Title - Election Sentimental Analysis

Presented by
Rohit Bhomkar-8
Himanshu Goyal-18

Content

Introduction

In today's digital age, social media platforms have become a powerful tool for gauging public sentiment, especially in political contexts. With millions of users sharing their opinions on platforms like Twitter, it is possible to analyze and predict public opinion trends based on the content of these posts.

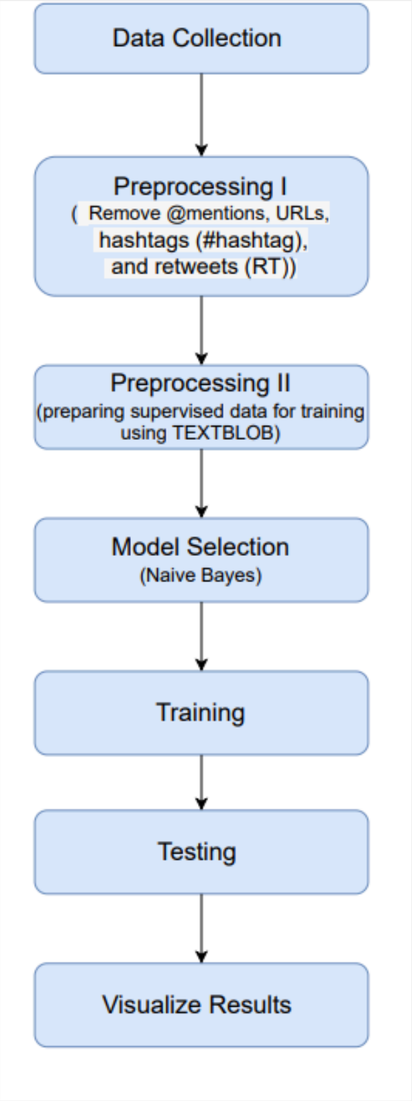
This project focuses on sentiment analysis, a branch of natural language processing (NLP), to analyze the public sentiment around two major Indian political figures, Rahul Gandhi and Narendra Modi. By analyzing tweets related to these politicians, we aim to understand the general mood of the electorate and even predict electoral outcomes based on the sentiment of social media discussions.

Dataset

1. User: This column represents the Twitter handle or username of the individual who posted the tweet. It provides information about the source of the tweet, but in most cases, it is used as metadata to track who is saying what.
2. Tweet: This column contains the actual text of the tweet posted by the user. It is the main content that will be analyzed for sentiment. Each row corresponds to a unique tweet made by a user, making this column critical for extracting insights such as sentiment (positive, negative, neutral) and public opinion on the topic.

The User column can help you identify different patterns of sentiment based on who is tweeting, while the Tweet column is where the core sentiment analysis will take place.

System Architecture



We use two primary techniques for sentiment analysis in this project:

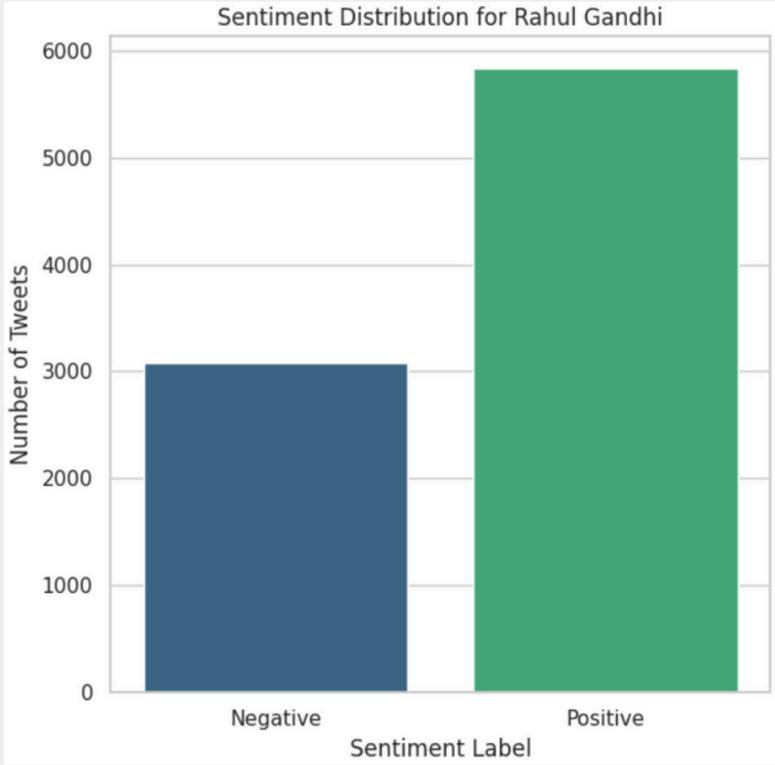
- TextBlob, which is used to assign sentiment polarity to individual tweets (positive, negative, or neutral).

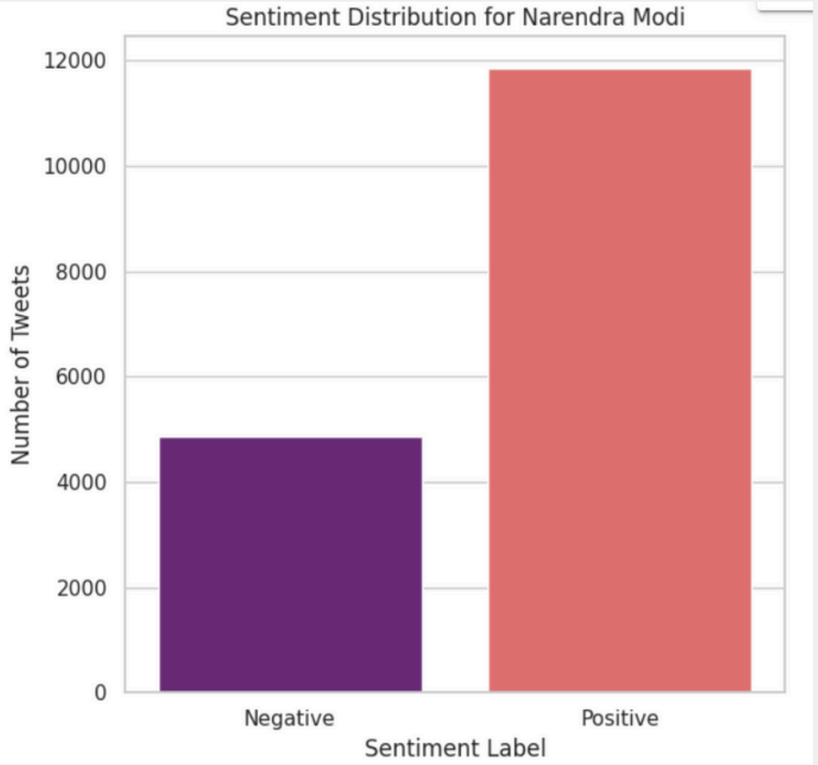
TextBlob assigns a polarity score to each tweet, ranging from -1 (most negative) to +1 (most positive). Neutral sentiments receive a score close to 0.

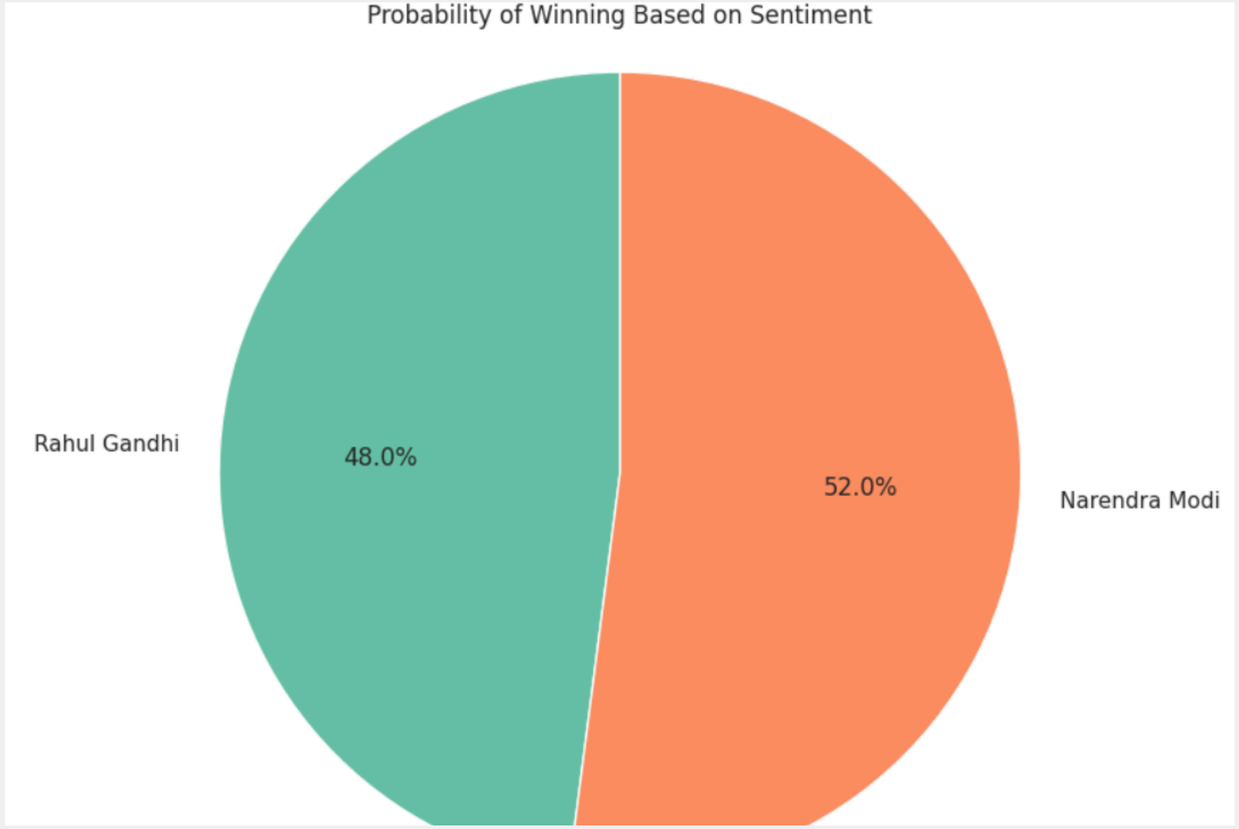
- Naive Bayes, a machine learning algorithm, to classify tweet sentiments and train a model that can predict sentiment on new data.Naive Bayes is then trained on the labeled tweets and used for more robust classification and prediction of sentiment, allowing you to apply it to new, unseen data with better accuracy and learning from the patterns in the data.

Results

```
Rahul Gandhi Sentiment Analysis:
              precision    recall  f1-score   support

         0.0       0.79      0.56      0.66       581
         1.0       0.81      0.93      0.87      1204

    accuracy                           0.81      1785
   macro avg       0.80      0.75      0.76      1785
weighted avg       0.81      0.81      0.80      1785

Accuracy: 0.8095238095238095

Narendra Modi Sentiment Analysis:
              precision    recall  f1-score   support

         0.0       0.81      0.53      0.64       971
         1.0       0.83      0.95      0.89      2373

    accuracy                           0.83      3344
   macro avg       0.82      0.74      0.76      3344
weighted avg       0.83      0.83      0.82      3344

Accuracy: 0.8280502392344498
```







The analysis reflects the public's current sentiment on social media, which may provide insights into voter mood and preferences in a political context. However, it’s important to remember that sentiment on social media doesn’t always directly translate to election outcomes, as these platforms capture only a portion of the overall electorate. Factors such as the demographic using Twitter, political biases, and regional influences must also be considered.