A Project Report on

“ Election Result Prediction By Analyzing Tweets ”

SUBMITTED TO

G H RAISONI COLLEGE OF ENGINEERING MANAGEMENT

DEPARTMENT OF ARTIFICIAL INTELLIGENCE

SUBMITTED BY-

ONKAR MANE	**TY-AI-40**
RANI MANWAR	**TY-AI-47**
RUTUJA OHAL	**TY-AI-51**
SANKET CHAUDHARY	**TY-AI-56**

UNDER THE GUIDANCE OF

Prof. *SMITA BATTALWAR*

DEPARTMENT OF ARTIFICIAL INTELLIGENCE

G H RAISONI COLLEGE OF ENGINEERING MANAGEMENT, WAGHOLI,PUNE.

2022-23

CERTIFICATE

This is to certify that the minor project report entitles

Submitted By

Onkar Mane	**A40**
Rani Manwar	**A47**
Rutuja Ohal	**A51**
Sanket Chaudhary	**A56**

are the bonafide students of this institute and the work has been carried out by them under the supervision of **Prof. Smita Battalwar** and it is approved for the partial fulfilment of the requirement.

Prof. Smita Battalwar
Guide

Prof. Rachna Sable
Head of Department

Dr. R. D. Kharadkar
Campus Director
GHRCEM

ACKNOWLEDGEMENT

It gives us great pleasure and satisfaction in presenting this project report on **“Election Result Prediction By Analyzing Tweets ”.**

We are thankful to and fortunate enough to get constant encouragement, support and guidance from all Teaching staffs of AI Department which helped us in successfully completing our project work. Also, we would like to extend our sincere esteems to all staff in laboratory for their timely support.

We have furthermore to thank AI Department HOD Prof. R. Y. Sable and Guide Prof. Smita Battalwar to encourage us to go ahead and for continuous guidance. We would also like to thank our project team members who showed immense patience and understanding throughout the project.

We would like to thank all those, who have directly or indirectly helped us for the completion of the work during this project.

1. Onkar Mane
2. Rani Manwar
3. Rutuja Ohal
4. Sanket Chaudhary

TABLE OF CONTENTS

1.**INTRODUCTION:**

Social media has become a powerful tool for sharing opinions. There are social media platforms like Facebook, Twitter to share opinions, reviews and ratings. All major political parties and their members all over the world have their official accounts on Twitter with millions of followers. They consider this platform as a medium to connect with young people who might vote them. With significant rise of Indian users on Twitter during the pandemic, people have been more vocal to criticize or appreciate a political decision.

Sentimental Analysis is a method to teach a machine to extract emotion from a given text. A text can be anything, a simple review, a social statement, tweets or messages. Twitter Sentiment Analysis of tweets regarding elections can be used by the general public as well as the political parties to understand the positive or negative views of people regarding a particular political party, thus, helping to predict the election results during that period.

Elections play an important role in a democratic country. US parliamentary system gives its people the right to decide who will govern them for the next four year.

2.Related Work: -

In this section, we are going to discuss related works about predicting the result of an election using Twitter. We noticed that researchers use a different approach regarding this problem. There are researchers who try to discover the political or ideology preference of a user, then relate it to the election and there are others who use selected tweet related to the upcoming election and figure out vote preference of the user using that data. Different strategies such as profile details, user behaviour, Twitter specific feature (reply/re-tweet), user graph and sentiment from tweet content can be used for inferring political leaning. For example, the authors used tweet containing parties' name in several political events to assign a political/ideological leaning of the user who posted the tweets. Similar to the previous method, used the tweets and retweets of a user regarding a political party to infer the political leaning. Assigned a score to every congress member which a Twitter user is following, then a political preference is assigned based on that score. The authors compared several features such as user's bio and avatar, posting behaviour, linguistic content, follower, reply and retweet. They found out that the combination between user profile and linguistic outperform other feature. They then applied to classify the ethnicity of the user and whether the user is a Starbucks fan, but their result showed that information from user bio is more accurate for classifying Starbucks fan, and user's avatar for classifying user's ethnic.

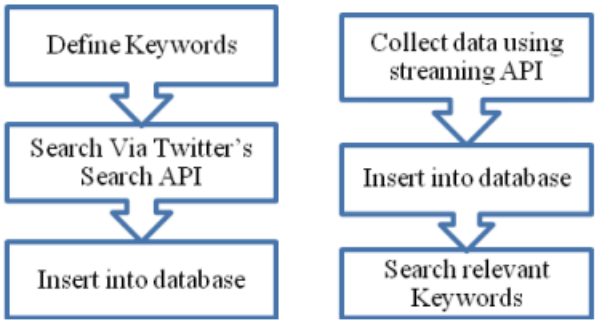
The second approach is by using selected data just days or weeks prior to the election. The prediction could be derived by comparing the number of tweets mentioning each candidate or by comparing the number of tweets that has positive sentiments towards each candidate. The earliest research stated that the number of tweets mentioning a party reflects the election result where they found out that the prediction result from Twitter were only better than other. While is the first research in which argued that sentiment detection approach from Twitter can replace the expensive and time intensive polling?

Researchers have tried to compare these two methods, for example, that tried to predict congress and senate election in several states of the US. They showed that though the method is the same, the prediction error can vary greatly. The research also showed that lexicon based sentiment analysis improves the prediction result, but the improvements also vary in different states. Same result was shown inwhere they predict the result of Irish general election using both methods and which predicts the Italian primary election. All of the research showed that sentiment detection does reduce the error of the prediction result. Because of that, several researchers focused on improving the sentiment analysis, such as and who used more sophisticated sentiment analysis than lexicon based in the US presidential election, France legislative election, and Italy primary election.

3.Methodology: -

1. Importing the Libraries.
2. Loading the Training Dataset.
3. Normalizing the Dataset
4. Reshape the data
5. Building the Model by Importing the Libraries.
7. Extracting the Actual Tweets.
8. Predicting the future Values.
9. Predicted Result.

- **Data Collection**
  The data collection step is the initial phase in the research, where data is collected from twitter. There are two methods on how to connect and collect tweets from Twitter. The first method is by searching tweets matching to the keywords. The second method is by collecting all the tweets provided by Twitter through streaming API, or all the tweets in a specific language, or all the tweets in a specific location then put all of them into the database.



Method 1 Method 2

- **Sentiment Analysis:**

  Lexicon Based Approach: There three main approaches to compile sentiment words. Three main approaches are: manual approach, dictionary-based approach, and corpus-based approach.in our research we used dictionary-based approach. We used eleven different variables for classification, that variables are sadness, tentativeness, anxiety, work, anger, certainty, achievement, positive words, negative words, positive hashtag, and negative hashtag. We collected various word related to that eleven variable and classified them.

4.SYSTEM REQUIREMENT: -

PC/Laptop

Google Colab

Stable Internet Connection

OS Requirement: -

Windows

Hardware Requirements: -

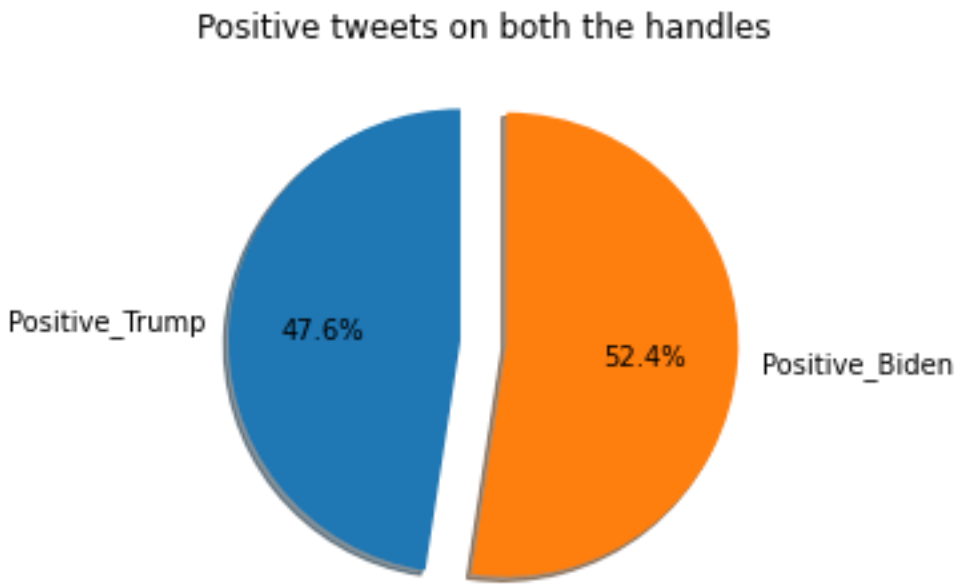
Windows Intel i5 10th Gen

GPU

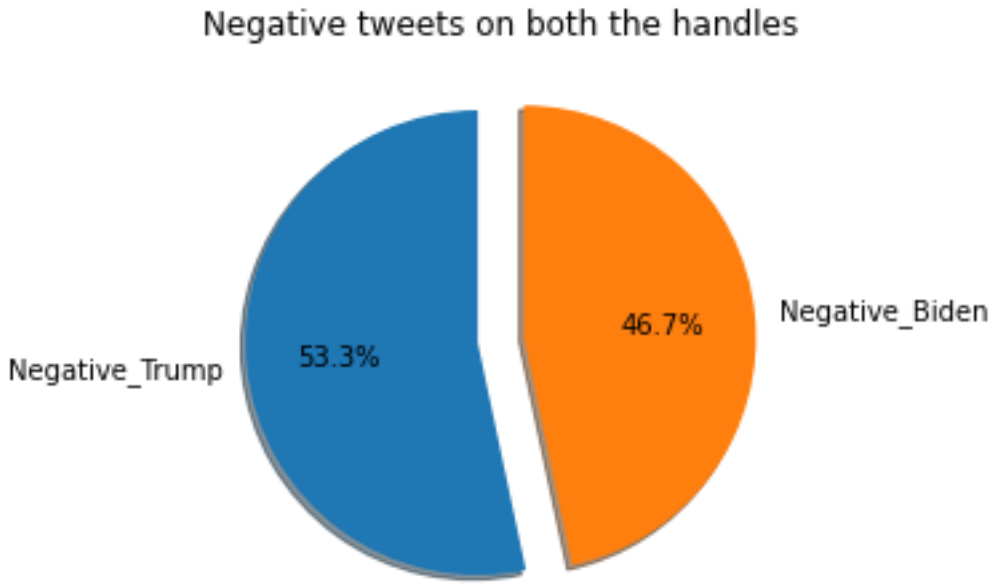
5.RESULTS: -

We collect data through twitter API, after that we performed preprocessing on that data. For collected data for US election we classified polarity. Classifying tweets in three categories positive, negative and neutral.

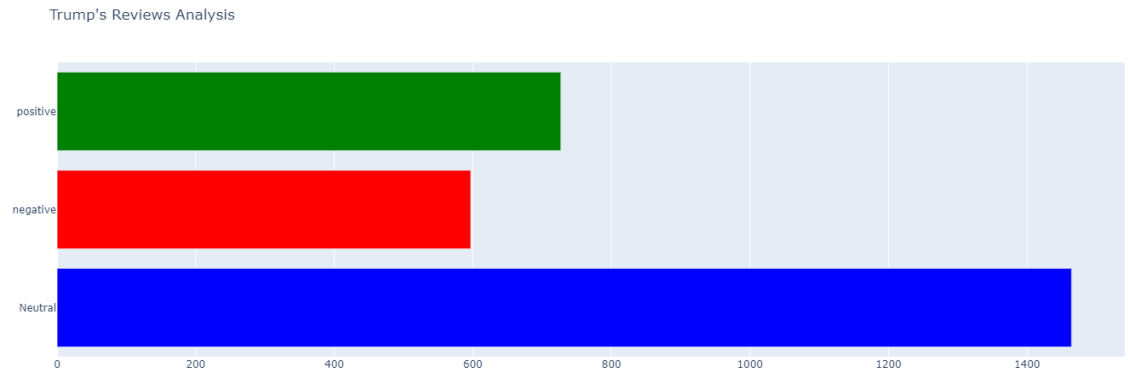
Comparison Between Positive Tweets: -



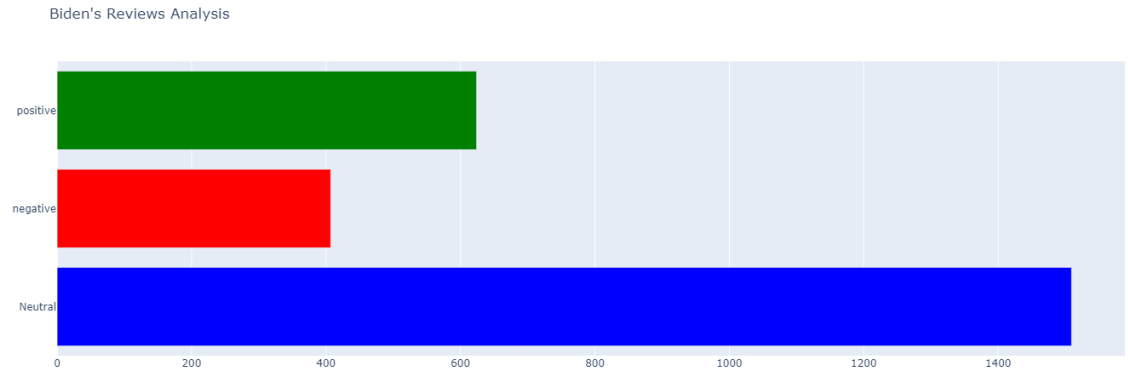
Comparison Between Negative Tweets: -



Trump’s Reviews Analysis: -



Biden’s Reviews Analysis: -



We observed that in sentiment analysis positive tweets were more in Biden data than that of trump and vice versa. Neutral tweets are also more for Biden. As prediction we can predict that Joe Biden will win elections from tweeter data we collected. we can see the comparison of positive, negative, and neutral tweets for Biden ad trump in piechart-1.

6.Conclusion: -

The proposed system can be used by political parties to improve their campaigning strategies during the election period. It can be used by them as a part of social media analytics to study the trends of other political parties as well. User can make informed decision in voting by seeing the current trends of political parties. Political analyst and strategist can use this methodology, as application, as a long-term plan for a political party to study the sentiments of people over a long time period. Observing the expanded use of social media platforms, this project concentrated on exploring of social platform (Twitter) as the chase for elections' campaign.

7.Reference: -

[1] Data Robot, "Introduction to Sentiment Analysis: What is Sentiment Analysis?" Data Robot, 26 March 2018. [Online]. Available: https://www.datarobot.com/blog/introduction-tosentiment-analysis-what-is-sentiment-analysis/.

[2] T.-S. M. Parul Sharma, "Prediction of Indian Election Using Sentiment Analysis," 2016 IEEE International Conference on Big Data (Big Data), pp. 1966-1971, 2016.

[3] S. U. S. K. M. S. R. G. C. U. J. Dr D Rajeswara Rao, "Result prediction for political parties using twitter sentiment analysis," International Journal of Computer Engineering and Technology, no. 11(4), 2020.

[4] F. J. J. Joseph, "Twitter Based Outcome Predictions of 2019 Indian General Elections Using Decision Tree," 2019 4th International Conference on Information Technology (InCIT), Bangkok, THAILAND, pp. 50-53, 2019.

[5] Y. W. M. K. a. N. R. Meng-Hsiu Tsai, "A machine learning based strategy for election result prediction," 2019 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 1408-1410, 2019.

[6] A. S. S. a. C. G. Payal Khurana Batra, "Election Result Prediction Using Twitter Sentiments Analysis," 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 182-185, 2020.

[7] "Uttar Pradesh Assembly Election 2022 Opinion Poll," Oneindia, 2022. [Online]. Available: https://www.oneindia.com/uttar-pradesh-election2022-opinion-poll-and-exit-poll/.

[8] GeeksforGeeks, 7 Oct 2021. [Online]. Available: https://www.geeksforgeeks.org/python-sentimentanalysis-using-vader/.

[9] Mamun, "Medium," 20 June 2019. [Online]. Available: https://medium.com/@imamun/creating-a-tf-idf-inpython-e43f05e4d424. [Accessed 20 May 2022].

[10] "Indian Elections 2022," Oneindia, 25 March 2022. [Online]. Available: https://www.oneindia.com/elections/