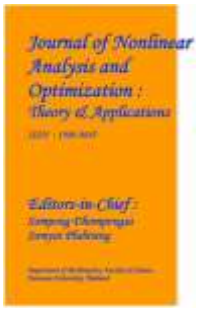
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 1: 2025
ISSN : **1906-9685**



MULTI-MODAL APPROACH FOR IDENTIFICATION AND DETECTION OF CYBERBULLYING IN SOCIAL NETWORKS

K.UDAY KIRAN[1], BELLAM HEMANTH KUMAR[2]

[1]Assistant Professor, Dept. of MCA, QIS College of Engineering and Technology, Ongole, Andhra Pradesh.

[2]PG Scholar, Dept. of MCA, QIS College of Engineering and Technology, Ongole, Andhra Pradesh.

ABSTRACT— Given the widespread use of social networks in people's everyday lives, cyberbullying has emerged as a major threat, especially affecting younger users on these platforms. This matter has generated significant societal apprehensions. Prior studies have primarily concentrated on analyzing text in relation to cyberbullying. However, the dynamic nature of cyberbullying covers many goals, communication platforms, and manifestations. Conventional text analysis approaches are not effective in dealing with the wide range of bullying data seen in social networks. In order to tackle this difficulty, our suggested multi-modal detection approach integrates data from diverse sources including photos, videos, comments, and temporal information from social networks. In addition to textual data, our approach employs hierarchical attention networks to record session features and encode various media information. The resulting multi-modal cyberbullying detection platform provides a comprehensive approach to address this emerging kind of cyberbullying. By conducting experimental analysis on two actual datasets, our framework exhibits greater performance in comparison to many state-of-the-art models. This highlights its effectiveness in dealing with the intricate nature of cyberbullying in social networks.

Index Terms – Cyberbullying, multi-modality, social media, hierarchy attention.

I. INTRODUCTION

Social media networks such as Facebook, Twitter, Flickr, and Instagram have become

the preferred online platforms for interaction and socialization among people of all ages. While these platforms enable people to communicate and interact in previously un thinkable ways, they have also led to malevolent activities such as cyber-bullying. Cyber bullying is a type of psychological abuse with a significant impact on society. Cyber-bullying events have been increasing mostly among young people spending most of their time navigating between different social media platforms. Particularly, social media networks such as Twitter and Facebook are prone to CB because of their popularity and the anonymity that the Internet provides to abusers. In India, for example, 14 percent of all harassment occurs on Facebook and Twitter, with 37 percent of these incidents involving youngsters.

Moreover, cyber bullying might lead to serious mental issues and adverse mental health effects. Most suicides are due to the anxiety, depression, stress, and social and emotional difficulties from cyber-bullying events. This motivates the need for an approach to identify cyber bullying in social media messages (e.g., posts, tweets, and comments). In this article, we mainly focus on the problem of cyber bullying detection on the Twitter platform. As cyber bullying is becoming a prevalent problem in Twitter, the detection of cyber bullying events from tweets and provisioning preventive measures are the primary tasks in battling cyber bullying threats. Therefore, there is a greater need to increase the research on social networks-based CB in order to get greater in sights and aid in the development of effective tools and approaches to effectively combat cyber bullying problem. Manually monitoring and controlling cyber bullying on Twitter platform is virtually impossible. Furthermore, mining social media messages for cyber bullying detection is quite difficult. For example, Twitter messages are often brief, full of slang, and may include emojis, and gifs, which makes it impossible to deduce individuals' intentions and meanings purely from social media messages. Moreover, bullying can be difficult to detect if the bully uses strategies like sarcasm or passive-aggressiveness to conceal it.

Despite the challenges that social media messages bring, cyber bullying detection on social media is an open and active research topic. Cyber bullying detection within the Twitter platform has largely been pursued through tweet classification and to a certain extent with topic modeling approaches. Text classification based on supervised machine

learning (ML) models are commonly used for classifying tweets into bullying and non-bullying tweets.

II. LITERATURE SURVEY

A. Cyber bullying Detection and Hate Speech Identification using Machine Learning Techniques

Bullying has been prevalent since the beginning of time, It's just the ways of bullying that have changed over the years, from physical bullying to cyberbullying. According to Williard (2004), there are eight types of cyberbullying such as harassment, denigration, impersonation, etc. It's been around 2 decades since social media sites came into the picture, but there haven't been a lot of effective measures to curb social bullying and it has become one of the alarming issues in recent times.Our paper presents an analytical review of cyberbullying detection approaches and assesses methods to recognize hate speech on social media. We aim to apply traditional supervised classification methods as well as some novel ensemble machine learning techniques using a manually annotated open-source dataset for this purpose. This paper does a comparative study of various Supervised algorithms, including standard, as well as ensemble methods. The evaluations of the result based upon the scores obtained by accuracy shows that Ensemble supervised methods have the potential to perform better than traditional supervised methods.

B. Cyber bullying detection on social media using machine learning

Usage of internet and social media backgrounds tends in the use of sending, receiving and posting of negative, harmful, false or mean content about another individual which thus means Cyberbullying. Bullying over social media also works the same as threatening, calumny, and chastising the individual. Cyberbullying has led to a severe increase in mental health problems, especially among the young generation. It has resulted in lower self-esteem, increased suicidal ideation. Unless some measure against cyberbullying is taken, self-esteem and mental health issues will affect an entire generation of young adults. Many of the traditional machine learning models have been implemented in the past for the automatic detection of cyberbullying on social media. But these models have not considered all the necessary features that can be used to identify or classify a statement or post as bullying. In this paper, we proposed a model based on various features that should be considered while detecting cyberbullying and

implement a few features with the help of a bidirectional deep learning model called BERT.

C. Aggression detection in social media from textual data using deep learning models

With the advancement of technology, social media such as Facebook, Twitter, etc. plays an important role in communication whether it is texting, sharing photos, audio-video calls or expressing views through comments. Along with these advantages, it has some negative sides as well which brings aggression towards some section of people. Such aggression, hatred in social media needs to be detected and prevented automatically which is the main objective of our work. We have worked on Hindi, English and Hindi-English (code-mixed) datasets. We used features like word vectors, aggressive words (manually created dictionary), sentiment scores, parts of speech and emojis for the classification task. We experimented with several machine learning and deep learning models and the results indicate that XGBoost Classifier, Gradient Boosting Classifier (GBM) and Support Vector Machine (SVM) are most suited for the task. Therefore the output of the three classifiers were used for majority voting which provides f-scores of 68.13, 54.82 and 55.31 for the English, Hindi and code-mixed datasets respectively.

III. PROPOSED SYSTEM

The overview of our proposed system is shown in the below figure.

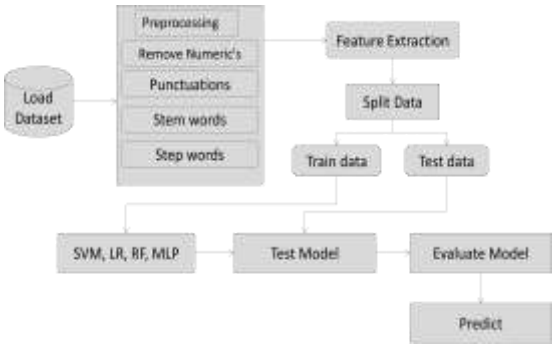


Fig. 1: System Overview

Implementation Modules

Service Provider

- In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Login, Train and Test Data Sets, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Cyber bullying Predict Type Details, Find Cyber bullying Prediction Ratio on Data Sets, Download Cyber Bullying Prediction Data Sets, View Cyber bullying Prediction Ratio Results, View All Remote Users.

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 admin authorizes the users.

Remote 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 REGISTER AND LOGIN, PREDICT CYBERBULLYING, VIEW YOUR PROFILE.

IV. RESULTS

Fig.2: Home Page

Fig.3: Service Provider Login

Fig.4: Model Accuracy

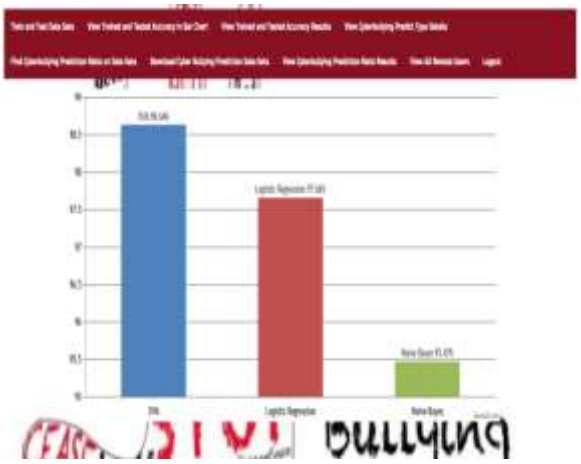


Fig.5: Model Accuracy Results

Fig.6: Predicted Results

V. CONCLUSION

This project developed an efficient tweet classification model to enhance the effectiveness of topic models for the detection of cyber-bullying events. DEA-RNN was developed by combining both the DEA optimization and the Elman type SVM, RF, and NB methods on a newly created Twitter dataset, which was extracted using CB keywords. The experimental analysis showed that the DEA-RNN had achieved optimal results compared to the other existing methods in all the scenarios with various metrics such as accuracy, recall, F-measure, precision, and specificity. This signifies the impact of DEA on the performance. Although the hybrid proposed model obtained higher performance rates than the other considered existing models, the feature compatibility of DEA-RNN reduces when the input data is increased greater than the initial input.

REFERENCES

[1] F. Mishna, M. Khoury-Kassabri, T. Gadalla, and J. Daciuk, ``Risk factorsfor involvement in cyber bullying: Victims, bullies and bullyvictims," Children Youth Services Rev., vol. 34, no. 1, pp. 6370, Jan. 2012, doi: 10.1016/j.childyouth.2011.08.032.

[2] K. Miller, ``Cyberbullying and its consequences: How cyberbullying is contorting the minds of victims and bullies alike, and the law's limited available redress," Southern California Interdiscipl. Law J., vol. 26, no. 2, p. 379, 2016.

[3] Mohammed Ali Shaik, Dhanraj Verma, (2021), “Agent-MB-DivClues: Multi Agent Mean based Divisive Clustering”, Ilkogretim Online - Elementary Education, Vol 20(5), pp. 5597-5603, doi:10.17051/ilkonline.2021.05.629

[4] H. Sampasa-Kanyinga, P. Roumeliotis, and H. Xu, ``Associations between cyberbullying and school bullying victimization and suicidal ideation, plan and attempts among Canadian schoolchildren," PLoS ONE, vol. 9, no.

7, Jul. 2014, Art. no. e102145, doi: 10.1371/journal.pone.0102145.

[5] Mohammed Ali Shaik, “Time Series Forecasting using Vector quantization”, International Journal of Advanced Science and Technology (IJAST), ISSN:2005-4238, Volume-29,Issue-4 (2020), Pp.169-175.

[6] A. S. Srinath, H. Johnson, G. G. Dagher, and M. Long, ``BullyNet: Unmasking cyberbullies on social networks," IEEE Trans. Computat. Social Syst., vol. 8, no. 2, pp. 332344, Apr. 2021, doi: 10.1109/TCSS.2021.3049232.

[7] A. Agarwal, A. S. Chivukula, M. H. Bhuyan, T. Jan, B. Narayan, and M. Prasad, ``Identication and classication of cyberbullying posts: A recurrent neural network approach using under-sampling and class weighting," in Neural Information Processing (Communications in Computer and Information Science), vol. 1333, H. Yang, K. Pasupa, A. C.-S. Leung, J. T. Kwok, J. H. Chan, and I. King, Eds. Cham, Switzerland: Springer, 2020, pp. 113120.

[8] M. A. Shaik, A. Rahim, V. Subhalakshmi, D. R. Ravi Kumar, R. Pasunuri and D. Verma, "Exploring Time Series Techniques in Production Function Modeling: ARIMA and VECM Applications," 2025 International Conference on Intelligent Computing and Control Systems (ICICCS), Erode, India, 2025, pp. 160-165, doi: 10.1109/ICICCS65191.2025.10985700

[9] N. Yuvaraj, K. Srihari, G. Dhiman, K. Somasundaram, A. Sharma, S. Rajeskannan, M. Soni, G. S. Gaba, M. A. AlZain, and M. Masud, ``Nature-inspired-based approach for automated cyberbullying classication on multimedia social networking," Math. Problems Eng., vol. 2021, pp. 112, Feb. 2021, doi: 10.1155/2021/6644652.

AUTHORS Profile

Mr. K.UDAY KIRAN has received his MCA from BAPATLA ENGINEE-RING COLLEGE, Bapatla in 2012 He is dedicated to teaching field from the last 2 years.His research area is Cloud Computing. At present he is working as Assistant Professor inMCA Department in QIS College of Engineering & Technology, Ongole, Andhra Pradesh, India.

Mr. BELLAM HEMAN-TH KUMAR has receiv-ed him B.C.A (Bachelor of

Computer Applications) and degree from ANU 2023 and pursuing MCA in QIS College of Engineering and Technology affiliated to JNTUK in 2023-2025.