International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7018
***Article Received**: 01 April 2023 **Revised:** 22 May 2023 **Accepted:** 08 June 2023*

Classification of Cyberbullying Detection in Social Networking with Audio using Machine Learning Approach

Bandari Saichandana[1], Dr. Pille Kamakshi[2]
[1]Dept. of Information Technology, Kakatiya Institute of Technology & Science
Warangal, India
M21ds001@kitsw.ac.in
[2]Dept. of Information Technology, Kakatiya Institute of Technology & Science
Warangal, India
pk.it@kitsw.ac.in

Abstract— Every day, more people use the internet and social media, which leads to an increase in cyberbullying vulnerabilities. By transmitting, posting, and spreading damaging, false, and bad stuff online, it is taking place. For those impacted, it causes psychological and emotional issues. Therefore, the development of automated tools for cyberbullying identification and prevention is essential. Recent research on identifying cyberbullying has largely focused on text-based analysis. The two most significant media in cases of cyberbullying are text and audio. In this paper, a machine learning model for detecting cyberbullying in two types of social data, namely text and audio is presented. This paper is focused on detecting majors form of Cyberbullying: cyberbullying dataset on Twitter and classify them as containing Cyberbullying or not. This paper used datasets, namely, 'cyberbullying Dataset'. In this paper, the implementation can be done with machine learning algorithms such as logistic regression, naive bayesian classifier, and support vector machine. Also, these three algorithms were compared and evaluated with performance metrics like accuracy, precision, recall, and f1 score. The main aim is to detecting cyberbullying messages on any type of social media platform.

Keywords- Cyberbullying; machine learning; support vector machine; naïve bayesian classifier; logistic regression; Twitter; social media platform.

I. INTRODUCTION

Among the millions of young people who frequent social networking sites, online information exchange is common. Social networks enable communication and information sharing with anybody, at any time, and with a large group of individuals all at once. Globally, there are more than 3 billion users of social media. The National Crime Security Council defines cyberbullying as the intentional harm or public humiliation of another person while utilizing a mobile device, a video game app, or any other method to communicate or share text, audio, photos, or videos online (NCPC). Every day of the week, at any hour, anyone can be the victim of cyberbullying. Cyberbullying can take the form of text, audio, images, or video that is posted in an anonymous way. Finding the author of this post is sometimes impossible and can be complicated. Additionally, it was impossible to delete these communications later. The most frequent bullying websites on the internet are Twitter, Instagram, Facebook, YouTube, Snapchat, Skype, and a number of social media platforms. There are certain social networking sites, like Facebook, that offer advice on how to stop bullying. It includes a section specifically explaining how to report cyberbullying and avoid user blocking. When someone posts something on Instagram that the user finds uncomfortable, they might be blocked or monitored. Users can also recommend changes to the app and report violations of our community.[1]

By the use of multi-mode input sources including text, photographs, and videos, cyberbullying can take many different forms. Most study studies have mostly focused on evaluating textual content, such as comments and text messages, because at first, cyberbullying was unstructured and the usage medium was purely text. Emoji, memes, text, and image characters are still used in communication, making it difficult to spot cyberbullying. Bullying incidents are now a well-organized, multi-media data source. Websites for social media networking place a strong focus on photo sharing. As a result of these tendencies, cyberbullying behaviours among victims switch from text to visual.[2]

Analysis and overview of the context of cyberbullying are required because social lifestyles now include unrestrained communication with strangers and go beyond the physical boundaries of human connection. The victim of cyberbullying feels constantly attacked since he has access to the internet at

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7018
***Article Received**: 01 April 2023 **Revised:** 22 May 2023 **Accepted:** 08 June 2023*

all times. The victim could experience physical, mental, and emotional effects. Social media is used in text- or image-based cyberbullying. A system can respond appropriately if bullying text can be discriminated from non-bullying communication. Social media platforms and other messaging services may benefit from an effective cyberbullying detection system in thwarting such attacks and lowering the prevalence of cyberbullying. The cyberbullying detection system seeks to recognise the cyberbullying text and take into account its meaning. One must first assess a text's many elements before using previous knowledge or images to deduce the context. A customised system that can quickly and simply extract such a text must be developed.[3]

II. REVIEW OF LITERATURE

M. Di Capua, et al. [4], suggested developing a cyberbullying model unsupervised by combining a variety of variables, including both normal textual elements and some "social features." Syntactic, semantic, emotive, and social aspects were split into four groups to represent the qualities. The input layer for the author's neural network was a Growing Hierarchical Self-Organizing Map (GHSOM) network with a grid of 50 50 neurons and 20 features. M. Di Capua et al. categorized the input dataset and GHSOM on the Forms pring dataset using the clustering method K-means. The findings of this unsupervised combination technique performed better than the preceding ones. Then, using the YouTube dataset, the author tested three different machine learning models: the Naive Bayes Classifier, the Decision Tree Classifier (C4.5), and the SVM with a Linear Kernel. The clustering results for the hate posts in the YouTube dataset were found to have lesser precision than the Form Spring testing because textual analysis and syntactical features function differently on each side. For the Twitter dataset, this hybrid approach achieved a subpar recall and F1 Score. It is possible to enhance the authors' methodology and use it to create beneficial applications that address the problem of cyberbullying.

J. Yadav, et al.[5], proposes A brand-new method that uses a single layer of a linear neural network as a classifier and combines the BERT model to detect cyberbullying on social media platforms. Using datasets from Wikipedia and Formspring forums, the model is trained and tested. When compared to the previously employed models, the suggested model's performance accuracy was 98% for the Form spring dataset and 96% for the Wikipedia dataset. The suggested strategy performed better than it did for the Form spring dataset given the size of the Wikipedia dataset without the requirement for oversampling.

R. R. Dalvi, et al.[6], proposed A method that uses machine learning methods for supervised categorization to keep an eye on Twitter for Internet exploitation and stop it. The live Twitter API is used in this study to gather tweets and build datasets. Using the collected datasets, the suggested model contrasts Support Vector Machine with Naive Bayes. They employed the TFIDF vectorizer to extract the feature. The findings indicate that while the accuracy of the Support Vector Machine-based cyberbullying model is close to 71.25%, that of the Naive Bayes model was close to 52.75%.

N. Tsapatsoulis, et al. [7], proposed On Twitter, cyberbullying is thoroughly examined. It emphasises how important it is to track down different Twitter violators. The paper gives a thorough explanation of a variety of practical procedures required to develop a programme that can rapidly and accurately detect cyberbullying. The case studies that made use of these technologies are explained, along with the patterns in the tagging and categorization of data platforms, machine learning models, and feature types. The initiative for applying machine learning to detect cyberbullying will start with this paper.

Natarajan Yuvaraj et al. [8] suggested a fresh classification model for Twitter data-based CB detection. For tweet classification, deep decision-tree classification using multi-feature-based AI was utilised. The decision tree classifier was combined with the deep neural network's hidden layers to create the deep decision tree classifier. Three feature selection methods were also used in this strategy: IG, Pearson Correlation, and Chi-Square. High-dimensional data, however, cannot be handled by it with the same accuracy.

III. PROPOSED METHOD

In recent years, numerous works linked to the detection of cyberbullying have been proposed but do not focus on audio. Because just one medium may identify the victim's motivation for engaging in cyberbullying in a meritorious manner in social media, our research focuses on both audio and text-based cyberbullying detection.

The major goal is to identify cyberbullying posts and cyberbullying information contained in a message on any social media platform.

The first audio dataset to be converted to text using Django. Django app allows users to convert their speech into text and send that text as a message and then predict post message working with both media like text and audio. This app record blobs in real time to the server! After every 10 seconds recorded blob is converted into text and send as a message to chat.

There are two following modules in this project.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7018
***Article Received**: 01 April 2023 **Revised:** 22 May 2023 **Accepted:** 08 June 2023*

User module: using this module users can create an account. Using account details they can login to the application and then send and view posts.

Administrator Module: The administrator can view all users who have registered and can then approve or disapprove new users. The admin is in charge of updating the machine learning training dataset with new cyberbullying and non-cyberbullying communications. To execute cyberbullying message identification from the user side, the administrator must run all or at least one SVM algorithm. All posts sent by all users can be viewed or tracked by the admin.

Create a database first by copying and pasting the contents of the "DB.txt" file into the MYSQL console. Install DJANGO, deploy the "Cyber" folder, start the server, and launch the browser by typing the URL "http://127.0.0.1:8000/index.html" to run this project. The below-screen URL will be displayed after execution.

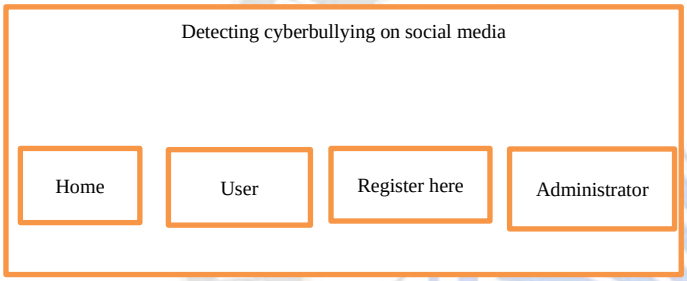

Figure 1 : Home page

Step 1: To establish an account, click the "Register Here" link on the previous screen and add the new user.

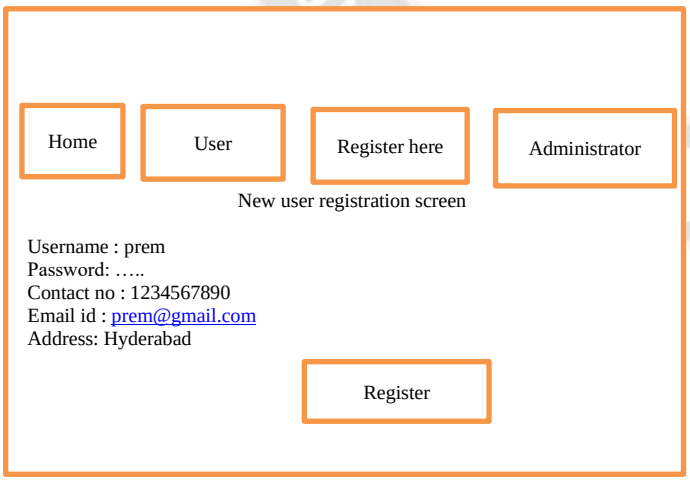

Figure 2: New user registration screen

Step 2: Click the "Register" button now to add details to our website.

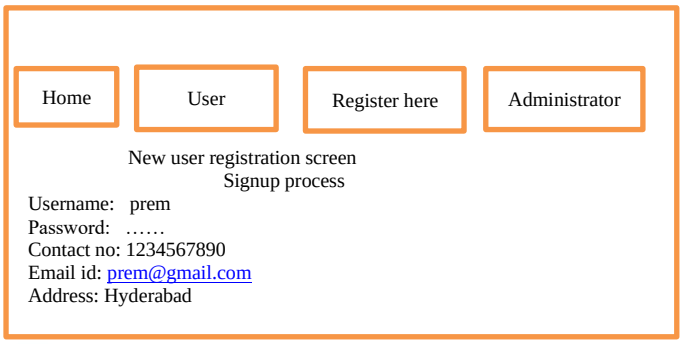

Figure 3: User registration process completed.

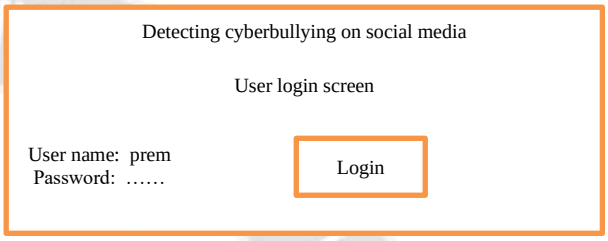

Figure 4: User login screen

Step 3: In the above screen, sign-up process is completed. prem user is logged in and will see the screen below after.

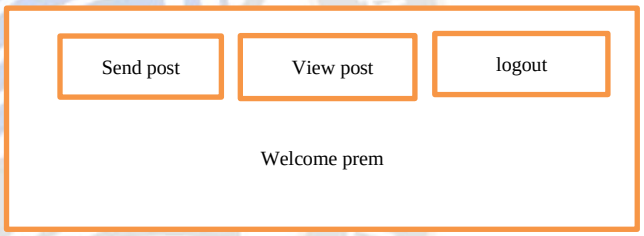

Figure 5: User home page

Step 4 : To view the screen below, click the “Submit Post” option in the below screen.

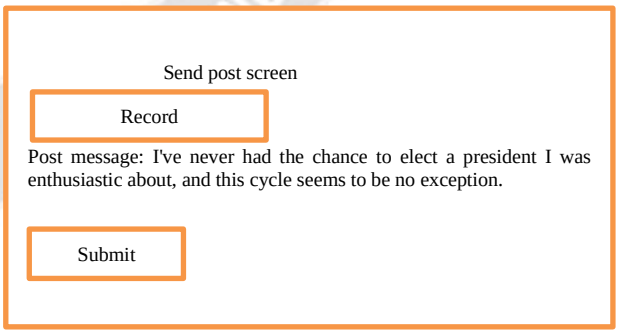

Figure 6 : Send post screen page

Step 5 : After clicking "submit," I added a few messages that included text and audio to the post in the above page and sent them to a friend account. The record option is used to record messages and convert them to text, which is present in post messages.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7018
***Article Received**: 01 April 2023 **Revised:** 22 May 2023 **Accepted:** 08 June 2023*

Step 6 : Now click on the 'Administrator' link to login as an admin and view new user details.

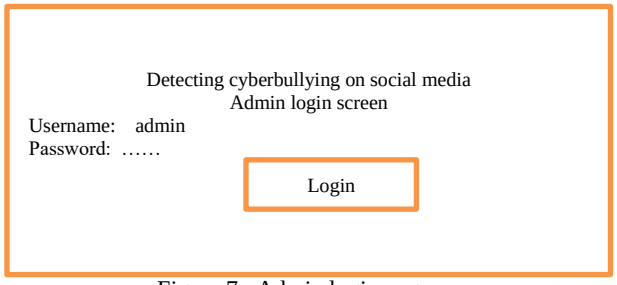

Figure 7: Admin login page

Step 7 : Enter "admin" as the username and "admin" as the password to log in as "admin" on the screen above. After login will get below screen.

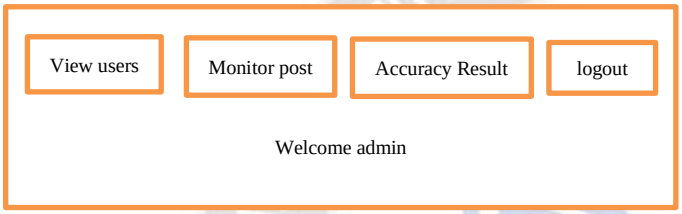

Figure 8: Admin home page

Step 8 : Now admin can click on 'View Users' link to view all user's list.

View all remote user

User name	password	Contact no	Email id	Address
prem	……	1234567890	prem@gmail.com	Hyderabad
chandu	……	2244556677	chandu@gmail.com	Khammam

Figure 9 : View all remote user

Step 9 : In above screen, we can see 'prem' account created. Now admin can click on 'Monitor Post' to view all posts from past users.

Sender name	Message	Status
Chandu	The issue is that there is just one police officer. You need more, more powerful weapons, maybe even a tank with soldiers.	Non-cyberbullying
Prem	White countries for everyone, Asia for the Asians, and Africa for the Africans are all code words for anti-white racism.	cyberbullying

Figure 10: Monitor post screen

Step 10 : In above screen application will automatically detect whether message is non-cyberbullying or cyberbullying from machine learning algorithms.

A. Classification

In this study, we forecast cyberbullying posts from social networks using a variety of machine learning methods, including Support vector machines, Naive Bayes, and logistic regression.

B. SVM: Support Vector Machine

In a (N)-dimensional space with many different properties, this theorem is essentially used to print a hyperplane that defines a boundary between data points. The margin value hinge function is one of the better loss functions for this. The following scenario makes use of linear SVM and is excellent for data that can be separated linearly. If there is no misclassification, which means that our model correctly predicted the class of the data point, we merely need to adjust the gradient based on the regularisation arguments [9]

C. Naive Bayes

This powerful machine learning method is built on the Bayes theorem [10]. The programme bases its predictions on how likely an object is to occur. The challenges of binary and multi-class classification can be quickly overcome using this technique. According to the Bayes' Theorem, the likelihood of one event happening given the likelihood that another event has already happened is as follows:

D. Logistic regression

One of the Machine Learning algorithms that is most frequently used in the Supervised Learning category is logistic regression. With a predetermined set of independent factors, it is used to forecast the categorical dependent variable. A categorical dependent variable's output can be predicted using logistic regression. As a result, the result must be a discrete or categorical value. Instead of the precise numbers between 0 and 1, it delivers the probabilistic values that fall between 0 and 1. Either True or False, 0 or 1, or Yes or No, are possible outcomes.

E. Usage of the cyberbullying Prediction System

we are seeing posts from all users and we can see with the help of machine learning SVM algorithms we are automatically able to predict messages as cyberbullying or non-cyberbullying. Similarly, you can sign up new users and then post new messages.

Machine learning will predict cyberbullying or non-cyberbullying message based on dataset records so you add all possible cyberbullying and non-cyberbullying message to

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7018
***Article Received**: 01 April 2023 **Revised:** 22 May 2023 **Accepted:** 08 June 2023*

dataset by using the 'adding dataset' module from admin. After adding dataset then run algorithms link to train model and then you can see the application will predict cyberbullying or non-cyberbullying automatically.

Sender name	Message	Status
Raj	it isn't to target illegals and/or criminals. it's just plain common sense.	Cyberbullying
Veeru	no one but you can make you #quote	Non-cyberbullying
Ashu	is not an opinion. racism is degrading people based on race to justify discrimination or open violence against them. Stop racism	cyberbullying

Figure 11 : Cyberbullying prediction result

The technique helped to identify cyberbullying message, and if a post was accidentally shared with another individual, the user could block the account by admin directly.

In above screen, the application will automatically detect whether the message is non-cyberbullying or cyberbullying from machine learning algorithms.

IV. Experiment and Result

To categorise comments as bullying or non-bullying, We used Naive Bayes (NB), Support Vector Machines (SVM), and logistic regression as our three machine learning techniques (LR). The datasets for the experiment are initially described in this part, followed by a discussion of the results.

A. Dataset

In this part, we explicitly clarify the dataset that was used in this paper, the "cyberbullying Dataset." Essentially, it consists of 2 columns and 31,962 rows. A unique comment and the categories to which it belongs are included in each row. There are several categories for the comment in the dataset, including "cyberbullying" and "non-cyberbullying."

B. Exploratory Analysis

Exploratory analysis involves charting a dataset for analysis in order to better understand its structure. To display the numerous comments that fit into several categories, a bar graph is created.

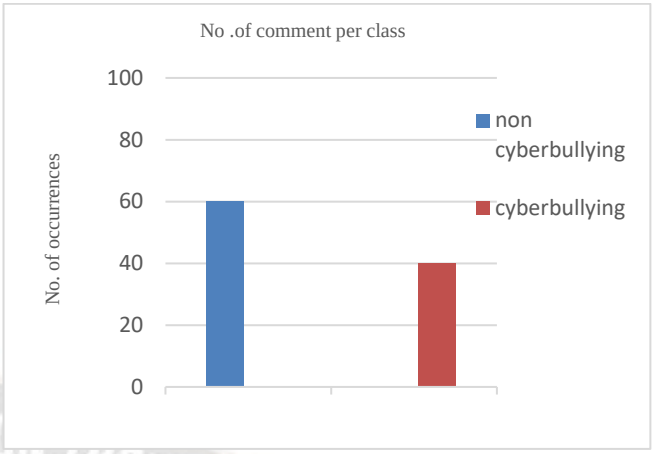

Figure 12: Number of Comments Per Class in a Bar Chart.

Figure 12: Shows that there are 12,762 cyberbullying messages and 19,200 non-cyberbullying messages.

C. Data Preparation

Data preparation is a procedure that involves converting unprocessed data into a format that can be used by other application models for machine learning. To improve the effectiveness of the machine learning models, the dataset's comments must be cleaned up because they contain characters that are not necessary for detecting whether a message is related to cyberbullying. To clean up the raw data, several pre-processing processes have been taken. They consist of removing letters and digits together, substituting white spaces for punctuation, changing comments to lowercase, removing escape sequences and substituting white spaces, removing non-ASCII characters, removing stop words, etc. [11]

D. Result for dataset

We used a tweet from Twitter with trace of bullying and applied it to our model. Table I shows the classification report based on our model. Here labels 0 and 1 represent non-cyberbullying and cyberbullying respectively.

Table II represents the accuracy of the support vector machine is 98.31%, naïve bayes is 95.16%, and logistic regression is 97.22%.

TABLE **I:** Result of cyberbullying prediction model

Algorithm	Label	Recall	Precision	F1-score
Support Vector machine	0	0.97	1.00	0.98
	1	1.00	0.97	0.98
Logistic regression	0	0.96	0.99	0.97
	1	0.99	0.96	0.97

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7018
***Article Received**: 01 April 2023 **Revised:** 22 May 2023 **Accepted:** 08 June 2023*

Naïve bayes	0	0.92	0.98	0.95
	1	0.98	0.92	0.95

The highest precision algorithms are support vector machine. The algorithms with the highest Recall are support vector machine. The algorithms with the highest F- measure are support vector machine exhibit high values for all performance metrics compared with all other algorithms.

TABLE II: Result of accuracy

Algorithm	Accuracy
Support vector machine	98.31%
Logistic regression	97.22%
Naïve bayes	95.16%

From the above resultant table II, the highest accuracy algorithms are support vector machine.

V. CONCLUSION AND FUTURE SCOPE

There is a need to restrict the growth of cyberbullying because it can be harmful and result in unfortunate events like depression and suicide. A technique for predicting cyberbullying based on a machine learning algorithm was presented in this paper. It was possible to develop forecasts for individual cases of cyberbullying by comparing the profiles of existing accounts contain with cyberbullying information. According to the result, the Support vector machine provided the highest, at 98.31%, while logistic regression and naïve bayes provided the lowest accuracy, at 97.22% and 95.16%. The textual content of tweets and audio is used in the proposed model to identify cyberbullying.

In the future, we plan to put our suggested method although other types of media, such as photos and video, are still under investigation.

REFERENCE

[1] Aditya Desai1 , Shashank Kalaskar2 , Omkar Kumbhar3 , and Rashmi Dhumal4 , Cyber Bullying Detection on Social Media using Machine Learning ITM Web of Conferences 40, 03038 (2021) ICACC-2021

[2] Dr. Vijayakumar V*1 , Dr Hari Prasad D2 , Adolf P3 Multimodal Cyberbullying Detection using Hybrid Deep Learning Algorithms International Journal of Applied Engineering Research ISSN 0973-4562 Volume 16, Number 7 (2021) pp. 568-574 © Research India Publications.

[3] Aditya Desai1 , Shashank Kalaskar2 , Omkar Kumbhar3 , and Rashmi Dhumal4 , Cyber Bullying Detection on Social Media using Machine Learning ITM Web of Conferences 40, 03038 (2021) ICACC-2021

[4] Mr. Dharmesh Dhabliya, Dr.S.A.Sivakumar. (2019). Analysis and Design of Universal Shift Register Using Pulsed Latches . International Journal of New Practices in Management and Engineering, 8(03), 10 - 16. https://doi.org/10.17762/ijnpme.v8i03.78

[5] M. Di Capua, E. Di Nardo and A. Petrosino, Unsupervised cyberbullying detection in social networks, ICPR, pp. 432-437, doi: 10.1109/ICPR.2016.7899672. (2016)

[6] J. Yadav, D. Kumar and D. Chauhan, Cyberbullying Detection using Pre-Trained BERT Model, ICESC, pp. 1096-1100, doi: 10.1109/ICESC48915.2020.9155700. (2020)

[7] R. R. Dalvi, S. Baliram Chavan and A. Halbe, Detecting A Twitter Cyberbullying Using Machine Learning, ICICCS, pp. 297-301, doi: 10.1109/ICICCS48265.2020.9120893. (2020)

[8] Annam, S. ., & Singla, A. . (2023). Estimating the Concentration of Soil Heavy Metals in Agricultural Areas from AVIRIS Hyperspectral Imagery. International Journal of Intelligent Systems and Applications in Engineering, 11(2s), 156 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2519

[9] N. Tsapatsoulis and V. Anastasopoulou, Cyberbullies in Twitter: A focused review, SMAP, pp. 1-6, doi: 10.1109/SMAP.2019.8864918. (2019)

[10] Mwangi, J., Cohen, D., Costa, R., Min-ji, K., & Suzuki, H. Optimizing Neural Network Architecture for Time Series Forecasting. Kuwait Journal of Machine Learning, 1(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/132

[11] N. Yuvaraj, V. Chang, B. Gobinathan, A. Pinagapani, S. Kannan, G. Dhiman, and A. R. Rajan, ''Automatic detection of cyberbullying using multi-feature based artificial intelligence with deep decision tree classification,'' Comput. Electr. Eng., vol. 92, Jun. 2021, Art. no. 107186, doi: 10.1016/j.compeleceng.2021.107186.

[12] Varun Jain1 Vishant Kumar 2 Dinesh Kumar Vishwakarma3 Department of Vivek Pal4 Detection of Cyberbullying on Social Media Using Machine learning Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021) IEEE Xplore Part Number: CFP21K25-ART

[13] Rish et al., "An empirical study of the naive bayes classifier," IJCAI2001 workshop on empirical methods in artificial intelligence, vol. 3,no. 22, pp. 41–46, 2001.

[14] Evangeline D1 , Amy S Vadakkan2 , Sachin R S3 , Aakifha Khateeb4 , Bhaskar C5 Cyberbullying Detection in Textual Modality (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 12, 2021

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7018
***Article Received**: 01 April 2023 **Revised:** 22 May 2023 **Accepted:** 08 June 2023*

Author's Profile:

B. Saichandana pursuing M. Tech in Data Science in the Department of Information Technology, Kakatiya Institute of Technology & Science, Warangal, Telangana, India. Her research interest includes Machine learning, Deep learning.

Dr.P.Kamakshi received the Ph. D in Computer Science and Engineering from Jawaharlal Nehru Technological University, Hyderabad. She is currently working as Professor in the Department of Information Technology, Kakatiya Institute of Technology & Science, Warangal, Telangana, India. Research interest is in the area of privacy preserving data mining.