Intelligent Chatbot Development for Text based Cyberbullying Prevention

Dr. Vijayakumar V
Professor& COE, Department of Computer Science,
Sri Ramakrishna College of Arts and Science, Coimbatore-641 006, Tamilnadu, India.

Dr Hari Prasad D
Professor and Head, Department of Computer Applications,
Sri Ramakrishna College of Arts and Science, Coimbatore-641 006, Tamilnadu, India,

Abstract- Now a day's social media network offer great communication opportunities, they increase the vulnerability of users to threatening situations online. They also lead to addiction; misleading information; scams & frauds and reduce family association. One of the main challenges and issues of social media is cyberbullying. Cyberbullying can occur through sending, posting, or sharing negative, hurtful, untrue, or cruel substance almost somebody else. Text is the most popular medium of social media as it is simple and shortest way of communication. Capturing information from text messages is the most important aspect of text-based cyberbullying identification, but it is still a challenge. Machine or deep learning algorithms, NLP algorithms are useful to predict the cyberbullying. In this paper, a deep learning based LSTM algorithm is proposed to detect the cyberbullying incident. The proposed method also interfaced with chatbot to manage and prevent the event effectively. The experimental conducted on real world, publicly available social media datasets. This paper aims to develop practical and impactful application for future systems.

Keywords – Cyberbullying Prediction, Intelligent Chat Bot, Deep Learning, cyber safety

I. INTRODUCTION

Social media are utilized by one-in-three people within the world, and more than two-thirds of all internet customers. Due to usage of social media and technology, the growths of online crimes are increasing. Most of the social media platforms like Facebook, WhatsApp, Instagram, snapchat etc. have reported higher levels of cyberbullying to the users [4]. Cyberbullying is posting personal information, pictures, or videos designed to hurt, embarrass, harass, threaten, or target every other individual through the use of technology. It can lead to serious problems such as Emotional Effects of Isolation, Anger; Mental Effects of Depression and Anxiety, Academic Issues, Suicidal Thoughts and Self-Harm; Behavioural Effects of Using drugs or alcohol and Skipping regular activities. Cyberbullying Prevention depends on the discovery of unsafe data such as messages, images and videos and providing the proper alerts for the victim. Although research on cyberbullying detection is more limited, majority of research is to identify the motivation behind bullying and looked at the problem from socio-psychological and educational perspectives. Automatically identifying bully words, emojis and audio/video features from online social platforms is an important research.

Cyberbullying detection requires effective intelligent systems to identify possible threats automatically. The main approach to cyberbullying detection involves natural language processing, machine learning and deep learning. These algorithms are finding a negative comment from the messages in text which received by a user. Cyberbullying happens in multiple platforms in different types. The detection of cyberbullying in social media is more challenging task due to the content information in social media is short, noisy and unstructured and people unaware of being cyberbullied. It is observed that cyberbullying involves text, image, audio, video and emoji types of data.

An artificial intelligence chatbot can be used to simulate a conversation with a user in natural language through messaging. It gives us the chance to engage in a more delightful way and create a better experience for the user. This technology can be strengthening in saving the victims to report their issues of harassment accurately and securely. In this paper, an intelligence chatbot was developed with text based Deep Neural Network (DNN) cyberbullying prevention.

The rest of the paper is organized as follows. The detailed review of literature is presented in the section II. Proposed method and detection algorithms are explained in section III. Experimental results are presented in section IV. Conclusion and future works are given in section V.

II. Review of Literature

In the recent years many works were proposed related to cyberbullying detection. Our research focuses on text based cyberbullying prevention because text is the most common form of social media. The literature review was prepared based on the dataset used, nature of detection and methods used for bullying prevention.

Most of the authors used Fromspring.Me, ask.fm, twitter, YouTube etc datasets. The author Ç. ACI et al. used multiple datasets [20] to investigate the effects of text mining methods such as pre-processing, feature extraction, feature selection and classification on automatic detection of cyberbullying. N. Lu et al. used Chinese weibo dataset, English tweet dataset [2] and proposed a Char - CNNS (Character - level Convolutional Neural Network with Shortcuts) model to identify whether the text in social media contains cyberbullying. C. Van Hee et al. [21] proposed automatic cyberbullying detection in social media by modelling posts written by bullies, victims, and bystanders of online bullying in English and Dutch dataset. L. Cheng et al. and M. Yao et al. used Instagram datasets [8][16] [18] for cyberbullying detection. V. Banerjee et al. detected the cyberbullying text in Twitter dataset using deep neural network [7]. M. A. Al-Ajlan et al. [24] and N. Tahmasbi et al. [26] also used Twitter dataset for detecting cyberbullying text. A. Bakshi et al. used YouTube dataset [19]. C. Emmery et al. used the Ask.fm and crowdsourced dataset for cyberbullying detection [10]. H. Rosa et al., conducted an in-depth analysis of 22 studies on automatic cyberbullying detection using formspring data [14]. P. Tata et al used Crime debate forum dataset for detection using hybrid techniques (multiple correlation coefficient – MCC and support vector machine – SVM) [27]. Another major source of cyberbullying data is gaming platforms. V. Balakrishnan et al., M. Garaigordobil et al and S. Murnion used various gaming platform data to detect cyberbullying [5][29][30].

Machine learning covers a wide range of methods that enables systems to quickly access and learn from data, and to make decisions about complex problems. Machine learning algorithms could help to timely and automatically detect cyberbullying. Talpur BA constructed machine learning based multi-classifier for classifying cyberbullying harshness into distinctive tiers [1]. V. Balakrishnan et al. proposed the Random Forest, machine-learning algorithm for cyberbullying classification (i.e. aggressor, spammer, bully and normal) [5]. A. Bakshi et al., [19], S. Murnion et al. [29] applied ML techniques to detect the cyberbullying. Feature extraction plays a crucial role in classification. There are new features are identified to improve the cyberbullying detection process. However, this approach might generate huge number of features that require careful feature extraction and selection stages. This leads to computational overhead and harder. Deep learning has the advantage of identifying a model of unstructured data such as images, sound, video, and text. It also automatically learns features from data. In recent studies, different models of deep learning algorithms proposed for cyberbullying detection. V. Banerjee et al., used the Deep neural network methods for cyberbullying detection [7]. M. A. Al-Ajlan et al. proposed a CNN-CB algorithm to exclude the need for features and yield better prediction [24]. H. Rosa et al., implemented a simple CNN, a hybrid CNN-LSTM and a mixed CNN-LSTM-DNN [14] [33] and trained via the word2vec model with Google-News, Twitter and Formspring data set.

The usage of Natural Language Processing (NLP) is very necessary in the case of cyberbullying detection process. It helps to understand the context of a conversation. It extracts grammatical structure and meaning from input. NLP is used by C. Ziems et al. [11] and N. Tahmasbi et al. [6] and Ibn Rafiq et al., [35] in their cyberbullying research. They examined predictive strength n-grams, part-speech information and sentiment information based on profanity lexicons used for detecting events related to cyberbullying.

A number of smartphone apps have been created to monitor interactions between teenagers and general social media users. Some applications give parents complete control over their children's online interactions with others such as BullyBlocker. Thun LJ et al., analysed different types of mobile applications that manage cyberbullying. They proposed a mechanism, which combines the best cyberbullying detection features to fill the gaps and limitations of existing applications [34]. Several types of applications are available to detect cyberbully online such as ReThink, Cyberbully Blocker and etc. These applications sends warning message when users try to send texts consisting of harmful words. In this proposed work, Cyberbullying detection model integrates into a Chatbot application. It alerts and prevents the social media users if any bully trying to bullying through text messages in social media text. The summary of the review of literature is presented in the Table 1.

Table -1 Experiment Result

Authors	Year	Method / Technology / Algorithm	Outcome	Challenges / Issues / Demerit / Limitation
N. Lu, G. Wu, Z. Zhang, Y. Zheng, Y. Ren, and K. K. R. Choo	2020	Character level-CNN	Cyberbullying Detection	can't identify types of cyberbullying
A. J. Sánchez-Medina, I. Galván-Sánchez, and M. Fernández-Monroy	2020	Random Forest and Adabag package	Cyberbullying Detection with neuroticism and psychopathy	A particular language based research
C. Ziems, Y. Vigfusson, and F. Morstatter	2020	Logistic regression	Cyberbullying Detection	lack of real world datasets
Ç. ACI, E. ÇÜRÜK, and E. S. EŞSİZ	2019	SGD and MLP classifiers, SVM-RFE algorithm	Cyberbullying Detection on selected features	soft computing and deep learning techniques are absent
A. Bakshi and A. K. Patel	2019	Random Forest, k-Nearest Neighbor, Sequential Machine Optimization, and Naive Bayes	Cyberbullying Detection	soft computing and deep learning techniques are absent
V. Balakrishnan, S. Khan, T. Fernandez, and H. R. Arabnia	2019	Random Forest	Cyberbullying Detection with neuroticism and psychopathy	Dataset is based on online game
V. Banerjee, J. Telavane, P. Gaikwad, and P. Vartak	2019	CNN	Cyberbullying detection using DNN	Single data source
L. Cheng, R. Guo, Y. Silva, D. Hall, and H. Liu	2019	Bidirectional GRU-RNN	Multi Model Cyberbullying detection with time series	Single data source
L. Cheng, J. Li, Y. N. Silva, D. L. Hall, and H. Liu	2019	Random Forest, Linear SVM, Logistic Regression	Multi Model Cyberbullying detection with time series	Less amout of data is used
L. Cheng, J. Li, Y. Silva, D. Hall, and H. Liu	2019	k∗-NN	Cyberbullying Detection	Dataset is based on few keywords onlyNA
Tanmayee Patange, Jigyasa Singh, Aishwarya Thorve, Yadnyashree Somaraj Madhura Vyawahare	2019	CNN, Word2Vec, Offensiveness	Multi input cyberbullying detection	Instagram only
R. Sprugnoli, S. Menini, S. Tonelli, F. Oncini, and E. Piras	2019	NA	Dataset created	Italian language oriented dataset
N. Tahmasbi and A. Fuchsberger	2019	NLP, Machine learning	Cyberbullying detection for helping parents	Targetted detection only
M. Yao, C. Chelmis, and D. S. Zois	2019	CONcISE-framework	Cyberbullying detection on Instagram	Single data source
M. A. Al-Ajlan and M. Ykhlef	2018	CNN	Cyberbullying Detection	Not Multilingual
M. Dadvar and K. Eckert	2018	CNN, LSTM, BLSTM and BLSTM with attention	Cyberbullying Detection	Not Multilingual
S. V. Georgakopoulos, A. G. Vrahatis, S. K. Tasoulis, and V. P. Plagianakos	2018	CNN	Cyberbullying Detection	Not Multilingual
S. Murnion, W. J. Buchanan, A. Smales, and G. Russell	2018	NA	Sentiment analysis	Based on game data
T. Pradheep, J. Sheeba, T. Yogeshwaran, and S. Pradeep Devaneyan	2018	Shot Boundary detection algorithm, Optical Character Recognition, Naive Bayes	Multi Model Cyberbullying detection	Need of new algorithms in detection
H. Rosa, D. Matos, R. Ribeiro, L. Coheur, and J. P. Carvalho	2018	simple CNN, a hybrid CNN-LSTM and a mixed CNN-LSTM-DNN	Cyberbullying Detection	Time consumption in model building
N. Tahmasbi and E. Rastegari	2018	Naïve Bayes, SVM, Random Forest, logistic, JRip, and J48		cross-context and cross-platform approach to automated cyberbullying problem
P. Tata, S. P. Devaneyan, and J. I. Sheeba	2018	Linear SVM	detection of cyberbully	Need of new algorithms in detection
C. Van Hee et al.	2018	SVM	detect signals of cyberbullying on social media	detection of threats, curses and expressions of racism and hate is not possible
H. Dani, J. Li, and H. Liu	2017	TF-IDF	Cyberbullying Detection, Sentiment Difference	Multi lingual is not possible

The contribution of the proposed paper are listed below:
- The cyberbullying dataset are manual labelled which are scraped from twitter tweets.
- A temporal based LSTM deep learning model was developed to detect cyberbullying.
- A user friendly chatbot was designed for interacting with user.
- The live cyberbullying texts / comments detection method was verified.
- A system was developed

III. PROPOSED METHOD

The data is collected from social media platform tweets of twitter by performing an initial manual search of common slurs and terms used pertaining to religious, sexual, gender, and ethnic minorities from different parts of the world. This dataset includes 16K annotated tweets, 5054 are labelled as positive for cyberbullying and the remaining are marked as neither sexist nor racist. The processed dataset containing mixed cyberbullying tweet and not. This data is manually labelled.

A temporal based Long Short Term Memory (LSTM) is proposed for text based cyberbullying detection. It is an advanced Recurrent Neural Network (RNN). While processing long sequence, the LSTM network permits information to persist. Gradients are values used to update a neural network weights. Vanishing gradient problem happens when the gradient shrinks as it back propagates through time. When gradient value becomes extremely small it is neglected. This stops the update of neural network. The proposed LSTM method is capable of handling the vanishing gradient problem. LSTM has forget gate, input gate and output gates. The gates can learn what information is relevant to keep or forget during training. The forget gate decides what information should be thrown away or kept. Input gate is used to update cell state. The output gate decides what the next hidden state should be. Hidden state contains information on previous inputs. This hidden state is also used for predictions.

After the trained deep neural network model to achieve prediction in cyberbullying content in text based using data set. Then, the model is deployed in a manner with registered people. They can test the detection process. An observation on incoming messages are done in social media is enabled after the parent is given permission. And if any bully events occurred, it is recorded into a database and fetched to the parent via chatbot. For the prevention of the cyberbullying an alert will be sent the parent and the children regarding the incident. There will be options to complaint to the official authorities on the incident. The proposed method is shown in the Figure 1 and LSTM Model for Cyberbullying Prediction is shown in the Figure 2.

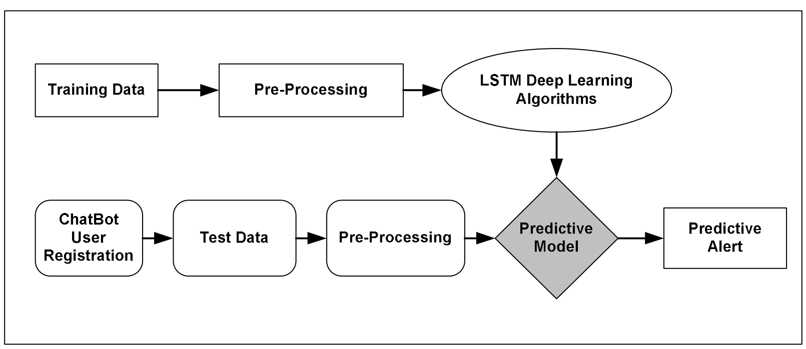

Figure 1. Proposed Cyberbullying Detection and Prevention Method

The detailed step by step explained below:

Step 1: Model Development

- Importing Libraries: Tensorflow pre-processing is used for vocabulary pre-processing. The Tensorflow, Keras, Numpy, Sklearn, libraries are imported to build the model.
- Load Dataset: The twitter tweets data is manually labelled with different classes. The data is classified into train and test with 10% test size
- Build Model :
  - The LSTM (Feedback Neural Network) model is prepared as an embedding layer followed by dropout function. Then, a dense layer is added.
  - The output class is defined as first class is categorized to not bully other classes is to bully.

- Train a LSTM model

Step 2: User Registration in Chatbot

Step 3: Deploy the model.

Step 4: Test the created model with chatbot to alert the bully.

Step 5: If any Cyberbullying text predicts, it give alert message to user to prevent the cyberbullying event..

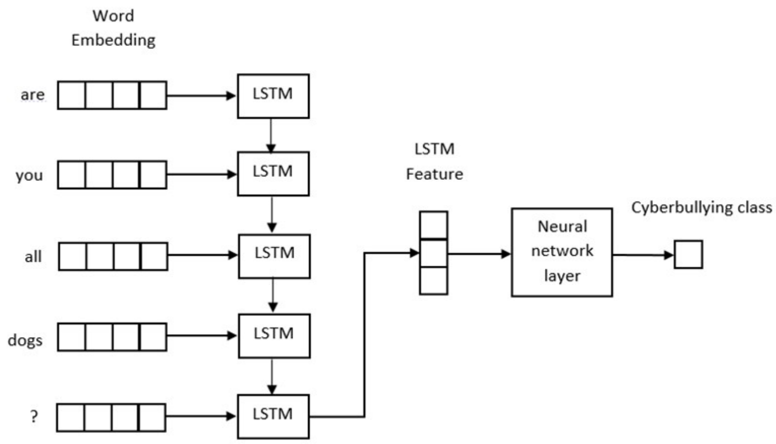

Figure 2. LSTM Model for Cyberbullying Prediction

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The LSTM based Deep Neural Network was built with Keras. The Google COLAB with GPU supports to help to process data and DNN modelling very fast. An Intel i5 processor system with 8 GB RAM was used to develop the model. Internet with at least speed of 1 MBPS is used for Google Colab GPU.

The dataset contains tweets from micro blogging platform twitter (www.twitter.com). It contains examples of racism and sexism. A total of 16,000 manual labelled tweets contains sexual, ethnic minority, religious, gender, neutral relates contents. Out of total tweets 5054 tweets are positive for cyberbullying. The DNN takes pre-processed tweets as input and gives prediction as output. The model summary is shown in the Figure 3. The model is observed as initial word embedding layer followed by a dropout function. Then, it is aligned in LSTM layer and again a dropout function is called followed by dense layer.

```
Layer (type)                 Output Shape              Param #
=================================================================
embedding_4 (Embedding)      (None, 38, 200)           1778800
_________________________________________________________________
dropout_7 (Dropout)          (None, 38, 200)           0
_________________________________________________________________
lstm_4 (LSTM)                (None, 200)               320800
_________________________________________________________________
dropout_8 (Dropout)          (None, 200)               0
_________________________________________________________________
dense_4 (Dense)              (None, 3)                 603
=================================================================
Total params: 2,100,203
Trainable params: 2,100,203
Non-trainable params: 0
```

Figure 3: Model Summary

4.1. Performance Metrics:–

The performance of the proposed framework was measured in terms of the quality measures, namely precision, F1 score and Recall. Precision is the ratio between the true positive (correct predictions) and the total predictions. Recall is the ratio of the correct predictions and the total number of correct items in the set. F1 score is the weighted harmonic mean of precision and recall. Accuracy calculates the proportion of correctly identified cyberbully words. The accuracy of the model is 94% working efficiency.

Precision	Recall	F1 score
0.96	0.88	0.92

The embedding layer output size and the LSTM layer size is equal. In each iteration this number is increased when the model is running. It varies from 50 to 200. So the accuracy is increased each time. The efficiency starting from 90% to 94%. This Deep Neural Network (DNN) model effectively detects the cyberbullying text in given messages.
Cyberbullying texts are identified individually from a group of text messages. The sample output predictions is shown in the Figure 4.

Text	Prediction
"you f**king faggot",	1
"you are my biggest weakness",	0
"you are a bottle neck",	1
"weddings are very fun",	0
"the connection is strong",	0
"that person is great",	0
"old people have diabetics",	1
"my job is make you simple",	0
"i don't make any profit",	0
"i am there for you, bro",	0
"i am the living proof that global warming exist",	1
"i am gonna do some jokes",	0
"how much for the shirt?",	0
"f**k off you bitch",	1
"everything has a price",	0
"be uncomfortable with that",	1
"are you all dogs?",	1

Figure 4. Sample inputs and Predictions

Cyberbullying monitoring chatbot was developed for monitoring and prevention. The Figure 5. Show the registration of user in the chatbot. The chatbot chat history is shown in the Figure 6. The detailed communication with user and alert management is shown in the Figure 7 and Figure 8.

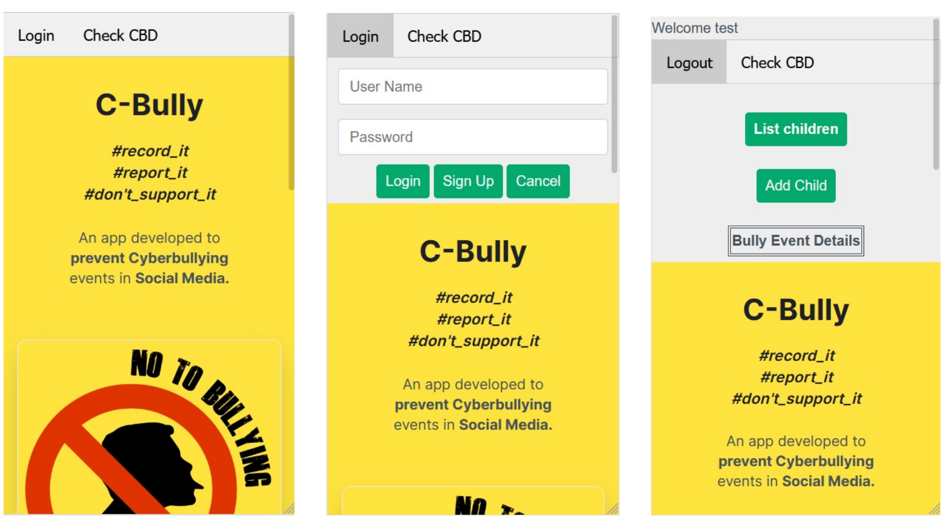

Figure 5. Chatbot Registration

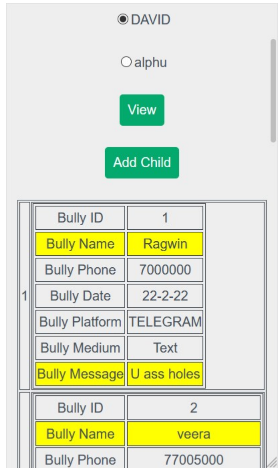

Figure.6. Chatbot Report

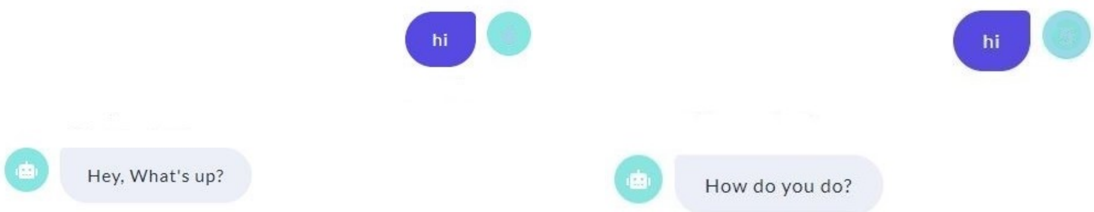

Figure 7. Chatbot Communication

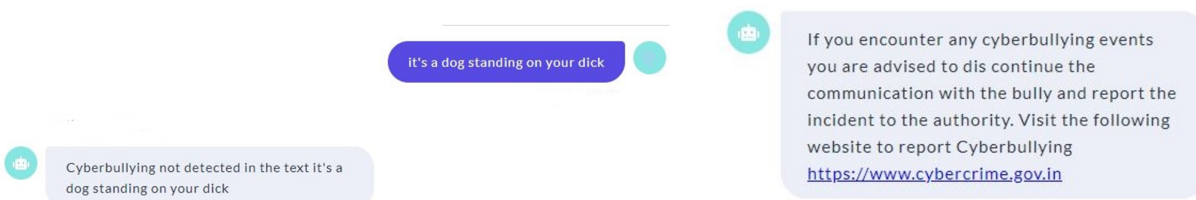

Figure 8. Cyberbully Detecion and Prevention

V. Conclusion and Future Work

Many active researches addressed cyberbullying problem in different approaches. In this paper, an intelligence chatbot was developed with text based Deep Neural Network cyberbullying prevention. Long short term memory is proposed for text based cyberbullying detection. The predicted results of this model is starting 90% to 94% and detects more accurately compared to machine learning models. LSTM based Deep Neural Network (DNN) model gives higher performance in prediction. Automatically detecting cyberbullying event and pro-actively acting upon it becomes of the utmost importance. It shows great advantage to the world, to fight against cyberbullying. The main challenges of cyber bullying detection are the scarcity of cyberbullying data, context of the conversation and time series data in nature. And also the presence of short forms and emoji's, multilingual data, Cross language, Mix language are some of the challenges in the prediction. In future, the proposed system can be extended with the use of multi-source data from (Reddit, Wikipedia, Formspring etc.) and Multi language support integration with Chabot.

Acknowledgment

This research work was funded by the ICSSR under IMPRESS Scheme. The author also thanks the ICSSR and Sri Ramakrishna College of Arts and Science who provided the support for doing the research.

REFERENCES

[1] Talpur BA, O'Sullivan D (2020) "Cyberbullying severity detection: A machine learning approach". PLoS ONE 15(10): e0240924. https://doi.org/10.1371/journal.pone.0240924

[2] N. Lu, G. Wu, Z. Zhang, Y. Zheng, Y. Ren, and K. K. R. Choo, “Cyberbullying detection in social media text based on character-level convolutional neural network with shortcuts,” Concurr. Comput. , no. July 2019, pp. 1–11, 2020, doi: 10.1002/cpe.5627.

[3] A. J. Sánchez-Medina, I. Galván-Sánchez, and M. Fernández-Monroy, “Applying artificial intelligence to explore sexual cyberbullying behaviour,” Heliyon, vol. 6, no. 1, 2020, doi: 10.1016/j.heliyon.2020.e03218.

[4] S. Salawu, Y. He, and J. Lumsden, “Approaches to Automated Detection of Cyberbullying: A Survey,” IEEE Trans. Affect. Comput., vol. 11, no. 1, pp. 3–24, 2020, doi: 10.1109/TAFFC.2017.2761757.

[5] V. Balakrishnan, S. Khan, T. Fernandez, and H. R. Arabnia, “Cyberbullying detection on twitter using Big Five and Dark Triad features,” Pers. Individ. Dif., vol. 141, no. January, pp. 252–257, 2019, doi: 10.1016/j.paid.2019.01.024.

[6] N. Tahmasbi and A. Fuchsberger, “ChatterShield – A multi-platform cyberbullying detection system for parents,” 25th Am. Conf. Inf. Syst. AMCIS 2019, pp. 1–5, 2019.

[7] V. Banerjee, J. Telavane, P. Gaikwad, and P. Vartak, “Detection of Cyberbullying Using Deep Neural Network,” 2019 5th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2019, pp. 604–607, 2019, doi: 10.1109/ICACCS.2019.8728378.

[8] L. Cheng, R. Guo, Y. Silva, D. Hall, and H. Liu, “Hierarchical attention networks for cyberbullying detection on the instagram social network,” SIAM Int. Conf. Data Mining, SDM 2019, pp. 235–243, 2019, doi: 10.1137/1.9781611975673.27.

[9] R. Sprugnoli, S. Menini, S. Tonelli, F. Oncini, and E. Piras, “Creating a WhatsApp Dataset to Study Pre-teen Cyberbullying,” pp. 51–59, 2019, doi: 10.18653/v1/w18-5107.

[10] Emmery, C., Verhoeven, B., De Pauw, G. et al. “Current limitations in cyberbullying detection: On evaluation criteria, reproducibility, and data scarcity.” Lang Resources & Evaluation (2020). https://doi.org/10.1007/s10579-020-09509-1

[11] T. Mahlangu, C. Tu, and P. Owolawi, “A review of automated detection methods for cyberbullying,” 2018 Int. Conf. Intell. Innov. Comput. Appl. ICONIC 2018, pp. 1–5, 2019, doi: 10.1109/ICONIC.2018.8601278.

[12] L. Cheng, J. Li, Y. Silva, D. Hall, and H. Liu, “Pi-Bully: Personalized cyberbullying detection with peer influence,” IJCAI Int. Jt. Conf. Artif. Intell., vol. 2019-Augus, pp. 5829–5835, 2019, doi: 10.24963/ijcai.2019/808.

[13] A. Singh and M. Kaur, “Content-based cybercrime detection: A concise review,” Int. J. Innov. Technol. Explor. Eng., vol. 8, no. 8, pp. 1193–1207, 2019.

[14] H. Rosa et al., “Automatic cyberbullying detection: A systematic review,” Comput. Human Behav., vol. 93, no. October 2018, pp. 333–345, 2019, doi: 10.1016/j.chb.2018.12.021.

[15] H. Vandebosch, “Cyberbullying Prevention, Detection and Intervention.” Springer International Publishing, 2019.

[16] M. Yao, C. Chelmis, and D. S. Zois, “Cyberbullying ends here: Towards robust detection of cyberbullying in social media,” Web Conf. 2019 - Proc. World Wide Web Conf. WWW 2019, pp. 3427–3433, 2019, doi: 10.1145/3308558.3313462.

[17] Y. S. M. V. Tanmayee Patange, Jigyasa Singh, Aishwarya Thorve, “DETECTION OF CYBERHECTORING ON INSTAGRAM,” pp. 5–8, 2019.

[18] L. Cheng, J. Li, Y. N. Silva, D. L. Hall, and H. Liu, “XBully: Cyberbullying Detection within a Multi-Modal Context,” pp. 339–347, 2019, doi: 10.1145/3289600.3291037.

[19] A. Bakshi and A. K. Patel, “Empirical Analysis of Supervised Machine Learning Techniques for Cyberbullying Detection,” vol. 2, no. January. Springer Singapore, 2019.

[20] Ç. ACI, E. ÇÜRÜK, and E. S. EŞSİZ, “Automatic Detection of Cyberbullying in Formspring.Me, Myspace and Youtube Social Networks,” Turkish J. Eng., vol. 3, no. 4, pp. 168–178, 2019, doi: 10.31127/tuje.554417.

[21] C. Van Hee et al., “Automatic detection of cyberbullying in social media text,” PLoS One, vol. 13, no. 10, pp. 1–22, 2018, doi: 10.1371/journal.pone.0203794.

[22] S. V. Georgakopoulos, A. G. Vrahatis, S. K. Tasoulis, and V. P. Plagianakos, “Convolutional neural networks for toxic comment classification,” ACM Int. Conf. Proceeding Ser., no. April, 2018, doi: 10.1145/3200947.3208069.

[23] T. Pradheep, J. . Sheeba, T. Yogeshwaran, and S. Pradeep Devaneyan, “Automatic Multi Model Cyber Bullying Detection from Social Networks,” SSRN Electron. J., pp. 248–254, 2018, doi: 10.2139/ssrn.3123710.

[24] M. A. Al-Ajlan and M. Ykhlef, “Deep learning algorithm for cyberbullying detection,” Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 9, pp. 199–205, 2018, doi: 10.14569/ijacsa.2018.090927.

[25] M. Dadvar and K. Eckert, “Cyberbullying Detection in Social Networks Using Deep Learning Based Models; A Reproducibility Study,” pp. 1–13, 2018, [Online]. Available: http://arxiv.org/abs/1812.08046.

[26] N. Tahmasbi and E. Rastegari, “A Socio-contextual Approach in Automated Detection of Cyberbullying,” Proc. 51st Hawaii Int. Conf. Syst. Sci., vol. 9, pp. 2151–2160, 2018, doi: 10.24251/hicss.2018.269.

[27] P. Tata, S. P. Devaneyan, and J. I. Sheeba, “Improved cyberbully detection techniques using multiple correlation coefficient from forum corpus,” Int. J. Auton. Comput., vol. 3, no. 2, p. 152, 2018, doi: 10.1504/ijac.2018.10018810.

[28] G. Campisi, S. Corchs, E. Fersini, F. Gasparini, and M. Mantovani, “Automatic detection of sexist content in memes,” p. 3247, 2018.

[29] S. Murnion, W. J. Buchanan, A. Smales, and G. Russell, “Machine learning and semantic analysis of in-game chat for cyberbullying,” Comput. Secur., vol. 76, pp. 197–213, 2018, doi: 10.1016/j.cose.2018.02.016.

[30] M. Garaigordobil and V. Martínez-Valderrey, “Technological resources to prevent cyberbullying during adolescence: The Cyberprogram 2.0 program and the cooperative Cybereduca 2.0 Videogame,” Front. Psychol., vol. 9, no. MAY, pp. 1–12, 2018, doi: 10.3389/fpsyg.2018.00745.

[31] D. Mario and P. G. P. Madhuri, “A Study to Learn Robust and Discriminative Representation to Tackle Cyber bullying Detection www.ijseat.com * corresponding Author,” vol. 5, no. 6, pp. 643–646, 2017.

[32] H. Dani, J. Li, and H. Liu, “Sentiment Informed Cyberbullying Detection in Social Media,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10534 LNAI, pp. 52–67, 2017, doi: 10.1007/978-3-319-71249-9_4.

[33] H. Rosa, D. Matos, R. Ribeiro, L. Coheur and J. P. Carvalho, "A “Deeper” Look at Detecting Cyberbullying in Social Networks," 2018 International Joint Conference on Neural Networks (IJCNN), 2018, pp. 1-8, doi: 10.1109/IJCNN.2018.8489211.

[34] Lee Jia Thun, Phoey Lee Teh, Chi-Bin Cheng, CyberAid: Are your children safe from cyberbullying?, Journal of King Saud University - Computer and Information Sciences, 2021, https://doi.org/10.1016/j.jksuci.2021.03.001. (https://www.sciencedirect.com/science/article/pii/S1319157821000537)

[35] Ibn Rafiq, R.; Hosseinmardi, H.; Han, R.; Lv, Q.; Mishra, S.; Mattson, S.A. Careful What You Share in Six Seconds.ss2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015—ASONAM '15; Association for Computing Machinery (ACM): New York, NY, USA, 2015; pp. 617–622.