[1]Sandip Bankar
[2]Surekha Janrao
[3]Preeti Gupta
[4]Rohini Patil
[5]Mansi Nandkar
[6]Avantika Jalote
[7]Karim Sohail Khan
[8]Kevin Ninan Mathew

Cyberbullying Detection on Twitter using Sentiment Analysis

Journal of Electrical Systems

Abstract: - Online social networking sites like Twitter are currently faced with a notable and widespread issue known as cyberbullying. The phenomenon of online harassment presents a grave danger to the mental and emotional well-being of individuals, thereby prompting an examination of its profound consequences on those affected. Additionally, there have been frequent changes in tactics used by internet bullies demonstrated the pressing need of effective strategies to minimize this evil. One way out is through application of more enhanced computerized systems for detecting cyber-bullying, which relies on sentiment analysis. The paper aims at providing a safe online environment both for Twitter users and the platform's management. Consequently, this study proposes an extensive comparative evaluation of different feature selection techniques (e.g. Bag of Words, TF-IDF—Term Frequency-Inverse Document Frequency and Count Vectorizer) used in text classification tasks coupled with machine learning algorithms i.e. Naive Bayes, Random Forests and Support Vector Machine. Conclusively, TF-IDF technique combined with SVM model proved to be highly accurate with 92.98% hence making it a very good choice for any future real-time tweet prediction applications. The objective of this comprehensive research is to provide insights into how effective these methods are and guide someone to choose the best one depending on certain contexts.

Keywords: Cyberbullying, Sentiment Analysis, Twitter, Machine Learning, Natural Language Processing, Support Vector Machine, Naïve Bayes, Term Frequency-Inverse Document Frequency, Bag of Words, Random Forest.

I. INTRODUCTION

Today's digital age has a wide spread case of cyberbullying that is very challenging especially on social media platforms such as Twitter where anonymity and ease of communication has led to high cases of cyberbullying. Thus, with the online platform's anonymity and easy access facility have made it possible for people to conduct bullying without fearing for an immediate repercussion. Consequently, these phenomena are becoming more common and include cases of harassment, intimidation, and assault based on such factors as race, sex or sexual orientation. The serious effects of this abuse range from emotional distress and psychological damage to self-destruction or suicide [1]. Therefore, dealing with twitter cyberbullying has become a critical societal concern.

Such things as minor and vulnerable groups are greatly impacted by cyber bullying. Often, targets of online harassment suffer great psychological torments such as low self-esteem, anxiety disorders e.g. depression and feeling of isolation. This effect may also be prolonged therefore affecting their mental health status hence makes it difficult for them to freely interact on the social media. It is because cyber bullying is widespread that key players must assume responsibility including other social network sites like Twitter

However, this problem still exists despite some efforts made by platforms such as Twitter to deal with cyberbullying through reporting mechanisms, community guidelines as well as content moderation. This underscores the necessity of innovative and proactive measures that make use of technology along with data analysis in order to effectively identify and mitigate cases of cyberbullying [2].

[1] *Corresponding author Email: sandipbnkr@gmail.com

[1,3,5,6,7,8],Computer Engineering, STME,SVKM's NMIMS Navi Mumbai, Maharashtra, India

[2] Computer Engineering , K J Somaiya College of Engineering Mumbai, Maharashtra, India

[4] Computer Engineering,Terna Engineering College, Nerul Navi Mumbai, Maharashtra, India

II. Related Work

This section consists of various objectives and their methodology implemented over the past few years. In order to examine the contributions made by earlier studies to the automatic identification of cyberbullying, cyberstalking, and other forms of cyber harassment, some relevant research publications were selected for the literature review.

Using machine learning approaches, Ghasem et al. [3] proposed a model for automatically identifying and regulating cyberstalking and cyberbullying. This strategy was mostly centered on documenting evidence to fight hackers and automating email-based cyberstalking detection.

Shekhar et al. [4] aimed to enhance text processing accuracy in cyberbullying detection by generating phonetic codes for misspelled and censored words using the Soundex Algorithm. Mathur et al. [5] explored the effectiveness of different machine learning algorithms in analyzing tweet content for cyberbullying indicators using Selenium for web scraping. Atoum [6] employed Naïve Bayes and Support Vector Machine models with different N-gram language models for sentiment analysis to detect cyberbullying. Dedeepya et al. [7] utilized TF-IDF, Count Vectorization, Bag of Words, SVM, and NLP techniques for cyberbullying detection. Islam et al. [8] evaluated the performance of various machine learning algorithms in classifying cyberbullying instances based on Bag of Words and TF-IDF features. Prabowo et al. [9] found SVM most effective with TF-IDF for cyberbullying classification.

Mody et al. [10] integrated rule-based sentiment analysis and machine learning methods like Naive Bayes and Linear SVC for cyberbullying detection. Prathap et al. [11] aimed to determine changes in public opinion about crime events and identify emotions associated with various types of cyberbullying. Yadav et al. [12] proposed a deep learning model utilizing a pre-trained BERT model with a single linear neural network layer for cyberbullying detection.

While scholars at [5-8] constructed the detection model using machine learning approaches, authors at [3,4] employed a lexicon-based approach in the literature. Hybrid approaches, such as lexicon-based and machine learning techniques for automated detection, have also been proposed by authors at [9–16]. Most studies used machine learning methods to automatically categorize tweets. Real-time automatic detection of cyberstalking on Twitter and other social media platforms remains a difficult challenge. Real-time, automated techniques for detecting cyberstalking that function admirably are still lacking.

III. Proposed Methodology

After thoroughly examining various methods, we developed a framework that forms the basis of our approach that we applied in our research and implementation as shown in Figure 1.

A. Proposed Methodology

The process of implementation includes:

- Data Collection: Data is collected from various sources which include Kaggle and IEEE papers.
- Pre-processing: The tweets are then pre-processed to clean the text data. This involved removing emojis, links, mentions, non-ASCII characters, contractions, hashtags and extra spaces. We also filtered special characters and carried out stemming.
- Balancing the dataset: Synthetic Minority Over-sampling Technique (SMOTE) is applied to generate synthetic samples for the minority class by interpolating between existing minority class samples, thereby balancing the class distribution.
- Feature Selection: Next, features are selected from the text using BoW with n-grams, TF-IDF and Count Vectorizer.
- Machine Learning Model Training: Different machine learning models are trained on the labelled data. These models you include: Random Forest, SVM, Naive Bayes
- Performance Scores: After training the models, their performance is evaluated using performance scores like precision, recall, and F1-score to measure the effectiveness of a machine learning model.
- Classification: Once a model is determined to have the best performance, it is used to classify new tweets as cyberbullying or not cyberbullying.
- Deployment: Finally, the system is deployed to be used in a real-world setting.

B. Data Collection

In data collection for cyberbullying research, there is a significant diversity in the data sources used in different research papers. Some studies draw on publicly available data from platforms such as Kaggle, while others prefer real-time data collection. In addition, many researchers are taking the painstaking route of manually preparing their data, while some use a combination of these approaches. These datasets, rich in textual content, often contain emojis, peculiar characters, and unconventional text, which require careful preprocessing to ensure their suitability for analysis.

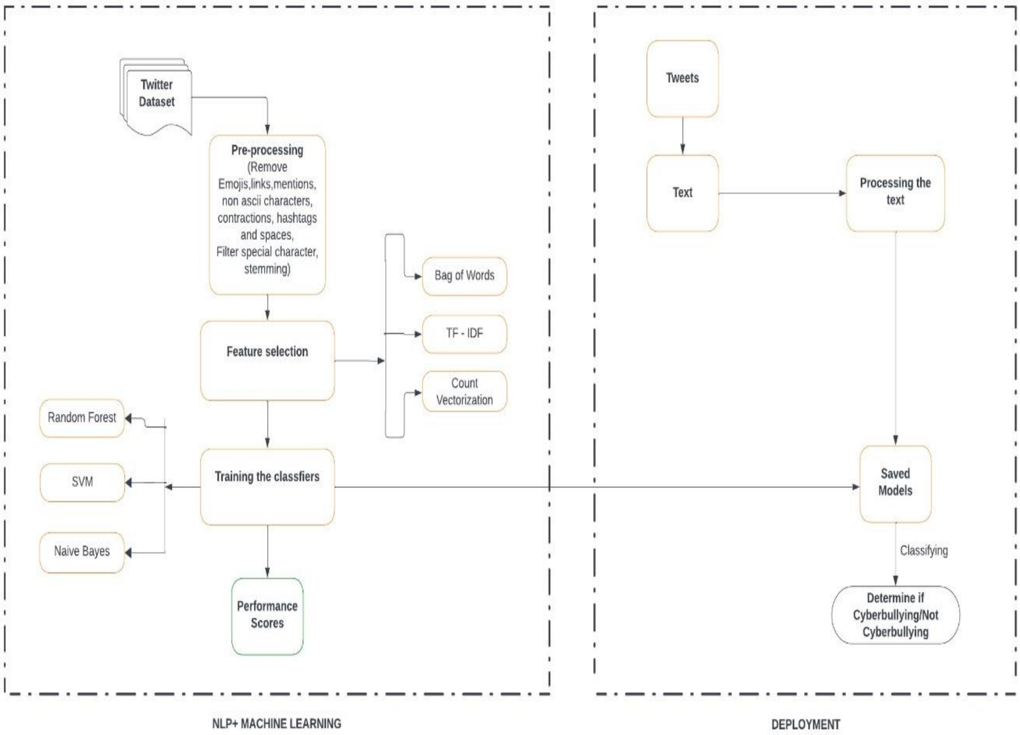

Figure 1: Proposed Framework

The data was taken from the paper published by Jason Wang, Kaiqun Fu, Chang-Tien Lu [10] in 2020. This dataset is a compilation of six datasets. After balancing the data, the dataset had 48000 tweets with tags as: Cyberbullying, not cyberbullying, age, ethnicity, gender, religion and other.

The classification was changed to binary classification of Cyberbullying and not cyberbullying. 5500 rows from another Kaggle dataset were further appended to add more instances of not cyberbullying class [11].

C. Data Preprocessing

The steps for preprocessing are as follows:

- Remove Emojis: The first step involved removing emojis from the data set.
- Convert text to lowercase: This step converts all uppercase letters in the text data to lowercase.
- Remove Hashtags: Hashtags are irrelevant here and are removed.
- Remove links, mentions, non-ASCII characters: Removed links embedded within the text data, as well as mentions (usernames preceded by an “@” symbol) and non-ASCII characters.
- Filter special characters: Removed special characters from the data set like punctuation marks, mathematical symbols, and other non-alphanumeric characters.
- Remove contractions: Contractions, like “don’t” or “can’t” are expanded into their full forms.
- Remove spaces: Extra spaces from the text data are removed.
- Stemming: Words are reduced to their base form using stemming.

D. Balancing the Data

There is a significant difference between the tweets available for the two classes: cyberbullying and not cyberbullying. SMOTE was used to balance the classes.

E. Feature Extraction

After SMOTE, we moved on to Feature extraction where we used three techniques: Bag of Words (n-grams), TF-IDF, and Count Vectorizer. When combined with diverse machine learning algorithms and data preprocessing methods, these yielded varying accuracies in sentiment analysis and text classification tasks.

- The Bag of Words (BoW): Bag of Words (BoW) is a text representation technique where a document is represented as a bag (multiset) of its words, disregarding grammar, and word order. N-grams are contiguous sequences of N items (words in this context) from a given text. So, BoW with N-grams considers not only single words but also sequences of words up to length N in a document.

In this project, BoW with N-grams is utilized as a feature extraction technique for text data. By converting text into a matrix of token counts (occurrences of words or N-grams), it creates a numerical representation of text data suitable for machine learning algorithms. This allows classification models to process and analyze text data effectively by capturing the presence and frequency of words and N-grams in each document.

- Term Frequency-Inverse Document Frequency (TF-IDF):TF-IDF is a quantitative measure that indicates the significance of a word in a document compared to a set of texts. It is frequently employed in tasks related to natural language processing and information retrieval.

TF-IDF transforms the raw text into numerical data by assigning weights to words based on their importance in the document and the overall collection of texts. This helps identify unique characteristics in each document, which is a valuable part of sentiment analysis. Under cyberbullying detection, the occurrence of specific phrases can serve as an indication of the sentiment conveyed in the text. TF-IDF is subsequently utilized as input for machine learning models in order to detect instances of cyberbullying on the Twitter platform.

- Count Vectorizer: The proposed system uses Count Vectorizer which is a technique employed in natural language processing to extract features. This function transforms a set of textual documents into a matrix that represents the frequency of each token in the texts. Every row in the matrix corresponds to a document, whereas each column corresponds to a distinct word in the corpus. Each cell contains the frequency of the associated term in the document. Here, it is used to transform the text data into numerical characteristics. Every tweet is converted into a vector of word frequencies, which represents the number of times each word appears in the tweet. Subsequently, these vectors representing the frequency of words are employed as input characteristics for machine learning techniques in order to train models to detect cyberbullying.

F. Machine Learning Models

The models selected for implementation, in conjunction with the three feature extraction methodologies, are as follows:

- Random Forest: Random Forest algorithm constructs several decision trees throughout the training process. Every tree in the forest functions autonomously and generates a forecast, and the ultimate prediction is chosen by a majority vote or average. The proposed system uses Random Forest algorithm as a machine learning model to detect instances of cyberbullying. It is employed to categories textual information (tweets) into several cyberbullying categories depending on their content. Through the utilization of pre-labelled data, Random Forest algorithm acquires knowledge of patterns within the textual characteristics and thereafter possesses the capability to accurately anticipate the sentiment (whether a tweet is classified as Cyberbullying or Not Cyberbullying) of previously viewed tweets.

- Support Vector Machine (SVM): Support Vector Machine (SVM) is a type of supervised machine learning technique that is utilized for both classification and regression applications. The goal is to identify the most effective hyperplane that can divide various classes in the feature space, while also maximizing the distance between the classes. The proposed work has used Support Vector Machines (SVM) as a classifier to differentiate various forms of cyberbullying by analyzing the content of tweets. The system acquires the ability to categories tweets into predetermined classifications, such as Cyberbullying or non-cyberbullying, by examining the characteristics derived from the textual information. The objective of SVM is to identify the optimal decision boundary that effectively distinguishes tweets as either cyberbullying or non-cyberbullying, hence facilitating accurate detection and classification of cyberbullying incidents on Twitter.

- Naïve Bayes: Naive Bayes is a rudimentary probabilistic classifier that relies on Bayes' theorem and assumes that the characteristics are independent of each other. Although it is simple, it frequently achieves high performance in text categorization problems. In this project, Naive Bayes is used as one of the machine learning models for sentiment analysis of Twitter data. It leverages the frequency of words in the cleaned text data to classify tweets whether they are cyberbullying or not. By assuming independence between the occurrence of words in tweets, Naive Bayes efficiently learns and predicts the sentiment of tweets based on their textual content

IV. Deployment and Result analysis

A. *Deployment*

After a comparative analysis to find which model combination yields the best result, the relevant feature extraction vectorizer and Machine Learning model is downloaded as joblib files. The deployment diagram is as follows:

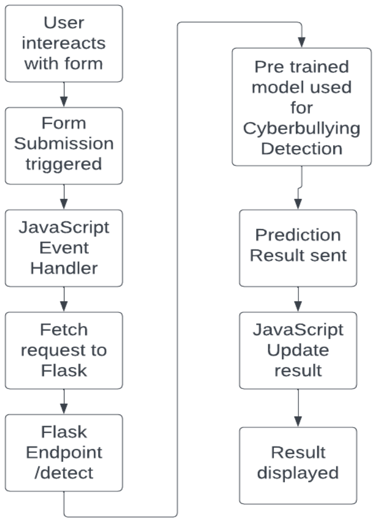

Figure 2: Deployment Process

Steps:

1. The procedure starts when a user provides an input by filling out the form.

2. The procedure of form submission is initiated when the user submits the form.

3. User events are handled using a JavaScript event handler.

4. The JavaScript function subsequently transmits the text inputted in the form to obtain a forecast. This process is submitting the material to a pre-existing model that has been specifically trained to detect instances of cyberbullying in text.

5. Initiate a fetch request to a Flask endpoint, typically hosted on a server. The web application is specifically built to do predictions on the server-side.

6. The request is sent to a designated Flask endpoint called /detect whose agenda is to accept text input and utilise a pre-trained model to detect instances of cyberbullying.

7. The JavaScript code receives the prediction result, indicating if the text is cyberbullying, as a response from the Flask endpoint. It modifies the outcome correspondingly.

8. Result is displayed by the code that exhibits the outcome on the user's webpage.

B. *Deployment*

After a comparative In all 9 different combinations are formed and in-depth analysis is done with these all 9 different combinations.

- Count vectorizer with random forest(CV and RF) was one of the better performing models with 90.14% accuracy score, 90.26% precision score, 90.14% recall score, 90.13% F1 score.In the Classification report for CV with RF has the following statistics: There are 6749 true negatives, 6448 true positives, 521 false positives, and 922 false negatives. A steeper ROC curve indicates better performance, and the Area Under the Curve (AUC) value, denoted as 0.97 here, quantifies the model's overall performance; a higher AUC suggests better discrimination between the classes. The high AUC value of 0.97 indicates that the classifier has excellent predictive performance.

- Count Vectorizer and SVM : Count vectorizer with SVM was the third best performing models with 91.62% accuracy score, 91.69% precision score, 91.62% recall score, 91.62% F1 score. The confusion matrix has got the following statistics: 6802 true negatives, 6611 true positives, 468 false positives, and 759 false negatives. The Area Under the Curve (AUC) value, denoted as 0.98, quantifies the model's overall performance with a higher AUC suggests better discrimination between the classes.

- Count Vectorizer and Naïve Bayes : Count vectorizer with Naïve Bayes was one of the poor performing models with 83.10% accuracy score, 83.67% precision score, 83.10% recall score, 83.02% F1 score. The confusion matrix has got the these statistics: 5522 true negatives, 6614 true positives, 1718 false positives, and 756 false negatives. The ROC curve for the Naive Bayes classifier with Count Vectorizer representation demonstrates strong performance, with an area under the curve (AUC) of 0.92.

- TF-IDF and Random Forest : TF-IDF with random forest was one of the better performing models with 92.31% accuracy score, 92.69% precision score, 92.30% recall score, 92.29% F1 score. The confusion matrix has got the following statistics: 7051 true negatives, 6463 true positives, 219 false positives, and 907 false negatives. The ROC curve for the Random Forest classifier with TF-IDF representation exhibits excellent performance, with an area under the curve (AUC) of 0.98.

- TF-IDF and SVM: TF-IDF with SVM was the best performing models with 92.98% accuracy score, 92.99% precision score, 92.98% recall score, 92.98% F1 score.The confusion matrix has the following statistics: there are 6801 true negatives, 6812 true positives, 469 false positives, and 558 false negatives. The ROC curve for the SVM classifier with TF-IDF representation illustrates excellent performance, with an AUC of 0.98. This indicates a high ability of the classifier to distinguish between true positive and false positive rates, reflecting strong predictive capabilities across various thresholds.

- TF-IDF and Naïve Bayes : TF-IDF with Naïve bayes had 84.85% accuracy score, 84.27% precision score, 84.85% recall score, 84.78% F1 score. The confusion matrix has the following statistics: there are 5757 true negatives, 6665 true positives, 1513 false positives, and 705 false negatives. The ROC curve for the Naive Bayes classifier with TF-IDF representation indicates strong performance for both Class 0 and Class 1, with an area under the curve (AUC) of 0.94 for each class. This suggests that the classifier has a high ability to distinguish between the positive and negative classes.

- BoW using n-grams and Random Forest : BoW using n-grams and random forest had an 89.25% accuracy score, 89.69% precision score, 89.25% recall score, 89.22% F1 score. The confusion matrix has the following statistics: there are 6868 true negatives, 6198 true positives, 402 false positives, and 1172 false negatives. With an AUC of 0.97, this classifier demonstrates excellent predictive performance.

- BoW using n-grams and SVM : BoW using n-grams and SVM had a 90.93% accuracy score, 90.94% precision score, 90.93% recall score, 90.93% F1 score. The confusion matrix has the following statistics: there are 6626 true negatives, 6687 true positives, 644 false positives, and 683 false negatives. With an AUC of 0.98, BoW using n-grams and SVM exhibits excellent predictive performance.

- BoW using n-grams and Naïve Bayes : This combination was one of the worst performing models with 76.52% accuracy score, 79.22% precision score, 76.52% recall score, 75.93% F1 score. The confusion matrix has the following statistics: there are 4431 true negatives, 6772 true positives, 2839 false positives, and 598 false negatives. With an AUC of 0.91, this classifier demonstrates good predictive performance.

After analyzing various combinations, their performance measures are combined and presented in table 4.1.

Table 4.1 Comparison table

Feature Selection	Count Vectorizer			TF-IDF			Bag of Words (n-gram)		
Metrics/ML	SVM	RF	NB	SVM	RF	NB	SVM	RF	NB
Accuracy	91.62	90.14	83.10	92.98	92.31	84.85	90.94	89.25	76.52
F1 Score	91.62	90.13	83.02	92.99	92.29	84.80	90.94	89.22	75.93
Precision	91.69	90.26	83.67	92.99	92.69	85.27	90.94	89.69	79.22
Recall	91.62	90.14	83.10	92.98	92.31	84.85	90.94	89.25	76.52

The bar graph charts provide the comparative study analysis of each performance measure :Accuracy Precision, Recall and F1 Score(Figure 3,4,5,and 6 respectively) to evaluate the best model.

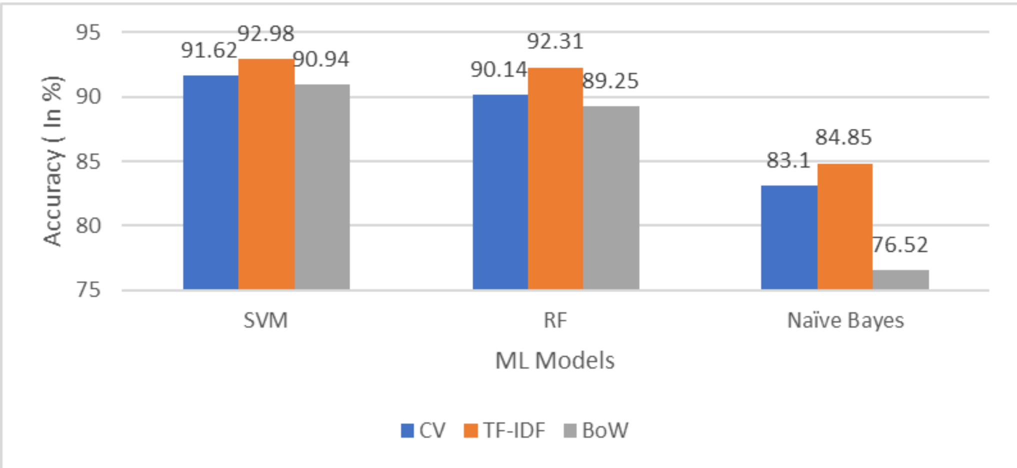

Figure 3: Accuracy Scores for all combinations

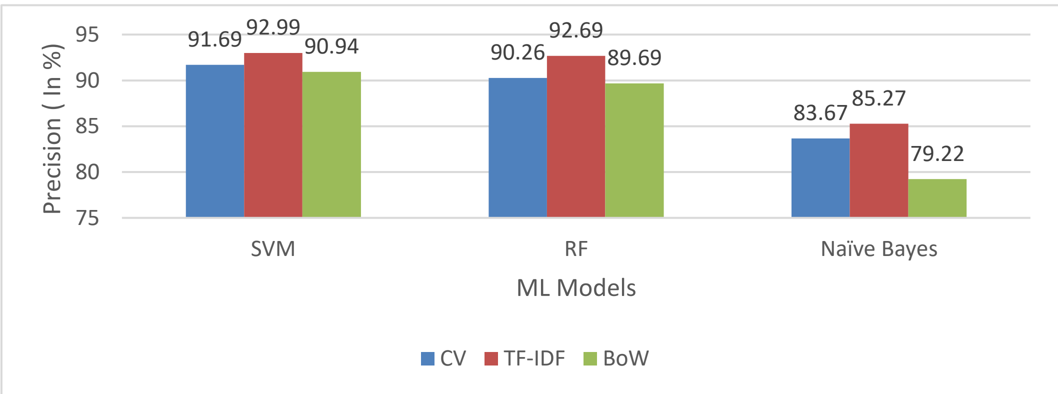

Figure 4: Precision Scores for all combinations

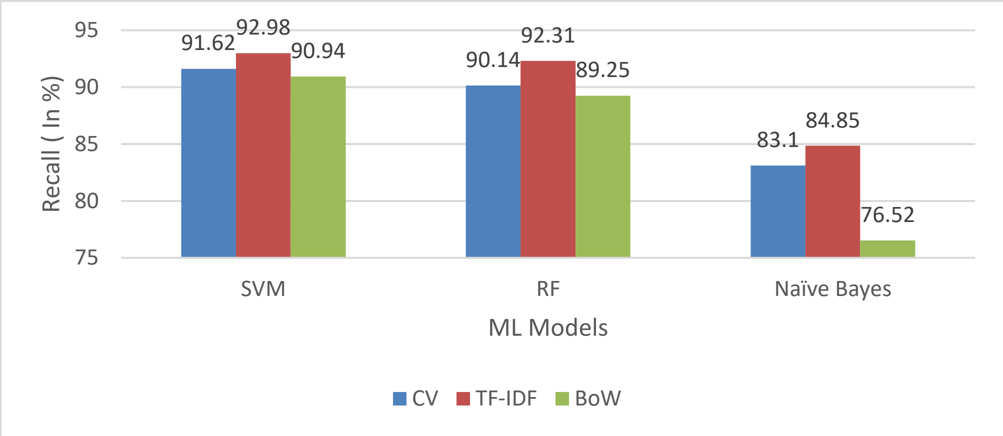

Figure 5: Recall Scores for all combinations

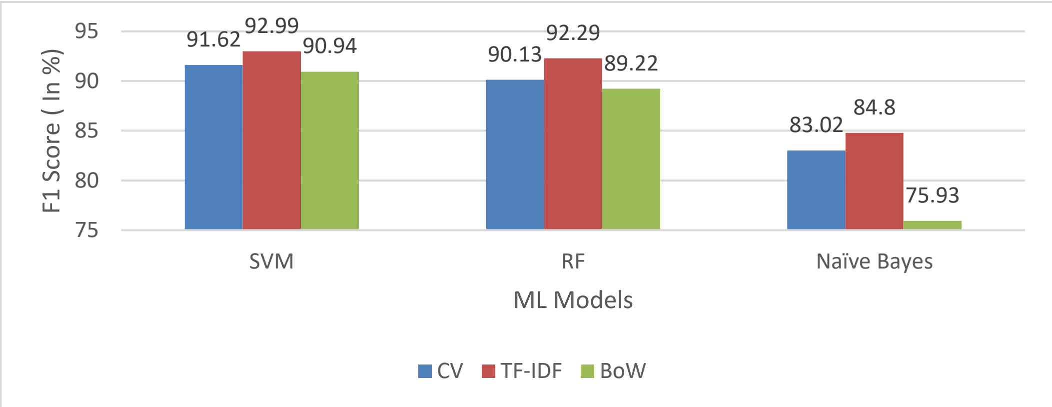

Figure 6: F1 Scores for all combinations

Following a comprehensive evaluation of multiple models for cyberbullying detection, SVM model utilizing TF-IDF for feature extraction emerged as the most effective choice with a 92.98% accuracy. Cross-validation was conducted to verify the accuracy of the model, with both accuracy and cross-validation scores showing close alignment. Consequently, the accuracy of the SVM model was duly validated. This model was then deployed for practical use in detecting cyberbullying. It accepts input in the form of strings and employs its trained knowledge to classify whether the input statement exhibits characteristics of cyberbullying or not.

V. Conclusion

This work represents a significant advancement in ongoing efforts to address cyberbullying on Twitter. A solution was discovered that effectively detects cyberbullying with a high level of accuracy, specifically 92.98%. This was achieved by employing sentiment analysis and machine learning techniques, using Support Vector Machine (SVM) with Term Frequency-Inverse Document Frequency (TF-IDF). Furthermore, the initiative beyond mere identification by offering a pragmatic and accessible resolution that can be executed on a community level, enabling immediate identification of cyberbullying content. This empowers both platform administrators and users to immediately and proactively address cases of cyberbullying, ultimately fostering a safer and more inclusive digital environment on Twitter. In essence, the goal is to create an environment of understanding, empathy, and mutual respect, so that Twitter remains a platform where people may freely express themselves without fear of being mistreated or harassed. By continuously exploring, cooperating, and advocating, we are dedicated to achieving the aim of avoiding cyberbullying and establishing a safer online environment for all users.

References

[1] M. Baer. "Cyberstalking and the Internet Landscape We Have Constructed," Virginia Journal of Law & Technology, 154(15), 2020, pp. 153-227.

[2] Z. Ghasem, I. Frommholz, and C. Maple, "Machine learning solutions for controlling cyberbullying and cyberstalking," International Journal of Information Security, 6(2), 2023, pp. 55-64.

[3] B. R. Prathap and K. Ramesha, "Twitter Sentiment for Analyzing Different Types of Crimes," 2018 International Conference on Communication, Computing, and Internet of Things (IC3IoT), Chennai, India, 2018, pp. 483-488, doi: 10.1109/IC3IoT.2018.8668140.

[4] S. A. Mathur, S. Isarka, B. Dharmasivam, and J. C. D., "Analysis of Tweets for Cyberbullying Detection," 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2023, pp. 269-274, doi: 10.1109/ICSCCC58608.2023.10176416.

[5] J. O. Atoum, "Cyberbullying Detection Through Sentiment Analysis," 2020 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2020, pp. 292-297, doi: 10.1109/CSCI51800.2020.00056

[6] P. Dedeepya, P. Sowmya, T. D. Saketh, P. Sruthi, P. Abhijit and S. P. Praveen, "Detecting Cyber Bullying on Twitter using Support Vector Machine," 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2023, pp. 817-822, doi: 10.1109/ICAIS56108.2023.10073658.

[7] M. M. Islam, M. A. Uddin, L. Islam, A. Akter, S. Sharmin and U. K. Acharjee, "Cyberbullying Detection on Social Networks Using Machine Learning Approaches," 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Gold Coast, Australia, 2020, pp. 1-6, doi: 10.1109/CSDE50874.2020.9411601.

[8] W. A. Prabowo and F. Azizah, “Sentiment Analysis for Detecting Cyberbullying Using Term Frequency Inverse Document Frequency and Support Vector Machine”, J. RESTI (Rekayasa Sist. Teknol.Inf.), vol. 4, no. 6, pp. 1142 -, Dec. 2020.

[9] A. Mody, S. Shah, R. Pimple, and N. Shekokar, "Identification of Potential Cyber Bullying Tweets using Hybrid Approach in Sentiment Analysis," 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Msyuru, India, 2018, pp. 878-881, doi: 10.1109/ICEECCOT43722.2018.9001476.

[10] A. Shekhar and M. Venkatesan, "A Bag-of-Phonetic-Codes Model for Cyber-Bullying Detection in Twitter," 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, India, 2018, pp. 1-7, doi: 10.1109/ICCTCT.2018.8550938.

[11] J. Yadav, D. Kumar, and D. Chauhan, "Cyberbullying Detection using Pre-Trained BERT Model," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 1096-1100, doi: 10.1109/ICESC48915.2020.9155700.

[12] J. Wang, K. Fu, C.T. Lu, “SOSNet: A Graph Convolutional Network Approach to Fine-Grained Cyberbullying Detection,” Proceedings of the 2020 IEEE International Conference on Big Data (IEEE BigData 2020), December 10-13, 2020

[13] N. Yuvaraj, V. Chang, B. Gobinathan, A. Pinagapani, S. Kannan, G. Dhiman, & A. R. Rajan, "Automatic detection of cyberbullying using multi-feature based artificial intelligence with deep decision tree classification," Computers & Electrical Engineering, 92, 2021.

[14] S. Sadiq, A. Mehmood, S. Ullah, M. Ahmad, G.S. Choi, "Aggression detection through deep neural model on twitter," Future Generation Computer Systems, 114, 2021, pp. 120-129.

[15] Sangwan, Saurabh Raj, and M. P. S. Bhatia, "D-BullyRumbler: a safety rumble strip to resolve online denigration bullying using a hybrid filter-wrapper approach," Multimedia Systems, 2020, pp. 1-17.

[16] S. Bailkar, K. Shenoy, A. Bedekar, S. Bankar and P. More, "Smart Inventory Optimization using Machine Learning Algorithms," 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India, 2024, pp. 1395-1400, doi: 10.1109/IDCIoT59759.2024.10467512.