OPINION MINING AND SENTIMENTAL ANALYSIS OF TWITTER DATA DURING INDIA'S GENERAL ELECTIONS

Sai Keerthiga Murugan

Student, Department of Computer Science and Engineering

National Institute of Technology Puducherry, Karaikal, India

Ankita Bhowmik

Student, Department of Computer Science and Engineering

National Institute of Technology Puducherry, Karaikal, India

Dr. V. Vani

Assistant Professor, Department of Computer Science and Engineering

National Institute of Technology Puducherry, Karaikal, India

Dr. N Karthik

Assistant Professor, Department of Computer Science and Engineering

National Institute of Technology Puducherry, Karaikal, India

Abstract: Social Networking like Twitter have become significant sources of political discourse, offering valuable insights into public sentiment and attitudes during elections. In this research, analyzing emotions in tweets related to political figures Party 1 and Party 2 with the goal of predicting public sentiment and potentially forecasting election outcomes. Our study focused on extracting emotional tones from tweets, classifying them as positive or negative, and then assessing how these emotions relate to each candidate's popularity. Data preprocessing was a crucial initial step in preparing the Twitter data for analysis, involving tasks such as text cleaning and tokenization. We used several machines learning algorithms, including k-Nearest Neighbors (KNN), Random Forest, Multinomial Naive Bayes, Gaussian Naive Bayes, Random Forest to build sentiment classifiers. These models were trained on labeled data to distinguish between positive, negative, and neutral emotions expressed in tweets. Our findings revealed that the Random Forest and KNN model yielded the more accurate in sentiment classification. For Party 1, this model achieved the accuracy of 71%, while for Party 2, it achieved the accuracy of 70%. Based on the sentiment expressed in tweets, these results imply that the Random Forest model could serve as a dependable predictor of public sentiment towards these political figures. This research not only shows the potential of analyzing emotions based on social media data to gather public sentiment during elections but also highlights the significance of emotion-driven insights for political campaigns and electoral strategies. The potential of the Random Forest model for prediction could be valuable in gaining insights into public sentiment and making informed choices for campaign adjustments or foreseeing election results.

Key words: Sentiment Analysis, Twitter Elections, Emotional Emotion, Data Preprocessing, Algorithms, k-Nearest Neighbors (KNN), Random Forest, Multinomial Naive Bayes, Gaussian Naive Bayes, Accuracy, Positive and Negative Sentiment, Twitter; tweet

1. Introduction

The intersection of politics and technology has witnessed a transformative evolution in lately, with social media platforms emerging as powerful tools that shape the political landscape. The 2019 General Elections in India stand as a poignant testament to the influence of digital platforms, particularly Twitter, in shaping the political narrative and

influencing public opinion. This research paper embarks on a comprehensive analysis of the pivotal role played by social media in these elections, focusing on Twitter as the primary medium for political discourse.

The 2019 Indian General Elections were not a mere electoral event but a watershed moment in the political history of the world's largest democracy. As over 900 million, the political contenders vied not only for qualified voters cast their votes, traditional votes but also for the attention, approval, and endorsement of the massive online populace. Social media platforms served as the agora where political debates, campaigns, and discussions unfolded in real-time, effectively reaching millions of citizens, including the urban youth and the rural masses.

Twitter, a microblogging platform with its characteristic 280-character limit, became a central arena for political communication. It enabled politicians, government officials, and opposition leaders to bypass traditional media gatekeepers and communicate directly with their constituents and the public. In this context, it is imperative to test the usage of Twitter as a political tool and assess how the emotional tone of tweets and the sentiments they convey the larger political discourse.

To undertake this analysis, this research paper employs a multi-faceted approach. It begins by addressing the crucial task of data preprocessing, encompassing text cleaning and tokenization, to be guarantee the integrity of the tweets suitable for rigorous analysis. Subsequently, a range of machine learning algorithms, including k-Nearest Neighbors (KNN), Random Forest, Multinomial Naive Bayes, and Gaussian Naive Bayes, are deployed to construct sentiment classifiers. These classifiers are trained on labeled data, enabling the distinction of diverse emotional tones found within the vast Twitter dataset.

The statement of this study revolves around the significance of sentiment analysis in the sphere of political discourse on social media. We aim to shed light on how Twitter users' emotional responses to political events, candidates, and issues reflect broader public sentiment. Furthermore, by examining the predictive potential of sentiment analysis, we intend to explore whether the emotional content of Twitter info can serve as a tool for forecasting election outcomes or as a barometer for gauging public sentiment.

During our analysis, the research paper uncovers intriguing insights, with the Random Forest model emerging as the most accurate sentiment classifier, achieving a precision rate of 71% for tweets associated with the incumbent Prime Minister, Narendra Modi. Additionally, the KNN model demonstrates its effectiveness, achieving a 70% precision rate for tweets is related to the leader of the opposition, Rahul Gandhi.

The Discovery of this study not only emphasize the profound impact of sentiment analysis on understanding public sentiment in the context of political discourse on social media but also underscore the potential of these insights to influence political strategies, campaign approaches, and possibly, the forecasting of election outcomes. They reaffirm the inextricable link between technology, politics, and public opinion in the digital age and underline the need for a deeper understanding of these dynamics.

This research paper strives to contribute to the ongoing discourse on the evolving nature of political communication in the age of social media. It illuminates the transformative potential of digital platforms, the utility of sentiment analysis in analyszing the politics, and the need for political actors.Effectively utilize the potential of social media in an ever-evolving political environment.

1.1.

The primary achievements of this research can be outlined as follows:

- Insights into the 2019 Indian General Elections through Twitter data analysis.

- Application of sentiment analysis to political context.

- Demonstration of sentiment analysis's predictive capabilities.

- Relevance of emotional insights for political campaigns and strategies.

2. The Related Literature

In the literature review, we looked at what researchers have found about using sentiment analysis to understand how people feel about politics on social media, especially on Twitter. We found that sentiment analysis is important for figuring out public opinions. Social media, like Twitter, has changed how people talk about politics. We studied how sentiment analysis was used in elections, including India's 2019 General Elections. We also checked out different computer methods used in sentiment analysis. Some studies showed that sentiment analysis can predict election results or help politicians make better decisions. We also looked at the ethical concerns and problems with using sentiment analysis in politics. Overall, we found that understanding public opinions on social media, like Twitter, is a big deal in politics. It helps politicians and researchers like us understand what people want. Our study fits into this field by focusing on sentiment analysis in the context of political discourse on social media, especially during elections.

[1]. Predicting election outcomes using Twitter is a popular yet challenging trend in research. While appealing, it often prioritizes positive results over sound methods. Recent papers tend to favor studies supporting Twitter's predictive power, neglecting balanced reviews. In response, this paper critically assesses all relevant research, concluding that Twitter's predictive ability has been overstated, emphasizing the need for more rigorous research and methodology.

[2]. Exploring the challenges of extracting valuable insights from social media data for election predictions is done in this research. We introduce a unique data collection approach that combines RSS news feeds and Twitter trends. Using tweet volume and sentiment analysis, we construct a predictive model. We review existing techniques, analyze factors impacting predictions, and found that social media data surpasses traditional exit polls in accuracy. This study underscores the importance of data collection methods and timing for improved electoral predictions and meaningful insights.

[3]. this research uses Twitter data to analyse the Indonesian Presidential election. Tweets from candidates Jokowi and Prabowo, along with relevant hashtags, were analyzed for sentiment from March to July 2018. An algorithm was developed to extract essential data and train a sentiment model. Results in R language show Jokowi leading, aligning with four Indonesian survey institutes, demonstrating the method's reliability.

[4]. this study assesses the effectiveness of sentiment analysis and social network methods in forecasting important decisions based on social media data. It specifically concentrates on utilizing sentiment analysis to predict election results by gauging the public mood on platforms such as Facebook and Twitter. The paper examines previous research contributions, tackles obstacles, and offers recommendations for future developments in using social media content for election prognostication.

[5]. In this study, 42,235 Hindi tweets related to Indian political parties during the 2016 general state elections campaign were evaluated using sentiment analysis. Various methods, including Naive Bayes, SVM, and Dictionary-Based approaches, were used to categorize the tweets into positive, negative, or neutral sentiments. The SVM approach predicted a 78.4% likelihood of the BJP's victory, which was subsequently confirmed when the BJP won sixty out of 126 constituencies, surpassing other political parties.

[6]. in this research, Twitter data and analysis of emotions are employed to forecast election outcomes by examining the balance between positive and negative messages. Through the use of an aging estimation technique, it achieves remarkably precise predictions for the 2020 United States presidential election within a span of four months.

[7]. this research uses sentiment analysis to assess public sentiments during the COVID-19 pandemic by analyzing social media content. The research employs keywords like "COVID," "Corona virus," and "COVID-19" to determine whether sentiments are positive, negative, or neutral.

[8]. this survey explores the extensive research conducted on Twitter, the third most popular Online Social Network (OSN) globally. It highlights three key areas: social graph analysis, sentiment analysis, and addressing threats like spam, bots, fake news, and hate speech. The survey covers Twitter's data model, data access best practices, and computational techniques, including Graph Sampling, Natural Language Processing, and Machine Learning. By summarizing key findings and discussing the current state of research, this paper serves as a guide for researchers delving into Twitter's research landscape.

[9]. in this research, Twitter served as a source of opinionated data, and tweets were gathered from January 2019 to March 2019. Sentiment analysis was then conducted on these tweets to determine public sentiment regarding the Indian general elections. The study found that Candidate-1 was more popular and well-received compared to Candidate-2. Importantly, the research results aligned closely with the actual election outcomes in May 2019. Limited Discussion of Limitations, No Machine Learning Approach, Limited Scope of Sentiment Dimensions, Limited Scope of Sentiment Dimensions, Limited Scope of Sentiment Dimensions, Lack of Contextual Analysis.

[10]. While pre-election polls have long been used to predict outcomes, this paper argues for the effectiveness of Twitter-based election forecasting, particularly in developing countries. Using Indonesia's 2014 presidential elections as a case study, Twitter predictors outperform most traditional polls, with the best predictor surpassing all traditional methods on a national scale. This research demonstrates the cost-effective potential of Twitter in predicting election results in the developing world.

[11]. in this study, we present a method for predicting election results using Twitter data. We first identify social media users' stances through retweets and then employ four different counting methods for prediction: simple user counting, city-based weighted counting, closest-city-based prediction, and using former election results. Our evaluation using data from the 2018 Turkish presidential election demonstrates that incorporating domain-specific information and weighted counting by location helps reduce bias. The city-based and former election results-based methods outperform traditional polls, highlighting the potential of social media as an alternative means for election polling.

[12]. this survey paper provides a comprehensive overview of recent advancements in Sentiment Analysis (SA) in the text mining field. It covers enhancements in SA algorithms, various applications, and related areas such as transfer learning, emotion detection, and resource building. The paper categorizes recent articles, offering a detailed snapshot of the state of SA research and its trends.

[13]. Democracy allows people to choose their leaders through elections, as in India. Today, people use social media platforms like Facebook, Twitter, and WhatsApp to predict election outcomes. These platforms provide a space for

sharing views, news, and discussions, with Twitter being particularly influential during events. Analyzing tweets on Twitter can help predict election results through sentiment analysis, often before official exit polls are released.

[14]. U.S. President Joe Biden took office after the 2020 controversial elections, which involved extensive postal voting due to the pandemic. The sentiment analysis of Twitter data examined public opinions before, during, and after the elections, comparing them with actual results and the 2016 election. Using tweet data, TF-IDF features, and the Naive Bayes Classifier, we identified outliers, analyzed swing states, and cross-validated results against social media sentiments. The findings indicate that election outcomes often align with social media sentiment. Pre- and post-election sentiment analysis reveals changes in sentiment. The sentiment classifier achieved 94.58% accuracy and 93.19% precision.

[15]. we conducted sentiment analysis on tweets from German politicians and party accounts during the 2021 German federal election. We collected 58,000 tweets from seven parties, with 2,000 annotated by three annotators. Transformer-based models, like BERT, outperformed traditional methods, achieving 93.3% accuracy and a macro f1 score of 93.4%. Overall, negative sentiment was most common, with significant sentiment shifts before and after the election. Opposition party tweets tended to have more negative sentiment compared to governing parties.

[16]. Twitter is a platform for sharing thoughts and connecting with over 206 million daily users in 2022. Analyzing online information, like sentiment on Twitter, helps understand changes in people's perceptions and behaviors. This has increased interest in Natural Language Processing (NLP) and Artificial Intelligence (AI) for text analysis.

[17]. we collected 153,528 tweets from 18,730 Twitter users over a year, totaling 2,840,024 words. We examined these tweets using two sentiment lexicons: one translated from English to Greek using the Vader library and one in Greek. Our analysis focused on tracking the impact of COVID-19 in terms of positive and negative sentiments, six specific sentiments (Surprise, Disgust, Anger, Happiness, Fear, and Sadness), and correlations between COVID-19 cases, sentiments, and data volume.

[18]. this research examined the collective behavior of Twitter users throughout the Indian General Election of 2019 by conducting weekly sentiment analysis. The primary objective was to determine if it was possible to predict election outcomes in India through the analysis of social media activity. Social media content was used as a factor for forecasting election results, and the accuracy of the predictions was evaluated by comparing them to exit poll results and the actual election results, using mean absolute error. The study's forecasts closely mirrored the final election results, underscoring the promise of social media analysis in predicting election outcomes.

[19]. this study aimed to assess the persuasive power of social media during an election by comparing user sentiments with actual election outcomes. Using tweets from the 2019 UK General Elections as a case study, it found that tweet sentiment is an unreliable predictor of election results. Additionally, the study suggests that social media sentiment polarity does not significantly impact voting decisions.

[20]. the increase in the number of social media users results in the generation of extensive unstructured text containing various opinions. Social media platforms, such as Twitter, serve as valuable resources for gaining political insights, enabling the use of sentiment analysis to forecast election patterns. This research centers on the analysis of tweets related to India's 2019 General Elections, with a particular focus on understanding public sentiments towards major

political parties. It establishes a classification model based on sentiment using Long Short-Term Memory (LSTM) and makes a comparison with conventional machine learning models.

3. Methodology

This study followed an e methodology to conduct the research. The initial phase involved collecting tweets, followed by annotation, pre-processing, word embedding, fine-tuning hyper-parameters, conducting comparative analysis, incorporating semi-supervised learning, and, finally, evaluating performance.
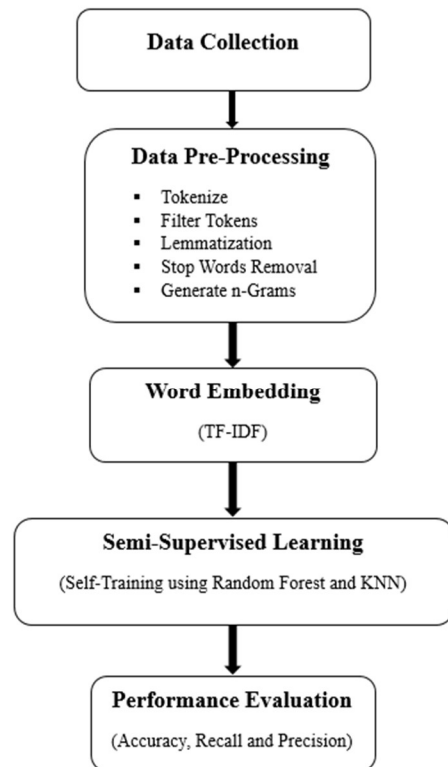
This entire process is visually represented in Figure 1.

Fig 1. Flow chart involves Data collection, Data Pre- Processing, Word Embedding, Semi-Supervised Learning, and Performance Evaluation

3.1 Data collection

Collecting data for this project involved gathering and analyzing Twitter data related to Indian political figures Narendra Modi and Rahul Gandhi during the 2019 Indian General Elections. The data collection process consisted of extracting tweets, user profiles, date information, and emotional tones expressed in the tweets. This data was obtained from a 4GB dataset sourced from Kaggle.

3.2 Data preprocessing

Twitter data related to political figures Rahul Gandhi and Narendra Modi. The preprocessing involves multiple steps, including removing links, mentions, special characters, numbers, and stop words, as well as applying lemmatization and lowercasing to the text. These steps are designed to clean and standardize the text data, making it ready for

analyzing the emotions or further analysis of public emotions toward Rahul and Modi during elections or other political events.

3.3 Transforming the words into vectors using TF-IDF

TF-IDF Vectorizer from scikit-learn to transform text data associated with political figures Rahul Gandhi and Narendra Modi into TF-IDF (Term Frequency-Inverse Document Frequency) features. This process quantifies the importance of words within the text data making it a valuable step in understanding how individuals articulate their viewpoints and emotions concerning Rahul and Modi in the context of elections or political discussions.

3.4 Model selection

Model selection is a critical step in sentiment analysis for political figures like Rahul Gandhi and Narendra Modi. In this context, Random Forest demonstrates a strong performance with an accuracy of 70% for Rahul, while K-Nearest Neighbors (KNN) achieves 71% accuracy for Modi in classifying sentiment within tweets. These results suggest that Random Forest and KNN are robust choices for sentiment analysis, highlighting their potential to provide valuable insights into public sentiment during election campaigns and political discussions.

4. Results and Discussion

In this project, Python, Jupyter Notebooks, Twitter API, Scikit-learn are used as tools. Python served as the primary programming language for data processing, analysis, and Machine learning model development. Python's extensive libraries for NLP and machine learning. Jupyter Notebooks were used to create an interactive and organized environment for code development, data exploration, and documentation. Access to Twitter data was facilitated through the Twitter API, which allowed the collection of real-time tweets relevant to the 2019 Indian General Elections. Scikit-learn, a powerful machine learning library in Python, provided the necessary tools for constructing and assessing the machine learning models. Algorithms like k-Nearest Neighbors, Random Forest, Multinomial Naive Bayes, and Gaussian Naive Bayes were implemented using this library. Matplotlib for pie charts plotting.

For this reasearch, Microsoft Windows 11, Processor: 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz, 2304 Mhz, 8 Core(s), 16 Logical Processor(s), Memory: 16GB DDR4 RAM, Storage: 512GB NVMe SSD has been used.

The data set for this research consisted of tweets collected from Twitter during the period of the 2019 Indian General Elections. Dataset size is 4GB. The dataset comprises Twitter data related to Indian political leaders Narendra Modi and Rahul Gandhi. Sourced from Kaggle, it includes tweets, user profiles, date, and emotions. This dataset can be used for research and analysis, offering insights into public sentiment, political trends, and social media discussions. When using this data, it's important to follow the rules of Twitter and give credit to Kaggle as the source of the data.

The results of our sentiment analysis provide valuable insights into public sentiment towards two prominent political figures, Rahul Gandhi, and Narendra Modi, based on Twitter data. The analysis was conducted using Random Forest for Rahul and K-Nearest Neighbors (KNN) for Modi. For Narendra Modi, the K-Nearest Neighbors (KNN) model achieved an accuracy of 71% in sentiment classification. The model's predictions were heavily skewed towards negative sentiment, with 80% of predictions falling into the negative category, while only 20% were categorized as positive. In the case of Rahul 70% accuracy in classifying sentiment. Similar to Modi, the model's predictions were weighted toward negative sentiment, with 89.1% of tweets being categorized as negative and only 11.9% as positive. The prevalence of negative sentiment in both cases might reflect the nature of political discourse on social media, where critical opinions and critiques are more prominently expressed. It's essential to acknowledge that sentiment analysis accuracy depends on the quality of training data and the complexity of the sentiment nuances, which can vary greatly in political discussions. These results serve as a valuable starting point for understanding public sentiment on Twitter, but further analysis is needed to uncover the subtleties and variations within public sentiment towards these political figures. In the context of political campaigns and decision-making, these results highlight the challenges and opportunities of understanding public sentiment on social media platforms. They underscore the importance of considering the prevalence of negative sentiments and the need for strategies that address and engage with such sentiments.

Overall, our sentiment analysis offers a glimpse into how Twitter users express their sentiments about Rahul Gandhi and Narendra Modi in the context of political discourse and elections, shedding light on both the predictive potential and challenges of sentiment analysis in this domain.

Table 1. Rahul tweets

	Accuracy	Precision (mean)	Recall (mean)	F1-Score (mean)
Multinomial Navies Bayes	0.682171	0.840467	0.506024	0.416997
Gaussian Navies Bayes	0.647287	0.568246	0.553150	0.551972
Logistic Regression	0.686047	0.841797	0.512048	0.429562
K-NearestNeighbors	0.713178	0.671237	0.604888	0.607225
Support Vector Machine	0.686047	0.620524	0.550052	0.528932
Decision Tree	0.585271	0.485859	0.488434	0.482442
Random Forest	0.701550	0.785500	0.539312	0.486335

Table 2. Modi tweets

|	Accuracy	Precision (mean)	Recall (mean)	F1-Score (mean)
Multinomial Navies Bayes	0.636719	0.763008	0.541329	0.465121
Gaussian Navies Bayes	0.632812	0.609235	0.600192	0.601273
Logistic Regression	0.652344	0.817623	0.559406	0.494666
K-NearestNeighbors	0.644531	0.622785	0.613318	0.614910
Support Vector Machine	0.679688	0.705896	0.611306	0.596928
Decision Tree	0.640625	0.617321	0.591121	0.587328
Random Forest	0.703125	0.777959	0.630661	0.615889
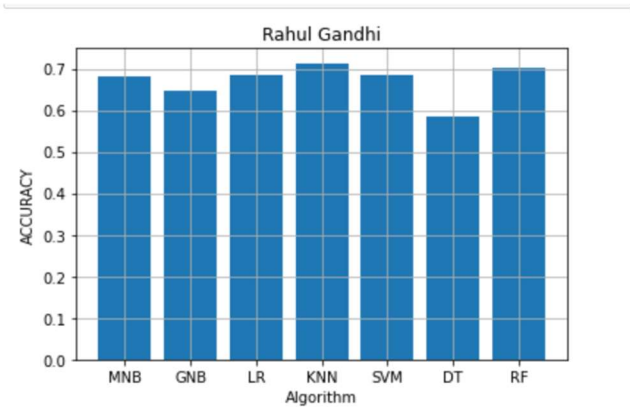
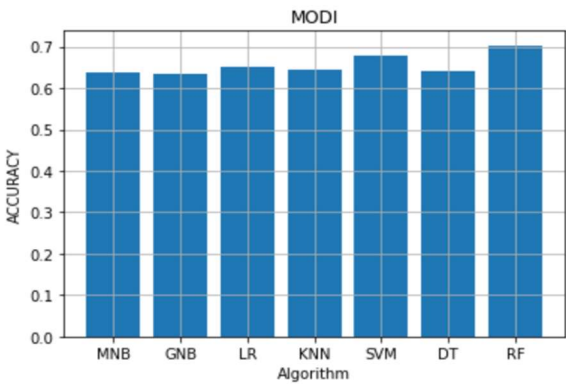Fig 2. Graphical representation of Modi   Fig 3. Graphical representation of Rahul

Table 1 & 2 , It is showing the Accuracy, Precision, Recall and F1-Score for different algorithms such as Multinomial Naives Bayes, Gaussian Naives Bayes, Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Decision Tree, and Random Forest. Among all the algorithms, for Modi Tweets K-Nearest Neighbors is more accurate and for Rahul Tweets Random Forest is more accurate. It is graphically represented in Fig 2 and Fig 3.
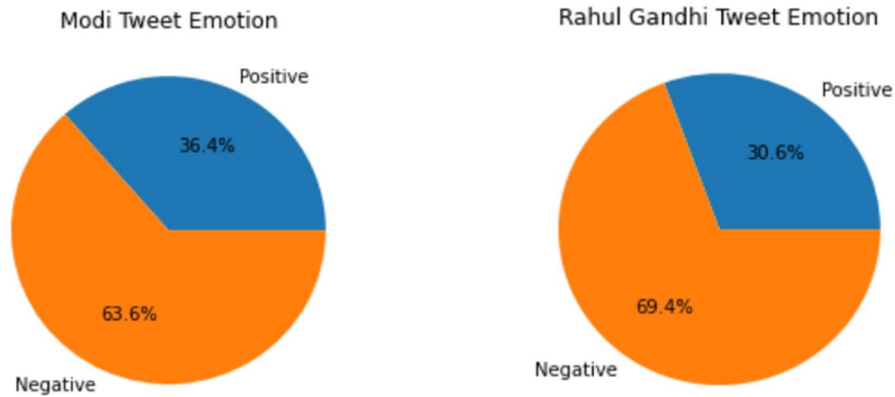
Fig 4. Data visualization of Modi and Rahul

Based on the tweets of trained dataset, the ratio of positive and negative emotions of publics about the political figures Rahul Gandhi and Narendra Modi are visually represented in Fig 4.
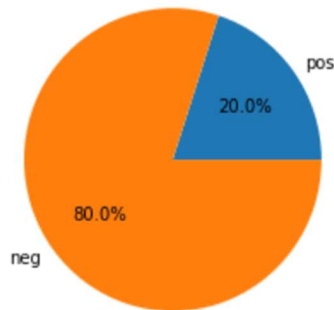

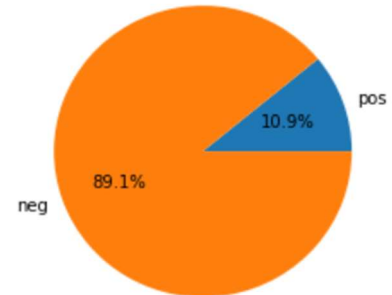Fig 5. Rahul Gandhi Prediction Representation   Fig 6. Modi Prediction Representation

Based on the tweets of testing dataset, the ratio of predicted positive and negative emotions of publics about Rahul Gandhi and Modi are represented by the pie chart in Fig 5 and Fig 6.

5. Conclusion

Our project has advanced the domain of sentiment analysis in political discussions on social media. By utilizing machine learning models on Twitter data related to Rahul Gandhi and Narendra Modi, we have deepened our comprehension of public sentiment during electoral processes. The suggested approach, using K-Nearest Neighbors (KNN) and Random Forest models, yielded accuracy rates of 71% and 70%, respectively, effectively categorizing sentiments within tweets. Despite the prevalence of negative sentiments, it's vital to recognize that sentiment analysis alone cannot forecast election results, as elections are influenced by a multitude of factors. Our project contributes to the body of knowledge in political sentiment analysis, emphasizing the significance of comprehensive data sources for more precise predictions and establishing a foundation for further research in this field.

6. Future Work

In this project we tried to predict election result based on sentimental analysis using tweets where the average accuracy is 71% only. In near future we will try to predict based on all social media with higher accuracy.

7. References

[1]. Gayo-Avello, Daniel. "A balanced survey on election prediction using twitter data." Department of Computer Science-University of Oviedo, Spain (2012).

[2]. Jain, Vinay K., and Shishir Kumar. "Towards prediction of election outcomes using social media." International Journal of Intelligent Systems and Applications 9.12 (2017): twenty.

[3]. Budiharto, Widodo, and Meiliana Meiliana. "Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis." Journal of Big data 5.1 (2018): 1-10.

[4]. Sharma, Ankita, and Udayan Ghose. "Sentimental analysis of twitter data with respect to general elections in India." Procedia Computer Science 173 (2020): 325-334.

[5]. Chauhan, Priyavrat, Nonita Sharma, and Geeta Sikka. "The emergence of social media data and sentiment analysis in election prediction." Journal of Ambient Intelligence and Humanized Computing 12 (2021): 2601-2627.

[6]. Sharma, Parul, and Teng-Sheng Moh. "Prediction of Indian election using sentiment analysis on Hindi Twitter." 2016 IEEE international conference on big data (big data). IEEE, 2016.

[7]. Satti, Ayush, Deepanshu Bhardwaj, and Lupin Garg. "Prediction of Election Result by Analysing Twitter Data via Constituency Model."

[8]. Vijayaraj, A., et al. "Twitter based sentimental analysis of Covid-19 observations." Materials Today: Proceedings 64 (2022): 713-719.

[9]. Antonakaki, Despoina, Paraskevi Fragopoulou, and Sotiris Ioannidis. "A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks." Expert Systems with Applications 164 (2021): 114006.

[10]. Dwi Prasetyo, Nugroho, and Claudia Hauff. "Twitter-based election prediction in the developing world." Proceedings of the 26th ACM Conference on Hypertext & social media. 2015.

[11]. Bayrak, Cansin, and Mucahid Kutlu. "Predicting Election Results via Social Media: A Case Study for 2018 Turkish Presidential Election." IEEE Transactions on Computational Social Systems (2022).

[12]. Medhat, Walaa, Ahmed Hassan, and Hoda Korashy. "Sentiment analysis algorithms and applications: A survey." Ain Shams engineering journal 5.4 (2014): 1093-1113.

[13]. Ramteke, Jyoti, et al. "Election result prediction using Twitter sentiment analysis." 2016 international conference on inventive computation technologies (ICICT). Vol. 1. IEEE, 2016.

[14]. Chaudhry, Hassan Nazeer, et al. "Sentiment analysis of before and after elections: Twitter data of us election 2020." Electronics 10.17 (2021): 2082.

[15]. Schmidt, Thomas, et al. "Sentiment analysis on twitter for the major German parties during the 2021 German federal election." Proceedings of the 18th Conference on Natural Language Processing (KONVENS 2022). 2022.

[16]. Qi, Yuxing, and Zahratu Shabrina. "Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach." Social Network Analysis and Mining 13.1 (2023): thirty-one.

[17]. Samaras, Loukas, Elena García-Barriocanal, and Miguel-Angel Sicilia. "Sentiment analysis of COVID-19 cases in Greece using Twitter data." Expert Systems with Applications (2023): 120577.

[18]. Samaras, Loukas, Elena García-Barriocanal, and Miguel-Angel Sicilia. "Sentiment analysis of COVID-19 cases in Greece using Twitter data." Expert Systems with Applications (2023): 120577.

[19]. Rita, Paulo, Nuno António, and Ana Patrícia Afonso. "Social media discourse and voting decisions influence: sentiment analysis in tweets during an electoral period." Social Network Analysis and Mining 13.1 (2023): forty-six.

[20]. Ansari, Mohd Zeeshan, et al. "Analysis of political sentiment orientations on twitter." Procedia Computer Science 167 (2020): 1821-1828.