBDC.text_str.str.contains('|'.join(k
eep_words))]
```

The best-known application of DLT is Bitcoin, which is a decentralized digital currency. The l
currencies are "permissionless," which means that anyone with a computer can download t
ledger and start authorizing transactions.

The decentralized aspect of the technology is why some predict that widespread applicatio
revolutionize entire industries. They contemplate alternative futures, such as one in which t
disintermediation of banks and even central banks, with state currencies being replaced by
currencies.

I see this as highly unlikely. People have not widely embraced digital currencies like Bitcoin

---

```
df_strcly_CBDC = df_CBDC[~df_CBDC.text_str.str.contains('|'.join(kee
p_words))].copy()
```

central bank digital currency on the ledger. But why stop here?

What if this concept is broadened? What if PSP or even individuals could access the central bank digital currency on the ledger via their bank accounts? This scenario would preserve the current relationship model of banks with central banks and to some extent also the existing organisational structure of the financial ecosystem.

3.1.2. Results

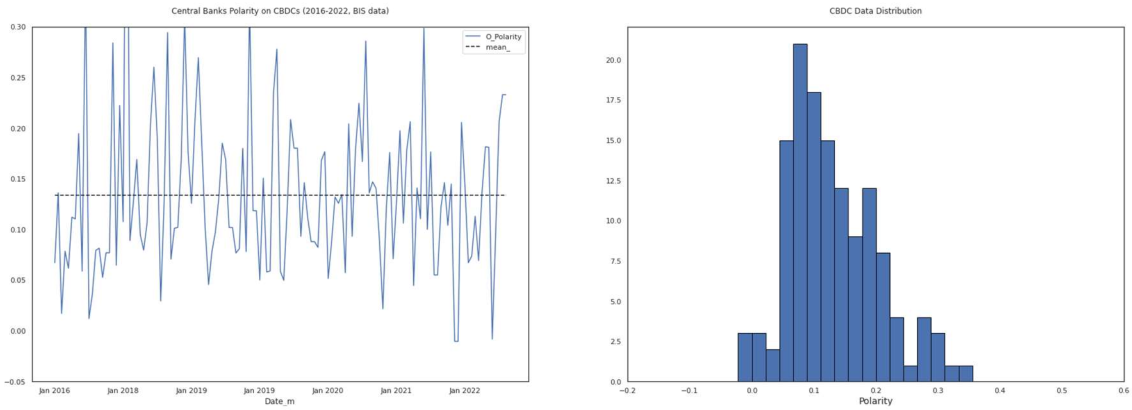
a. Polarity Over Time

- On CBDCs.

  Even though there is a small number of observations, the results indicate that central banks are generally more positive in speeches that contain the word "CBDC" with a mean polarity of +0.13 with a standard deviation of 0.08.

Chart 2:

Polarity of Speeches from Central Banks on CBDC-related themes

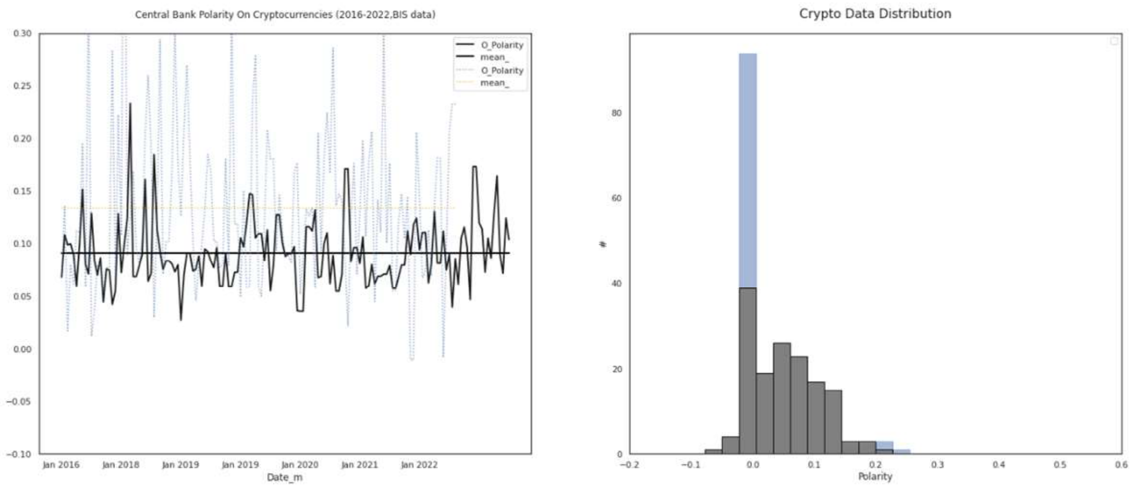


- On Cryptocurrencies.

  On the contrary, speeches containing the string "crypto" have a mean polarity of +0.09. with a standard deviation of 0.03.

Chart 3:

Polarity of Speeches from Central Banks Mentioning "crypto"

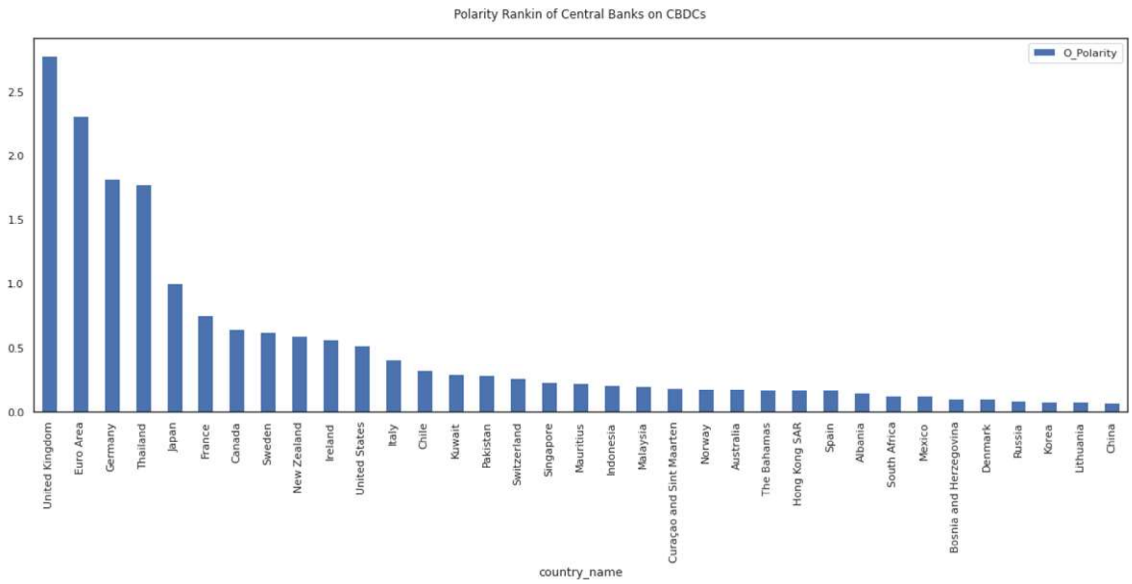


b. Most Positive and Negative Banks Rankin

The most positive CBs seem to be the Bank of England, followed by The European Central Bank and the Deutsche Bundesbank.

Chart 4:

Rankin of Most Positive Central Banks on CBDCs



3.1.3. What moves Central banks being favorable or not towards DLT and Crypto assets?

From the sentiment column we have sorted the most negative and positive speeches to see which the most repeated words that are contained in the "freq_wrds" column. The aim here is to find the most negative and positive speeches to see what the most mentioned ideas and concerns about CBDCs are. Additionally, we will consider the country of the central banks speaking and compare the degree of development from their CBDC projects to classify the most innovative central banks.

a. 3 Most Positive Speeches

Chart 5: Most Frequent Words from The Most Positive Speeches

As a side note, from the metric given by the BIS, the speech_stance in the three most positive Speeches is 1,0, and 1 which is in line with our results.

- Eddie Yue: Join us and be part of the change, Hong Kong SAR in Jan 2019.

> Most frequent words:
>
> {'Hong': 25, 'Fintech': 25, 'Kong': 21, 'We': 18, 'technology': 17, 'new': 16, 'This': 13, 'financial': 13, '': 12, 'banks': 12…}
>
> BIS speech_stance: +1

Yue is the Chief Executive of the Hong Kong Monetary Authority; during the Hong Kong Fintech Week in 2019, he reviews the latest innovations in the fintech space. In this conference he mentions how the use of blockchain technologies for cross-border payments or for connecting digital trade finance platforms looks promising.

In his most recent speech (16th of December 2022) Yue talks about 3 main points that better reflect his view on CBDCs and Crypto:

- Risks Posed by Cryptos and Defi. In this point Yue does not position himself in a completely negative way, instead he starts by stating that innovation for DeFi will grow stronger and that the benefits of it should not be overlooked. Nevertheless, some risks that Yue considers when speaking about DeFi are:
    - Market integrity
    - Financial stability threatened by code and cyber risks
    - Spillovers to the traditional financial system from the interconnectedness with the conventional financial system.
- How to Manage Risks with Regulation. Here Yue states four areas for regulation:
    - Assets: specially stablecoins should be regulated under the principal of "same risks, same regulations".
    - Trading Platforms: the place where the general public interact with Defi and crypto. He has also announced

that Hong Kong will soon have a new law regulating three vectors of risk: Money-laundering, financing terrorism and market integrity.
        - On-ramping and off-ramping activities: regulate banks exposure to direct exposure from providing trading and custody services.
        - Investor protection: through KYC assessments, or advertising restrictions to retail customers.
    - Role of Central Banks Digital Currency (CBDC)
        - CBDCs would make DeFi more trusted and widely used. Yuan states that: “DeFi already has two channels to connect back to the financial system, through stablecoins and banks. For these on-ramp/off-ramp activities, CBDCs should in theory make it easier and also provide more confidence for users”
- Ignazio Visco: An overview of the work of the G20 under the Italian presidency, Italy in Jan 2021.

<u>Most frequent words</u>: {'Financial': 6, 'digital': 6, 'work': 5, 'Italy': 5, 'best': 5, 'important': 5, 'On': 5, 'main': 4, 'Finance': 4,} <u>BIS speech_stance</u>: 0

- In the part relative to CBDCs Visco reviews how during the presidency of Italy in the G20 they have paid special attention to stable coins and CBDCs and their regulations.
  During this period the Financial Stability Board issued 10 reccomendations for the regulation, supervision, and oversight of global stable coin arrangements (FSB, 2022) . Some recommendations are: “Authorities should have and utilize the necessary or appropriate powers and tools, and adequate resources, to comprehensively regulate, supervise, and oversee a GSC arrangement”, “Comprehensive oversight of GSC activities and functions”, “Cross-border cooperation,

coordination and information sharing", "Governance structures and decentralized operations", etc.

- Philip Lowe: Opening remarks at the Melbourne Business Analytics Conference, Australia in March 2021.

> <u>Most frequent words</u>:
>
> {'data': 18, 'digital': 15, 'also': 14, 'economy': 9, 'investment': 8, 'make': 8, 'recovery': 7, 'support': 7, 'investments': 7…}
>
> <u>BIS speech_stance</u>: +1

During this speech, Lowe mentions that DLT was being studied to implement a Wholesale central banks digital currency and gives as an example the use of DLT to support the settlement of transactions in the interbank payment system.

b. 3 Most Negative Speeches

The BIS speech_stance metric states that the speeches we obtained as most negative speeches are all at 0, meaning our results are also in line.

Chart 6: Most Frequent Words from The Most Negative Speeches*

**Negativity close to Neutral or 0*

- Yves Mersch: Digital Base Money - an assessment from the European Central Bank's perspective, Euro Area in Jan 2017.

> <u>Most frequent words</u>:
>
> {DBM': 60, 'bank': 49, 'central': 39, 'would': 36, 'non-banks': 18, 'rate': 18, 'commercial': 17, 'deposits': 15, 'This': 14, 'may': 14, 'cash': 13, 'banks': 10, 'negative': 10}
>
> <u>BIS speech_stance</u>: 0

- In this conference, Mersch focuses mainly on Digital Base Money, a general term for CBDC. Mersch saw two main reasons why DBM was being considered:
  - First, the increasing popularity of digital payments
  - And Second, Technological Developments, mainly DLT, that would make it easier and less expensive to introduce a CBDC. Additionally, Mersch remarked that DLT could be a solution but that as it was an emerging technology it would have to wait.

  One of the main concerns for Mersch was the impact on interest rates from allowing households to hold deposits in central banks directly. As conclusion he gives four principles to assess the materialization of DBM: 1. technological safety, 2 efficiencies, 3. technological neutrality, and 4. freedom of choice for users.
- François Villeroy de Galhau: The role of banking in a sustainable global economy, France in Jan 2019.

Most frequent words: {'financial': 15, 'European': 10, 'central': 9, 'green': 9, 'climate': 9, 'digital': 6, 'Let': 6, 'We': 6, 'banks': 6, 'This': 6, 'de': 6, 'two': 5, 'common': 5, 'new': 5, 'policy': 5} BIS speech_stance: 0

In his speech Villeroy starts by highlighting how Bigtechs have the potential to redefine financial intermediation and how this is a big challenge for regulators. He also talks about stable coin projects which he defines as completely different to “speculative assets like Bitcoins” and that as the G7 said, they pose a sovereignty issue. Above all Villeroy insists in the need for antimony-laundering regulation for stable-coin projects due to the anonymity they provide.
After speaking of decentralized assets, he seems concerned about the lack of competitiveness in Europe to provide innovative payment

solutions and calls for the creation of a pan-european payment solution.

When speaking about a CBDC Villeroy thinks that central banks should first study all consequences from implementing it.

- Encik Abdul Rasheed Ghaffour: Optimal balance of paper and digital, cash and cashless; and next page for physical currency, Malaysia in Jan 2017.

> Most frequent words:
>
> {'cash': 20, 'This': 14, 'digital': 10, 'central': 9, 'recent': 9, 'banks': 9, 'mobile': 8, 'around': 8, 'We': 7, 'However,': 7, 'payments': 7, 'physical': 6, 'transactions': 6, 'cashless': 6, 'currency': 6, 'cost': 6}
>
> BIS speech_stance : 0

Abdul states that central banks will never stop using paper money but rather have a mix of digital and paper. As the main subject of the speech is this equilibrium between paper and a digital form of money , he states the 3S regulatory principals to modulate the balance between paper and digital money:

- Security,
- Social Cost
- Stability

3.2. Sentiment Analysis of Commercial Bank News

In this section we will study the sentiment over time of Commercial banks towards cryptocurrencies and CBDCs and get an approximation of what extent returns from cryptocurrencies change their posture.

3.2.1. Methodology

1. Obtaining daily news from commercial banks.

The first step consists of taking the 50 biggest banks worldwide (Wikipedia, 2022) and obtaining the sentiment of all news from Google since 2016 for the search queries: "[bank] on cryptocurrency" and "[bank] on CBDCs". To achieve this, we also use code to automate the searches.

Procedure 7:

Getting Daily Commercial Banks News on CBDC and Crypto Since 2016

```
for i in [2016,2017,2018,2019,2020,2021,2022]:
    url = "https://www.google.com/search"
    params = {"q": '', 'tbs':f'cdr:1,cd_min:1/1/{i},cd_max:7/12/{i+1}',
        "hl":"en"}
    headers = {"User-Agent": "Mozilla/5.0 (X11; Ubuntu; Linux x86_64,
      rv:89.0) Gecko/20100101 Firefox/89.0"}
    values = []
    links = []
    print(i)
    for b in Banks:
        params['q'] = f'{b} on crypto'
        soup = BeautifulSoup(
        requests.get(url, params=params,
        headers=headers).content,'html.parser')
        year_news_values = {}

        for a in soup.find('div',{'data
         hveid':'CBsQAA'}).find_all_next('div',{'class':'SoaBEf'}):
            link = ''.join(a.text)
            values.append(link)
        year_news_values[i] = '/'.join(values)
    text_titles[i] = values
with open('C:/Users/Ramon/Documents/CODING/Projects/NLP/Sentiment
analysis/Cryptocurrency/bydate_Crypto_CBanksHeadlines.txt','w') as f:
     f.write(json.dumps(text_titles))
```

2. Creating Sentiment Measures

a. Yearly Sentiment.

Procedure 8:

Getting Yearly Sentiment for Commercial Banks News

```
with open('/content/drive/MyDrive/Colab Notebooks/NLP CBDC/Google Searches ana
lysis/bydate_CBDC_ComBanksHeadlines.txt') as f:
    bydate_cbdc_headlines = json.load(f)
    f.close()

with open('/content/drive/MyDrive/Colab Note-
books/NLP CBDC/Google Searches analysis/bydate_Crypto_ComBanksHead-
lines.txt') as f:
    bydate_crypto_headlines = json.load(f)
    f.close()
```

```
Text - headlines: {"2016": [in Headlines: The Year's 13 Biggest Blockchain
StoriesNot Just Bitcoin: The Top 7 Cryptocurrencies All Gained in
2016Here's Why Bitcoin Boomed in 2016 - Business InsiderThought Bitcoin Was
Dead? 2016 Is the Year It Goes…]
```

```
df_ComBanks_sentiment_bydate = pd.DataFrame(bydate_crypto_headlines.items(),co
lumns=['Date','crypto_headlines'])
df_ComBanks_sentiment_bydate['cbdc_headlines'] = bydate_cbdc_headlines.values(
)
df_ComBanks_sentiment_bydate['cbdc_headlines'] = df_ComBanks_sentiment_bydate[
'cbdc_headlines'].apply(lambda x: ''.join(x))
df_ComBanks_sentiment_bydate['crypto_headlines'] = df_ComBanks_sentiment_bydat
e['crypto_headlines'].apply(lambda x: ''.join(x))
```

```
#applying sentiment analysis
df_ComBanks_sentiment_bydate['cbdc_Sentiment'] = df_ComBanks_sentiment_bydate[
'cbdc_headlines'].apply(lambda x: TextBlob(''.join([x for i in bydate_cbdc_hea
dlines if x not in stop_words])).sentiment)
df_ComBanks_sentiment_bydate['crypto_Sentiment'] = df_ComBanks_sentiment_bydat
e['crypto_headlines'].apply(lambda x: TextBlob(''.join([x for i in bydate_cryp
to_headlines if x not in stop_words])).sentiment)
df_ComBanks_sentiment_bydate[['crypto_Polarity','crypto_Subjectivity']] = df_C
omBanks_sentiment_bydate['crypto_Sentiment'].tolist()
df_ComBanks_sentiment_bydate[['cbdc_Polarity','cbdc_Subjectivity']] = df_ComBa
nks_sentiment_bydate['cbdc_Sentiment'].tolist()
```

Date	crypto_headlines	crypto_Polarity	cbdc_Polarity
2016	2016 in Headlines: The Year's 13 Biggest Block...	0.101351	0.000000

b. Sentiment by Bank

Procedure 8:

Getting Yearly Sentiment for Commercial Banks News

```
Text - headlines:
"Bank of America": "Bank of America Says Cryptocurrencies Continue to Act
as .../Bank of America's Crypto Users Shrink by 50% in Bear Market/Latest
News on Bank of America - Cointelegraph/Bank of America Has No Plans to
Offer Crypto Services Says .../Crypto May Be the Future of Investing, Says
Bank of America .../Bank of America Has Lost Half of Its Active Crypto
Users/Bank of America \"disagrees\" that crypto has no intrinsic value/Bank
of America is Bullish on Digital Assets; Believes Bitcoin .../Bank of
America CEO to focus on digital investments, crypto at .../BofA Global
Research Launches Coverage of Digital Assets"
```

```
with open('/content/drive/MyDrive/Colab Notebooks/NLP CBDC/Google Searches
analysis/Crypto_CommercialBanksHeadlines.txt','r') as f:
    crypto_headlines = json.load(f)
    f.close()

with open('/content/drive/MyDrive/Colab Notebooks/NLP CBDC/Google Searches
analysis/CBDC_CommercialBanksHeadlines.txt') as f:
    cbdc_headlines = json.load(f)
    f.close()

#Creating Dataframe
df_google = pd.DataFrame(crypto_headlines.items(),columns=['Banks','crypto_
Headlines'])
df_google['crypto_Sentiment'] = df_google['crypto_Headlines'].apply(lambda
x: TextBlob(''.join([x for i in crypto_headlines if x not in stop_words])).
sentiment) #apply texblob to cleaned str
df_google['cbdc_Headlines'] = cbdc_headlines.values()
df_google['cbdc_Sentiment'] = df_google['cbdc_Headlines'].apply(lambda x: T
extBlob(''.join([x for i in crypto_headlines if x not in stop_words])).sent
iment)

#Abreviating names
df_google.loc[[1],['Banks']] = 'ICBC'
df_google.loc[[25],['Banks']] = 'ANZ BG'

#Expading sentiment measures into different columns
df_google[['crypto_Polarity','crypto_Subjectivity']] = df_google['crypto_Se
ntiment'].tolist()
df_google[['cbdc_Polarity','cbdc_Subjectivity']] = df_google['cbdc_Sentimen
t'].tolist()
```

Banks	crypto_Polarity	cbdc_Polarity
JPMorgan Chase	-0.042857	0.000000

3.2.2. Results

a. Top 50 Worldwide Commercial Banks Polarity Positions

- In the case of CBDCs the most negative bank is the Bank of China and the most positive one is HSBC

Chart 7: News Used To Calculate Sentiment For HSBC on CBDCs

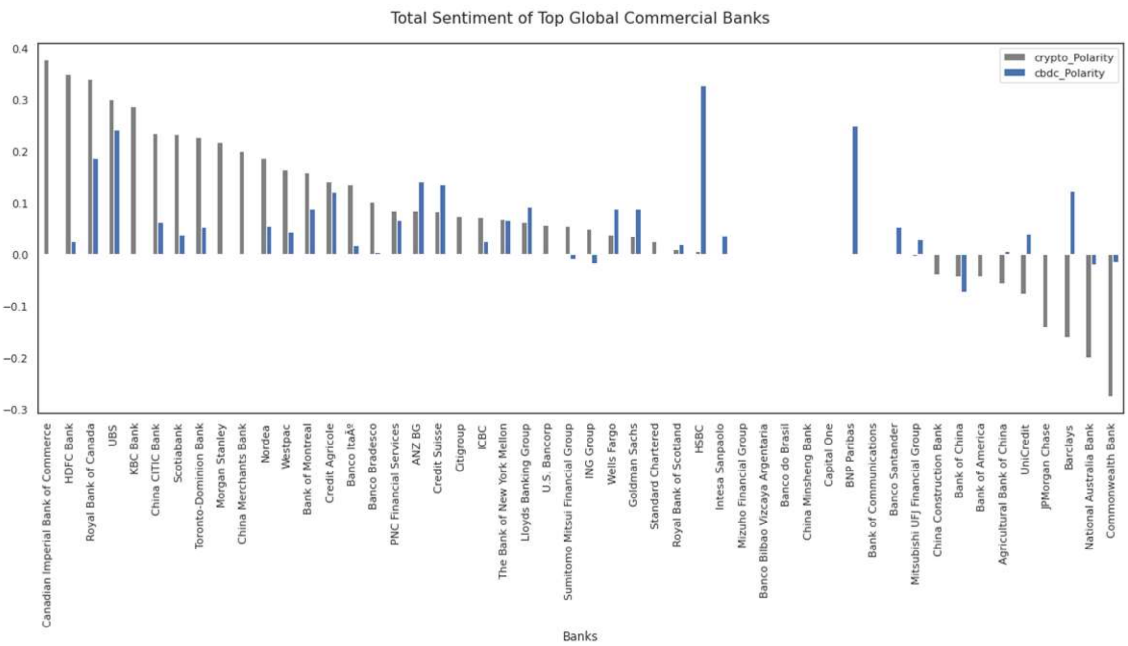
"HSBC": "Central Bank Digital Currencies | Insights | HSBC/Central bank digital currencies | Insights | HSBC/HSBC And IBM Successfully Design And Test Interoperable .../HSBC, Visa, G+D winners of Hong Kong's Global Fast Track .../HSBC supportive of CBDC, working with 8 nations/HSBC Says it Supports CBDC with Regulations - PYMNTS.com/HSBC and IBM create successful multi-ledger CBDC demo/Sibos 2022: No bank is an island \u2013 joining up digital currencies/Digital currencies and fintech: projects - Bank of Canada/HSBC and IBM successfully test multi-ledger CBDC, bond and ..."

- For cryptocurrency, the most negative bank is the Commonwealth Bank and the most positive is the Canadian Imperial Bank of Commerce.

Chart 8: News Used To Calculate Sentiment For CIB on Crypto

"Canadian Imperial Bank of Commerce": "Tokenized Canadian Imperial Bank of Commerce - CMca shares .../Invest in Canadian Imperial Bank of Commerce (cm) - Bitpanda/Canadian Imperial Bank of Commerce Archives \u2013 Bitcoin News/How To Buy Crypto With CIBC Bank In 2022 - 4 Easy Steps/CIBC CEO: We are not afraid of Bitcoin - NewsBTC/Cryptocurrencies may become 'legitimate stores of value': Dodig/Canadian Imperial Bank of Commerce Class A (CM_pr)/Buy Bitcoin cheap using CAD via Canadian Imperial Bank of .../Sell Bitcoin to get CAD via Canadian Imperial Bank of .../Five Canadian banks turn to blockchain for identity verification"

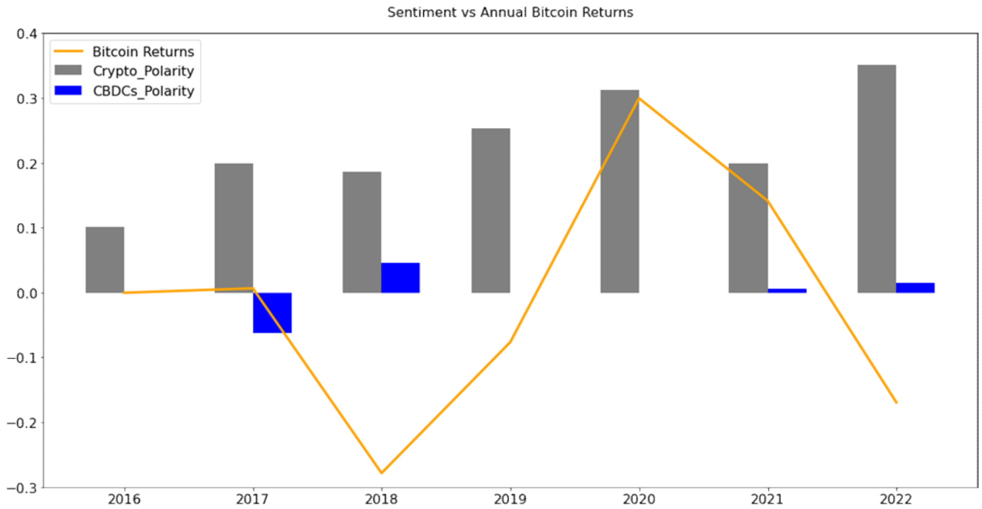
Chart 7: Top 50 Worldwide Commercial Banks' sentiment on CBDC and Crypto



b. <u>Historical view of sentiment and returns</u>

In the first chart we can already see a clear trend that indicates that whenever Bitcoin returns increase, commercial banks become more positive towards "cryptocurrency"; the opposite occurs for CBDCs.

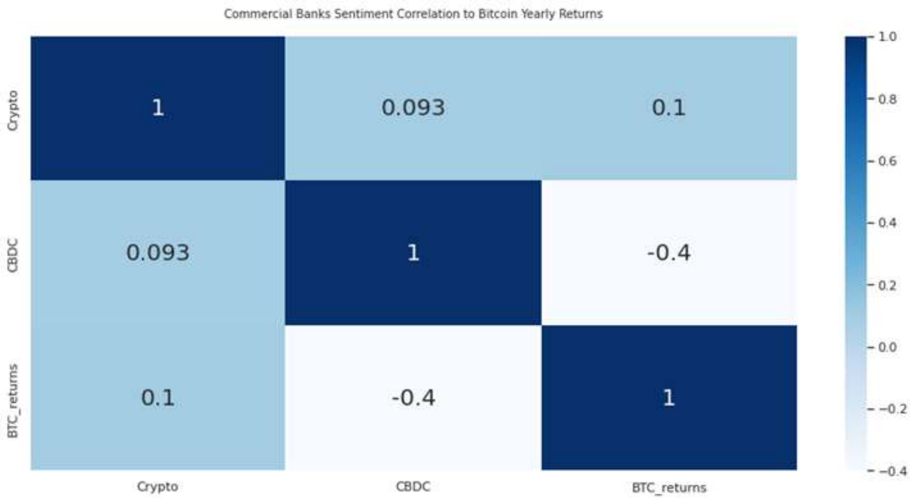
Chart 8: Historical Overview of Sentiment and Returns



c. <u>Correlation of Sentiment and Returns</u>

To get more clear measure we can apply a Pearson correlation to cryptocurrency and CBDCs sentiment and returns.

Chart 9: Correlation Matrix of Sentiment and Returns

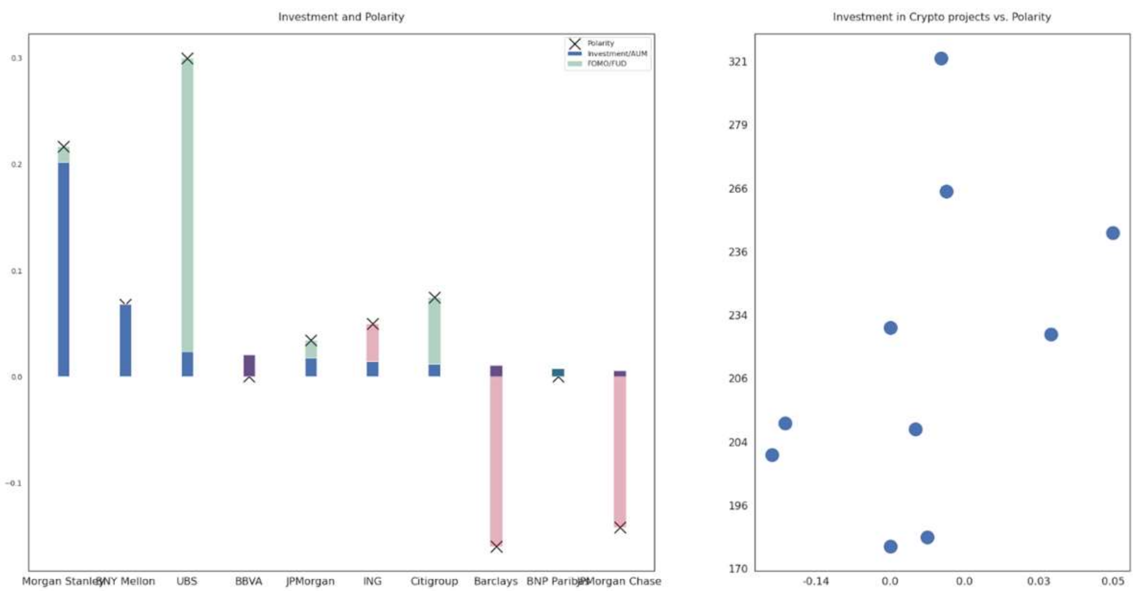


The Pearson correlation measure shows a strong negative correlation between the sentiment from Commercial Banks on CBDCs and the Yearly returns. For example, if Bitcoin returns increase then Commercial Banks tend to Become 10% more positive towards cryptocurrencies and 40% more negative towards CBDCs. With this we can state the preliminary conclusion that commercial Banks' sentiment moves with Bitcoin returns in a direct relationship.

d. Investment vs. Sentiment

Additionally, we have considered interesting to compare investment in the crypto industry and sentiment from top investing banks globally. For this, we take the top 10 commercial banks that invest the most in the cryptocurrency industry. We have compared individual sentiment to the ratio investment/AUM. The idea behind is to approximate the level of confidence from top investing banks in the cryptocurrency industry and further see if the reason for their large investments is due to FOMO, FUD or else. For example, we could say that UBS is positive but not confident while Morgan Stanley or BNY Mellon are positive and confident. Nevertheless, confidence does not mean risk taking and this data can give us a measure of the risk aversion from each bank; for example, UBS may show a risk-taking position in this industry, while Morgan Stanley or BBVA can be more risk averse.

Chart 10: Relationship Between Investment and Sentiment from Top 10 Crypto Investing Banks



From these two plots we can make a hypothesis about the FOMO or FUD of Banks. For example, UBS appears to be very positive but invests a lot less than Morgan Stanley, which could mean that they consider it risky or are still skeptical about it.

4. Predicting Sentiment trends and Price of Decentralized Assets

After building an understanding around the positions towards centralized and decentralized digital assets from main institutional parties, we have proceeded to conduct a more in-depth analysis to find a practical use of Sentiment Data. For this work a simple price prediction model for Bitcoin is built to prove if Sentiment has a predictive value over the price decentralized digital assets.

Although the model's results are acceptable it only takes one feature variable to make predictions and it has room for further improvements. Therefore, next steps and improvements are proposed to create more complex models that could presumably predict price more accurately; for example, making use of each bank's sentiment to calibrate news sentiment or using Twitter to add the general public's sentiment. The main of this point is to serve as starting point for more complex models that could take in all the insights from the previous points where we analyzed sentiment data.

4.1. Data inputs for the MCE model

The data used for the model will be, on one hand, the daily polarity from news that correspond to the query: "Bitcoin News", and on the other, the daily price of the predicted asset, in this case, Bitcoin.

4.1.1 Methodology

a. Getting the Data

As well as with previous analysis, to get the daily news a web scrapping automation is built. Once we have daily news, we will be able to create a table with Date, Sentiment and Price.

Procedure 9:

Getting Daily "Bitcoin News" Google Search Results

```
"11/11/2022": ["Top cryptocurrency to explode in 2023 | HeraldScotland", "Sam Bankman-Fried steps down, FTX files for bankruptcy", "Scaramucci talks FTX, Sam Bankman-Fried and 'the ... - CNBC", "Australian investors in limbo after collapse of FTX ... - ABC", "Cryptocurrency exchange FTX files for bankruptcy protection ...", "Cryptocurrency exchange FTX now worthless, says key investor", "Cryptocurrency News And Price Weekly Wrap-Up For Nov. 11", "Latest Cryptocurrency News | Nasdaq", "Cryptocurrency News | Seeking Alpha"]
```

```
#Reading json object containing daily news
with open('/content/drive/MyDrive/Colab Notebooks/NLP CBDC/Regression Model
/cryptocurrency_news.txt') as f:
    dct = json.load(f)

dates = list(dct.keys())
filter_for_BTC = [i for i in dct.values()]
cleaned_news = []

for i in filter_for_BTC:

    j = (' '.join(i)).replace('...','').replace('-','').replace(':','')

    cleaned_news.append(' '.join([i for i in j.split(' ') if i not in stop_
words]))

Polarity = []

for i in df['Day_news']:

    Pol = list(TextBlob(i).sentiment)[0]
    Polarity.append(Pol)

df['Sentiment'] = [i for i in Polarity]
```

	Date	Day_news	Sentiment
0	2016-01-01	Bitcoin's Big Challenge 2016 Reaching 100 Mill...	0.213095
1	2016-01-02	CRYPTO20 Tokenized Cryptocurrency Index Fund ...	0.146259
2	2016-01-03	Ethereum Announces Launch Homestead Cointeleg...	0.028571
3	2016-01-04	Vatican Slated State Adopters Cryptocurrency A...	0.241667
4	2016-01-05	ELSAGB Societatea Energetica Electrica SA CNB...	0.144540

b. Getting Daily Bitcoin Price

The daily bitcoin price is obtained from the Yfinance module provided by python, this module requests data to the yahoo finance API. Once we have these two data series, we can compute daily sentiment and separate data into train and test data.

Procedure 8:

Getting The Daily Close Price of BTC

```
#Getting price of bitcoin for each date
import yfinance as yf

ticker = 'BTC-USD'
daily_price = []

price = pd.DataFrame(yf.download(tickers=ticker,interval='1d',start=dates[0
],end=dates[-1])['Close'])
price.reset_index(inplace=True)
price['Date'] = price['Date'].astype(str)
```

	Date	Close
0	2016-01-01	434.334015
1	2016-01-02	433.437988
2	2016-01-03	430.010986
3	2016-01-04	433.091003
4	2016-01-05	431.959991

c. Cleaning Noise from Sentiment Data and transforming Price into Logarithmic Price

When previewing the prediction data, Sentiment and Close price did not show a linear relationship at first. This was because of the vast amount of data. The daily sentiment as is, is not sufficient to show a trend.

To solve it and uncover a trend from sentiment data we used an Exponential Moving Average of 360 periods, meaning that each day's sentiment is the exponential average of the previous 360 days. Additionally, "Close" is transformed into logarithmic scale to be able to have all models .

Chart 11:

Before and After Transforming Sentiment

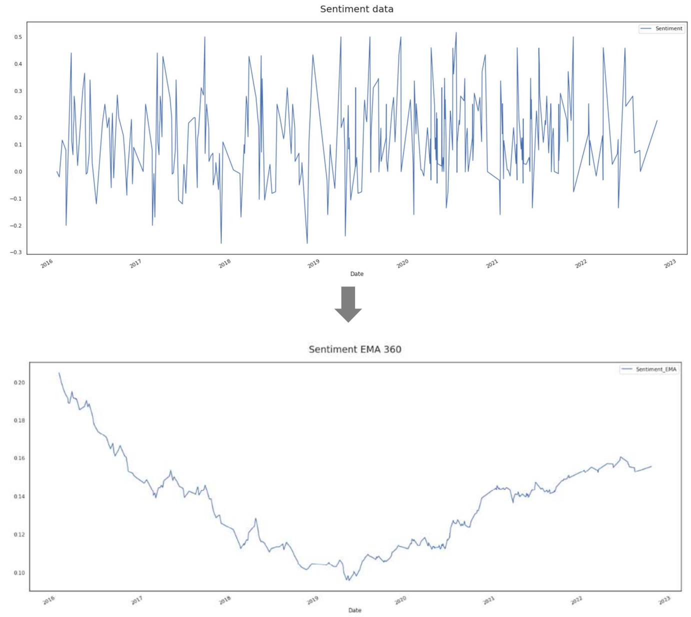
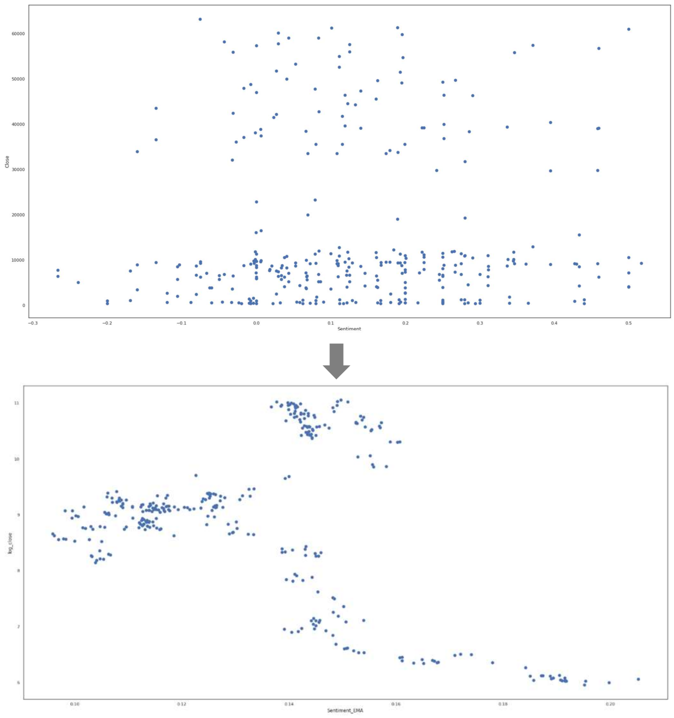


Chart 12:

Before and After Transforming x and y



4.2. Train and Test periods

Now that the data is transformed and shows a relationship, we separate the whole dataset table into to two sets: train and test. The train period will be from the 1st of January of 2018 to the 1st of January of 2022. The 2016 year is discarded with the premise that information about Bitcoin was not accurate, and 2017 is also discarded as we want to get only EMA that take 360 days and not less; for example, if we took as starting point 2017 the sentiment's exponential MA would not show such as clear trend.

Procedure 9:

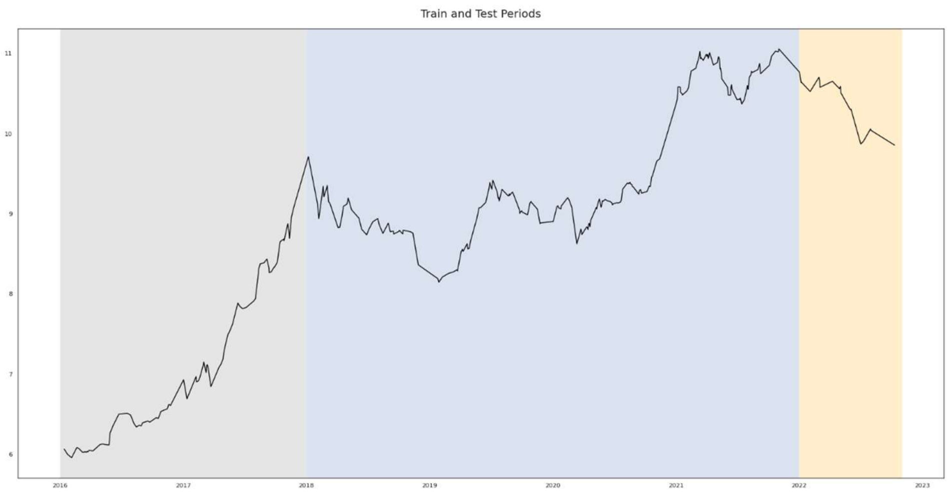
Creating The Train Data

```
#Train data
df_model_train = df_model[(df_model.Date > '2018-01-
01') & (df_model.Date < '2022-01-01')]
```

Date	Day_news	Sentiment_EMA	log_close
2018-01-07	Inside bitcoin bubble Which? investigates craz...	0.122600	9.709757
2018-02-03	Will Cryptocurrency Replace National Currencie...	0.114572	9.124228
2018-02-07	Over 800 cryptocurrencies dead bitcoin CNBC ...	0.112614	8.938702
2018-02-20	Coders Are Trying Connect Bitcoin's Lightning ...	0.115413	9.341693
2018-02-22	world's mostused cryptocurrency bitcoin Offlin...	0.114648	9.210840

Chart 13:

Train and Test Periods (blue for train, orange for test)



4.3. Linear Regression MCE Model

Once we have the train data, we fit a linear regression model using python. The model fitted is written as:

$$\log_{close} = \beta_0 + \beta_1 \times s + \varepsilon$$

4.3.1 Methodology

To train the linear model we use the "df_model_train" table. This data is then fitted into a linear regression function from the TensorFlow library , which is a well-known Machine Learning library within python.

Procedure 10: Training Linear Model

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, mean_squared_error
#X
features = 'Sentiment_EMA'
X_train = df_model_train[features]
X_train = X_train.to_numpy().reshape(-1,1)
#y
target = 'log_close'
y_train = df_model_train[target]
#BaseLine
y_mean = y_train.mean()
y_pred_baseline = [y_mean] * len(y_train)
```

⬇

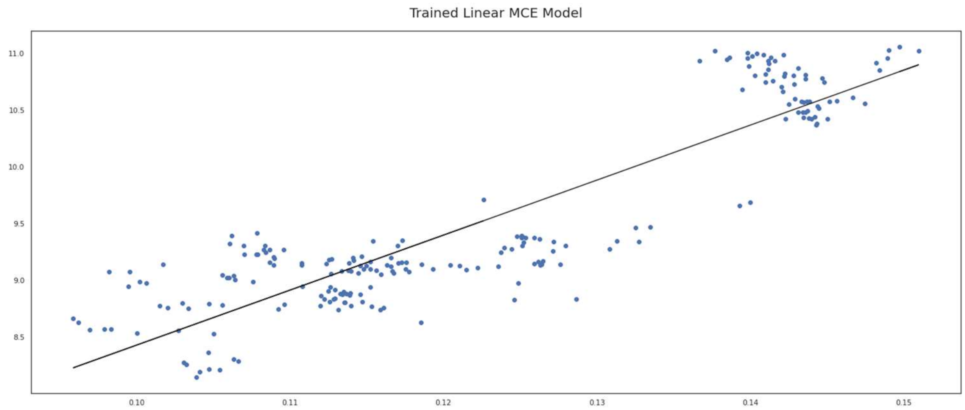
```
linear_model = LinearRegression()

linear_model.fit(X_train,y_train)

y_pred_training = linear_model.predict(X_train) #yhat predictions

mae_training = mean_absolute_error(y_train, y_pred_training)
```

4.3.2 Training Results

After training the model we get the following results:

R^2	0.8096249136850192
MAE	9.391325641714495
MSE	88.8540391345836

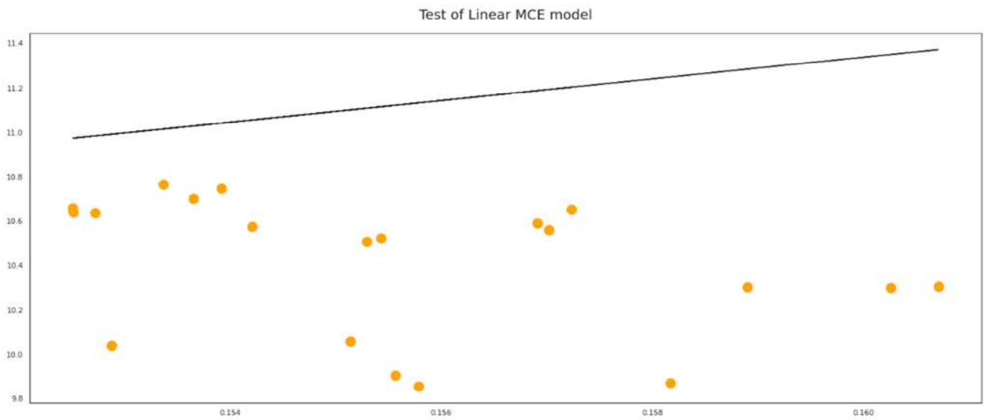
Chart 14: Training Linear Regression



4.3.3 Test Results

Test results from the linear regression show that the linear model fails to predict price observations from 2022. These observations correspond to a high volatility period.

Chart 15: Testing Linear Regression



4.4. Non-Linear Regression MCE

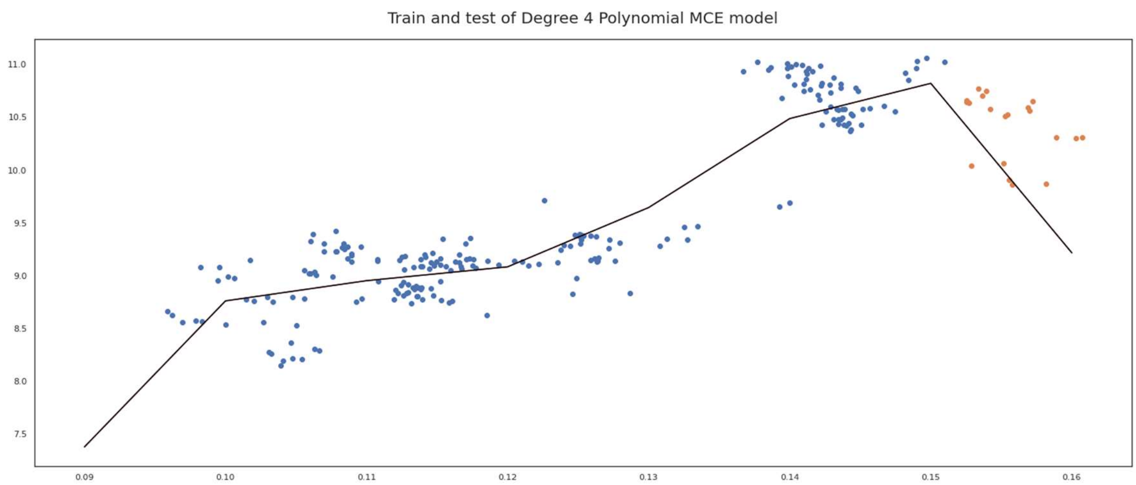
After Observing that the Linear Model was not able correctly predict test data due its lower values from all test observations, a polynomial regression is calculated to find a closer fit to the train and test data. For this a degree 4 polynomial regression is calculated. The degree 4 was found to best capture abrupt changes while also following the data better.

Procedure 11: Training Polynomial Model

$$\log_{close} = \beta_0 + \beta_1 \times s + \beta_2 \times s^2 + \beta_3 \times s^3 + \beta_4 \times s^4 + \varepsilon$$

```
X = X_train
y = y_train

poly_features = PolynomialFeatures(degree=4,include_bias=False)
X_poly = poly_features.fit_transform(X)

reg = LinearRegression()

reg.fit(X_poly, y)

X_vals = np.arange(0.09,0.16,0.01).reshape(-1,1)
X_vals_poly = poly_features.transform(X_vals)
y_vals = reg.predict(X_vals_poly)
coefficient = list(reg.coef_)
```

Chart 16: Polynomial Regression



4.5. Ensemble Final Model

The final model is an ensemble model that takes a weight of the predictions from the linear and polynomial models. Ensemble is a way of improving predictions by combining several predictions or regressions that are weaker. The final prediction of the bitcoin price is obtained by weighting both linear and polynomial predictions as follows:

$$\log_{close} = 0.15 \times Linear_{Predictions} + 0.85 * Polynomial_{Predictions}$$

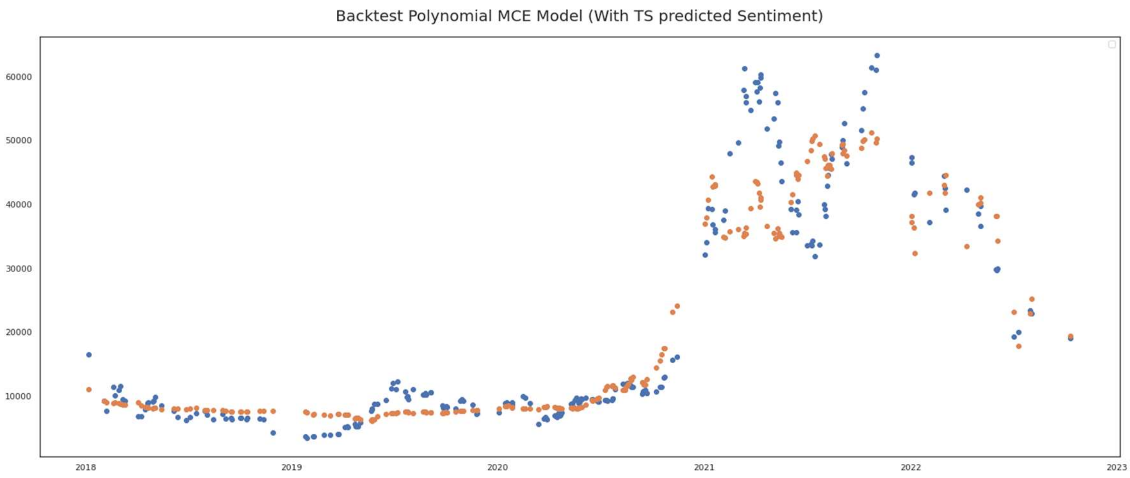
4.6. Back Testing the Model

For this test we will compare how the model it is able to predict past values. Also, we will test how it predicts price using time series predicted sentiment vs. predictions made with the actual Sentiment exponential moving average.

4.6.1 Using Sentiment Forecast Values from Facebook Prophet.

After already conducting all regressions and computing the ensemble model predicted values, we can test how this model could have performed with the predicted sentiment from the Facebook Prophet model.

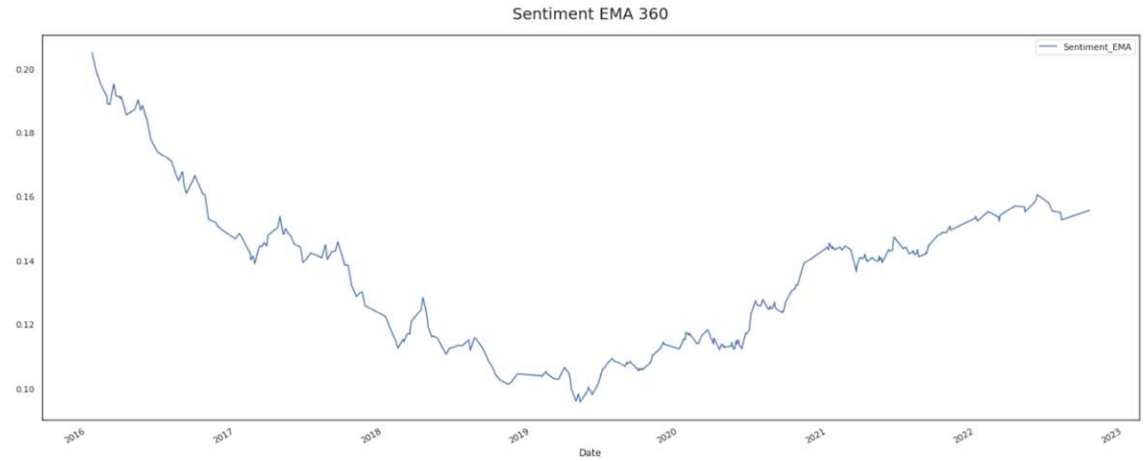
Chart 17: Backtest of Ensemble Model with Facebook Prophet Predictions

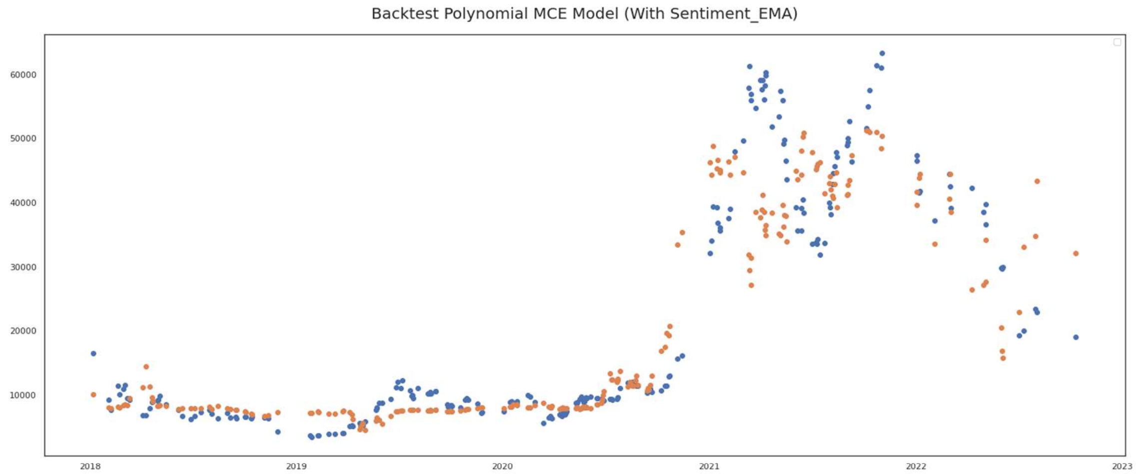


4.6.2 Using Actual Sentiment Exponential Moving Average Data

Instead of using sentiment predictions from a Time series model to later predict Price we tested to make predictions with the exponential moving average of sentiment that was calculated before the linear regression. The results below show that the model can still predict price but is less smooth when there are no major changes in sentiment; meaning that it is possible that using a shorter exponential moving average results will be more accurate in case we want to predict without the sentiment time series predictions.

Chart 18: Backtest of Ensemble Model with Facebook Prophet Predictions



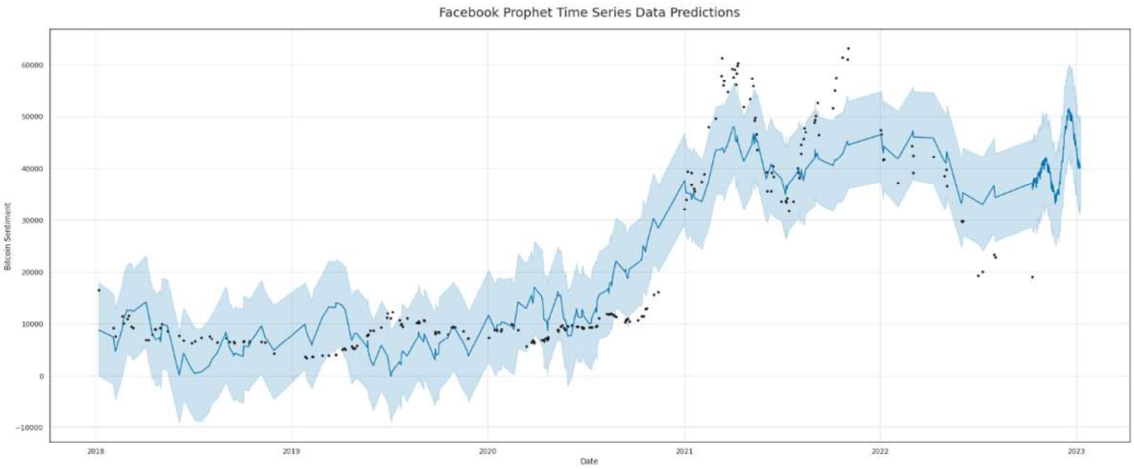


4.7. Comparing to a Time Series Model with No Sentiment Data

If we compare the model built that uses sentiment data to a time series model that takes only the price of Bitcoin, we can see that the second is less precise. The Facebook Prophet takes in the following feature variables: Trend values that do not repeat g(t), repeated seasonal changes s(t), and irregular changes h(t). To obtain predictions on the price of bitcoin we pass to the model all close prices since 2016 to get its predictions.

Chart 19:

Time Series Prediction with No Sentiment Data (FB Prophet)



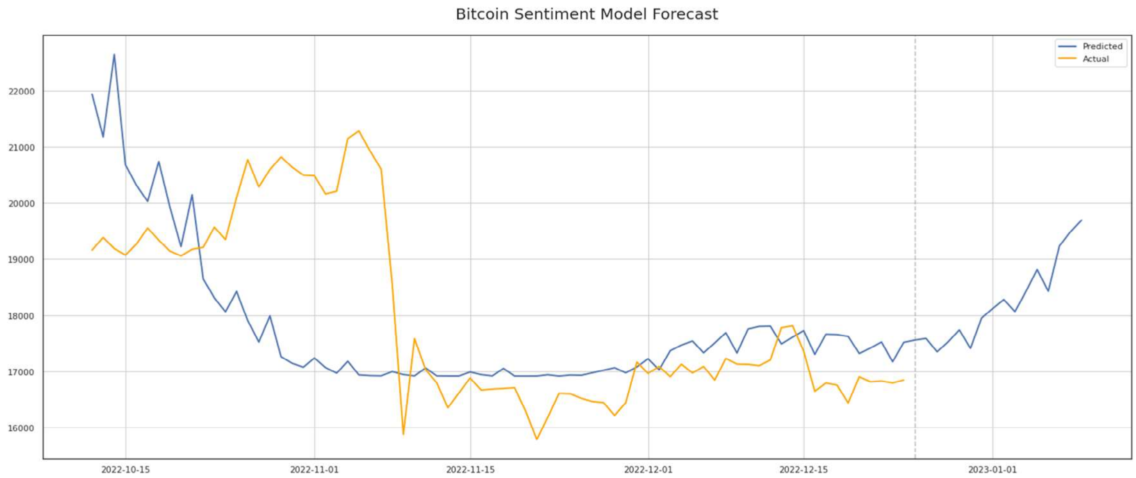
The predictions show that with no sentiment data the model fails to predict at times were price changes rapidly due to higher volatility. This shows that when sentiment data is used to predict price guesses are more accurate.

4.8. 90-day Forecast

As a mere curiosity we used the Model and the FB Prophet forecast of 90 days to plot our price prediction of BTC for the next 90 days. The results show that the model jointly with time series predictions gives reasonable predictions.

Chart 20:

90 Days Forecast of Bitcoin Price with Sentiment Data



4.9. Future Model Improvements

a. Train the model with more features.

For example: open, high, low, spot volume, derivatives volume, future contracts open interest, returns from opposite assets, mining difficulty, number of cold Bitcoin, tweet volume, Bitcoin dominance (BTC capitalization vs rest of crypto assets) etc.

b. Use Piece-Wise Regressions.

This regression are made by partitioning data in individual smaller regressions. The result is a set system equation of different regressions. This could be used instead of the polynomial regression to fit the different areas of the data.

c. Use different model ensemble techniques to improve the model's accuracy

5. Conclusions

a. Central Banks will become more positive towards CBDCs if Policies are implemented.
b. Central Banks Think Digital Assets can be a source for financial instability.
c. Meta Libra project has had an unprecedented impact in central bankers, as they realise that bigtechs' can directly act as central banks given that they have all the tools and resources.
d. Central Banks Are very concerned about the effects on monetary policy from implementing CBDCs.
e. Most negative central banks do not mention DLT while positive mention frequently.
f. Overall, most frequent significant word is "policy" in central banks speeches.
g. European central banks and the ECB are more negative than Asian CBs.
h. Most positive central banks do contemplate a coexistence of decentralized assets with CBDCs.
i. Usually central or commercial banks that are positive on CBDCs are positioned in the opposite way towards cryptocurrencies.
j. Australian commercial banks are among the most negative about cryptocurrencies.
k. Each time yearly returns of Bitcoin increased, overall sentiment from commercial banks on cryptocurrency increased, while it decreased for CBDCs.
l. Commercial Banks become positive by 10% if Bitcoin returns increase.
m. Commercial Banks become negative by 40% towards CBDCs if Bitcoin returns increase.
n. BNP Paribas or BBVA are quite neutral when investing in the Crypto industry.
o. UBS invests a lot less than it is positive.
p. There is a positive linear relationship between the investment in cryptocurrency assets and the sentiment from commercial Banks.
q. Daily news from Google can predict price trends of BTC.
r. Sentiment data along time series predictions can accurately make price trend forecasts.

6. Bibliography

Lamport, L. (1983). The Byzantine Generals Problem. *Journal of the ACM*, [online] 30(3), pp.668–676. doi:10.1145/2402.322398.

Rhodes, D. (2020). *The Byzantine Generals Problem, Explained*. [online] Komodo Academy | En. Available at: https://komodoplatform.com/en/academy/byzantine-generals-problem/.

Weidmann, J. (2020). Jens Weidmann: Welcome address for Christine Lagarde. [online] Available at: https://www.bis.org/review/r200925b.pdf [Accessed 23 Dec. 2022].

Lagarde, C. (2020). Payments in a digital world. www.ecb.europa.eu. [online] Available at: https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp200910~31e6ae9835.en.html.

BIS Monetary and Economic Department (2022). The future of the monetary system Monetary and Economic Department. [online] Available at: https://www.bis.org/publ/bppdf/bispap131.pdf

Review of the FSB High-level Recommendations of the Regulation, Supervision and Oversight of ‘Global Stablecoin’ Arrangements Consultative report. (2022). [online] Available at: https://www.fsb.org/wp-content/uploads/P111022-4.pdf

Central Bank Digital Currency tracker (2022) Atlantic Council. Available at: https://www.atlanticcouncil.org/cbdctracker/

Alhamid, M. (2021). Ensemble Models. [online] Medium. Available at: https://towardsdatascience.com/ensemble-models-5a62d4f4cb0c.