University of Pisa

Master's Degree in

Artificial Intelligence and Data Engineering

Sentiment Analysis

on Metaverse's Tweets

Candidates

Del Seppia Matteo

Minniti Federico

Terranova Franco

ACADEMIC YEAR 2021/2022

INDICE

INTRODUCTION

Data collected from social media can reveal important information about the thoughts of society towards a certain topic, event, or situation. One of the platforms where people typically express feelings is Twitter, and tweets' analysis can be a useful tool to determine people's sentiments about topics. We've focused on determining feelings and sentiments around Metaverse.

By definition, the metaverse is a vast network where individuals can interact socially and professionally via their avatars, invest in cryptocurrencies and NFTs to own unique meta-wearables and virtual lands, take classes, work, and travel in 3-D virtual reality.

Twitter sentiment analysis is typically handled as a text classification problem.

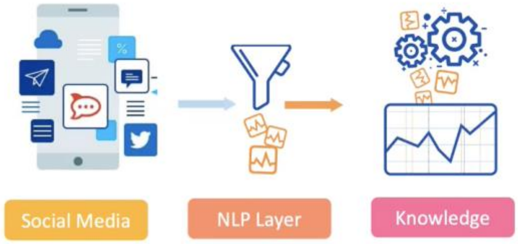

Figure 1. Sentiment Analysis WorkFlow

1.1. HOW CAN COMPANIES USE OUR TOOL?

Over the past years, the metaverse is growing fastly and while the metaverse hype continues to rise, companies interested in activating virtually should take precautions on satisfying correctly potential customers.

Sentiment analysis can be used to understand the opinions and concerns of customers and receive their feedback. Using our tool, companies can understand key actions of competitors, already entered in this new ecosystem, that obtained positive sentiment.

Moreover, understanding the concerns of users on companies in the market could allow a company to avoid their same errors.

1.2. DATA COLLECTION

In our study, we’ve started collecting Twitter posts containing the hashtag #*metaverse*. For this aim, *snsscrape* tool has been used, a Python library used to scrape tweets through Twitter's API.

1.3. BOTOMETER

Social media provide support for anonymity and easy ways to automate information production and diffusion. A social bot is a social media account controlled by a software in an automatic way. Bots were easily recognizable at the beginning, then become less recognizable with the improvement of bot techniques. Bot detection can be accomplished in different ways, one of which is *Individual Bot Detection*.

The data scraped using the metaverse topic contained a lot of tweets probably produced by bots. For this reason, we’ve used *Botometer*, an ensemble of Random Forest classifiers taking in input 1200 features of different kinds, with an accuracy close to 90%.

This tool checks the activity of a Twitter account and gives it a score. Higher scores mean more bot-like activity. We’ve used a score threshold of 0.6 to determine the propensity of the users to be bots and ignored their tweets as a consequence, limiting us though to the requests limit of the Botometer Free Plan.

1.4. NLP PIPELINE STAGES

Here we highlight the stages in our NLP pipeline.

1) In our *data cleaning* step, we’ve used different regular expressions to clean our text, including operations such as the removal of URLs, twitter handlers, hashtags, special characters, and single characters, and the substitution of multiple spaces with single spaces.

2) *Tokenization* is the process of breaking down a text into the smallest unit of sentence, or token.

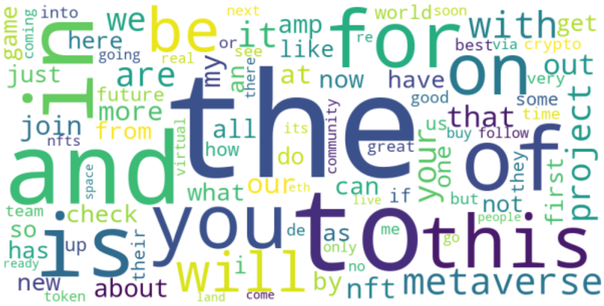

Figure 2. WordCloud after Tokenization

3) *Normalization* is the translation in lowercase text and the removal of accented characters.

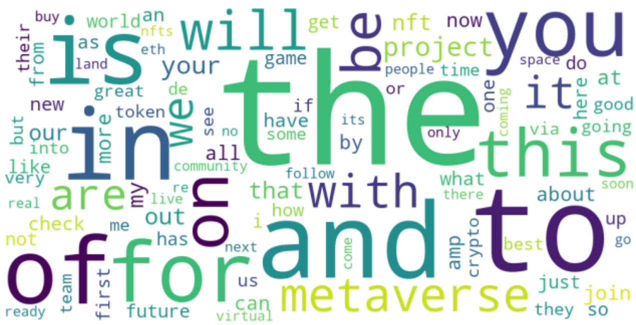

Figure 3. WordCloud after Normalization

4) *Filtering* is the stage involving the removal of stop words and short words.

Figure 4. WordCloud after Filtering

5) *Stemming* is the process of reducing inflected words in their stem form.

Figure 5. WordCloud after Stemming

2. SENTIMENT ANALYSIS

Sentiment analysis is a natural language processing (NLP) technique used to determine whether the sentiment of data is positive, negative or neutral. Sentiment analysis can be used on textual data to help businesses monitor customers' sentiments and understand customer needs.
We've used three different techniques to determine the sentiments of our tweets.

2.1. POLARITY AND SUBJECTIVITY

TextBlob is a Python library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks. When a sentence is passed into Textblob, it returns the polarity and the subjectivity of the sentence. Polarity is a metric that lies between [-1,1], where -1 refers to negative sentiment and +1 refers to positive sentiment. Subjectivity is a metric that lies within [0,1] and refers to personal opinions and judgments.

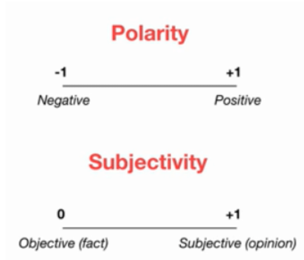

Figure 6. Polarity and Subjectivity

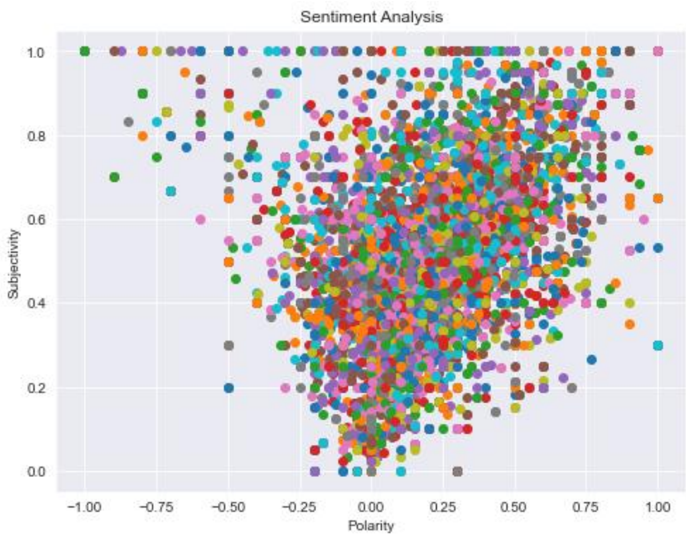

Figure 7. Polarity and Subjectivity of our dataset

We've noticed how most of the tweets concentrates in the zone with positive polarity.

We've then used the following rule to determine the sentiment of the tweets.

$$Sentiment = \begin{cases} Positive & if\ polarity > 0 \\ Neutral & if\ polarity == 0 \\ Negative & if\ polarity < 0 \end{cases}$$

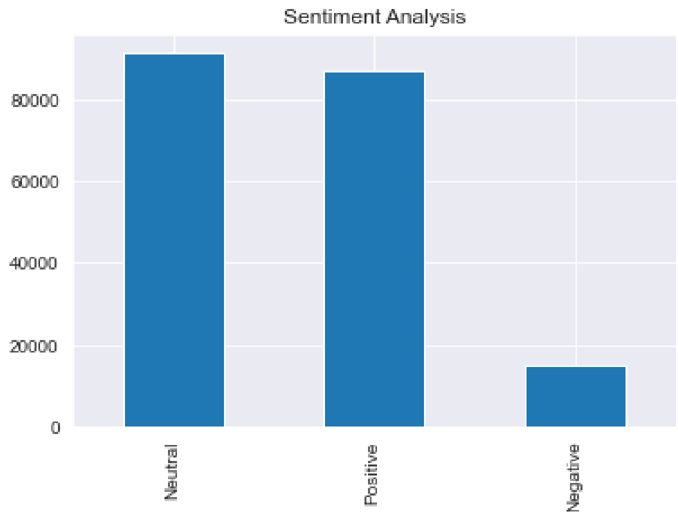

Figure 8. Sentiment Distribution of TextBlob

The previous bar chart shows the distribution of the sentiment of the tweets. We've then analyzed the sentiment change over time using the polarity returned by TextBlob, focusing on the daily sentiment mean and standard deviation.

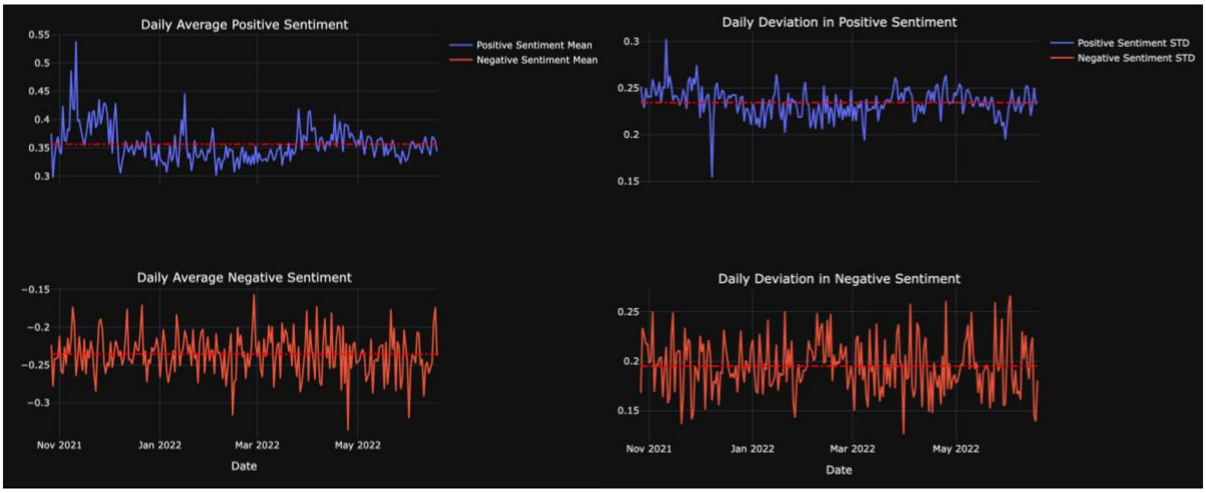

Figure 9. Daily Sentiment Variation of TextBlob's Polarity

2.2. SENTIMENT INTENSITY ANALYZER

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media.

It returns for each tweet post a polarity score that can then be divided into a Positive Sentiment Score, Negative Sentiment Score and Neutral Sentiment Score, each varying in the range [0,1]. The following chart shows the distribution of these three scores in our tweets.

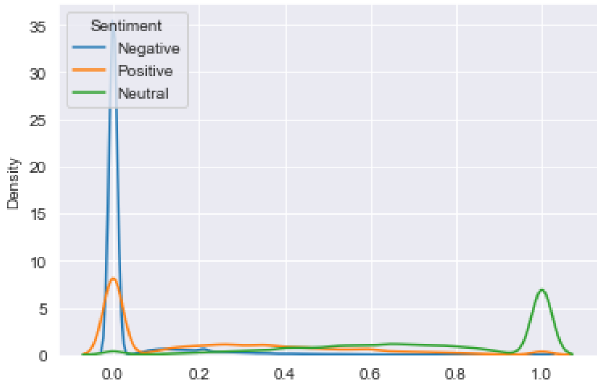

Figure 10. Sentiment Distribution of SIA

We can notice how most of the negative and positive scores have an average value close to 0, while most of the neutral scores are close to 1.

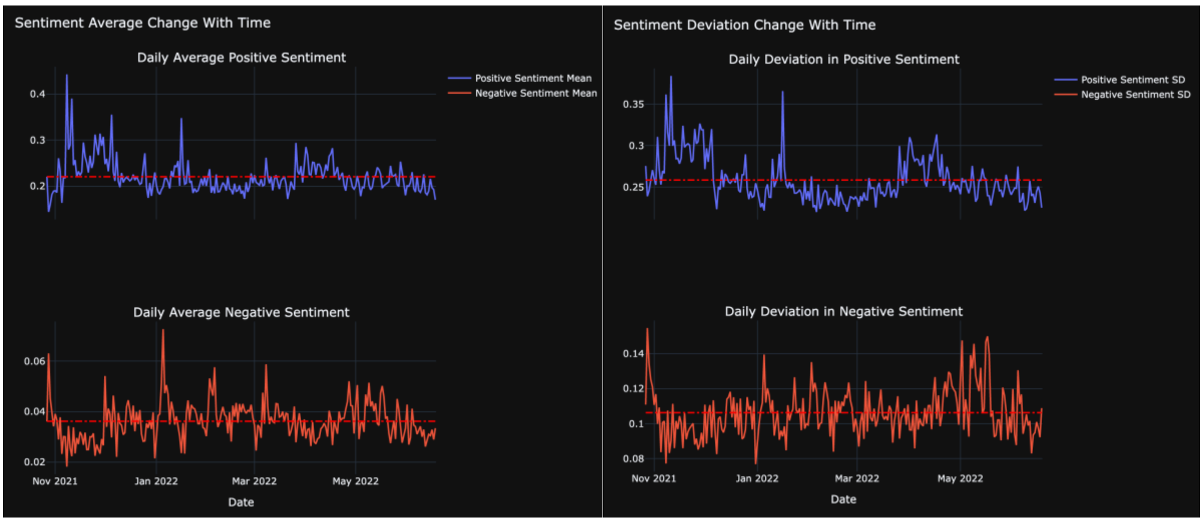

Figure 11. Daily Sentiment Variation of SIA

We've then analyzed the sentiment change over time using the positive and negative score, and in particular their daily sentiment mean and standard deviation.

We can notice how results are similar to the ones provided by TextBlob.

Since most of the negative and positive scores are close to 0, we've decided to set a cut-off to determine which tweets could be considered as negative or positive. We've decided to select as cut-off the mean value of the sentiment.

Figure 12. Cut-off for most negative tweets

2.3. CLASSIFICATION

We've decided also to consider a classification task in order to implement also a domain-specific classifier. For this aim, we've manually labeled a certain number of tweets to have a ground-truth. A python script was used to automatize the labelling task, reaching almost 1000 labelled tweets.

2.3.1. REBALANCING

Imbalanced classifications pose a big challenge as most of the machine learning algorithms used for classification were designed around the assumption of an equal number of examples for each class. Since an unequal distribution of classes was present within our dataset, we've started with a rebalancing task, using SMOTE as technique of oversampling.

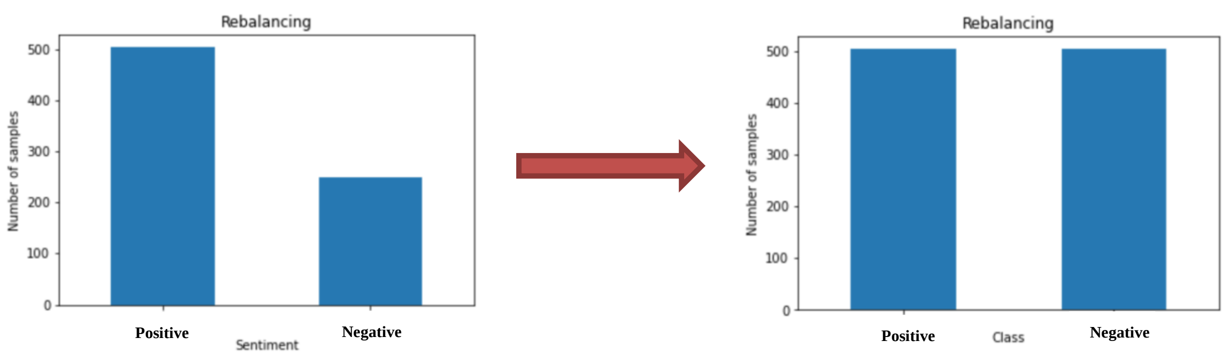

Figure 13. Rebalancing

2.3.2. FEATURE SELECTION

Feature selection improves the machine learning process by selecting the most important variables and eliminating redundant and irrelevant features, facing the problem of curse of dimensionality.

Univariate feature selection works by selecting the best features based on univariate statistical tests. We've used this technique, with *F-test* for feature scoring and the *ANOVA F-value*, to select the most significant features with *SelectKBest* as transform method, removing all but the k highest scoring features.

$$Univariate\ Score = -Log(pvalue)$$

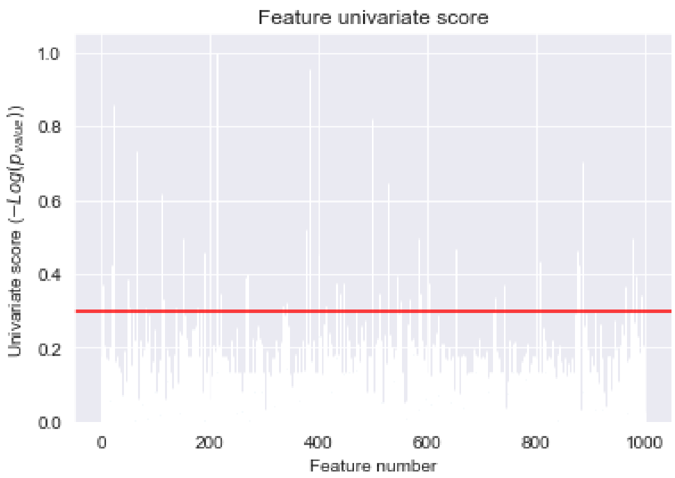

Figure 14. Features Selection

Using a score threshold equal to 0.3, we were able to select 56 features.

2.3.3. CLASSIFIERS COMPARISON

Once applied our preprocessing steps, we've evaluated different classifiers using the accuracy as evaluation metric and a K-Fold Cross Validation with K = 10, validation techniques for assessing how the results of a statistical analysis will generalize to an independent data set.

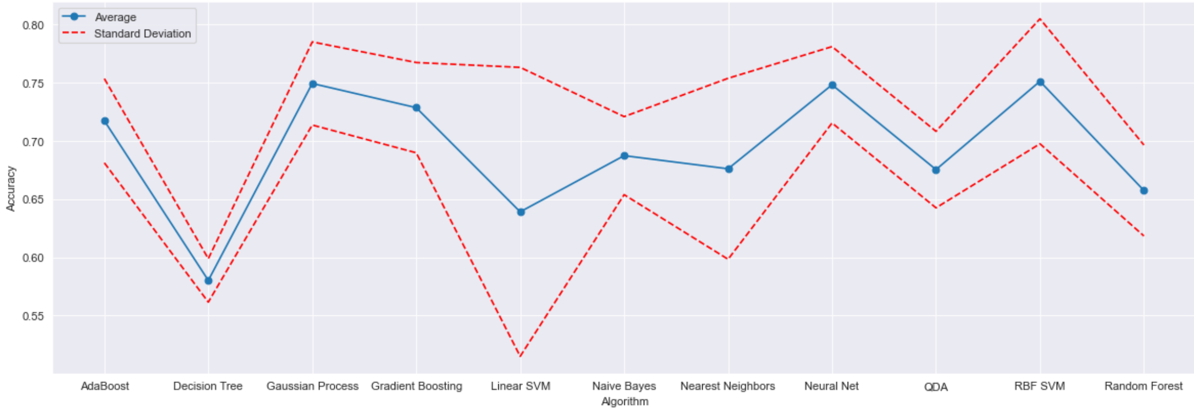

Figure 15. Classifiers Accuracy Comparison

Algorithm	Mean Accuracy	Std Accuracy
AdaBoost	0.72	0.04
Decision Tree	0.58	0.02
Gaussian Process	0.75	0.04
Gradient Boosting	0.73	0.04
Linear SVM	0.64	0.13
Naive Bayes	0.69	0.03
Nearest Neighbors	0.68	0.08
Neural Net	0.75	0.04
QDA	0.68	0.03
RBF SVM	0.75	0.05
Random Forest	0.65	0.03

Considering this metric, we've decided to use the Gaussian process model as classifier.

2.3.4. CLASSIFICATION RESULTS

Using this classifier, we've determined the number of positive and negative tweets over the number of tweets of each day and highlighted the following results.

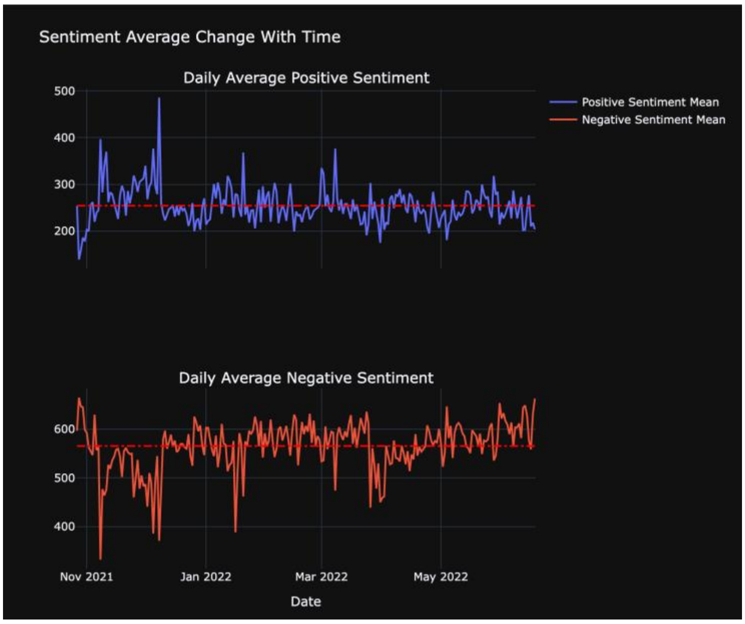

Figure 16. Sentiment Average Variation Gaussian Process

2.3.5. ENSEMBLE LEARNING

In the subsequent steps of the study, we've decided to use the sentiment results provided by the three models, combining them into a single model in an ensemble learning fashion.

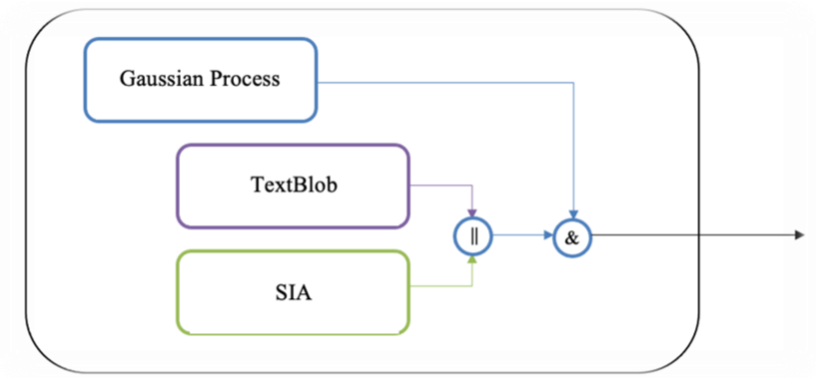

Figure 17. Ensemble Learning Model

Using the Ensemble Learning Model highlighted we were able to give a more accurate result of our classifier, merging its result with the ones provided by the other two models. Following this scheme, our classifier will confirm its classification only if one of the other two models agree.

3. TOPIC MODELLING AND CLUSTERING

Focusing on intervals of days in which we can identify peaks of daily average sentiments deviation, we've tried to extract topics in order to determine the main reasons that could have driven the sentiments' change.
Topic modeling is a statistical modeling technique for discovering the abstract topics that occur in a collection of documents. In this technique we end up with a list of topics each containing a set of words. We've also used *clustering* techniques to cluster documents in different groups based on a similarity measure.
For all our algorithms we've determined a suggested choice of parameters according to different heuristic methods and then explored a range of values close to the suggested point to determine a valid set of topics.

3.1. LATENT DIRICHLET ALLOCATION (LDA)

Latent Dirichlet Allocation (LDA) is a popular topic modeling technique to extract topics from a given corpus. It classifies or categorizes the text into a document and the words per topic, modeled based on the Dirichlet distributions. It describes the pattern of words that are repeating together, occurring frequently, and are similar to each other.

3.2. LATENT SEMANTIC ANALYSIS (LSA/LSI)

Latent Semantic Analysis, or LSA is an important technique in topic modeling. The core idea is to take a matrix of documents and terms and decompose it into a separate document-topic matrix and a topic-term matrix. It is a dimensionality reduction technique that projects documents to a lower-dimensional semantic space causing documents with similar content to be close to one another in the resulting space.

3.3. UMASS METRIC

For evaluating the optimal number of topics to extract using our Topic Modelling techniques, we've used the UMass Metric, defining a score based on document co-occurrence.

$$score(v_i, v_j, \epsilon) = \log \frac{D(v_i, v_j) + \epsilon}{D(v_j)}$$

where *D(x)* is the number of documents containing the word *x*.

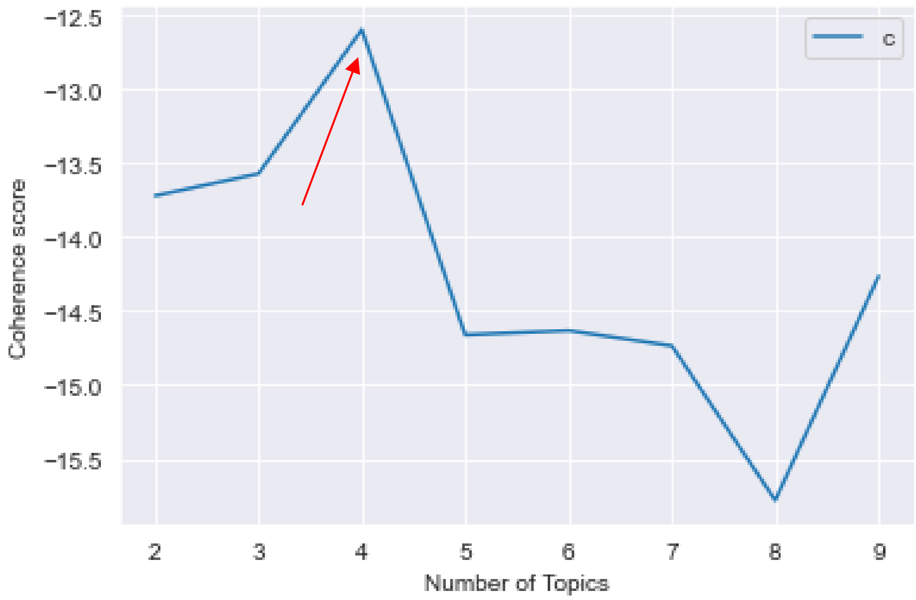

Figure 18. Example of evaluation of the coherence score

3.4. K-MEANS

K-Means algorithm is an iterative clustering algorithm that tries to partition the dataset into *k* non-overlapping clusters. To determine the best number of clusters to extract, we've used the Elbow method, which uses the sum of squared distance to determine the ideal value of *k* based on the distance between data points and their centroid. The most relevant turning point can be used as the optimal number of clusters.

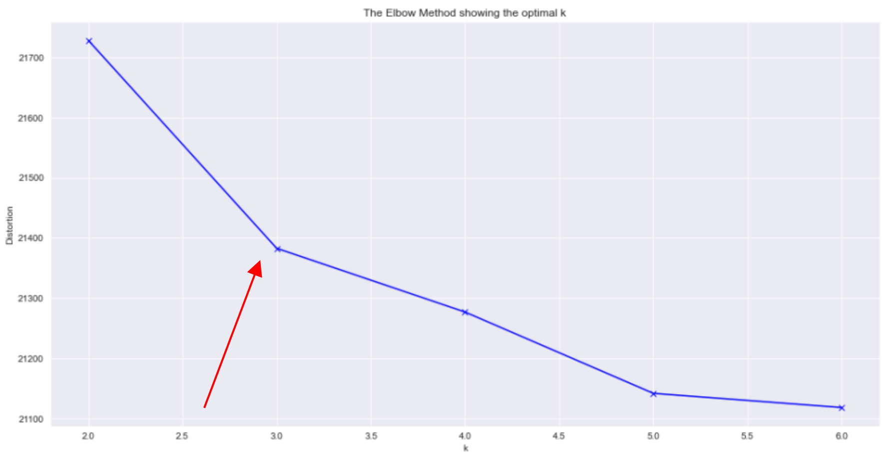

Figure 19. Example of evaluation of the Elbow Method

3.5. DBSCAN

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) relies on a density-based notion of cluster, which is defined as a maximal set of density-connected points and can be determined fixing two parameters, eps and MinPts, and searching for clusters by checking the eps-neighborhood of each object in the dataset. A typical suggested choice of MinPts is the following:

$$MinPts = 2 * NumFeatures$$

To determine the optimal value of eps we've used a heuristic approach called *k-dist*, mapping each point to the distance from its k-th nearest neighbor. We've sorted the points in descending order of their k-dist values and the graph of this function gives some hints concerning the density distribution. The threshold point is the first point in the first "valley" of the sorted k-dist graph.

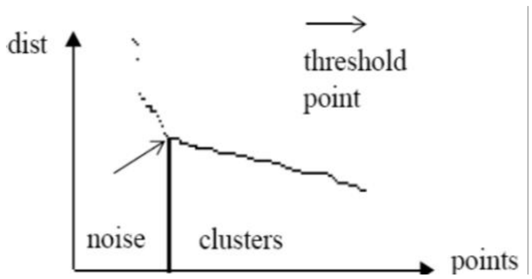

Figure 20. Example of evaluation of the k-dist function

4. RESULTS

Focusing on the peaks highlighted by the sentiment analysis, we've used positive and negative classified tweets, respectively, depending on the type of peak, provided by our ensemble model and used topic modelling and clustering techniques to determine the main reasons of this deviation of sentiments, obtaining the following results.

4.1. FACEBOOK'S REBRANDING IN META

We've started to analyze the peak pointed out from both negative sentiments' time series produced by our classifier and SIA, in the range of dates from 27/10/2021 to 29/10/2021.

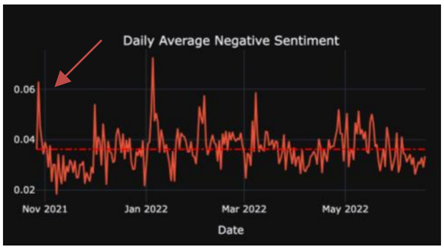

Figure 21. Facebook's Rebranding Peak

As highlighted from our topic modelling results provided by LDA, we've been able to identify the most relevant topic of that period that created negative sentiments in the community, Facebook's rebranding in Meta.

Figure 22. Topic Result from LDA about Facebook's Rebranding

Examples of tweets in that period are the following:

> *"Is @Facebook rebranding following a series of negative stories about itself or because it's building a #Metaverse?"*
>
> *"Either meta or facebook, it still doesn't eliminate the mess the vulnerability and the lack of user's privacy protection. Mark zuckerberg is ignoring the real issues and doing name rebranding"*

“#meta is a distraction and barely newsworthy. Facebook has died. They just trying to raise that dinosaur from the grave. 🦕🦖🦕🦖🦕”

“#MetaVerse as a caution not as a future opportunity for #MarkZuckerberg to change subjects and evade responsibility for @Facebook damaging democracy.”

4.2. ENS’ AIRDROP, DISNEY BUSINESS IDEA, IOEN’S PEAK

In the range of dates between 08/11/2021 and 12/11/2021, a convergence of positive events in the world of cryptos and mainstream entertainment contributed to the rise of positive sentiment towards the Metaverse.

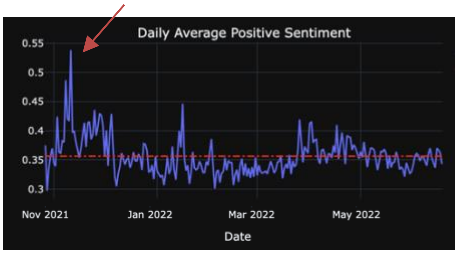

Figure 23. ENS' Airdrop, Disney and IOEN'S Positive Sentiment

LDA showed that the main topics in those days were:

- Disney’s CEO Bob Chapek announcement that the company would invest “to connect the physical and digital worlds even more closely, allowing for storytelling without boundaries in our own Disney metaverse”
- The Ethereum Name Services, a DAO whose primary goal is to give people and applications an easy way to read and share crypto addresses (similar to DNS), started the airdrop of ENS tokens to the holders of Ethereum wallet addresses. ENS tokens allow users to vote for executable and constitutional proposals inside the organization
- The IOEN, a crypto token whose goal is to be used as a staking asset to unlock new microgrid economies, reached its maximum in price leading to a consistent number of positive reactions on Twitter. The token has dropped in price and never recovered shortly after

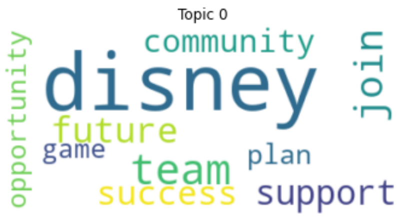

Figure 24. Disney's LDA Topic

Figure 25. ENS's Tokens Airdrop LDA’s Topic

Figure 26. IOEN LDA's Topic

Figure 27. CoinMarketCap IOEN's Price

4.3. NFT.COM REACHES 10K WHITELISTED USERS

On 30/03/2022 the NFT.com marketplace for non-fungible tokens announced that more than ten thousand users asked to be part of their whitelist. More than twenty thousand users joined the marketplace's Discord server. Being part of the whitelist has been necessary for users in April that wanted to participate in the Genesis Key Whitelisted auction. Genesis Keys are a collection of 10.000 unique fully animated NFTs, and their owners had first access to the NFT.com platform and can provide participation in its governance. Each key grants a holder to create two unique NFT.com profiles, also represented as NFTs.

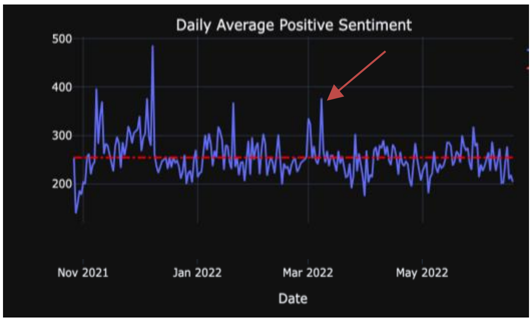

Figure 29. NFT.COM Positive Sentiment

Figure 28. K-Means NFT.COM Result

4.3. TERRA’S ECOSYSTEM CRASH

We’ve started to analyze the series of peaks pointed out from both negative sentiments’ time series produced by SIA, in the range of dates from 14/05/2022 to 18/05/2022. As highlighted from our topic modelling results provided by LDA, we’ve been able to identify the most relevant topic of that period that created negative sentiments in the community, Terra’s Ecosystem crash with UST and LUNA.

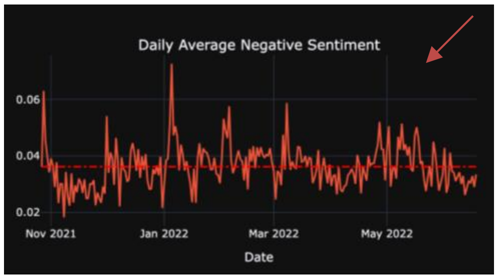

Figure 30. Terra's Ecosystem Crash Peak

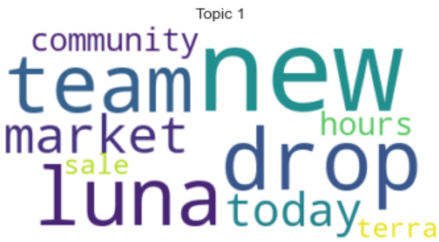

Figure 31. Topic's Results from LDA about Terra's ecosystem crash

Figure 32. CoinMarketCap UST and LUNA’s price

Examples of tweets in that period are the following:

“@SynthwaveManaic @Domi14253851 @stablechen #luna this luna project is faulty, (luna2.0. Llet's take out a new one, if not, we will release luna3.0 … there is no end to it... Because of luna, people now view crypto as a scam #web3 #metaverse etc... the whole market crashed”

“Everbody write Luna is scam binance is scam. Because you are Real power #binance #luna #Web3 #Terra_Luna #BSC #NFTs #BTC #Metaverse”

“These @terra_money idiots can't see the #LUNA community is committed to buy this project back up to respectability because they have no desire to make wise choices“

CONCLUSIONS

Sentiment analysis is a method nowadays used in many areas and able to provide effective results. The concept of Metaverse has increased its popularity in the last years. Our sentiment analysis tool may be useful for new companies ready to launch in this new ecosystem understanding in advance needs and concerns of users. This study has been conducted on the general topic of Metaverse. More detailed results may be obtained from a company by conducting an analysis on metaverse’s posts within their business area.