CYBERBULLYING IN SOCIAL NETWORKING SITES USING DEEP LEARNING MODEL

E.Elanchezhiyan.M.E.,
Assistant Professor
Department of Computer science and Engineering
Paavai Engineering College
Pachal,Namakkal
elanchezhiyan2014@gmail.com

Rajadurai.M
UG student
Department of Computer science and Engineering
Paavai Engineering College
Pachal,Namakkal
rajadurai1324@@gmail.com

Satish.J
UG student
Department of Computer science and Engineering
Paavai Engineering College
Pachal,Namakkal
satishj8248@gmail.com

Prabudeva.B
UG student
Department of Computer science and Engineering
Paavai Engineering College
Pachal,Namakkal
prabudeva031999@gmail.com

***Abstract*— Cyberbullying is bullying that takes place over digital devices like cell phones, computers, and tablets. Cyberbullying can occur through SMS, Text, and apps, or online in social media, forums, or gaming where people can view, participate in, or share content. Cyberbullying includes sending, posting, or sharing negative, harmful, false, or mean content about someone else. It can include sharing personal or private information about someone else causing embarrassment or humiliation. The content an individual share online – both their personal content as well as any negative, mean, or hurtful content – creates a kind of permanent public record of their views, activities, and behaviour. To avoidor detecting cyberbullying attacks, many existing approaches in the literature incorporate Machine Learning and Natural Language Processing text classification models without considering the sentence semantics. The main goal of this project is to overcome that issue. This project proposed a model LSTM - CNN architecture for detecting cyberbullying attacks and it used word2vec to train the custom of word embeddings.This model is used to classify tweets or comments as bullying or non-bullying based on the toxicity score.LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. A convolutional neural network (CNN) is a type of artificial neural network and it has a convolutional layer to extract information by a larger piece of text and by using this model LSTM- CNN achieve a higher accuracy in analysis, classification and detecting the cyberbullying attacks on posts and comments.**

Keywords—convolutional neural network (CNN), Long short term memory (LSTM).

I.INTRODUCTION

Cyberbullying is bullying that takes place over digital devices like cell phones, computers, and tablets. Cyberbullying can occur through SMS, Text, and apps, or online in social media, forums, or gaming where people can view, participate in, or share content. Cyberbullying includes sending, posting, or sharing negative, harmful, false, or mean content about someone else. It can include sharing personal or private information about someone else causing embarrassment or humiliation. Some cyberbullying crosses the line into unlawful or criminal behaviour.

Different Kinds of Cyberbullying:

There are many ways that someone can fall victim to or experience cyberbullying when using technology and the internet. Some common methods of cyberbullying are:

Harassment – When someone is being harassed online, they are being subjected to a string of abusive messages or efforts to contact them by one person or a group of people. People can be harassed through social media as well as through their mobile phone (texting and calling) and email. Most of the contact the victim will receive will be of a malicious or threatening nature.

Doxing – Doxing is when an individual or group of people distribute another person's personal information such as their home address, cell phone number or place of work onto social media or public forums without that person's permission to do so. Doxing can cause the victim to feel extremely anxious and it can affect their mental health.

Cyberstalking – Similar to harassment, cyberstalking involves the perpetrator making persistent efforts to gain contact with the victim, however this differs from harassment – more commonly than not, people will cyberstalk another person due to deep feelings towards that person, whether they are positive or negative. Someone who is cyberstalking is more likely to escalate their stalking into the offline world.

Revenge porn – Revenge porn, is when sexually explicit or compromising images of a person have been distributed onto social media or shared on revenge porn specific websites without their permission to do so. Normally, images of this nature are posted by an ex-partner, who does it with the purpose of causing humiliation and damage to their reputation.

Swatting – Swatting is when someone calls emergency responders with claims of dangerous events taking.

1.1 PROJECT 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 unthinkable ways, they have also led to malevolent activities such as cyber-bullying. Cyberbullying 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 Interne 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, cyberbullying 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 cyberbullying in social media messages (e.g., posts, tweets, and comments).In this project, we mainly focus on the problem of cyberbullying detection on the Twitter platform. As cyberbullying is becoming a prevalent problem in Twitter, the detection of cyberbullying events from tweets and provisioning preventive measures are the primary tasks in battling cyberbullying threats. Therefore, there is a greater need to increase the research on social networks-based CB in order to get greaterinsights and aid in the development of effective tools and approaches to effectively combat cyberbullying problem. Manually monitoring and controlling cyberbullying on Twitter platform is virtually impossible. Furthermore, mining social media messages for cyberbullying 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, cyberbullying detection on social media is an open and active research topic. Cyberbullying detection within the Twitter platform has largely been pursued through tweet classification and to a certain extent with topic modelling approaches. Text classification based on supervised machine learning (ML) models are commonly used for classifying tweets into bullying and non-bullying tweets.Also, it may be suitable only for a pre-determined collection of events, but it cannot successfully handle tweets that change on the fly. Considering these limitations, an efficient tweet classification approach must be developed to bridge the gap between the classifier and the topic model so that the adaptability is significantly proficient.The deep learning model has been successfully used by current researchers to address various issues related to the text domain, including sentiment analysis, question answering, document classification, sentence classification, spam filtering and others. By following them, this project also uses a deep learning algorithm to address the cyber bullying detection issue. BiLSTM is capable of capturing the semantics of the sentence by performing the convolution operations over the tweets.

1.2 OBJECTIVE

[1] This project aims to automatically detect cyberbullying from tweets by using deep learning approaches.

[2] The aim of this project is to identify the maximum number of cyberbullying related tweets from Twitter as soon as it is posted by users.

[3] The objective of our solution is to identify the bullies from raw Twitter data based on the context as well as the contents in which the tweets exist. To warn and block the bully.

II.EXISTING SYSTEM

In this chapter existing machine learning classifiers utilized for tweet classification will be discussed. This chapter analysed five supervised machine learning algorithms: Support Vector Machines (SVM), Naive Bayes (NB), Random Forest (RF), Decision Tree (DT), Gradient Boosting model (GBM), Logistic Regression (LR) and Voting Classifier (Logistic Regression C Stochastic Gradient Descent classifier).

Random Forest-RF is a tree based classifier in which input vector generated trees randomly. RF uses random features, to create multiple decision trees, to make a forest. Then class labels of test data are predicted by aggregating voting of all trees. Higher weights are assigned to the decision trees with low value error. Overall prediction accuracy is improved by considering trees with low error rate.

Support Vector Machine-The Support vector machine (SVM) is understood that executes properly as sentiment analysis. SVM typifies preference, confines and makes usage of the mechanisms for the assessment and examines records, which are attained within the index area. Arrangements of vectors for every magnitude embody crucial details. Information (shown in form of vector) has been arranged in type to achieve this target. Next, the border is categorized in two training sets by stratagem. This is a long way from any area in the training samples. Support-vector machines in machine learning includes focused learning models connected to learning evaluations which inspect material that is exploited to categorize, also revert inspection.

Naive Bayes-Ordering approach, Naive Bayes(NB), with sturdy (naive) independent assumptions among stabilities, depends on Bayes' Theorem. NB classifier anticipates that the proximity of a specific element of class that is confined to the closeness of a couple of different variables. For instance, a natural organic product is presumably viewed as an apple, if its shading is dark red, if type of it is round and it is roughly 3 creeps in expansiveness. In machine learning, Naive Bayes classifiers are a gathering of essential "probabilistic classifiers" considering applying Bayes' speculation with gullible opportunity assumptions between the features. They are considered as the minimum problematic Bayesian network models.

2.1.1 DISADVATAGES

[1] Process of reporting such cases is long, tedious job.[2] Difficult to track.

[3] Low accuracy.[4] Response time is slow.[5] Data are manually labelled using online services or custom

applications.[6] Usually data limited only to a small percentage.

2.1 PROPOSED SYSTEM

Bidirectional LSTMs are an extension of LSTMs that can improve model performance on sequence classification problems. In problems where all time steps of the input sequence are available, Bidirectional LSTMs train two instead of one LSTMs on the input sequence. This can provide additional input context to the network and result in faster and even fuller learning on the problem. It involves duplicating the first periodic layer in the network so that there is now two layers' side-by-side, then providing the input sequence as-is as input to the first layer and providing a reversed copy of the input sequence to the second layer. The use of sequence bi-directionally was initially justified in the domain of speech recognition because there is evidence that the input context of the whole utterance is used to interpret what is being said rather than a simple interpretation. The use of bidirectional LSTMs may not make sense for all prediction problems but can offer benefits in terms of better results to those domains where it is appropriate.

2.1.1 ADVANTAGES

[1] It successfully classifies the tweets in various classes. [2]Auto report generator generates a simple report for probable accusers. [3] Several analytics and report can be sent to the crime department.[4] Foul language on any given page, removes it, and can highlight words as well.[5] This method detects the offensive post or messages it block that user id.[6] An automatically generate a report for each incident is also provided.

III.ALGORITHMS USED

A.CONVENTIONAL NEURAL NETWORK

[1]A conventional neural network (also known as feedforward neural network) is a type of machine learning algorithm that is designed to recognize patterns and make predictions.[2] It consists of multiple layers of interconnected nodes (neurons) that process information and pass it on to the next layer until the final output layer is reached.[3] In a typical feedforward neural network, information flows in one direction, from input nodes through hidden layers to output nodes.[4] Each node in a layer receives input from the previous layer and applies a mathematical transformation to produce an output, which is then passed to the next layer.[5] The weights and biases of the network are adjusted during the training process to optimize the performance of the network on a particular task. [6]Conventional neural networks are used in a wide variety of applications, including image recognition, speech recognition, natural language processing, and predictive modeling. [7] However, they have limitations in dealing with complex sequential data or long-term dependencies.[8] To overcome these limitations, recurrent neural networks (RNNs) and other more advanced architectures have been developed.[9] For some updating feature the conventional Neural Network (CNN) the Deep learning concept BiLSTM is used.

B.BiLSTM

[1] BiLSTM stands for Bidirectional Long Short-Term Memory. It is a type of recurrent neural network (RNN) that is designed to process sequential data in both forward and backward directions.[2] In a traditional LSTM (Long Short-Term Memory) network, the input sequence is processed in a single direction, from the first element to the last.[3] However, this approach may not be sufficient for some tasks where information from both directions is important.[4] For example, in natural language processing, the meaning of a word may be influenced by the words that come before and after it. A BiLSTM network consists of two LSTM layers, one processing the input sequence in the forward direction, and the other processing it in the backward direction.[4] The outputs of these two layers are then concatenated to produce the final output of the network.[5] By processing the sequence in both directions, the BiLSTM network can capture dependencies between elements of the sequence that may be missed by a unidirectional LSTM.

IV.PROBLEM DEFINITION

In the proposed framework shown in Figure 4.1., the process of detecting cyberbully activities begins with input dataset from social network. Input is text conversation collected from twitter. Input is given to data pre-processing which is applied to improve the quality of the project data and subsequent analytical steps, this includes removing stop words, extra characters and hyperlinks. After performing pre-processing on the input data, it is given to Feature Extraction. Feature Extraction is done to obtain features like Noun, Adjective and Pronoun from the text and statistics on occurrence of word (frequency) in the text. The cyberbullying words are given as training dataset. With the training dataset the preprocessed online social network conversation is tested for bullying word presence. Feature Vector distance algorithm detects the cyberbully words present in the conversation and displays it. For cyberbully Classification, BiLSTM is used.

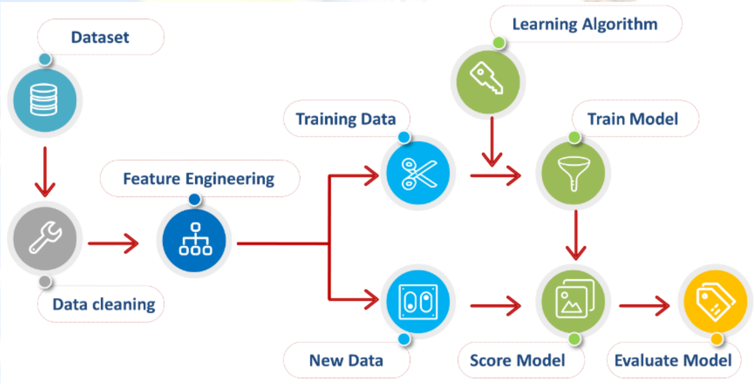

Figure 4.1. Cyberbullying Detection Proposed Model

V.OVERVIEW OF THE PROJECT

[1] This project aims to automatically detect cyberbullying from tweets by using deep learning approaches.[2] The aim of this project is to identify the maximum number of cyberbullying related tweets from Twitter as soon as it is posted by users.[3] The objective of our solution is to identify the bullies from raw Twitter data based on the context as well as the contents in which the tweets exist. To warn and block the bully.

VI.SYSTEM DESIGN

6.1 ARCHITECTURE DIAGRAM

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. System architecture can comprise system components, the externally visible properties of those components, the relationships (e.g. the behavior) between them. It can provide a plan from which products can be procured, and systems developed, that will work together to implement the overall system. There have been efforts to formalize languages to describe system architecture, collectively these are called architecture description languages (ADLs).

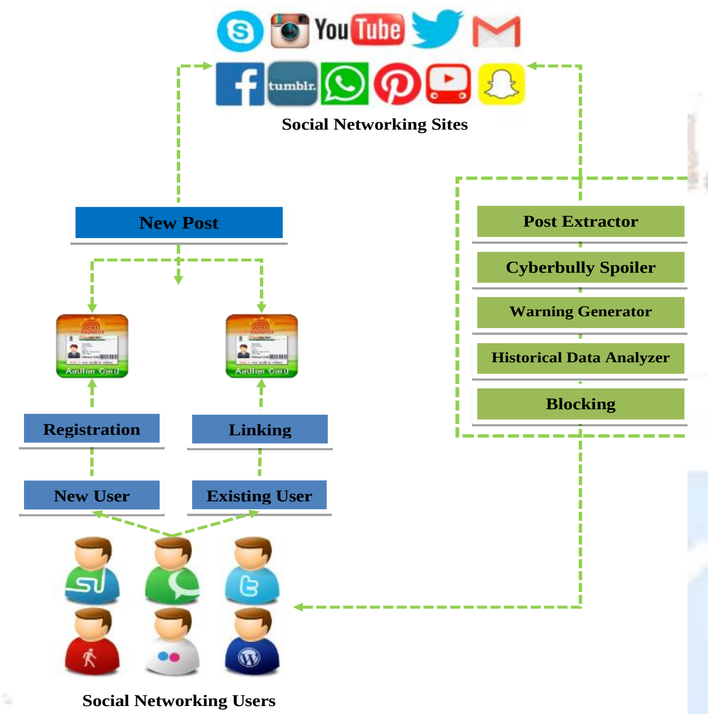

Figure no.6.1 Architecture diagram

Here, the user enters the login parameter, and the system automates to identify the compatibility of which type of attack happens in the system are find and gives the exact set of information about the type of attack happens over the system are analyzed by using CNN algorithm. And gives the solution for the attack are defined.

6.2 SEQUENCE DIAGRAM

A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios.

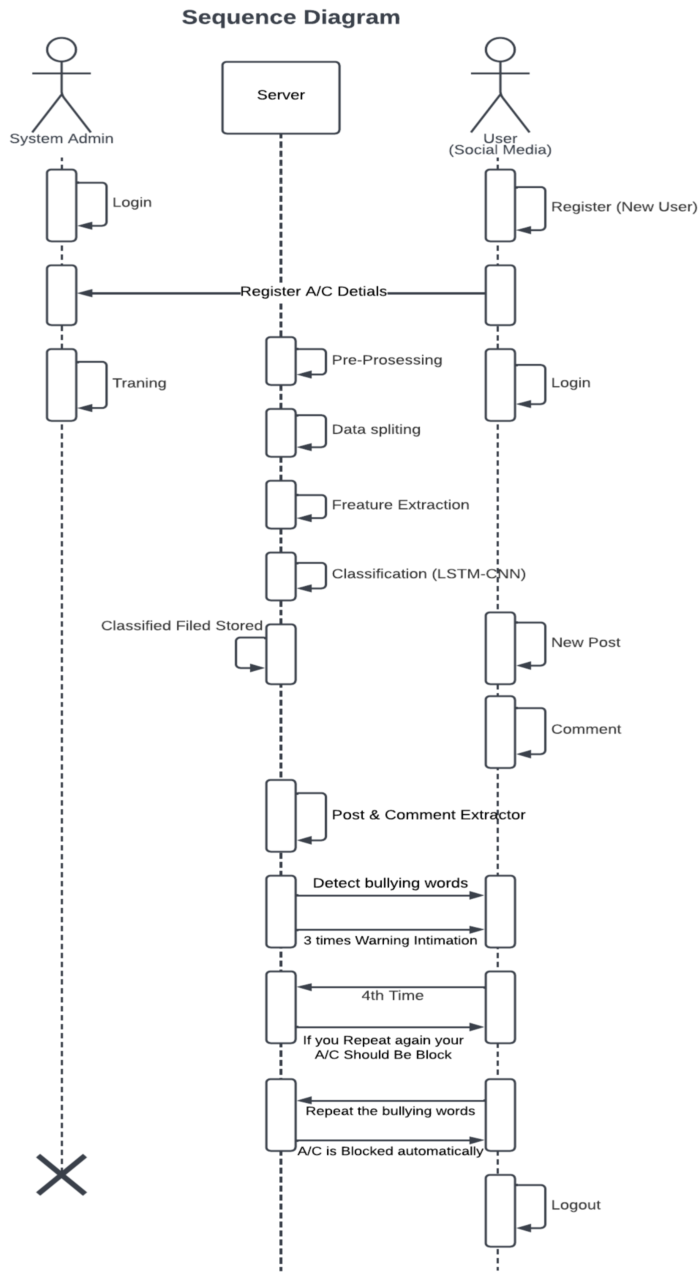

Figure no.6.2 Sequence Diagram

6.3 ACCURACY DIAGRAM

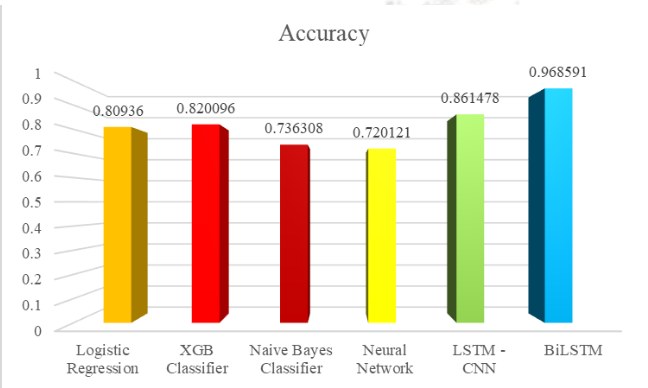

Figure no.6.3 Accuracy Diagram

6.4 DATAFLOW DIAGRAM

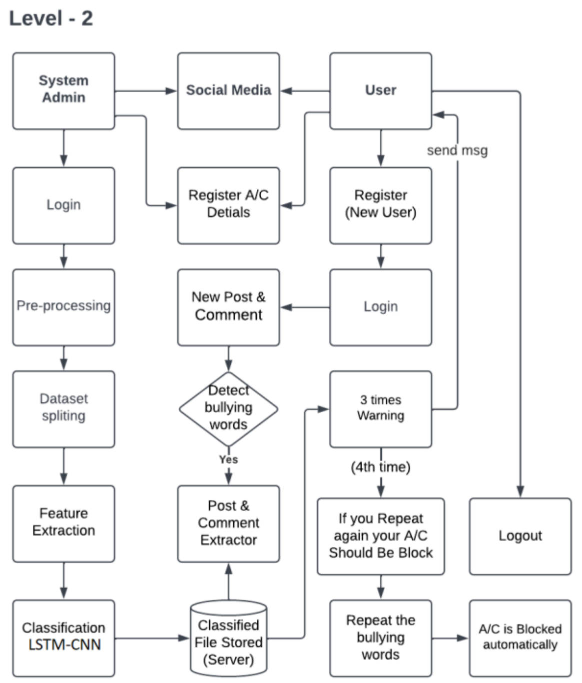

Figure no 6.4 Dataflow Diagram

VII.CYBERBULLYING TWEET CLASSIFICATION

We used cyberbullying data from Kaggle. The dataset in consisted of two labels, positive and negative, while was composed of three labels of positive, neutral, and negative. Furthermore, the dataset in was composed of five labels of positive, somewhat positive, neutral, somewhat negative, and negative.

Cyberbullying Data Set Annotation

A. Pre-processing

Datasets contain unnecessary data in raw form that can be unstructured or semi-structured. Such unnecessary data increases training time of the model and might degrades its performance. Pre-processing plays a vital role in improving the efficiency of DL models and saving computational resources. Text pre-processing boosts the prediction accuracy of the model. The preprocessing step is essential in cyberbullying detection. It consists of both cleaning of texts (e.g., removal of stop words and punctuation marks), as well as spam content removal. In the proposed model, it has been applied to remove and clean unwanted noise in text detection. For example, stop words, special characters, and repeated words were removed.

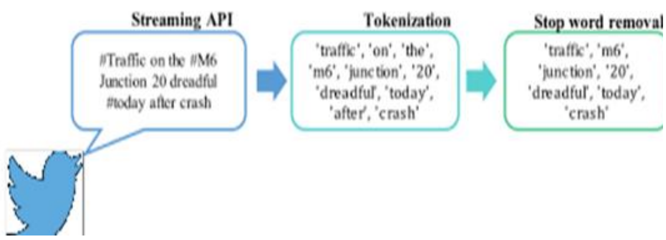

Figure no.7.1 Pre-processing

B. Feature Extraction

After the data pre-processing step, the next essential step is the choice of features on a refined dataset. Supervised deep learning classifiers require textual data in vector form to get trained on it. The textual features are converted into vector form using TF and TF-IDF techniques
in this project. Features extraction techniques not only convert textual features into vector form but also helps to find significant features necessary to make predictions. For the most part all features do no contribute to the prediction of the target class. That is the reason feature extraction is the important part in the recognition of happy and unhappy related tweets.

$$TF(t)=\frac{\text{No. of times term t shows in a document}}{\text{Total no. of terms inside document}}$$

$$IDF(t)= D\log(e)\,\frac{\text{Total No. of documents}}{\text{No. of documents through term t in it}}$$

Term frequency (TF) is utilized regarding data recovery and shows how regularly an articulation (term, word) happens in a tweets.

C. Word Embedding's

The semantic meanings of words are provided by word embedding in this project, which is first used in the semantic word cloud generation to the best of our knowledge. We prepare the related text corpus and then train our word embedding by using the continuous bag-of-words (CBOW) model, which is implemented in the open-source toolkit word2vec. CBOW is orders of magnitude faster than the others for training datasets and yields significant gains for dependency parsing. After training, we extract the semantic meanings of all important key words from the word embedding. By using the pre-trained word embedding, each word corresponds a vector in the low dimensional space, typically 50-500.

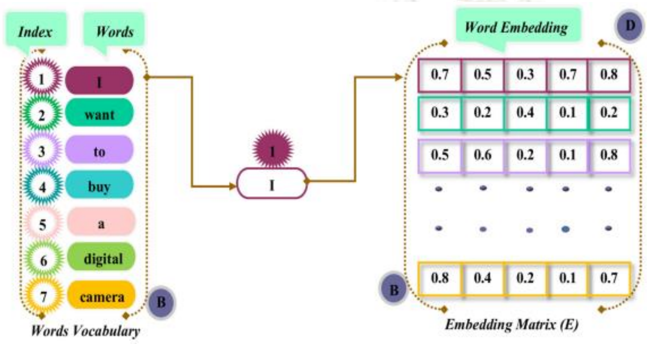

Figure no 7.2 Embedding Layer

VIII.WORKING OF THIS MODULE

The detection of spam emails can be evaluated by different performance measures. Confusion Matrix is being used to visualise the detection of the emails for models Several measurements are used for performance evaluation of classifiers like accuracy, precision, recall, and f-score. These measurements are computed by a confusion matrix, which is composed of four terms. Confusion matrix can be defined as below:

- True positive (TP): are the positive values correctly classified as positive.
- True Negative (TN): are the negative values correctly classified as negative.
- False Positive (FP): are the negative values incorrectly classified as positive.
- False Negative (FN): are the positive values incorrectly classified as negative.

For the performance evaluation of our proposed model, we use the following metrics.

8.1. Accuracy

The accuracy measure is the ratio of the number of bully users detected to the total number of bullies. It does not perform well with imbalanced data sets

$$\text{AccuracyCM} = \frac{\text{\# of detected bullies}}{\text{total number of bullies}}$$

training score :0.991249719542293
testing score :0.979372197309417

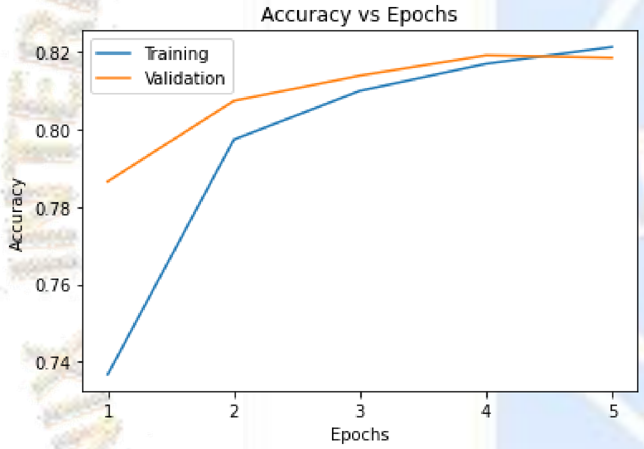

Figure no 8.1 Embedding Layer

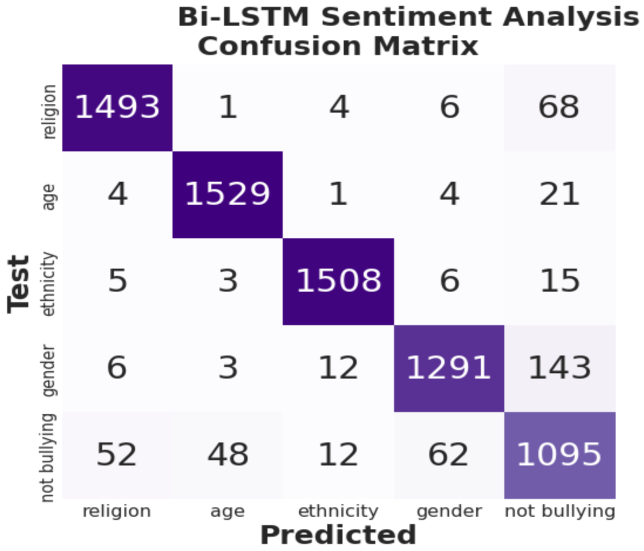

Figure no 8.2 Confusion Matrix

8.2. Precision

Precision is evaluation metrics used in binary classification tasks.Precision is the measure of exactness.

$$\text{Precision} = \frac{\text{\# of true bullies detected}}{\text{total number of detected users}}$$

In simple terms, high precision means that an algorithm returned substantially more bully users

8.3. Recall

The recall is a fraction of the predicted correctly classified applications to the total number of applications classified correctly or incorrectly. Recall is the measure of completeness.

$$\text{Recall} = \frac{\text{\# of true bullies detected}}{\text{total number of true bullies}}$$

whereas high recall means that an algorithm returned most of the bullies.

8.4.Score

F-score is the harmonic mean of precision and recall. It symbolizes the capability of the model for making fine distinctions. F1 Measure is the harmonic mean between precision and recall. The range for F1 is [0, 1]. It measures how many bullies are identified correctly and how robust it is. Mathematically, it can be expressed as

F1-Score=2*(Precision*Recall)/ (Precision + Recall)

F1 Measure attempts to find a balance between precision and recall. The greater the F1 Measure, the better is the performance of our approach.

IX.CONCLUSIONS

Cyberbullying is the harassment that takes place in digital devices such as mobile phones, computers and tablets. The means used to harass victims are very diverse: text messages, applications, social media, forums or interactive games.One of the things that complicates these types of situations that occur through the Internet, is the anonymity this environment allows. Since this facilitates cyberbullying can cover almost all areas of the victim's life, that is: educational environment, work, social or loving life. When the identity of the harasser is not known, even if the facts are reported, in many cases it is not enough to open an investigation, identify it and pay for the crime committed. This project proposed a deep learning model Bidirectional Long Short Term Memory (BiLSTM. Thus, this projecthasdesigned a method of automatically detecting the Cyberbullying attack cases. Identifies the messages or comments or posts which the BiLSTM model predicts as offensive or negative then it blocks that person id,then the admin can create automated reports and send to the concern department.Experiments are conducted to test three machine learning and 2 deep learning models that are; (1) GBM, (2) LR, (3) NB, (4) LSTM-CNN and (5) BiLSTM. This project also employed two feature representation techniques Tfand TF-IDF. The results showed that all models performed well on tweet dataset but our proposed BiLSTM classifier outperforms by using both TF and TF-IDF among all. Proposed model achieves the highest results using TF-IDF with 96% Accuracy, 92% Recall and 95% F1-score.

9.1 FUTURE ENHANCEMENTS

[1] While the deep learning approach is robust and flexible, there are certain steps which can be taken to improve their performance and better classify the data.[2] Integration with other security tools: MalFree can be enhanced to integrate with other security tools, such as firewalls and intrusion detection systems, to provide a more comprehensive cybersecurity solution.[3] Support for multiple operating systems: Currently, MalFree is designed to work with a specific operating system. Future enhancements can include support for multiple operating systems, such as Windows and Linux, to provide a more comprehensive cybersecurity solution. Integration with threat intelligence feeds.

9.2 FUTURE SCOPE

For the present, the bot works for Twitter, so it can be extended to various other social media platforms like Instagram, Reedit, etc. Currently, only images are classified for NSFW content, classifying text, videos could be an addition. A report tracking feature could be added along with a cross-platform Mobile / Desktop application (Progressive Web App) for the Admin. This model could be implemented for many languages like French, Spanish, Russian, etc. along with India languages like Hindi, Gujarati, etc.

REFERENCE

[1] 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. 332-344, Apr. 2021.

[2] Z. L. Chia, M. Ptaszynski, F. Masui, G. Leliwa and M. Wroczynski, "Machine learning and feature engineering-based study into sarcasm and irony classification with application to cyberbullying detection", Inf. Process. Manage., vol. 58, no. 4, Jul. 2021.

[3] N. Yuvaraj, K. Srihari, G. Dhiman, K. Somasundaram, A. Sharma, S. Rajeskannan, et al., "Nature-inspired-based approach for automated cyberbullying classification on multimedia social networking", Math. Problems Eng., vol. 2021, pp. 1-12, Feb. 2021.

[4] R. R. Dalvi, S. B. Chavan and A. Halbe, "Detecting a Twitter cyberbullying using machine learning", Ann. Romanian Soc. Cell Biol., vol. 25, no. 4, pp. 16307-16315, 2021.

[5] N. Yuvaraj, V. Chang, B. Gobinathan, A. Pinagapani, S. Kannan, G. Dhiman, et al., "Automatic detection of cyberbullying using multi-feature based artificial intelligence with deep decision tree classification", Comput. Electr. Eng., vol. 92, Jun. 2021.

[6] A. Al-Hassan and H. Al-Dossari, "Detection of hate speech in Arabic tweets using deep learning", Multimedia Syst., Jan. 2021. mechanism", Information, vol. 12, no. 4, pp. 171, Apr. 2021.

[7] Y. Fang, S. Yang, B. Zhao and C. Huang, "Cyberbullying detection in social networks using bi-GRU with self-attention

[8] B. A. Talpur and D. O'Sullivan, "Multi-class imbalance in text classification: A feature engineering approach to detect cyberbullying in Twitter", Informatics, vol. 7, no. 4, pp. 52, Nov. 2020.

[9] A. Agarwal, A. S. Chivukula, M. H. Bhuyan, T. Jan, B. Narayan and M. Prasad, "Identification and classification of cyberbullying posts: A recurrent neural network approach using under-sampling and class weighting" in Neural Information Processing, Cham, Switzerland:Springer, vol. 1333, pp. 113-120, 2020.

[10] C. Iwendi, G. Srivastava, S. Khan and P. K. R. Maddikunta, "Cyberbullying detection solutions based on deep learning architectures", Multimedia Syst., 2020.

[11] L. Cheng, J. Li, Y. N. Silva, D. L. Hall, and H. Liu, "XBully: Cyberbullying detection within a multi-modal context," in Proc. 12th ACM Int. Conf. Web Search Data Mining, Jan. 2019, pp. 339–347.

[12] C. Van Hee et al., "Automatic Detection of Cyberbullying in Social Media Text." 2018.

[13] M. Rezvan, S. Shekarpour, L. Balasuriya, K. Thirunarayan, V. Shalin, and A. Sheth, "A Quality Type-aware Annotated Corpus and Lexicon for Harassment Research," pp. 33–36, 2018.

[14] A. H. Alduailej and M. B. Khan, "The challenge of cyberbullying and its automatic detection in Arabic text," 2017 Int. Conf. Comput. Appl. ICCA 2017, pp. 389–394, 2017.

[15] A. Power, A. Keane, B. Nolan, and B. O. Neill, "A lexical database for public textual cyberbullying detection," Rev. Lenguas Para Fines Específicos, vol. 2, pp. 157–186, 2017.

[16] M. Drahošová and P. Balco, "ScienceDirect The analysis of advantages and disadvantages of use of social media the analysis of advantages and disadvantages of use of social media in European Union in European Union," Procedia Comput. Sci., vol. 109, pp. 1005–1009, 2017.

[17] R. Zhao, A. Zhou, and K. Mao, "Automatic detection of cyberbullying on social networks based on bullying features," in Proc. 17th Int. Conf. Distrib. Comput. Netw., Jan. 2016, pp. 1–6.

[18] V. K. Singh, Q. Huang, and P. K. Atrey, "Cyberbullying detection using probabilistic socio-textual information fusion," in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), Aug. 2016, pp. 884–887.

[19] A. Squicciarini, S. Rajtmajer, Y. Liu, and C. Griffin, "Identification and characterization of cyberbullying dynamics in an online social network," in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining, Aug. 2015, pp. 280–285.

[20] 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 J. IGPL, vol. 24, no. 1, pp. 42–53, 2015.