Journal of Metaverse
Research Article
Received: 2021-11-05 **|** **Reviewing:** 2021-11-05 & 2021-12-26 **|** **Accepted:** 2021-12-26 **|** **Online:** 2021-12-27 **|** **Issue Date:** 2021-12-31
Year: 2021**, Volume:** 1**, Issue:**1**, Pages:**25-30

Tweet Classification and Sentiment Analysis on Metaverse Related Messages

Özgür Ağralı
Muğla Sıtkı Koçman University
Graduate School of Natural and Applied Sciences
Department of Artificial Intelligence
Muğla, Türkiye
agrali.ozgurr@gmail.com
0000-0001-6974-011X

Ömer Aydın
(Corresponding Author)
Manisa Celal Bayar University, Faculty of Engineering
Department of Electrical and Electronics Engineering
Manisa, Türkiye
omer.aydin@cbu.edu.tr
0000-0002-7137-4881

***Abstract*— The data obtained from social media platforms is a popular study subject nowadays. These studies give important information about the thoughts of the society towards an event, situation, or concept. For this purpose, several studies have been carried out with different methods in the literature. These studies mainly try to obtain meaningful results by applying various methods according to the language of the social media content. One of these platforms where people freely express their feelings and ideas is Twitter. It is a popular and useful study to examine people's feelings and tendencies about a topic by doing tweet analysis. In this study, feelings about Metaverse are tried to be evaluated. We evaluate the tweets posted one week ago and later of the date Mark Zuckerberg announced that her company would change its name to Meta. Tweets sent in English with the "metaverse" hashtag on Twitter were used as the dataset. These tweets were analysed by the Sentiment Analysis method. Obtained findings and results are shared comparatively.**

***Keywords*— Sentiment analysis, Metaverse, tweet classification, Twitter, NLP**

I. INTRODUCTION

The concept of Metaverse was put forward in the early 1990s and with the development of technology, this concept has become applicable. The concept of Metaverse has increased its popularity as the internet connection has reached a sufficient level, its prevalence has increased, and the hardware that will provide access to this environment has been produced and accessible. This concept, which we can also call a virtual universe, allows people to exist mentally in a virtual world. Unlike today's systems, it promises to experience the feeling of reality in a virtual world while being physically in any environment with virtual reality equipment. It has similarities with the situation in which people are mentally in a different universe from their physical existence, such as daydreaming, dreams and hallucinations. However, in the Metaverse, signals received from people's bodies in the real world can be transferred to the virtual world. Its current core application necessitates physical activity in the real world. It can be predicted that in the future, there will be the possibility of existence in the Metaverse by processing thoughts directly without the need for physical activity. There are those who see this universe as the next version of the internet. Others see it as a great opportunity and developmental step for humanity. In this sense, the number of people who have positive ideas is quite high. Especially the elimination of physical boundaries, the possibility of people to realise what they want freely and without harming other people, and even the feeling of being immortal in this environment attracts people. Despite all these reasons, serious reactions have emerged that show that it will cause serious damage to the future of humanity and that humanity will even come to an end.

In the current technological environment, social networks are used extensively by people. Due to their nature, social networks are an important opportunity for people to share their ideas and feelings. People share ideas for all kinds of events, situations, and news in these environments. By examining what is shared in this environment, the tendencies, ideas, and possible actions of large masses can be predicted. Therefore, the processing and examination of social media data is an important issue. Every day, different applications are put forward on different data sets.

There are many studies in the literature to make sentiment analysis from data obtained from social networks. It is possible to have an idea about the future of these topics by examining current or popular topics with sentiment analysis over social network data. This study is set out to examine the tweets shared on Twitter about the reactions and thoughts about Metaverse and to evaluate them according to whether they are positive or negative.

II. RELATED WORKS

There are many studies in the literature examining sentiment analysis. Pang and Lee reviewed the studies conducted for the sentiment analysis in 2008 [1]. In their early study, classified the texts primarily according to whether they contain sentiment or not. They then proposed a hierarchical scheme that classifies them as positive or negative [2]. In another study, emotion icons were used to analyse sentiment [3]. Successful results can be obtained in such a method because many tweets contain icons and icons are designed for emotional expression [4].

Twitter sentiment analysis is basically handled as a text classification problem. The results of the pre-processing technique used for text classification may vary. In some studies, methods such as information retrieval, text classification and document filtering can be used. In 2018 Çoban and Tümüklü-Özyer examined term weighted methods which are newly proposed for various research areas. They suggested a term weighting method to be used in Twitter sentiment analysis. In this study, Bag of Words (BoW) and character level (N-gram) models were used for feature extraction. Turkish and English tweets were used as dataset. Latent Dirichlet Allocation (LDA) based subject model was applied for sentiment classification. In addition, the Support Vector Machine (SVM) algorithm was used in classification. They claim that this method is the most effective method by looking at the experimental results they obtained [5].

In some studies, sentiment analysis was used on twitter to detect customer opinions. Thus, companies and organisations have the opportunity to receive feedback on their success in the market. Sarlan et al. did a study that examined a large number of tweets and used prototyping. As a result of their analysis, they found 84.1% "null", 6.5% "negative" and 9.4% "positive" tweets. As can be understood from the results, the study has serious shortcomings and needs to be improved [6].

Karayiğit et al explained the relationship between the concepts of opinion mining.and sentiment analysis. The terms used for these two concepts are explained and examples are given for Turkish sentiment analysis. In addition, the problems encountered while carrying out Turkish studies were examined and solutions were offered for these problems [7]. Abali et al conducted a study using Turkish tweets from the Aegean Region of Turkey. With this study, it is aimed to identify the problems of people and to create an intelligent system that find out the location of the problems [8].

Ayan et al. tried to determine the situation of Islamophobia on social media with sentiment analysis. For this reason, tweets obtained from Twitter were examined. Linear ridge regression and Naive Bayes Classifier were used for training the models. Afterwards, estimations were made using precision, recall, F1 measures [9].

Fadel and Öz applied sentiment analysis to understand the comments and thoughts. In their study, they proposed a method to automatically classify tweets sent on Twitter after a terrorist attack. They used a machine learning approach. They used a lexicon approach to generate labelled training datasets. Naive Bayes applied majority voting among the Support Vector Machine and Logistic regression machine learning classification algorithms. The results show that the model achieved 94.8% accuracy with an F1 score of 95.9% [10].

Alqaraleh proposed a new and efficient sentiment analysis system. This system also supports the Turkish language. Due to the language structure of Turkish, a preprocessing model has been applied in order to work for agglutinative language. He made experiments using the "Turkish movie reviews" [11] data set. He reported that the approach he came up with gave an efficient and good performance for Turkish [12]. However, since the dataset used belongs to 2013, it could be worked with a more up-to-date dataset.

Kemaloglu et al. proposed a sentiment analysis system using data from RSS feeds, Youtube, Facebook, Instagram and Twitter. Classification and deep learning algorithms such as Logistic Regression, Random Forest and Long Short-Term Memory (LSTM) were used. The LSTM model provided the highest accuracy in their study [13].

Ergül Aydın et al. conducted a sentiment analysis (SA) on Turkish tweets collected about an Open and Distance education (ODE) System to monitor students' views and feelings about the system. In the study, they collected 63699 tweets. It is aimed to develop strategies that will increase the quality of education and training services by quickly learning the ideas of the students about the system [14].

Dalkilic and Cam analysed the tweets shared on Twitter about current series and movies with sentiment analysis. In this way, they will be able to make suggestions about TV series and movies. Turkish TV series and movie reviews were examined. Naive Bayes (NB), Support Vector Machines (SVM) and Random Forest (RF) from machine learning models were used. It is also trained for sentiment analysis by taking polarity values obtained by the general ensemble algorithms such as Bagging and Voting. The voting algorithm's accuracy is 87%. On the other hand, SVMs give the best area under receiver operating characteristics curve (AUC) of 0.96 [15].

Tokcaer examined sentiment analysis studies that used Turkish texts and also applied methods, databases, and open source libraries used in previous studies [16].

III. Method and Implementation

In the study, tweets in English shared with the word "metaverse" were examined and classified as positive, negative and neutral. Exploratory data analysis and data pre-processing stages are very important for performing accurate analyses on social media data. Different data exploration and pre-processing will result in different analysis results. For this reason, the application of standard methods for analysis produces an optimum result. In this study, Twitter posts containing the word "metaverse" in two different date ranges will be examined. While doing this review, the first step is to obtain the data. For this, the snscrape tool will be used. The collected data will be kept in two different csv files. Different notebooks were used for each data set. The second step is to transform the obtained data into a structure suitable for analysis. For this, data pre-processing processes covering 80% of data science processes will be applied. These processes are discussed under two headings, data and text pre-processing. After the data is brought into a structured state, sentiment analysis inferences will be made with Textblob, Vader and Afinn sentiment analysis tools. The applications in this study were Python 3.7.12 and Colab Notebook. The flow of the methods to be applied is given in Fig 1.

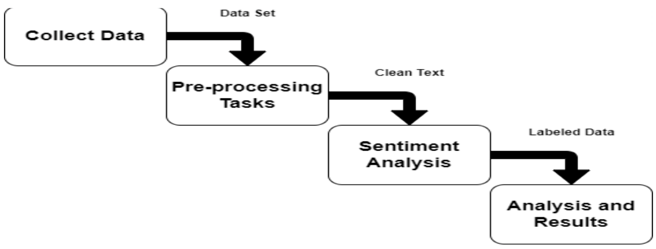


Fig 1. Flow chart of the method to be applied

A. Dataset

Since no ready-made data on the studied subject could be obtained, these data were obtained using specific web mining and web scraping tools. There are many methods to obtain Twitter data. Some of these methods are using the Twitter API and some are methods that do not need the API. Tweepy, TwitterAPI and Orange frameworks are scraped with Twitter API information. But Twint, GetOldTweets3, Selenium, BeautifulSoup, twitter-scrapper and snscrape frameworks allow scraping without any API knowledge. These methods are all free and most of them are python libraries. Among these methods, snscrape is the best method to pull both specific date ranges and unlimited data. The snscrape interface is also very easy to use.Therefore, in the study, data was collected via twitter with snscrape, a library of the Python programming language.

TABLE I. DATASET DESCRIPTION

Colum Name	Description
Datetime	The date the tweet was posted
Tweet Id	Unique tweet identifier
Text	The content of the tweet as text
Username	The person who shared it
Language	The language in which the sharing is made

Within the scope of this study, tweets with metaverse keywords were collected from Twitter. These posts were pulled using the snscrape library without the need for the Twitter API. The tweets were exported to files with the ".csv" extension. Exporting over csv is because csv is easy to access and manage with Python. Tweets were collected in two stages. The first phase is between 2021-10-21 and 2021-10-28, and the second part is between 2021-10-28 and 2021-11-04, each of which is a one-week process. Facebook CEO made statements about the metaverse on 2021-10-28. The first stage covers one week before this date. The second stage covers a one-week period, including 2021-10-28 and the following 6 days. In the first stage, 84,803 tweets were collected. In the second stage, 350,988 tweets were collected. In total, two different data sets were collected, one before and one after. At the same time, there are 5 columns on tweets collected with snscrape as we can see in Table 1. These columns are: 'Datetime', 'Tweet Id', 'Text', 'Username' and 'Language'.

B. Pre-processing Tasks

Tweet messages contain some parts that should not be used in our analyses, which we can call noise in the text data. It is an important step to clear the data to be analysed from noise. After this stage, the texts can be analysed in a way that gives results with higher success.

In this section, data pre-processing and manipulation steps will be explained. First, the process of removing repetitive tweets is performed. Afterwards, the datetime column was converted to a just day date format. We also removed non-English tweets because we only aimed to analyse tweets in English in our study. Finally, we remove all columns except Text. These steps will include the following which will be described in the following paragraphs:

- Remove duplicate tweets
- Convert datetime column as a just day date format
- Remove non-English tweets
- Remove all columns except text

Tweets with the same content cause biased analysis results. Therefore, it is necessary to clean the content of repetitive tweets. For this, the repetitive tweets will be deleted according to the "Text" column in the dataset. A total of 84,803 tweets were sent between 2021-10-21 and 2021-10-28. The number of tweets sent after deduplication was determined as 82,602. In other words, 2,201 tweets with the same content were deleted from the dataset. At the same time, although the total number of tweets sent on 2021-10-28 and 2021-11-04 was 350,988, this number was determined as 345,086 after deduplication. In other words, 5,902 tweet lines were deleted from the dataset.

The format of the "Datetime" column in the dataset must be edited to analyse tweets based on time. The date format in the dataset is yyyy-mm-ddThh:mm:ss.s+zzzzzz. This format has been converted to yyyy-mm-dd. These transformations were applied for two different datasets.

After deduplication, there are 82,602 tweets in the first dataset and 345,086 tweets in the second dataset. There is a "Language" column for these datasets. In the first dataset, there are tweets written in 48 different languages, one of which is undefined. In the second dataset, there are tweets written in 64 different languages, one of which is undefined. Tweets, where the "Language" column equals "en" for both datasets, will be used for analysis. In other words, only tweets written in English will be reviewed. In the first dataset, the number of tweets in English is 67,024. In other words, 15,578 tweets are tweets posted outside of English. In the second dataset, there are 262,561 tweets in English. In other words, 82,525 tweets are tweets written outside of English. As a result, in the first dataset, a total of 67,024 and in the second dataset, 262,561 rows of tweets will be analysed.

The datasets consist of 5 different columns. Unnecessary columns will be deleted from the data set after data analysis is done. What is needed for sentiment analysis is the content of the tweets. In other words, all tweets in the dataset will be deleted except for the "Text" column. In this way, we can now also do text analysis. Both datasets have only a "Text" column. In the last case, there are 67,024 rows and 1 column in the first data set. The second dataset has 262,561 rows and 1 column.

After the necessary data manipulation operations are performed on the obtained data set, the obtained data becomes text data. Many social media data consists of unstructured text data. Unstructured text data contains non-significant expressions. This "dirty" data needs to be cleaned for NLP work to be done effectively. In addition, non-English characters should be cleaned in tweet texts to carry out NLP work on English texts. At the same time, the words must go through the normalisation process to carry out the emotion analysis work correctly. These steps will include the following which will be described in the following paragraphs:

- Noise removal in Tweet Texts
- Normalization
- Tokenization
- Stemming
- Lemmatization
- Remove Stopwords

There are many different symbols, emojis and characters among the tweets. Among these, words and characters containing numbers, punctuation, special characters, emojis, link addresses (http), tagging (#) and mention (@) have been cleaned. Also, non-English characters have been removed. Since sentiment analysis work is a work based on words, these need to be cleared. After these operations applied on the "Text" column, a new column named "clean_text" was added to the data set.

Before starting the analysis, it is necessary to perform word-based normalization processes. These normalization stages will be applied so that words with the same meaning and spelling are not perceived differently in the sentence. This process allows us to perform a standard and acceptable analysis. These operations are divided into 4 parts, respectively, Tokenization, Stemming, Lemataziation and

Remove stopwords. The results of each of these stages are shown in Fig 2. Natural Language Toolkit (NLTK) will be used to do all these steps quickly. The most used library in the natural language processing stages of the Python programming language is NLTK. On the other hand, one of the most important normalization operations is the conversion of letters to lowercase. This process was applied after the Tokenization step.

Tokenization is one of the first text normalization operations to be implemented. The purpose of this step is to divide the text or paragraphs into smaller sections. In this way, more accurate transactions and analyzes are made. Tokenization can be used in two different ways, either word-based or sentence based. In this study, a word-based tokenization process will be applied. Tokenization has been applied to the texts in the "clean_text" column of the dataset. The words in the text for each line formed the new "tokenized" column as a list.

Stemming is applied to remove the inflections (prefix or suffix) of words. Words that have the same meaning and spelling are evaluated as different words by taking a prefix or a suffix. Stemming process is used to prevent this. After the tokenization process, the words kept as a list will be converted into root words. The results of this operation are added to the data set as the "stemmed" column.

Lemmatization is appliedas stemming can sometimes fail to find root words. Lemmatization, which considers the morphological analysis of words and appropriately separates the meaningful word into its roots, can be used as an alternative. Lemmatization is a very important method to find the smallest root form of a word. In this way, each word becomes able to represent itself. The lemmatization process has been applied to the "stemmed" column and added as a new column to the dataset as "lemmatized".

Remove stopwordsstep is applied afterwards. While creating sentences, words that do not mean anything in terms of emotions and meanings are used. These words are the most common words in a language (such as “the”, “a”, “in”), which are usually helpful in sentence construction. These words are called stopwords. There are different stopwords for each language. In this study, English stopwords are discussed. Removing these words from the sentence will not have any effect in terms of sentiment analysis in the sentence. For this reason, stopwords in the "lemmatized" column with root words will be removed.

The new texts obtained after all these text preprocessing stages were added to the data set as the "processed_text" column. The results of each of these stages are shown in Fig 2.

Text	clean_text	tokenized	stemmed	lemmatized	processed_text
@WuBlockchain #BlockchainNative #Metaverse #NFTgaming Because they don't learn the second law of hacker metaverse. @FoundrunS @hackernoon https://t.co/IgePwHEeSX	Because they dont learn second hacker metaverse	[Because, they, dont, learn, second, hacker, metaverse]	[becaus, they, dont, learn, second, hacker, metavers]	[becaus, they, dont, learn, second, hacker, metavers]	becaus dont learn second hacker metavers
@BrianDEvans Hey Brian with the @SVSNFT x #888innercircle collab just announced for Halloween, $blood token utility, SVS video game, metaverse @WilderWorld, and gen 2 around the corner… I'd say grab a vampire! Don't forget to stake it for $blood 🩸🧛🩸 https://t.co/Jh1f24GQr6	Brian with collab just announced Halloween blood token utility video game metaverse around corner grab vampire Dont forget stake blood	[Brian, with, collab, just, announced, Halloween, blood, token, utility, video, game, metaverse, around, corner, grab, vampire, Dont, forget, stake, blood]	[brian, with, collab, just, announc, halloween, blood, token, util, video, game, metavers, around, corner, grab, vampir, dont, forget, stake, blood]	[brian, with, collab, just, announc, halloween, blood, token, util, video, game, metavers, around, corner, grab, vampir, dont, forget, stake, blood]	brian collab announc halloween blood token util video game metavers around corner grab vampir dont forget stake blood

Fig 2. Text preprocessing step results in the dataset

In addition, tweets with a length of more than 3 texts in the "processed_text" column obtained were discussed. This process has taken place because texts of less than 3 letters or characters do not provide any emotion or meaning. These operations were made according to the texts in the "processed_text" column. As explained above, 67,024 rows were found in the first data set and 262,561 rows in the second data set. In addition, row data with a text length less than 3 was removed from both datasets. Accordingly, 8,598 rows in the first data set and 38,616 rows in the second data set were deleted. As a result, there are 58,426 rows in the first data set and 223,945 rows in the second data set.

C. Sentiment Analysis

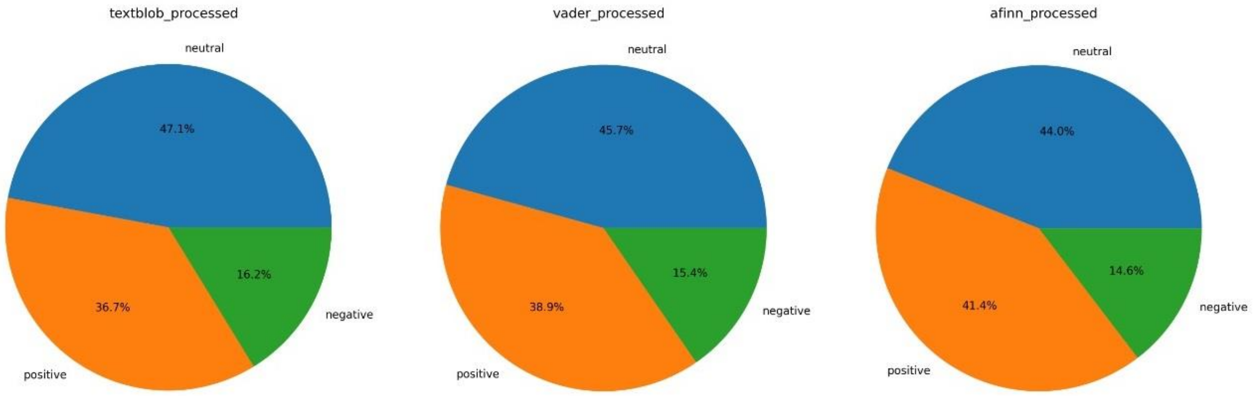
There are many methods for sentiment analysis in English texts. Among these methods, unsupervised sentiment analysis methods such as Bert, IBM Watson, Textblob, and Vader are used for sentiment analysis in English with rule-based and dictionary-based Afinn libraries. Bert, TextBlob, and Vader are open source and free, while IBM Watson is a paid library. IBM Watson provides API for understanding natural language and performing sentiment analysis. However, it has not been used because it has a limited and paid use opportunity. . In addition, sentiment analysis work with the deep learning NLP Bert model developed by google was not performed in this study due to slower processing power. As a result, sentiment analysis will be performed with Textblob and VADER, which are unsupervised learning models. Apart from these, sentiment analysis will be done with the dictionary based Afinn library. They are models that have a high accuracy rate, especially for social media data. For preprocessed tweets, emotion classification will be made with the three models above. In our study, we used Textblob, Vader and Afinn methods for sentiment analysis. We will share the details of these methods under this heading.

Textblob is a completely free and open-source Python library. In addition, Textblob is an NLP package that contains many functions such as classification, translation, parsing and sentiment analysis on texts. It will be used as sentiment analysis in this study. It has a rule-based structure for sentiment analysis. Using the NLTK base, it uses a corpus where emotion values were previously registered. It makes sentiment analysis based on the frequency of the words used. The sum of the polarity values of each word gives the sentiment analysis result of the sentence. Sentiment analysis processes of each sentence sent to Texblob are classified according to the estimated polarity value. Polarity consists of decimal numbers between [-1,1]. -1 indicates negative emotion and +1 indicates positive emotion. It will label the tweet as neutral if the sentiment score, that is, polarity value is equal to 0, negative if it is less than 0, and positive if it is greater than 0. The emotion tag of each tweet was determined with Texblob separately for both data sets. This application was applied on pre-processed tweet rows.

Vader (Valence Aware Dictionary and Sentiment Reasoner) is an open-source and free library maintained under the MIT license. It is also a dictionary and rule-based sentiment analysis tool specifically tailored to the emotions expressed on social media. That's why Vader is optimized for social media data. Also, it is a very important analysis tool for NLP. It is available in the NLTK package and can be applied directly to unlabeled text data. The emotional intensity score is in 4 different ways: ('neg', 'neu', 'pos', and 'compound'). The negative, positive and neutral percentiles of a sentence are shown in the labels 'neg', 'neu' and 'pos'. The sum of these three possibilities is 1. In emotion classification, the feature with the highest probability is accepted. However, classification was made based on "compound", that is, compound density scale. Based on this score, emotion classification was determined in three different ways. The 'compound' value is labelled as negative if less than 0, neutral if equal to 0, and positive if greater than =0. In this way, Vader emotion labelling was performed for both data sets.

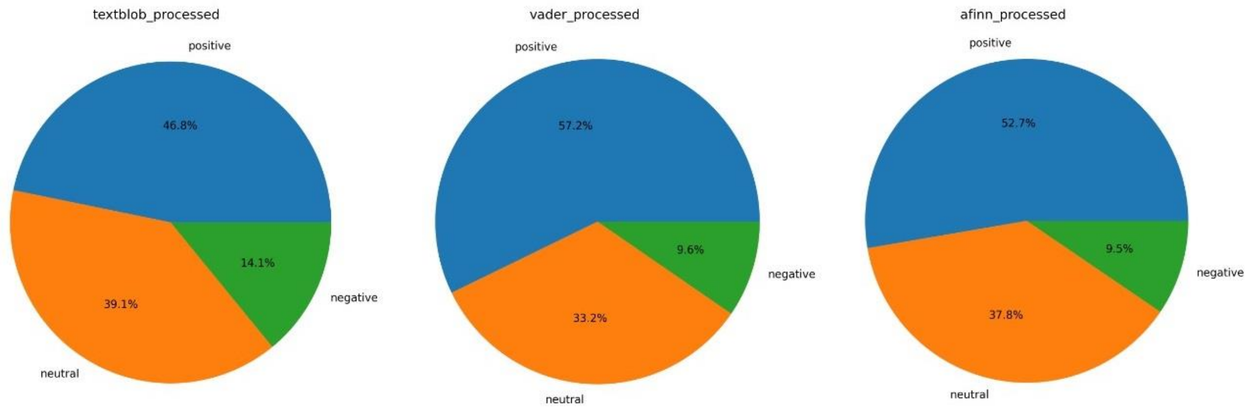
Afinn is one of the simplest and most popular dictionaries that can be used extensively for sentiment analysis. This dictionary is completely free and open source. Contains more than 3,300 words with a polarity score associated with each word in the texts. In the emotion definition of a tweet, words not included in the Afinn word list are assumed to have a "0" emotion score. Afinn is a list of words rated as a score with an integer from -5 (negative) to +5 (positive). The score values of the words in the sentence are found by adding the values of the words in the Afinn dictionary. As a result of this sum, classification is made according to the triple ranking scale. If this summed result is less than 0, the sentence is labelled negative, equal to 0, neutral, and greater than 0, the sentence is labelled as positive. These processes were applied to all pre-processed tweet rows in both datasets.

IV. Discussion



***Cite:** Agrali, Ö., Aydin, Ö. (2021).* Tweets Classification and Sentiment Analysis on Metaverse Related Messages. Journal of Metaverse. 1(1). pp.25-30.

Fig 3. Results of the sentiment analysis for the tweets posted before Mark's speech



***Cite:** Agrali, Ö., Aydin, Ö. (2021).* Tweets Classification and Sentiment Analysis on Metaverse Related Messages. Journal of Metaverse. 1(1). pp.25-30.

Fig 4. Results of the sentiment analysis for the tweets posted after Mark's speech

Results of the three different sentiment analysis model for the tweets posted before Mark Zuckerberg's speech about change Facebook name to Meta can be seen in Fig. 3 and also results for the tweets posted after it can be seen in Fig. 4.

When the tweets sent in the one-week period before Mark's speech on the change of his company's name on October 28, 2021 were analysed with 3 different models, the Textblob results were 46.8% positive, 14.1% negative and

39.1% neutral. In the analysis made with the Vader method, 57.2% positive, 9.6% negative and 33.2% neutral analysis results were obtained. Results were obtained as 52.7% positive, 9.5% negative and 37.8% neutral with the Afinn method. Our analysis of the tweets sent the week after Mark's statement yielded the following results. With the Textblob method, 47.1% were neutral, 36.7% positive, 16.2% negative. In the analysis made by Vader method, 45.7% neutral, 38.9% positive and 15.4% negative results were obtained. With the Afinn method, 41.4% positive, 14.6% negative and 44% neutral results were obtained. You can see these results comparatively in Table 2.

TABLE II. SENTIMENT ANALYSIS RESULTS

	A week before Mark's Speech			**A week after Mark's Speech**		
	Textblob	Vader	Afinn	Textblob	Vader	Afinn
Positive (%)	46.8	57.2	52.7	36.7	38.9	41.4
Negative (%)	14.1	9.6	9.5	16.2	15.4	14.6
Neutral (%)	39.1	33.2	37.8	47.1	45.7	44

When the results in Table 2 are examined, it is seen that the tweets are generally positive in the analysis made with 3 methods before the Mark's speech. After Mark's speech, positive tweet rates are decreasing and both neutral and negative tweet rates are increasing. The main reason for this may be that the topic became more popular after Mark's speech and negative-minded people who did not tweet about it started tweeting.

V. CONCLUSION

Sentiment analysis is a method that is used in many areas and provides effective results. With this method, data from different data sources are processed and sentiment analysis is attempted. For these operations, preliminary operations are performed on the dataset and then inferences are made on the data that is ready to be processed.

Today, an old topic can become popular again or a popular topic can lose its importance. The concept of metaverse is also old, but it is a concept that has become popular again today. It gained a serious momentum especially when Facebook CEO Mark Zuckerberg changed the name of his company to Meta. Of course, many different companies are also working in this field. In this study, we examined this newly popular concept with sentiment analysis. In particular, tweets posted on Twitter before and after Mark Zuckerberg's speech were evaluated. The obtained data are shared in the article comparatively.

In the future, the relevant dataset can be expanded to cover wider time periods so that the findings can be examined. It would be appropriate to realise this with a more convenient study in terms of time and resources.

REFERENCES

[1] Pang, B., & Lee, L.J. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval, 2*(1-2), 1-135. Doi: 10.1561/1500000011

[2] Pang, B., & Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. arXiv preprint cs/0409058

[3] Read, J. (2005, June). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. *In Proceedings of the ACL student research workshop* (pp. 43-48).

[4] Go, A., Huang, L., & Bhayani, R. (2009). Twitter sentiment analysis. *Entropy, 17*, 252.

[5] Çoban, Ö., & Tümüklü-Özyer, G. (2018). Twitter duygu analizinde terim ağırlıklandırma yönteminin etkisi. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 24*(2), 283-291. doi: 10.5505/pajes.2017.50480

[6] Sarlan, A., Nadam, C., & Basri, S. (2014, November). Twitter sentiment analysis. *In Proceedings of the 6th International conference on Information Technology and Multimedia* (pp. 212-216). IEEE.

[7] Karayiğit, H., Çiğdem, A. C. I., & Akdağlı, A. (2018). A Review of Turkish Sentiment Analysis and Opinion Mining. *Balkan Journal of Electrical and Computer Engineering, 6*(2), 94-98. DOI: 10.17694/bajece.419547

[8] Abalı, G., Karaarslan, E., Hürriyetoğlu, A., & Dalkılıç, F. (2018, April). Detecting citizen problems and their locations using twitter data. *In 2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)* (pp. 30-33). IEEE.

[9] Ayan, B., Kuyumcu, B., & Ciylan, B. (2019). Detection of Islamophobic Tweets on Twitter Using Sentiment Analysis. *Gazi Üniversitesi Fen Bilimleri Dergisi, 7*(2), 495-502. DOI:10.29109/gujsc.561806

[10] Fadel, I.A., & Öz; C. (2020), A Sentiment Analysis Model for Terrorist Attacks Reviews on Twitter. *Sakarya University Journal of Science, 24*(6), 1294-1302, DOI: 10.16984/saufenbilder.711612

[11] Demirtas, E., & Pechenizkiy, M. (2013, August). Cross-lingual polarity detection with machine translation. *In Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining* (pp. 1-8). DOI:10.1145/2502069.2502078

[12] Alqaraleh, S. (2020). Turkish Sentiment Analysis System via Ensemble Learning. *Avrupa Bilim ve Teknoloji Dergisi*, 122-129. DOI:10.31590/ejosat.779181

[13] Kemaloğlu, N., Küçüksille, E., & Özgünsür; M.E. (2021), Turkish Sentiment Analysis on Social Media. *Sakarya University Journal of Science, 25*(3), 629-638, DOI: https://doi.org/10.16984/saufenbilder.872227

[14] Ergül Aydın, Z., Kamışlı Özturk, Z., & Erzurum Ciçek, Z. İ. (2021) Turkish Sentiment Analysis for Open and Dıstance Educatıon Systems. *Turkish Online Journal of Distance Education, 22*(3) , 124-138 . DOI: 10.17718/tojde.961825

[15] Dalkilic, F., & Cam, A. (2021). Automatic Movie Rating by Using Twitter Sentiment Analysis and Monitoring Tool. *Journal of Emerging Computer Technologies, 1*(2), 55-60.

[16] Tokcaer, S. (2021). Türkçe Metinlerde Duygu Analizi. *Yaşar Üniversitesi E-Dergisi, 16*(63), 1516-1536. DOI:10.19168/jyasar.928843