**Indian Politics Tweets & Reactions - Sentiment Analysis Model**

**Overview**

The proposed work identifies and classifies a specific collection of public tweets regarding the Indian political scene, in order to assess and understand how the public’s opinion changes in regard to different issues, events or political campaigns. The Indian Politics Tweets & Reactions dataset consists of a number of tweets that have engagement metrics (likes, retweets, comments) attached to them, and can be used to form opinions on political parties, leaders, their policies, and elections and associated controversies, among others.

In order to enhance the accuracy of sentiment classification in the polarized environment, the advanced machine learning algorithms along with a deep learning model which are utilized in this endeavor and the tweet in consideration, based on its content, is classified into one of three broad categories based on its sentiment. This analysis cohesively brings together, the processing of text data as well as feature engineering in order to analyze and predict sentiments from social media interactions.

**Goal**

The main target of credits is to make prediction as to the sentiments into which a given tweet may be placed for instance in this project the targets include:

* Investigate the emotions and relationships as well as the target of social actors towards Indian political issues. Social media data will come handy in gauging the mood of the masses.
* Utilize the high-class model to anticipate sentiments that go with the concept in question.
* Address the questions of how different tweets appeal to the masses in order to understand the level of political involvement as well as the effectiveness of the messages communicated.

**Dataset Description**

The Indian Politics Tweets & Reactions dataset includes:

* Tweet Text: Publicly shared messages related to Indian politics on Kaggle.
* User Reactions: Engagement data, such as likes, retweets, and comments, to gauge tweet impact and reach.
* Topics Covered: Political campaigns, government policies, elections, controversies, leaders, and more.
* Features for Analysis: Cleaned tweet text, polarity score, timestamp, and other numerical features.

This dataset offers a rich source for sentiment analysis, allowing researchers and analysts to examine trends in public opinion and the effectiveness of political messaging.

**Tools Used and Their Purpose**

This project relied on several essential Python libraries to transform, model, and evaluate the dataset effectively. **Pandas** was used extensively for data processing, cleaning, and organization, allowing efficient handling of the tweet dataset, including removing noise, handling missing values, and structuring the data for feature extraction. **NumPy** facilitated efficient numerical computations and array manipulation, making it easier to process and combine TF-IDF vectorized text features with additional numerical features like Polarity.

**Analysis of Model Performance on Sentiment Classification**

Each model’s performance is evaluated based on its ability to classify tweets into three sentiment categories: Negative, Neutral, and Positive. Key metrics, such as accuracy, precision, recall, and F1-score, offer insights into each model's strengths and weaknesses.

**1. Decision Tree Classifier**

Accuracy: 72%

Class-Level Performance:

Negative: Precision = 67%, Recall = 59%, F1-score = 63%

Neutral: Precision = 67%, Recall = 76%, F1-score = 71%

Positive: Precision = 78%, Recall = 77%, F1-score = 77%

Interpretation: The Decision Tree Classifier demonstrates relatively strong performance on the Positive class with an F1-score of 77%. However, it struggles with the Negative class, showing a low recall of 59%, indicating that it often misclassifies negative tweets as other sentiments. The overall accuracy of 72% suggests moderate performance, likely due to the simplicity of the model.

**2. Random Forest Classifier**

Accuracy: 78%

Class-Level Performance:

Negative: Precision = 85%, Recall = 60%, F1-score = 70%

Neutral: Precision = 73%, Recall = 83%, F1-score = 78%

Positive: Precision = 79%, Recall = 86%, F1-score = 82%

Interpretation: With an accuracy of 78%, the Random Forest Classifier outperforms the Decision Tree, especially in handling the Positive class, where it achieves a recall of 86% and an F1-score of 82%. However, its recall for the Negative class is still relatively low (60%), indicating persistent challenges in identifying negative sentiment accurately. This improvement over Decision Tree results from Random Forest's ensemble nature, which aggregates multiple trees to make more balanced predictions.

**3. Logistic Regression Model**

Accuracy: 82%

Class-Level Performance:

Negative: Precision = 84%, Recall = 69%, F1-score = 75%

Neutral: Precision = 80%, Recall = 86%, F1-score = 83%

Positive: Precision = 84%, Recall = 88%, F1-score = 86%

Interpretation: Logistic Regression achieves an accuracy of 82%, showing balanced performance across all classes. It performs well with the Positive class, achieving an F1-score of 86%, but slightly underperforms on the Negative class with a recall of 69%. This model’s ability to capture linear relationships within the features contributes to its overall effectiveness, particularly in distinguishing between Neutral and Positive tweets.

**4. Support Vector Classifier (SVC)**

Accuracy: 84%

Class-Level Performance:

Negative: Precision = 84%, Recall = 72%, F1-score = 77%

Neutral: Precision = 81%, Recall = 91%, F1-score = 86%

Positive: Precision = 87%, Recall = 87%, F1-score = 87%

Interpretation: The SVC model performs robustly, achieving the highest accuracy of 84%. Its performance is particularly notable in the Neutral class, with a recall of 91% and an F1-score of 86%, indicating strong predictive capability for neutral sentiment. Additionally, it maintains a balanced precision and recall in the Positive class, both at 87%. This model’s high accuracy and consistent performance across classes make it the strongest candidate among the models tested, likely due to SVC's ability to effectively separate classes in higher-dimensional space.

**5. Gradient Boosting Classifier**

Accuracy: 68%

Class-Level Performance:

Negative: Precision = 86%, Recall = 41%, F1-score = 56%

Neutral: Precision = 69%, Recall = 58%, F1-score = 63%

Positive: Precision = 64%, Recall = 89%, F1-score = 75%

Interpretation: Gradient Boosting shows a relatively low accuracy of 68%, with significant variability in class performance. Although it achieves a high precision of 86% for the Negative class, its recall is notably low at 41%, suggesting that it struggles to correctly identify many Negative tweets. Conversely, it achieves high recall in the Positive class (89%) but with reduced precision. The low overall accuracy and inconsistent performance across classes indicate that this model may not be well-suited for this dataset without further tuning.

**6. Convolutional Neural Network (CNN)**

Accuracy: 81%

Class-Level Performance:

Negative: Precision = 76%, Recall = 70%, F1-score = 73%

Neutral: Precision = 83%, Recall = 77%, F1-score = 80%

Positive: Precision = 82%, Recall = 88%, F1-score = 85%

Interpretation: The CNN model achieves an accuracy of 81%, showing relatively balanced performance across classes. It performs well with the Positive class, reaching an F1-score of 85%, and maintains reasonable precision and recall in the Neutral class. However, it slightly underperforms in the Negative class with a recall of 70%. The CNN’s architecture, which is effective for complex data patterns, captures some nuances within the text data but doesn’t outperform simpler models like SVC in this particular context.

**Conclusions and Suggestions**

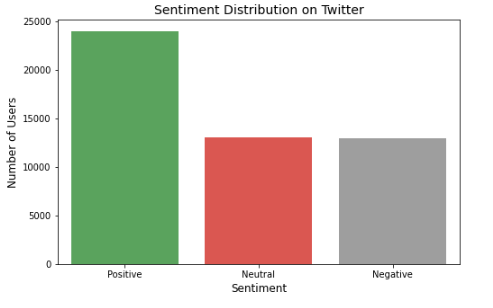
In this study, the model that performed best in providing accuracies surpassing the 80% threshold was Support Vector Classifier. It demonstrated strong balanced performance across all classes while doing particularly well in positive and neutral sentiment recognition. It was also established that Logistic Regression and CNN can also make suitable alternatives to SVC as they achieved accuracies of 81% and 82%, respectively.

On the literacy tests effectively predicting three of the classes decision trees and gradient boosting algorithms performed the weakest estimating a single class with slight fluctuation, possibly indicating their parameters are not optimal for this dataset. As for further enhancements, it could be beneficial to focus on boosting the performance of the previously tested SVC or investigate per-class focus ensemble approaches such as stacking.

Citation for Dataset:

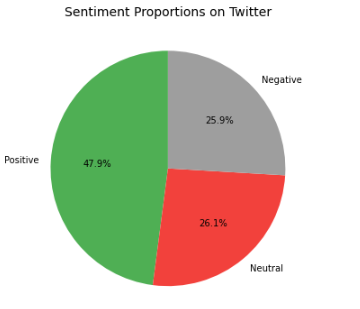
*Adritpal. (2023).* Dataset of Indian politics tweets and reactions *[Data set]. Kaggle.* [*https://www.kaggle.com/datasets/adritpal08/dataset-of-indian-politics-tweets-and-reactions*](https://www.kaggle.com/datasets/adritpal08/dataset-of-indian-politics-tweets-and-reactions)

**Appendix:**

****

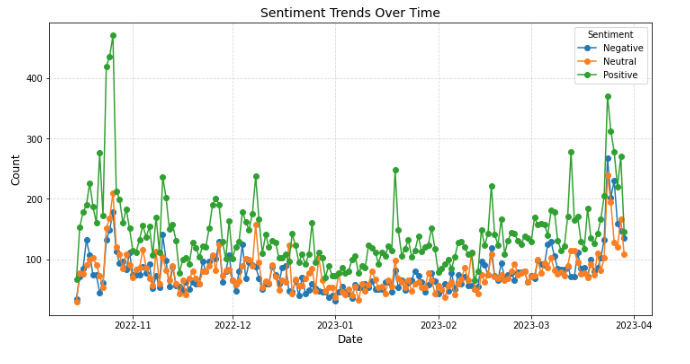
**Figure 1: Sentiment Distribution**

This bar plot illustrates the distribution of different sentiments (Positive, Negative, Neutral) among Twitter users. It shows the number of users expressing each sentiment, with colors representing positive (green), negative (red), and neutral (grey) reactions.

****

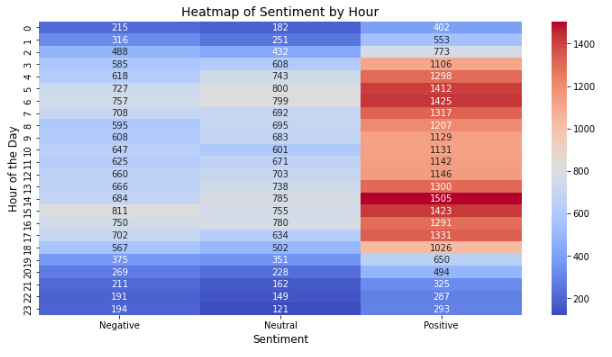
**Figure 2: Pie-chart for Sentiment Allocation**

This pie chart displays the proportions of different sentiments expressed on Twitter. Each slice represents a sentiment category, showing its percentage share of the overall sentiment distribution among users.

****

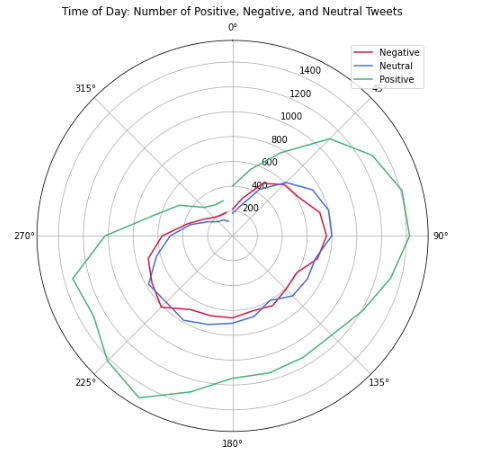
**Figure 3: Time Series Plot for Sentiment Trends**

This time series plot illustrates the daily trends in over a specified period. Each line represents the count of tweets per sentiment, showing how the volume of each sentiment fluctuates over time, potentially indicating shifts in public opinion on political topics.

****

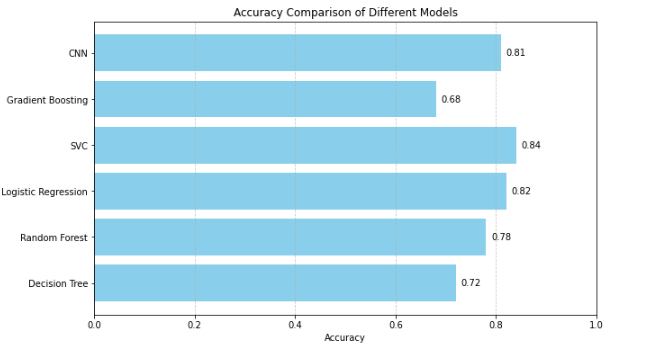
**Figure 4: Heat map of Sentiment by Hour**

This heat map displays the distribution of sentiments by each hour of the day. The intensity of color indicates the frequency of each sentiment across different hours, helping to identify peak times for certain sentiments and patterns in user sentiment throughout the day.

****

**Figure 5: Polar Line Plot**

This polar line plot illustrates the distribution of tweets by sentiment across different hours of the day in a circular format. Each line represents a sentiment, showing the frequency of tweets at each hour, with colors indicating positive (green), negative (red), and neutral (blue) sentiments. The plot provides an intuitive view of when specific sentiments are most prominent throughout the day.



**Figure 6: Plot for Model Accuracies**

This horizontal bar plot visualizes the accuracy of different machine learning models used for sentiment classification. Each bar represents a model, with its length indicating the accuracy level. By displaying accuracy values directly on the bars, the plot provides an at-a-glance comparison, highlighting which models performed best (with higher accuracy) and helping to identify the most effective model for sentiment analysis. The highest accuracy is achieved by the SVC model at 0.84, while the Gradient Boosting model shows the lowest accuracy at 0.68.

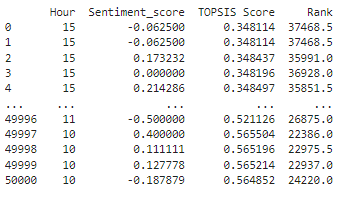


Table 1: Topsis Score

The TOPSIS ranking results provide an insightful perspective on how posting times and sentiment scores contribute to the effectiveness of tweets. In this analysis, the Hour and Sentiment score are the key criteria, with lower Hour values being more desirable and higher Sentiment score indicating more positive sentiment. Observing the table, tweets with similar posting times Hour 15 have varying sentiment scores, resulting in different TOPSIS scores and ranks. For instance, a sentiment score of 0.214286 at Hour 15 results in a higher TOPSIS score (0.348497), indicating better desirability compared to negative sentiment scores (-0.062500) at the same hour (0.348114). Tweets posted at earlier hours, such as Hour 10 or 11, generally achieved higher TOPSIS scores, which implies that posting at optimal times, combined with positive sentiment, has a significant impact on engagement potential. This analysis complements our sentiment analysis by quantifying the influence of sentiment polarity and timing on the overall effectiveness of the tweets, highlighting the importance of aligning both positive sentiment and strategic posting hours.