Identification of Cyberbullying in Social Media using Machine Learning

[,1] Prof.Sagar Dhanake, [2] Adesh Dorge , [3] Ganesh Hande, [4] Akshay Gaykar

[1] Professor, Dr. D. Y. Patil College of Engineering and Innovation, Varale, Pune, India

[2]Student, Dr. D. Y. Patil College of Engineering and Innovation, Varale, Pune, India

[3]Student, Dr. D. Y. Patil College of Engineering and Innovation, Varale, Pune, India

[4] Student,Dr. D. Y. Patil College of Engineering and Innovation, Varale, Pune, India

Abstract : In the modern era, the usage of internet has increased tremendously which in turn has led to the evolution of large amount of data .Cyber world d has its own pros and cons. One of the alarming situations in web 4.0 is cyber bullying a type of cyber-crime. When the bullying occurs on line with the aid of technology it is known as cyber bullying. This research paper have surveyed the work done by 30 different researchers on cyber bullying, and elaborated on different methodologies adopted by them for the detection of bullying, and how you protect the community from online evil act of cyber bullying. Cyber- crimes involve all the crimes where internet is used as an access medium and committed through some electronic device such as computers and mobile phones. Unavailability of datasets, hidden identity of predators and the privacy of the victims are the main factors for limiting the past research in cyberbullying detection.

Index Terms - cyberbullying; machine learning; victim; social media; dataset.

Introduction

The largest Social media platforms have become integral parts of our daily lives, serving as channels for communication, information sharing, and community building. While these platforms offer numerous benefits, they also present a darker side 3cyberbullying. Cyberbullying is a pervasive and harmful issue that involves the use of digital technologies to harass, threaten, or demean individuals. It not only damages the emotional well-being of victims but also undermines the positive aspects of online interaction. In this paper, we delve into the field of cyberbullying detection using machine learning techniques. We explore the various challenges associated with identifying cyberbullying in social media content, including the ever-evolving nature of online harassment and the nuances of language and context. Additionally, we investigate the different approaches, datasets, and features commonly used in the development of cyberbullying detection models.

However our goal is to provide a comprehensive overview of the state of the art in cyberbullying identification through machine learning. We discuss the potential impact of these systems on creating safer and more inclusive online environments. Furthermore, we evaluate the effectiveness of different machine learning models and highlight the need for ongoing research in this critical area. By the end of this paper, readers will have a solid understanding of the advancements made and the challenges that persist in the fight against cyberbullying on social media platforms.

The battle against cyberbullying is an ongoing and ever-evolving struggle, but with the integration of machine learning techniques, we aim to make significant strides in identifying and mitigating this detrimental behavior, fostering a more respectful and secure online world for all users.

Literature survey:

Sr.No	Paper Name/Year	AuthorName	Description
1	Rapid Cyber-bullying detection method using Compact BERT Models (2021)	Mitra Behzadi, Ian G.Harris,Ali Derakhshan	Nowadays, some people use their social media platform to spread hate online and that is why the problem of cyber-bullying detection has been the focus of many researchers over the past decade.
2	Identification and Classification of Cybercrimes using Text Mining Technique (2021)	ShizaAndleeb,Dr.Rashid Ahmed,Zaheer Ahmed,Maira Kanwal	• This research paper have surveyed the work done by 30 other researchers on cyber bullying, and elaborated on other methodologies adopted by them for the detection of bullying, and how you protect the community from online evil act of cyber bullying.
3	Identification and Classification of Cybercrimes using Text Mining Technique	ShizaAndleeb, Dr.Rashid Ahmed,Zaheer Ahmed,Maira Kanwal	cyber-crimes involve all the crimes where internet is used as an access medium and committed through some electronic device such as computers and mo-bile phones.
4	Identification of Potential Cyber Bullying Tweets using Hybrid Ap-proach in Sentiment Analysis	Akankshi Mody , Shreni Shah , Reeya Pimple ,Narendra Shekokar	The aim of this study is to do the same by using sentiment analysis. We perform cyber bullying detection using a novice approach on Tweets using Natural Language Processing and Machine Learning techniques.

Project Aim:

Cyber bullying is the problem in current Situation of the world Because All the Student or humans use social media. we try to prevent Cyber bullying through our project . The propose is to find an efficient way to detect sarcastic tweets, and study how to use this information (i.e., whether the tweet is sarcastic or not) to enhance the accuracy of Cyber bullying .

Algorithm & System Model:

Module

• Admin

• In this module,the Admin has to log in by using valid user name and password. After login successful he can do some operations such as View All Users and Authorize, View All E-Commerce Website and Authorize,View All Products and Reviews,View All Products Early Reviews, View All Keyword Search Details, View All Products Search Ratio, View All Keyword Search Results, View All Product Review Rank Results.

• View and Authorize Users

• In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and College Short Form Name, Department of Computer Engineering 2021 22 admin authorizes the users.

• View Charts Results

• View All Products Search Ratio,View All Keyword Search Results,View All Product Review Rank Results.

• Ecommerce User • In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful,he has to login by using authorized user name and password Once Login is successful user will do some operations like Add Products, View All Products with reviews, View All Early Product's reviews, View All Purchased Transactions.

• End User

• In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will best or to the database. After registration successful,he has to login by

using authorized user name and password. Once Login is successful user will do some operations like Manage Account, Search Products by keyword and Purchase, View Your Search Transactions, View

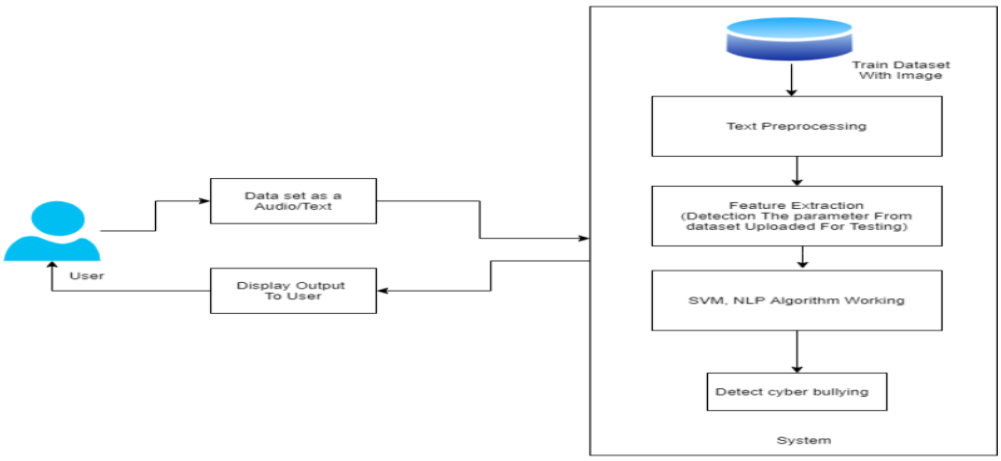

Fig.1

Working Methodology:

A cyber bullying detection system utilizing Support Vector Machine (SVM) and Natural Language Processing (NLP) algorithms employs a multi-step methodology to effectively identify and combat online harassment. Initially, the system gathers textual data from various online sources, such as social media platforms and communication channels. The NLP algorithm plays a pivotal role in preprocessing this textual information, extracting relevant features, and converting the unstructured data into a format suitable for machine learning analysis. NLP techniques, including sentiment analysis and language pattern recognition, are applied to discern the underlying tone and context of the text.

Following the preprocessing stage, the SVM algorithm comes into play. SVM, a supervised learning model, is trained on labeled datasets to distinguish between instances of cyberbullying and normal online communication. It works by creating a hyperplane that best separates different classes of data in a high-dimensional space. In the context of cyber bullying detection, SVM learns to discriminate between harmful and non-harmful content based on the features extracted through NLP.

Once trained, the SVM model can then classify new, unseen data, effectively identifying instances of cyber bullying. The system can be fine-tuned and updated with additional labeled data to continually improve its accuracy and adapt to evolving patterns of online harassment. By combining the strengths of SVM for classification and NLP for linguistic analysis, this methodology provides a robust approach to detecting cyber bullying in online communication, contributing to a safer and more positive digital environment.

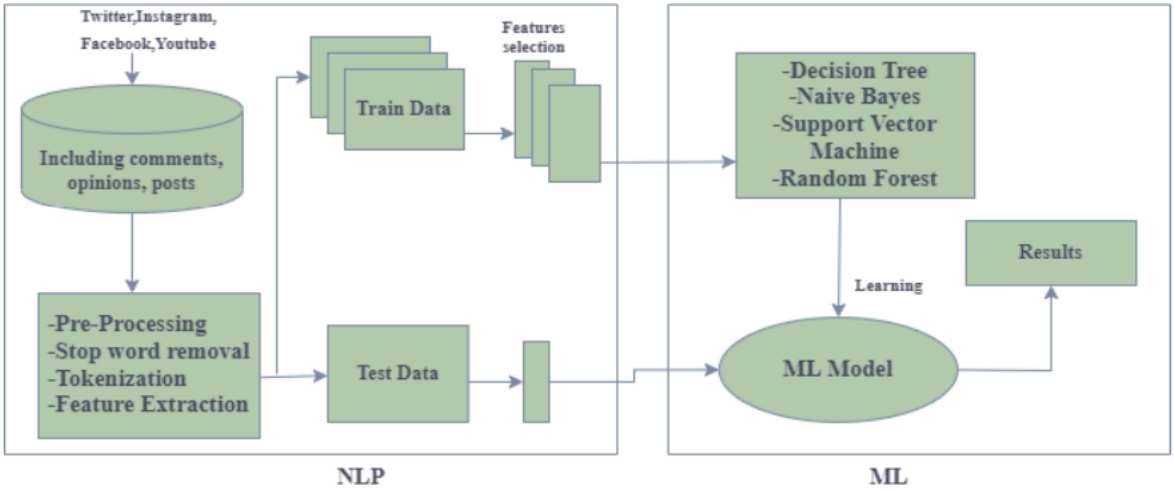

Fig 2: Framework of Cyberbullying Detection System

Conclusion:

In this work, a system is proposed which detects on English as well as on Hindi tweets in Twitter. Cyber bullying is very dependent and highly contextual; therefore, sentiment and other contextual clues to help detect the Cyber bullying . The system uses sarcastic tweets, 9,111 tweets containing Cyber bullying , and not dataset. The system uses the LR Algorithm . The approach has shown good results and it is observed that LR classifier has more accuracy than other classifier. All patterns for sarcastic detection are not covered in the extracted patterns. From the survey it can be concluded that the traditional machine learning algorithms are incapable of handling the enormous amount of data being generated in Web 4.0 moreover the cyber bullying content cannot be detected accurately .Recently Deep learning techniques, NLP, deep recurrent neural network,CNN, stacked auto-encoder,has gained the attention of some researchers.Future work can target on usage of these deep learning techniques for precise detection of cyberbullying in social media.Lot of research is being conducted in the field of cyberbullying. It is an emerging issue which needs to be addressed in Web 4.0 .After reviewing the 30 research papers it found that there is a lack of proper dataset, collecting huge dataset is a major challenge, integrating social, contextual, sentiment features can improve the accuracy of detection of bullying content. For future work, data from multiple social media platforms can be considered , apart from text, image , video must be taken into account for experimentation.

References:

[1] D. Yin, Z. Xue, L. Hong, B. D. Davison, A. Kontostathis, and L. Edwards, "Detection of harassment on web 2.0." Proceedings of the Content Analysis in the WEB 2, 2009, pp. 1-7.

[2] K. Dinakar, R. Reichart, and H. Lieberman. "Modeling the detection of textual cyberbullying." In fifth international AAAI conference on weblogs and social media. 2011.

[3] J. Bayzick, A. Kontostathis, and L. Edwards. "Detecting the presence of cyberbullying K. Elissa, “Title of paper if known,” unpublished.

[4] K. Reynolds, A. Kontostathis, and L. Edwards. "Using machine learning to detect cyberbullying." In 2011 10th International Conference on Machine learning and applications and workshops, IEEE, vol. 2, 2011, pp. 241-244.

[5] K. Dinakar, B. Jones, C.Havasi, H. Lieberman, and R. Picard. "Common sense reasoning for detection, prevention, and mitigation of cyberbullying." ACM Transactions on Interactive Intelligent Systems (TiiS) 2, no. 3, 2012, p. 18.

[6] V. Nahar, S. Unankard, X. Li, and C. Pang. "Sentiment analysis for effective detection of cyber bullying." In Asia-Pacific Web Conference, Springer, Berlin, Heidelberg, 2012, pp. 767-774.

[7] V. Nahar, X. Li, C. Pang, and Y. Zhang. "Cyberbullying detect ion based on text-stream classification." In The 11th Australasian Data Mining Conference (AusDM 2013), 2013.

[8] M. Dadvar, D. Trieschnigg, R. Ordelman, and F. de Jong. "Improving cyberbullying detection with user context." In European Conference on Information Retrieval, Springer, Berlin, Heidelberg, 2013, pp.

[9] V. Nahar, S. Al-Maskari, X. Li, and C. Pang. "Semi-supervised learning for cyberbullying detection in social networks." In Australasian Database Conference, Springer, Cham, 2014, pp. 160- 171.

[10] V. Nahar, X. Li, H. L. Zhang, and C. Pang. "Detecting cyberbullying in social networks using multi-agent system." Web Intelligence and Agent Systems: An International Journal 12, no. 4, 2014, pp. 375- 388.

[11] Q. Huang, V. K. Singh, and P. K. Atrey. "Cyberbullying detection using social and textual analysis." In Proceedings of the 3rd International Workshop on Socially-Aware Multimedia, ACM, 2014, pp. 3-6.

[12] A. Mangaonkar, A. Hayrapetian, and R. Raje. "Collaborative detection of cyberbullying behavior in Twitter data." In 2015 IEEE international conference on electro/information technology (EIT). IEEE, 2015, pp. 611-616.

[13] C. Van Hee, E. Lefever, B. Verhoeven, J. Mennes, B. Desmet, G. De Pauw, W. Daelemans, and V. Hoste. "Automatic detection and prevention of cyberbullying." In International Conference on Human and Social Analytics (HUSO 2015), IARIA, 2015, pp. 13-18.

[14] H. Hosseinmardi, S. A. Mattson, R. I. Rafiq, R. Han, Q. Lv, and S. Mishra. "Detection of cyberbullying incidents on the instagram social network." arXiv preprint arXiv:1503.03909, 2015.

[15] P. Galán-García, J. G. de la Puerta, C. L. Gómez, I. Santos, and P. G. Bringas. "Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying." Logic Journal of the IGPL 24, no. 1, 2016, pp. 42-53.