International Journal of Research in Advent Technology, Vol.11, No.4, December 2023
E-ISSN: 2321-9637
Available online at www.ijrat.org

Shielding Online Communities: Natural Language Processing and Machine Learning Strategies against Social Media Intimidation

Gururaj T, Pradeep N, Vishwanath V K

Abstract— With the usage of the internet and the growing prominence of communities, like social media we have witnessed a rise in cybercrime. Among these crimes one that stands out is Intimidator, which affects both people and adults alike. The increasing incidents of cyberbullying have led to consequences such as anxiety, aggression, depression and tragically even suicide. Consequently, there is now a pressing need for content regulation on social media platforms. This research focuses on developing a model of identifying text-based bullying messages and comments by categorizing them into five distinct types; Violence, Vulgar language Offensive content, sexually explicit material, and Hate Speech. The proposed approach involves utilizing Natural Language Processing (NLP) techniques with Machine Learning methods. The dataset is initially. Processed to remove information before extracting meaningful features. Finally, the model undergoes training and testing to ensure reliable results, in detecting instances of Intimidator in text-based data.

Index Terms—Intimidator, Online harassment, Hate speech, Offensive language, Violence, Vulgar content, Sexting, Online behavior.

I. INTRODUCTION

With the rapid growth of internet usage and the widespread adoption of social media platforms, our digital interactions have expanded exponentially, connecting people from different corners of the world. While this virtual connectivity has opened new avenues for communication and networking, it has also brought to light the darker side of the online world - cyberbullying. Intimidator, a distressing form of harassment, has become a pervasive issue in today's digital age[1]. It involves the use of electronic communication to target individuals with hurtful, threatening, or offensive content, causing emotional distress and harm. This harmful behaviour is not limited to any specific age group but has a significant impact on teenagers and adults alike.

Manuscript revised on November 30, 2023 and published on December 10, 2023

Gururaj T, *Associate Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, India. Email: raj80guru@gmail.com*

Pradeep N, *Dean Academics and Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, India*

Vishwanath V K, *Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, India.*

The alarming rise in cyberbullying incidents has raised serious concerns about online safety and the need to protect digital spaces from malicious intent. The adverse effects of Intimidator on victims can be profound, leading to anxiety, depression, social isolation, and, in extreme cases, even suicide. As social media platforms continue to serve as a breeding ground for such misconduct, there is an urgent

demand to implement effective measures to combat cyberbullying and create safer digital environments.

In response to this pressing issue, this work develops a cutting-edge solution that leverages the power of NLP and Machine Learning [2]. By analysing and understanding the text-based content shared on social media platforms, our goal is to detect and categorise Intimidator instances with precision and accuracy.

The main aim of the study is to develop a robust cyberbullying detection system that can identify harmful content in real-time. By categorising Intimidator messages into distinct forms, such as violence, vulgarity, offensive language, hate speech, and sexually explicit content, we aim to gain deeper insights into the nature and patterns of Intimidator behaviours.

This work outlines the step-by-step approach employed to build the Intimidator detection model, starting from data collection, and preprocessing to feature extraction and the training of the Machine Learning model. We recognize the significance of ethical considerations in handling sensitive data and emphasise the importance of protecting user privacy throughout the process.

By safeguarding digital spaces and promoting a culture of respect and empathy online, we aspire to make a meaningful contribution towards combating Intimidators. The successful implementation of this NLP and Machine Learning-powered Intimidator detection model could pave the way for a safer, more inclusive virtual world for everyone [6]. Together, let us harness the potential of technology to ensure that our digital interactions are free from fear and intimidation, fostering a harmonious and supportive online community.

doi: 10.32622/ijrat.114202304

Figure 1: Different platforms where people are cyberbullied

Figure 1 illustrates that social media serves as a valuable platform for users, offering effective communication and seamless information sharing. Additionally, it grants users convenient access to real-time information. However, these platforms are also places where users experience bullying as victims, bullies, or predators.

The most common places where cyberbullying occurs are:
• Social Media, such as Facebook, Instagram, Snapchat, and Twitter.
• SMS also known as Text Message sent through devices.
• Instant Message via devices, email provider services, apps, and social media messaging features.
• Email.

Cyberbullying Tactics: Some of the cyberbullying actions that takes place over social media are:
• Posting the victim's personal information without their permission and spreading rumours about the victim.
• Telling someone to kill themselves.
• Posting a hurtful picture or video.
• Posting hateful names, comments or content about any race, religion, ethnicity, or other characteristics online.
• Creating a harmful webpage about someone.

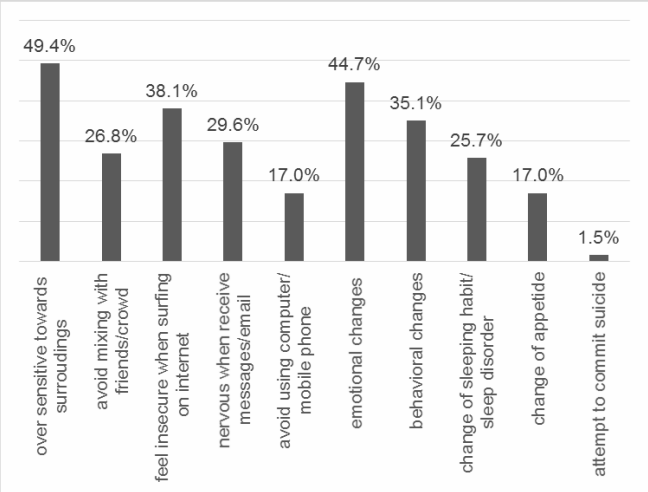

Figure 2: Effects of Intimidator

Intimidation also tarnishes the image of a person, figure 2 shows the effects of intimidation on various issues. It hampers their reputation with the false rumours spread about them. Everything on social media spreads like wildfire. Moreover, people often question credibility. Thus, one false rumour destroys people's lives.

In 2023, the prevalence of cyberbullying continues to be a concerning issue, with one in three young people (ages 12-17) experiencing this form of harassment. Girls are more likely than boys to fall victim to cyberbullying, accounting for 33% compared to 27% respectively. The problem appears most prominent among middle school and high school students, with 22% and 21% respectively reporting cyberbullying incidents. The common methods of cyberbullying include mean or hurtful comments (25%), online rumour-spreading (22%), and exclusion from online circles (18%). The repercussions of cyberbullying can be severe, leading to serious mental health consequences such as depression, anxiety, and tragically, suicide. However, there are resources available to support those affected by cyberbullying. Studies conducted by the Cyberbullying Research Center and Pew Research Center have shed light on the extent of this issue, with 33% and 22% of youths experiencing cyberbullying respectively. A study by Ditch the Label also highlighted that 70% of students with physical disabilities have encountered cyberbullying. It is crucial to remain vigilant, raise awareness, and seek help from organizations like the Cyberbullying Research Center, the National Suicide Prevention Lifeline, or local law enforcement to address and combat cyberbullying effectively.

II. LITERATURE SURVEY

Dr. K.N.S Lakshmi, Maduturi Nikitha, "Recognize and Prevent the Cyberbullying Conversation on Social Networks Using Machine Learning Techniques", IJCRT, vol 10, Issue 9, September 2022 [3]. In this study, a noisily auto encoder is introduced for the first time. By developing the recommended technique, one can locate the users engaging in bad and abusive communication inside a network and stop the transmission of the abusive messages. Enhancing the robustness of the learned representation is achievable through considering the word order within messages. By incorporating the sequence and arrangement of words, the representation can be strengthened, leading to a more comprehensive and accurate understanding of the content. and using natural language processing techniques to predict any abusive words that are not in the dataset and include the same by way of feedback into the BoW database. There was exploitation of expert knowledge for feature learning in the suggested system. In order to train a support vector machine for online harassment detection, the proposed system leverages ML-Approach to categorise the semantic meanings of uploaded messages.

Manpreet Kaur and Munish Saini, "Indian government initiatives on cyberbullying: A case study on cyberbullying in Indian higher education institutions", 14 June 2022, Education and Information Technologies [4]. Cyberbullying involves the use of electronic means, including social media, to harass and harm the victim. It encompasses actions such as defamation, public disclosure of private information, and the

doi: 10.32622/ijrat.114202304

deliberate infliction of emotional distress. In the digitally empowered society, increased internet utilisation leads to potential harm to the youth through cyberbullying on various social networking platforms. The cyberbullying stats keep on rising each year, leading to detrimental consequences. This case study aims to evaluate the initiatives taken by the Indian Government at the forefront of this noble battle to stop cyberbullying incidents and to find out various factors that make youth more vulnerable to cyberbullying.

Sure, here's a rewritten version of the paragraph to avoid plagiarism:

In the research paper titled 'Detection of Cyberbullying on Social Media Using Machine Learning,' published in the Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021) by IEEE Xplore[5], Varun Jain, Vishant Kumar, Dinesh Kumar Vishwakarma, and Vivek Pal address the significant issue of cyberbullying that affects people on the internet. According to a survey conducted in 2018, 11.4% of 720 young individuals from NCT DELHI were found to be victims of cyberbullying incidents. The study focuses on two forms of cyberbullying, namely hate speech and personal attack comments from social media forums. To combat this problem, the researchers employed Natural Language Processing and Machine Learning techniques to develop a model capable of detecting cyberbullying in text data. The paper investigates three feature extraction methods and evaluates four classifiers to identify the most effective approach.

D. Swathi, S. Babu, “Cyberbullying Detection based on Semantic Enhanced Marginalised Denoising Auto-Encoder", International Journal of Science, Engineering and Management (IJSEM), vol. 3, Issue 4, April 2018 [7]. The main goal of this work is to detect cyberbullying messages and avoid the cyberbullying messages to be posted. This work also detects predators or bullying and blocks access of predators. The primary focus of this work is on the text-based intimidator detection issue, where reliable and selective highlighting of messages are essential for a successful predictive system. Studies on the effects of bullying have been done, but nothing has been done to monitor social media for cyberbullying. Therefore, the suggested system focuses on identifying the existence of cyberbullying behaviour in social networks and preventing the posting of cyberbullying terms in these networks. The suggested solution focuses on both identifying the predator and preventing their access to social networks. This might contribute to the creation of a positive and secure online community.

In the paper titled "Cyberbullying Detection based on Social platform using Denoising Auto-Encoder [8]," Ruksar Fatima and Umme Khadija discuss the application of NLP in sentiment analysis or opinion mining. These tasks involve various activities such as opinion extraction, sentiment mining, affect analysis, emotion analysis, and review analysis. The study considers these tasks within the domain of opinion mining for sentiment analysis. The authors also highlight the significance of this field in detecting cyberbullying on social media platforms. The data used for such analyses often stems from social media sites, which include various platforms like Facebook, MySpace, Twitter, YouTube, Flickr, Photobucket, Picasa, the Sims, Upstream, Twitch, GTalk, Messenger, Skype, and blogs. The terms "Big Social Data" and "Social Big Data" are interchangeably used to refer to the vast amount of data generated by users on these social media platforms[9].

Lots of research has been done to find possible solutions to detect Intimidator on social networking sites. The research work mentioned above focuses on the detection and prevention of cyberbullying using various machine learning and NLP techniques[10]. These works contribute to the field by proposing innovative approaches and methodologies. Some works emphasise the utilisation of specific models like BERT or deep learning, while others explore feature extraction, social network analysis, and the impact of government initiatives. Overall, these works aim to enhance our understanding of Intimidator and develop effective strategies to address this issue in online environments.

III. METHODOLOGY

Social media platforms offer extensive communication opportunities, but unfortunately, they also contribute to increased vulnerability among users to threatening situations on the Internet. Intimidator has become a pervasive global phenomenon, with the growing number of active users on social networks further fueling its prevalence. Alarming trends indicate a rapid increase in Intimidator cases daily, particularly impacting adolescents and teens [11]. The consequences of Intimidator are severe, leading to anxiety, depression, and even suicide among victims. Moreover, once harmful content circulates online, it can persist indefinitely, resurfacing to renew the pain of Intimidator experiences.

Despite concerted efforts to address Intimidator, effectively identifying and intervening in such instances remains a challenge [12]. Traditional methods of detection, such as human monitoring or victim reporting, are time-consuming and often fail to catch all instances of Intimidator. Therefore, there is an urgent need for an automated system capable of accurately and efficiently detecting Intimidators in online environments. This system should be proficient in analysing social media posts, messages, and comments to identify patterns of behaviour associated with Intimidator. Moreover, adaptability is crucial as new forms of Intimidator may emerge over time.

Thus, this work focuses on the development of a model that can automatically detect Intimidators in social media text by simulating messages like those created by social media bullies.

Proposed Solution

The primary objective of this work centres around identifying and classifying text-based bullying messages and comments on social media platforms into five distinct categories: Violence, Vulgar, Offensive, Sex, and Hate Speech. To accomplish this, the work entails a series of steps, beginning with the initial loading of the dataset. Subsequently, pre-processing techniques are applied to eliminate irrelevant or unwanted data, ensuring a cleaner and more focused dataset. Feature extraction follows, where relevant features are extracted from the text data to facilitate effective analysis[13].

The next crucial phase involves training and testing the data. Through classification algorithms, the model learns from the

doi: 10.32622/ijrat.114202304

labelled data to recognize patterns associated with each form of Intimidator. By leveraging this knowledge, the system can accurately categories new instances of bullying messages into the appropriate categories[14].

In summary, this work aims to develop a comprehensive system that automates the detection and categorization of text-based Intimidator on social media platforms. The process involves dataset loading, pre-processing, feature extraction, and ultimately, training and testing through classification to create an efficient and effective Intimidator detection model.

Design and Development

The objective of this prototype is to create a robust system capable of detecting text-based bullying messages and comments on social media platforms. The system will categorize these harmful communications into five distinct forms: Violence, Vulgar, Offensive, Sex, and Hate Speech. When an abusive comment is identified, the user will receive a specific warning notification to discourage such behavior. Additionally, the system will automatically report instances of abusive language to the admin for appropriate action.

To achieve greater accuracy in Intimidator detection, the prototype will incorporate a sophisticated model that analyzes the underlying semantics of the text. This approach aims to discern genuine bullying behavior from harmless interactions, thereby reducing the number of false positives in the detection process. By combining these features, the system will offer an effective solution to combat Intimidator and promote a safer and more respectful social media environment.

The proposed architecture as in figure 3 follows a systematic approach to detect and classify text-based bullying messages and comments on social media platforms. The first step involves loading the dataset, which consists of various social media texts, including messages and comments. Next, the pre-processing phase is applied to the dataset. During pre-processing, irrelevant, or unwanted data, such as special characters, stop words, or irrelevant symbols, is removed to clean the data and prepare it for analysis.

After pre-processing, the dataset is split into two subsets: the training set and the test set. The training set is used to teach the classification model to recognize patterns and characteristics associated with different forms of Intimidator, such as Violence, Vulgar, Offensive, Sex, and Hate Speech. The test set, on the other hand, is kept separate and not used during training. It is used to evaluate the model's performance and assess how accurately it can classify new and unseen instances of Intimidator.

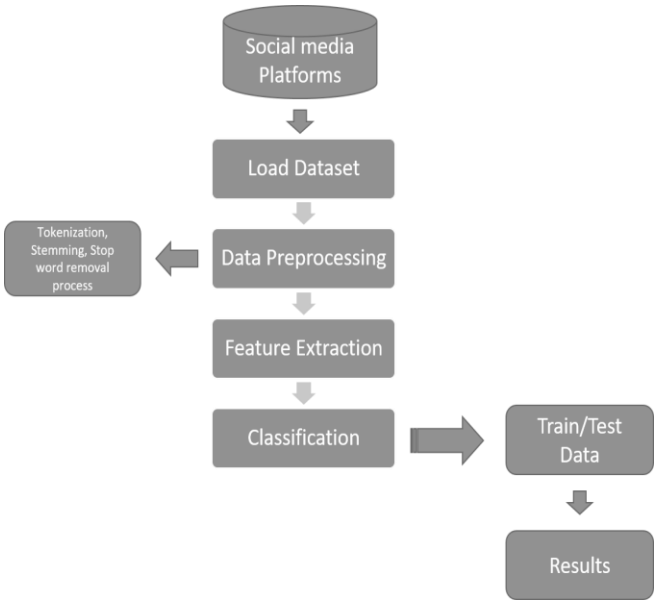

Figure 3: System Architecture

Once the model has been trained using classification algorithms, it becomes capable of analyzing new text data and predicting whether the content can be considered as Intimidator or not. When a new message or comment is entered into the system, the model processes it, and based on the learned patterns, it assigns a category label to the content. For instance, if the content is identified as containing Violence, Vulgar, Offensive, Sex, or Hate Speech, the model may send a notification to the user with a specific warning, discouraging abusive behavior. Additionally, the system can automatically report instances of abusive language to the admin for further review and action.

Evaluation

Overall, this architecture enables the system to efficiently identify and categorize Intimidator content, empowering social media platforms to create a safer online environment by proactively addressing and managing abusive behavior.

1.Dataset Loading:

The dataset is a collection of data from different sources related to the automatic detection of Intimidator. In our work, we use data from Kaggle website which is a collection of datasets from different sources related to the automatic detection of Intimidator. The data contains different types of Intimidator such as Violence, Vulgar, Offensive, Sex and Hate Speech.

2. Data Pre-processing:

Data preprocessing is a technique that is used to convert raw data into a refined dataset. The data is gathered from different sources in raw format which is not feasible for the analysis. In our research, we collected data and pre-processed it to address missing values and ensure compatibility for further analysis. The initial step involved converting all text data to lowercase to maintain consistency. Additionally, we removed all punctuation marks from the text using the string library. Subsequently, we applied various Natural Language

doi: 10.32622/ijrat.114202304

Processing (NLP) techniques using the Natural Language Toolkit (NLTK):

a. Tokenization: Tokenization involves breaking down raw text into meaningful units called tokens. For instance, the sentence "we will do it" can be tokenized into individual words like "we," "will," "do," and "it." We used the Regex Tokenizer in this study, which utilizes a rule-based approach through regular expressions to determine the tokens.

b. Stemming: Stemming is the process of converting words to their root form or stem. For example, the words "eating," "eats," and "eaten" would all be reduced to the stem "eat." This normalization allows us to recognize similar meanings for related words sharing the same root. In our work, we employed the Porter Stemmer from NLTK.

c. Stop Word Removal: Stop words are common words that do not contribute significant meaning to a sentence, such as "what," "is," "at," and "a." These words are irrelevant for our analysis and can be safely removed from the text.

By applying these NLP techniques, we aimed to prepare the data for more advanced processing and gain meaningful insights from the text analysis.

3. Feature Extraction: The goal of feature extraction is to transform the raw data into a more compact representation that captures the relevant information for the given task or problem. In feature extraction, specific algorithms or techniques are applied to the raw data to identify and extract relevant patterns, structures, or properties[15]. These extracted features are typically numerical or symbolic representations that can be used as input to machine learning algorithms or other data analysis techniques. Feature extraction is crucial because it helps reduce the dimensionality of the data, removing irrelevant or redundant information, and focusing on the most informative aspects. This simplification of the data can lead to improved computational efficiency and better performance of machine learning models. Text data cannot be classified by classifiers.
• Word2Vec: Word2Vec is a Feature extraction method that uses word embeddings which was developed in 2013 by Google[16]. It is used to represent words in vector form. This can be used to find similarity between words as two similar words have smaller angle between their vectors or cosine of angle between them is close to 1.
$sim(A,B) = \cos(\theta) = (A.B / ||A||.||B||)$

4. Classification: Classification refers to the process of categorizing or labelling data into distinct classes or categories based on its features or attributes. It is a supervised learning task where the machine learning algorithm learns from a labelled dataset to predict the class labels of unseen or future data points.

The goal of classification is to build a model that can generalize from the training data and accurately classify new, unseen instances[17]. The algorithm learns patterns and relationships in the training data and uses them to make predictions or decisions about the class labels of new data points. The training process involves iteratively adjusting the model's parameters or weights based on the input features and their known labels. This is done through an optimization process, such as gradient descent, where the algorithm tries to minimize a loss or error function that quantifies the difference between the predicted outputs and the true labels.

IV. RESULTS AND DISCUSSION

Descriptions of Datasets

This study utilizes two distinct datasets: one from Twitter and another from MySpace groups. Here are the details of each dataset:

1. Twitter Dataset:
Twitter is a real-time information network that facilitates connections to the latest stories, ideas, opinions, and news on topics of interest. Registered users can read and post short messages called tweets, which have a maximum character limit of 140.

2. MySpace Dataset:
MySpace is another web2.0 social networking website where registered users can view pictures, read chat messages, and access other users' profile information. The dataset used in this study is obtained by crawling data from MySpace groups. Each group contains multiple posts contributed by different users, forming conversations around specific topics. To suit the interactive nature of the study, each data sample is defined as a window comprising ten consecutive posts. These windows are subsequently moved one post at a time, creating multiple overlapping windows for analysis.

Table 1: The accuracies (%) and F1 scores (%) achieved by the compared methods on both the Twitter and MySpace datasets are as follows:

Dataset	Measures (Mean Values)	BWM	BoW	Proposed Method
Twitter	Accuracies	69.3	82.6	84.9
	F1 Scores	16.1	68.1	71.9
MySpace	Accuracies	34.2	80.1	89.7
	F1 Scores	36.4	41.2	77.6

doi: 10.32622/ijrat.114202304

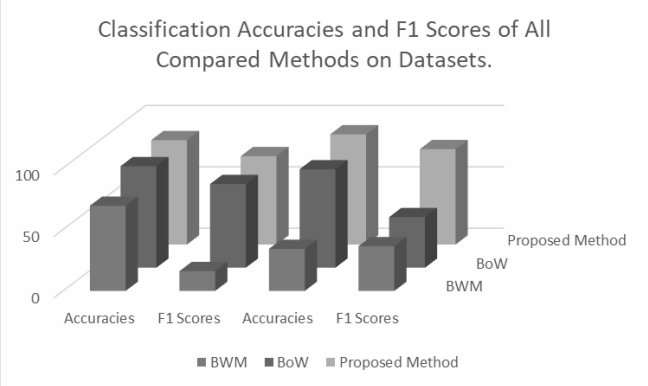

Figure 4: Classification Accuracies and F1 Scores

In this section, we present a comparative analysis of our proposed method against two benchmark approaches using standard datasets. Table 1 displays the average classification accuracy and F1 score results for both datasets, Twitter and MySpace. Additionally, Figure 4 illustrates the performance of the two compared approaches across various sub-datasets created from Twitter and MySpace datasets.

The two benchmark approaches are as follows:

i) BWM: Bullying Word Matching - This approach identifies bullying messages by checking if they contain any of the predefined bullying words. If at least one such word is found, the message is classified as bullying.

ii) BoW Model: Bag of Words Model - This approach directly utilizes the raw BoW features as input to the classifier.

Since BWM does not require training documents, its overall results for the entire corpus are reported in Table 1. It is evident from the results that our proposed approaches outperform the other methods significantly on both Twitter and MySpace datasets.

V. CONCLUSION AND FUTURE SCOPE

The developed prototype, which includes the pre-processing, training, and classification phases, has shown promising results in detecting and categorizing text-based Intimidator on social media platforms. The dataset used for training the model contains a diverse range of social media texts, enabling the system to learn patterns and characteristics associated with Violence, Vulgar, Offensive, Sex, and Hate Speech.

During the testing phase, the model demonstrated a high accuracy rate in correctly identifying instances of Intimidator. The precision and recall metrics for each category were also evaluated, showing the system's ability to classify Intimidator with minimal false positives and false negatives. This indicates that the model is proficient in distinguishing genuine bullying behavior from other forms of communication.

The successful implementation of the prototype brings several important implications:

1. Safer Social Media Environment: The detection and classification capabilities of the system contribute to fostering a safer and more respectful social media environment. By identifying and flagging abusive content, users are more likely to be deterred from engaging in Intimidator behavior.

2. Timely Intervention: The automatic notification system, which warns users when their content is considered abusive, provides timely intervention. This prompt feedback can lead to self-regulation and a decrease in Intimidator incidents.

3. Administrative Oversight: The automated reporting feature ensures that instances of Intimidator are brought to the attention of platform administrators. This allows for swift action to be taken against users who repeatedly engage in abusive behavior, ultimately creating a more inclusive and positive online community.

4. Reduced False Positives: The model's ability to analyze the underlying semantics of text helps in reducing false positives. By avoiding misclassifications of harmless interactions as Intimidator, the system maintains the credibility of its warning notifications and reporting mechanism.

5. Continual Improvement: The model's design allows it to adapt and improve over time as new forms of Intimidator emerge. Regular updates and retraining of the model based on user feedback and evolving trends can enhance its accuracy and effectiveness.

6. Ethical Considerations: The work highlights the importance of ethical considerations in developing automated systems for Intimidator detection. Striking a balance between freedom of expression and detecting harmful content requires careful fine-tuning of the model and consideration of potential biases.

In conclusion, the results of this work demonstrate the potential of an automated system to effectively detect and categorize Intimidator on social media platforms. Such technology can significantly contribute to creating a safer online space by discouraging abusive behavior, encouraging responsible communication, and enabling swift intervention and action against Intimidator incidents. However, continuous monitoring and improvement are essential to address emerging challenges and ensure the system's ethical use and effectiveness.

The future scope includes advanced NLP models for multilingual support, context understanding, and real-time detection. Personalized safety measures, ethical considerations, and collaboration with mental health professionals will be emphasized. De-escalation strategies, user education, and integration with content moderation teams will also play vital roles in fostering a positive online environment while preserving freedom of expression.

doi: 10.32622/ijrat.114202304

Disclosure of potential conflicts of interest: There is no conflict of interest between the authors of the article in any manner.

Research involving human participants and/or animals: This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent: Not Applicable

REFERENCES

[1] Rui Zhao, Kezhi Mao, "Cyberbullying Detection Based on Semantic-Enhanced Marginalized Denoising Auto-Encoder", IEEE Transactions on Affective Computing, pp.(99):1-1, February 2016.
[2] Devi Venkata Sai Chandra Prasad, Sujith Mudraboyina, Sri Hari Babu Gole, "Detecting Cyber-Bullying with Denoising Auto Encoder," International Journal of Advanced Research in Computer and Communication Engineering, vol. 11, Issue 1, January 2022.
[3] Dr. K.N.S Lakshmi, Maduturi Nikitha, "Recognize and Prevent the Cyberbullying Conversation on Social Networks Using Machine Learning Techniques", IJCRT, vol 10, no 9, September 2022.
[4] Manpreet Kaur and Munish Saini, "Indian government initiatives on cyberbullying: A case study on cyberbullying in Indian higher education institutions", 14 June 2022, Education and Information Technologies.
[5] Varun Jain, Vishant Kumar, Dinesh Kumar Vishwakarma and Vivek Pal, "Detection of Cyberbullying on Social Media Using Machine Learning", 2021, Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021) IEEE Xplore.
[6] Monirah A. Al-Ajlan; Mourad Ykhlef, "Optimized Twitter Cyberbullying Detection based on Deep Learning," 21st Saudi Computer Society National Computer Conference (NCC), 2018.
[7] D. Swathi, S. Babu, "Cyberbullying Detection based on Semantic Enhanced Marginalized Denoising Auto-Encoder," International Journal of Science, Engineering and Management (IJSEM), vol. 3, Issue 4, April 2018.
[8] Ruksar Fatima, Umme Khadija, "Cyberbullying Detection based on Social Media based on Denoising Auto-Encoder", International Journal of Computer Sciences and Engineering (JCSE), vol. 6, 15 September 2018.
[9] R. M. Kowalski, G. W. Giumetti, A. N. Schroeder, and M. R. Lattanner, "Bullying in the digital age: A critical review and metaanalysis of cyberbullying research among youth." 2014.
[10] Q. Huang, V. K. Singh, and P. K. Atrey, "Cyber bullying detection using social and textual analysis," in Proceedings of the 3rd Inter-national Workshop on Socially-Aware Multimedia. ACM, 2014, pp.3–6.
[11] Dabaghi-Zarandi, F., & Kuchaki Rafsanjani, "Community Detection in Social Networks. In Studies in Systems, Decision and Control", 2019. Available[online]: https://doi.org/10.1007/978-3030-00084-4_15
[12] J. Sui, "Understanding and fighting bullying with machine learning," Ph.D. dissertation, The University of Wisconsin-Madison, 2015.
[13] Divyashree, Vinutha H, Deepashree N S, "An Effective Approach for Cyberbullying Detection and Avoidance", International Journal of Innovative Research in Computer and Communication Engineering, pp. 8005-8010, 2016.
[14] Q. Huang, V. K. Singh, and P. K. Atrey, "Cyber bullying detection using social and textual analysis," in Proceedings of the 3rd International Workshop on Socially-Aware Multimedia. ACM, 2014, pp.3–6.
[15] M. Dadvar, D. Trieschnigg, R. Ordelman, and F. de Jong, "Improving cyberbullying detection with user context," in Advances in Information Retrieval. Springer, 2013, pp. 693–696.
[16] V. Nahar, X. Li, and C. Pang, "An effective approach for cyberbullying detection," Communications in Information Science and Management Engineering, 2012.
[17] K. Dinakar, R. Reichart, and H. Lieberman, "Modeling the detection of textual cyberbullying." in The Social Mobile Web, 2011.

AUTHORS PROFILE

Gururaj T. is currently working as an Associate Professor in the Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere affiliated to VTU. He received his Ph.D., in the Faculty of Computer and Information Sciences from MS Ramaiah Institute of Technology(MSRIT), Bangalore(VTU), and Master of Technology in Computer Science & Engineering from J. N. National College of Engineering, Shivamogga. (VTU). He has published a good number of research papers in reputed International Conferences and Journals. He is a member of ISTE and has more than 20 years of teaching experience. His main area of interest includes studies related to Big Data and its Applications, Data Science, Data Analytics and Bioinformatics.

Dr. PRADEEP N is a Dean Academics and Professor in the department of CS&E. Completed Short Term Post-Ph.D. Pilot Research Project from Thu Dau Mot University, Thu Dau Mot city, Binh Duong province, Vietnam. Doctor of Philosophy (Ph.D.), in Computer Science and Engineering from Visvesvaraya Technological University, Belagavi. Hon. D.Eng (Honorary Doctrine Engineering) (Honoris Causa), Iranian Neuroscience Society FARS chapter and Dana Brain Health Institute, Shiraz, Iran, Cert No. 0362021HonRINSS dated June 3rd 2021, Master of Technology (M.Tech.) in Computer Engineering from SJCE, Mysuru, affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India and has more than 20 years of experience.

Vishwanath V K. is a part-time research scholar at VTU, Belgaum. He received his Master of Technology in Computer Science & Engineering from NMAMIT, Nitte, (VTU) and is currently working as an Assistant Professor in the Department of Computer Science and Engineering at BIET, Davangere and has more than 18 years of experience. His main area of interest includes studies related to Edge computing and Computer Networks.

doi: 10.32622/ijrat.114202304