Original Article

Detection of Cyberbullying in Social Media to Control Users' Mental Health Issues Using Recurrent Neural Network Architectures

FHA. Shibly[1], Dr. Uzzal Sharma[2], Prof. HMM. Naleer[3]

[1]PhD, Research Scholar, Assam Don Bosco University, Guwahati, India.

Senior Lecturer, South Eastern University of Sri Lanka. E-mail: shiblyfh@seu.ac.lk

[2]Assistant Professor, Department of Computer Applications, Assam Don Bosco University, Guwahati, India. E-mail: uzzal.sharma@dbuniversity.ac.in

[3]Professor & Head, Dept. of Computer Sciences, Faculty of Applied Sciences, South Eastern University of Sri Lanka. E-mail: drnaleer@seu.ac.lk

Abstract

Cyberbullying is a kind of bullying that takes place via social media with digital gadgets. It may cause significant, long-lasting trauma and difficulties with mental and physical issues. This condition may also result in difficulties in personal lives. Because of this, detecting and reporting such offensive posts may help avoid the harmful repercussions of cyberbullying. Some studies are available in this context, and it was found that deep learning algorithms are more efficient in detecting abusive texts on social media. Some studies were carried out on deep learning algorithms. The issue was not solved, and more studies are needed to find the most efficient model to control this problem. It was observed from the literature survey that the performance of Recurrent Neural Network (RNN) architectures is good. Therefore, the objective of this study was to find and develop an efficient model to detect cyberbullying texts on social media. For this research, the Twitter dataset was used, which was uploaded by Munki Albright in Kaggle. The dataset consists suspicious, cyberbullying, hate and suicidal classifications. This research was focused on cyberbullying texts. Sentiment analysis was implemented to classify the cyberbullying posts. For that, sexism (Class 2), racism (Class 1) and either (Class 0) classification were analyzed.

Researchers applied the LSTM, GRU and Bidirectional LSTM architectures of RNN to perform the sentiment analysis. All models gave somewhat similar results. When we consider f1, GRU shows a better F1 score for all the classes when compared to others (Class 0 was 95%, Class 1 – 70%, and Class 0 was 56%). GRU shows the best results for this data set compared to other models. Based on the Accuracy, all three architectures were obtained at around 90%. However, GRU has outperformed the other two models in accuracy as well. After that, researchers developed an ensemble model using all three models. In the ensemble, each model has been given weight. Since the GRU is the best performing model, it was given 0.4, and the other two models were assigned 0.3 each. The ensemble model performed well in F1 measurement (57%) and accuracy (91.17%). When we compared the previous works against the proposed model, our ensemble model obtained the highest values and performed well in detecting cyberbullying in sentiment analysis.

Keywords: Cyberbullying, Deep Learning, Recurrent Neural Network, Detection and Sentiment Analysis.

DOI: 10.47750/pnr.2022.13.S03.072

INTRODUCTION

In the rapid development of the internet era, there are many advantages and disadvantages for the people and society by using technologies. The rapid growth of information technology and social media has not gone unnoticed in promoting a variety of community-based crimes. Social media crimes known as cyberbullying have been on the rise in recent times. In research from Diganto Deb Barma et al. [1], Cyberbullying is an act of intentionally & repeatedly hurting others by using abusing & disrespecting words.

Recently, cyberbullying has become widespread on social media around the world. Such activities, which are harmful to many on social media, are increasing daily as politics, economics, sports, literature and so on are multiplied and diverse. There have been several serious incidents of cyberbullying in recent times. In recent years, cyberbullying, which may be described as "an aggressive, purposeful act carried out by a group or person utilizing electronic forms of interaction repeatedly and over time against a victim who cannot readily defend him or herself" [2], has become a significant societal issue.

Repetition, power difference, and purpose are all considered essential components in cyberbullying [3]. Cyberbullying victims had greater levels of despair and anxiety [4], suicide thoughts and attempts [5], poor academic performance [6], poor work performance [7], and poor physical and mental health [8]. Cyberbullying has more severe consequences than conventional bullying because of the larger audience on the internet and the rapid dissemination of words. Cyberbullying has more severe consequences than traditional bullying because of the larger audience on the internet [9]. Thus, cybercrime inflicts a wide range of physical and psychological damage on individuals, communities, and professionals.

Therefore, various studies are being carried out on these and their controls because their effects are very high. Cyberbullying, hate speech [10, 11], trolling [12], false news [13], rumours [14], counterfeit profile detection [15], sexism [16], and other social abnormalities are all becoming more common in recent social media-based studies. Users, on the other hand, can obscure words and phrases. Human readers can readily identify these mixed words, but an automatic system will struggle [17]. However, there are ways to control such cybercrime through artificial intelligence. Studies related to them are being carried out continuously. There is little question that significant areas of artificial intelligence, such as deep learning and machine learning, should be used to monitor and regulate these types of cybercrimes.

Our study proposes the best algorithm for automatic detection and control of cybercrime using appropriate artificial intelligence techniques based on a solid literary review, primarily through a deep learning mechanism. It is to be hoped that this study will help control the widespread use of cyberbullying crimes rampant on social media. So, our study aims to identify practical algorithms for controlling cybercrime shared on social media and use them to create a suitable classifier through ensemble technology. Researchers applied the LSTM, GRU and Bidirectional LSTM architectures of RNN to perform the sentiment analysis.

The summary of this work is structured as follows. Section 2 analyzes the relevant literature in the area; Section 3 discusses the methods we utilized to address the classification issue with dataset specifics; Section 4 offers the experiments and findings, and Section 6 draws conclusions and summarizes the current study.

RELATED WORKS

Cyberbullying is described as "a hostile, intentional act done by a group or person using internet means of contact against a defenceless victim" [18]. It is deliberate and repeated damage performed via electronic technology [19]. It is typically characterized by the posting of defamatory remarks online and the public disclosure of private information to inflict mental distress on the victim [20] intentionally.

Detecting cyberbullying has been a prominent topic in natural language processing [21]. The goal of cyberbullying detection, like other NLP jobs, is to preprocess the text and extract meaningful information so that the machine learning system can interpret and categorize each text. Traditional text classification algorithms employ a methodology for simplifying text representation followed by a machine learning classifier [22].

Natural language processing is a challenging undertaking in and of itself since it requires dealing with unclear text. Depending on the context, the same sentence might have several meanings. The task becomes much more difficult when dealing with internet material, which frequently contains misspellings, uncommon abbreviations, slang, and short words. Despite the challenges, researchers have used several machine learning algorithms to analyze emotion and attitudes [23], forecast online harassment and cyberbullying [24], crisis response and emergency scenario awareness [25], and anticipate domestic abuse crises [26].

In machine learning, Logistic regression, Nave Bayes, Support Vector Machine, Decision Tree, K-Nearest Neighbors, and Random Forest are among the most often used text categorization techniques. The success of afforested classifiers is highly reliant on feature engineering [27], even though each of these strategies may apply to various situations.

Deep learning is a relatively new phenomenon in machine learning methods that have recently slowed the development of neural networks for several years. Deep learning [28] has shown amazing breakthroughs in computer vision, pattern recognition, and image processing. Recurrent Neural Networks [29] and Convolutional Neural Networks [30] are the two most used designs for deep learning. These two approaches use text word embeddings as inputs and generate feature vectors, manipulable numerical representations. Convolutional Neural Networks have surpassed standard machine learning techniques in question categorization and sentence-level sentiment analysis.

The usage of Recurrent Neural Networks to represent text sequences in a corpus has been proven to enhance multiclass classification performance [31]. RNN variants include Long Short-Term Memory networks (LSTMs) [32], Bidirectional LSTM [33], and Gated Recurrent Units [34]. These RNN variants have improved performance in Natural Language Processing applications [35] because of their integrated memory architecture for storing long-range associations and historical data.

Based on the literature review, it is observed that the RNN architectures are a more suitable technique to detect cyberbullying. Very few studies were carried out in this context. Therefore, the researchers apply LSTM, GRU and Bidirectional LSTM models from the RNN architecture to train and test the performance over to a selected dataset and develop an ensemble classifier combining all three algorithms, which can help detect and control cyberbullying posts in social media more efficiently.

METHODS AND MATERIALS

This research includes two major phases to detect cyberbullying posts and develop an efficient classifier to identify them. Figure 1 illustrates the proposed methodology.

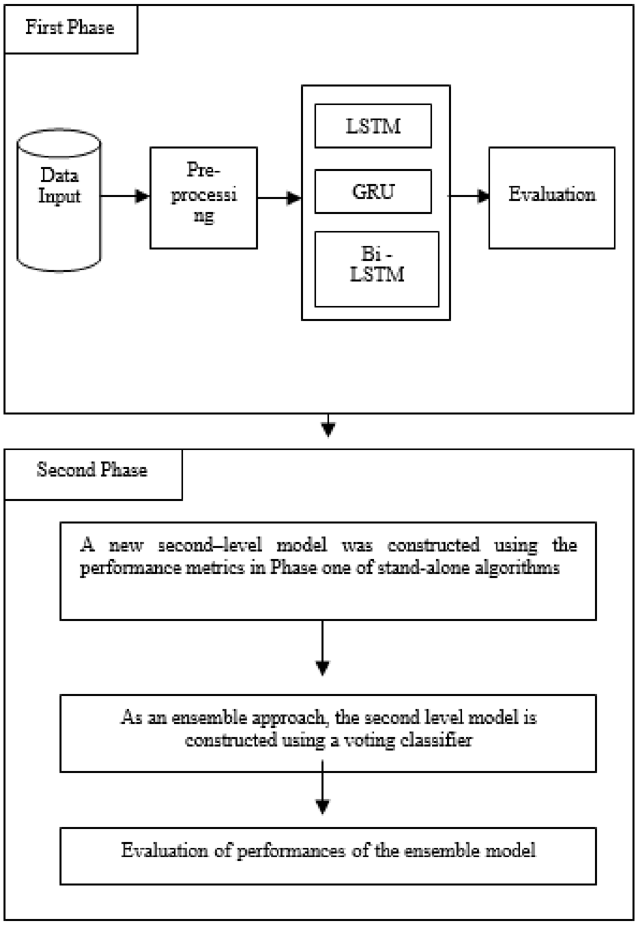

Fig. 1. Proposed Research Architecture

A. Dataset

Munki Albright uploaded the Twitter dataset on Kaggle [36]. The dataset consists of 19,934 tweets, including suspicious, cyberbullying, hate and suicidal classifications. This research was focused on cyberbullying texts. Sentiment analysis was implemented to classify the cyberbullying posts. For that, sexism (Class 2), racism (Class 1) and either (Class 0) classification were analyzed. Table 1 shows the description of the dataset in detail. This dataset is the most suitable one to carry out our research.

TABLE I. DATASET DESCRIPTION

	Suspicious	Cyberbullying	Hate	Suicidal
Count	19,934	19,934	19,934	19,934
Mean	0.361944	0.221280	0.221280	0.052573
Std	0.48.575	0.588393	0.588393	0.223186
Min	0.000000	0.000000	0.000000	0.000000
Max	1.000000	2. 000000	2. 000000	1.000000

B. Preprocessing and Feature Extraction

When a text is obtained based on Fig. 1, the data is extensively evaluated using and according to the following procedures: Stopword elimination, tokenization, sentence segmentation, and punctuation removal are some of the accessible features. These measures were used to reduce the quantity of the data, and as a consequence, we were able to delete any unnecessary information. In support of this method, we built a preprocessing program that removes punctuation and some non-letter characters from each text. Finally, the letter case of each document was decreased. This approach produced a sliced document text using an n-gram word-based tokenizer based on the length of n.

Tokenization

This approach use tokenization to deal with situations in which a given text will be separated into tokens. The objects listed below are also considered tokens. There are letters, numbers, and punctuation marks present. In addition, a nonsensical sensitive data element was swapped with a non-sensitive comparable element. We assured that the tokenization procedure was protected and tested following the strictest standards for the protection of sensitive data.

Stemming

After going through the tokenization method, the next step is to translate the tokens into another standard format. Simply said, stemming means that we may now convert the words back to their original form but with fewer word types and classes in the data. For example, "Winning," "Won," and "Winner" have been abbreviated to "win." It indicates that stemming may be used to build classes.

C. Classifiers

Long Short Term Memory

RNN, a family of artificial neural networks with node-to-node connections, has severe disadvantages due to numerous network layers. A recent study has found that the LSTM network offers a solution to this problem due to its chain-like topology, which is similar to that of many neural network modules in RNN. Figure 2 displays the LSTM architecture, which comprises many gates, including the input gate, the output gate, and a forget gate employed in the LSTM model [37].

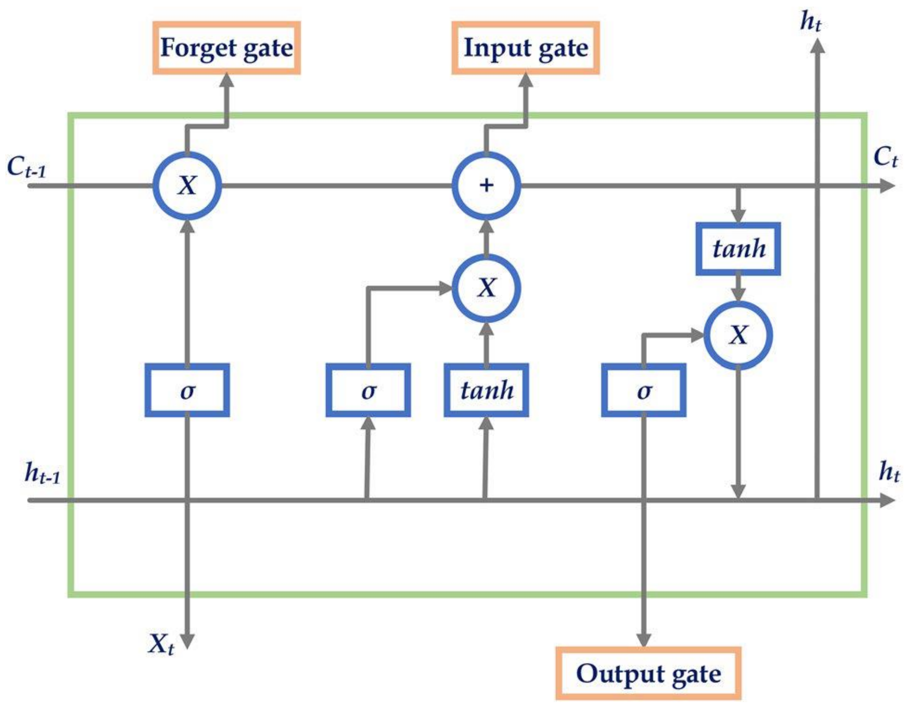

Fig. 2. LSTM Architecture

Gated recurrent unit (GRU)

GRU is the most current RNN variant designed to overcome similar issues with short-term memory as LSTM. Note that GRU lacks a cell state and instead transmits information via a hidden state. Given that GRU has fewer gates than LSTM, which speeds up the training process [37], the reset gate determines how much of the past data is forgotten. It consists of two gates: a reset gate, represented by rt, and an update gate, represented by zt and Eqs. (1). (2).

$$Z^t = \sigma\, (W_{z^t} x^t + U_{z^t}\, h_{t-1} + b_{z^t} \quad (1)$$

$$r^t = \sigma\, (W_{r^t} x^t + U_{r^t}\, h_{t-1} + b_{r^t} \quad (2)$$

Bidirectional Long Short Term Memory (Bi - LSTM)

The bidirectional LSTM (BLSTM) model maintains and advances two distinct input states given by two distinct LSTMs. The first LSTM is a typical sequence beginning at the beginning of the paragraph, whereas the second LSTM is a typical sequence with the inputs in reverse order. The purpose of the bidirectional network is to collect information about the inputs around it. It often learns more quickly than a one-way technique, although this relies on the objective, as seen in Fig. 3, which illustrates the constraint of the Bi - LSTM model [37].

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding (Embedding)        (None, None, 300)         4569300
_________________________________________________________________
bidirectional (Bidirectional (None, None, 128)         186880
_________________________________________________________________
bidirectional_1 (Bidirection (None, 128)               98816
_________________________________________________________________
dense (Dense)                (None, 3)                 387
=================================================================
Total params: 4,855,383
Trainable params: 286,083
Non-trainable params: 4,569,300
```

Fig. 3. Bi - LSTM Architecture

Evaluation

The findings of the experiments are reported in this part, as suggested by the authors in [38, 39, 40]. Google Colab, Google's online Graphical Processing Unit, is used for the project (GPU). We used Python 3.7 as our programming language in this study, and an Intel Core i5 8265 CPU @ 1.60 GHz computer with a Windows 10 pro system and CPU were used to train and effectively test the algorithms.

In data mining and data retrieval, evaluating the accuracy of selected classifiers is one of the most critical phases. Error rate and F-measure are widely used to determine the accuracy of a classifier's ability to locate the proper category or class of unknown causes. The error rate is the instances of the test set that were erroneously categorized. We will call this set of data "X" and let "m" represent how many occurrences were misclassified by a classification model C. You can calculate the accuracy of C in selecting the correct classes of X instances using the following formula:

$$Accuracy\ (m) = \frac{m}{n} \quad (3)$$

The error rate approach ignores the cost of inaccurate predictions in ML. For the most part, F-measure is used to solve this problem. To determine the value of the F-measure, two basic metrics are used: precision and recall. Imagine that some of the data in the test set belong to a particular class or category S. It assigns a category label to each test data. There will be four kinds of forecasts for the test set S:

The percentage of accurately forecast data for category S is known as precision. The percentage of correctly forecasted accurate data for category S is called recall. It is possible to calculate the F-measure based on precision and recall (4-6).

Precision= TP/(TP+FP) (4)

Recall = TP/(TP+FN) (5)

$$F - \text{Measure} = 2\ X\ \frac{\text{Precision X Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Results and Discussions

In this analysis, all algorithms were tested by using precision, recall, F measures and accuracy. Based on the dataset, Hate Speech (HS), Offensive Languages (OL) and No Hate Speech (N) classifications were tested.

Table II shows that LSTM testing accuracy is 90.74%, while testing loss is 0.27 after 20 epochs. Similarly, GRU and Bi-LSTM achieved 90.82% and 90.66% testing accuracy. Precision, recall and F1 measures for all three classes are shown in Table II.

TABLE II. Parameter evaluation

Algorithms	Class	Parameter Evaluation			
		Precision	Recall	F1	Accuracy
LSTM	0	0.93	0.97	0.95	90.74%
	1	0.68	0.74	0.71	
	2	0.72	0.42	0.53	
GRU	0	0.93	0.96	0.95	90.82%
	1	0.69	0.71	0.70	
	2	0.71	0.47	0.36	
Bi LSTM	0	0.93	0.97	0.95	90.66%
	1	0.69	0.74	0.71	
	2	0.71	0.41	0.52	

Based on the results in the above Table, GRU has outperformed the other two algorithms in terms of accuracy. However, when we checked the precision, recall and F1 metrics, all three algorithms performed well, and there was a narrow margin.

Receiver Operating Characteristic (ROC) curve for STM is represented in Fig. 4. GRU and Bi – LSTM ROC curves are represented in Fig. 5 & 6. It seems GRU and Bi – LSTM are performed well than LSTM. However, all values are very marginal. So, it can be observed that all the algorithms work better. So, such algorithms better race the cyberbullying-related posts widely shared on social media. This further demonstrates the fact that RNN algorithms can work better.

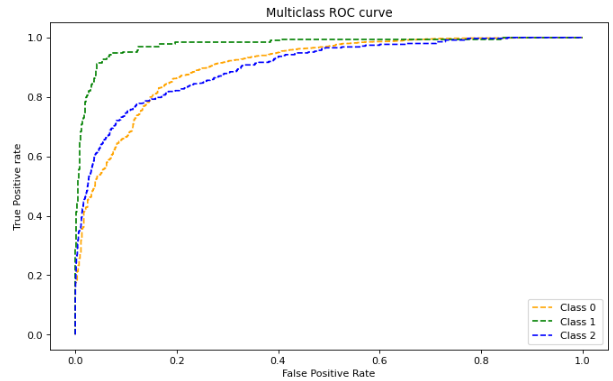

Fig. 4. ROC of LSTM Architecture

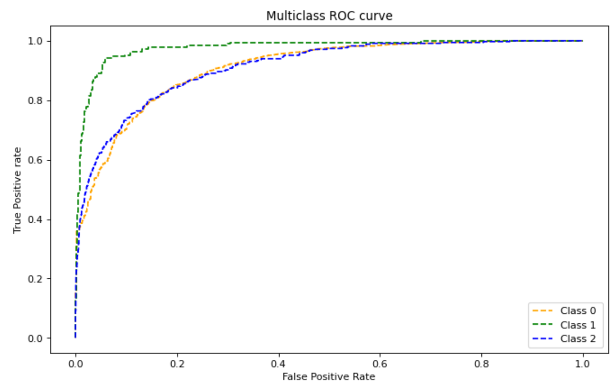

Fig. 5. ROC of GRU

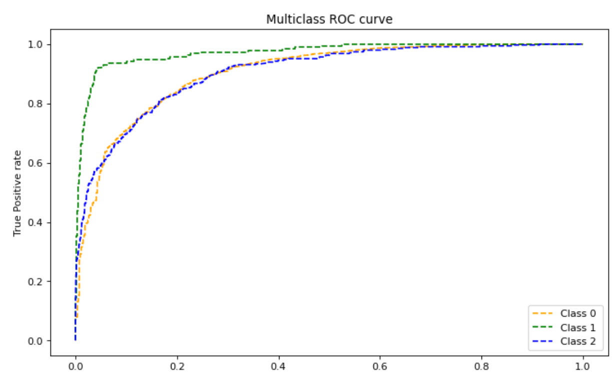

Fig. 6. ROC of Bi - LSTM

Ensemble model

After phase one, researchers developed an ensemble model using all three models. In the ensemble classification, each model has been given weight. Since the GRU is the best performing model, it was given 0.4 weight, and the other two models were assigned 0.3 weight. Here the reason for assigning more weight to the GRU model is that it is the model which gave the highest f1, precision and recall for class 3. Other models gave fewer values for those matrices for class 3. The ensemble model could increase the f1 score for class 3. The evaluation of the ensemble model is given in Table III:

TABLE III. PARAMETER EVALUATION FOR ENSEMBLE MODEL

Algorithms	Class	Parameter Evaluation			
		Precision	Recall	F1	Accuracy
Ensemble Model	0	0.93	0.97	0.95	91.17%
	1	0.68	0.72	0.70	
	2	0.78	0.45	0.57	

It clearly shows that the proposed model has the highest values in all metrics, including accuracy (91.17%). When we check class 3, the proposed model has performed well in precision (0.78). Recall (0.45) and F measures (0.57), and it can calculate cyberbullying posts more efficiently than other stand-alone classifiers (Table III).

According to the proposed model's confusion matrix (Figure 7), the model correctly identified the number of positive and negative class data points. Additionally, the model mistakenly categorized a small number of negative class data points as belonging to the positive class and a small number of positive class data points as belonging to the negative class. It demonstrates that the suggested methodology effectively detects hate speech against women.

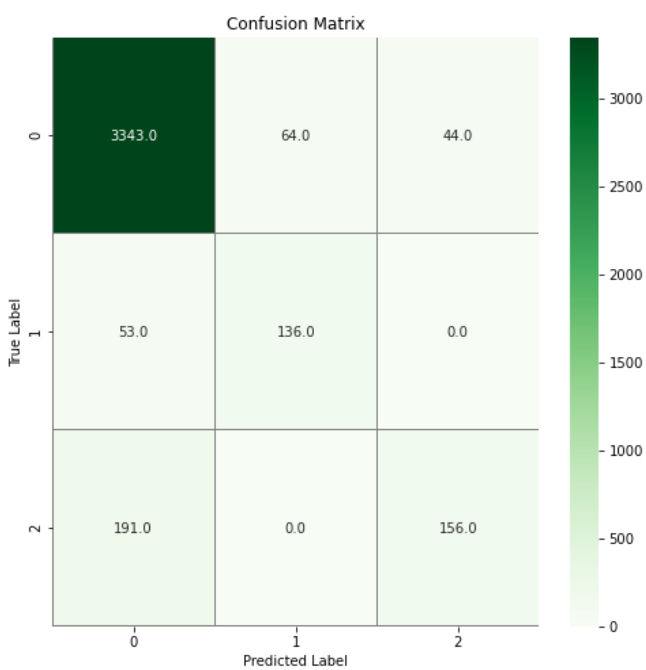

Fig. 7. Confusion Matrix of Ensemble Model

Comparative Analysis

The researchers have developed an ensemble model with 91.17% accuracy in detecting cyberbullying content in social media using RNN architectures. A comparison analysis must be applied between the proposed model and existing models. Therefore, the researchers conducted a comparative analysis and tabulated it in Table IV.

TABLE IV. COMPARATIVE ANALYSIS

Paper ID	Authors & Year	Model	Accuracy
38	Iwendi, C.(2020)	BiLSTM	82.18%
41	Alotaibi, M. et.al (2021)	CNN+BiGRU	87.99%
42	Singh, Ravinder.et.al (2020)	RNNs	76.36%
43	Van Bruwaene, D. et.al (2020)	XGBoost_LIWC	88.10%
44	N. Yuvaraj et al (2021)	ANN + DRL	80.70%
45	R. Zhao (2017)	SmSDA	84%
46	A. T. Aind (2020)	Reinforcement Learning	89%
	Our Model (2022)	LSTM+GRU+BiLSTM	91.17%

Researchers compared the similar type of research works and performances along with the cyberbullying datasets. It is observed that there are limited research works available related to the detection of cyberbullying in social media. However, some remarkable works have been carried out, and some authors have produced good results recently. Machine learning algorithms and deep learning algorithms were used to tackle this issue. According to Table IV, our proposed model has obtained the highest accuracy with 91.17% and outperformed existing models within the research context.

Conclusion and Recommendations

The range of the fastest growing social media is vast and widespread. There is no denying that the independence of social media has many advantages and disadvantages. In recent times, there have been various incidents that have mainly affected one's posts and comments to another extent or a community. While social media is responsible for controlling this, researchers are also paying close attention to it. Thus, our study forms one of the studies that find mechanisms to identify cyberbullying crimes on social media and control them.

In this study, a comparison is made between our unique technique and the most recent research on employing deep learning-based models to detect cyberbullying episodes. Our suggested LSTM employs twice as many input gates, output gates, and forget gates as the conventional LSTM. This accomplishment improves your accuracy. However, our suggested approach has a higher computational complexity and performance cost. This article focuses on fixing the constraints of prior experiments with improved identification efficiency when compared directly to standard versions. In this study, a comparison is made between our unique technique and the most recent research on employing deep learning-based models to detect cyberbullying episodes. Our suggested LSTM employs twice as many input gates, output gates, and forget gates as the conventional LSTM. This accomplishment improves our accuracy. However, our suggested approach has a higher computational complexity and performance cost. This article focuses on fixing the constraints of prior experiments with improved identification efficiency when compared directly to standard versions.

In addition, an empirical evaluation was conducted to establish the efficacy and performance of deep learning algorithms in identifying insults in Social Commentary. Experiments employ four deep learning models: LSTM, GRU, and Bi - LSTM. As data are preprocessing steps, text cleaning, tokenization, stemming, and lemmatization are applied to delete and prevent phrases in the communication chain from reaching trusting individuals. The data from the preprocessing stage is then sent via textual data that has been cleaned and directly into deep learning prediction algorithms. The Twitter dataset published by Munki Albright to Kaggle was utilized. The dataset contains suspicious, cyberbullying, hatred, and suicide behaviour categories. This study focused on cyberbullying posts.

This research was carried out in two major phases. The first phase is to evaluate the stand–alone performances in detecting cyberbullying posts. In the second phase, an ensemble model was developed based on the performance of all three algorithms in phase one. Researchers can conclude that the GRU model achieved higher accuracy and F1-measure scores than LSTM and Bi - LSTM. All models gave somewhat similar results. When we consider f1, GRU shows a better F1 score for all the classes compared to others (Class 0 was 95%, Class 1 – 70%, and Class 0 was 56%). GRU shows the best results for this data set compared to other models. Based on the Accuracy, all three architectures were obtained at around 90%.

Nevertheless, GRU has outperformed the other two models in accuracy as well. After that, researchers developed an ensemble model using all three models. In the ensemble, each model has been given weight. Since the GRU is the best performing model, it was given 0.4, and the other two models were assigned 0.3 each. The ensemble model performed well in F1 measurement (57%) and accuracy (91.17%). When we compared the previous works against the proposed model, our ensemble model obtained the highest values and performed well in detecting cyberbullying in sentiment analysis. Therefore, our model can be incorporated into social media applications to control cyberbullying, limiting so many unnecessary personal and social issues aligned with physical and mental trauma.

REFERENCES

Diganto deb barma, dwvedi, v., & ballal, a. (2020). Cyberbullying detection prevention on social networks. Int. J. Adv. Res., 8(7), 87–90.

Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. Journal of Child Psychology and Psychiatry, 49(4), 376–385.

C. Langos, "Cyberbullying: The challenge to define," Cyberpsychol. Behav. Soc. Netw., vol. 15, no. 6, pp. 285-289, 2012.

D. Zhang, E. S. Huebner, and L. Tian, "Longitudinal associations among neuroticism, depression, and cyberbullying in early adolescents," Comput. Hum. Behav., vol. 112, p. 106475, 2020.

M. C. Martínez-Monteagudo, B. Delgado, Á. Díaz- Herrero, and J. M. García-Fernández, "Relationship between suicidal thinking, anxiety, depression and stress in university students who are victims of cyberbullying," Psychiatry Res., vol. 286, p. 112856, 2020.

C. E. Torres, S. J. D'Alessio, and L. Stolzenberg, "The effect of social, verbal, physical, and cyberbullying victimization on academic performance," Vict. Offender, vol. 15, no. 1, pp. 1-21, 2020.

A. Oksanen, R. Oksa, N. Savela, M. Kaakinen, and N. Ellonen, "Cyberbullying victimization at work: Social media identity bubble approach," Comput. Hum. Behav., vol. 109, p. 106363, 2020, doi: 10.1016/j.chb.2020.106363.

T. Vaillancourt, R. Faris, and F. Mishna, "Cyberbullying in children and youth: implications for health and clinical practice," Can. J. Psychiatry, vol. 62, no. 6, pp. 368-373, 2017.

Arif, M. (2021). A systematic review of machine learning algorithms in cyberbullying detection: future directions and challenges. Journal of Information Security and Cybercrimes Research, 4(1), 01-26.

P. Fortuna and S. Nunes, "A survey on automatic detection of hate speech in text," ACM Comput. Surv. (CSUR), vol. 51, no. 4, pp. 1-30, 2018.

S. MacAvaney, H.-R. Yao, E. Yang, K. Russell, N. Goharian, and O. Frieder, "Hate speech detection: Challenges and solutions," PLOS One, vol. 14, no. 8, p. e0221152, 2019.

M. Tomaiuolo, G. Lombardo, M. Mordonini, S. Cagnoni, and A. Poggi, "A survey on troll detection," Future Internet, vol. 12, no. 2, p. 31, 2020.

T. Jiang, J. P. Li, A. U. Haq, A. Saboor, and A. Ali, "A Novel Stacking Approach for Accurate Detection of Fake News," IEEE Access, vol. 9, pp. 22626-22639, 2021.

A. Kumar, V. Singh, T. Ali, S. Pal, and J. Singh, "Empirical evaluation of shallow and deep classifiers for rumor detection," in Proc. ICACM 2019, in Advances in Computing and Intelligent Systems, in Algorithms for Intelligent Systems, 2020, pp. 239-252.

S. R. Sahoo and B. Gupta, "Real-time detection of fake account in twitter using machine-learning approach," in Proc. CICT 2019, in Advances in computational intelligence and communication technology, in Advances in Intelligent Systems and Computing, vol. 1086, 2021, pp. 149-159.

E. W. Pamungkas, V. Basile, and V. Patti, "Misogyny detection in twitter: a multilingual and cross-domain study," Inf. Process. Manag., vol. 57, no. 6, p. 102360, 2020.

H.-S. Lee, H.-R. Lee, J.-U. Park, and Y.-S. Han, "An abusive text detection system based on enhanced abusive and non-abusive word lists," Decis. Support Syst., vol. 113, pp. 22-31, 2018.

Peter K. Smith, Jess Mahdavi, Manuel Carvalho, Sonja Fisher, Shanette Russell, and Neil Tippett. 2008. Cyberbullying: Its nature and impact in secondary school pupils. Journal of Child Psychology and Psychiatry and Allied Disciplines, 49:376–385.

Sameer Hinduja and Justin W. Patchin. 2008. Cyberbullying: An exploratory analysis of factors related to offending and victimization. Deviant Behavior, 29:129–156.

N. Willard. 2005. An educator's guide to cyberbullying and cyberthreats: Responding to the challenge of online social aggression, threats, and distress.

Jacob, S.S.; Vijayakumar, R. Sentimental analysis over twitter data using clustering based machine learning algorithm. J. Ambient Intell. Human. Comput. 2021, 1–12.

Sadiq, S.; Mehmood, A.; Ullah, S.; Ahmad, M.; Choi, G.S.; On, B.W. Aggression detection through deep neural model on Twitter. Future Gener. Comput. Syst. 2021, 114, 120–129.

M. S. Neethu and R. Rajasree, ''Sentiment analysis in Twitter using machine learning techniques,'' in Proc. 4th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2013, pp. 1–5.

E. V. Altay and B. Alatas, ''Detection of cyberbullying in social networks using machine learning methods,'' in Proc. Int. Congr. Big Data, Deep Learn. Fighting Cyber Terrorism (IBIGDELFT), Dec. 2018, pp. 87–91.

M. A. Cameron, R. Power, B. Robinson, and J. Yin, â€˜â€˜Emergency situation awareness from Twitter for crisis management,â€™â€™ in Proc. 21st Int. Conf. World Wide Web, 2012, pp. 695â€“698.

S. Subramani, H. Wang, H. Q. Vu, and G. Li, ''Domestic violence crisis identification from Facebook posts based on deep learning,'' IEEE Access, vol. 6, pp. 54075–54085, 2018,

doi: 10.1109/access.2018.2871446.

Z.-G. Chen, Z.-H. Zhan, H. Wang, and J. Zhang, ''Distributed individuals for multiple peaks: A novel differential evolution for multimodal optimization problems,'' IEEE Trans. Evol. Comput., vol. 24, no. 4, pp. 708–719, Aug. 2020.

Y. LeCun, Y. Bengio, and G. Hinton, ''Deep learning,'' Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

I. Sutskever, J. Martens, and G. E. Hinton, ''Generating text with recurrent neural networks,'' in Proc. 28th Int. Conf. Mach. Learn., 2011, pp. 1017–1024.

Y. Kim, ''Convolutional neural networks for sentence classification,'' 2014, arXiv:1408.5882.

[Online]. Available: http://arxiv.org/abs/1408. 5882

P. Liu, X. Qiu, and X. Huang, ''Recurrent neural network for text classification with multi-task learning,'' 2016, arXiv:1605.05101. [Online]. Available: https://arxiv.org/abs/1605.05101

A. Graves, ''Generating sequences with recurrent neural networks,'' 2013, arXiv:1308.0850.

[Online]. Available: http://arxiv.org/abs/1308.0850

A. Graves, N. Jaitly, and A.-R. Mohamed, ''Hybrid speech recognition with deep bidirectional LSTM,'' in Proc. IEEE Workshop Autom. Speech Recognit. Understand., Dec. 2013, pp. 273–278.

K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, ''On the properties of neural machine translation: Encoder-decoder approaches,'' 2014, arXiv:1409.1259.

[Online]. Available: http://arxiv.org/abs/1409.1259

A. Graves, S. Fernández, and J. Schmidhuber, ''Bidirectional LSTM networks for improved phoneme classification and recognition,'' in Proc. Int. Conf. Artif. Neural Netw., 2005, pp. 799–804.

Munki Albright. (2021, August 13). Classified Tweets [Dataset]. https://www.kaggle.com/datasets/munkialbright/classified-tweets

Iwendi, C., Srivastava, G., Khan, S., & Maddikunta, P. K. R. (2020). Cyberbullying detection solutions based on deep learning architectures. Multimedia Systems, 1-14.

Iwendi, C., Jalil, Z., Javed, A.R., Reddy, T., Kaluri, R., Srivastava, G., Jo, O.: Keysplitwatermark: zero watermarking algorithm for software protection against cyber-attacks. IEEE Access 8, 72650–72660 (2020)

Kumar, N.: Analyzing the effectiveness and contribution of each axis of tri-axial accelerometer sensor for accurate activity recognition. Sensors 20(8), 2216 (2020)

Mittal, M., Iwendi, C., Khan, S., Rehman Javed, A.: Analysis of security and energy efficiency for shortest route discovery in lowenergy adaptive clustering hierarchy protocol using levenbergmarquardt neural network and gated recurrent unit for intrusion detection system. Trans. Emerg. Telecommun. Technol. (2020). https ://doi.org/10.1002/ett.3997

Alotaibi, M.; Alotaibi, B.; Razaque, A. A Multichannel Deep Learning Framework for Cyberbullying Detection on Social Media. Preprints 2021, 1, 0. https://doi.org/

Singh, Ravinder, Subramani, Sudha, Du, Jiahua, Zhang, Yanchun, Wang, Hua, Ahmed, Khandakar and Chen, Zhenxiang (2020) Deep Learning for Multi-Class Antisocial Behavior Identification From Twitter. IEEE Access, 8. pp. 194027-194044. ISSN 2169-3536

Van Bruwaene, D., Huang, Q., & Inkpen, D. (2020). A multi-platform dataset for detecting cyberbullying in social media. Language Resources and Evaluation, 54(4), 851-874.

N. Yuvaraj et al., "Nature-Inspired-Based Approach for Automated Cyberbullying Classification on Multimedia Social Networking," Math. Probl. Eng., vol. 2021, 2021, doi: 10.1155/2021/6644652.

R. Zhao and K. Mao, "Cyberbullying detection based on semantic-enhanced marginalized denoising autoencoder," IEEE Trans. Affect. Comput., vol. 8, no. 3, pp. 328-339, 1 July-Sept. 2017, doi: 10.1109/TAFFC.2016.2531682.

A. T. Aind, A. Ramnaney, and D. Sethia, "Q-Bully: A Reinforcement Learning based Cyberbullying Detection Framework," in 2020 Int. Conf. Emerg. Technol. (INCET), India, June 5-7, 2020, pp. 1-6, doi: 10.1109/INCET49848.2020.9154092.

Ibrahim, S. (2022). DISCRETE LEAST SQUARE METHOD FOR SOLVING DIFFERENTIAL EQUATIONS. Advances and Applications in Discrete Mathematics, 30, 87-102.