Cyberbully Detection On Social Media

Dr. P.Nagaraj[1]*, Dasugari Tharun Goud[2], G Anirudh Goud[3], Mandaloju Shravya[4], Pisati Hima Reddy[5]

Abstract
The use of derogatory and combative language has dramatically expanded in the social media and networking era. Young people are largely responsible for it. More than half of the young people who use social media are victims of cyber bullying. Insults in social media websites create negative interactions within the network. These comments foster a disrespectful atmosphere in internet. Online harassment, including the dissemination of private chats, rumours, and sexual insults, has recently been the cause of numerous instances all around the world. As a result, academics are paying more and more attention to the detection of bullying text or message on social media. In order to classify such comments in a practical way, the study aims to uncover techniques to recognise bullying in text by analysing and experimenting with various methodologies. The goal of this research is to combine machine learning and natural language processing to create a powerful method for identifying online bullying and abusive messages. In order to discover bullying tests and hostile comments, we devised an effective algorithm, and we analysed these comments to ensure their validity.

Keywords: Cyber bullying, cyberspace, harassment, Machine Learning.

[1]*Associate Professor, Dept of CSE, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad.
[2,3,4,5]B.Tech students, Dept of CSE , Sreyas Institute of Engineering and Technology, Nagole, Hyderabad.

***Corresponding Author:** Dr. P. Nagaraj
*Associate Professor, Dept of CSE, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad.

I. INTRODUCTION

Social media is a platform that enables users to engage with others and share a wide variety of content, including images, videos, and documents. People use their computers or smart phones to access social media. Facebook1, Twitter2, Instagram3, TikTok4, and other social media platforms are among the most widely used. Social media is now used in a variety of fields, including business, education, and good causes. Social networking is boosting the global economy by opening up a lot of new job prospects. Social media offers many advantages, but it also has certain disadvantages. Through the usage of these platforms, malicious persons commit unethical and dishonest behaviors in order to offend other people and harm their reputations. Cyber bullying has emerged as one of the most pressing social media problems recently. An electronic form of bullying or harassment is referred to as cyberbullying or cyber-harassment. Online bullying also refers to cyberbullying and cyber-harassment. Cyberbullying has increased in prevalence throughout time, especially among young people as the digital world and technology have expanded.

Project Aim

The cyberbullying detection system's goal is to recognise the cyberbullying text while also taking its context into account. To determine the context of a text, one must first analyze its numerous components before using the prior knowledge or images. The goal of this research is to combine machine learning and natural language processing algorithms to provide a powerful method for identifying online bullying and abusive messages.

Problems with Existing System

Cyberbullying affects over 50% of kids in America. The victim of this bullying suffers both psychologically and physically. The pain of cyberbullying, which is difficult to sustain, leads the victims to pursue self-destructive behaviours like suicide. Thus, it is crucial to recognize and stop cyberbullying if we want to protect teenagers. Decision tree techniques are used in the current machine learning application for cyberbullying detection, but this algorithm is ineffective in categorizing the cyberbullying texts. Decision tree techniques are used in the current machine learning application for cyberbullying detection, but this algorithm is ineffective in categorising the cyberbullying texts.

1. An unstable decision tree might result from a little change in the data that has a huge impact on the decision tree's structure.
2. In comparison to other algorithms, a decision tree's calculations may become far more complicated.
3. Training a decision tree typically takes more time.
4. Decision tree training is relatively expensive as the complexity and time has taken aremore
5. Regression and the prediction of continuous values cannot be done effectively with the Decision Tree algorithm.

II. PROBLEM STATEMENT

social media Networks offer us fantastic options for communication, but they also make young people more susceptible to dangerous circumstances online. Because so many people use social media, cyberbullying on it is a worldwide phenomenon. Cyber bullying on social networks is on the rise, according to the trend. Cyberbullying is becoming a bigger issue among young people, according to recent studies. The appropriate detection of potentially hazardous messages is essential for effective prevention, and the Internet's information deluge necessitates intelligent systems that can swiftly identify possible threats. As a result, the main goal of this research is to develop a model that can automatically identify cyberbullying in social media text by simulating communications that are the result of social media bullying.

III. PROPOSED SYSTEM

The cyberbullying detection framework, which has two main components as depicted in this section, is described here. 1. The first component is known as "Natural Language Processing," and the second component is known as "Machine Learning." First, datasets including bullying texts, messages, or posts are gathered and prepared using natural language processing for machine learning

techniques. The machine learning algorithms are then trained to recognise any harassing or bullying messages on social media, like Facebook and Twitter, using the processed datasets. Multiple decision tree classifiers make up the classifier known as Random Forest. Individual class predictions are provided for each tree. Our ultimate finding is the maximum number of the projected class. This classifier is a supervised learning model that yields correct results because the output is created by combining numerous decision trees. The random forest makes decisions about the ultimate result based on the majority votes of predictions, rather than depending just on one decision tree.

IV. METHODOLOGY

Natural Language processing: The real world posts or text contain various unnecessary characters or text. For example, numbers or punctuation is irrelevant to bullying detection. Before applying the machine learning algorithms to the comments, we need to clean and prepared them for the detection phase.

In this phase, various processing task including removal of all irrelevant characters like stop-words, punctuation and numbers, tokenization's, stemming etc. After the preprocessing, we prepare the two important features of the texts as follows:

1) **Bag-of-Word:** The machine learning algorithms cannot work directly with the raw text. Sobefore applying the algorithms,
2) We must convert them to vectors or numbers. So, the processed data is converted to Bag-of-Words (BoW) for the next phase.
3) **TF-IDF:** This is another feature that we consider for our model. TF-IDF (Term Frequency- Inverse Document Frequency) is a statistical measure that can evaluate how relevant a word isto a document in a collection of documents. In bag of words, every word is given equal importance while in TF-IDF the words that occur more frequently should be given more importance as they are more useful for classification.

- **Machine Learning:** This module involves in applying various machine learning approaches Random Forest to detect the bullying message and text. The classifier with the highest accuracyis discovered for a particular public cyberbullying dataset.

DESIGN ARCHITECTURE

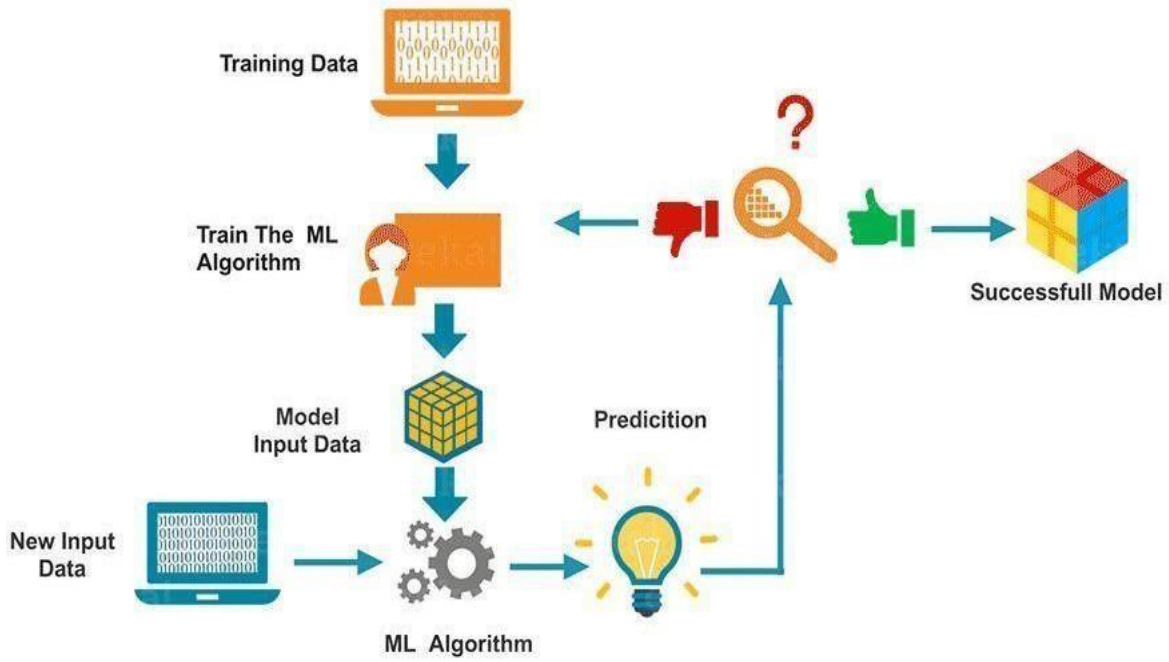

Fig 4.1: Architecture *of proposed Methodology*

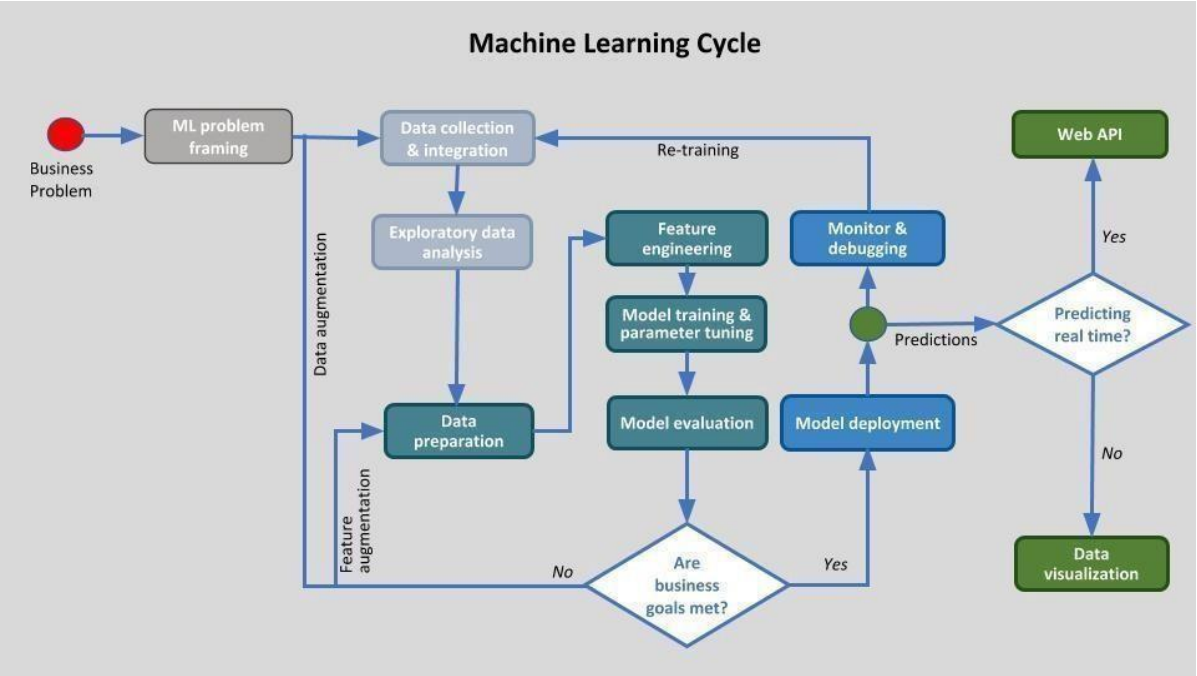

Fig 4.2: Technical Architecture

V. RESULTS

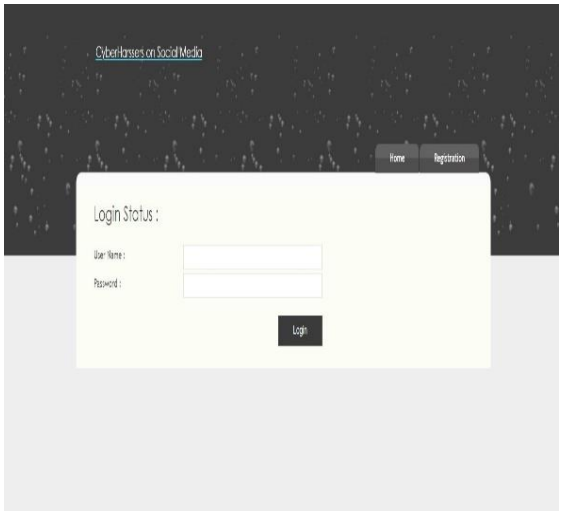

Fig 5.1: Login Page

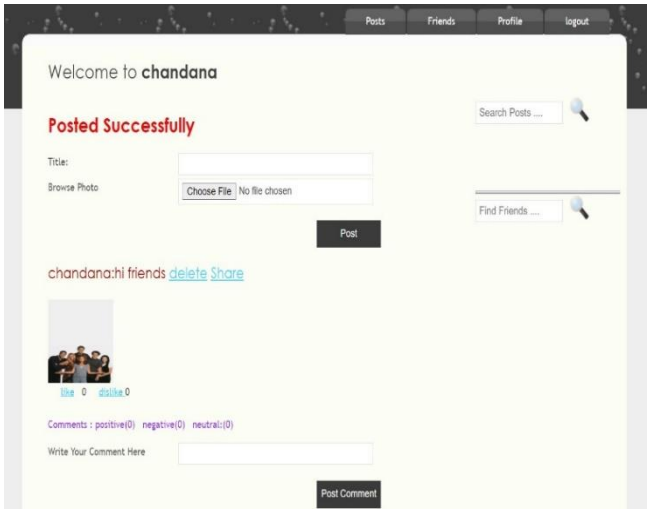

Fig 5.2: Picture Posted Successfully

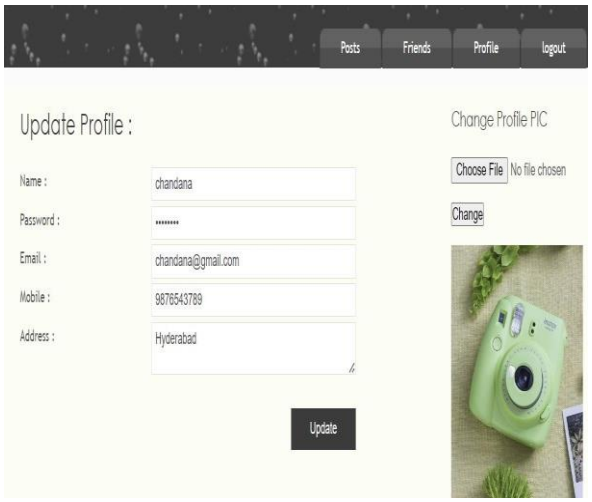

Fig 5.3: Update Profile

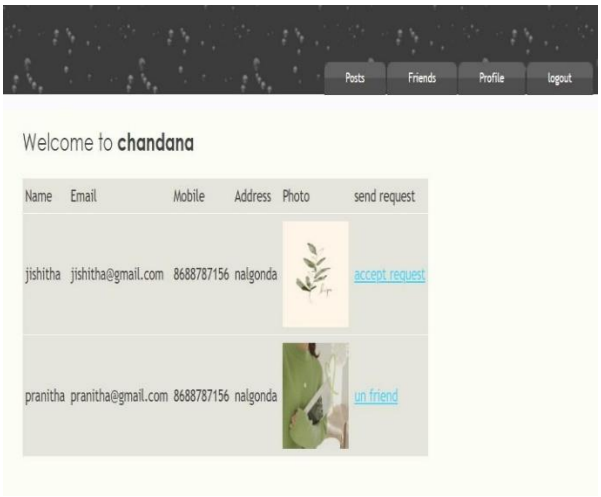

Fig 5.4: Friend Request

VI. CONCLUSION

With the growing popularity of social media

sites and the rise in teen social media use, in particular, cyberbullying has increased in frequency and has started to cause serious social problems. To prevent negative effects from cyberbullying, a system for automatically detecting it must be developed. Given the importance of cyberbullying detection, we looked into the automated detection of social media posts that were related to cyberbullying by taking into account two traits, BoW and TF-IDF. Four machine learning algorithms are used to identify bullying text and SVM for both BoW and TF-IDF. In future we are planningto design a framework for automatic detection and classification of cyberbullying from Bengalitexts using deep learning algorithms.

VII. FUTURE WORK

Improving the model assessment metrics by combining various machine learning models utilizing ensemble learning, such as the Random Forest classifier with the TF-IDF word vectors and the Support Vector Machine with our unique word vectors. generating more precise bespoke word vectors to train the SVM generating more intricate bespoke word vectors to train the SVM. As we all know, social media has become a popular venue for people to express their opinions, but there aren't many kids who engage in bullying on these types of platforms, which can be upsetting to someone's mental and emotional health. So, we created our project to identify cyberbullying using machine learning in order to stop this kind of taunting or bullying. Cyberbullying and other harmful social media behaviors must be identified early on in order to recognize threatening online aberrations and stop them from spreading. The comments from the subeditor were successfully downloaded for this project using paw, and I was also able to detect the obscene remarks using three machine learning algorithms: bag of words, term frequency, inverse document frequency, and support vector machine.

VIII. REFERENCES

W. Akram and R. Kumar, “A study on positive and negative effects of social media on society,” International Journal of Computer Sciences and Engineering, vol. 5, no. 10, pp. 351–354, 2017.

D. Tapscott et al., The digital economy. McGraw-Hill Education,, 2015.

S. Bastiaensens, H. Vandebosch, K. Poels, K. Van Cleemput, A. Desmet, and I. De Bourdeaudhuij, “Cyberbullying on social network sites. an experimental study into bystanders’behavioural intentions to help the victim or reinforce the bully,” Computers in Human Behavior, vol. 31, pp. 259–271, 2014.

D. L. Hoff and S. N. Mitchell, “Cyberbullying: Causes, effects, and remedies,” Journal of Educational Administration, 2009.

S. Hinduja and J. W. Patchin, “Bullying, cyberbullying, and suicide,” Archives of suicide research, vol. 14, no. 3, pp. 206–221, 2010.

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, vol. 2, pp. 1–7, 2009.

K. Dinakar, R. Reichart, and H. Lieberman, “Modeling the detection of textual cyberbullying,”in In Proceedings of the Social Mobile Web. Citeseer, 2011.

K. Dinakar, R. Reichart, and H. Lieberman, “Modeling the detection of textual cyberbullying.” in The Social Mobile Web, 2011.

P. Nagaraj, Dr A. V. Krishna Prasad, Dr. M. Venkat Dass, Kallepalli Rohit Kumar, “Swine Flu Hotspot Prediction In Regions Based on Dynamic Hotspot Detection Algorithm” Journal of Theoretical and Applied Information Technology(JATIT), 30 th Nov-2022, ISSN:1992-8645, Vol-100, N0.22, Pages-6535 to 6544.

Nagaraj, Dr. A. V. Krishna Prasad, Dr. V. B. Narsimha, Dr. B. Sujatha “ A swine flu Detection and Location using Machine Learning Techniques and GIS”, International Journal of Advanced Computer Science and Application (IJACSA),Vol-13,No.9, 2022.Pages 1001 to 1009.

P. Nagaraj , Rajesh Banala and A. V. Krishna Prasad, “Real Time Face Recognition using Effective Supervised Machine Learning Algorithms”, Journal of Physics: Conference Series 1998 (2021)

012007 IOP Publishing
doi:10.1088/1742-6596/1998/1/012007

P. Nagaraj, Dr. M. Venkat Dass, E. Mahender "Breast Cancer Risk Detection Using XGB Classification Machine Learning Technique", IEEE International Conference on Current Development in Engineering and Technology (CCET)-2022,Sage university, Bhopal, India, 23-24,Dec 2022.

P. Nagaraj, Gunta Sherly Phebe, Anupam Singh, "A Novel Technique to Classify Face Mask for Human Safety", 2021 Sixth ICIIP Published in: 2021 Sixth International Conference on Image Information Processing (ICIIP),26-28 Nov. 2021, 10 February 2022 DOI: 10.1109/ICIIP53038.2021.9702607 Publisher: IEEE Conference Location: Shimla, India

P. Nagaraj and Dr A. V. Krishna Prasad, "A Novel Technique to Detect the Hotspots Swine Flu Effected Regions", Published in: 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO),15 November 2021 DOI:10.1109/ICRITO51393.2021.9596422, Electronic ISBN:978-1-6654-1703-7 CD:978-1-6654-1702-0.