Study of the Problems of Determining Public Opinion of the Israeli-Palestinian War in Social Networks

Roman Lynnyk[1], Victoria Vysotska[1], Vasyl Andrunyk[1]

1 Lviv Polytechnic National University

Abstract

This paper explores the use of neural networks for determining public opinion on the Israeli-Palestinian conflict. In today's world, where numerous user opinions on various topics are posted daily on social media platforms, it is important to be able to analyze this public sentiment. The method of sentiment analysis and weighting of votes in comments is proposed, allowing for the identification of overall audience moods and assessment of their reactions to different events or content. The research results indicate the potential usefulness of these methods for marketing strategies, reputation management, and decision-making in various fields. However, there are challenges associated with data verification and consideration of manipulative interventions in public opinion. This work aims to review the current state of research in this area and identify prospects for further development. In today's digital world, analyzing public opinion on social media is becoming increasingly important. Understanding how users react to various topics and events is crucial for various fields, from marketing to politics. This paper discusses methods of sentiment analysis and vote weighting in comments using the example of analyzing public opinion on the Israeli-Palestinian conflict on the Reddit social network. The use of natural language processing tools, machine learning, and data visualization allows for valuable insights into users' emotional attitudes toward discussed topics and trends in public opinion. However, to achieve the best results, it is necessary to consider not only textual content but also social interaction and user status. The development of effective social media analysis tools opens the way to a deeper understanding of public opinion and making informed decisions in various areas of social life.

Roman Lynnyk, **Victoria Vysotska**, and **Vasyl Andrunyk**

Lviv Polytechnic National University, Lviv, Ukraine

Keywords: Public opinion, Sentiment analysis, Vote weighting, Social media, Israeli-Palestinian conflict, Machine learning, Data visualization.

Introduction

In today's world, where numerous user opinions on a variety of topics are published daily, from global events such as wars and international conflicts to domestic political processes including parliamentary elections and presidential campaigns, the ability to analyze public opinion takes on special significance. Social networks provide unprecedented opportunities to detect moods, views, and priorities of society, but they also present challenges associated with processing vast volumes of unstructured data.

The issue of determining public opinion on social networks is relevant and requires continuous updating of approaches and methodologies due to the rapid development of digital technologies. By analyzing comments and posts of users, one can not only understand general trends but also identify changing motives and emotional underpinnings that influence the formation of public opinion. This opens a wide field for research and the development of tools that would allow effective collection, classification, and analysis of data, ensuring accurate and objective representation of societal sentiments.

Work on decision support systems based on the analysis of public opinion in social networks is particularly relevant in the context of the need for a rapid response to social changes. The use of machine learning, natural language processing, and other artificial intelligence methods enables not only the detection of general trends but also the prediction of possible shifts in public opinion, which can be useful during electoral campaigns, social campaigns, or crisis management.

However, along with the great potential, the analysis of social networks opens up a range of challenges and difficulties. The main problem that arises in the context of social media data analysis is related to information verification. It is not always possible to definitively determine whether the opinions published reflect the real beliefs of users or are the result of manipulative interventions such as trolling or the use of bots. This challenge becomes even more relevant when we consider the potential of social networks for the mass dissemination of information and its impact on public opinion. Additionally, another characteristic problem in determining public opinion is considering not only the comment itself but also how many people support or oppose it, and how well-verified the user's profile is to be able to take their opinion into account.

Taking these aspects into account, this article is intended to examine the current state of research in this field, assess the potential and challenges of existing software and hardware systems, and identify prospects for further development directions. The presentation of the main material of the article includes an analysis of key methodologies and approaches, the application of machine learning algorithms for text analysis, and a discussion of problems that arise in the context of using these technologies

Formulation of the problem

In the modern world, where numerous user opinions on a variety of topics are published daily, from global events like wars and international conflicts to domestic political processes including parliamentary elections and presidential campaigns, the ability to analyze public opinion becomes particularly relevant. The widespread accessibility of the internet

and social networks, where every user can express their opinion on various events, phenomena, products, services, or ideas, enhances this relevance. However, despite the obvious value of this information, analyzing and interpreting large volumes of textual data from social networks requires the development of effective methods and tools. This involves not only text processing technologies but also an understanding of the social processes occurring within online communities.

The main goal of social network analysis is to determine the sentiment score of a specific comment, that is, to identify its emotional colouring or mood. This task is fundamental because the emotional tone of the text can provide valuable information about users' attitudes towards the topics discussed. Sentiment analysis technologies are used for the automatic classification of texts by emotional charge, allowing conclusions to be drawn based on large datasets without the need for manual review.

A second task involves considering the weight of likes and dislikes received by the comment. These indicators can play an important role in determining the overall community response to an opinion or message. They allow not only for assessing the popularity or unpopularity of certain statements but also for identifying those that have provoked the most controversy or discussion.

The third task includes analyzing the karma of the user to consider the weight of their opinion depending on their status in the community (for example, whether the user is verified or a regular participant in discussions). Verified users or those with high karma are often considered more reliable information sources, and thus their opinions may carry more weight in public opinion analysis. Determining the influence of karma on the perception of a comment helps to understand how much the community trusts the opinion of a particular user and assess its significance in the context of the overall discussion.

The fourth task, analyzing the trends of public opinion for each post and its changes according to the day, is key to determining the dynamics of topic perception by the audience. This allows us not only to identify general trends in reactions to publications but also to track how public opinion changes over time, responding to new events, information, or arguments. Using dated metadata allows these trends to be associated with specific events or publications, providing a deeper context for interpreting the data obtained.

Addressing these tasks requires the integration of various methodologies and technical solutions, including but not limited to natural language processing algorithms, machine learning, statistical analysis, and data visualization techniques.

Framing the problem in analyzing public opinion in social networks requires a deep understanding of both the technical capabilities of modern data analysis tools and the social processes occurring in online communities.

Developing effective methods to address the above tasks will not only improve understanding of public opinion but also facilitate the development of tools for making informed decisions in fields such as politics, marketing, social research, and more.

In this work, we will analyze comments under posts on the social network Reddit and explore trends in support of the Israeli-Palestinian war and how public opinion has changed over a certain period. In this context, integrating analytical

tools that consider not only the textual content of comments but also the social interaction of users (such as likes and dislikes) and the status of users, such as karma (or whether the user is verified), is key to developing comprehensive and reliable social network analysis systems.

Thus, the problem statement includes not only the development of technical solutions for analyzing textual data but also an understanding of how these data can represent public opinion regarding the Israeli-Palestinian war.

Analysis of recent research and publications

In today's world, where millions of user opinions on various topics such as wars and country support, parliamentary and presidential elections, and discussions on issues critical to society are published daily, the importance of quickly and accurately determining public opinion is growing. Social networks play a special role in this process, having become a powerful tool for gauging public sentiment. Platforms like Twitter, Facebook, Instagram, and others are used daily by millions of people to express their opinions and positions on current events. This opens up new opportunities for big data analysis and the application of machine learning technologies to determine general trends in public opinion.

However, despite the importance of social networks as a source of information about public opinion, there are several challenges and difficulties associated with their analysis. First and foremost, the vast amount of data generated daily by users requires the development of effective text processing and analysis algorithms. Additionally, it is necessary to consider the diversity of languages, slang, irony, and other linguistic features that can complicate the correct interpretation of posts. Another complexity is the identification and filtering of fake accounts and bots that can distort the real picture of public opinion.

Despite these challenges, the application of artificial intelligence and machine learning technologies shows significant potential in solving public opinion analysis tasks. In particular, neural networks and deep learning algorithms allow for the effective processing of large volumes of textual data, identifying key themes, emotional assessments, and trends in user opinions. Natural language processing (NLP) methods also play an important role, helping to better understand the semantics of text and its context.

The relevance of studying public opinion on social networks is increasing every year, considering their role in shaping political, social, and economic narratives. Therefore, developing new approaches and tools for analyzing social networks becomes a key task for researchers, marketers, political analysts, and anyone who seeks to better understand and predict the dynamics of public sentiments.

In the article titled "Monitoring Public Opinion through Collective Semantic Analysis of Tweets," published in the journal Social Network Analysis and Mining (2022), the authors present a new automated mechanism for monitoring public opinion using Twitter. This method employs advanced natural language processing (NLP) technologies and deep neural networks (DNN) for semantic analysis of tweets. The mechanism consists of a semantic descriptor that quantitatively determines the polarity, offensiveness, bias, and figurativeness of each tweet independently [1].

These descriptors are then aggregated according to the desired strategy and goal, facilitating various applications such as forecasting trends in public opinion or evaluating the effectiveness of political campaigns.

The innovative aspect of the mechanism is its multidimensional approach to semantic analysis, which is a significant advancement compared to previous methods that primarily focused on sentiment analysis. By incorporating additional dimensions such as bias and figurativeness, this approach offers a richer and more nuanced understanding of public opinion. The article demonstrates the application of the mechanism to datasets of tweets from the 2016/2020 U.S. presidential elections, showing its potential to provide insightful and concise descriptions of public opinion.

Furthermore, the article discusses the challenges and limitations of modern NLP and DNN technologies in accurately capturing and interpreting the multifaceted nature of public opinion. The authors emphasize the importance of developing more complex algorithms that can handle the complexity and subtleties of human language, especially in the context of social networks, where expressions of opinion are often concise and colloquial.

The proposed mechanism is a significant contribution to the fields of political science, social media analysis, and computational linguistics. It highlights the potential of artificial intelligence technologies to transform our understanding of public opinion, offering tools that can analyze large volumes of data quickly and accurately. However, the authors also note ethical considerations and the need for transparency and responsibility in the use of these technologies, especially considering privacy and the potential for manipulation.

In the article titled "Intelligent Analysis of Ukrainian-Language Tweets for Researching Public Opinion Based on NLP Methods and Machine Learning Technologies," published in the journal I.J. Modern Education and Computer Science (2023), researchers present a new automated approach to analyzing public opinion through social media. This method applies advanced natural language processing (NLP) technologies and deep neural networks to analyze the linguistic content of tweets. The research includes a comprehensive analysis of data collection methods, their filtering, cleaning, and preprocessing for further study of public sentiments.

An innovative aspect is the use of technologies such as BERT vectorization and HDBSCAN classification, which allow for more accurate identification of thematic and emotional clusters in large volumes of textual data. The article also demonstrates how the application of these methods to the analysis of tweets in Ukrainian can help detect trends in public opinion and responses to socially significant events [2].

However, as emphasized in the article, the use of such complex technologies requires significant computational resources and may involve data selectivity, as Twitter users may not fully represent all demographic groups in society. The authors of the article discuss challenges related to natural language processing, such as the need to process unstructured text that includes colloquial language, irony, and jargon, which are characteristic of social media.

The proposed methods and approaches are a significant contribution to the development of tools for analyzing social networks and can be used to support decisions in politics, social affairs, and business. However, the authors also point to the need for further research to optimize algorithms and methods to reduce computational resource requirements and increase the accuracy of results.

In the article titled "Classification of Sentiments in Contemporary Public Opinion on BREXIT: A Comparison of the Naive Bayes Classifier Model with Python's TextBlob Approach," a study is presented that compares the effectiveness of two sentiment analysis methods using social media data. This research utilizes natural language processing (NLP) and machine learning technologies to analyze public sentiments regarding BREXIT, using data from Twitter.

The authors of the article focus on the fact that with a large number of users on social platforms such as Twitter, there arises the opportunity to classify public opinions about various events or products. In particular, the article discusses the use of the Naive Bayes algorithm and Python's TextBlob library for determining the emotional tone of tweets. The study found that the Naive Bayes model demonstrates better results compared to the TextBlob approach in terms of accuracy, which is significant for the British and Irish governments in shaping domestic and foreign policies.

A distinctive feature of the work is the use of visualization tools, such as Tableau, to represent the collected data, allowing for a deeper analysis of changes in public opinion and reactions to political events. The article also discusses future research directions, including the use of more complex NLP models and machine learning to improve the accuracy of public opinion analysis.

The authors assert that the research results can be useful for governments and organizations that aim to understand public sentiments and respond more effectively to public challenges and changes in societal mood.

Further, we will consider the research presented in the article "Managing Student Public Opinion on Campus Based on Deep Learning" by Ling Shen and Minhong Xu, which focuses on analyzing student opinions and crisis management of public opinion in colleges and universities. The article discusses the challenges faced by higher education institutions due to the informal structure and lack of standardization of short text comments on social media, which complicate the extraction of textual features for sentiment analysis.

The authors propose an LSTM model for the deep semantic characteristics of the text, which makes it possible to accurately determine emotional trends [3]. The main tasks of the model are to use CNN to better extract local features of the text and LSTM to preserve information about the history of the text, which contributes to the effective highlighting of global features of the strategy.

The application of deep learning proves to be particularly useful in the context of multi-campus institutions, where managing student organizations and public opinion among students is becoming an increasing challenge. The article demonstrates how improved methodologies using deep learning models can help manage public opinion, ensuring timely and effective tracking of information.

In addition to implementing LSTM and CNN, the article explores the use of the CBOW model to optimize the focus on feature vectors that impact classification outcomes. The enhanced model in the article is compared with traditional methods using metrics such as accuracy, recall, loss, and F1 score, which indicate the performance evaluation of the model.

This work is important for understanding how deep learning can be applied to improve the management of public opinion in academic institutions, enhancing the quality and accuracy of the analysis of student comments and feedback.

Additionally, we can consider the methodology proposed in the research by Baojun Ma and co-authors, which uses probabilistic topic modelling and deep learning to analyze public opinion published on social media. This methodology includes several important steps: from gathering information to analyzing it using time series, which allows tracking the dynamics of changes in the emotional state of the public [4].

The first step involves collecting data from social networks and preprocessing this data, which includes removing noise and extraneous information that could distort the results of the analysis. After data processing, the Latent Dirichlet Allocation (LDA) model is used to identify the main topics discussed by users. This allows the identification of key issues and areas of interest across a large amount of textual data.

The second step focuses on analyzing the emotional intensity of texts using the word2vec model, which trains word vectors and determines the emotional colouring of texts. This approach allows not only to recognise the general emotional tone of statements but also to detect changes in emotions that may indicate a shift in public moods or a response to certain events.

The final step includes a time analysis of emotional intensity, which allows observing the dynamics of public opinion and emotions in response to the development of social events. This aspect is particularly important for understanding how certain events affect public opinion and how moods in society change over time.

The use of this comprehensive methodology within the analyzed study has shown its effectiveness and usefulness for a deep understanding of the structure and dynamics of public opinion in social networks. Thanks to this approach, it is possible not only to analyze the current state of affairs in public consciousness but also to predict possible changes, which is valuable for developing strategies in the field of social management and communications.

Purpose of the article

The goal of the research is to develop and apply a comprehensive methodological approach that uses natural language processing algorithms and social network analysis to accurately determine and interpret public opinion based on the dynamics of social reactions to the Israeli-Palestinian conflict. The objective of the study is to identify, classify, and analyze public opinions, emotions, and attitudes that are expressed in comments on social networks, with a special focus on their evolution over time.

The subject of the study is comments and posts on the social network Reddit related to the Israeli-Palestinian war. In this work, we analyze public opinion using sentiment analysis to determine the emotional colouring of comments. We also consider user interactions, such as likes and dislikes, as well as the status of users, including their karma, to understand the weight of their opinions within the community.

The main tasks include determining trends and dynamics of public opinion, through time analysis of reactions to publications and identifying changes in public opinion that occur over time, and the relationship of these changes to specific events or publications.

To visualize the change in user opinion trends, we will create a graph that will show how public opinion about the Israeli-Palestinian war changes over a certain period linked to various comments under posts on Reddit. This will allow for a deeper understanding of the structure and dynamics of public moods, to improve tools for making informed decisions in politics, marketing, social research, and other fields.

The main algorithm process part

In this work, a sentiment analysis algorithm using machine learning, specifically text classification methods, was utilized to analyze comments on Reddit. The main idea was to train the model to recognize the tone of the comments (positive, negative, or neutral).

The first step involved preparing a training dataset, which consisted of the text of the comments and corresponding sentiment labels (positive, negative, or neutral). This dataset was divided into training, validation, and test samples for effective model training and evaluation.

For this study, a dataset on the Israeli-Palestinian war was used, with the following key columns:

1. Text of the comment (self_text): This column contains the text of the comment itself, which will be analyzed to determine its sentiment.
2. Date and time of the comment (created_time): This column contains information about the date and time the comment was published. This information can be useful for analyzing the dynamics of sentiment over time or understanding the impact of events on sentiment.
3. Comment identifier (comment_id): A unique identifier for each comment, which can be used to reference a specific comment if necessary.
4. Post identifier (post_id): Used to link comments to a specific post, allowing us to see classified information on whether the Israeli-Palestinian war is supported or not according to each post.
5. Likes under the comment (ups): Allows us to analyze the bias of the corresponding weights according to how many people support a particular opinion under the comment.
6. Dislikes under the comment (downs): Similar to ups, but with the weighting shifted in the opposite (negative) direction.
7. User verification (user_is_verified): This allows us to more accurately analyze public opinion, as we can filter out users who are not verified.
8. User karma (user_awardee_karma): Also one of the key attributes of this dataset, which allows us to shift weights according to the karma in the user's profile.

After analyzing this dataset, we will proceed with building the actual algorithm to assess opinions about the Israeli-

Palestinian war over a certain period and how this opinion has changed through sentiment analysis.

The main algorithm involves the following steps:

1. Load data about the Israeli-Palestinian war and read it into a pd dataframe.
2. Preprocess the data, filling missing values in the columns 'self_text', 'post_self_text', 'post_title' with empty strings, and converting the 'created_time' column into a datetime object.
3. Perform sentiment analysis using the SentimentIntensityAnalyzer from the NLTK library to determine the compound sentiment of each comment, normalizing the results to the range [0,1].
4. Account for the vote weight (likes/dislikes) under comments, calculating it as the proportion of likes to the total number of votes, or 0 if the total number of votes is negative or zero.
5. Group data by posts and dates, grouping data by the identifier of the post by the date the comment was created.
6. Calculate the average sentiment for each post: compute the average normalized sentiment for each post based on grouped information.
7. Visualize the graph, where a line graph is created for each post showing the change in sentiment according to the time series. Each post is marked with its unique identifier.

Next, we will build a diagram of activities, illustrating the main actions performed by the algorithm and displaying it in Fig.1.

The algorithm starts with data reading, then preprocesses it, performs sentiment analysis on the comments, normalizes the sentiment scores, calculates the vote weight, aggregates the sentiment ratings by comments, sorts the posts by date, creates time-series sentiment graphs for the messages, and finally completes the process. Each action transitions sequentially from one to the next until the process is completed.
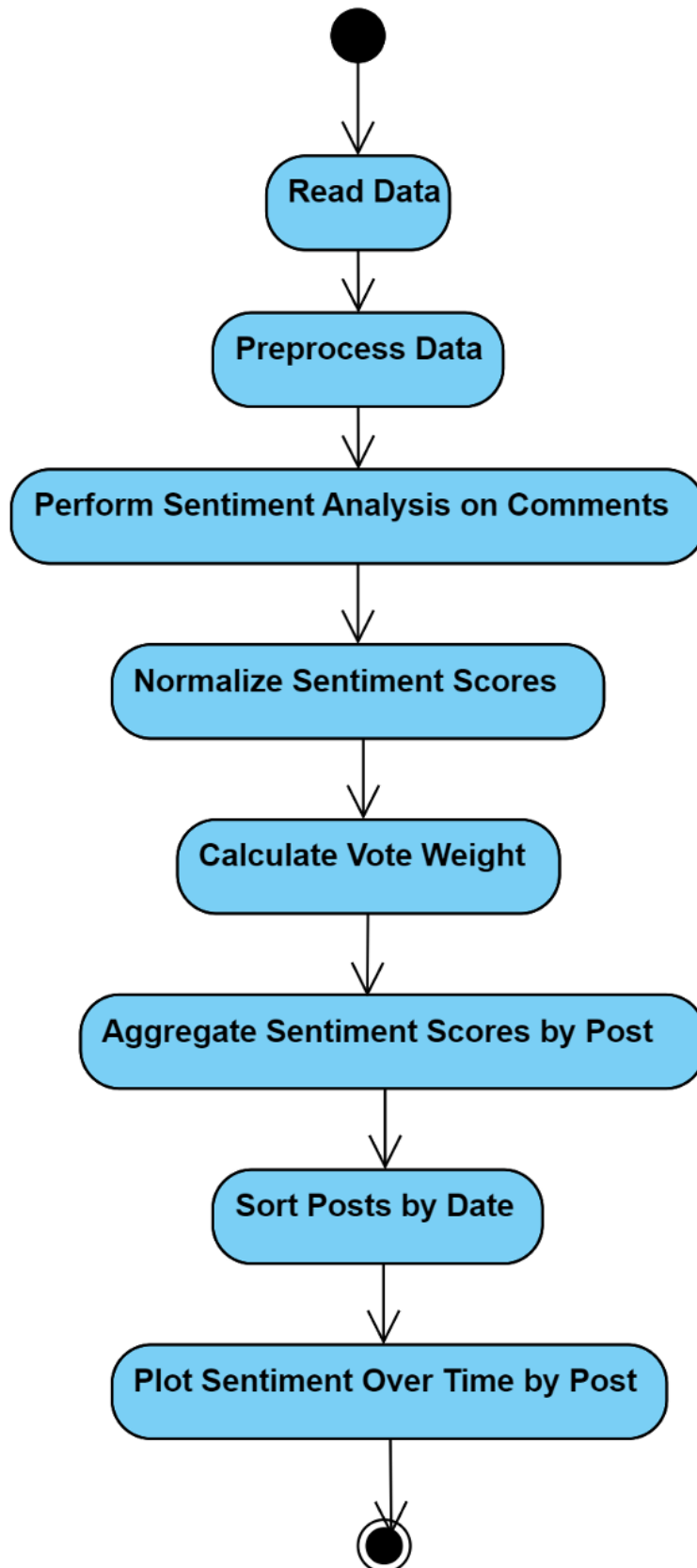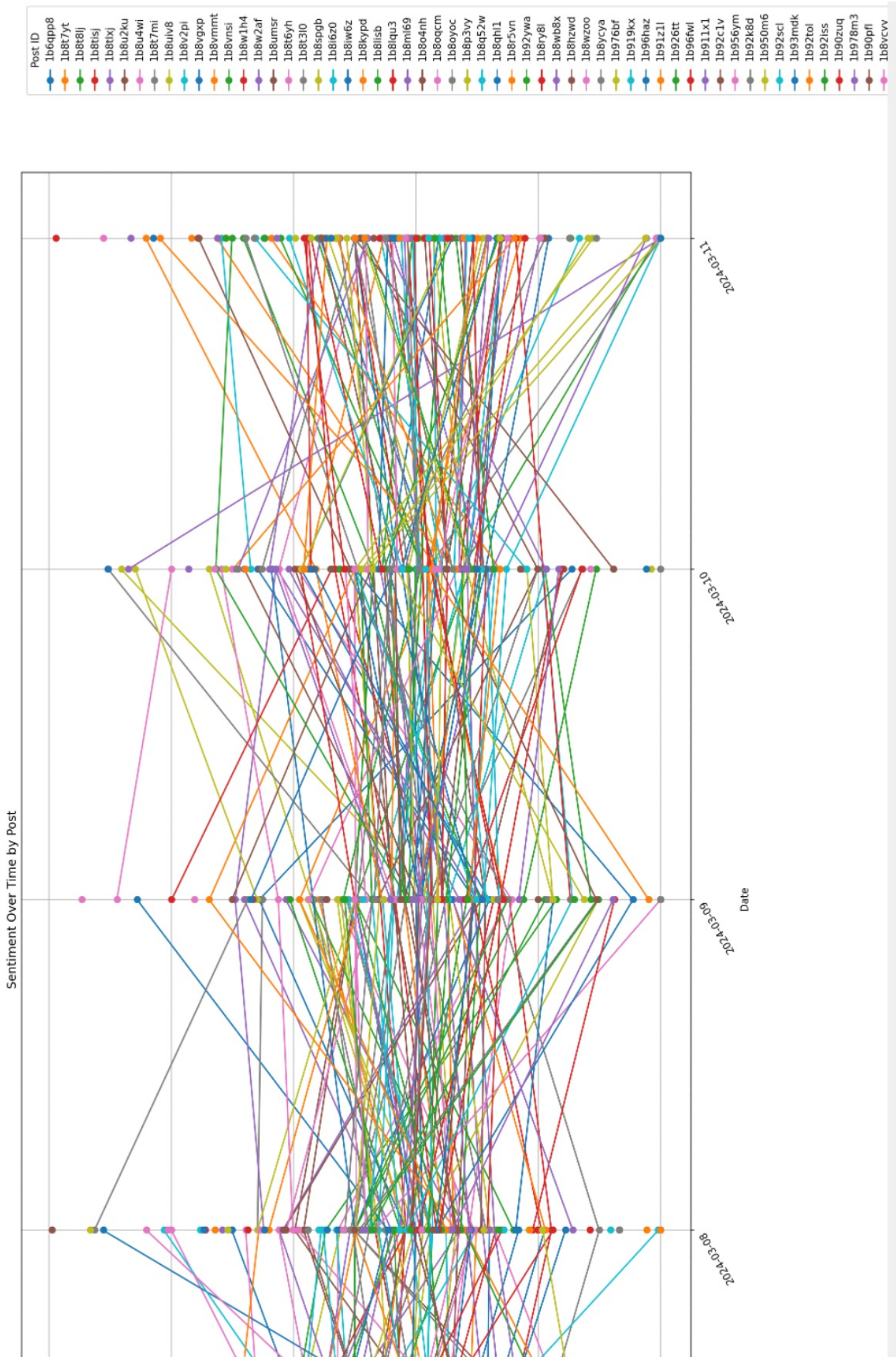
Fig. 1. *Diagram of the operation of the algorithm*

After executing this algorithm, we will get the following result, which is shown in Fig. 2.:
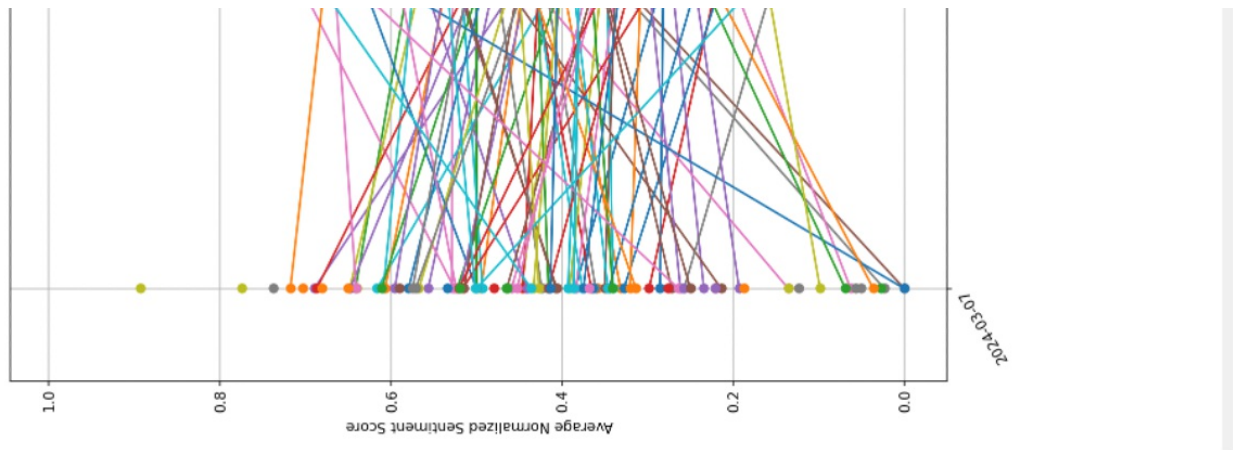
Sentiment Over Time by Post

Fig. 2. *Sentiment value in accordance with the time series*

In this chart, we can see the average normalized sentiment value according to the aggregated data of each comment according to the post, each post is marked with its color and has a unique identifier, and we can see the time series and how the general trend of changing opinion about the Israeli-Palestinian war has changed. In general, we can conclude that the majority of people either do not support this cruel and unjustified war or have a neutral attitude towards it. But, unfortunately, there are of course those who support it.

Of course, we can also analyze one post at a certain time interval, we will confirm it for the post with the unique identifier "1bc67nd" in Fig. 3.
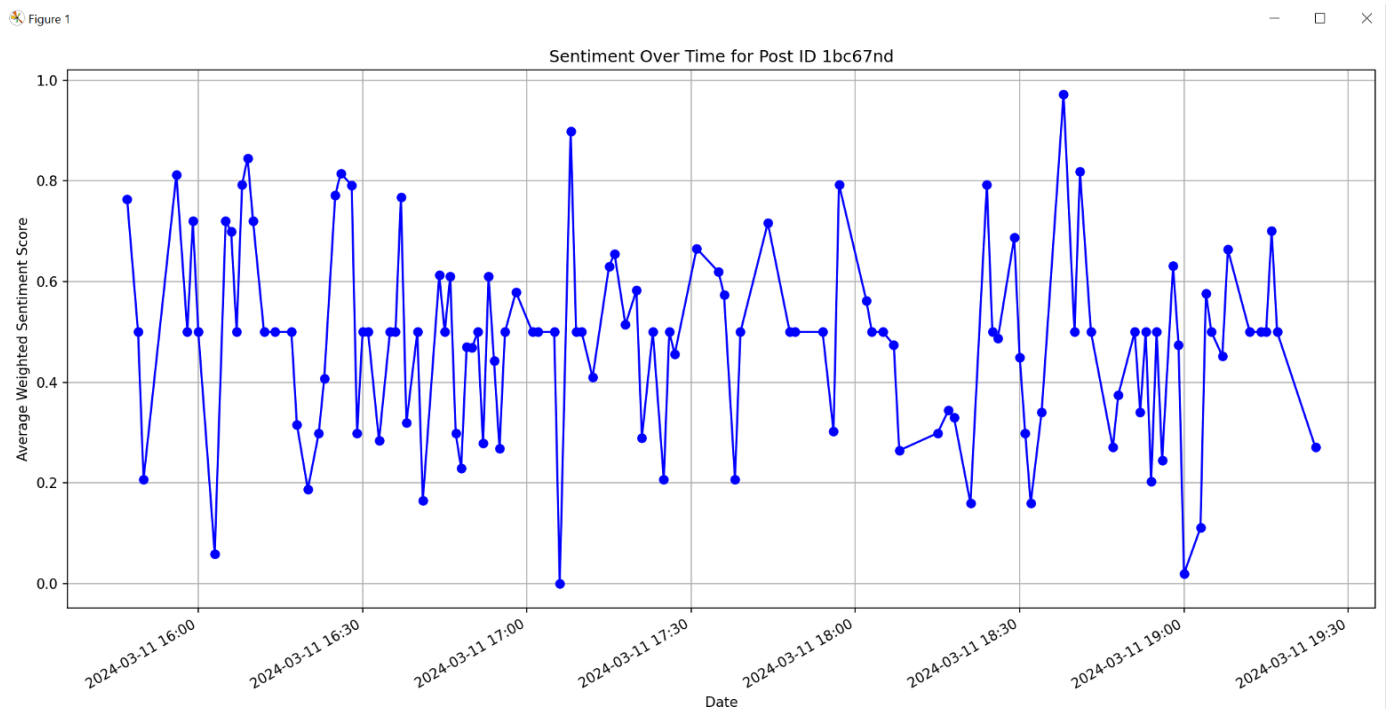

Fig. 3. *Sentiment value at a certain time interval for a specific post*

As we can see, opinions could differ quite often, which means that the post itself could be quite controversial or there were

external factors that influenced the opinions of the users themselves.

Of course, this issue is important and requires further in-depth research, but this work is a good basis for further research.

Conclusions

During the study, the possibility of using user feedback analysis on social networks to assess sentiment regarding specific topics or events was examined. The work included processing textual data, using the VADER algorithm for sentiment analysis, and calculating the weight of votes for each comment. The results showed that normalized sentiment scores along with vote weight can be an effective tool for determining the overall user sentiment regarding specific content.

By analyzing the data, it was found that this algorithm not only allows tracking changes in sentiment over time but also compares sentiment across different posts or topics. It was also discovered that including the weight of votes enhances the objectivity of the results, as it considers not just the act of commenting but also the users' activity in shaping the feedback.

Thus, the results of this work indicate the potential usefulness of applying sentiment analysis and vote weighting in social networks for understanding audience moods and evaluating their reactions to various events or content, in this case, the level of support for the Israeli-Palestinian war, its analysis, and understanding how the world views the events between these two countries.

Other References

- Bhupender Singh Shekhawat, Vladimir Milosavljevic (2019). Sentiment Classification of Current Public Opinion on BREXIT: Naïve Bayes Classifier Model vs Python's TextBlob Approach.

References

1. ^*Karamouzas, D., Mademlis, I. & Pitas, I. Public opinion monitoring through collective semantic analysis of tweets. Soc. Netw. Anal. Min. 12, 91 (2022). https://doi.org/10.1007/s13278-022-00922-8*

2. ^*Oleh Prokipchuk, Victoria Vysotska, Petro Pukach, Vasyl Lytvyn, Dmytro Uhryn, Yuriy Ushenko, Zhengbing Hu, "Intelligent Analysis of Ukrainian-language Tweets for Public Opinion Research based on NLP Methods and Machine Learning Technology", International Journal of Modern Education and Computer Science(IJMECS), Vol.15, No.3, pp. 70-93, 2023. DOI:10.5815/ijmecs.2023.03.06*

3. ^*Ling Shen, Minghong Xu, "Student Public Opinion Management in Campus Commentary Based on Deep Learning", Wireless Communications and Mobile Computing, vol. 2022, Article ID 2130391, 12 pages, 2022. https://doi.org/10.1155/2022/2130391*

4. ^*Baojun Ma, Hua Yuan, Yan Wan, Yu Qian, Nan Zhang " PUBLIC OPINION ANALYSIS BASED ON PROBABILISTIC*

TOPIC MODELING AND DEEP LEARNING", PACIS 2016 PROCEEDINGS. http://www.pacis2016.org/