Sentiment Analysis of Indian Political Tweets

CIS8045 – Group 7

Nikhil Anne

Srivatsav Busi

Haripriya Janardhan Rao

Sai Chaitanya Nagandla

Sahithi Nallani

EXECUTIVE SUMMARY

Introduction:

Sentiment Analysis of Indian Political Tweets is a critical area of study that harnesses the power of natural language processing and data analytics to gauge public sentiment towards political figures, parties, and issues in India. With the widespread adoption of social media, Twitter has emerged as a prominent platform for political discourse and expression of public opinion. This executive summary provides an overview of the key findings, implications, and recommendations from our comprehensive analysis.

Problem Statement:

From the existing dataset, we want to analyze the sentiment of the Indian political tweets as the elections in India are round the corner happening in Summer 2024. Hence, this becomes an important analysis to gauge people's current sentiment which in turn could be used by political parties in their campaigns as well as media to understand the current situation.

Solution Methodology:

Initially, we wanted to understand the content in the dataset and proceed with the data cleaning and preprocessing using stemming and lemmatization methods. We then want to do the exploratory data analysis using Tableau wherein we want to understand the current trends from the data and to try and identify any interesting points.

Our next steps would be to plot the word clouds to understand the most frequently used words and to derive positive, negative, and neutral sentiments from them. We will then proceed to do the Topic modeling using Latent Dirichlet Allocation (LDA) and identify some of the top topics of discussions among the people and understand to what topics they belong.

Finally, we would perform machine learning using Support Vector Machine (SVM) and Naïve Bayes models to find the accuracies and determine the suitable model for our dataset.

Expected Conclusion:

Sentiment Analysis of Indian Political Tweets offers a valuable means of gauging public sentiment in a rapidly evolving political landscape. As India continues to engage in vibrant political discourse on social media, we would like to leverage sentiment analysis which can provide key insights to inform political strategies, media coverage, and decision-making processes.

INTRODUCTION

In the interconnected global village of the 21st century, the nature of political discourse has evolved profoundly. Gone are the days when political insights were exclusively sourced from rallies, town hall meetings, or newspaper op-eds. Today, the digital realm, led predominantly by social media platforms, serves as a significant influencer in shaping public sentiment and political inclinations. Among these platforms, Twitter stands out as a vibrant hub for real-time discussions, debates, and deliberations.

The power of Twitter in influencing public sentiment is undeniable. However, with millions of tweets pouring in every day, sifting through this colossal data to discern clear patterns, themes, and sentiments becomes a daunting task. Our study, "Sentiment Analysis of Indian Political Tweets," helmed by Group 7, ventures into this intricate maze of digital expressions. Through a systematic and rigorous analytical approach, we aim to unearth the predominant themes that resonate with the Indian public, understand the overarching sentiment of political discourse, and delve deeper into public reactions through likes, retweets, and comments.

The very essence of our project is timely and pivotal. With India gearing up for its elections, understanding the prevailing political sentiment becomes crucial. Not only does it aid political parties in refining their campaigns, but it also offers insights into public policy recommendations, predicting potential election outcomes, and understanding real-time sentiment shifts.

Relying on a vast dataset sourced from Kaggle, comprising many unique tweets, our research harnesses advanced methodologies ranging from Natural Language Processing (NLP) techniques, LDA topic modeling, to machine learning classifiers like the Multinomial Naive Bayes and the Support Vector Machine. Each step of our analysis, from meticulous data cleaning to comprehensive sentiment analysis, aims to present a holistic picture of the current political sentiment in India, as voiced on Twitter.

This report unfolds the journey of our analysis. As you delve deeper, you will witness the oscillations of public sentiment, the highs and lows of political discourse, and the intricate tapestry of topics that dominate the Twitter space. We invite you to join us on this insightful voyage as we decode the sentiments, understand the narratives, and grasp the essence of Indian political discourse on Twitter.

PROJECT BACKGROUND

As the world moves deeper into the digital age, the political landscape isn't left behind. The discussions and debates that once echoed in town halls, tea stalls, and living rooms have now found a new arena: social media platforms. Twitter, with its concise format and real-time interaction, stands out as a crucial platform in this digital revolution. It provides an environment where politicians, activists, influencers, and the common man converge to voice opinions, debate issues, and set narratives. Recognizing this paradigm shift, our team embarked on a journey to decode the intricacies of political sentiment on Twitter, aiming to distill patterns, emotions, and predominant themes from the vast ocean of tweets.

Description of the Problem

While Twitter offers a real-time reflection of public sentiment, its sheer volume and dynamism pose challenges. Every day, millions of tweets are shared, each carrying its own sentiment, context, and subtext. Filtering genuine public sentiment from noise, discerning patterns amidst the chaos, and understanding the ebb and flow of political sentiment in real-time are complex tasks. The challenge is multifold:

1. Volume & Diversity: The sheer number of tweets, coupled with the vast diversity of opinions, makes it challenging to identify predominant themes and sentiments.
2. Brevity & Nuance: Twitter's character limit means tweets are often concise, packed with nuances, internet jargon, and evolving lexicons that can make sentiment analysis intricate.
3. Authenticity: With the rise of bots, fake accounts, and orchestrated digital campaigns, distinguishing genuine sentiments from manufactured narratives becomes imperative.

Given these challenges, our primary problem statement was clear: How can we effectively analyze Twitter data to extract genuine political sentiments, identify predominant discussion themes, and offer actionable insights into the Indian political landscape?

DATA DESCRIPTION

<u>Sources of Data:</u>

Our study's backbone is a comprehensive dataset sourced from Kaggle. As a global platform for data science and analytics, Kaggle is renowned for its vast repository of authentic datasets spanning various domains. Choosing a dataset from such a reputable platform ensures the authenticity and reliability of our data, setting a solid foundation for our analysis.

Description of Data Fields:

1. Date: More than just a timestamp, the date provides chronological context. This allows us to understand temporal trends, correlate tweet sentiments with real-world events, and track the progression of political narratives over time.

2. User: This field offers more than just identification. By analyzing user data, we can identify influential figures, detect patterns among frequent tweeters, and even uncover potential bots or fake accounts that might skew our analysis.

3. Tweet: At the heart of our dataset lies the content of the tweets. Each tweet is a potential goldmine of sentiment, opinion, and insight. Using advanced Natural Language Processing techniques, we delve into these tweets to extract emotions, identify main discussion themes, and categorize sentiments.

4. Likes: This metric offers a direct insight into a tweet's acceptance and popularity. A high number of likes can indicate alignment with public sentiment, while a low count might hint at controversial or niche opinions.

5. Retweets: A measure of a tweet's reach and impact. The number of retweets can indicate the influence of the message, its viral nature, and its potential to shape public discourse.

With this robust dataset spanning 49,570 unique tweets, we embark on our analytical journey, aiming to unravel the rich tapestry of Indian political sentiment on Twitter. Through our rigorous methodologies, we hope to offer a comprehensive picture of the political landscape as seen through the tweets of its netizens.

	Date	User	Tweet	Likes	Retweets
0	2023-03-29 15:42:36+00:00	AnandPatni8	@vinodkapri @RahulGandhi Respected Indian Citizens, Namaskaar🙏I am the original Gandhi. I have no branches or franchises. None of my relatives are in politics. Beware of fake Gandhi's. Take care🙏 Mohandas Karamchand Gandhi.	0	0
1	2023-03-29 15:42:05+00:00	dhinamum	*Respected Indian Citizens,* Namaskaar I Am The Original Gandhi. I Have No Branches Or Franchises. None Of My Relatives Are In Politics. Beware Of Fake Gandhi's. Take Care,,,_ https://t.co/0OFay52fqP	0	0

METHODOLOGY

1. Data Cleaning:

To ensure the reliability and accuracy of our analysis, it's crucial that our dataset is pristine. Our process identifies and removes any missing values, duplicates, and non-essential elements like URLs. By cleaning the data, we create a robust foundation for the subsequent stages of our analysis.

2. Text Normalization:

Given the diverse ways users express themselves on Twitter, text normalization is essential. By implementing techniques like lemmatization and stemming, we ensure that variations of the same word are standardized. This crucial step ensures consistency and aids in achieving a more accurate sentiment analysis.

3. Sentiment Analysis:

At the heart of our project is sentiment analysis. Leveraging the capabilities of TextBlob, each tweet is categorized as 'Positive', 'Negative', or 'Neutral'. This categorization paints a preliminary picture of the prevailing sentiments in the vast political discourse on Twitter.

4. Exploratory Analysis:

Beyond mere sentiment, we venture deeper to understand patterns. This phase involves examining the frequency of tweets, identifying top influencers in the political sphere, and tracking the ebb and flow of sentiments over time. Such insights provide context & depth to the raw sentiment data.

5. Topic Modeling:

Politics is multifaceted, and Twitter discussions span myriad topics. Using the LDA (Latent Dirichlet Allocation) technique, we discern core discussion themes. This allows us to pinpoint primary subjects and issues resonating among the public, offering a lens into the heart of political discussions.

DATA CLEANING

The essence of a robust sentiment analysis lies not just in sophisticated algorithms, but also in the quality of the data fed to these algorithms. The initial stage of our methodology is dedicated to data cleaning, ensuring our dataset's purity. We employ the `df.isna().sum()` method to rapidly detect and evaluate any missing data within our records. Such missing or null data can be detrimental, introducing noise and inaccuracies. To combat this, we utilize `df.dropna(inplace=True)`, efficiently discarding these incomplete records, thus maintaining the integrity of our dataset.

However, maintaining data integrity does not stop at handling missing values. The presence of duplicated entries can skew our analysis, emphasizing certain sentiments or topics inaccurately. By deploying `df.duplicated().sum()`, we swiftly identify and eradicate these duplicate entries, ensuring that each tweet in our dataset is unique and counts just once.

Lastly, the aim is to dive deep into the text of each tweet. To that end, we employ processes to enhance the readability and relevance of tweets. By systematically excluding Twitter handles and URLs, we ensure our analysis is centered solely on the core content of each tweet, free from distractions or redundant information. This meticulous data cleaning approach sets the stage for an informed, precise, and reliable sentiment analysis.

Text Preprocessing & Refinement

Tweets often contain extraneous characters, punctuations, and widespread words like "and" or "the" that don't add sentiment value. Discarding irrelevant characters and punctuations declutters each tweet. Meanwhile, eliminating stopwords - common words without sentiment significance - amplifies the core message of the tweet. This process ensures that our sentiment analysis is based on the most meaningful and relevant content.

User	Tweet	Likes	Retweets	Original_Tweet	DateTime	date	month	year	hour
AnandPatni8	respected indian citizens namaskaar I original...	0.0	0.0	@vinodkapri @RahulGandhi Respected Indian Citi...	2023-03-29 15:42:36	2023-03-29	3	2023	15
dhinamum	respected indian citizens namaskaar I original...	0.0	0.0	*Respected Indian Citizens,* Namaskaar I Am Th...	2023-03-29 15:42:05	2023-03-29	3	2023	15
PrincetonCGI	1 meet filmmaker prakash jha new jersey talkin...	0.0	0.0	1/n-Meet Filmmaker Prakash Jha in New Jersey t...	2023-03-29 15:34:29	2023-03-29	3	2023	15
RishiJoeSanu	would politicians stop using religion politics...	0.0	0.0	@MrinalWahal Why would politicians stop using ...	2023-03-29 15:31:43	2023-03-29	3	2023	15
itweetsensee	state level president knows policy pm union mi...	0.0	0.0	@annamalai_k @narendramodi A state level presi...	2023-03-29 15:26:48	2023-03-29	3	2023	15

TEXT NORMALIZATION

A.WordNetLemmatizer Implementation:

The diverse nature of language means that a single idea or action can be represented by various word forms. The WordNetLemmatizer tool addresses this challenge by converting each word to its base or dictionary form. For instance, “running,” “runs,” and “ran” all reduce to the base word “run.” This process, known as lemmatization, ensures that our data is consistent, enabling more accurate sentiment analysis. By standardizing word variations, the WordNetLemmatizer promotes clarity and uniformity in textual data.

B.Snowball Stemmer Application:

While lemmatization focuses on using a dictionary base for word consistency, stemming takes a slightly different approach. The Snowball Stemmer, renowned for its efficiency, truncates words to a root form by removing prefixes or suffixes. This might not always produce dictionary-standard words, but it's effective for reducing word variations. For example, “happily” and “happiness” might both be stemmed to “happi.” Through this process, we can streamline the content and size of our dataset, ensuring that similar sentiments aren't scattered across different word variations but are consolidated for a more robust analysis.

User	Tweet	Likes	Retweets	Original_Tweet
AnandPatni8	respect indian citizens namaskaar I original g...	0.0	0.0	@vinodkapri @RahulGandhi Respected Indian Citi...
dhinamum	respect indian citizens namaskaar I original g...	0.0	0.0	*Respected Indian Citizens,* Namaskaar I Am Th...
PrincetonCGI	1 meet filmmaker prakash jha new jersey talk s...	0.0	0.0	1/n-Meet Filmmaker Prakash Jha in New Jersey t...
RishiJoeSanu	would politicians stop use religion politics i...	0.0	0.0	@MrinalWahal Why would politicians stop using ...
itweetsensee	state level president know policy pm union min...	0.0	0.0	@annamalai_k @narendramodi A state level presi...

SENTIMENT ANALYSIS

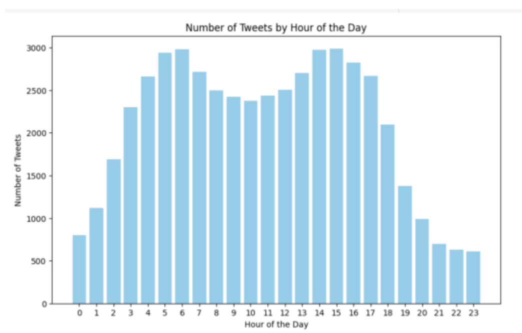
In the digital realm of social media, the sentiment behind each tweet can offer valuable insights into public opinion. To decipher this sentiment, we employ TextBlob, a powerful Python library tailored for textual data processing. Upon feeding each tweet from the 'Tweet' column to TextBlob, it calculates a sentiment polarity score. This score, ranging from -1 to 1, represents the emotional tone of the tweet. A score closer to -1 indicates a negative sentiment, a score near 1 implies positivity, and scores around 0 suggest neutrality.

Once the polarity for each tweet is computed, the task shifts to classification. By setting specific thresholds for the polarity scores, tweets are categorized. Those with a score greater than 0.05, for instance, might be marked as 'Positive,' while scores less than -0.05 could be tagged as 'Negative.' Tweets with scores falling between these bounds are labeled 'Neutral.' This classified sentiment is then stored in a dedicated 'Sentiment' column. This systematic categorization not only simplifies the data interpretation process but also paves the way for further analyses like trend identification or sentiment shifts over time.

User	Tweet	Likes	Retweets	Polarity	Sentiment
AnandPatni8	respect indian citizens namaskaar l original g...	0.0	0.0	-0.062500	Negative
dhinamum	respect indian citizens namaskaar l original g...	0.0	0.0	-0.062500	Negative
PrincetonCGI	1 meet filmmaker prakash jha new jersey talk s...	0.0	0.0	0.173232	Positive
RishiJoeSanu	would politicians stop use religion politics i...	0.0	0.0	0.000000	Neutral
itweetsensee	state level president know policy pm union min...	0.0	0.0	0.214286	Positive

EXPLORATORY DATA ANALYSIS

1. Number of tweets posted by hour of the day



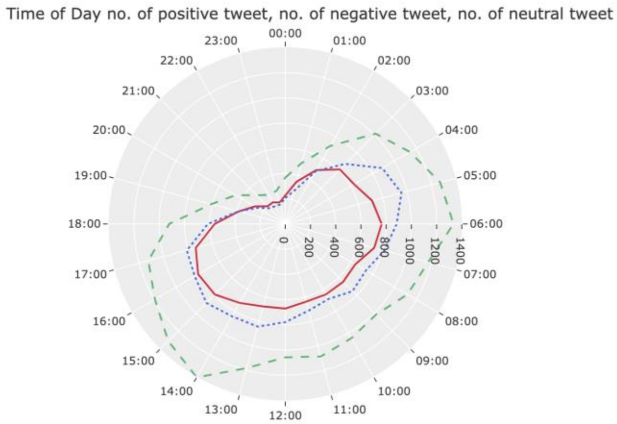
1. Activity Peak in Mid-Day: One of the most striking patterns observed is the surge of activity during the mid-day hours. Between 14:00 (2 PM) and 15:00 (3 PM), there's a pronounced increase in tweets.

2. Dip in Early Hours: The lowest number of tweets occur in the early hours of the day, particularly around midnight (hour 0).

3. Gradual Evening Decline: Post mid-day peak, the number of tweets gradually declines as the evening progresses, reaching a trough after midnight.

4. Consistent Morning Rise: Starting from the early hours, there's a consistent increase in tweets until it reaches the aforementioned mid-day peak. This indicates morning to afternoon as the active period for users.

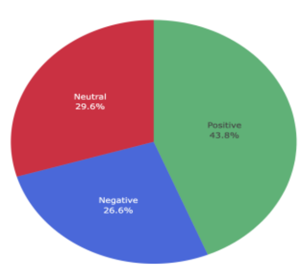
2. Daily Sentiment Trends in Tweets



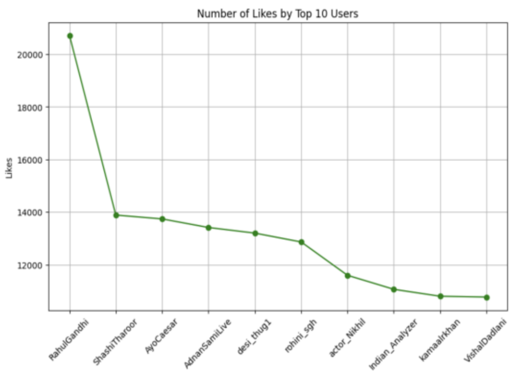
1. Sentiment hourly Peak Positive Sentiment in Midday: The highest number of positive tweets are observed during the 14:00 (2 PM) hour, indicating a possible peak of positive sentiment in the early afternoon.
2. Decreasing Sentiments at Night: All sentiment categories (Positive, Neutral, Negative) show a decline in tweet numbers during late-night hours, indicating fewer people tweeting or fewer instances of sentiment-specific content.
3. Overall Sentiment Distribution: The frequency of positive tweets seems to be higher compared to negative and neutral during most parts of the day. This might suggest a generally positive outlook or positive content trending on the platform during these times.

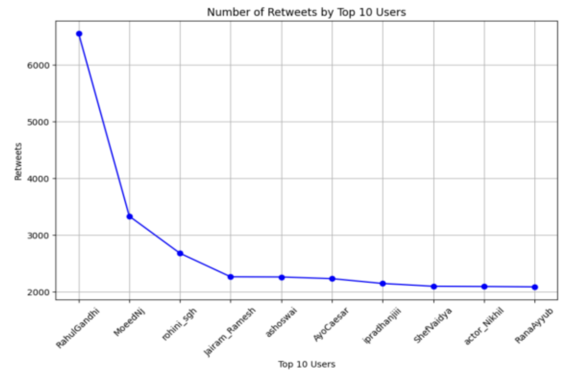
3. Distribution of tweet sentiments

We have generated a pie chart to understand the sentiment distribution. From the below chart, we can see that our dataset is well-balanced with 44% of the tweets carrying a positive sentiment and Neutral sentiments at nearly 30%, slightly surpass negative sentiments which stand at about 26%.



4. Analysis of Top User Interactions





The first line graph visualizes the sum of likes received by the top 10 users, highlighting their popularity. The subsequent graph showcases the aggregate retweets of the top 10 users, offering insight into the users' influential content. We can clearly observe ‘Rahul Gandhi’ who is the opposition leader has the greatest number of likes & retweets signifying his positive presence on Twitter.

Word Cloud Generation for Tweets

We have generated word clouds to visually represent the most used words for different sentiments in tweets. Words such as "BJP," "Congress," "cricket," and "Modi" are recurrent across all sentiments, indicating widespread discussions on these topics.

- In the negative sentiment, words like "politician" and "need" appear, possibly highlighting dissatisfaction or concerns with present politicians and perceived shortcomings.

- The positive sentiment prominently features "good" and "team," hinting at positive acknowledgments or praises, possibly in contexts like teamwork or sports achievements.

- In the Neutral sentiment, common verbs like "go," "get," "make," and "say" are used across different sentiments, making them neutral in conveying emotion on their own.

The word cloud you provided shows the most common words in a set of political tweets. The largest words in the word cloud are "leader," "state," "one," "make," "BJP," "get," "go," "Like," ,"congress”, ”cricket," "time," "want," "politician," "party," "modi," "know," "also," "team," "say," and "history."

This word cloud suggests that the tweets are focused on the following topics:

- The leadership of the BJP party
- The current state of the country
- The need for change

TOPIC MODELING OF TWEETS USING LDA

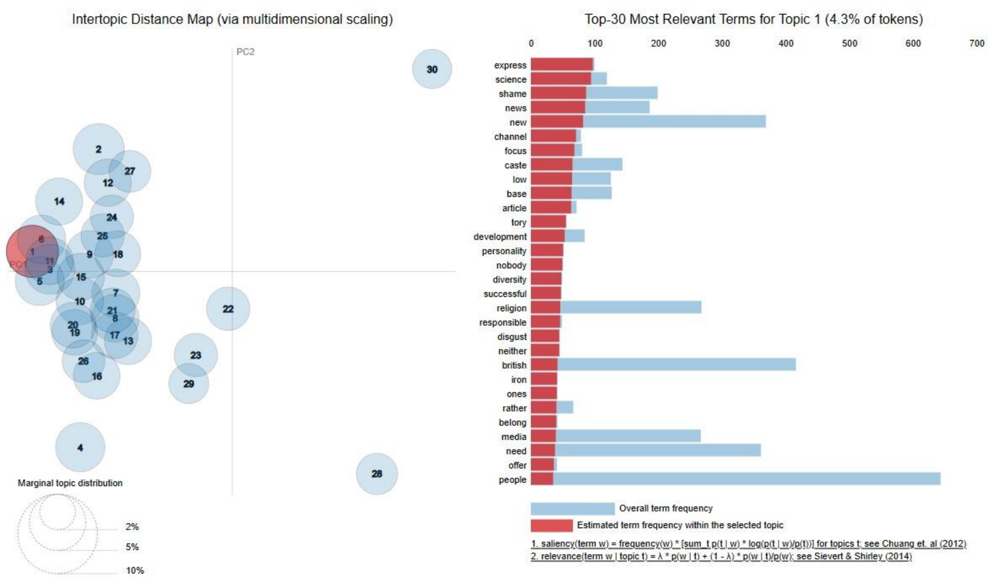
Latent Dirichlet Allocation (LDA) – a widely employed technique for uncovering latent topics within a corpus of text. The initial step involved preprocessing the tweets, a process that included tasks such as removing special characters, stopwords, and converting the text to lowercase. Subsequently, the preprocessed tweets were transformed into a structured list of words, facilitating further analysis. To facilitate the modeling process, a dictionary and a corpus were constructed using the Gensim library. The dictionary served as a comprehensive list of unique words present in the corpus, while the corpus represented a numerical representation of the text data, pivotal for subsequent LDA modeling. The final phase of the project involved implementing the LDA model, configured to identify 30 distinct topics within the corpus. Each topic was then characterized by a collection of 20 most significant words, providing valuable insights into the underlying themes present in the tweet dataset. This comprehensive approach enabled a nuanced exploration of the diverse array of topics prevalent in the dataset, ultimately contributing to a deeper understanding of the content and sentiment within the Twitter data.

- Preprocessing the tweets from a dataframe, converting them into a list of words
- Creating a dictionary and corpus for the words using the Gensim library
- An LDA model is built to identify 30 topics from the corpus, with each topic showcasing a set of 20 most significant words.

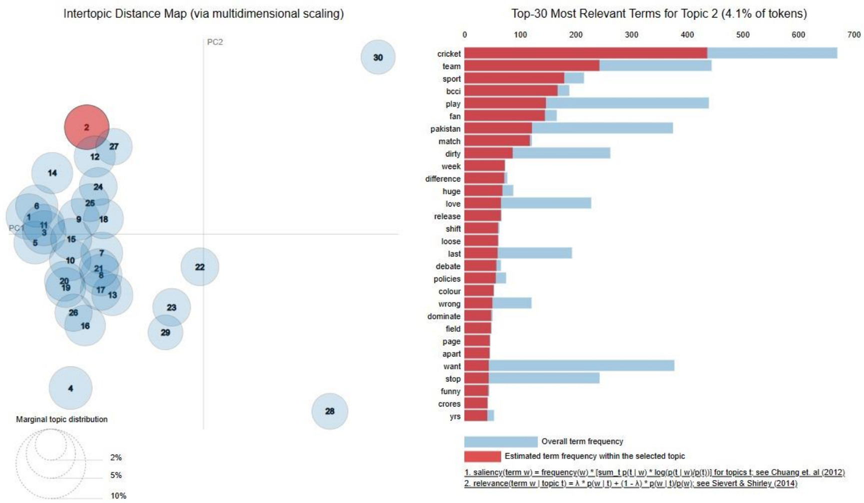
Inferences from LDA

Overall, this topic model visualization provides a valuable overview of the main topics of the corpus, as well as the relationships between those topics. This information can be used to inform a variety of tasks, such as document summarization, text classification, and question answering.

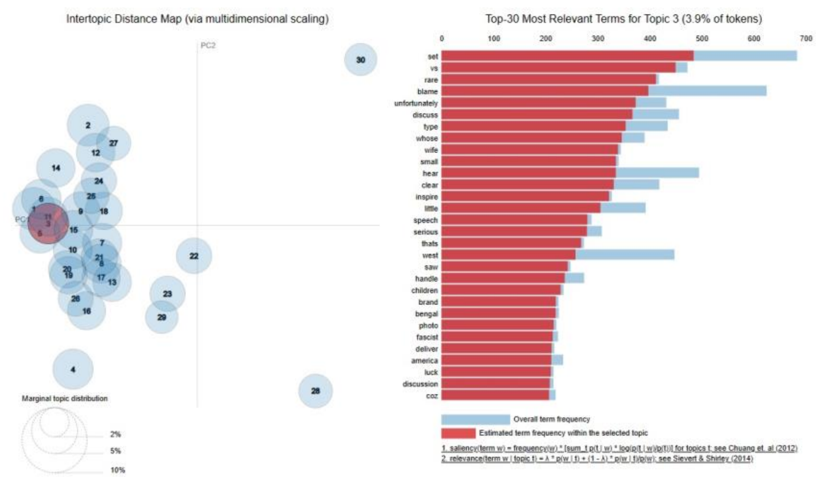
Topic-1: Social Issues and Media Influence- The topic highlights societal discussions shaped by media narratives. It delves into modern challenges like technology and its implications.



Topic-2: Sports, Events, and National Relations -The conversation is heavily centered around cricket, emphasizing its cultural significance in India. The term 'debate' suggests potential discussions on controversial sports events or decisions.



Topic- 3: Personal and Regional Relations- This topic encapsulates a fascinating interplay between discussions on personal relationships and the intricacies of regional identities. Within this topic, the narratives woven by the text data suggest a profound connection between individual experiences and the broader context of one's regional affiliations.



TF-IDF (term frequency-inverse document frequency) transformation

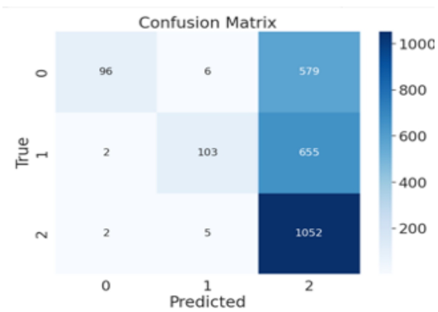
We have performed Sentiment analysis using TF-IDF representation of preprocessed data with Naïve Bayes and Random Forest Classifier.

```
In [143]: # Transform data by applying term frequency inverse document frequency (TF-IDF)
          from sklearn.feature_extraction.text import TfidfTransformer
          tfidf = TfidfTransformer() #by default applies "l2" normalization
          X_tfidf = tfidf.fit_transform(X_vec)
          X_tfidf = X_tfidf.todense()
          X_tfidf = np.asarray(X_tfidf)
          X_tfidf
```

TfidfTransformer class from the scikit-learn library is used to apply TF-IDF transformation to a vectorized text data represented by X_vec. The TF-IDF transformation calculates the importance of each term in each document based on its frequency and inverse document frequency and normalizes the scores. The resulting transformed data is stored in the X_tfidf variable.

Multinomial Naive Bayes

- Preprocessing text data using tokenization and TF-IDF, sample and split into training and testing sets.
- Training a Multinomial Naive Bayes classifier and calculating a confusion matrix to evaluate the performance.
- Amidst upcoming elections, the classifier accurately detected most tweets as positive, with 1052 correct picks**.**



	precision	recall	f1-score	support
0	0.96	0.14	0.25	681
1	0.90	0.14	0.24	760
2	0.46	0.99	0.63	1059
accuracy			0.50	2500
macro avg	0.77	0.42	0.37	2500
weighted avg	0.73	0.50	0.41	2500

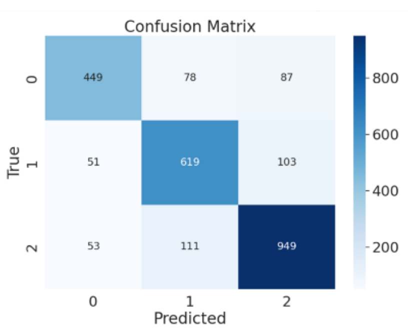
- The classifier reliably predicts negative and neutral sentiments but often misses most of them (high precision, low recall for classes 0 and 1).
- It heavily favors positive sentiment, correctly identifying most but with many false positives (high recall, lower precision for class 2).
- The overall accuracy stands at 50%, indicating mixed reliability.
- Given the upcoming elections, the results suggest a skewed perception towards positive sentiment, possibly influenced by dataset composition.

Support Vector Machine Classifier

The Support Vector Machine (SVM) Classifier employed in this project has demonstrated commendable performance across various sentiment categories. The classifier exhibits a remarkable consistency in its predictions, achieving an impressive overall accuracy rate of 81%.

- Notably, it excels in discerning positive sentiments (class 2), boasting both high precision and recall scores. This indicates a proficiency in identifying positive sentiment instances.
- Furthermore, the SVM exhibits a balanced performance in classifying negative (class 0) and neutral (class 1) sentiments, with comparable precision and recall values for both categories.
- This balanced classification performance is particularly noteworthy, as it suggests that the dataset encapsulates a diverse spectrum of political opinions and sentiments. Given the forthcoming elections, this balanced classification is indicative of the dataset's potential to provide valuable insights into the multifaceted landscape of public sentiment and political discourse surrounding the event.

The SVM classifier's robust performance underscores its effectiveness as a valuable tool for sentiment analysis within the context of political discussions.



	precision	recall	f1-score	support
0	0.81	0.73	0.77	614
1	0.77	0.80	0.78	773
2	0.83	0.85	0.84	1113
accuracy			0.81	2500
macro avg	0.80	0.79	0.80	2500
weighted avg	0.81	0.81	0.81	2500

Model Performance Evaluation (Naive Bayes vs SVM)

1. **Higher Accuracy:** SVM boasts 81% accuracy, significantly outperforming Naive Bayes at 50%.
2. **Balanced Predictions:** SVM shows consistent precision and recall, whereas Naive Bayes displays a positive sentiment bias**.**
3. **Better Classification**: SVM's confusion matrix reveals more accurate sentiment categorizations compared to Naive Bayes.
4. **F1-Score:** SVM has a superior F1-score, indicating a harmonized balance between precision and recall, essential for real-world applications.

This comprehensive evaluation unequivocally establishes the superiority of the SVM classifier over Naive Bayes in the context of sentiment analysis for this dataset. The robust performance of the SVM model positions it as the preferred choice for extracting nuanced sentiment insights from textual data in similar applications.

Key learnings from our project

- The frequent mentions of "BJP," "Congress," and "Modi" in discussions highlight their central role in political discourse, suggesting that BJP and Congress are the primary contesting parties, with Modi being a key political figure.
- A significant surge in tweet activity is observed between 2 PM and 3 PM, indicating a specific peak in engagement during that hour.
- The consistent reference to "cricket" alongside major political terms underscores the cultural significance of cricket in India, even within political discussions.
- SVM demonstrates effectiveness in establishing a clear decision boundary, especially when dealing with well-separated data, whereas Naive Bayes relies on probabilistic calculations that may be influenced by class imbalances in the training dataset.
- Interestingly, Rahul Gandhi, the opposition leader in India, has garnered the most likes and retweets, possibly indicating a prevailing anti-incumbent sentiment in anticipation of upcoming elections.

Future Scope & Managerial Implications

- Real-time Political Opinion Monitoring: Developing systems for real-time monitoring of political sentiment can provide insights into public opinion trends during elections, policy changes, and political events. This can be valuable for political parties and government agencies.
- Election Forecasting: Sentiment analysis can be used to predict election outcomes based on the sentiment expressed in tweets. By analyzing a large volume of political tweets, patterns and trends in public sentiment can be identified and correlated with election results.
- Sentiment Analysis for Journalism: Journalists can use sentiment analysis tools to track public sentiment on political issues and use this information to inform their reporting. It can also help in identifying trending topics.