Proceedings of the 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT-2024)
IEEE Xplore Part Number: CFP24CV1-ART; ISBN: 979-8-3503-2753-3

2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT) | 979-8-3503-2753-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/IDCIOT59759.2024.10467658

A Review on Cyberbullying Detection on Twitter Dataset

Avantika Jalote
School of Technology Management & Engineering, SVKM's NMIMS University
Navi Mumbai, Maharashtra, India – 410210
Email: jalote.avantika@gmail.com

Karim Sohail Khan Patan
School of Technology Management & Engineering, SVKM's NMIMS University
Navi Mumbai, Maharashtra, India – 410210
Email: karimsohailkhan@gmail.com

Kevin Ninan Mathew
School of Technology Management & Engineering, SVKM's NMIMS University
Navi Mumbai, Maharashtra, India – 410210
Email: mathew.kevin0502@gmail.com

Mansi Nandkar
School of Technology Management & Engineering, SVKM's NMIMS University
Navi Mumbai, Maharashtra, India – 410210
Email: mansinandkar@gmail.com

Prof. Pranav More
School of Technology Management & Engineering, SVKM's NMIMS University
Navi Mumbai, Maharashtra, India – 410210
Email: pranavmore2530@gmail.com

***Abstract*— Cyberbullying on Twitter has emerged as a pervasive and distressing issue in the digital age, posing significant threats to the well-being of users. This review study delves into the critical issue of cyberbullying on Twitter, highlighting its devastating consequences on victims, from emotional distress to offline repercussions. While Twitter has taken commendable steps to combat cyberbullying, the ever-evolving tactics of cyberbullying pose ongoing challenges.**

The objective of this review study is to explore the potential for mitigating cyberbullying through the implementation of a robust cyberbullying detection system. By utilizing sentiment analysis, the study proposes the development of a tool that can proactively identify and address instances of cyberbullying. This tool aims to empower Twitter users and the platform itself to collaboratively create a safer and more respectful online environment.

This study underscores the urgent need to address cyberbullying on Twitter and emphasizes the importance of technology-driven solutions to combat this issue. By fostering a culture of digital inclusivity and kindness, to the study contributes to the broader conversation about online safety and inspire innovative approaches for a more secure and respectful online environment.

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

Cyberbullying on Twitter is a universal issue in the digital/social age, with users targeting individuals based on race, gender, appearance, political inclination, or beliefs. The outcome of this harassment is harsh, causing emotional distress, mental health issues and even leading to physical harm. Even though Twitter has made efforts to fight against cyberbullying through reporting mechanisms and anti-abuse policies, the battle continues as technology advances and methods for cyberbullying change, making it harder to maintain a safe and respectful online environment for all users.

This review **study** delves into the complex issue of cyberbullying on Twitter, exploring its evolution from previous research **study**. It also explores the platform's reporting mechanisms and assesses their effectiveness in addressing cyberbullying by the usage of different techniques in implementing sentiment analysis for its prediction. It emphasizes the need for ongoing research, policy development, and community involvement to create a safer and more inclusive digital environment. Addressing cyberbullying effectively is crucial to protecting the mental and emotional well-being of all users and fostering a more respectful online community.

Sentiment analysis, also known as opinion mining, is a natural language processing (Natural Language Processing) technique used to determine and evaluate the sentiment, emotions, and subjective information expressed in text data. This process involves classifying text as positive, negative, or neutral, providing insights into the opinions and emotions of individuals towards a specific topic, product, or event. Sentiment analysis is a valuable tool with various applications, including detecting cyberbullying on social media platforms like Twitter which is implemented using various machine learning and natural language processing models using Python libraries. It can be used in the form of applications, bots, and more.

It can be invaluable in the detection of cyberbullying on Twitter by automatically identifying and categorizing harmful content based on sentiment. By combining sentiment analysis with cyberbullying classification, it becomes possible to discern the emotional intensity of abusive messages, enabling platforms to take prompt action in moderating or blocking harmful accounts. This approach helps protect users from cyberbullying, creates a safer online environment, and offers researchers valuable insights into the nature and prevalence of online harassment. While sentiment analysis is a powerful tool, it is important to continually refine and adapt the model to keep up with evolving online behaviours and abusive language to effectively combat cyberbullying.

II. RELATED WORK

After a brief introduction to various approaches to Cyberbullying detection using Sentiment Analysis. This section consists of various objectives and their methodology implemented by various users over the past few years.

TABLE I. OBJECTIVES AND RESEARCH ANALYSIS ON CYBERBULLYING

S.No.	**Research Analysis of Papers on Cyberbullying Detection**			
	Description/ Objectives	***Author Name***	***Year***	***Research Analysis***
1	A Bag-of-Phonetic-Codes model, using the pronunciation of words as features to rectify misspelled and censored words. [1]	A. Shekhar and M. Venkatesan	2018	The model is inspired by the Bag-of-Words model for extracting textual features. Phonetic code generation is done using the Soundex Algorithm.
2	Analysis of Tweets with the help of Selenium for web scrapping [2]	S. A. Mathur, S. Isarka, B. Dharmasivam and J. C. D.	2023	Different algorithms like Random Forest classifier, gradient boosting classifier, Ada boost and more were tried.
3	Cyberbullying Detection Through Sentiment Analysis by Naïve Bayes and Support Vector Machine Models [3]	Jalal Omer Atoum	2020	Naïve Bayes and Support Vector Machine Measures for Different N-gram Language Models were tried.
4	Detecting Cyber Bullying using Support Vector Machine and Natural Language Processing Techniques [4]	P. Dedeepya, P. Sowmya, T. D. Saketh, P. Sruthi, P. Abhijit and S. P. Praveen	2023	Term Frequency Inverse Document Frequency, Count Vectorization, Bag of Words, Support Vector Machine, and Natural Language Processing were used
5	Detecting Cyberbullying using Machine Learning Approaches. [5]	M. M. Islam, M. A. Uddin, L. Islam, A. Akter, S. Sharmin and U. K. Acharjee	2020	Two notable features were used for the analysis: Bag of Words and Term Frequency Inverse Document Frequency. Various machine learning algorithms were applied.
6	Detecting Cyberbullying using Machine Learning Approaches using Term Frequency Inverse Document Frequency and Support Vector Machine [6]	W. A. Prabowo and F. Azizah	2020	Support Vector Machine was found most applicable to classify cyberbullying in applications programs.
7	Cyberbullying detection using a Hybrid Approach in Sentiment Analysis using POS tagging and Support Vector Machine [7]	A. Mody, S. Shah, R. Pimple, and N. Shekokar	2018	The textual data in each tweet is processed using a two-pronged strategy: a knowledge-based approach using SentiWordNet and a machine learning approach using Naive Bayes and Linear Support Vector Classifier (SVC)
8	Analysis of different types of cyberbullying [8]	B.R. Prathap and K. Ramesha	2018	The approach aims to determine the change in public opinion about crime events and identify emotions associated with diverse types of crimes.
9	Cyberbullying Detection by Pre-Trained BERT Model - A deep learning model [9]	J. Yadav, D. Kumar and D. Chauhan J. Yadav, D. Kumar, and D. Chauhan	2020	It utilizes a pre-trained BERT model with a single linear neural network layer as a classifier.

After a detailed literature survey of the above research papers that we have analyzed. Here is a detailed discussion of these methodologies with some of the most common observations given below:

A. 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 [1] 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. In a notable instance, the research paper dealt with the arduous task of manually marking comments as cyberbullying or not, a labor-intensive job.

Web scraping is a powerful and versatile approach that provides access to a wide range of information from websites. This method proves [2] invaluable, particularly when real-time or current data are required. Web scraping allows researchers to extract a variety of data types, including textual content, images, pricing information, user reviews, and more. One of its outstanding advantages is automation, which allows the collection of data from multiple websites without the need for constant manual intervention, thus saving valuable time and effort. It also provides us the opportunity to define the parameters like target elements, URL or website, and more for finer extraction of data.

B. Data Pre-Processing

The diversity of data preprocessing approaches used by various models underscores the absence of standardized procedures. This variability in data preparation arises from the inherent subjectivity in determining what constitutes relevant information for a given task. Emojis, for instance, serve as valuable emotional indicators for some models, leading to their retention, while others opt for elimination due to their potential to introduce noise. Nonetheless, a consensus appears to emerge around several fundamental data preprocessing steps of Natural language Processing (NLP), such as the removal of stop words, links, and the standardization of letter cases, alongside techniques like stemming and lemmatization.

The importance of data cleaning in the context of sentiment analysis cannot be overstated, as it profoundly influences the subsequent predictive capabilities of these models. This divergence in data preprocessing requirements mirrors the dynamic nature of sentiment analysis, wherein tailored strategies become imperative to accommodate the idiosyncrasies of various applications and datasets. Therefore, the multifaceted nature of data preprocessing continues to pose intriguing challenges and opportunities in the ever-evolving field of sentiment analysis, underscoring the need for a flexible and nuanced approach to text data preparation. In this regard, researchers and practitioners must exercise discernment in selecting the most appropriate preprocessing techniques that align with the specific objectives and intricacies of their sentiment analysis tasks.

An alternative approach, such as the Soundex algorithm and the Bag of Phonetics [1], has been explored. The Soundex algorithm encodes words based on their phonetic similarity, aiding in tasks like name matching. Its advantage lies in facilitating the identification of phonetically similar terms, but it may struggle with words from languages outside its design scope and exhibit imprecision in some cases. The Bag of Phonetics represents words as phonetic tokens, allowing for phonetic-based similarity analysis. The accuracy of the Soundex algorithm [1] is good for its intended purpose, which is to find phonetic matches, especially in the context of names and similar-sounding terms. However, its accuracy can be compromised in cases involving homophones, misspellings, complex phonetic variations, or languages other than English. These alternative approaches exemplify the pursuit of innovative methods to address specific challenges in text analysis, as shown by the research documented in referred papers

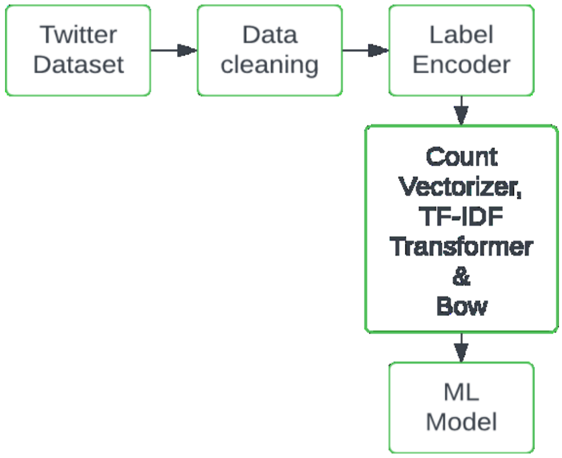

Fig. 1. Preprocessing

C. Feature Extraction Models

In feature extraction, three of the most common Natural language Processing(NLP) techniques are the Bag of Words, Term Frequency Inverse Document Frequency, and Count Vectorizer. When combined with diverse machine learning algorithms and data preprocessing methods, these techniques yield varying accuracies in sentiment analysis and text classification tasks.

1. The Bag of Words (BoW)

This method represents a document by creating a vocabulary of its constituent words, disregarding their order, and measuring their frequency. One of its major advantages is simplicity and efficiency in capturing the document's term frequency information. However, a notable drawback is that it completely ignores word semantics and context, which can limit its accuracy in capturing the subtleties of language.

2. Term Frequency-Inverse Document Frequency (Term Frequency Inverse Document Frequency)

It stands for Term Frequency-Inverse Document Frequency, which evaluates the importance of a term in a document relative to its occurrence in a larger corpus. It excels in highlighting essential terms and is sensitive to the uniqueness of words across documents. However, Term Frequency Inverse Document Frequency may still struggle with capturing context and semantics, and its effectiveness can diminish when applied to exceptionally large and diverse datasets.

3. Count Vectorizer

Akin to Bag of Words, creates a matrix of term frequency counts. Its strength lies in its ability to maintain word frequency information and contextual relationships. Nonetheless, it faces the disadvantage of generating large, sparse matrices, which can pose computational challenges, particularly with high-dimensional data.

Other methods like Latent Sentiment Analysis, Senti Word net [7], named entity recognition, and Jaccard similarity [2] are some of the well-known Natural Language Processing techniques. A few papers have also implemented the same but are not commonly used due to the limitations and compatibility with the models.

D. Machine Learning Models

Many machine learning algorithms have been investigated in cyberbullying detection to recognize and classify offensive or harmful content on online platforms. Support Vector Machines (Support Vector Machine) [4], Naive Bayes (NB), Decision Trees, and hybrid models that integrate aspects such as Gradient Boosting [2] and AdaBoost [2] are notable examples. Support Vector Machine, Naïve Bayes, Random Forest, and Naive Bayes stand out as the most used classifiers in literature. These models are often trained on manually labelled datasets, with each tweet classified as cyberbullying or not. Some research, on the other hand, uses processed and cleaned datasets to train these models, emphasizing the relevance of data pretreatment in the classification process.

The analysis of the literature demonstrates a wide range of model performance, with varied degrees of accuracy reported in different research studies. While promising, the maximum reported accuracy stays around 70%, indicating that the efficiency of cyberbullying detection algorithms is dependent on several parameters, including the quality and amount of the training data, feature engineering, and model complexity. Attained accuracy frequently decreases as model complexity increases, illustrating the trade-off between sophistication and reliability in real-world applications. Surprisingly, many research publications are concerned with the implementation method, with little discussion of the nuanced issues and complexities of model performance, highlighting the need for a more comprehensive knowledge of the practical aspects of cyberbullying detection.

Other deep learning models were also found to be implemented. Models like BERT [9], while highly effective in various natural language processing tasks, are not as commonly used for cyberbullying detection when compared to sentiment analysis with machine learning models for a few reasons. Cyberbullying detection typically involves identifying nuanced and context-specific forms of online harassment, which might not be adequately addressed by models trained in general language understanding tasks. Additionally, BERT and similar deep learning models are computationally intensive and require large, labelled datasets for fine-tuning, which can be challenging to acquire in the domain of cyberbullying due to the sensitive and context-dependent nature of such content. Machine learning models, with their simplicity and interpretability, often suffice for sentiment analysis, while deep learning models may be overkill for the task of distinguishing between positive and negative sentiment, making them less commonly adopted for this specific application. Hence, deep learning models were not the most effective choice to be used for such models to date.

E. Cyberbullying in various domains

Effectively addressing the multifaceted nature of cyberbullying detection demands a comprehensive and adaptable strategy. The persistent challenge of effectively tackling abusive behavior spanning diverse domains necessitates a reimagined approach. Existing models have demonstrated limitations in their ability to concurrently identify cyberbullying across various contexts. These constraints can be attributed, in part, to the constraints of previous datasets and the manual labor required for domain categorization. A promising remedy lies in the creation of a unified model designed to proficiently manage the nuanced attributes of multiple domains, encompassing racial, caste-based, gender-related, and location-specific forms of harassment [2] [8]. To achieve this, researchers are strongly encouraged to diligently curate extensive and domain-diverse datasets, explore adaptable machine learning frameworks, and establish automated categorization mechanisms.

TABLE II. EVALUATION METRICS OF SUPPORT VECTOR MACHINE CLASSIFIER

Category	*Precision*	*Recall*	*F1 score*
Neutral	0.96	1.00	0.98
Racist	0.98	0.87	0.97
Sexist	1.00	0.86	0.93
Avg	0.97	0.97	0.97

[1] In this research endeavor, the task of ternary classification was executed using a dataset 'C' (labeled dataset of 26,000 tweets) with the aid of a Support Vector Machine classifier, employing training to test data split of 7:3. Impressively, this approach yielded an overall accuracy of 97%, a noteworthy achievement in the domain of text classification. The dataset was categorized into the themes of sexism, racism, and neutral content. These labels were procured through a crowdsourcing platform, which facilitated the meticulous labeling of 16,914 tweets, with 3,383 tweets earmarked for sexist content, contributed by 613 users, and 1,972 tweets representing racism, submitted by 9 users. An additional 11,559 tweets were categorized as neither sexist nor racist, originating from 614 users. Subsequent data cleansing and the removal of inaccessible tweets culminated in a refined dataset of 16,000 tweets. To gauge the model's performance, essential evaluation metrics such as precision, accuracy, and recall were meticulously calculated and presented in a comprehensible tabular format.

A limitation arises when the model is applied in a multi-domain context. It exhibits high evaluation metrics when focused on a singular domain, such as racism or sexism, as seen in Table No. II. It has also been observed that the overall efficiency diminishes when ensembles of multiple domains are considered. Support Vector Machine classifier with linear kernel gave an accuracy of 55-57% over multiple iterations. Multinomial Naive Bayes with ternary classification gave an accuracy of 57 %. However, an overall accuracy of 80 % was achieved in the task of binary classification (bullying/non-bullying) These observations underscore the importance of expanding the labeled dataset and refining the model's capabilities to address multi-domain scenarios, enhancing its robustness and real-world applicability in combating cyberbullying.

To devise a model capable of recognizing cyberbullying instances across diverse domains like race, caste, gender, and location, several key steps should be considered. Firstly, one should compile or curate a broad and varied dataset that encompasses examples from all relevant domains. This dataset should be meticulously annotated with domain-specific information, facilitating the model's ability to learn and generalize effectively. Subsequently, the adoption of a versatile machine learning framework becomes imperative, one that can accommodate multiple classification tasks within a single model. Techniques such as multi-task learning or ensemble methods can be leveraged to simultaneously address different domains. Lastly, it is crucial to institute an automated categorization mechanism capable of discerning the domain context and classifying content as either constituting cyberbullying or not. By harnessing advanced methodologies in natural language processing and feature engineering, such a model can seamlessly adapt to varying domains, making a substantial contribution to the comprehensive mitigation of cyberbullying across a spectrum of contexts.

III. REVIEWS

After examining and analyzing the insights gleaned from the extensive research papers explored. These papers encompass a wide array of perspectives and approaches to the topic, offering valuable observations and methodologies. We will be delving into common themes and divergences, providing a comprehensive overview of the state of the field discussed in each paper and its model.

1. The imbalanced dataset and limited distinction between neutral tweets challenge the accuracy of traditional machine learning models like Support Vector Machine (55%) and Naïve Bayes (57%), urging the need for improved models and feature engineering in cyberbullying detection [1]

2. Random Forest's remarkable accuracy of 94% after tuning underscores its classification efficiency, whereas AdaBoost and Gradient Boosting exhibit lower accuracies, emphasizing the significance of selecting the appropriate ensemble method for specific classification tasks [2].

3. Support Vector Machine classifiers outperform Naïve Bayes in various evaluation metrics across n-gram models, with improved performance measures compared to previous studies, affirming their effectiveness in enhancing classification accuracy and reliability [3].

4. Support Vector Machine faces challenges in outlier misclassification, with a lack of an implementation solution provided, underscoring the importance of addressing outlier-related issues to enhance Support Vector Machine-based classification model robustness [4].

5. The Term Frequency Inverse Document Frequency method consistently outperforms the Bag of Words, while the Support Vector Machine stands out as the top-performing machine learning algorithm, highlighting its effectiveness in text classification tasks and potential for enhancing model accuracy and robustness [5].

6. The model's impressive accuracy (93%), precision (95%), and recall (97%), particularly when applied to Indonesian user data, signifies its relevance and adaptability in specific regional contexts, crucial for tailored Natural Language Processing solutions [6].

7. The study reports an overall classification accuracy of 70.3% for identifying cyberbullying tweets, emphasizing the ongoing challenge of achieving higher accuracy in this complex and evolving field of research [7].

8. The paper's failure to provide specific implementation and data analysis details raises concerns about reproducibility and hampers comprehensive evaluation of the methodology and results, indicating the need for improved transparency [8].

9. The comparative analysis reveals the superior accuracy of deep learning models like BERT-based approaches in cyberbullying detection, emphasizing the value of advanced deep learning techniques in addressing complex Natural Language Processing tasks [9].

These collective findings provide a solid foundation for the ongoing evolution of cyberbullying detection methodologies. It has provided valuable insights into the diverse landscape of cyberbullying detection, encompassing both its challenges and opportunities. The observations underscore the need for improved models and feature engineering to tackle issues.

IV. RESEARCH GAPS AND DISCUSSION

The exploration of research gaps in cyberbullying detection, as highlighted by the preceding reviews, reveals crucial areas that demand further investigation. A few of them have been discussed below:

A. *Comparative Analysis of Various Models*

A significant gap exists in a comprehensive comparative analysis of feature selection methods commonly used in text classification tasks, namely Bag of Words [2], Term Frequency-Inverse Document Frequency (TF-IDF) [2] [6], and Count Vectorizer. While these methods have been separately examined in multiple studies, a holistic assessment that directly compares their effectiveness in cyberbullying detection is notably absent. Such an analysis would not only offer insights into the relative performance of these approaches but also inform the selection of the most suitable method based on the specific requirements of a given context.

TABLE III. ACCURACY SCORES OF DIFFERENT ALGORITHMS [1]

Model	***Accuracy***
Vanilla Random Forest	80.75%
AdaBoost	90.12%
Gradient Boost	83.09%
Tuned Random Forest	94.06%

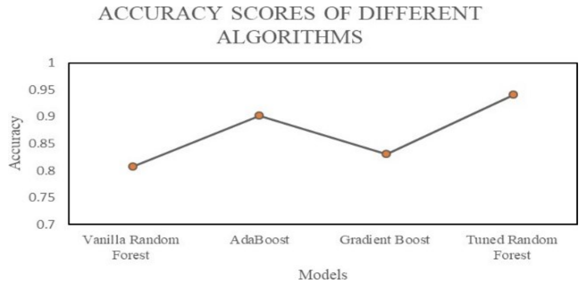

Fig. 2. Accuracy Scores of Different Algorithms

In a research model [2], other machine learning-based classifiers were implemented and compared. Classifiers like AdaBoost and Gradient Boost were experimented with, but none managed to outperform the tuned Random Forest classifier. Interestingly, Random Forest's simplicity and tunability emerged as a key advantage. This robustness is attributed to the model's use of random sampling in training each tree independently, effectively mitigating overfitting risks. The research's comprehensive accuracy analysis, summarized in Table No. III underscores the efficiency of different Machine Learning models in cyberbullying detection.

B. Feature Selection – Machine Learning combinations

There is a notable deficiency in research that explores the combinations of feature selection models with machine learning algorithms and subsequently identifies the combinations that yield the most promising results. To address this gap, future research should embark on comprehensive comparative studies that systematically pair various feature selection techniques with a range of machine learning models. Such an endeavour would provide a nuanced understanding of which combinations are best suited for cyberbullying detection tasks, paving the way for more effective and accurate models.

C. Overall Summary of Machine Learning Models

TABLE IV. COMPARISON OF RESULTS

Model Name	***Parameters***	***Accuracy***
SVM [1]	Manually tagged	55%-57%
Naïve Bayes [1]	Manually tagged	57%
SVM [1]	Labelled	97%
Vanilla Random Forest [2]	Labelled	80.75%
AdaBoost [2]	Labelled	90.12%
Gradient Boost [2]	Labelled	83.09%
Tuned Random Forest [2]	Count Vectorizer and TF-IDF in Preprocessing and 50 estimators and Gini index for information gain	94.06%

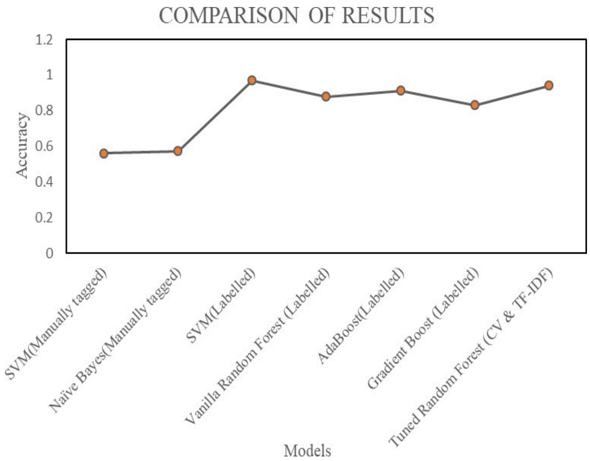

Fig. 3. Comparison of Results

The Table IV summarizes the accuracy scores of the Machine Learning models used by various research papers covered in our literature review. It is found that Support Vector Machine performs the best, giving 97% accuracy.

V. CONCLUSION

This review **study** presents a comprehensive overview of diverse methodologies and approaches to enhance the accuracy of cyberbullying detection. It acknowledges the challenges posed by manually flagging data, which can be a laborious task. Instead, it emphasizes the significance of obtaining a well-suited dataset or utilizing web scraping techniques to acquire specific data requirements efficiently. Furthermore, the **study** underscores the critical role of data preprocessing, highlighting that the most pivotal step in this process is ensuring thorough data cleaning, as it directly impacts the quality of data extraction. By employing a variety of combinations of machine learning models and feature extraction methods, the **study** suggests that optimal results can be achieved. The **study** provides valuable insights into the key components of an effective cyberbullying detection strategy, from data sourcing to preprocessing and the selection of machine learning models and feature extraction techniques.

The study plans to implement and present a comparative analysis of all the feature extraction techniques and Machine Learning models and analyze which combination gives the best result.

REFERENCES

[1]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.

[2]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.

[3]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

[4]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.

[5]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.

[6]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.

[7]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.

[8]B. R. Prathap and K. Ramesha, "Twitter Sentiment for Analysing 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.

[9]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.