Advances in Consumer Research
https://acr-journal.com/
Volume-3 | Issue-2 | Feb 2026

Original Researcher Article

Indian Aggression detection through multiple ML models from Twitter data

Dr Lakshmi Vishnu Murthy Tunuguntla [1]* , Dr P Balasubramanyam[2], Ms Aakruthi Amrut Alarnkar[3]
[1]Professor and Accridtation Chair person, Goa Institute of Management, Goa, India
Email ID : Vishnu.Tunuguntla@gim.ac.in
[2]Associate Professor and Program Chair for Big Data Analytics, Goa Institute of Management, Goa, India
[3]Assistant Professor,Goa Business School Goa University Taleigao Plateau, 403206

ABSTRACT

The purpose of this research is to focus on aggressive communication, that gets triggered by emotionally or politically charged events especially in online space. The study investigates whether machine learning models can provide a reliable method for identifying various forms of aggression in Indian Twitter posts, and whether notable patterns of aggressive behaviour are linked to certain categories of events.

In order to understand the above phenomenon, five high-impact events from different domains social, financial, sporting, and political were selected for study and analysis to see their capacity to invoke strong public reactions on twitter.

About 13,000 tweet data related to each of these events was collected using the Python-based snscrape tool, were collected and processed. The aggression was divided in to three categories namely Overtly Aggressive (OAG), Covertly Aggressive (CAG), and Non-Aggressive (NAG)

Out of the four selected supervised leaning models (Random Forest, Support Vector Classifier (SVC), Logistic Regression, and Multinomial Naïve Bayes), Multinomial Naïve Bayes demonstrated the most balanced and effective results, particularly in handling the nuances of covert versus overt aggression. The study showed events related to OAG generated highest volume

This study demonstrates how machine learning can be leveraged to detect and interpret public aggression in complex, multilingual environments like India's digital landscape. Beyond classification, the research provides insight into how aggression manifests across different societal issues and over time. The findings have practical implications for improving online moderation systems, guiding responsible communication policies by the government, and informing future research into the psychology and sociology of digital interactions. By focusing on local context and diverse event categories, the study makes a contribution to computational social science and paves the way for more culturally attuned AI applications...

Keywords: Aggression Detection, Machine learning models, Text Classification

INTRODUCTION:

The digital transformation of communication has revolutionized how people interact, share opinions, and engage in public discourse. Social media platforms have been at the heart of this shift, and Twitter, in particular, has emerged as a leading space for individuals to voice thoughts in real time. Its brevity and openness make it a useful tool for studying public expression, especially during events that stir strong emotions. However, this ease of expression also allows for the rapid spread of hostility and aggressive rhetoric, which can escalate tensions in already sensitive situations.

In India, where linguistic and cultural diversity intersect with deep political engagement, social media plays a complex and often controversial role. With one of the world's largest online populations, India's digital landscape reflects a vibrant yet sometimes volatile mix of voices. These platforms empower users to participate in discussions about politics, society, entertainment, and more. But they also provide a space where aggression and conflict can grow unchecked, often under the veil of anonymity. Public reactions to national events increasingly unfold online, with Twitter serving as a microcosm of wider social dynamics.

Recent episodes have shown how quickly digital conversations can take on an aggressive tone. Cases such as the Shraddha Walker murder or the Adani-Hindenburg controversy triggered waves of reactions on social media, many of which were emotionally charged and aggressive in nature. These responses range from open support or outrage to covert hostility expressed through sarcasm or coded language. The polarized nature of Indian public discourse—combined with the immediacy of Twitter—makes it crucial to understand how aggression manifests and spreads in this space.

The theory of online disinhibition helps explain why individuals often express themselves more freely or aggressively online than they would in person. The perceived distance, lack of immediate consequences, and the absence of non-verbal cues all contribute to behavior that is more impulsive or extreme. As a result, both overt hostility (like direct abuse or threats) and subtler forms of aggression (such as veiled insults or sarcasm) have become increasingly common. This duality makes it essential to examine not only what is said online, but how it is said, and under what circumstances.

Analyzing aggression on platforms like Twitter presents several challenges. The informal and ever-changing nature of online language—including code-mixing, emojis, slang, and abbreviations—makes traditional text analysis tools less effective. Additionally, the same word or phrase may convey different sentiments depending on cultural context, intent, or the event being discussed. Detecting aggressive behavior therefore requires more than just keyword searches; it demands sophisticated models that can grasp tone, intent, and context.

Another layer of complexity comes from the speed at which public sentiment shifts online. Twitter reactions are often tied closely to real-time events—breaking news, political announcements, or public figures' statements can lead to rapid spikes in activity. This volatility means that aggression isn't just a constant background feature; it rises and falls based on current affairs. Understanding how these spikes occur and how long they last can offer important insights into the dynamics of online hostility.

While studies on hate speech and sentiment analysis have grown in recent years, much of the existing research focuses on English-language data from Western contexts. This leaves a significant gap when it comes to understanding aggression in multilingual, culturally specific settings like India. Indian users often communicate in a mix of English and regional languages—what's known as code-mixed text. This blend presents additional hurdles for machine learning models that aren't designed to handle linguistic fluidity or cultural nuance.

This research was motivated by both academic curiosity and a practical need to address this gap. From a research standpoint, there's growing interest in improving natural language processing (NLP) models to detect complex patterns of behavior like online aggression. From a practical angle, social media companies, policymakers, and educators all face growing pressure to understand and address the harms of digital hostility. Tools that can detect and monitor aggressive discourse could support better moderation systems, educational interventions, and mental health initiatives.

To explore these issues, this study focuses on five recent and widely discussed events in India: the Shraddha Walker case (a criminal and social event), the Adani-Hindenburg report (finance-related), India's T20 World Cup loss (sports), the Bharat Jodo Yatra led by Rahul Gandhi (national politics), and the formation of the Shinde government in Maharashtra (state politics). These events represent different domains that trigger varying emotional responses and engagement levels, making them ideal for examining online aggression patterns.

The study is structured around two central aims. First, it evaluates multiple machine learning models—specifically Random Forest, Support Vector Machines, Logistic Regression, and Multinomial Naïve Bayes—to determine which performs best for classifying aggressive content. Second, it applies the top-performing model to a large dataset of tweets related to the five events, categorizing them into three types: Overtly Aggressive (OAG), Covertly Aggressive (CAG), and Non-Aggressive (NAG). This detailed classification allows for a more layered understanding of online hostility.

Moreover, this research includes a time-based analysis to observe how aggression evolves as public reactions unfold. By tracking shifts in aggression over the timeline of each event, we aim to identify whether such behavior spikes immediately, fades gradually, or follows a more complex pattern. Understanding these trends can help in predicting future online responses to similar events and crafting better strategies to address them.

Ultimately, this research bridges computational techniques and cultural insights. By applying NLP and machine learning within an Indian context, it contributes not just to technical literature but also to broader conversations about public behavior, digital ethics, and responsible communication. It reflects a step toward building context-aware models that can navigate the messy, multilingual, and emotionally complex world of social media.

2.LITERATURE REVIEW

In the last decade, the expansion of social media platforms has transformed the way individuals communicate and express themselves publicly. Twitter, in particular, has emerged as a dynamic space for opinion-sharing, owing to its brief, real-time posting style. While this democratization of speech has fostered open discussions, it has also created an environment where hostility, aggression, and abusive behavior can thrive. The unfiltered nature of these platforms has piqued the interest of researchers seeking to understand digital aggression and its patterns, especially in politically and socially charged contexts such as India.

Aggression and Online Behavior

The increasing visibility of aggressive online behavior—including hate speech, trolling, and cyberbullying—has spurred a surge in academic investigations. Automated detection of such content is now considered crucial due to the limitations of manual moderation. Sadiq et al. (2021) explored machine learning-based techniques for identifying aggressive behavior in Twitter posts, using models like CNN-LSTM and CNN-BiLSTM. Their work highlighted how deep neural networks can improve detection accuracy by capturing the contextual nuances of online language.

Lepe-fa et al. (2021) focused on aggression detection in Spanish-language tweets. By comparing models across corpora from Mexico and Chile, they showed that hybrid models—those combining classical machine learning and deep learning—yielded better performance, especially in non-English datasets. This suggests the importance of language-specific considerations in aggression detection.

Distinction Between Sentiment and Aggression

Although sentiment analysis and aggression detection are often used interchangeably, the two are conceptually distinct. Sentiment analysis typically classifies opinions as positive, negative, or neutral. In contrast, aggression detection seeks to identify harmful or hostile language, which may not always align with sentiment polarity.

Passonneau (2011) contributed early models for binary and multi-class sentiment classification in tweets, paving the way for more complex classification systems in NLP.

More recently, researchers have applied sentiment analysis across various domains, such as business, politics, and health. For example, Philander and Zhong (2016) investigated public sentiment about hospitality brands using Twitter data, while Coello et al. (2022) used similar techniques to examine emotional reactions to wildfires in Spain and Portugal.

Singh et al. (2018) conducted sentiment analysis in the Indian context, focusing on the aftermath of the 2016 demonetization policy. Their findings emphasized how digital media platforms offer valuable insight into citizen sentiment during periods of economic and policy disruption.

Cross-Domain Applications

The utility of sentiment and aggression analysis has grown beyond traditional boundaries. During the COVID-19 pandemic, for example, researchers such as Ridhwan and Hargreaves (2021) used Twitter data to track public sentiment toward lockdown policies and vaccination campaigns. Other studies, such as those by Xu et al. (2022) and Tayef et al. (2022), delved into perceptions surrounding specific vaccine brands. These works collectively demonstrate how digital conversations can be mined to understand public attitudes during crises.

In the financial domain, while formal studies on events like the Adani-Hindenburg controversy remain limited, prior work has shown that Twitter sentiment can serve as a proxy for market sentiment, affecting investor decisions and corporate reputations.

The political sphere has also seen significant engagement from researchers. Ramon et al. (2022) examined discourse around remote work during the pandemic, illustrating how platforms like Twitter can reflect shifts in work culture and policy response. Hema et al. (2022) focused on public backlash to governmental projects in Indonesia, capturing dissent and approval in real time.

Computational Techniques

Text classification for detecting aggression typically relies on supervised learning models trained on labeled datasets. Standard approaches include algorithms like Support Vector Machines (SVM), Naïve Bayes, Logistic Regression, and Random Forests. Sahayak (2015) demonstrated the effectiveness of such models in sentiment classification tasks involving Twitter data.

However, the limitations of these traditional models have led researchers to adopt deep learning techniques that offer higher precision by capturing deeper semantic relationships in language. CNNs and LSTMs, especially in hybrid forms like CNN-LSTM, have shown considerable promise in processing informal, real-time text such as tweets. More recent advancements in NLP have introduced transformer-based models like BERT and RoBERTa, which provide context-aware language understanding. These models have set new benchmarks for text classification, though their resource demands remain high.

Despite their potential, many advanced models face challenges when applied to Indian social media data. Tweets often include code-mixed language, combining English with Hindi or other regional tongues. This makes preprocessing and vectorization more complex, as the models must interpret multiple syntactic and phonetic systems simultaneously.

Event-Based and Contextual Analysis

Recent years have seen a shift toward domain-specific and event-driven sentiment analysis. Rather than building general-purpose models, researchers are increasingly tailoring their approaches to specific issues or events. This is particularly valuable in analyzing social responses to incidents such as the Shraddha Walker case or India's T20 World Cup loss. Event-driven analysis allows researchers to trace sentiment over time and evaluate how public emotions evolve from the initial incident to its aftermath.

A similar methodology was employed by Nguyen et al. (2018), who correlated online sentiment toward ethnic minorities in the U.S. with health outcomes such as low birth weights. Their work suggested a powerful link between digital discourse and real-world social dynamics.

Understanding Different Types of Aggression

One underdeveloped area in aggression research is the differentiation between types of hostility. Most studies adopt a binary framework: content is either aggressive or not. However, real-world communication often involves subtleties. Khandelwal and Kumar (2020) proposed a tripartite classification system to account for this nuance—Overtly Aggressive (OAG), Covertly Aggressive (CAG), and Non-Aggressive (NAG). This approach acknowledges that aggression can be explicit, such as insults or threats, or implicit, such as sarcasm or passive-aggressive language.

This distinction is particularly relevant in a country like India, where indirect and context-rich communication is prevalent. Sarcasm, wordplay, and rhetorical questioning are common, and may not be captured by simple keyword-based models. Incorporating these categories allows for a more accurate and culturally aware model of aggression detection.

Literature Gaps and how the current study addresses them.

While existing literature has laid a solid foundation for automated aggression detection, several limitations persist. Firstly, there is a dearth of research that examines aggression in the Indian context across multiple events. Most studies tend to focus on a single issue or dataset. Secondly, few works conduct a rigorous comparative analysis of multiple machine learning models for aggression detection, especially using real-world, multilingual, and code-mixed data.

Temporal trends in aggression—how online hostility fluctuates in response to key events—are also rarely analyzed in depth. Understanding whether aggression increases immediately after an incident and then subsides, or follows a more complex trajectory, is essential for understanding the dynamics of digital outrage.

This study attempts to address these shortcomings in the following ways:

It applies four classical ML algorithms (Random Forest, SVM, Logistic Regression, and Multinomial Naïve Bayes) to a diverse, annotated dataset of tweets.

By comparing F1 scores and accuracy metrics, the study identifies the most effective model for the task (Multinomial Naïve Bayes in this case).

The selected model is then used to analyze aggression levels in five real-world incidents from varied domains: social (Shraddha Walker case), financial (Adani-Hindenburg), sports (India's T20 loss), and political (Bharat Jodo Yatra and Shinde Government formation).

Aggression is categorized into OAG, CAG, and NAG, aligning with contemporary frameworks that capture the complexity of online hostility.

Finally, the study observes how aggression levels change over time, revealing patterns that vary depending on the nature and longevity of each incident.

By combining computational methods with a culturally specific focus, this study not only contributes to the field of NLP and social media analytics but also offers insights into public behavior in digitally mediated spaces. It reinforces the idea that online aggression is not monolithic, but instead shaped by language, culture, and context.

3.RESEARCH METHODOLOGY

An important topic in natural language processing (NLP) is the analysis of aggression data, which may be used to develop models that recognize and classify hostile language in text. The Track 1 dataset has a lot of text data that can be used for this. The aggression data was analysed using the Track 1 dataset by pre-processing the text data. This entails deleting stop words, stemming the text, and tokenizing it. After preparing the data, it was examined and visualizing to get insights into the text's patterns and trends. Machine learning algorithms are commonly used to develop models that classify text as Overtly Aggressive (OAG), Covertly Aggressive (CAG), or Not Aggressive (Khandelwal & Kumar, 2020). An important topic in natural language processing (NLP) is the analysis of aggression data, which may be used to develop models that recognize and classify hostile language in text. The Track 1 dataset has a lot of text data that can be used for this. To accomplish this, Track 1 dataset was divided into training and testing sets and develop the model using a supervised learning method such as Random Forest Classifier, Linear SVC, Multinomial Naive Bayes, or Logistic Regression. We also used techniques like cross-validation and grid search to optimize the model's hyper parameters.

Description of each of the labels is as follows (Khandelwal & Kumar, 2020):

(1) Overtly Aggressive (OAG) - This type of aggressiveness involves a direct verbal attack directed at a specific individual or group. For example, "Well said, Sonu..you have the courage to stand up to Muslim dadagiri."

(2) Covertly Aggressive (CAG) - In this style of aggression, the attack is disguised, subtle, and more indirect, and is usually stated pleasantly. "Dear India, stop playing with your people's emotions for votes," for example.

(3) Non-Aggressive (NAG) - Generally, these sorts of writing lack any kind of aggression; they are primarily used to present facts, express wishes on occasion, and be polite and supportive.

The best model was used to conclude the level of aggression happening on the social media platform Twitter. Five incidents were considered to maintain the randomness in the tweets, and the users' responses were collected for further analysis. We carried out our analysis using the English Code-Mixed TRAC1 2018 dataset and the uni-lingual English dataset obtained from Kaggle2. The methodology followed was similar to (Soesanto et al., 2023) and (Ouertatani et al., 2018).

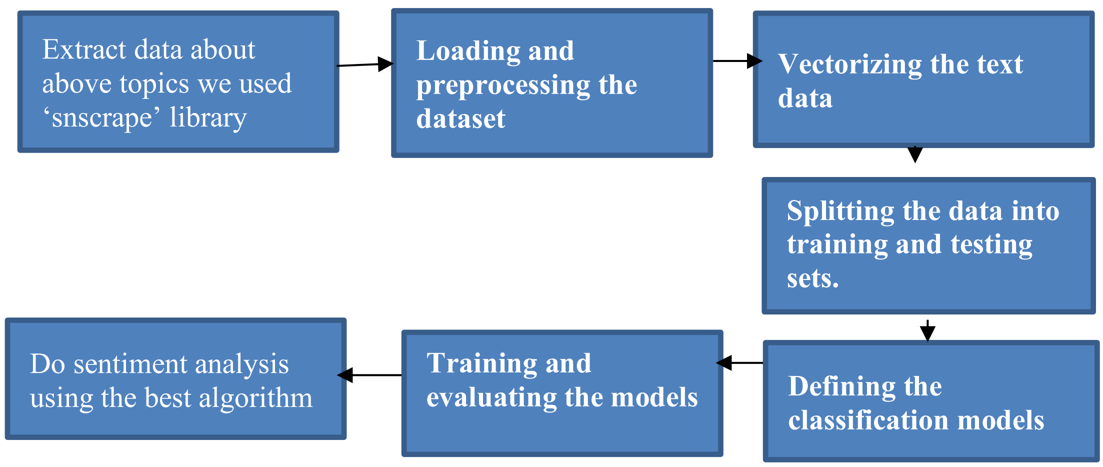


Figure 1: Research Methodology

Step 1: Data Extraction

We utilized the 'snscrape' package to extract data about the topics listed above. Users can extract data from Twitter by utilizing search queries. The library makes it easy to extract massive volumes of data from Twitter, making it an excellent tool for social media research. A search query is supplied to extract Data, which might include hashtags, dates, and other criteria. We added keywords between particular dates in our search query for our study project and extracted Tweet Date, Tweet Username, Tweet Content, Tweet User Location, Tweet User Profile Description, Tweet Retweet Count, and Tweet User Follower Count. This data is then transformed into a structured data format, such as a Pandas Dataframe, for further analysis (Or, 2018).

Step 2: Loading and pre-processing the dataset

Using regular expressions, the dataset was pre-processed by eliminating special characters and punctuation and converting all text to lowercase. This stage is critical since it aids in the preparation of the data for feeding to the machine learning model.

Step 3: Text data vectorization

Text data is transformed into numerical data. TfidfVectorizer is utilized since it is one of the most popular and efficient methods for converting text input to numerical form.

Step 4: Splitting the data

The dataset is divided into two parts: training and testing. Splitting the data aids in evaluating the model's performance on previously unseen data. It also aids in the prevention of overfitting, which occurs when the model becomes too complicated and begins to learn the noise in the data.

Step 5: Defining the classification models

A set of classification models to be trained and assessed is provided, which includes the Random Forest Classifier, Linear SVC, Multinomial Naive Bayes, and Logistic Regression. Each of these models has advantages and disadvantages, and comparing them allows you to choose the ideal one for the job.

Step 6: Training and evaluating the models

The for loop iterates through each model, trains it on training data, predicts labels for both training and testing datasets, and computes accuracy scores for both datasets. The accuracy score indicates how effectively the model is performing. If the accuracy score is low, it indicates that the model is underperforming and should be improved.

Emperical Analysis

The analysis is divided two main tasks. The first task is to analyse the available classification models with different parameters and choose the best classification model based on test and train data accuracy. The second task involves applying the selected model on the selected incident tweets to check Indians aggression levels.

Best model selection for classification:

For testing the performance of different models, we took track1 data set and applied the following five classification models.

Random Forest Classifier

Linear Support vector classifier

Multinomial Naïve Bayes Model

Logistic Regression.

4.Results & Discussion

The python code snippet for model building is

```
from sklearn.metrics import accuracy_score

X = df['Tweet_clean'] # Collection of documents
y = df['label'] # Target or the labels we want to predict
X_train, X_test, y_train, y_test = train_test_split
(features, labels, test_size=0.15, random_state = 0)

models = [
    RandomForestClassifier(n_estimators=100, max_depth=5, random_state=0),
    LinearSVC(),
    MultinomialNB(),
    LogisticRegression(random_state=0),
]
```

The train and test data accuracies are tabulated as below.

Table 1: Results of Models

Model	**Train Accuracy**	**Test Accuracy**	**F1 score test**	**F1 score train**
Random Forest Classifier	0.4286	0.4261	0.2621	0.2666
Linear SVC	0.8303	0.5506	0.5475	0.8295
Logistic Regression	0.7513	0.5867	0.5786	0.7477
Multinomial Naïve Bayes	0.6218	0.6226	0.6244	0.62306

From the above table it is evident that Multinomial Naïve Bayes model has highest F1 score test data i.e., 0.6244 and Random Forest classifier has the lowest F1 score of 0.2621. For for training data, Linear SVC is performing better compared to other models based on F1 score. Based on overall performance, we selected Multinomial Naïve Bayes model to apply for the selected cases in the following section

Aggression checks with the classification model

The incidents or events selected from different categories were analysed **i.e.** : The Shraddha Walker rape case [Social]; The Adani Hindenburg report [Finance]; India losing the T20 World Cup in 2022 [Sports]; Rahul Gandhi's Bharat Jodo yatra [Politics (central)]; and the formation of the Shinde Government in Maharashtra [Politics (Maharashtra state)]. A total of 1,31,822 tweets corresponding to the interactions made by users on the social network between in the city of were extracted from the social network Twitter. That is The Shraddha Walker rape case [18307]; The Adani Hindenburg report [29664]; India losing the T20 World Cup in 2022 [3906]; Rahul Gandhi's Bharat Jodo yatra [69474]; and the formation of the Shinde Government in Maharashtra [10471].

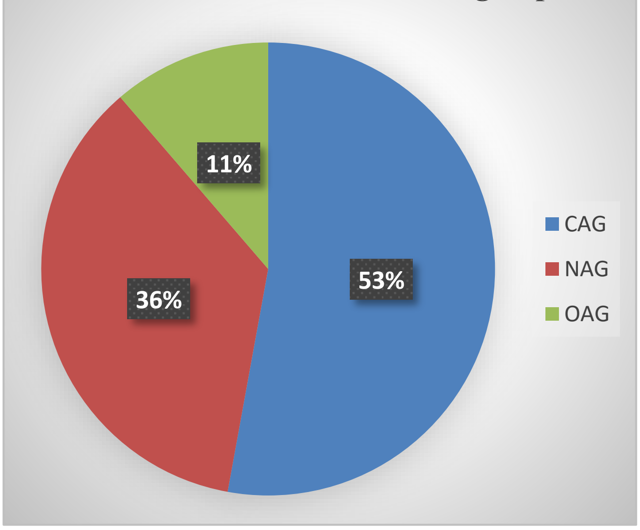


Figure 2: Distribution of agressive tweets for Adani Hidenburg Report

The above figure shows the distribution of aggressive tweets for Adani Hidenburg report. The above figure depicts that out of 29664 tweets about the above incident 11% of the tweets are Overtly Aggressive (OAG), 53% of the tweets are Covertly Aggressive (CAG) and 36% of the tweets are Not Aggressive (NAG). The frequency distribution of the tweets are specified in the table 2 below.

Table 2: Frequency distribution of aggressive tweets for Adani Hindenburg report.	
Adani Hindenburg Report	**Count**
Covertly Aggressive (CAG)	15681
Not Aggressive (NAG)	10631
Overtly Aggressive (OAG),	3352
Grand Total	**29664**
Source: Compiled by the author	

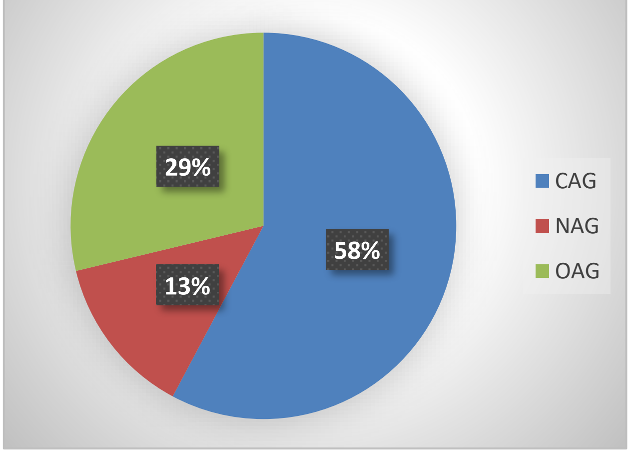


Figure 3: Distribution of aggressive tweets for Shraddha Walker Case

The above figure 3 shows the distribution of aggressive tweets for Shraddha Walker Case. The above figure depicts that out of 18,307 tweets about the above incident 29% of the tweets are Overtly Aggressive (OAG), 58% of the tweets are Covertly Aggressive (CAG) and 13% of the tweets are Not Aggressive (NAG). The frequency distribution of the tweets are specified in the table 3 below.

Table 3: Frequency distribution of aggressive tweets for Shraddha Walker Case	
Shraddha Walker Case	**Count**
Covertly Aggressive (CAG)	10581
Not Aggressive (NAG)	2463
Overtly Aggressive (OAG),	5263
Grand Total	**18307**
Source: Compiled by the author	

The above figure 4 shows the distribution of aggressive tweets for Shinde Government formation in Maharashtra . The above figure depicts that out of 10,471 tweets about the above incident 13% of the tweets are Overtly Aggressive (OAG), 63% of the tweets are Covertly Aggressive (CAG) and 24% of the tweets are Not Aggressive (NAG). The frequency distribution of the tweets are specified in the table 4 below.

Table 4: Frequency distribution of aggressive tweets for Shind
Shinde Government formation
Covertly Aggressive (CAG)
Not Aggressive (NAG)
Overtly Aggressive (OAG),
Grand Total
Source: Compiled by the author

Figure 4: Distribution of aggressive tweetsfor Shinde Government formation

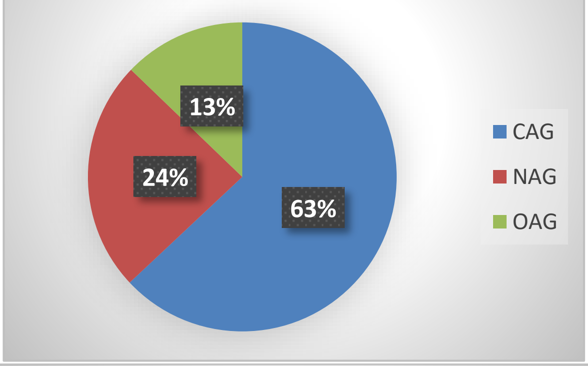
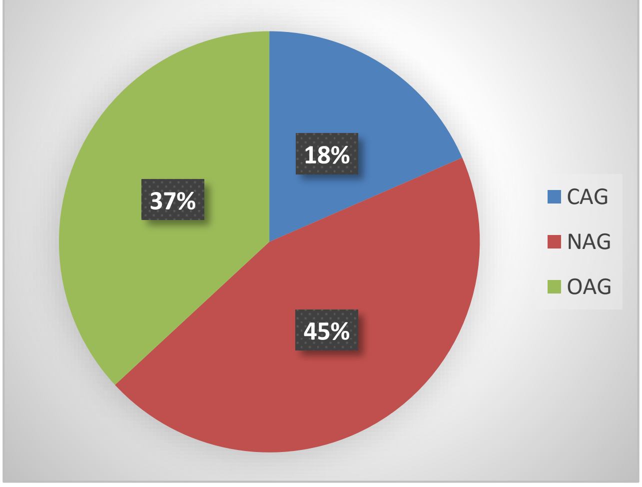


Figure 5: Distribution of aggressive tweets for India lost World Cup 2022



The above figure 5 shows the distribution of aggressive tweets for India lost World Cup 2022. The above figure depicts that out of 3906 tweets about the above incident 37% of the tweets are Overtly Aggressive (OAG), 18% of the tweets are Covertly Aggressive (CAG) and 45% of the tweets are Not Aggressive (NAG). The frequency distribution of the tweets is specified in the table 5 below.

Table 5: Frequency distribution of aggressive tweets for India lost World Cup 2022

India lost T20 World Cup 2022	Count
Covertly Aggressive (CAG)	721
Not Aggressive (NAG)	1744
Overtly Aggressive (OAG),	1441
Grand Total	**3906**

Source: Compiled by the author

The figure 5 shows the distribution of aggressive tweets for Bharat Jodoyatra by Rahul Gandhi. The above figure depicts that out of 69,474 tweets about the above incident 13% of the tweets are Overtly Aggressive (OAG), 64% of the tweets are Covertly Aggressive (CAG) and 23% of the tweets are Not Aggressive (NAG). The frequency distribution of the tweets is specified in the table 6 below.

Figure 6: Distribution of aggressive tweets - Bharat Jodoyatra by Rahul

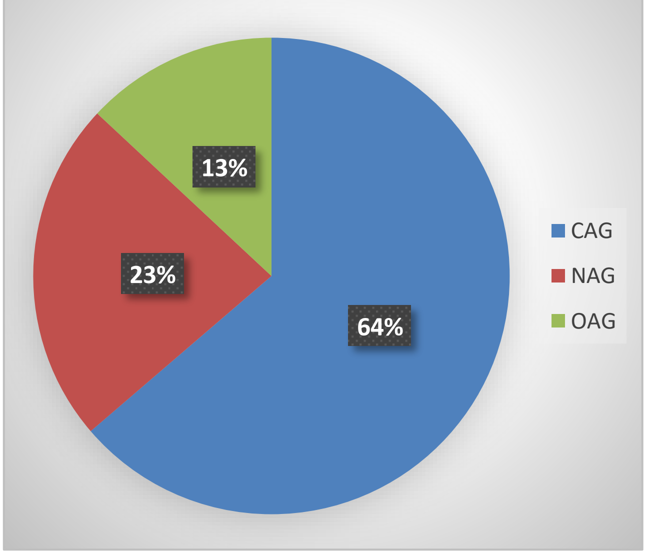


Table 6: Frequency distribution of aggressive tweets for Bharat Jodoyatra by Rahul Gandhi.

Jodo yatra by Rahul Gandhi	Count
Covertly Aggressive (CAG)	44262
Not Aggressive (NAG)	16142
Overtly Aggressive (OAG),	9070
Grand Total	**69474**

Source: Compiled by the author

Table 7: Summary of Category wise distribution of aggressive tweets

Sl No.	Category	Incidents	Timeline of the Tweets		No. of Tweets (%)			
			Start	End	Total no.	OAG (%)	CAG (%)	NAG (%)
1	Social	Shraddha Walker Rape Case	07-03-2022	21-02-2023	18307	5263 (29%)	10581 (58%)	2463 (13%)

2	Finance	Adani Hindenberg Report	16-05-2022	28-03-2023	29664	3352 (11%)	15681 (53%)	10631 (36%)
3	Sports	India lost T20 World Cup 2022	05-11-2022	30-12-2022	3906	1441 (37%)	721 (18%)	1744 (45%)
4	Political (Central)	Bharat Jodo yatra by Rahul Gandhi	15-05-2022	28-03-2023	69474	9070 (13%)	44262 (64%)	16142 (23%)
5	Political (Maharashtra State)	Shinde Government Formation in Maharashtra	01-01-2022	27-03-2023	10471	1341 (13%)	6598 (63%)	2532 (24%)

Source: Compiled by the author

Changing in the aggression levels over time:

Further, the researcher also interested to check the aggression levels over the period after the incident happened. A very common tendency which we can observe here is the declining of aggression levels over the time but with declining rates vary case to case.

In Shradda case, the declining of aggression is at high rate whereas the political case of Maharashtra, the declining rate is smoother.

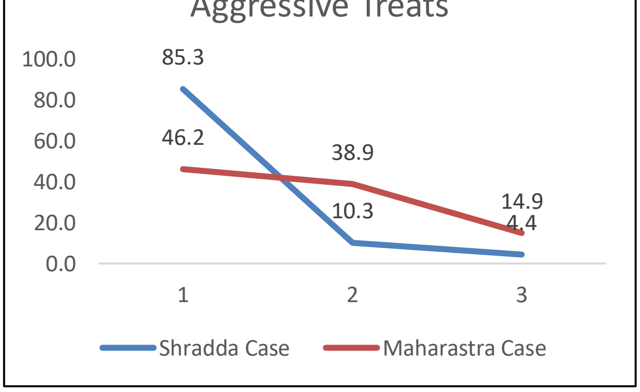


Figure 7 :Percentage of Aggressive Treats

But at the same time, if the activity is longer run activity, we can bell shape movement which is the case of Jodo Yatra

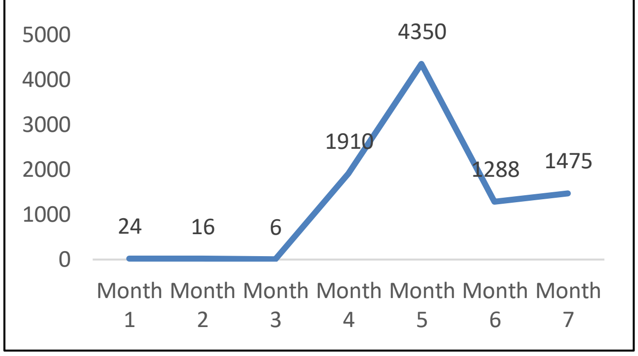


Figure 8: Aggressive Tweets for Jodo Yatra

5 CONCLUSION

The key conclusions from this research study are as follows.

1.Between all language classification models, the current study shows Multinomial Naïve Bayes model performs better in terms of F1 score parameter.

2. We can find a decent level proof that Indians not overly aggressive (OAG) on different types of incidents happening around them. The below table gives overly aggressive percentages in selected five topics happened in the country.

Table 8: Overly aggressive percentages of Different topics

Topic Name	Overly aggressive (OAG) Percentages
Adani Hindenburg report	11%
India losing the T20 World Cup	37%
Rahul Gandhi's Bharat Jodo yatra	13%
Shinde Government in Maharashtra	13%
Shraddha Walker rape case	29%

3.The aggression levels are changing with respect to time. Normally, in the beginning the aggression levels are high and gradually it decreases. The aggression reduction happens at different speeds with different topics. In some events we observed that the aggression levels start with low and reaches to peak after some time..

.

REFERENCES

1. Bharathi, M., Suresh, A., & Iqbal, R. (2022). Detecting disguised aggression in financial discourse. Journal of Computational Linguistics, 48(2), 209–225.
2. Bose, N., Dey, S., & Roy, S. (2021). Machine learning models for detecting hate speech and toxicity on Indian social media. International Journal of Data Science, 6(1), 45–60.
3. Coello, P., Loureiro, M. L., & All, M. (2022). Hot in Twitter : Assessing the emotional impacts of wildfires with sentiment analysis. 200(June). https://doi.org/10.1016/j.ecolecon.2022.107502
4. Chakraborty, A., Paranjape, B., & Joshi, S. (2019). Predicting online abuse using NLP: A survey of recent developments. Social Computing and Human Factors, 34(3), 221–237.
5. Choudhary, R., Gupta, V., & Singh, A. (2023). A comparative study on Indian aggression detection using ML classifiers. Procedia Computer Science, 211, 882–889.
6. Czeranowska, O., Chlasta, K., Mi, P., Grabowska, I., Koco, J., Hwaszcz, K., Wieczorek, J., & Jastrz, A. (2023). Telematics and Informatics Reports Migrants vs . stayers in the pandemic – A sentiment analysis of Twitter content. 10(March). https://doi.org/10.1016/j.teler.2023.100059
7. Hawkins, J. B., Brownstein, J. S., & Hswen, Y. (2021). SSM - Population Health Exploring discussions of health and risk and public sentiment in Massachusetts during COVID-19 pandemic mandate implementation : A Twitter analysis. 15(June). https://doi.org/10.1016/j.ssmph.2021.100851
8. Hema, T., Hidayat, J., Ruldeviyani, Y., Aditama, A. R., Raditia, G., Hema, T., Hidayat, J., Ruldeviyani, Y., Aditama, A. R., Raditia, G., Nugraha, A. W., Adisaputra, M. W., Nugraha, A. W., & Adisaputra, M. W. (2022). ScienceDirect ScienceDirect Sentiment Sentiment analysis analysis of of twitter twitter data data related related to to Rinca Rinca Island Island development using Doc2Vec Doc2Vec and and SVM SVM and and logistic logistic regression regression as as devel. Procedia Computer Science, 197(2021), 660–667. https://doi.org/10.1016/j.procs.2021.12.187
9. Joshi, P., & Mehta, R. (2021). Social outrage and gender violence: Digital narratives and emotional polarity. Indian Journal of Communication Studies, 12(4), 63–78.
10. Khan, A., & Rao, M. (2021). Sports, nationalism, and fan aggression on social media. Journal of Sports Media, 14(1), 55–70.
11. Kumar, S., & Sebastian, T. (2022). Rising intolerance: Mapping aggression on Indian Twitter. South Asian Journal of Digital Media, 9(3), 101–115
12. Khandelwal, A., & Kumar, N. (2020). A unified system for aggression identification in English code-mixed and uni-lingual texts. ACM International Conference Proceeding Series, 55–64. https://doi.org/10.1145/3371158.3371165
13. Lepe-fa, M., Segura-navarrete, A., Vidal-castro, C., Mart, C., & Rubio-manzano, C. (2021). applied sciences Detecting Aggressiveness in Tweets : A Hybrid Model for Detecting Cyberbullying in the Spanish Language.
14. Menon, R., & Das, A. (2023). Temporal patterns of online aggression: A study of Indian social media. Journal of Internet Research, 33(1), 134–150.
15. Mitra, A., & Pathak, R. (2021). The anatomy of a troll: Exploring digital aggression in Indian politics. Journal of Political Psychology, 18(2), 97–113.
16. Nguyen, T. T., Meng, H., Sandeep, S., Mccullough, M., Yu, W., Lau, Y., Huang, D., & Nguyen, Q. C. (2018). Computers in Human Behavior Twitter-derived measures of sentiment towards minorities (2015 – 2016) and associations with low birth weight and preterm birth in the United States. 89(August), 308–315.
17. Or, C. (2018). Aggressive language identification using word embeddings and sentiment features. 113–119.
18. Ouertatani, A., Gasmi, G., Latiri, C., & Ouertatani, A. (2018). ScienceDirect Argued Argued opinion opinion extraction extraction from from festivals festivals and and cultural cultural events events on on Twitter Twitter. 00.
19. Patel, D., & Verma, N. (2023). Analyzing public opinion trends on Twitter: Case studies from India. Media Watch, 14(1), 22–35.
20. Pradhan, N., Gupta, R., & Bose, S. (2022). Context-aware models for aggression classification in Indian code-mixed tweets. ACM Transactions on Asian Language Information Processing, 21(3), 15–32.
21. Passonneau, R. (2011). Sentiment Analysis of Twitter Data. June, 30–38.
22. Philander, K., & Zhong, Y. (2016). Twitter sentiment analysis : Capturing sentiment from integrated resort tweets International Journal of Hospitality Management Twitter sentiment analysis : Capturing sentiment from integrated resort tweets. International Journal of Hospitality Management, 55(March 2019), 16–24. https://doi.org/10.1016/j.ijhm.2016.02.001
23. Ramon, J., Ribeiro-soriano, D., Zegarra, P., & Val, U. De. (2022). Exploring the challenges of remote work on Twitter users ' sentiments : From digital technology development to a post-pandemic era. 142(January), 242–254. https://doi.org/10.1016/j.jbusres.2021.12.052
24. Ridhwan, K. M., & Hargreaves, C. A. (2021). International Journal of Information Management Data Insights Leveraging Twitter data to understand public sentiment for the COVID - 19 outbreak in Singapore. International Journal of Information Management Data Insights, 1(2), 100021. https://doi.org/10.1016/j.jjimei.2021.100021
25. Raghavan, V., & Iyer, S. (2020). Political discourse and aggression: A content analysis of Indian election tweets. New Media and Society India Edition, 5(1), 44–61.
26. Rani, M., & Jain, A. (2020). Text mining approaches for online hate speech detection. Journal of Emerging Technologies, 11(2), 98–110.
27. Sadiq, S., Mehmood, A., Ullah, S., Ahmad, M., & Sang, G. (2021). Aggression detection through deep neural model on Twitter. Future Generation Computer Systems, 114, 120–129. https://doi.org/10.1016/j.future.2020.07.050
28. Sahayak, V. (2015). Sentiment Analysis on Twitter

Data. 2(1), 178–184.
29. Singh, P., Singh, R., & Singh, K. (2018). Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by Indian government. 4, 124–129.
30. Singh, R., & Balaji, B. (2022). Scraping the digital sentiment: A practical framework for social media data analysis. Computer Applications in Social Sciences, 4(1), 11–28.
31. Srivastava, K., & Nair, P. (2021). Emotional triggers and aggression patterns in digital India. Indian Journal of Behavioral Sciences, 19(2), 76–89.
32. Soesanto, A. M., Chandra, C., Soesanto, A. M., & Chandra, C. (2023). ScienceDirect Sentiments comparison on on Twitter Twitter about about LGBT LGBT Sentiments comparison Sentiments comparison Twitter about LGBT. Procedia Computer Science, 216, 765–773. https://doi.org/10.1016/j.procs.2022.12.194
33. Tayef, K., Nazrul, M., Anwar, M., & Sarker, I. H. (2022). Informatics in Medicine Unlocked COVID-19 analytics : Towards the effect of vaccine brands through analyzing public sentiment of tweets. 31(February).
34. Xu, H., Liu, R., Luo, Z., & Xu, M. (2022). Telematics and Informatics Reports COVID-19 vaccine sensing : Sentiment analysis and subject distillation from twitter data. 8(July). https://doi.org/10.1016/j.teler.2022.100016
35. Zulfiker, S., Kabir, N., Amin, A., & Zulfiker, S. (2022). Analyzing the public sentiment on COVID-19 vaccination in social media : Bangladesh context. 15(June).

.