Computers, Materials & Continua
DOI: 10.32604/cmc.2022.024488
Article

Tech Science Press

Optimal Deep Learning-based Cyberattack Detection and Classification Technique on Social Networks

Amani Abdulrahman Albraikan[1], Siwar Ben Haj Hassine[2], Suliman Mohamed Fati[3], Fahd N. Al-Wesabi[2,4], Anwer Mustafa Hilal[5,*], Abdelwahed Motwakel[5], Manar Ahmed Hamza[5] and Mesfer Al Duhayyim[6]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, 11671, Saudi Arabia
[2]Department of Computer Science, College of Science and Arts, King Khalid University, Mahayil Asir, Saudi Arabia
[3]Department of Information Systems, College of Computer and Information Sciences, Prince Sultan University, Saudi Arabia
[4]Faculty of Computer and IT, Sana'a University, Sana'a, Yemen
[5]Department of Computer and Self-Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia
[6]Department of Natural and Applied Sciences, College of Community-Aflaj, Prince Sattam bin Abdulaziz University, Saudi Arabia
*Corresponding Author: Anwer Mustafa Hilal. Email: a.hilal@psau.edu.sa

Received: 19 October 2021; Accepted: 29 December 2021

Abstract: Cyberbullying (CB) is a distressing online behavior that disturbs mental health significantly. Earlier studies have employed statistical and Machine Learning (ML) techniques for CB detection. With this motivation, the current paper presents an Optimal Deep Learning-based Cyberbullying Detection and Classification (ODL-CDC) technique for CB detection in social networks. The proposed ODL-CDC technique involves different processes such as pre-processing, prediction, and hyperparameter optimization. In addition, GloVe approach is employed in the generation of word embedding. Besides, the pre-processed data is fed into Bidirectional Gated Recurrent Neural Network (BiGRNN) model for prediction. Moreover, hyperparameter tuning of BiGRNN model is carried out with the help of Search and Rescue Optimization (SRO) algorithm. In order to validate the improved classification performance of ODL-CDC technique, a comprehensive experimental analysis was carried out upon benchmark dataset and the results were inspected under varying aspects. A detailed comparative study portrayed the superiority of the proposed ODL-CDC technique over recent techniques, in terms of performance, with the maximum accuracy of 92.45%.

Keywords: Cybersecurity; cyberbullying; social networks; parameter tuning; deep learning; metaheuristics

1 Introduction

There has been a tremendous increase observed in online activities in the recent years among teenagers. This phenomenon is prevalently found in social media platforms that consistently expose teenagers to Cyber Bullying (CB). Bullying can be determined as an aggressive and intentional behaviour/act which is performed repeatedly by an individual or a group against victim(s) who cannot protect themselves over a period of time [1]. Cyberbullying has been documented as a national health problem and is recognized as one of the main mental health problems owing to its continuous increase in social media and other online modes of communication [2]. The negative effect of cyberbullying shares multiple similarities with conventional bullying. However, cyberbullying can be increasingly dangerous due to high propagation and frequency empowered by technology at the same time [3]. Surveys conducted earlier have related cyberbullying with negative impacts on academic performance, high risk of suicidal ideation, depression psychological and physical health [4]. So, early recognition of cyberbullying in social networking sites is paramount to reduce and mitigate its negative impacts upon the victim. Furthermore, the repetitive nature of cyberbullying makes it highly challenging to terminate and detect immediately. Cyber aggression, on one side, helps in detecting the aggressor and on the other side, assist the victim.

From the perspectives of cyber world, the applications with CB pose significant challenges in terms of lack of straightforward communications, consequences associated over others, self and victim's identity, and the ignorance of aggressors [5]. The failure in straightforward communications disrupts the nature of message and partial interpretation leads to confusion on the interaction messages. Further, the intention of the individual with exchange of messages becomes complicated. Though there are difficulties in identification, when detecting an individual's behavioral intention, the main reason behind the transition of a harasser from aggression to CB is their intent to harm oneself [6]. At present, authors have attempted to innovate technologies for resolving real-time challenges in cyberbullying. Nowadays, the social networking platforms are developed with an automatic behaviour that alerts the moderator to analyze the content [7]. But the present architecture lacks an intelligent automation process that can alter and detect the CB content quickly with precision, when compared to conventional reporting method.

In case of an alert, the moderators respond and take the required action whereas alerts are taken against the user content [8]. This study was conventionally performed on the surveyed data/accessible datasets, in which the victims/perpetrators report the impression. Highly accurate and automated recognition process is one another challenge faced in the mitigation of cyber bullying from the accessible dataset. Hence, the difficulty in recognizing activity must be improved [9]. The need for advanced tools that can incorporate the features with automatic decision method is increasing. Several studies conducted upon smart CB detection used machine learning methods as well as adapted psychological and common features. However, this smart system is limited one, confined to the comments of an individual, leaving behind their context [10]. Another research work reported about the usage of user contexts in action which includes the history of a user's comments and their characteristics to enhance CB classification or detection accuracy. Currently, the studies conducted earlier have established novel technologies for automatic recognition of CB with real-time challenges.

Ptaszynski et al. [11] proposed a new approach for automated recognition of cyberbullying entries on internet. In this model, the researchers utilized seed words from three classes to analyze a semantic orientation score and later maximized the significance of classes. The presented model outperformed benchmark settings in real world and laboratory conditions. In practice, the proposed method was tested and deployed. Bozyiğit et al. [12] developed the significance of social networking features in cyberbullying recognition. In this study, a balanced dataset was prepared that consists of 5,000 labelled contents from various social networking attributes. Next, the relationships amid cyberbullying and social networking attributes were investigated using chi-square test. For example, a user with more followers on social media are disinclined towards posting online bullying content. Consequently, ML methods were tested on two distinct kinds of the prepared dataset.

Chia et al. [13] clarified and reviewed the description of sarcasm and irony by discussing several research terms. Then, the researchers conducted early experiments by comparing different kinds of classification algorithms that involve some common classifiers for text classification process. Next, various kinds of data pre-processing models were analyzed and compared. Lastly, the relationships among cyberbullying, irony, and sarcasm were discussed. Eronen et al. [14] conducted a study using numerous datasets including common datasets like Yelp business review datasets which were utilized to train conventional SA method. Newer datasets attempted to address the problems of cyberbullying since it is a crucial social problem and a complicated problem from the perspective of linguistic representations.

Balakrishnan et al. [15] proposed an automated cyberbullying recognition method to tap the psychological characteristics of twitter users involving sentiment, emotion, and personalities. In this study, user personality has been defined by Big Five and Dark Triad methods, where ML classifiers like NB, RF, and J48 were utilized in the classification of twitters under four classes namely, normal, bully, aggressor, and spammer. This twitter dataset contained 5,453 tweets with the hashtag #Gamergate and was automatically annotated by human expertise.

1.1 Objectives

The major objective of this study is to develop Optimal Deep Learning based Cyberbullying Detection and Classification (ODL-CDC) technique for effective identification and classification of different kinds of cyberbullying that exist in social networking sites.

1.2 Paper Contribution

The current research work presents an Optimal Deep Learning based Cyberbullying Detection and Classification (ODL-CDC) technique for identifying cyberbullying on social networks. The proposed ODL-CDC technique involves different processes such as pre-processing, prediction, and hyperparameter optimization. In addition, GloVe approach is also employed for the generation of word embedding. Besides, the pre-processed data is fed into Bidirectional Gated Recurrent Neural Network (BiGRNN) model for prediction process. Moreover, hyperparameter tuning of BiGRNN model is carried out with the help of Search and Rescue Optimization (SRO) algorithm. In order to validate the improved classification performance of ODL-CDC technique, a comprehensive experimental analysis was carried out on benchmark dataset and the results were inspected under varying aspects.

1.3 Paper Organization

Rest of the sections in this paper is arranged as follows. Section 2 elaborates the proposed ODL-CDC technique while the simulation results are discussed in Section 3. Finally, the concluding remarks are identified in Section 4.

2 The Proposed Model

In this study, a novel ODL-CDC technique is developed to identify the presence of cyberbullying in social networking websites. The proposed ODL-CDC technique comprises of pre-processing, Glove-based word embedding, BiGRNN-based prediction, and SRO-based hyperparameter optimization. SRO algorithm is utilized for proper hyperparameter tuning of BiGRNN parameters which helps in improving the classification performance substantially. Fig. 1 illustrates the overall process of ODL-CDC model.

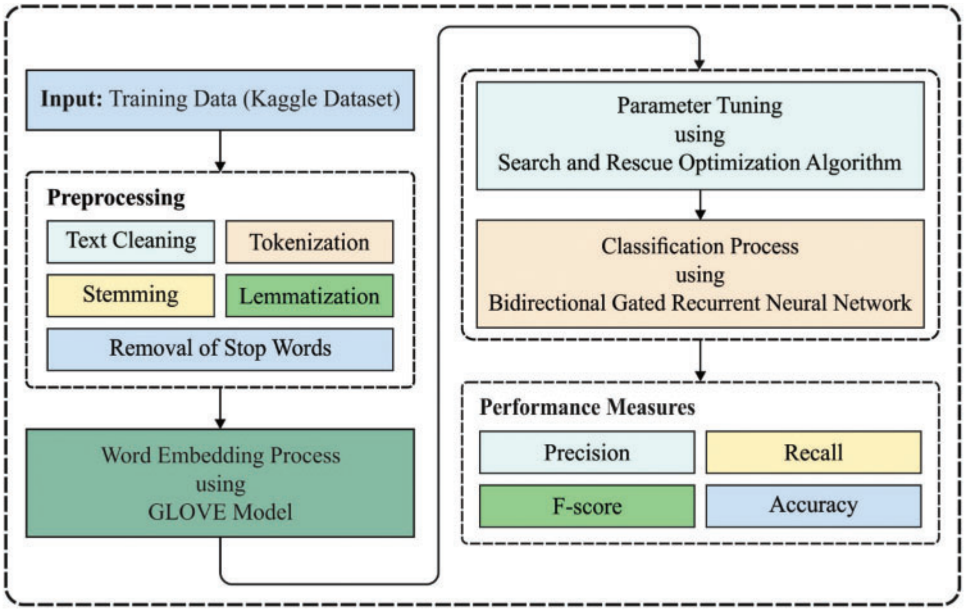

Figure 1: Overall process of ODL-CDC model

2.1 *Pre-processing*

In this primary stage, social networking data is pre-processed under different stages [16] such as text cleaning, tokenization, stemming, lemmatization, and stop word removal.

2.2 *Glove's Word Embedding Technique*

GloVe word-embedded model is utilized in the presented technique to extract semantic features in text-on-web page. GloVe is an abbreviation of Global Vectors while GloVe embedded model has unsupervised learning technique to distributed word demonstration of the text removed in this web page. GloVe technique is easy to train the information because of its parallel execution method. It takes the semantic connections of words from vector spaces. Global co-occurrence matrix X is generated based on the words established in Wikipedia dataset to train GloVe word-embedded technique [17].

During co-occurrence matrix, X: X_{ij} implies the amount of context words i that act as word, j. GloVe technique minimizes the subsequent main function.

$$I = Minimize \sum_{i=1}^{V} \sum_{j=1}^{V} (W_i^t W_j + b_i + b_j - \mathrm{logx_{ij}}) \tag{1}$$

2.3 BiGRNN-Based Prediction

At this stage, the social networking data is classified into normal and insult by BiGRNN model. GRU-NN is a basic form of LSTM while the procedure mimics RNN. In contrast with LSTM, GRU has a combination of input as well as forgetting gates to upgrade gate. Its fundamental framework is demonstrated in Fig. 2.

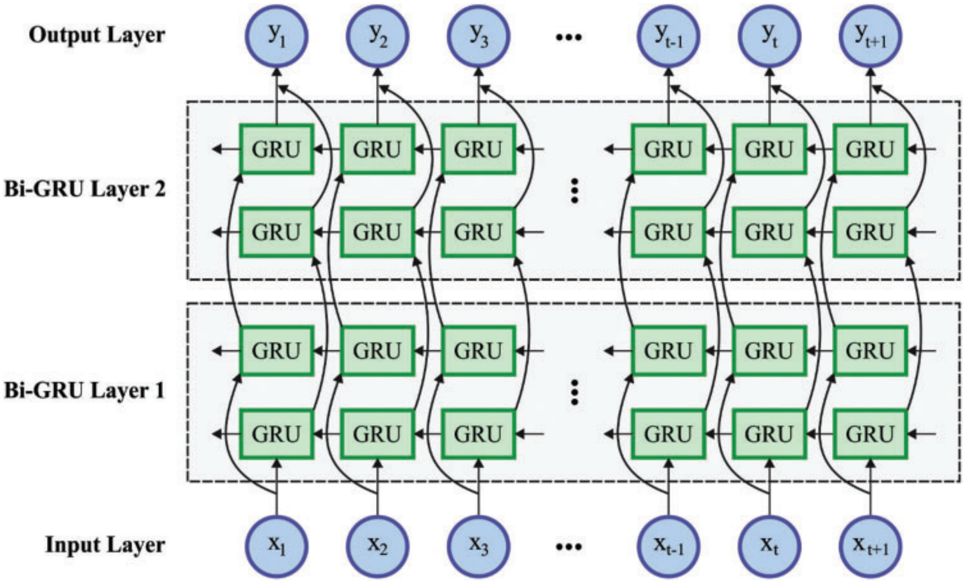

Figure 2: BiGRU structure

Considering that the amount of hidden units is h, a small-batch input with provided time step of t is $X_t \in \mathbb{R}^{n*d}$ (the amount of instances is n, the amount of inputs is d), and the hidden state at preceding time step t1 is $H_{t-1} \in \mathbb{R}^{n*h}$. The resultant hidden state h of single GRU at present time step t is as follows [18]:

$$R_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r) \tag{2}$$

$$Z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \tag{3}$$

$$\bar{H} = \tan h(X_t W_{xh} + (R_t E \odot H_{t-1}) W_{hh} + b_h) \tag{4}$$

$$H_t = (1 - Z_t) \odot H_{t-1} + Z_t E \odot \tilde{H}_t \tag{5}$$

where σ refers to sigmoid activation function, for instance, $\sigma(x) = 1/1 + e^{-x}, \cdot W_{xr}$, W_{hr}, W_{XZ}, and W_{hz} demonstrate the weights of linking input layer and reset gate, hidden layer and reset gate, input layer and update gate, and hidden layer and update gate correspondingly; b_r and b_z signify the bias of reset as well as update gates; H_t implies the candidate hiding state at present time step, t; $\odot$ stands for matrix multiplication of two elements whereas *Tanh* represents the hyperbolic tangent activation function for which the equation is as follows.

$$\tanh(x) = 1 - \frac{2}{1 + e^{-2x}} \tag{6}$$

If water quality time series parameter is forecasted, the value of present time is nearly connected with the values of preceding time and next time. But, GRU is a one-way NN framework due to which BiGRU is utilized. BiGRU is a bidirectional NN which is a collection of collected of forward-propagating GRU and backward-propagating GRU units. The present hidden layer state H_t of BiGRU is defined as present input X_t, the output $\vec{H}$ of forwarding hidden layer, and the output $\overleftarrow{\mathrm{H}}_t$ of backward hidden layer at time step $t - 1$, afterward,

$$\vec{H}_{\mathrm{t}} = GRU\left(X_t, \vec{H}_{\mathrm{t-1}}\right) \tag{7}$$

$$\overleftarrow{\mathrm{H}}_t = GRU\left(X_t, \overleftarrow{\mathrm{H}}_{\mathrm{t-1}}\right) \tag{8}$$

$$H_t = w_t\vec{H}_{\mathrm{t}} + v_t\overleftarrow{\mathrm{H}}_t\ + b_t \tag{9}$$

where GRU (.) function designates that GRU network is utilized for conducting non-linear alteration on input data of water quality whereas the input vector encodes to equivalent GRU hidden state; w_t and v_t correspondingly are the weights of state $\vec{H}_t$ of forward hidden layer and the state $\overleftarrow{\mathrm{H}}_t$ of backward hidden layer equivalent to BiGRU at time t, and b_t indicates the bias of state of the hidden layer, at time t.

2.4 SRO Based Hyperparameter Optimization

In order to have an appropriate fine tuning of the hyperparameters in BiGRNN model, SRO algorithm is utilized. Shabani et al. [19] presented SAR inspired from the research conducted upon humans at search and rescue processes. Search and rescue process comprises of two stages such as individual phase and social phase. The clues left at the time of search by a group member get stored in memory matrix (O) where the humans' position are stored from position matrix (W). The clue matrix B using size N^*D left a clue while the humans' position can be expressed in the matrix given below.

$$B = \begin{bmatrix} W \\ O \end{bmatrix} = \begin{bmatrix} W_{11} & \cdots & W_{1D} \\ \vdots & \ddots & \vdots \\ W_{N1} & \cdots & W_{ND} \\ O_{11} & \cdots & O_{1D} \\ \vdots & \ddots & \vdots \\ O_{N1} & \cdots & O_{ND} \end{bmatrix} \tag{10}$$

The two stages of human search are modelled as given herewith i) social stage: The search direction is represented as $SD_i = (W_i - B_k)$ whereas $k \neq i$. The novel solution can be attained as given herewith.

$$W'_{ij} = \begin{cases} \begin{cases} B_{ij} + r_1(W_{ij} - B_{ij}), & if\ f(B_i) > f(W_i) \\ W_{ij} + r_1(W_{ij} - B_{ij}), & otherwise \end{cases} & if\ r_2 > SE \\ W_{ij},\ otherwise & \end{cases} \tag{11}$$

Now $f(B_i)$ & $f(W_i)$ represent the fitness function values for $B_i \& W_i, r_1$ & r_2 denote the arbitrary values in the interval of [−1, 1] and [0, 1], SE denotes the model variable which lies in the range of 0 & 1.

ii) Individual Stage: According to the present human position, their novel location is found and i^{th} human can be represented as follows.

$$W'_i = W_i + r_3(B_k - B_m), i \neq k \neq m \tag{12}$$

Each solution must be found in the solution space, when a novel location is present on the outside of solution space as mentioned herewith.

$$W_{ij}^{\wedge} = \begin{cases} \dfrac{W_{i,j} + W_j^{\max}}{2} & if\ W_{ij}^{\wedge} > W_j^{\max} \\ \dfrac{W_{i,j} + w_j^{\min}}{2} & if\ W_{ij}^{\wedge} < W_j^{\min} \end{cases} \tag{13}$$

whereas $W_j^{\max}$ & W_j^{min} denote the maximal and minimal values of the threshold [20]. The efficacy of detecting global optimum solution can be improved as given herewith.

$$ME_n = \begin{cases} W_i\ if\ f(W_i') > f(W_i) \\ ME_n\ otherwise \end{cases} \tag{14}$$

$$W_i = \begin{cases} W_i'\ if\ f(W_i') > f(W_i) \\ W_i\ otherwise \end{cases} \tag{15}$$

Here, ME_n represents the n_{th} stored clues location in memory matrix and n indicates a random integer number in the range of 1 and N.

At the time of clues' search operation, when better clues are not found near the existing location after some searches, humans move to a novel location.

$$USN_i = \begin{cases} USN_i + 1\ if\ f(W_i') > f(W_i) \\ 0\ otherwise \end{cases} \tag{16}$$

While the USN value is high than the maximal unsuccessful search number, the human gets an arbitrary location in searching space as per Eq. (17) while the values of USN_i are fixed as 0 for that human.

$$w_{ij} = w_j^{\min} + r_4(w_j^{\max} - w_j^{\min});\ i = 1, \ldots D \tag{17}$$

Here, r_4 lies in the range of 0 and 1.

3 Experimental Validation

The performance validation of the proposed ODL-CDC technique was carried out using Kaggle dataset which contains instances under two classes namely, neutral/normal (class 0) and insult (class 1). Most of the words that exist in neural statements are 'people', 'like', 'just', 'make', 'now', 'right', 'can', 'one' and 'think'. Similarly, the words that exist in insult statements include 'people', 'shit', 'think', 'idiot', 'life', 'bitch', 'back' and 'dumb' which are the most frequent words used in bad sentences. The results were examined without hyperparameter tuning, i.e., BiGODL-CDC technique, and with hyperparameter tuning, i.e., ODL-CDC technique under different measures.

Tab. 1 shows the results of classification obtained by BiGODL-CDC technique on test Kaggle dataset. The results portray that BiGODL-CDC technique accomplished an effective performance under all classes. For instance, under run-1 with normal class, BiGODL-CDC technique obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 91.23%, 90.7%, and 92.98% respectively. At the same time, under run-1 with insult class, the BiGODL-CDC technique has attained Pre_n, Rec_l, and $F1_{measure}$ of 92.61%, 92.42%, and 92.48% correspondingly. At the same time, under run-2 with normal class, BiGODL-CDC technique attained Pre_n, Rec_l, and $F1_{measure}$ values such as 91.49%, 92.26%, and 91.68% correspondingly. In line with these, under run-2 with insult class, BiGODL-CDC methodology reached Pre_n, Rec_l, and $F1_{measure}$ values like 91.39%, 92.33%, and 90.63% respectively. Followed by, under run-3 with normal class, BiGODL-CDC technique obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 92.87%, 92.27%, and 91.91% correspondingly. Simultaneously, under run-3 with insult class, BiGODL-CDC method attained Pre_n, Rec_l, and $F1_{measure}$ values such as 92.35%, 91.14%, and 90.76% correspondingly. Moreover, under run-4 with normal class, BiGODL-CDC technique obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 90.12%, 92.61%, and 91.47% respectively. In addition, under run-4 with insult class, BiGODL-CDC scheme obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 90.01%, 90.22%, and 91.52% correspondingly. Eventually, under run-5 with normal class, BiGODL-CDC approach obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 91.52%, 91.08%, and 91.32% correspondingly. Then, under run-5 with insult class, BiGODL-CDC technique gained Pre_n, Rec_l, and $F1_{measure}$ values like 92.51%, 92.76%, and 90.6% respectively. Meanwhile, under run-6 with normal class, BiGODL-CDC technique achieved Pre_n, Rec_l, and $F1_{measure}$ values being 92.65%, 91.47%, and 91.06% correspondingly. Afterward, under run-6 with insult class, BiGODL-CDC system attained Pre_n, Rec_l, and $F1_{measure}$ values such as 90.42%, 90.22%, and 90.04% respectively.

Table 1: Results of the analysis of BiGODL-CDC model under different runs

No. of runs	Classes	Precision	Recall	F1-measure
Run-1	Normal	91.23	90.7	92.98
	Insult	92.61	92.42	92.48
Run-2	Normal	91.49	92.26	91.68
	Insult	91.39	92.33	90.63
Run-3	Normal	92.87	92.27	91.91
	Insult	92.35	91.14	90.76
Run-4	Normal	90.12	92.61	91.47
	Insult	90.01	90.22	91.52
Run-5	Normal	91.52	91.08	91.32
	Insult	92.51	92.76	90.6
Run-6	Normal	92.65	91.47	91.06
	Insult	90.42	90.22	90.04
Run-7	Normal	92.17	91.48	90.28
	Insult	92.75	90.26	91.97
Run-8	Normal	90.08	91.54	90.99
	Insult	92.65	92.29	90.58
Run-9	Normal	92.46	92.95	92.56
	Insult	92.08	90.97	92.67
Run-10	Normal	91.5	92.07	91.74
	Insult	91.51	90.23	91.45

Concurrently, under run-7 with normal class, BiGODL-CDC technique obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 92.17%, 91.48%, and 90.28% correspondingly. Besides, under run-7 with insult class, BiGODL-CDC technique attained Pre_n, Rec_l, and $F1_{measure}$ values such as 92.75%, 90.26%, and 91.97% respectively. Simultaneously, under run-8 with normal class, BiGODL-CDC technique obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 90.08%, 91.54%, and 90.99% respectively. At the same time, under run-8 with insult class, BiGODL-CDC technique attained Pre_n, Rec_l, and $F1_{measure}$ values such as 92.65%, 92.29%, and 90.58% correspondingly. Furthermore, under run-9 with normal class, BiGODL-CDC technique obtained Pre_n, Rec_l, and $F1_{measure}$ values namely 92.46%, 92.95%, and 92.56% correspondingly. Then, under run-9 with insult class, BiGODL-CDC technique attained Pre_n, Rec_l, and $F1_{measure}$ values namely 92.08%, 90.97%, and 92.67% respectively. Finally, under run-10 with normal class, BiGODL-CDC technique obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 91.5%, 92.07%, and 91.74% correspondingly. At last, under run-10 with insult class, BiGODL-CDC technique attained Pre_n, Rec_l, and $F1_{measure}$ values like 91.51%, 90.23%, and 91.45% respectively.

Tab. 2 and Fig. 3 provide the comprehensive classification results offered by BiGODL-CDC technique under ten distinct runs. The results show that the proposed BiGODL-CDC technique accomplished maximum cyberbullying detection performance. For instance, with run-1, BIGODL-CDC technique achieved Pre_n, Rec_l, $F1_{measure}$, and acc_y of 91.92%, 91.56%, 92.73%, and 91.64% correspondingly. In addition, with run-4, BiGODL-CDC method reached Pre_n, Rec_l, $F1_{measure}$, and acc_y of 90.07%, 91.42%, 91.50%, and 91.27% correspondingly. In line with these, with run-6, BIGODL-CDC algorithm obtained Pre_n, Rec_l, $F1_{measure}$, and acc_y values such as 91.54%, 90.85%, 90.55%, and 90.90% respectively. Likewise, with run-8, BiGODL-CDC technique reached Pre_n, Rec_l, $F1_{measure}$, and acc_y such as 91.37%, 91.92%, 90.79%, and 91.56% correspondingly. Lastly, with run-10, BiGODL-CDC methodology achieved Pre_n, Rec_l, $F1_{measure}$, and acc_y values such as 91.51%, 91.15%, 91.60%, and 91.48% correspondingly.

Table 2: Classification analysis results of BiGODL-CDC model under varying measures

No. of runs	Precision	Recall	F1-measure	Accuracy
Run-1	91.92	91.56	92.73	91.64
Run-2	91.44	92.30	91.16	91.55
Run-3	92.61	91.71	91.34	91.86
Run-4	90.07	91.42	91.50	91.27
Run-5	92.02	91.92	90.96	92.01
Run-6	91.54	90.85	90.55	90.90
Run-7	92.46	90.87	91.13	92.09
Run-8	91.37	91.92	90.79	91.56
Run-9	92.27	91.96	92.62	92.21
Run-10	91.51	91.15	91.60	91.48
Average	91.72	91.57	91.44	91.66

Fig. 4 illustrates the results for accuracy analysis of BiGODL-CDC technique. The figure implies that the accuracy values increased with an increase in epoch count. It is also noted that the validation accuracy also seemed to be higher than the training accuracy.

Loss graph analysis results accomplished by BiGODL-CDC technique is illustrated in Fig. 5. The figure infers that the loss values are minimal with increasing epoch counts. It is also observed that the validation loss is considerably lower than the training loss.

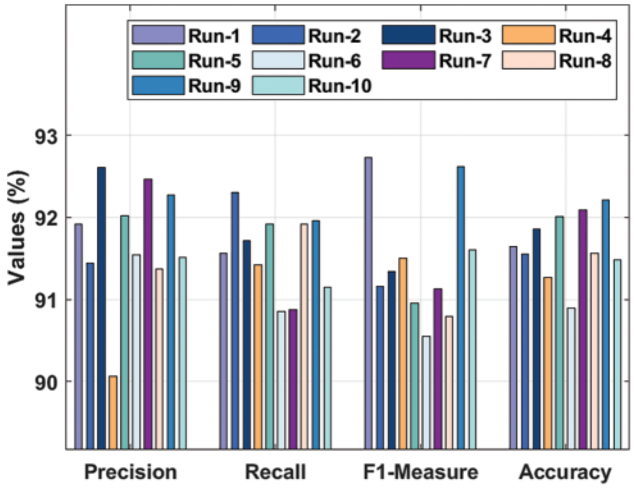

Figure 3: Classification analysis results of BiGODL-CDC model under different measures

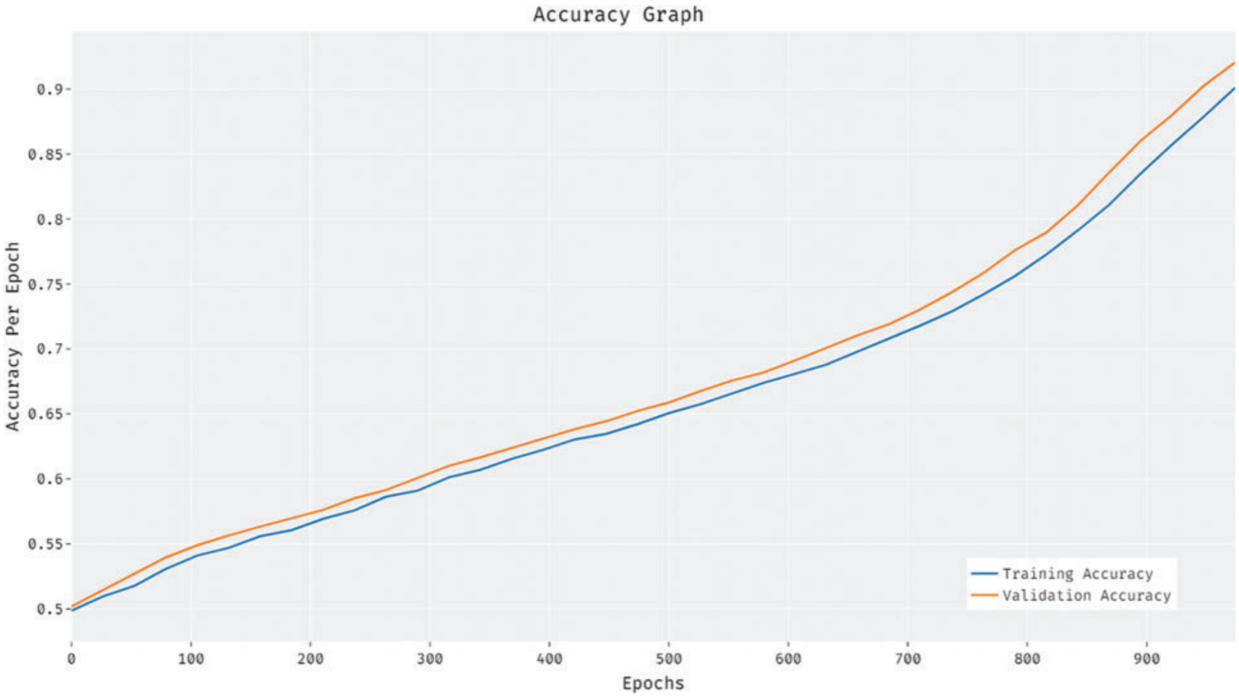

Figure 4: Accuracy analysis of BiGODL-CDC model

Tab. 3 shows the classification outcomes obtained by the proposed ODL-CDC system on test Kaggle dataset. The outcomes show that the proposed ODL-CDC technique outperformed all other methods and accomplished an effectual performance under all classes. For instance, under run-1 with normal class, ODL-CDC technique attained Pre_n, Rec_l, and $F1_{measure}$ values such as 92.04%, 91.23%, and 93.61% correspondingly. At the same time, under run-1 with insult class, ODL-CDC technique attained Pre_n, Rec_l, and $F1_{measure}$ values such as 93.59%, 93.05%, and 93.20% respectively. Simultaneously, under run-2 with normal class, ODL-CDC approach obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 92.02%, 92.96%, and 92.33% correspondingly. Concurrently, under run-2 with insult class, ODL-CDC approach gained Pre_n, Rec_l, and $F1_{measure}$ values such as 91.95%, 92.91%, and 91.54% correspondingly. Afterwards, under run-3 with normal class, ODL-CDC method obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 93.84%, 93.02%, and 92.48% respectively. At the same time, under run-3 with insult class, ODL-CDC technique reached Pre_n, Rec_l, and $F1_{measure}$ values such as 93.27%, 91.74%, and 91.55% correspondingly. Along with that, under run-4 with normal class, ODL-CDC methodology obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 90.79%, 93.36%, and 92.30% respectively. Also, under run-4 with insult class, ODL-CDC system attained Pre_n, Rec_l, and $F1_{measure}$ values such as 90.63%, 90.99%, and 92.26% correspondingly. In addition, under run-5 with normal class, ODL-CDC methodology achieved Pre_n, Rec_l, and $F1_{measure}$ values such as 92.42%, 92.06%, and 91.89% respectively. Then, under run-5 with insult class, ODL-CDC technique attained Pre_n, Rec_l, and $F1_{measure}$ values such as 93.50%, 93.45%, and 91.47% respectively. In the meantime, under run-6 with normal class, ODL-CDC system gained Pre_n, Rec_l, and $F1_{measure}$ values such as 93.40%, 92.38%, and 91.80% correspondingly. Afterward, under run-6 with insult class, ODL-CDC algorithm attained Pre_n, Rec_l, and $F1_{measure}$ values such as 91.09%, 90.81%, and 90.55% respectively.

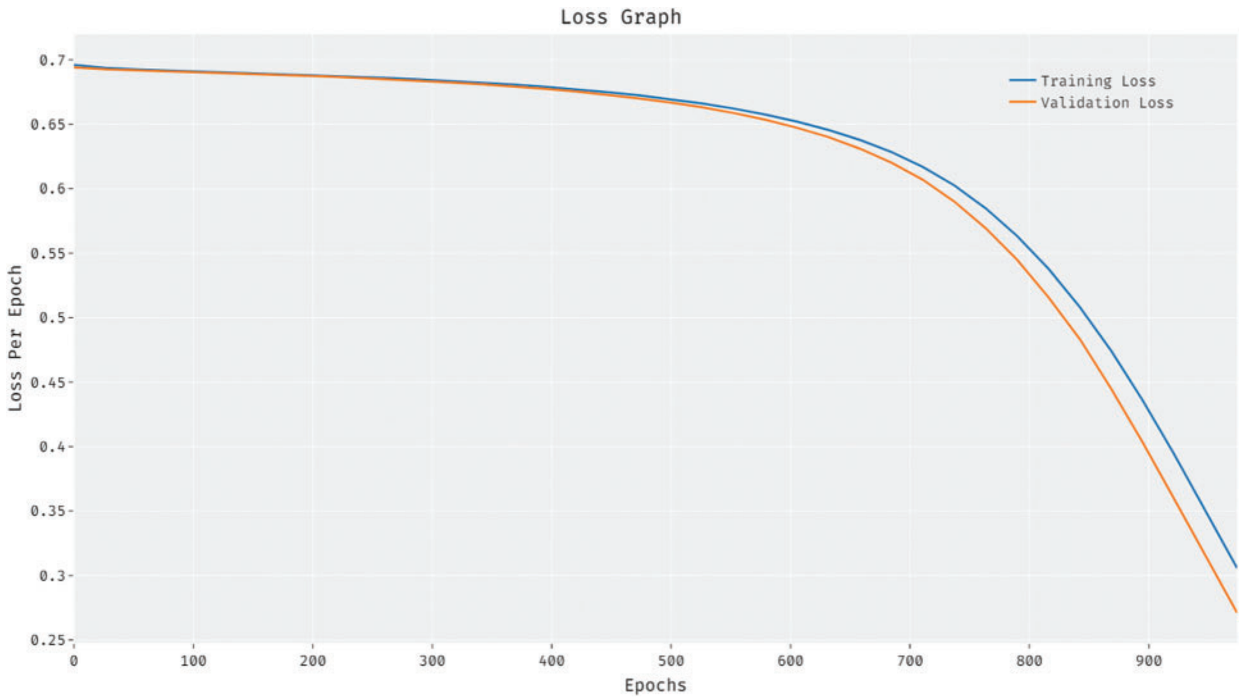

Figure 5: Loss analysis of BiGODL-CDC model

Table 3: Analysis results of ODL-CDC model with different runs

No. of runs	Classes	Precision	Recall	F1-measure
Run-1	Normal	92.04	91.23	93.61
	Insult	93.59	93.05	93.20
Run-2	Normal	92.02	92.96	92.33
	Insult	91.95	92.91	91.54
Run-3	Normal	93.84	93.02	92.48
	Insult	93.27	91.74	91.55
Run-4	Normal	90.79	93.36	92.30
	Insult	90.63	90.99	92.26
Run-5	Normal	92.42	92.06	91.89
	Insult	93.50	93.45	91.47
Run-6	Normal	93.40	92.38	91.80
	Insult	91.09	90.81	90.55
Run-7	Normal	92.72	92.14	91.27
	Insult	93.32	91.03	92.47
Run-8	Normal	90.96	92.07	91.59
	Insult	93.36	93.02	91.31
Run-9	Normal	93.44	93.77	93.47
	Insult	93.08	91.49	93.55
Run-10	Normal	93.44	93.77	93.47
	Insult	93.08	91.49	93.55

Also, under run-7 with normal class, ODL-CDC technique gained Pre_n, Rec_l, and $F1_{measure}$ values such as 92.72%, 92.14%, and 91.27% correspondingly. Then, under run-7 with insult class, ODL-CDC algorithm attained Pre_n, Rec_l, and $F1_{measure}$ values such as 93.32%, 91.03%, and 92.47% respectively. At the same time, under run-8 with normal class, ODL-CDC technique obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 90.96%, 92.07%, and 91.59% correspondingly. Also, under run-8 with insult class, ODL-CDC system attained Pre_n, Rec_l, and $F1_{measure}$ values such as 93.36%, 93.02%, and 91.31% respectively. Furthermore, under run-9 with normal class, ODL-CDC method obtained Pre_n, Rec_l, and $F1_{measure}$ values such as 93.44%, 93.77%, and 93.47% respectively. At the same time, under run-9 with insult class, ODL-CDC method achieved Pre_n, Rec_l, and $F1_{measure}$ values such as 93.08%, 91.49%, and 93.55% correspondingly. At last, under run-10 with normal class, ODL-CDC methodology attained Pre_n, Rec_l, and $F1_{measure}$ values such as 93.44%, 93.77%, and 93.47% respectively. Besides, under run-10 with insult class, ODL-CDC technique reached Pre_n, Rec_l, and $F1_{measure}$ values such as 93.08%, 91.49%, and 93.55% correspondingly.

Tab. 4 and Fig. 6 offer the results for comprehensive classification obtained by the proposed ODL-CDC manner under ten distinct runs. The outcomes showcase that the proposed ODL-CDC technique accomplished the maximal cyberbullying detection performance. For instance, with run-1, ODL-CDC approach gained Pre_n, Rec_l, $F1_{measure}$, and acc_y values such as 92.82%, 92.14%, 93.41%, and 92.39% correspondingly. Followed by, with run-4, ODL-CDC system achieved Pre_n, Rec_l, $F1_{measure}$, and acc_y values such as 90.71%, 92.18%, 92.28%, and 92.16% respectively. Likewise, with run-6, ODL-CDC methodology obtained Pre_n, Rec_l, $F1_{measure}$, and acc_y values such as 92.25%, 91.60%, 91.18%, and 91.67% correspondingly. At the same time, with run-8, ODL-CDC technique achieved Pre_n, Rec_l, $F1_{measure}$, and acc_y values such as 92.16%, 92.55%, 91.45%, and 92.44% respectively. Eventually, with run-10, ODL-CDC method reached Pre_n, Rec_l, $F1_{measure}$, and acc_y values such as 93.26%, 92.63%, 93.51%, and 92.93% correspondingly.

Table 4: Classification analysis results of ODL-CDC model with varying measures

No. of runs	Precision	Recall	F1-measure	Accuracy
Run-1	92.82	92.14	93.41	92.39
Run-2	91.99	92.94	91.94	92.00
Run-3	93.56	92.38	92.02	92.83
Run-4	90.71	92.18	92.28	92.16
Run-5	92.96	92.76	91.68	92.76
Run-6	92.25	91.60	91.18	92.09
Run-7	93.02	91.59	91.87	91.67
Run-8	92.16	92.55	91.45	92.44
Run-9	93.26	92.63	93.51	93.22
Run-10	93.26	92.63	93.51	92.93
Average	92.60	92.34	92.28	92.45

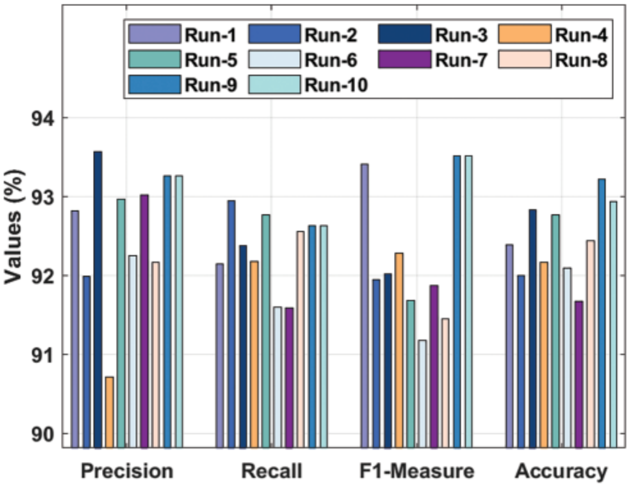

Figure 6: Classification analysis results of ODL-CDC model with different measures

The results for accuracy graph analysis obtained by ODL-CDC approach is demonstrated in Fig. 7. The figure infers that the accuracy values increased with increasing number of epoch counts. It can be also stated that the validation accuracy is superior to training accuracy.

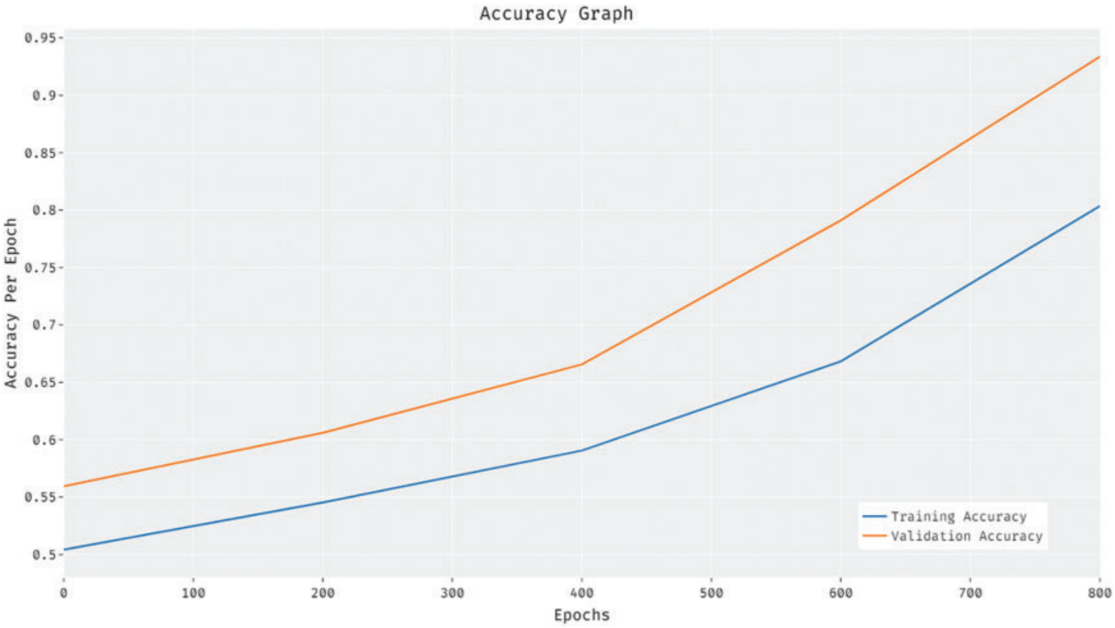

Figure 7: Accuracy analysis results of ODL-CDC model

Fig. 8 shows the results for loss graph analysis obtained by ODL-CDC manner. The figure infers that the loss values got established to a low, with an increase in epoch count. It is also clear that the validation loss is considerably lesser than the training loss.

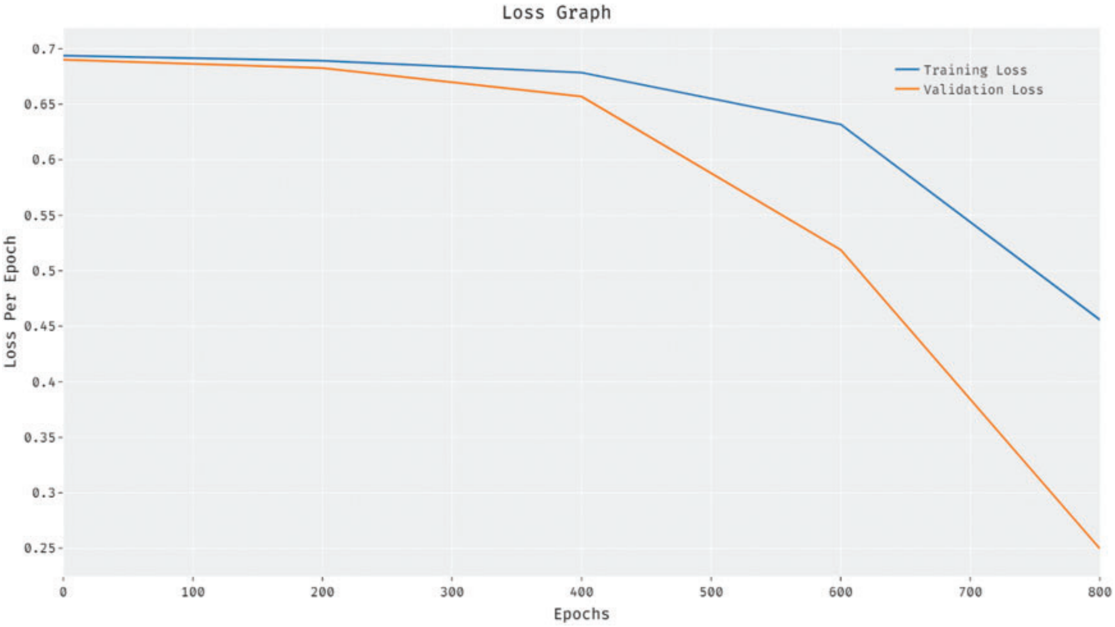

Figure 8: Loss analysis results of ODL-CDC model

Finally, a brief comparative accuracy analysis was conducted between ODL-CDC and recent approaches and the results are shown in Tab. 5 and Fig. 9. The results portray that BLSTM, GRU, LSTM, and RNN techniques obtained the least accuracy values such as 82.18%, 81.46%, 80.86%, and 81.01%. Followed by, BiGRNN technique accomplished a moderately reasonable accuracy of 91.66%. However, the proposed ODL-CDC technique resulted in maximum accuracy of 92.45%. These results validate the supremacy of the proposed ODL-CDC technique in tersm of detection and classification of cyberbullying.

Table 5: Testing accuracy analysis results of ODL-CDC model with existing approaches

Methods	Testing accuracy
ODL-CDC	92.45
BiGRNN	91.66
BLSTM	82.18
GRU	81.46
LSTM	80.86
RNN	81.01

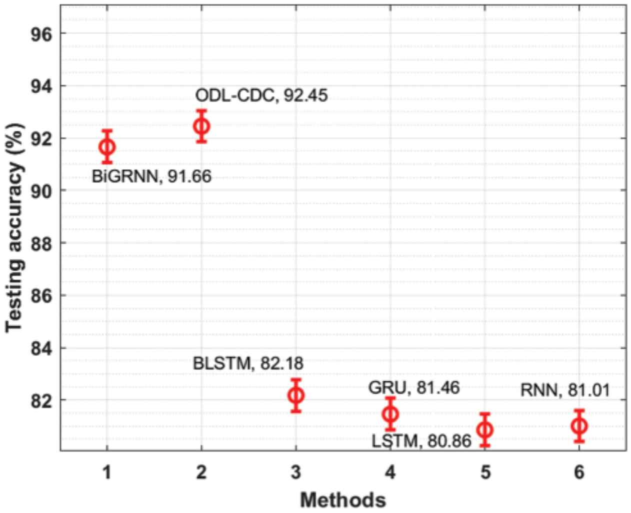

Figure 9: Testing accuracy analysis results of ODL-CDC model

4 Conclusion

In this study, a novel ODL-CDC technique is developed to identify the presence of cyberbullying in social networking data. The proposed ODL-CDC technique comprises of pre-processing, Glove-based word embedding, BiGRNN-based prediction, and SRO-based hyperparameter optimization processes. The utilization of SRO algorithm for proper hyperparameter tuning of BiGRNN parameters helps in boost the classification performance substantially. In order to ensure an improved classification performance of ODL-CDC technique, a comprehensive experimental analysis was carried out on benchmark dataset. The results were inspected under varying aspects while the proposed model achieved a maximum accuracy of 92.45%. A detailed comparative study portrayed the better performance of ODL-CDC technique over recent techniques under different measures. In future, hybrid metaheuristics and outlier detection concepts can be integrated with ODL-CDC technique to increase the detection performance.

Acknowledgement: The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

Funding Statement: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under Grant Number (GPR/303/42). Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2022R191), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] A. S. Medina, I. G. Sánchez and M. F. Monroy, "Applying artificial intelligence to explore sexual cyberbullying behaviour," *Heliyon*, vol. 6, no. 1, pp. e03218, 2021.

[2] K. Shankar, A. R. W. Sait, D. Gupta, S. K. Lakshmanaprabu, A. Khanna *et al.,* "Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model," *Pattern Recognition Letters*, vol. 133, pp. 210–216, 2020.

[3] M. F. L. Vizcaíno, F. J. Nóvoa, V. Carneiro and F. Cacheda, "Early detection of cyberbullying on social media networks," *Future Generation Computer Systems*, vol. 118, pp. 219–229, 2021.

[4] K. Shankar, Y. Zhang, Y. Liu, L. Wu and C. H. Chen, "Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification," *IEEE Access*, vol. 8, pp. 118164–118173, 2020.

[5] H. C. Chan and D. S. W. Wong, "Traditional school bullying and cyberbullying in Chinese societies: Prevalence and a review of the whole-school intervention approach," *Aggression and Violent Behavior*, vol. 23, pp. 98–108, 2015.

[6] A. Bozyigit, S. Utku and E. Nasiboglu, "Cyberbullying detection by using artificial neural network models," in *2019 4th Int. Conf. on Computer Science and Engineering (UBMK)*, Samsun, Turkey, pp. 520–524, 2019.

[7] R. J. S. Raj, S. J. Shobana, I. V. Pustokhina, D. A. Pustokhin, D. Gupta *et al.,* "Optimal feature selection-based medical image classification using deep learning model in internet of medical things," *IEEE Access*, vol. 8, pp. 58006–58017, 2020.

[8] M. A. A. Ajlan and M. Ykhlef, "Optimized twitter cyberbullying detection based on deep learning," in *2018 21st Saudi Computer Society National Computer Conf. (NCC)*, Riyadh, pp. 1–5, 2018.

[9] M. A. Hashedi, L. K. Soon and H. N. Goh, "Cyberbullying detection using deep learning and word embeddings: An empirical study," in *Proc. of the 2019 2nd Int. Conf. on Computational Intelligence and Intelligent Systems*, Bangkok Thailand, pp. 17–21, 2019.

[10] B. Haidar, M. Chamoun and A. Serhrouchni, "Multilingual cyberbullying detection system: Detecting cyberbullying in Arabic content," in *2017 1st Cyber Security in Networking Conf. (CSNet)*, Rio de Janeiro, pp. 1–8, 2017.

[11] M. Ptaszynski, F. Masui, T. Nitta, S. Hatakeyama, Y. Kimura *et al.,* "Sustainable cyberbullying detection with category-maximized relevance of harmful phrases and double-filtered automatic optimization," *International Journal of Child-Computer Interaction*, vol. 8, pp. 15–30, 2016.

[12] A. Bozyiğit, S. Utku and E. Nasibov, "Cyberbullying detection: Utilizing social media features," *Expert Systems with Applications*, vol. 179, pp. 115001, 2021.

[13] 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," *Information Processing & Management*, vol. 58, no. 4, pp. 102600, 2021.

[14] J. Eronen, M. Ptaszynski, F. Masui, A. S. Pohl, G. Leliwa *et al.,* "Improving classifier training efficiency for automatic cyberbullying detection with feature density," *Information Processing & Management*, vol. 58, no. 5, pp. 102616, 2021.

[15] V. Balakrishnan, S. Khan and H. R. Arabnia, "Improving cyberbullying detection using Twitter users" psychological features and machine learning," *Computers & Security*, vol. 90, pp. 101710, 2020.

[16] C. Iwendi, G. Srivastava, S. Khan and P. K. R. Maddikunta, "Cyberbullying detection solutions based on deep learning architectures," *Multimedia Systems*, vol. 13, pp. 1–17, 2020.

[17] A. K. Nandanwar and J. Choudhary, "Semantic features with contextual knowledge-based web page categorization using the GloVe model and Stacked BiLSTM," *Symmetry*, vol. 13, no. 10, pp. 1772, 2021.

[18] J. Yan, J. Liu, Y. Yu and H. Xu, "Water quality prediction in the luan river based on 1-DRCNN and BiGRU hybrid neural network model," *Water*, vol. 13, no. 9, pp. 1273, 2021.

[19] A. Shabani, B. Asgarian, S. A. Gharebaghi, M. A. Salido and A. Giret, "A new optimization algorithm based on search and rescue operations," *Mathematical Problems in Engineering*, vol. 2019, pp. 1–23, 2019.

[20] C. Muppala and V. Guruviah, "Detection of leaf folder and yellow stemborer moths in the paddy field using deep neural network with search and rescue optimization," *Information Processing in Agriculture*, vol. 8, no. 2, pp. 350–358, 2021.