Политическая лингвистика. 2023. № 2 (98).
Political Linguistics. 2023. No 2 (98).

УДК 81`42:81`38
ББК Ш105.51+Ш105.551.5 *ГСНТИ 16.21.27; 16.31.21* *Код ВАК 5.9.8*
doi: 10.26170/1999-2629_2023_02_06

Maria Yu. Ilyushkina
Ural Federal University named after the first President of Russia B.N.Yeltsin, Ekaterinburg, Russia,
ilyushkina_maria@mail.ru, https://orcid.org/0000-0003-4612-3025

Automated Sentiment Analysis of Social Networks During Epidemic Outbreaks

ABSTRACT. *Nowadays, the coronavirus infection COVID-19 caused by the SARS-CoV-2 virus is a global concern as it leads to a dramatic increase in new infections and deaths among the population. Obviously, COVID-19 is not the first pandemic in the world. Doctors and researchers have seriously tackled various viruses such as Ebola, Mers-Cov, SARS, etc.*

Since the first headline about the coronavirus disease outbreak was published in December 2019, social networks have become a favourable ground for the spread of information about new COVID-19. And the impact of social media during such an unprecedented pandemic crisis is to be defined in this study. The author believes many outbreaks and pandemics could have been controlled promptly if experts had considered the social media data.

In the current paper a social media platform Twitter was subjected to thorough analysis as it has the most accessible data sources currently available. Moreover, people in this social media are free to discuss and share their opinions about events in their daily life or express emotions associated with the pandemic. Thus, in the current study social data serves as the basis for analyzing the opinions and emotional attitude (sentiment) of the author of the text to some objects, processes or events. The aim of the research lies within the field of word processing in natural language. The automated sentiment analysis of messages from Twitter was carried out on the basis of modern computer systems and platforms. The large-scale identification of human emotions in social media is essential for international public influence, business decisions and policy development.

KEYWORDS: *Internet, Internet communication, Internet technologies, Internet discourse, Internet texts, Internet resources, social networks, analysis of emotional tone, COVID-19, pandemic, coronavirus, emotional attitude, automated text procession.*

AUTHOR'S INFORMATION: *Ilyushkina Maria Yur'evna, Candidate of Philology, Head of Department of Linguistics and Professional Communication in Foreign Languages, Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia*.

ACKNOWLEDGMENTS. The project is carried out by the winner of the grant competition among teachers of the master's degree courses 2021/2022 within V. Potanin scholarship № 22-000761.

FOR CITATION: *Ilyushkina M. Yu.* (2023). Automated Sentiment Analysis of Social Networks During Epidemic Outbreaks. In *Political Linguistics*. No 2 (98), pp. 54-62. (In Engl.). DOI: 10.26170/1999-2629_2023_02_06.

Мария Юрьевна Илюшкина
Уральский федеральный государственный университет имени первого Президента России Б. Н. Ельцина, Екатеринбург, Россия, ilyushkina_maria@mail.ru, https://orcid.org/0000-0003-4612-3025

Автоматизированный анализ тональности текстов в социальных сетях во время пандемий

АННОТАЦИЯ. *В настоящее время стремительное распространение по всему миру коронавирусной инфекции КОВИД-19, вызываемой вирусом SARS-CoV-2, является проблемой глобального масштаба, так как приводит к резкому росту новых случаев инфицирования, а также смертности среди населения. Очевидно, что это далеко не первый случай пандемии. В прошлом врачи и ученые серьезно занимались исследованием различных вирусов, таких как вирус Эбола, Mers-Cov, SARS и др.*

С тех пор, как в декабре 2019 года был опубликован первый заголовок о вспышке коронавирусной инфекции (КОВИД-19), социальные сети стали благоприятной средой для распространения информации о новом вирусе. В данной статье предпринята попытка определить влияние социальных сетей во время такого беспрецедентного пандемического кризиса. Автор полагает, что многие вспышки и пандемии можно было бы оперативно контролировать, если бы эксперты учитывали данные социальных сетей.

В данном исследовании подверглась анализу платформа социальной сети «Твиттер», так как именно она является самым доступным из существующих в настоящее время источником получения данных. Более того, в социальных сетях люди свободно обсуждают и высказывают свои мнения о событиях, происходящих в повседневной жизни, или выражают эмоции, связанные с пандемией. Таким образом, социальные данные служат основой для анализа мнений и эмоционального отношения автора текста к некоторым объектам, процессам или событиям. Целью работы является проведение автоматизированного анализа тональности сообщений в «Твиттере» с применением современных компьютерных программ и технологий в русле обработки естественного языка. Широкомасштабное определение человеческих эмоций в социальных сетях имеет важное значение для международного влияния на общество, принятия деловых решений и разработки политических стратегий.

КЛЮЧЕВЫЕ СЛОВА: *Интернет, интернет-коммуникация, интернет-технологии, интернет-дискурс, интернет-тексты, интернет-ресурсы, социальные сети, анализ эмоциональной окраски, КОВИД-19, пандемия, коронавирус, эмоциональное отношение, автоматическая обработка текста.*

ИНФОРМАЦИЯ ОБ АВТОРЕ: *Илюшкина Мария Юрьевна, кандидат филологических наук, зав. кафедрой лингвистики и профессиональной коммуникации на иностранных языках, Уральский федеральный университет имени первого Президента России Б. Н. Ельцина (Екатеринбург, Россия); 620007, Россия, Екатеринбург, пр-т Ленина, 51, к. 117; email: ilyushkina_maria@mail.ru.*

БЛАГОДАРНОСТИ. Проект реализуется победителем грантового конкурса для преподавателей магистратуры 2021/2022 Стипендиальной программы Владимира Потанина № 22-000761.

ДЛЯ ЦИТИРОВАНИЯ: *Илюшкина, М. Ю.* Автоматизированный анализ тональности текстов в социальных сетях во время пандемий / М. Ю. Илюшкина. — Текст : непосредственный // Политическая лингвистика. — 2023. — № 2 (98). — С. 54-62. — DOI: 10.26170/1999-2629_2023_02_06. — (На англ. яз.).

INTRODUCTION

During the COVID-19 pandemic social networks have been used as platforms by healthcare institutions, government organizations, hospitals and journals to distribute information in a timely manner. Social networks have also been the platform for the doctors treating COVID-19 patients to promptly inform the medical community about their experiences.

Many healthcare professionals are empowered to analyze emerging data, which helps them and government institutions understand how population reacts to events, benefiting the public interest [Alsaeedi & Khan 2019]. However, current trends in technology should be considered as well, as they have shown their contribution to medical decision-making related to infectious diseases and their outbreaks [Bhat et al. 2020; Soliman et al. 2020: 92]. Social networks have developed platforms that facilitate communication between human communities and help them exchange ideas, information, knowledge and other data through forms of electronic communication [Kim & Yoo-Lee 2014]. Social media platforms are considered to be the fastest growing information systems for public use [Appel et al. 2020: 79-95]. A distinctive feature of Twitter is the public availability of messages posted there. Twitter has become not only one of the most popular ways of communication, but also has infiltrated in the vocabulary and the way of thinking.

Thus, given the time that users spend on such social networks, from a few hours a week to daily use, these applications undoubtedly influence the behavior of people. Despite the large amount of data presented on these social media platforms, their content can have contradictory effects that range from negative to positive psychological effect on people's lives.

Sentiment analysis in this context can become a useful tool for assessment of public opinion. Sentiment analysis contributes to the understanding of human emotions as it can interpret the behavior of people who interact with social networking applications. In addition, these multimedia applications are used in various fields, including health [Rodrigues et al. 2016: 80], business, tourism [Ainin & Abdullah 2020], education, such as opinion mining and enabling people to express their emotions freely [Zarrad et al. 2014: 664-670]. Thus, sentiment analysis is a powerful tool for understanding the most important global trends and events. Especially when an epidemic becomes such an event. After all, immediate and effective communication is necessary to reduce uncertainty, since it allows getting a clear idea of the situation [Bordia et al. 2004].

THEORY AND METHODS

The increased interest in the analysis of subjective opinions in user texts dates back to the early 2000s [Pang & Lee 2008]. The book “Sentiment Analysis and Opinion Mining” [Liu 2012] shows the importance of applying this analysis in all spheres of business and social life. Moreover, it describes the problem of automatic sentiment detection of the text, as well as the problems that researchers most often face.

Thus, sentiment analysis of texts today can be considered as one of the most popular research topics in the field of natural language processing [Sailunaz & Alhajji 2019]. Many scientists are working on the analysis of sentiment of texts from social networks, for example, concerning the disease COVID-19. Thus, C. Kaur and A. Sharma analyze the sentiment of messages-tweets in the social network Twitter and determine the attitude of people towards the coronavirus disease COVID-19 [Kaur & Sharma 2020]. R. Medford compiled a list of COVID-19-related hashtags, that were used to distribute messages in blogs, and then analyzed their sentiment to determine the emotional significance of each message and identify the dominant emotion [Medford 2020]. K. Sailunaz and R. Al-Hajj are analyzing the sentiment of Arab tweets, Twitter posts containing hashtags related to government-imposed public health measures, and measuring the number of positive and negative messages [Sailunaz & Alhajj 2019].

The above mentioned researchers investigate the question of the relationship between

social networks and human behavior during the COVID-19 pandemic, emphasizing the practical importance of analyzing the sentiment of texts in solving the assigned tasks.

Sentiment analysis of texts finds its application, starting from assessing the quality of goods and services and finishing with the compilation of texts with predetermined emotional characteristics. The object of emotional assessment can be a proper name, the brand of a product, organization, service or profession, in relation to which the opinion is expressed. However, the daily number of published reviews in social networks is so huge that their manual processing is impossible and requires automation. Therefore, sentiment analysis every year becomes an increasingly urgent task from both theoretical and practical points of view due to the development of the Internet and the change in the format of communication between people.

The purpose of this research is to analyze the sentiment of messages in the social network Twitter during the pandemic period from December, 2020 to January, 2022. We consider “the sentiment” to be an emotional assessment expressed by the author regarding the COVID-19 disease.

This study does not target a specific continent, country or city to collect data as coronavirus is an almost ubiquitous health problem. The existing tools and approaches for word processing and automatic determination of their emotional component have been analyzed. To achieve the goals an application programming interface (Twitter API) was implemented as it allows one computer program to interact with others, load the Tweepy data library, and upload the necessary information by keywords. Throughout the research new words in the period of 2020-2021 related to the studied topic were defined. Moreover, the dynamics and frequency of their use was analyzed as well.

THE PROCEDURE AND ALGORITHM OF THE AUTOMATED SENTIMENT ANALYSIS

The study is aimed at understanding not only the interconnection between the data obtained from social media and human behavior during the outbreak of the coronavirus known as COVID-19, but also at showing how people, government organizations and news agencies transmit such situations.

Many people and government authorities want to present their views to the public about the coronavirus and use social media for this purpose. This data can be analyzed in terms of emotional coloration of utterances to observe large-scale communities at low cost [Choi et al. 2017].

Sentiment analysis is a class of computational linguistics methods for the automated identification of emotionally charged vocabulary and opinions on the topics mentioned in texts. In the analysis of the sentiment of the text, another term that is associated with the concept of opinion — subjectivity, is used. Pang and Lee define an objective sentence as expressing factual information about something, while a subjective one expresses someone's personal feelings or assumptions [Pang & Lee 2008: 6].

The algorithm for classifying texts is usually called a classifier. In the current research we have launched a probabilistic classifier such as a Naive Bayesian classifier, which is based on the maximum entropy method. The Naive Bayesian classifier is traditionally used in text classification tasks, such as spam filtering, automatic categorization, or document sentiment identification.

Sentiment analysis, also known as opinion mining, is a powerful tool that can be used to gain insight into mentions, brands, and products. It is a natural language processing algorithm that provides an overview of the positive, neutral and negative characteristic of the text, and that companies rely on to monitor social data.

At the first stage, operation with social media platforms is carried out. The user chooses the source from which he wants to extract an opinion; for example, when choosing social networks, the user can choose from various online sources such as Twitter and Reddit.

The second stage involves data collection with using specific keywords or hashtags (#) to get the desired information based on their preferences. This information takes various forms, such as tweets, messages, news and texts.

The third stage is a preliminary processing of the extracted information. It is processed to prepare the data for the next stage. This step includes feature extraction.

The fourth stage is data analysis when all preprocessed data is used for its designated purpose, for example, to determine the polarity, the sentiment analysis or the frequency analysis [Kim et al. 2016: 763].

Every stage of the sentiment analysis involves the usage of computer technology. For example, one of the most popular modern programming languages, Python, with its libraries is a universal tool for solving many applied problems, from automated information collecting to data analysis.

To achieve the research goal, two main methods were used:

1) based on the analysis of mass data of the social network Twitter: collection, processing and analysis of data;

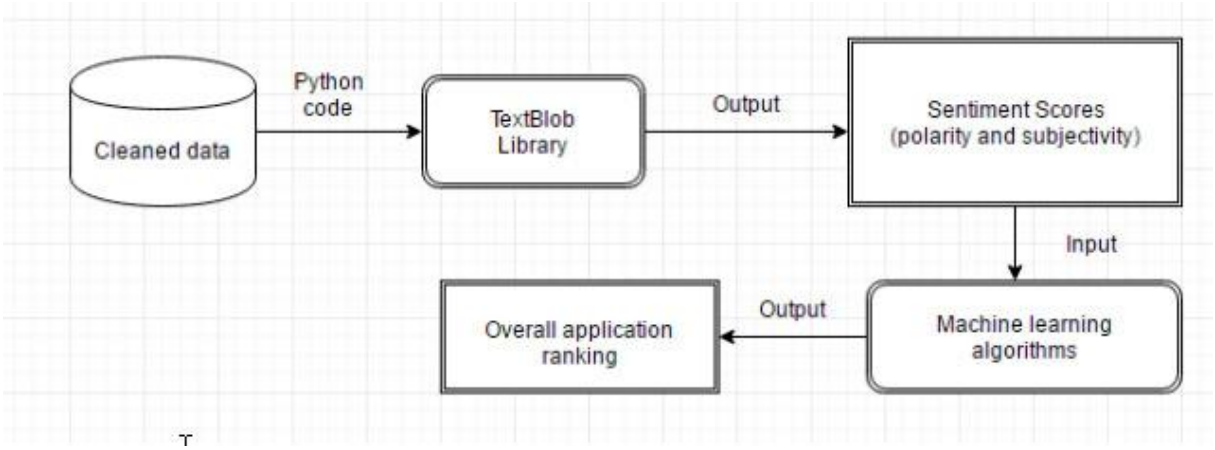


Fig. 1. TextBlob library operation

2) based on extracted tweets, natural language processing (hereinafter NLP — Natural language processing) to determine the sentiment of users' mood.

The data-driven methodology uses the Twitter API to access the Tweepy library through the Python programming language, which allows you to retrieve the relevant data using keyword and hashtag (#) searches. The extracted tweet methodology uses TextBlob, a text processing library written in Python, and also uses a Naive Bayes classifier for machine learning in the English language.

SENTIMENT ANALYSIS IN PYTHON WITH TEXTBLOB

Advanced technologies in natural language processing (NLP) allow us to analyze natural languages on different layers: from simple segmentation of textual information to more sophisticated methods of sentiment categorizations. Thus, for example, Python library — TextBlob — was created for processing textual data which provides a simple API interface for diving into common NLP tasks such as part-of-speech tagging, noun phrase extraction and sentiment analysis.

The approach that the TextBlob package applies to the sentiment analysis differs in that it is rule-based and, therefore, requires a predefined set of categorized words. These words can, for example, be uploaded from the NLTK database — a package of libraries and programs for symbolic and statistical natural language processing written in the Python programming language.

Once a Python model is fed by the necessary input data, a user can obtain the sentiment scores in the form of polarity and subjectivity. We can see how this process works in Figure 1.

The first step is to analyze the input file, which is the users` feedback file, i.e. a text format designed to present tabular data and pass it as an input to the sentiment analysis model. A Python library named TextBlob is used to create this model. This library contains two implementations of the sentiment analysis, that is, the NaiveBayes analyzer, which is part of the NLTK, and the template analyzer, which depends on template libraries. The Python code analyzes the file and uses a sentiment model, after that it gives an output in the form of polarity and subjectivity.

TextBlob output for a polarity task is a float within the range [-1.0, 1.0] where -1.0 is a negative polarity and 1.0 is positive. This score can also be equal to 0, which stands for a neutral evaluation of a statement as it doesn't contain any words from the training set. Polarity and subjectivity help define the sentiment of a comment and understand the opinions of different users. If the estimation decreases towards the positive range, then the comment is more likely to be positive. If the estimation equals to zero, the comment is neutral, and if the estimation lies in the negative range, the comment tends to be negative. Subjective evaluation helps to understand the objectivity of the sentence. If this estimation drops to 0.0, we can say that the comment is objective in nature, and if the estimation is close to 1.0, then the comment is subjective.

To change the default settings, we specify a Naïve Bayes analyzer in the code. Our Python model gives the sentiment evaluation. The TextBlob library is used to perform sentiment analysis on imported sentences as well as tweets.

SENTIMENT ANALYSIS OF COVID-19 MESSAGES

Sentiment analysis is a machine learning algorithm that automatically determines the emotional coloring of online conversations, classifying them as positive, negative, or neutral. Sentiment analysis tools are a cost-

effective and fast way to analyze huge amounts of opinion data that people simply cannot process efficiently. Equipped with natural language understanding feature (NLU), sentiment analysis tools can automatically understand text in the same way as a human. In our paper, sentiment analysis means finding opinions in the text and determining the emotional attitude to the pandemic events.

There are many open source code sentiment analysis resources used to create a sentiment analysis tool that can produce accurate results. We will apply one of these options for sentiment analysis evaluation in this research, using a database with messages from Twitter.

The platform Brand24 is a social tool designed to track and analyze brand status in real time, and will be used in the current research. The “mentions tab” filters and displays all mentions by sentiment, the “analysis tab” shows things like reposts, interactions, and likes, and the “summary tab” shows the overall coverage of a brand, product, or campaign. Brand24 allows us to export information from various sources and provides a 14-day free trial, allows up to 100,000 text messages for each study subject and has a built-in «sentiment analysis» function, which is perfectly suited for solving the problems in our research.

During pandemic time people spend more time on social media than usual. Twitter transmits news faster than official news channels and emergency response agencies [Imran et al. 2020]. With proper planning and implementation, social media data can be analyzed and processed to extract situational information that can be used to develop effective responses to a pandemic. Therefore, in this research, we use Twitter social media posts, known as tweets, that contain statements and opinions about COVID-19. We search for the required tweets by keywords of the year 2020 referring to coronavirus topic, for example, "Lockdown" — blocking, "social distancing" — maintaining social distance, etc. Based on Coronavirus Corpus, designed to reflect social, cultural and economic impacts of coronavirus (COVID-19) in 2020 and beyond, we evaluate the dynamics and frequency of use of identified words from January, 2020 to December, 2021 and view how things have changed over time. We also use an online platform Brand24, designed to track and analyze social opinions from Twitter in real time, allowing users to have negative and positive mentions, and analyze tweets for a month from December, 2020 to January, 2021, which was indicated as the peak period. We are using the Python programming language as it also provides many easy-to-use libraries for accessing the social networking platform Twitter. With the resulting data, we analyze the sentiment of the tweets (or, as it is often called "the analysis of emotional coloring"). Sentiment analysis should, first of all, serve the purposes of describing and predicting social phenomena in order to prevent victims and panic during pandemics and disasters.

CONSTRUCTING SAMPLES FOR ANALYSIS

We started constructing the samples for analysis by defining the keywords by which the search and export of the tweets we need will be carried out. For this reason, we search several well-known online sources such as the Global Language Monitor and Merriam-Webster, as well as dictionaries such as the Oxford English Dictionary and Collins Dictionary, which provide a list of 2020 new words on coronavirus topics, such as self-quarantine, lockdown, self-isolate, social distancing, Zoombombing, Coronacation, covidiot, etc.

To analyze the dynamics of the use of new words appeared in 2020, a corpus Coronavirus Corpus (*https://www.english-corpora.org*) was investigated. The corpus shows what people actually say in online newspapers and magazines in 20 different English-speaking countries. The corpus currently has around 814 million words and keep on growing with 3-4 million words every day. We can also compare different periods to see how things have changed over time. To do this, in the search bar, we indicate the search word and the time period for analysis (see Figure 2).

As can be seen in Figure 2, the search word is indicated in the search bar, we indicate the time interval and click "search for matching lines" of the search for matches. We begin our analysis with the word “Coronavirus”. Picture 1 demonstrates the dynamics of the use of this word in 2020.

As we can see in Picture 1, the word “coronavirus” was used according to the corpus 20 thousand times in January, 2020. This is due to the fact that the first death from the virus was confirmed, and in February, 2020 the virus began to spread rapidly in various countries, despite the quarantine measures taken by the Chinese authorities.

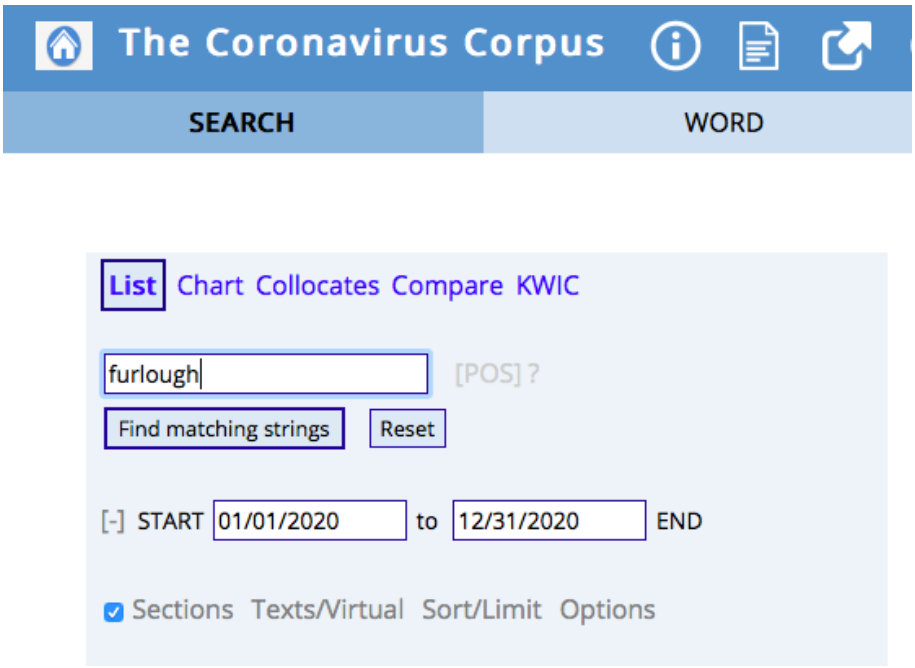
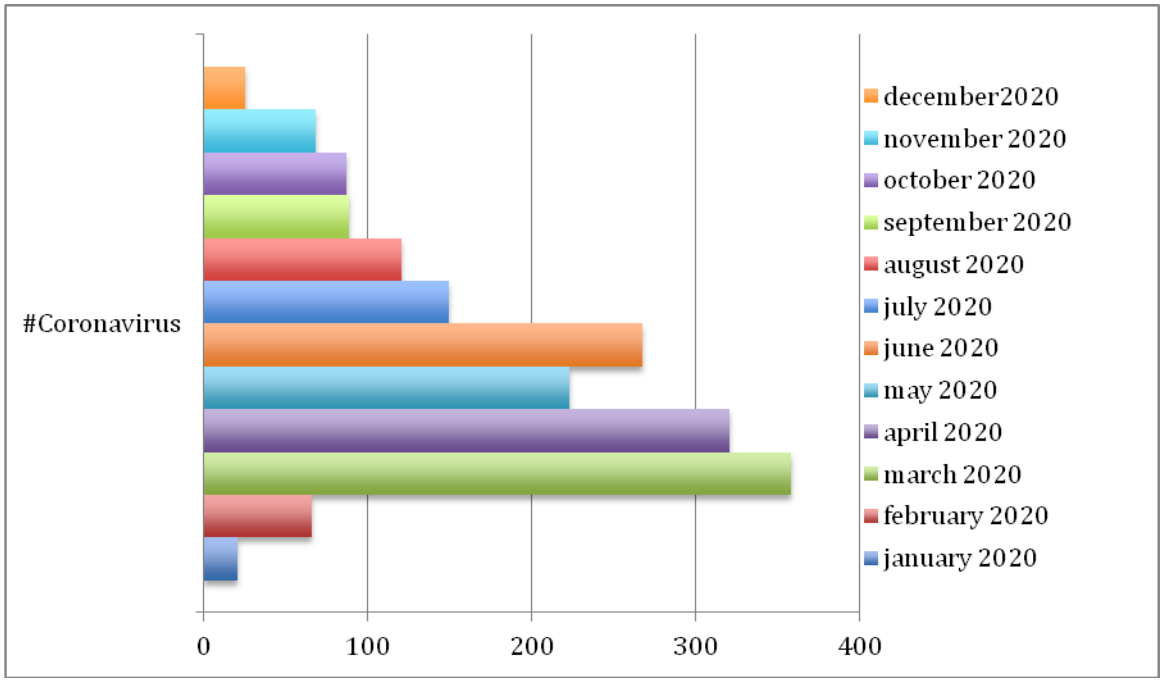


Fig. 2. Coronavirus Corpus interface



Picture 1. Dynamics of the use of the word “Coronavirus”

The next step was to estimate the frequency of use of new words in 2020.

Based on the results of Picture 2, we see that the words “Coronavirus” and “COVID-19” are on the list of the main contenders to become the words of the year. Coronavirus was used 1794.188 thousand times, COVID-19 — 2138.989 thousand times. Further descending are the words “Pandemic”, “Outbreak”, “Lockdown”, “Quarantine”, “Isolation”. The least popular words were “Furlough” — 14,248 thousand times, and “Flatten the curve” — 7,629 thousand times. This may be due to the fact that “Flatten the curve” was only discussed by health organizations and government authorities, as it was their job to develop a strategy to slow down the spread of the virus.

WORD	FREQUENCY, thousand times
#Coronavirus	1794,188
#COVID-19	2138,989
#Epidemic	66,390
#Pandemic	1148,138
#Stayathome	32,289
#Workfromhome	28,467
#Furlough	14,248
#Facemask	30,085
#Social distancing	218,105
#Physical distancing	30,825
#Flattenthecurve	7,629
#Symptoms	212,204
#Outbreak	356,715
#Lockdown	451,369
#Zoom	40,533
#Quarantine	206,255
#Isolation	107,397
#War	88,615
#Fighting	46,210

Picture 2. Comparison of the frequency of use of new words for 2020

TWITTER SENTIMENT ANALYSIS IN PYTHON USING TEXTBLOB

Sentiment analysis of tweets included the stages of preparation, app authentication and processing of statements. We needed two libraries to start sentiment analysis of tweets. The first one was Tweepy, a Python library for accessing the Twitter API. The other one, TextBlob, is a library for processing textual data. It provides a simple API for solving basic tasks in natural language processing: sentiment analysis, morphological parsing, nominative extraction, etc.

With the use of an authenticated account (API), we entered a search for specific words required for sentiment analysis. After choosing the keywords and compilation of tweets, we processed them using TextBlob.

For a start, we found the latest tweets associated with our keyword. Then we used TextBlob to determine the sentiment of each individual tweet and display it as shown in picture 3.

```
#Searching through tweets public_tweets = api.search('Coronavirus')
for tweet in public_tweets:

#Displaying tweets
print(tweet.text)

#Using TextBlob to determine the sentiment of a tweet
analysis = TextBlob(tweet.text)
print(analysis.sentiment)
print('\n')
```

Picture 3. Defining the sentiment through TexrBlob

Table 1. Twitter data on coronavirus per week

Date	Number of tweets #Coronavirus	Number of tweets #Covid-19	Total number
Thu 01/01/2021	44,357	49,183	93,540
Fri 02/01/2021	71,053	80,781	151,834
Sat 03/01/2021	16,399	26,231	42,630
Sun 04/01/2021	28,391	26,761	55,152
Mon 05/01/2021	31,775	26,747	58,522
Tue 06/01/2021	22,389	35,987	58,376
Wed 07/01/2021	34,229	35,949	70,178
Total per week			530,232

The program displays several tweets and analyzes their sentiment. Polarity determines how positive or negative the emotional colouring of the text is. Subjectivity measures to what extent the author's personal opinion is expressed in the text.

Despite numerous limitations of Twitter API, we have made consistent attempts to access as many posts as possible. About 530,000 tweets from around the world with the keywords #COVID-19 and #CORONAVIRUS were retrieved from Twitter (see Table 1).

Data collection was performed within one week, and the data for each day was stored in different files. Next, the tweets received from the Twitter were sent for sentiment analysis using the Python TextBlob library.

The sentiment polarity of tweets for seven consecutive days from January 1, 2021 to January 7, 2021, was observed. The data is displayed for two keywords, namely coronavirus and Covid-19. On average, more than 36% of users posted optimistic messages, while only about 14.5% of the tweets were negative. However, the neutral attitude towards the keywords coronavirus and COVID-19 was significantly high (49%).

The figures for different days are relatively similar. A large number of neutral tweets indicates that the bulk of the information included facts, not opinions. The next stage was to analyse the subjectivity of data in detail. For both keywords, most of the entries were objective, which accounted for approximately 64%. Meanwhile, about 22% of them were subjective, expressing feelings and opinions. Finally, 14% of tweets did not keep clear characteristics: neither subjective nor objective. This study unveiled the fact that Twitter users were being more optimistic when they used the word #COVID-19 rather than #Coronavirus in their tweets.

CONCLUSION

In this study, automated sentiment analysis of tweets on the topic of COVID-19 collected from Twitter database was applied. 818,224 coronavirus-related messages were collected during the high period of the pandemic from December, 2020 to January, 2022, and analyzed using Brand24, an online social data analysis service. 530,232 tweets that mention two hashtags #Coronavirus and #COVID-19 were collected during the week from January, 1 to January 7, 2021. Using the Twitter API and the flexible Python Tweepy library, the polarity and subjectivity of the tweets were identified. TextBlob, a library of Python sentiment analysis methods, was applied to the collected tweets. The results showed neutrality of most of the tweets, with a significant portion of entries being objective, which accounted for approximately 64%. However, the tone of the rest of the messages is negative. This data on the outbreak of the coronavirus known as COVID-19 showed us how people, state organizations, and media outlets interpreted the events.

One of the limitations of the current study is that Twitter users do not represent the entire population and only indicate the views and reactions of online users towards COVID-19. However, the Twitter dataset is a valuable source for understanding Twitter's real-time user-generated content in relation to actions against COVID-19.

A similarity was found between the results of text sentiment analysis using the Brand24 online service and the results obtained with Python using the TextBlob library: most of the messages were negative, i.e., statements on coronavirus were negative; however, it is worth noting that over 50% of the tweets were neutral, expressing no opinions, and, thus, not emotionally coloured.

However, many aspects are held aside for further research. In regard to sentiment analysis, in our paper we have chosen TextBlob, which uses the Naive Bayes classifier model, but it would be interesting to try other models that may lead to interesting results, such as vocabulary-based algorithms.

REFERENCES

1. Ainin, Sulaiman, Feizollah, Ali, Anuar, Nor Badrul, & Abdullah, Nor Aniza (2020). Sentiment analyses of multilingual

tweets on halal tourism. *Tourism Management Perspectives*, *34*, 100658.

2. Alsaeedi, Abdullah, & Khan, Mohammad Zubair (2019). A Study on Sentiment Analysis Techniques of Twitter Data. *International Journal of Advanced Computer Science and Applications*, *10*(2), 361–374. DOI: 10.14569/IJACSA.2019.0100248

3. Appel, Gil, Grewal, Lauren, Hadi, Rhonda, & Stephen, Andrew T. (2020). The future of social media in marketing. *Journal of the Academy of Marketing Science*, *48*, 79–95.

4. Bhat, Muza, Qadri, Monisa, Noor-ul-Asrar Beg, Majit Kundroo, Ahanger, Naffi, & Agarwal, Basant J. B. (2020). Behavior, & Immunity. Sentiment analysis of social media response on the Covid19 outbreak. *Brain, Behavior, and Immunity*, *87*, 136–137. DOI: 10.1016/j.bbi.2020.05.006

5. Bordia, Prashant, Hunt, Elizabeth, Paulsen, Neil, Tourish, Dennis, & DiFonzo, N. (2004). Uncertainty during organizational change: Is it all about control? *European Journal of Work and Organizational Psychology*, *13*(3), 345–365.

6. Choi, Sungwoon, Lee, Jangho, Kang, Min Gyu, Min, Hyeyoung, Chang, Yoon Seok, & Yoon, Sungroh J. M. (2017). Large-scale machine learning of media outlets for understanding public reactions to nation-wide viral infection outbreaks. *Methods*, *129*, 50–59.

7. Kaur, C., & Sharma, A. (2020). Twitter Sentiment Analysis on Coronavirus using Textblob. *EasyChair*, *2974*, 2516–2314.

8. Kim, Kyung-Sun, Sin, Sei-Ching Joanna, & Yoo-Lee Eun Young (2014). Undergraduates' use of social media as information sources. *College & Research Libraries*, *75*(4), 442–457. DOI: 10.5860/crl.75.4.442

9. Kim, Erin Hea-Jin, Jeong, Yoo Kyung, Kim, Yuyoung, Kang, Keun Young, Song, Min (2016). Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news. *Journal of Information Science*, *42*, 763–781.

10. Liu, Bing (2012). Sentiment Analysis and Subjectivity. In *Handbook of Natural Language Processing* (Second Edition).

11. Medford, Richard J., Saleh, S., Sumarsono, Andrew, Perl, Trish M., & Lehmann, Christoph U. (2020). An "Infodemic": Leveraging High-Volume Twitter Data to Understand Public Sentiment for the COVID-19 Outbreak. *Open Forum Inflectious Diseases*, *7*(7).

12. Pang, Bo, Lee, Lilian (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135.

13. Rodrigues, Ramon Gouveia, das Dores, Rafael Marques, Camilo-Junior, Celso G., & Rosa, Thierson Couto (2016). SentiHealth-Cancer: A sentiment analysis tool to help detecting mood of patients in online social networks. *International Journal of Medical Informatics*, *85*, 80–95.

14. Sailunaz, Kashfia, & Alhajj, Reda (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, *36*.

15. Soliman, Hatem, & Tabak, Fatema (2020). Sciences P. Deep learning framework for RDF and knowledge graphs using fuzzy maps to support medical decision. *Journal of International Research in Medical and Pharmaceutical Sciences*, *14*(3), 92–97.

16. Zarrad, A., Jaloud, A., Alsmadi, I. (2014). The evaluation of the public opinion — a case study: Mers-cov infection virus in KSA. In *IEEE International Conference on Semantic Computing* (pp. 664–670).