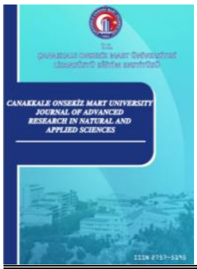
e-ISSN 2757-5195



Çanakkale Onsekiz Mart University

Journal of Advanced Research in Natural and Applied Sciences

Open Access

doi.org/10.28979/jarnas.1087772 2023, Vol. 9, Issue 1, Pages: 107-126 dergipark.org.tr/tr/pub/jarnas

Which Emotions of Social Media Users Lead to Dissemination of Fake News: Sentiment Analysis Towards Covid-19 Vaccine

Maide Feyza Er[1], Yonca Bayrakdar Yılmaz[2,*]

[1]Department of Software Engineering, Faculty of Engineering and Natural Sciences, Bandırma Onyedi Eylül University, Bandırma, Türkiye
[2]Department of Computer Engineering, Faculty of Engineering, Çanakkale Onsekiz Mart University, Çanakkale, Türkiye

Article History

Received: 15.03.2022

Accepted: 10.10.2022

Published: 05.03.2023

Research Article

Abstract − The use of social media as a news source is quite common today. However, the fact that the news encountered on social media are accepted as true without questioning or checking their validity is one of the main reasons for the dissemination of fake news. For the social media ecosystem, the question arises as to which emotion is more effective in spreading fake news, as the accuracy and validity of the news are under the control of opinions and emotions rather than evidence-based data. From this point of view, our study investigates whether there is a relationship between users' reaction to the news and the prevalence of the news. In our study, sentiment analysis was conducted on the reactions of Twitter users to fake news about the COVID-19 vaccine between December 31, 2019 and July 30, 2022. To fully assess whether there is a relationship between the reactions and the prevalence of the news, the spread of real news published in the same period in addition to fake news is also taken into consideration. Fake and real news comments, which were selected in different degrees of prevalence from the most to the least, were examined comparatively. In the study, where text mining techniques were used for text pre-processing, analysis was carried out with NLP techniques. In 83% of the fake news datasets and 91% of the overall news datasets considered in the study, negative emotion was more dominant than other emotions, and it was observed that as negative comments increased, fake news spread more as well as real news. While neutral comments have no effect on prevalence, users who comment on fake news for fun significantly increase the prevalence. Finally, to reveal bot activity NLP techniques were applied.

Keywords *− Bot detection, fake news, nlp, sentiment analysis, text mining*

1. Introduction

Today, it is very common that society uses social media as a news source. In 2021, while the rate of getting news from social media is over 40% in European countries, it is 42% in America and above 50% in Asian countries (Newman et al., 2021). This situation has various negative effects such as accepting the news that encountered as true without questioning or checking the validity of them. While distorting existing news or producing non-existent news, supporting it with elements such as text, audio, images or video, and the fact that the news is extremely difficult to confirm, turns social media into an ecosystem that rapidly spreads fake news, users also play a major role in the spread of fake news by liking and re-sharing the news. In order to see how serious the effects of fake news have reached, it is enough to look at the effects of fake news and conspiracy theories on people, about the so-called harms and side effects of the COVID-19 vaccine in the past period. The decrease in the rate of vaccination with vaccine opposition and vaccine indecision is attributed to exposure to misinformation about the COVID-19 vaccine on social media (Islam et al., 2021).

[1] mer@bandirma.edu.tr
[2] yonca@comu.edu.tr
[*]Corresponding Author

For the social media ecosystem, the question arises as to which emotion is more effective in spreading fake news, as the accuracy and validity of the news are under the control of opinions and emotions rather than evidence-based data. With this viewpoint, our study investigates whether there is a relationship between users' reaction to the news and the prevalence of the news. "When users encounter fake news, what type of emotion drives them to spread that fake news?" or in other words, "Is there a relationship between the reaction to fake news and the prevalence of the news?" are the research questions of our study. To fully reveal whether there is a relationship between users' reactions to the news and the prevalence of the news, the spread of real news is also discussed in addition to fake news. The emotional impact of real news, as well as fake news, on people has been examined and it has been evaluated whether there is a similarity between fake news and the emotion that spreads the real news. In this study, which is unique in its own right with its research questions and content, the comments on the Twitter platform, which is known as the first source of the news and the first place to go to learn the agenda, were used. A sentiment analysis was conducted on the reactions of Twitter users to fake news about the COVID-19 vaccine between December 31, 2019, and July 30, 2022. Fake and real news comments selected in different degrees of prevalence from very common to less common were examined comparatively.

Sentiment analysis is considered as a classification problem in the literature. Go, Bhayani & Huang (2009), who applied sentiment analysis to tweets for the first time, used sentiment expressions in tweets to classify sentiment, and classified tweets as positive or negative. Pak & Paroubek (2010), similar to the study of Go et al. (2009), by looking at the emotion expressions at the end of the tweets, handled the problem as a multi-class classification task and classified the tweets into three classes as positive, negative and neutral. It is important to determine which features are more effective in revealing the emotion in the text. To this end, a number of researchers (Agarwal et al., 2011; Hamdan, Bechet & Bellot, 2013; Kouloumpis, Wilson & Moore, 2011; Saif, He & Alani, 2012) examined the effect of different features on the Twitter sentiment analysis problem. Saif, He & Alani (2012) presented a novel approach to add semantic features in addition to the features in the training set. Neethu & Rajasree (2013) suggested feature extraction after applying two-step pre-processing to the text to process typos and slang words in their study. Another important study belongs to Hutto & Gilbert (2014). In their study, they recommended VADER as a lexicon and rule-based sentiment analysis tool. Thanks to the ability of easily accessing the thoughts of many users from all over the world on Twitter, sentiment analysis on Twitter has recently become a subject of study in the social and psychological fields. Dzogang, Lightman & Cristianini (2018), who conducted one of such studies, measured the daily variation of 73 psychometric variables in Tweets they analyzed from the UK, and found two main factors, "Categorical Thinking" and "Existential Thinking", which peaked at opposite time points throughout the 24-hour day. Another study was conducted by Snefjella, Schmidtke & Kuperman, (2018). They measured the emotional variation of tweets between Canada and the United States, with the aim of measuring the use of certain types of language across different cultures.

The increase in the spread and social impact of fake news as a result of the popularization of social networks has led to a growing interest in examining fake news on social networks. One of the key tricks that fake news producers use to support the success of fake news is to excite the emotions of the readers. Therefore, sentiment analysis turned out to be a very useful method for examining fake news when applied to both news items and related information such as user comments (Alonso et al., 2021). In this context, Del Vicario et al. (2019) provided a framework for early warning of potential misinformation targets in social media. In their study, they fed several classic machine learning classifiers with text features and emotion-based features of user behavior to detect fake news. The best performance was obtained with a logistic regression classifier. In another study, the authors investigated whether the hidden sentiments found in user comments on social media help distinguish fake news from credible content. Experimental results have shown that the component that contributes the most to the performance of the system is sentiment analysis (Cui, Wang & Lee, 2019). Anoop, Deepak & Lajish (2020) worked on the detection of fake news in the field of health. In their study, they looked at different types of emotional traits displayed in fake and real health news to detect fake news. They extracted emotion features from a lexicon to feed classical and deep learning classifiers and stated that the performance of all classifiers increased with emotion information.

Zhang et al. (2021) noted that most of the existing studies on fake news, even for fake news that go viral, focus on the emotional signals of the content conveyed by publishers rather than the emotions the news evokes in the crowd. Therefore, in their study, they investigated whether the emotions and content in news comments are useful for fake news detection. They stated that their proposed feature set could be well compatible with existing fake news detectors and effectively improve fake news detection performance. Iwendi et al. (2022) proposed a model in which they use 39 features from multimedia texts to detect fake news about COVID-19. They used information fusion in their work to extract social media data, where they applied cutting-edge deep learning models such as Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Recurrent Neural Network (RNN). RNN achieved a high precision of 85% for the model they used, while using GRU model for fake news, the best F1-Measure and recall were 83%. Another study examining the spread of fake news is by Bodaghi & Oliveira (2022). In their study, which examined users in different roles during the spread of fake news, they analyzed the real and fake news spread on Twitter. For this purpose, they analyzed 5 fake news collected from Twitter, which includes 8 million nodes and 28 million links in total, using graphs. With their work, they aimed to provide a better understanding of the nature of fake news spread on social media (Bodaghi & Oliveira, 2022). Studies in the literature used sentiment analysis to detect fake news and suggested different methods to examine its spread. There is no study we came across which deals with the spread of fake news using sentiment analysis.

In addition to being a research topic that has not been studied before in the literature, our study, which will shed light on the nature of fake news that causes its rapid spread, will provide a deeper understanding of human psychology and behavior, and raise awareness of individuals on how they can recognize when they come across a piece of news, whether it is fake news or not, before they spread it. The rest of the study organizes as follows. In Section 2 data collection, normalization of the text and analysis processes are explained. In Section 3, the findings are given and the results are discussed.

2. Materials and Methods

In this section, the steps performed in the study are given. The data and analysis methods of the study, which was carried out with three basic processes, namely data collection, text normalization and text analysis, are explained in this section. The system architecture proposed in our study is given in Figure 1.

2.1. Data Collecting

Fake and real news about the COVID-19 vaccine used in our study were selected from news verification sites such as Snopes (2018), FactCheck.org (2008) and Check Your Fact (2019), which are independent confirmatory. The fake and real news headlines we collected in our study and the number of comments made on these news are given in Table 1 and Table 2, respectively. The prevalence was determined by the number of comments made on fake/real news.

It is known in the literature that most of the researchers use Twitter APIs to collect tweets (Giachanou & Crestani, 2016). However, since Twitter's free access policy, prices, and different access options change over time, there are limitations on the number of requests per minute, tweets per request, and access to historical data in Twitter API (Antonakaki, Fragopoulou & Ioannidis, 2021). To overcome these limitations, Twint was used as a tweet collection tool in our study. Although Twint is open-source software, it does not rely on Twitter's API and allows unlimited tweet downloads (Zacharias & Poldi, 2018).

Table 1
Fake news circulated on social media about the COVID-19 vaccine between December 31, 2019 and July 30, 2022.

Fake news headlines	Number of comments
COVID-19 vaccine impairs fertility	62722
COVID-19 vaccine can monitor the human	56052
COVID-19 vaccine will reduce the planet's population	30792
COVID-19 vaccine can change human DNA	21073
COVID-19 vaccine contains aborted fetal cells	13164
COVID-19 vaccine makes our body magnetic	11206
Thousands of Americans died after receiving their COVID-19 vaccine	5435
Bill Gates admits that a COVID-19 vaccine could kill up to 700K people	3145

Table 2
Real news circulated on social media about the COVID-19 and COVID-19 vaccine between December 31, 2019 and July 30, 2022

Real news headlines	Number of comments
US Deaths from COVID Hit 1 Million, Less Than 2 1/2 Years in	51783
Biden Say "You Won't Get COVID if You're Vaccinated"	31673
US Opens COVID-19 Vaccine to Little Kids; Shots Begin This Week	17732
COVID-19 Shots Might Be Tweaked if Variants Get Worse	14888
FDA Advisers Recommend Updating COVID Booster Shots for Fall	14579
Pfizer to Discuss COVID-19 Vaccine Booster with US Officials	8717
Dr. Anthony Fauci Said US Is 'Out of the Pandemic Phase'	2565
Masks Could Return As COVID Surges Nationwide	2053

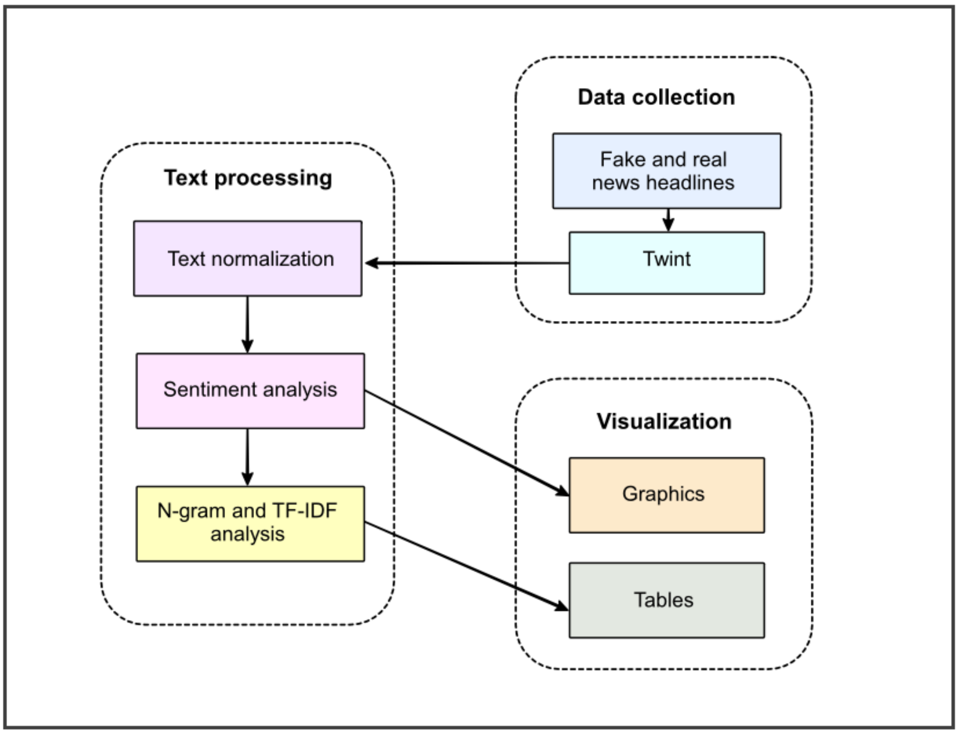


Figure 1. The proposed system architecture.

Comments were collected by giving parameters, such as search keywords, the display of replies to tweets and the first time the fake news was seen to CLI (Command Line Interface) commands specific to Twint. For example, a false news title such as “COVID-19 Vaccine Impairs Fertility” was questioned with the keywords “covid vaccine impact fertility”, “covid vaccine cause infertility” and “covid vaccine cause miscarriage”.

2.2. Text Normalizing

Tweets collected with Twint; along with the body of the tweet, contain the username, mentions, tweet ID, the date and time of the tweet, and the UTC (Coordinated Universal Time) time difference identifier. In addition to the noise brought by the data collection tool, the informal language and style used in social media also necessitate the text normalization process. The aggregated data, which appears extremely complex even with human reading, are given in the “Raw Text” column of Table 3. In this context, in our study, the data were subjected to a detailed text cleaning and normalization process. For this, text mining techniques and NLP tools were used for text normalizing. The workflow followed for text cleaning is shown in Figure 2.

Examples of raw data and the text after normalization are given in Table 3. Since the characters, numbers, date and time information belonging to Twitter are not related to the analysis we applied to the text, it was appropriate to delete them from the data set. Since the sentiment analysis tool we applied to the data is lexicon-based, “happy 😀” and “happy😀” are perceived as two separate words. To overcome this challenge, emoticons, emojis and other characters are separated from the words they are attached to. Since emojis are the most contributing component in revealing the emotion in the text, emojis have been replaced with their textual counterparts in the data set. For example, the “😀” emoji has been replaced by the phrase “smiling_face_with_smiling_eyes”. This transformation was achieved by using the open-source emoji module (Kim & Wurster, 2015).

It is very common to use abbreviations because of the character limit of Twitter. In order to capture the emotion that abbreviations add to the text, the data sets were subjected to a three-stage process for the full forms of abbreviations. Respectively, the full forms of official abbreviations used in English, unofficial abbreviations used in daily speech, and internet slang abbreviations were carried out. Abbreviations are created manually and kept in dictionary data structure type. Examples of abbreviation dictionaries are given in Table 3.

Table 3
Three randomly selected tweets from the “COVID-19 Vaccine Contains Aborted Fetal Cells” dataset. The first version after being collected from Twitter with Twint is given in the “Raw Text” column, the text after applying normalization is given under the “Normalized Text” column. Usernames and mentions have been replaced with “username” in accordance with the personal data protection law

Raw Text	Normalized Text
1389044057830273027 2021-05-03 05:28:05 +0300 <username> @username Covid vaccine contains chimpanzee bodyparts and aborted fetuses to steralise and kill society. https://t.co/tf1SgL4Rg5	covid vaccine contains chimpanzee bodyparts and aborted fetuses to steralise and kill society
1387590968019353600 2021-04-29 05:14:01 +0300 <username> A colleague at my school—a teacher-actually said today aborted fetuses are in the Covid vaccine...I'm done... 🤦🏼‍♀️🤦🏼‍♀️	a colleague at my school a teacher actually said today aborted fetuses are in the covid vaccine i am done person facepalming light skin tone female sign person facepalming light skin tone female sign
1377049899787825152 2021-03-31 03:07:35 +0300 <username> My mother won't get the Covid vaccine because she says it contains aborted fetuses. Can some of you medically knowledgeable people please help set her straight? I'm so tired of the bullshit and she might listen if it comes from someone else. She's vulnerable! 🥺 #CovidVaccine	my mother will not get the covid vaccine because she says it contains aborted fetuses can some of you medically knowledgeable people please help set her straight i am so tired of the bullshit and she might listen if it comes from someone else she is vulnerable pleading face covid vaccine

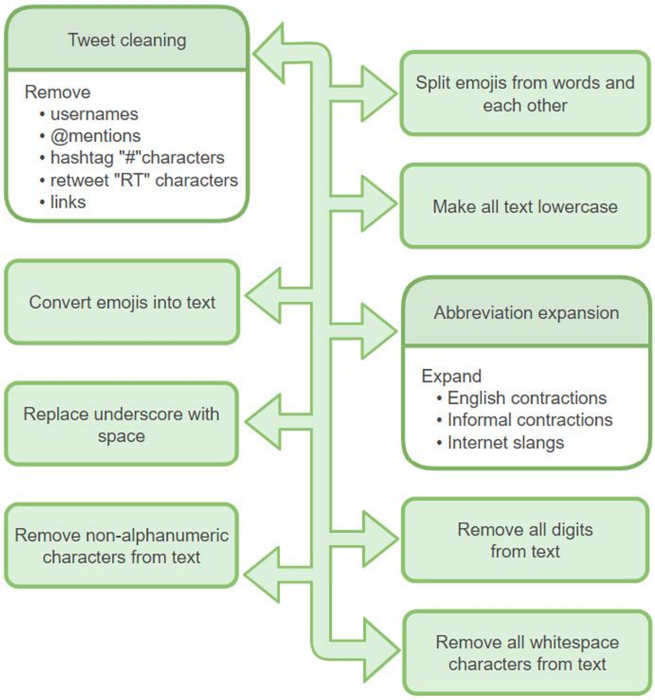


Figure 2. Workflow for text normalization

Table 4
Respectively; Sample abbreviations from English Contractions (2022), Informal Contractions List in English (2020), and Internet Slang Terms (2021) dictionaries.

English Contractions		Informal Abbreviations		Internet Slang	
Abbreviation	Full form	Abbreviation	Full form	Abbreviation	Full form
‘m	am	c’mon	come on	csl	can’t stop laughing
‘r	are	dija	did you	lol	laughing out loud
‘s	is	sorta	sort of	jk	just kidding
‘ve	have	dunno	don’t know	smh	shaking my head
‘ll	will	lemme	let me	qq	crying
‘d	had	innit	isn’t it	w8	wait

After replacing the underscores between words with a space character, deleting all characters except numbers and alphanumeric characters, and removing the white space character, the text normalization process was completed. Checking if the text contains the specified search pattern and replacing it with the desired character was performed using Python’s re module.

2.3. Analyzing the Text

In our study, dictionary and rule-based open-source VADER sentiment analysis tool was used to classify the emotional polarity of the texts (Hutto & Gilbert, 2014). VADER includes a dictionary of emotions labelled by multiple independent human reviewers, which considers the average of these labels as the word’s “valence” score. While calculating the perceived emotional intensity in the text at sentence level, it applies the grammatical and syntactic rules explained in their study. VADER, which is sensitive to both the polarity and the intensity of the expressed emotions, can be applied for sentiment analysis in all areas in general, as well as adapting to the emotions expressed in social media (Hutto & Gilbert, 2014).

When sentiment analysis is applied to a sentence, the composite score, referred to as "compound", is calculated by summing the valence scores of each word in the dictionary, adjusting them according to the rules, and then normalizing them from -1 (most negative) to +1 (most positive). The "pos", "neu" and "neg" scores indicate the proportions of texts falling into each category and therefore their sum should be 1 (Hutto & Gilbert, 2014). Based on this context, in our study, standard thresholds were used, and the sentences were classified as positive, negative or neutral according to the composite score.

After applying the sentiment analysis separately for each data set, the rate of positive, negative, and neutral polarity in the comments emerged. N-grams are used to count how often consecutive word strings occur in the datasets. N-gram is basically a string of N words that occur together in a particular document. The n-gram analysis was performed using the N-grams module of the NLTK (Natural Language Toolkit) library in two different ways (Bird, Klein & Loper, 2009). In the first process, n-grams were produced without deleting the stop words. Stop words (such as the, and, but, if, as in English) do not have a meaning on their own and are the most commonly used words such as adverbs, prepositions and conjunctions in a language. English stop words corpus of the NLTK library was used in our study. In the second process, n-grams were produced after the stop words were deleted. This enabled us to catch important abstract word groups.

Finally, in our study, TF-IDF (term frequency-inverse document frequency) analysis of the data sets was made under their own news headlines. TF- IDF is obtained by multiplying TF and IDF values, and TF, that is, "term frequency", refers to the frequency of occurrence of the term in the document. It is obtained by dividing the number of occurrences of the term in the document by the total number of words in the document. IDF is the inverse document frequency and is the logarithm of the ratio of the total number of documents to the number of documents containing the relevant word. The equation of IDF is given in 2.1.

$$idf_t = \log \frac{N}{df_t} \tag{2.1}$$

In 2.1, N represents the number of documents, and df_t represents the number of documents containing the t term. From this point of view, if a word occurs frequently in other documents, the IDF value will decrease. In other words, the higher the IDF value, the more unique the word is. The equation for the TF-IDF obtained by combining the term frequency and the inverse document frequency to produce a combined weight for each term in each document is given in 2.2.

$$tf - idf_{t,d} = tf_{t,d} \times idf_t \tag{2.2}$$

In 2.2, $tf_{t,d}$ shows the frequency of the term t in the document d; idf_t is the inverse document frequency calculated in 2.1. The TF-IDF value will take the lowest value when the term occurs in almost all documents (Manning, Raghavan & Schutze, 2008). To measure the diversity and uniqueness of the comments in the data set, TF-IDF analysis was performed at the last stage. In Section 3, the sentiment classification and polarity rates of fake news were compared in ascending order, and the diversity and uniqueness of the comments in the data set were interpreted in the light of the most common n-grams and TF-IDF values.

3. Results and Discussion

After the comments on the fake and real news headlines given in Table 1 and Table 2 were collected, they were converted into a dataset under separate headings and sentiment analysis was applied by isolating them from each other. The titles of the data sets, the number of data they contain, the rates of negative, neutral and positive comments are given in Figure 3 and Figure 4 for fake news, and Figure 5 and Figure 6 for real news, in ascending order. Data sets from Set 1 to Set 8 represent fake news datasets, while those from Set 9 to Set 16 represent real news datasets.

In Figure 3 and Figure 4 in which the emotional polarity rates and data numbers of the datasets are given, it is seen that the rate of negative comments is always higher in all fake news datasets than the other emotional polarities (positive and neutral). For the fake news datasets, all of the eight datasets had a higher rate of negative polarity than the others. That is, the predominant rate of negative polarity is 100% for fake news. This shows that fake news - whether its prevalence is more or less - is largely fed by negative mood. However, when Figure 5 and Figure 6 are examined, we see that the same is not valid for real news data sets. Four of the eight datasets had a higher rate of positive polarity than the others. Thus, the predominant rate of positive polarity is 50%. This shows that there are more positive comments on real news than fake news. However, for real news, negative polarity is predominantly found at 50%, while neutral polarity is dominant at 0%. That is, neutral polarity is not dominant in any data set. In this respect, real news and fake news are similar to each other.

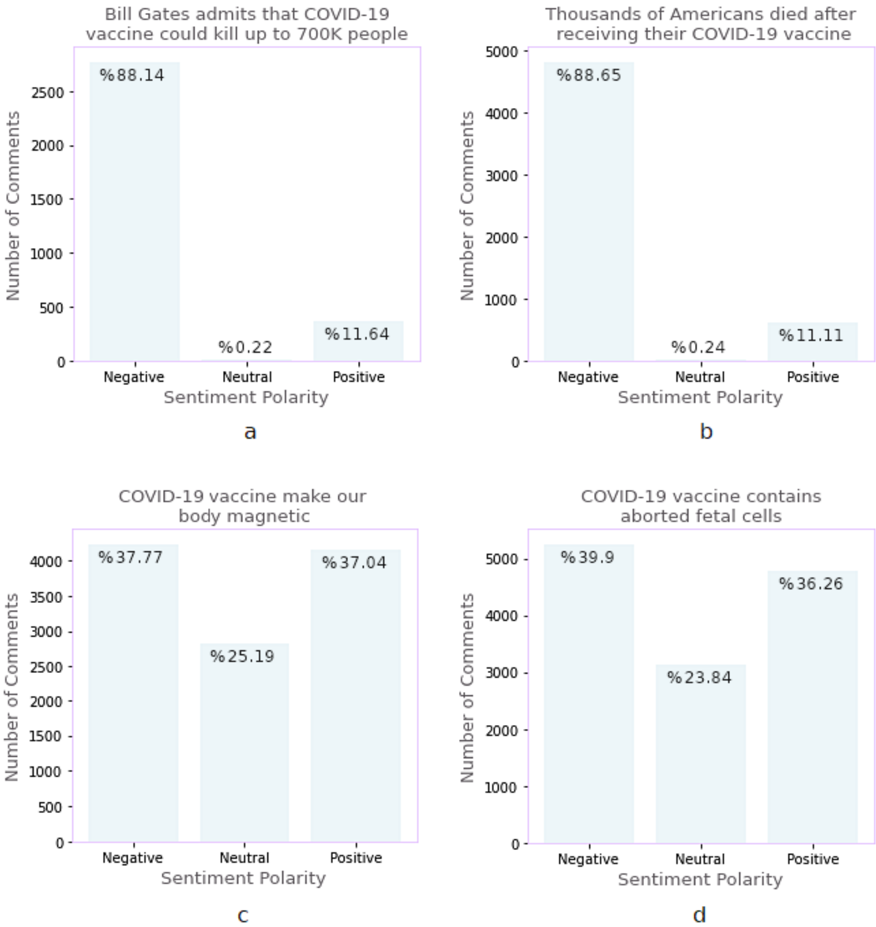


Figure 3. Polarity analysis of fake news datasets 1-4 (a) Set 1 “Bill Gates admits that a COVID-19 vaccine could kill up to 700K people”, number of data is 3145. (b) Set 2 “Thousands of Americans died after receiving their COVID-19 vaccine”, number of data is 5435. (c) Set 3 “COVID-19 vaccine makes our body magnetic”, number of data is 11206. (d) Set 4 “COVID-19 vaccine contains aborted fetal cells”, number of data is 13164.

However, except for Set 1 and Set 2, which had the lowest number of data, the interpretation rates in the other sets showed consistent changes. For this reason, it was appropriate that Set 1 and Set 2 were not included in the interpretation rate change graphs in Figure 7 in order not to prevent generalization, considering that they are insufficient to reflect the characteristics of the data due to the small number of data. However, while Set1 and Set2 are very similar to each other, displaying a completely different characteristic from the others is a situation that should be examined in particular. To understand the reason, the specific analyzes applied to Set 1 and Set 2 are explained later in the article.

When Set 1 and Set 2 are not included in the comment rate change graphs in Figure 7, the data set with the lowest data number of the fake news data sets is Set 3 with 11206 comments. Since the dataset with the least spread among fake news contains 11206 comments, it is necessary to provide equal conditions for real news. Therefore, datasets with less than 11206 comments (Set9, Set10, and Set11) from real news datasets are not included in the comment rate change graphs in Figure 7. Thus, the datasets with the lowest spread for fake and real news were brought to the same level.

When Figure 7 (a) and (b) are examined together, while the number of data increases in both, an increase in the rate of negative comments is observed. From this point of view, it can be said very clearly that as negative comments increase, the prevalence of news, whether it is fake news or real news, increases. In Figure 7(c), though not as much as the rate of change in Figure 7(a) a decrease in the rate of positive comments is observed. It is interesting, however, that the rates of positive comments range from 40% to 30% for all datasets given in the graph. In Figure 7(d), the decrease in the rate of positive comments is clearly seen. It is not possible to make any generalizations for Figure 7(e) and Figure 7(f) where the change in the neutral interpretation rate is given. Therefore, it turns out that neutral comments have no effect on prevalence. In 83% of the fake news datasets and 91% of the overall news datasets considered in the study, it is observed that the prevalence of the news increases as the negative emotional comments increase.

Another common element seen in Figure 7(a) and Figure 7(c) is that Set 7 behaves contrary to the general data set characteristic. The rate of negative comments decreases when it should increase, and the rate of positive comments increases when it should decrease. To understand the reason for this behavior, n-gram analysis was applied to Set 7. The most frequent words and word groups in the data set are given in Table 5.

When Table 5 is examined, the most common words in the data set are “vaccine” and “microchip”. The first two 5-gram structures that are most common in the document appear as emojis that are converted into words to add emotion to the text in the text normalization step. These two emojis corresponding to 5-grams are given in parentheses. When the other 5-grams are examined, it is seen that the word groups “a microchip in the vaccine” are encountered again and at higher frequency when going from specific to general, that is, from 5-grams to 1-grams. Especially when 2-grams are examined, the words “microchip” and “vaccine” are frequently used in the document with different conjunctions and prepositions. This indicates that the comments are diverse and unique. The words “5g”, “5g chip” and “laughing” are also among the most common words that appear singularly in the document. When this document, which has a high positive polarity rate, is examined in the light of n-grams; it can be clearly stated that the rumor of “There is a microchip in the vaccine” is dominant and therefore fake news is not taken seriously, and users tweet under this title for fun. However, the “mark of the beast” word group, which is common in 4-gram and 5-gram, is not found among other common n-grams. This word, which is linked to the fake news “COVID-19 Vaccine Is Linked to the Mark of the Beast” (Conversation, 2021), which emerged at about the same time as the other news headlines used in the study, indicates that this word group was not accepted among Twitter users, and therefore this word group was put into circulation as a result of a bot activity.

A detailed analysis was needed as the polarity ratios of Set 1 and Set 2, given in Figure 3 (a) and (b), were very different from the general trend. Even if it is a small dataset, it is not necessary to read all the comments in the dataset one by one to understand why it is so different from the general trend. Instead, the most frequently occurring word groups in the data set can be identified using n-grams. Therefore, an n-gram analysis was performed to count how frequently consecutive word strings in the data set were seen. After the most common word groups and the number of their use together were noted, it was manually examined whether there were potential sentences starting with the same word groups. These potential sentences were tabulated manually, and a visual representation of the word groups most frequently found together in the data set is given in Table 6 and Table 7. In Tables 6 and 7, a sentence is obtained by taking one word from each column from left to right. In this way, it is possible to see the frequently encountered sentences and the number of their occurrences in a summarized way, without the need to read all the comments in the relevant data set. For example, looking at Table 7, we see that the phrase "Americans have died from the vaccine" is used 105 times in Set 2. The word group summary for sentences starting with the “Bill Gates” in Set 1 is given in Table 6. This group of words was chosen because the 2-gram number has the highest value for “Bill Gates”. A similar analysis is given in Table 7 for Set 2. The word "Americans" was chosen to start n-grams because it is at the highest frequency in Set 2. The most common sentences in the data set are given in each row from the first column to the other columns, and the numbers of frequent word groups up to 6-grams are summarized. In the 3rd and 5th rows where the n-gram representation of the sentences starting with “Bill Gates vaccine” and “Bill Gates said”, the 4-gram and 5-gram columns were left blank because the same word groups could not be found that repeat often enough to aggregate. The sequence in the continuation of these word groups is seen in the 6-gram+ column. The first column is 2-grams, which means that the ‘Bill’ ‘Gates’ pair is found side by side 2635 times.

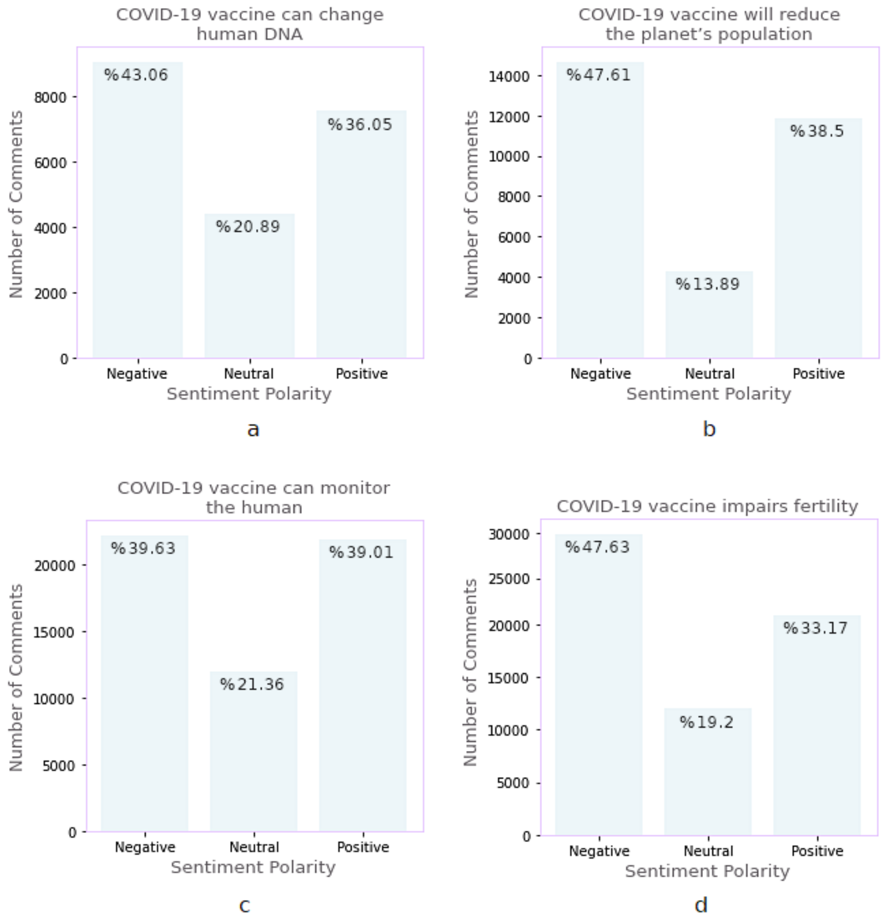


Figure 4. Polarity analysis of fake news datasets 5-8 (a) Set 5 "COVID-19 vaccine can change human DNA", number of data is 21073. (b) Set 6 "COVID-19 vaccine will reduce the planet's population", number of data is 30792. (c) Set 7 "COVID-19 vaccine can monitor the human", number of data is 56052. (d) Set 8 "COVID-19 vaccine impairs fertility", number of data is 62722

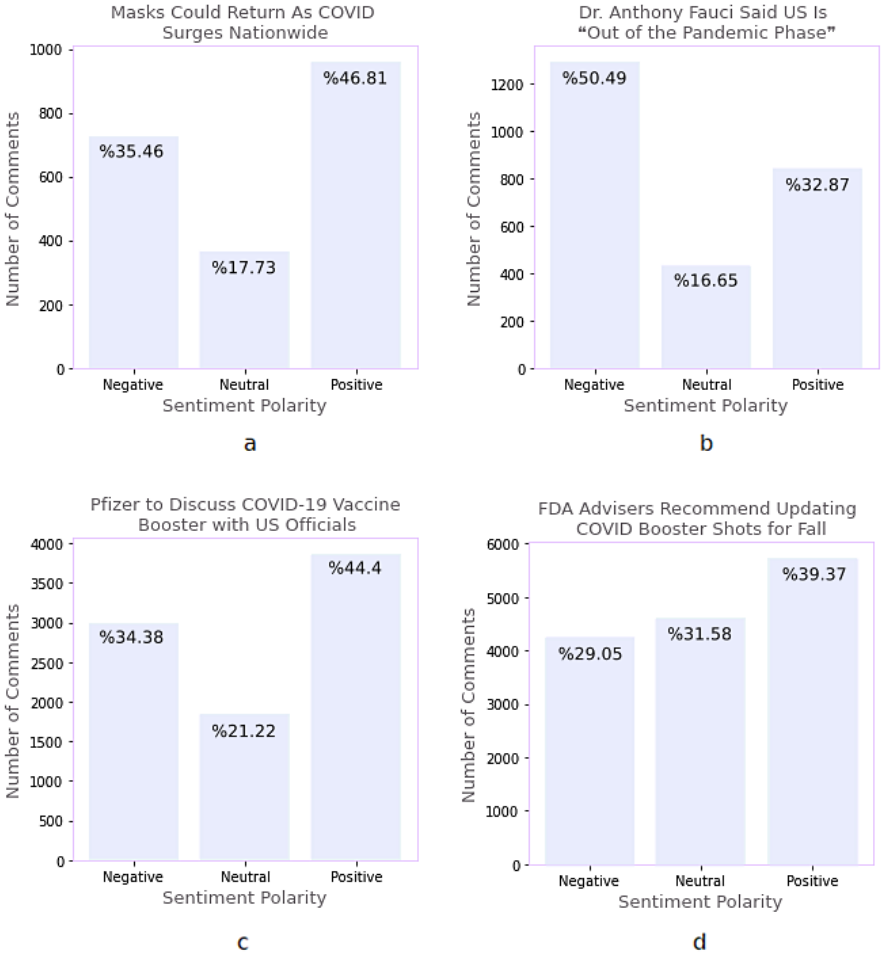


Figure 5. Polarity analysis of real news datasets 9-12 (a) Set 9 "Masks Could Return As COVID Surges Nationwide", number of data is 2053. (b) Set 10 "Dr. Anthony Fauci Said US Is 'Out of the Pandemic Phase'", number of data is 2565. (c) Set 11 "Pfizer to Discuss COVID-19 Vaccine Booster with US Officials", number of data is 8717. (d) Set 12 "FDA Advisers Recommend Updating COVID Booster Shots for Fall", number of data is 14579.

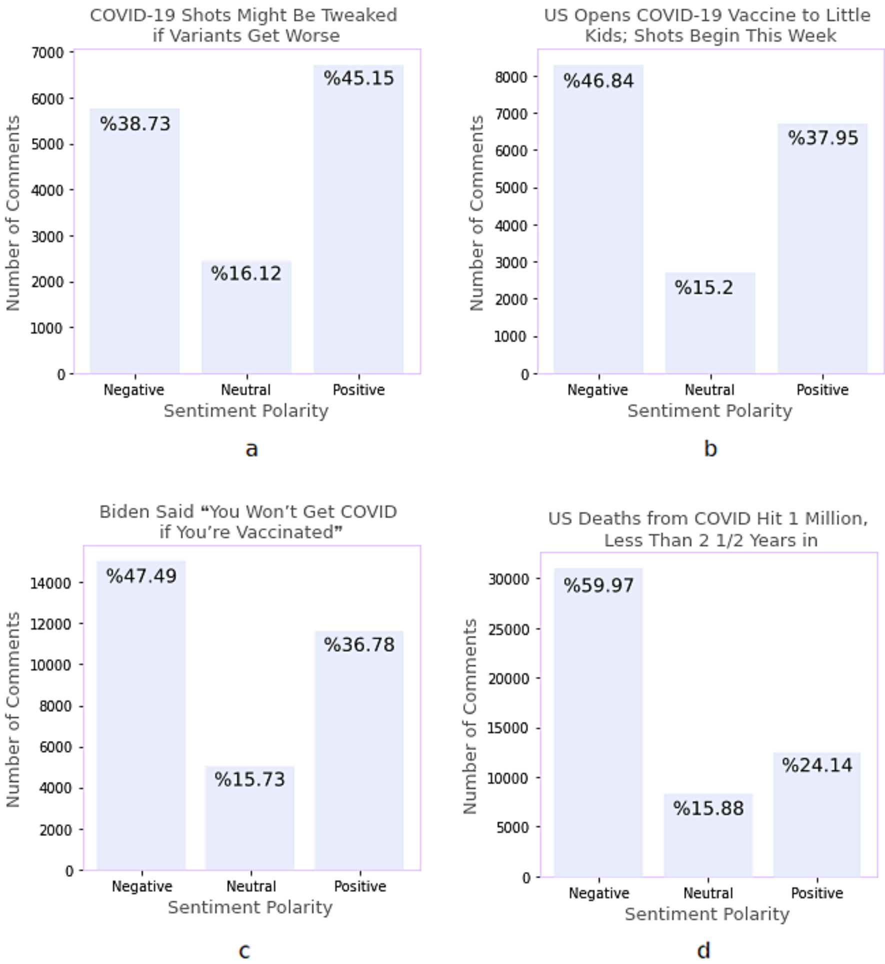


Figure 6. Polarity analysis of real news datasets 13-16 (a) Set 13 "COVID-19 Shots Might Be Tweaked if Variants Get Worse", number of data is 14888. (b) Set 14 "US Opens COVID-19 Vaccine to Little Kids; Shots Begin This Week", number of data is 17732. (c) Set 15 "Biden Say 'You Won't Get COVID if You're Vaccinated'", number of data is 31673. (d) Set 16 "US Deaths from COVID Hit 1 Million, Less Than 2 1/2 Years in", number of data is 51783.

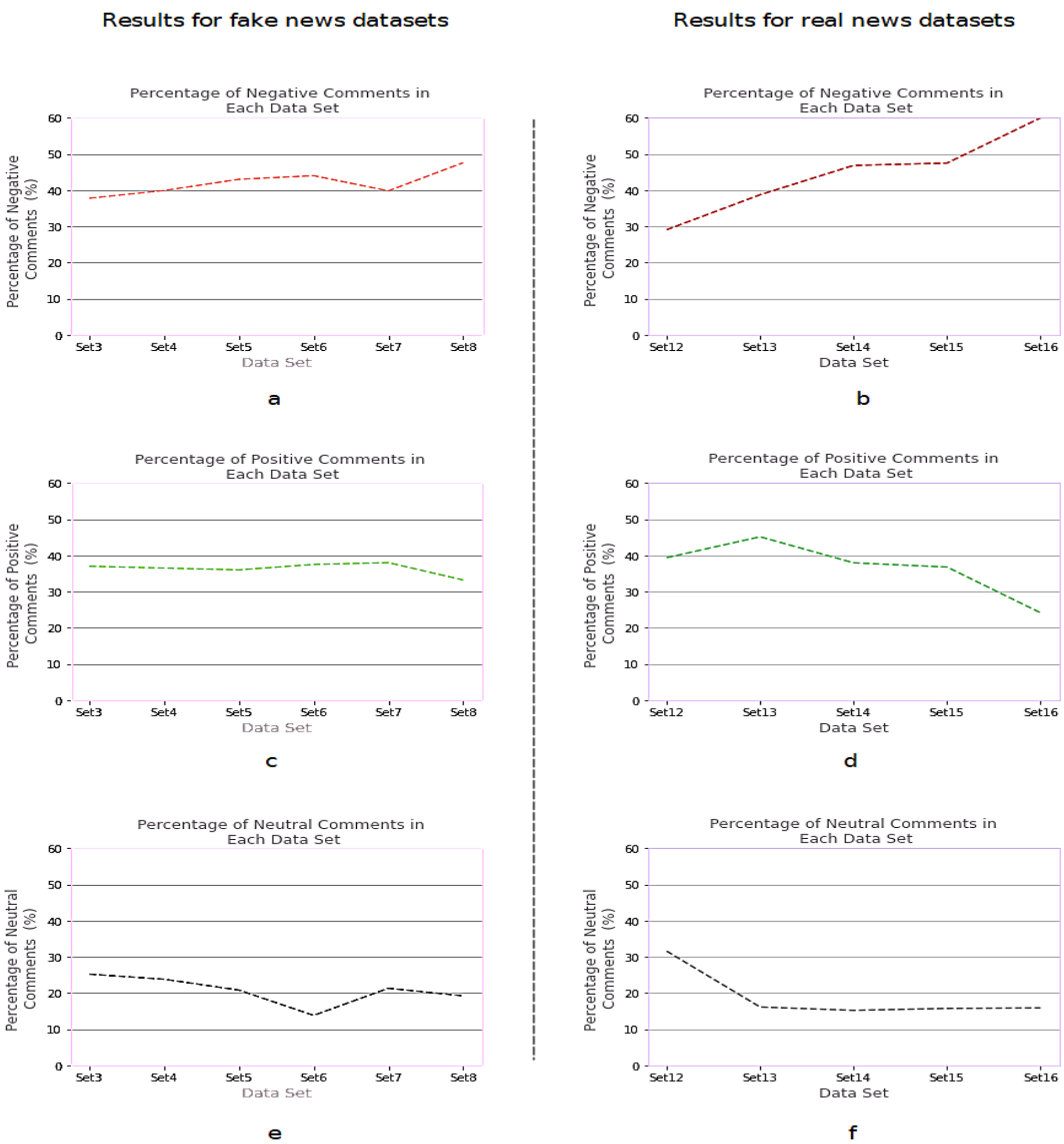


Figure 7. (a) The change in the rate of negative comments in the fake news data set according to the increasing number of data from Set 3 to Set 8. (b) Change in the rate of negative comments in the real news dataset with increasing data number from Set 12 to Set 16. (c) The change in the rate of positive comments in the fake news dataset according to the increasing number of data from Set 3 to Set 8. (d) Change in the rate of positive comments in the real news dataset with increasing data number from Set 12 to Set 16. (e) Neutral comment rate change in fake news dataset with increasing data number from Set 3 to Set 8. (f) Neutral comment rate change in real news dataset with increasing number of data from Set 12 to Set 16.

For example, if we read the table up to the 4th column of the 2nd row, which is 5-grams, the phrases “Bill Gates wants to kill” are found 87 times in the Set 1 data set, or if we read up to the 3rd column of the 4th row (4-gram) the phrases “Bill Gates will kill” appear side by side 54 times. The words in the last column of the table are the aggregated representation of the sentences in the data set from top to bottom such as “Bill Gates is tryingto kill people”, “Bill Gates is tryingto kill more people”, “Bill Gates is tryingto kill poor people”.

When we examine the n-gram analysis for Set 1, it is interesting to see that, comments can be so aggregated and have so many common word groups. From the last column showing 6-grams and more, it is understood that new phrases are derived using the adjectives of the same word (people, poor people, white people, old people, millions of people) and certain word groups (us all, them) are used in many of the sentences. These results suggest that the sentences are automatized and derived, and bots are used. In addition, the presence of the “Bill Gates” word group in 2635 of a dataset with 3145 comments is another important result that shows that the comments are not diverse and unique.

A similar situation applies to Set 2. It is interesting that prepositions come especially after the “Americans, have, died” group; like “from, of, due, and, in” or after “Americans, who, died”; like “from, without, of, because”. It appears that all possible sentences that can be written with a given set of words are produced. This situation regards that there are artificial sentence experiments produced to spread the existing word group.

In Table 8 and Table 9, the TF, IDF and TF-IDF scores of the six most frequent words in Set 1 and Set 2 are given, respectively. In Table 8, it is seen that the TF-IDF values of the words “vaccine”, “people” and “kill” with the lowest IDF values are also very low. For Table 9, the TF-IDF values of the words "americans" and "dead" are zero. This indicates that these words appear in most of the comments, and their uniqueness is extremely low. In the aggregated word group summary in Table 6 and Table 7, we see that these words are found in almost every comment. The frequent use of the word “kill” and “dead”, which has a high negative emotion intensity, explains the negative polarity rate of Set 1 to 88.14% and Set 2 to 88.65%. The fact that the IDF, which is the uniqueness score of words with high term frequency, is very low – or even zero for Set 2 – shows that similar sentences containing the same words are constantly tweeted.

Table 5
The most common words and phrases of Set 7

n-gram	Most frequent n-grams	Numbers
5-gram	'face', 'with', 'tears', 'of', 'joy' (😂)	2736
	'rolling', 'on', 'the', 'floor', 'laughing' (🤣)	1675
	'a', 'microchip', 'in', 'the', 'vaccine'	1355
	'the', 'mark', 'of', 'the', 'beast'	1241
4-gram	'microchip', 'in', 'the', 'vaccine'	2725
	'mark', 'of', 'the', 'beast'	2356
3-gram	'a', 'microchip', 'in'	5969
	'the', 'covid', 'vaccine'	3675
	'in', 'the', 'vaccine'	3637
	'there', 'is', 'a'	1608
2-gram	'a', 'microchip'	35055
	'the', 'vaccine'	26723
	'microchip', 'in'	20994
	'a', 'vaccine'	14293
	'to', 'microchip'	12810
	'the', 'microchip'	11508
	'microchip', 'vaccine'	4910
	'microchip', 'and'	3575
	'microchip', 'that'	3512
	'and', 'microchip'	3183
	'5g ', 'chip'	1713
1-gram	Vaccine	62572
	Microchip	54270
	5g	7763
	Laughing	2459

This result indicates that there is an effort to deliberately spread fake news and the use of bots. Since it is fake news that has not been spread, organic users' comments are in the data set, but since they are in the minority, the derived comments of bot accounts are especially evident as a result of such an analysis, and the low number of data in Set 1 turns into an advantage. The use of bots in the comments of this fake news, which could not reach Twitter users, is an intervention to the natural flow. Therefore, the interpretation rates of Set 1 and Set 2 were not considered within the scope of the analysis in the study. Thus, the results obtained in these data sets do not affect the fact that the increase in negative comments increases the prevalence of fake news.

When the data sets were examined, it was observed that negative emotions were intense in fake news comments, and positive and negative polarity were found to be balanced in real news. In our study, positive polarity is dominant at the rate of 50% in real news datasets, while negative polarity is dominant at 100% in fake news datasets. This observation is consistent with some previous studies (Dey et al., 2018; Dai, Sun & Wang, 2020). Dey et al. (2018) showed that credible tweets mostly have positive or neutral poles, while tweets with fake content have a strong tendency towards negative emotions. Dai, Sun & Wang (2020) made sentiment analysis on health news, and they found that responses to real news were more positive than fake news. The rise of negative comments as the prevalence of the news increases is common to both fake and real news reactions. This situation shows that the emotion that activates the desire to spread the news is negative emotions. This result is in parallel with the result obtained in the study of Bodaghi & Goliaei (2018). Bodaghi & Goliaei (2018), stated that an individual will feel a cognitive dissonance when they see that their belief is not known by others, or when they encounter an opposing view (true for those who believe in fake news, or fake news for those who believe in real news), and will tend to spread fake or true news by acting according to the backlash effect. In their study, where they provide a general framework for describing bots, Dickerson, Kagan & Subrahmanian (2014) stated that people tend to disagree more with the entire Twitter population than bots. This supports our analysis of bots in our study. While the rate of negative comments is over 88% in datasets inflated only by bots (Set 1 and Set 2), negative, positive and neutral comment rates in other datasets show a balanced distribution, which indicates differences of opinion in accordance with human nature.

Table 6
Word group summary of Set 1

2-gram	3-gram	4-gram	5-gram	6-gram+
Bill Gates (2635)	is (350)	tryingto (70)	kill (69)	people more people poor people black people us all them
	wants (159)	to (139)	kill (87)	people poor people millions of people white people old people us all them
	vaccine (103)			could kill more people kills enough people kill people will kill us all will kill millions will change our dna actually does kill did kill people
	will (80)	kill (54)	people (49)	
	said (77)			that the vaccine will kill people a vaccine would kill people the vaccine will kill million people a vaccine will kill almost a million people he was going to kill billion people
	admits (54)	his (21)	covid (19)	vaccine will kill people vaccine will kill millions of people vaccine might kill nearly people

Table 7
Word group summary of Set 2. (The number of words is less than 20 and the parts that cannot be divided further are given as N/A in order not to increase the complexity.)

1-gram	2-gram	3-gram	4-gram	5-gram	6-gram
Americans (6180)	have (2194)	died (1917)	from (946)	covid (870)	as (77)
					and (67)
					with (50)
				the (423)	vaccine (105)
					flu (89)
					covid (29)
					virus (20)
				it (100)	N/A
			of (300)	covid (396)	N/A
				the (108)	N/A
			due (160)	to (159)	the (100)
			and (143)	N/A	N/A
			in (103)		
		already (115)	died (100)	N/A	N/A
	died (1739)	from (385)	covid (152)	N/A	N/A
			the (114)		
		because (259)	of (144)	N/A	N/A
		of (195)	covid (98)	N/A	N/A
	who (531)	died (278)	from (62)	covid (24)	N/A
			without (55)	getting (55)	
			of (30)	covid (28)	
			because (26)	N/A	
		have (196)	died (167)	from (62)	covid (31)
				of (21)	covid (16)

Table 8
tf, idf and tf-idf values for the six words of Set 1 with the highest term frequency.

words	tf	idf	tf-ifd
vaccine	11.0	0.000318	0.003498
admits	9.0	3.749504	33.745535
gate	9.0	0.137126	1.234137
bill	9.0	0.137126	1.2345137
people	9.0	0.000636	0.005725
kill	9.0	0.000318	0.002862

Table 9
tf, idf and tf-idf values for the six words of Set 2 with the highest term frequency.

words	tf	idf	tf-ifd
americans	10.0	0.000000	0.000000
died	10.0	0.000000	0.000000
vaccine	9.0	1.340795	8.044771
another	9.0	3.804824	11.414473
virus	9.0	1.908531	5.725593
covid	9.0	0.684902	2.054705

4. Conclusion

In this study, sentiment analysis was applied on the reactions to fake and real news at different prevalence rates. Although there are minor differences, the changes in the emotional states of the comments made on real news and fake news show parallelism with each other. Fake news is successful as long as it resembles real news. Therefore, the success of fake news depends on how good it presents itself as real news. The better the fake news copies the real news and the more it makes an impression that it is real, the more attention it receives as if it were real news. However, it has been revealed that negative emotion is a very effective and key emotion in spreading the news, whether for fake news or real news. However, especially for fake news, negative polarity was seen to predominate in all datasets without exception. This clearly reveals that the strongest emotion in spreading fake news is negative emotions.

When it comes to real news, we can say that people take a much more serious approach by looking at the obvious decrease in the positive comment rates of real news as the prevalence increases. However, it is found that users who comment on fake news for fun significantly increase the prevalence. Even if there is a widespread fake news with the highest rate of positive comments among all data sets, the fact that the rate of negative comments is higher than the rate of positive comments shows that negative emotion is extremely effective in spreading fake news.

Another factor that we come across about fake news is bot activities. It is seen that bot activities come into play to spread the fake news after it is revealed. To reveal bot activity NLP techniques were applied in the study. The presence of word groups pointing to a different fake news headline in the comments made on a fake news headline makes us think that the fake news came out deliberately from a single source.

Author Contributions

Maide Feyza Er: Collected data and performed the analysis.

Yonca Bayrakdar Yılmaz: Performed statistical analysis and wrote the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. J. (2011, June). Sentiment analysis of twitter data. In Proceedings of the workshop on language in social media (LSM 2011) (pp. 30-38). Retrieved from: https://aclanthology.org/W11-0705

Alonso, M. A., Vilares, D., Gómez-Rodríguez, C., & Vilares, J. (2021). Sentiment analysis for fake news detection. Electronics, 10(11), 1348. DOI: https://doi.org/10.3390/electronics10111348

Anoop, K., Deepak, P., & Lajish, V. L. (2020). Emotion cognizance improves health fake news identification. In IDEAS (p. 24). DOI: https://doi.org/10.48550/arXiv.1906.10365

Antonakaki, D., Fragopoulou, P., & Ioannidis, S. (2021). A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks. Expert Systems with Applications, 164, 114006. DOI: https://doi.org/10.1016/j.eswa.2020.114006

Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc.. Retrieved from: https://www.jstor.org/stable/40925581

Bodaghi, A., & Goliaei, S. (2018). A novel model for rumor spreading on social networks with considering the influence of dissenting opinions. Advances in Complex Systems, 21(06n07), 1850011. DOI: https://doi.org/10.1142/S021952591850011X

Bodaghi, A., & Oliveira, J. (2022). The theater of fake news spreading, who plays which role? A study on real graphs of spreading on Twitter. Expert Systems with Applications, 189, 116110. DOI: https://doi.org/10.1016/j.eswa.2021.116110

Check Your Fact. (2019). Retrieved from checkyourfact website: https://checkyourfact.com/

Contractions. (2022). Retrieved July 26, 2022, from https://dictionary.cambridge.org/de/grammatik/britisch-grammatik/contractions

Conversation, T. (2021, April 8). COVID-19 vaccine is not linked to the mark of the beast. Retrieved June 12, 2021, from https://www.snopes.com/news/2021/04/08/no-the-covid-19-vaccine-is-not-linked-to-the-mark-of-the-beast

Cui, L., Wang, S., & Lee, D. (2019). Same: sentiment-aware multi-modal embedding for detecting fake news. In Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining (pp. 41-48). DOI: https://doi.org/10.1145/3341161.3342894

Dai, E., Sun, Y., & Wang, S. (2020, May). Ginger cannot cure cancer: Battling fake health news with a comprehensive data repository. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 14, pp. 853-862). DOI: https://doi.org/10.1609/icwsm.v14i1.7350

Dey, A., Rafi, R. Z., Parash, S. H., Arko, S. K., & Chakrabarty, A. (2018, June). Fake news pattern recognition using linguistic analysis. In 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition. (pp. 305-309). IEEE. DOI: 10.1109/ICIEV.2018.8641018

Dickerson, J. P., Kagan, V., & Subrahmanian, V. S. (2014, August). Using sentiment to detect bots on twitter: Are humans more opinionated than bots?. In 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014) (pp. 620-627). IEEE. DOI: 10.1109/ASONAM.2014.6921650

Dzogang, F., Lightman, S., & Cristianini, N. (2018). Diurnal variations of psychometric indicators in Twitter content. PloS one, 13(6), e0197002. DOI: https://doi.org/10.1371/journal.pone.0197002

FactCheck.org. (2008). Retrieved from FactCheck.org website: https://www.factcheck.org/

Giachanou, A., & Crestani, F. (2016). Like it or not: A survey of twitter sentiment analysis methods. ACM Computing Surveys (CSUR), 49(2), 1-41. DOI: https://doi.org/10.1145/2938640

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. CS224N project report, Stanford, 1(12), 2009. Retrieved June 22, 2021, from https://www-cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf

Hamdan, H., Béchet, F., & Bellot, P. (2013, June). Experiments with DBpedia, WordNet and SentiWordNet as resources for sentiment analysis in micro-blogging. In Second Joint Conference on Lexical and Computational Semantics (* SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013) (pp. 455-459). Retrieved from: https://aclanthology.org/S13-2075

Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8 No. 1, pp. 216-225). DOI: https://doi.org/10.1609/icwsm.v8i1.14550

Informal contractions list in English. (2020, December 25). Retrieved July 26, 2021, from https://7esl.com/informal-contractions-list

Internet Slang Terms. (2021, March 21). Retrieved July 28, 2021, from https://7esl.com/internet-slang

Islam, M. S., Kamal, A. H. M., Kabir, A., Southern, D. L., Khan, S. H., Hasan, S. M., ... and Seale, H., "COVID-19 vaccine rumors and conspiracy theories: The need for cognitive inoculation against misinformation to improve vaccine adherence." PloS ONE, 16 (5): e0251605, (2021). DOI: https://doi.org/10.1371/journal.pone.0251605

Iwendi, C., Mohan, S., Ibeke, E., Ahmadian, A., & Ciano, T. (2022). Covid-19 fake news sentiment analysis. Computers and electrical engineering, 101, 107967. DOI: https://doi.org/10.1016/j.compeleceng.2022.107967

Kim, T. & Wurster K. (2015). emoji (Version v1.7.0) [Computer software]. https://github.com/carpedm20/emoji

Kouloumpis, E., Wilson, T., & Moore, J. (2011). Twitter sentiment analysis: The good the bad and the omg!. In Proceedings of the international AAAI conference on web and social media (Vol. 5, No. 1, pp. 538-541). DOI: https://doi.org/10.1609/icwsm.v5i1.14185

Manning, C. D., Raghavan, P., Schutze, H., & Cambridge University Press. (2008). Introduction to information retrieval. Cambridge: Cambridge University Press.

Neethu, M. S., & Rajasree, R. (2013, July). Sentiment analysis in twitter using machine learning techniques. In 2013 fourth international conference on computing, communications and networking technologies (ICCCNT) (pp. 1-5). IEEE. DOI: 10.1109/ICCCNT.2013.6726818

Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C. T., & Nielsen, R. K. (2021). Reuters institute digital news report 2021. Reuters Institute for the Study of Journalism. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3873260

Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10). Retrieved from: http://www.lrec-conf.org/proceedings/lrec2010/pdf/385_Paper.pdf

Saif, H., He, Y., & Alani, H. (2012, November). Semantic sentiment analysis of twitter. In International semantic web conference (pp. 508-524). Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-35176-1_32

Snefjella, B., Schmidtke, D., & Kuperman, V. (2018). National character stereotypes mirror language use: A study of Canadian and American tweets. PloS one, 13(11), e0206188. DOI: https://doi.org/10.1371/journal.pone.0206188

Snopes. (2018). Retrieved from Snopes.com website: https://www.snopes.com/

Vicario, M. D., Quattrociocchi, W., Scala, A., & Zollo, F. (2019). Polarization and fake news: Early warning of potential misinformation targets. *ACM Transactions* on the Web (TWEB), 13(2), 1-22. DOI: https://doi.org/10.1145/3316809

Zacharias, C. & Poldi, F. (2018). Twint (Version 2.1.4) [Computer software]. https://github.com/twintproject/twint

Zhang, X., Cao, J., Li, X., Sheng, Q., Zhong, L., & Shu, K. (2021). Mining dual emotion for fake news detection. *In Proceedings of the Web Conference 2021* (pp. 3465-3476). DOI: https://doi.org/10.1145/3442381.3450004