IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 August 2023, accepted 13 October 2023, date of publication 23 October 2023, date of current version 8 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3326948

RESEARCH ARTICLE

Ensemble Learning With Tournament Selected Glowworm Swarm Optimization Algorithm for Cyberbullying Detection on Social Media

RAVURI DANIEL[1], T. SATYANARAYANA MURTHY[2], CH. D. V. P. KUMARI[3], E. LAXMI LYDIA[4], MOHAMAD KHAIRI ISHAK[5,6], MYRIAM HADJOUNI[7], AND SAMIH M. MOSTAFA[8,9]

[1]Department of Computer Science and Engineering, Prasad V. Potluri Siddhartha Institute of Technology, Vijayawada 520007, India
[2]Department of Information Technology, Chaitanya Bharathi Institute of Technology, Hyderabad 500075, India
[3]Department of CSE, Aditya College of Engineering, Surampalem 533437, India
[4]Department of Computer Science and Engineering, Vignan's Institute of Information Technology, Visakhapatnam 530049, India
[5]Department of Electrical and Computer Engineering, College of Engineering and Information Technology, Ajman University, Ajman, United Arab Emirates
[6]School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal, Penang 14300, Malaysia
[7]Department of Computer Sciences, College of Computer and Information Science, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[8]Computer Science Department, Faculty of Computers and Information, South Valley University, Qena 83523, Egypt
[9]Faculty of Industry and Energy Technology, New Assiut Technological University (NATU), New Assiut 71684, Egypt

Corresponding author: Myriam Hadjouni (mfhaojouni@pnu.edu.sa)

This work was supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R193), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

ABSTRACT Online social network (OSN) plays a crucial role to facilitate social connections; but, this social networking media increases antisocial behaviors, like trolling, cyberbullying, and hate speech. Cyberbullying has often resulted in serious physical and mental distress, especially for children and women, and even sometimes forces them to commit suicide. Conventional techniques for detecting cyberbullying, such as relying on users to report the instance of bullying, are not always effective. Deep learning (DL) and Machine learning (ML) techniques are trained to automatically recognize and flag potential cyberbullying content, along with identifying behavior patterns that are indicative of cyberbullying. Therefore, this study concentrates on the design and development of ensemble deep learning with tournament-selected glowworm swarm optimization (EDL-TSGSO) algorithm for cyberbullying detection and classification on Twitter data. The goal of the study is to examine social media data through the use of natural language processing (NLP) and ensemble learning process. This EDL-TSGSO technique preprocesses the raw tweets and then employs the Glove word embedding technique. In addition, the presented EDL-TSGSO technique utilizes ensemble long short-term memory with Adaboost (ELSTM-AB) model for effective cyberbullying detection and classification. The ensemble ELSTM-AB classifier integrates the prediction of LSTM and Adaboost models to enhance the overall classification performance. To further develop the cyberbullying detection performance of the EDL-TSGSO algorithm, the TSGSO algorithm is applied as a hyperparameter optimizer. The experimental validation of the EDL-TSGSO algorithm on the Twitter dataset demonstrates its promising performance over other state of art approaches in terms of different measures.

INDEX TERMS Cyberbullying detection, natural language processing, social media, ensemble learning, hyperparameter tuning.

I. INTRODUCTION

As social media users keep on increasing it has attracted the attention of researchers in examining a novel kind of creative language utilized over the Internet to best search the depth of communication and human thoughts. One most popular

The associate editor coordinating the review of this manuscript and approving it for publication was Mingbo Zhao.

social media is Twitter, a micro-blogging site that permits users to write up to 280 text characters simply called tweets. Developments in Twitter have changed the way individuals share their views and feelings with a large audience because of its easy accessibility and free format messages [1]. Twitter was a real-time information platform that collects the global opinions of the public and Twitter has been considered an outstanding channel to examine peoples' opinions and social interactions. Cyberbullying refers to the use of electronic communication, such as social media platforms, to harass, intimidate, or harm others. On Twitter, this could manifest in the form of abusive tweets, hate speech, or targeted harassment directed at specific individuals or groups [2]. Certainly, the students show symptoms of anxiety and depression, internalizing problems, and negative social relationships, with a risk of suicidal ideas as a function of the frequency of aggressions [3].

Given the significance of cyberbullying and bullying in society, many researchers have examined what can act as protective factors or risks in the involvement of phenomena, addressing the significance of implementing an ecological structure [4]. Cyberbullying through Twitter has gained attention in some years as its leads to several tragic, high-profile suicides. A conventional system was implemented for managing the problem of cyberbullying in Social Media platforms, with companies including guidelines that their users should follow along with using editors to check manually for bullying behaviour [5]. Fig. 1 represents the process involved in the area of intervention.

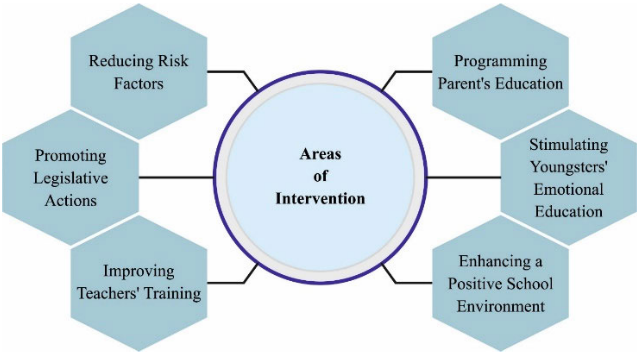
Reducing Risk Factors
Programming Parent's Education
Promoting Legislative Actions
Areas of Intervention
Stimulating Youngsters' Emotional Education
Improving Teachers' Training
Enhancing a Positive School Environment

FIGURE 1. Area of intervention.

Moreover, the significant growth in cyberbullying cases has emphasized the danger of cyberbullying, predominantly among adolescents and children [6], who can be juvenile and inconsiderate. Adolescents consider bullying as a serious problem without knowing how to handle social problems; this made them share their feelings on social networking sites in a way that could hurt others [7]. Many researchers have exposed that bullies undergo psychological states, which leads them to bully and inflict suffering on other people. Therefore, cyberbullying was the same epidemic, and can result in a violent society, predominantly considering high-tech university and school students. Thus, most of the global initiatives were modelled to tackle the issue of cyberbullying [8]. Detection of cyberbullying in social networking sites is highly essential and must be paid higher attention to so that society and children were protected from all those threats. Cyberbullying is hot a research topic among research communities aimed at deducting, controlling, and reducing cyberbullying on social networking sites [9]. One direction in this field was to find the intention of users to post aggressive content by examining offensive language related to different features, such as the unique content and structure, and the writing style of the users [10]. Another direction of cyberbullying research was to identify text content utilizing ML for offensive language classification and detection.

This study concentrates on the design and development of ensemble deep learning with tournament-selected glowworm swarm optimization (EDL-TSGSO) algorithm for cyberbullying detection and classification on Twitter data. This EDL-TSGSO technique preprocesses the raw tweets and then employs the Glove word embedding technique. In addition, the presented EDL-TSGSO technique utilizes ensemble long short-term memory with AdaBoost (ELSTM-AB) model for effective cyberbullying detection and classification. To further improve the cyberbullying detection performance of the EDL-TSGSO algorithm, the TSGSO algorithm is applied as a hyperparameter optimizer. The experimental validation of the EDL-TSGSO algorithm on the Twitter dataset demonstrates its promising performance over other existing systems in terms of different measures.

II. RELATED WORKS

Murshed et al. [11] presented a hybrid DL system termed DEA-RNN for identifying CB on Twitter social media networks. This DEA-RNN technique integrates Elman-type RNN with optimizer Dolphin Echolocation Algorithm (DEA) to fine-tune the Elman RNN's variables and minimalize trained hours. The author assessed DEA-RNN with datasets of 10,000 tweets and compared its performing ability to existing methods like RF, Bi-directional LSTM (Bi-LSTM), SVM, Multinomial NB (MNB), and RNN. Alotaibi et al. [12] introduced an automatic cyberbullying approach for identifying aggressive behavior utilizing a consolidated DL approach. This method leverages multichannel DL depending on 3 methods, namely, the CNN, the bidirectional GRU (BiGRU), and the transformer block for classifying Twitter comments into 2 classes not aggressive and aggressive.

Bharti et al. [13] intend to assess several approaches to mechanically find cyberbullying from tweets through DL and ML techniques. The authors implemented ML approaches and after analyzing the experimental outcomes, the authors postulated that DL approaches had superior performance. Word-embedding approaches have been utilized for word representation for this trained model. Bidirectional LSTM (BLSTM) has been utilized for the classifying process. Mahlangu and Tu [14] introduced a structure for detecting cyberbullying messages in the text form of data utilizing word embeddings and DNN. The author stacked together the existing Bert and Glove embeddings for enhancing the classifier performance. Al-Ajlan and Ykhlef [15] devised optimized Twitter cyberbullying recognition related to DL (OCDD), a new technique to solve the abovementioned difficulties. Dissimilar to prior work in this domain, OCDD could not be extracting features from tweets and giving them to classifiers; rather, it indicates a tweet as a group of word vectors.

Fang et al. [16] concentrated on text-related cyberbullying recognition as it was a typically utilized data carrier in social networks and was a broadly utilized feature in this work. Inspired by the success of NN, the author modelled a comprehensive model integrating the self-attention mechanism and bidirectional GRU (Bi-GRU). In-depth, the author presented the model of GRU cell and Bi-GRU benefit to learn underlying relations among words from both directions. On top of that, the author introduced the proposal of a self-attention system and the advantage of this joining to gain a superior performance of cyberbullying classifier errands. In [17], a sentiment detection mechanism was modelled through the text data. RNN can be utilized for CNN and text analysis for image analysis. The textual data was created on Twitter using Twitter API.

III. THE PROPOSED MODEL

In this study, we have established a novel EDL-TSGSO algorithm for the accurate cyberbullying detection and classification on Twitter data by the use of NLP and ensemble learning process. The presented EDL-TSGSO technique follows a series of processes namely preprocessing, Glove word embedding, ELSTM-AB classification, and TSGSO-based hyperparameter tuning. Fig. 2 demonstrates the overall workflow of the EDL-TSGSO approach.

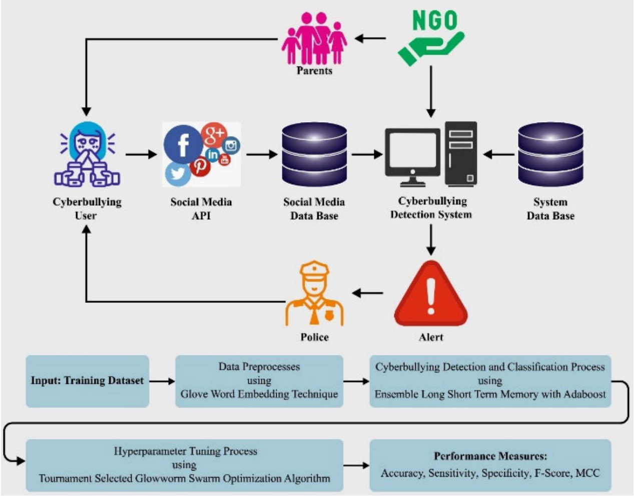
NGO
Parents
Cyberbullying User
Social Media API
Social Media Data Base
Cyberbullying Detection System
System Data Base
Police
Alert
Input: Training Dataset
Data Preprocesses using Glove Word Embedding Technique
Cyberbullying Detection and Classification Process using Ensemble Long Short Term Memory with Adaboost
Hyperparameter Tuning Process using Tournament Selected Glowworm Swarm Optimization Algorithm
Performance Measures: Accuracy, Sensitivity, Specificity, F-Score, MCC

FIGURE 2. Overall workflow of EDL-TSGSO system.

A. DATA PRE-PROCESSING

In the primary stage, the EDL-TSGSO method preprocesses the raw tweets. To preprocess the data, the Natural Language Toolkit (NLTK) is used [18]. Also, it is used to eliminate stop words in the text, tokenization of text patterns, and so on.

• Wordnet lemmatizer: Wordnet lemmatizer is used to find the synonyms of the word, meaning, etc., and connects them to one word.

• Tokenization: the input text was divided as separate words and they are added to the list. Initially, PunktSentenceTokenizer was utilized to tokenize text as sentences. Next, four distinct tokenizers can be utilized for tokenizing the sentence as words:

- TreebankWordTokenizer
- WhitespaceTokenizer
- PunctWordTokenizer
- WordPunctTokenizer

• Lowering Text: It lowers each letter of the words in the tokenization list. For instance: Before lowering "Hey There" after lowering "hey there".

• Removing Stop words: The major part of pre-processing. Stop words are meaningless words from the data. Stop words could be gotten rid of very easily by utilizing NLTK. During this phase, stop words like \u, \t, and https, are distant in the text.

B. GLOVE WORD EMBEDDING

Next to data pre-processing, the Glove word embedding technique is used. Word-embedded purposes are to convert textual data as a vector of real values. Semantic or language vector space methods of language signify all the words with real-valued vectors [19]. Many approaches are presented for representing words in real-valued vectors like neural network (NN), TF-IDF, and Latent Semantics Analysis (LSA) approaches namely GloVe and word2vec.

Word vectorization was separated as global matrix factorized like LSA and local context windows such as the Skip-Gram method. Global matrix factorized efficiently exploited the statistical data, however, it fails for capturing word analogy. Conversely, the Window-based model effectively captures word analogy and then worse utilizes global statistics.

The bag of words (BoW) method was utilized by TF-IDF techniques. TF-IDF depends on the statistical data of words in many files. Terms mention that word or set of words; TF refers to term frequency was fundamentally the word frequency in the documents; for normalizing the value, the frequency was separated by the word counts in a single file. IDF represents the inverse document frequency.

GloVe discovers the global representation of whole words and integrates the word's meaning. Word frequency and co-occurrence are essential metrics that the values of real-valued vectors of the specific corpus are computed. The gloVe is an unsupervised system, but there is no human for introducing ground-truth meaning as the group of words (corpus). The fundamental of computation is the utilization of the frequency of specific words and the neighboring corpus nearby every corpus.

In GloVe, a primary stage was gathering the most common corpus as the context. The secondary stage is to scan the word from the words for generating a *co*-occurrence matrix X. Assume i as the index of frequent corpus and j as the rest of the corpus from the corpus. P_{ij} implies the possibility of word j taking place with context word i.

$$P_{ij} = P\left(j \mid i\right) = \frac{X_{ij}}{X_i} \tag{1}$$

Assume that 2 words i and j, and a context word k; compute a ratio of *co*-occurrence possibility as:

$$F\left(w_{j,} w_j \tilde{W}_k\right) = \frac{P_{ij}}{P_{jk}} \tag{2}$$

Lastly, the loss function J is computed as:

$$J = \sum_{i,j=1}^{V} f(X_{ij})(W_i^T \tilde{W}_j + b_i + b_j - logX_{ij})^2, \tag{3}$$

In which f refers to the weighted function. The training purpose for minimizing the least square error. If the GloVe was trained, all the words are allocated to specific real-valued vectors.

C. ENSEMBLE CLASSIFICATION

In this work, the presented EDL-TSGSO technique employed the ELSTM-AB model for effective cyberbullying detection and classification. The LSTM network which is a special kind of RNN which consist of gate units and memory cells [20] was demonstrated that higher in mapping relationship among variables. In recent years, LSTM applications allocated to financial predicting tasks have accomplished remarkable performance. Gers et al. proposed LSTM which is the most common variant that is based on the pictorial representation using forget gate. The LSTM network preserves the prior state data over a long sequence utilizing the memory cell design C_τ that effectively enables the gradient to flow for a longer time, thus facilitating the gradient vanishing problems. The input data processed by the forget gate f_T and the input gate i_τ passes to the state information and memory cell C_τ that is controlled by the output gate 0_τ and later passes to another block of LSTM. The mathematical formula for the gate unit and memory cell outcomes are demonstrated in Eqs. (4) to (8):

In every time step, the fusion of hidden state $h_{\tau-1}$ and the current time input vector χ_t from the prior step is converted to LSTM cell units, and later evaluated by the logistic sigmoid function as follows:

$$i_\tau = \sigma\left(W_{xi}x_\tau + W_{hi}h_{\mathrm{t}-1} + W_{ci}C_{\tau-1} + b_i\right) \tag{4}$$

The above equation represents the logistic sigmoid function, W_{xi}, W_{hi}, and W_{ci} denote the separate weight vector for all the inputs connecting two components, the bias vector for the input gate unit, and the cell state from the prior step. The forget gate in the LSTM structure determines that data is detached from the cell state. The mathematical expression of the forget gate can be given as follows:

$$f_t = \sigma\left(W_{xf}x_\tau + W_{hf}h_{\tau-1} + W_{cf}C_{\tau-1} + b_f\right) \tag{5}$$

In Eq. (5), b_f indicates the bias for the forget gate, W_{xf}, W_{hf}, and W_{cf} represent the separate weight vector for all the inputs connecting two components. From the equation, the sigmoid function generates value within [0, 1]; the prior memory is forgotten when the resultant value for the forget gate was closer to 0, whereas value 1 specifies that everything saved in the prior memory block was remembered. Using Eq. (6), the cell state C_τ can be updated:

$$\begin{aligned} C_\tau &= f_t \odot C_{t-1} + i_t \\ &\odot\, tanh\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}C_{t-1} + b_c\right) \end{aligned} \tag{6}$$

In Eq. (6), b_c indicates the bias vector and the symbol $\odot$ represents the Hadamard (component-wise) products.

Lastly, the output of the LSTM block is produced using the following equations:

$$o_t = \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}C_{t-1} + b_o\right) \tag{7}$$

$$h_t = 0_t \odot tanh\left(C_t\right) \tag{8}$$

where b_o represents the bias for the output gate, W_{xo}, W_{ho}, and W_{co} denotes the separate weight vectors for all the inputs linking 2 elements. The AdaBoost was used for building a higher quality ensemble ELSTM-AB model that successively incorporates the different LSTM-based models and provides the well-functioning weaker classifier with a large voting weight. The major step to develop the presented DL ensemble classification method is given below:

Assume a tweet dataset that comprises of n training instances, $= \{(x_1, y_1), (x_2, y_2), \ldots, (x_i, y_i), \ldots, (x_n, y_n)\}$

Where x_i and y_i indicate the input feature space and class labels (offensive or non-offensive). Then, the LSTM network is utilized as a base learner in the ensemble module, with the resultant prediction being the base learner ensemble that is represented $F = \{L_1, L_2, \ldots, L_M\}$.

For simplification purposes, without losing generalization, assume that the weight distribution over this sample at m^{th} boosting iteration is represented as D_m that is endowed initially with the similar value $1/n$ at the initial iteration, with the overall predictive error for the present weaker classifiers on the trained dataset as follows:

$$\varepsilon_m = \sum_{i}^{n} D_m\left(i\right) \times \begin{cases} 1 & if\ L_m\left(x_i\right) \neq y_i \\ 0 & if\ L_m\left(x_i\right) = y_{i'} \end{cases} \tag{9}$$

In Eq. (9), y_i denotes the observed labels for input sample x_i, ε_m indicates the classification error for the present classifiers, and L_m shows the training LSTM at m iterations .

Next, the training data weight distribution was upgraded dependent upon the classifier outcome of the present hypothesis such that the misclassified sample is allocated a higher weight and the properly classified sample is allocated a lower weight. The updating process can be formulated in the following expression:

$$D_{m+1}\left(i\right)=\frac{D_m\left(i\right)}{Z_m}\exp\left(-\partial_m\times y_i\times L_m\left(x_i\right)\right), \tag{10}$$

In Eq. (10), Z_m indicates the normalization constant which ensures that the weight $D_{m+1}(i)$ has an appropriate distribution, and ∂_m denotes the voting weights for the training classification L_m. The Z_m and ∂_m are mathematically expressed as follows:

$$Z_m=\sum_1^n D_m\left(i\right)\exp\left(-\partial_m\times y_i\times L_m\left(x_i\right)\right) \tag{11}$$

$$\partial_m=\frac{1}{2}\ln\left(\frac{1-\varepsilon_m}{\varepsilon_m}\right). \tag{12}$$

Once the M iteration was processed, the ensemble encompassed of M weaker classifier. From Eq. (13), the classification outcome of AdaBoost is an integration of the classification outcomes weighted using ∂_m:

$$F\left(x\right)=sign\left(\sum_1^M \partial_m x L_m\left(x\right)\right), \tag{13}$$

In Eq. (13), $sign(x)$ characterizes the sign function and it can be expressed as follows:

$$sign(x)=\begin{cases}1, & if\ x>0\\ Q_p & if\ x=0.\\ -1_j & if\ x<0\end{cases} \tag{14}$$

D. HYPERPARAMETER TUNING

Finally, the TSGSO algorithm is applied as a hyperparameter optimizer. TSGSO is a kind of SI method that is based on the behavior of glowworms. The glowworm behavior pattern modifies the strength of Lucifer in release and glows at the different powers [21]. The TSGSO technique agent is regarded as glowworm viz., switching angle that transmits the luciferin luminescence. All the glowworms exploit the luciferin viz., neutral basis to transfer information from the existing location to the closet. Afterwards choosing the neighbor, implements the action. All the glowworms play a tournament selection method to select the surroundings with higher luciferin values and moved to them. It encompasses the space within the dynamic result area and the dynamic resultant gap with luciferin superior apart. The glowworms update the place to glowworm within the decision space radius and the dynamic result area. TSGSO comprises 2 major concepts:

The agent glows to the intensity proportional to the optimization of an objective task. Glowworm of bright intensity gets attracted by glowworms of less intensity.

TSGSO includes the dynamic decision space whereby the distant glowworm effects are discounted but the glowworm has a sufficient neighboring place.

1) DEPLOYMENT OF GLOWWORM PHASE

The glowworm viz., switching angle is scattered in the objective gap. Each glowworm includes a similar size of luciferin. The glowworms position and all the glowworms launch a similar luciferin rate according to the function rate by using the primary iteration. The rate changes frequently with the function rate in the existing location. During the luciferin update phase, all the glowworms load to the previous luciferin phase. The proportion of luciferin rate was eliminated to imitate decay in luciferin over time. In the present position, every glowworm changes the luciferin rate by using the main function rate and it could be expressed as:

$$LE_i\left(T\right)=\left(1-\rho\right)*LE_i\left(T-1\right)+\gamma*OF_i\left(T\right), \tag{15}$$

$$OF_i\left(T\right)=\min_{\delta_i}\left\{\left(\frac{100*\left(V_1^*-V_1\right)}{y_1*}\right)^4+\sum_{i=2}^{S}\frac{1}{h_s}\left(50*\frac{V_{h_s}}{V_1}\right)^2\right\}. \tag{16}$$

where “ρ” and “y” indicates the decay steady and proportion of luciferin. “$LE_i(T-1)$” denotes the preceding luciferin level for glowworm. “OF_i (T)” shows the objective task rate in glowworm. “ith” denotes the location at the time instant “T”.

2) MOVEMENT-PHASE

All the glowworms take the outcome with the tournament model for moving the neighboring through luciferin rate. TS is used to define the better glowworms utilizing the main function. Manhattan Distance (MD) was evaluated according to the present and the adjacent glowworm location.

$$Manhattan\ distance=\sqrt{\sum_{i,j=1}^{n}(g_j-g_i)^2}. \tag{17}$$

where “g_i” signifies the existing location of the glowworm and “g_j” characterizes the adjacent glowworm location. Once the distance is less, it has a high fitness value. The tournament selective method is utilized for selecting glowworms at random from the population with the highest fitness. The glowworm chosen was exploited for creating consecutive generations. By using the selection technique, the fitness of all the glowworms is evaluated as,

$$P=\frac{F_i}{\sum_{j=1}^{n}F_j}. \tag{18}$$

The above equation is the tournament selective Probability (P) of all the glowworms and F_i signifies the average fitness of populations in j^{th} glowworm. From the tournament, the initial better glowworm is selected with probability and the second selection probability is evaluated by

$$P*\left(1-P\right). \tag{19}$$

The third better glowworm with the probability was selected as

$$P*(1-P)^2 \tag{20}$$

Thus, the better glowworm was selected together with chosen probability. Once the size of the tournament is higher, weaker glowworm includes less probability for the selection of the tournament. Afterwards selection, the glowworm has included neighbors that glow brightly. The possibility of moving for the neighbor "q is given by

$$N_i(t) = \left\{q{:}D_{iq}(T) < R_d^i(T)\right\}, LE(T) \tag{21}$$

$$\rho_{iq}(T) = \frac{LE_q(T) - LE_i(T)}{\sum_{k\epsilon N_i(T)} LE_q(T) - LE_i(T)}. \tag{22}$$

where "T" denotes the index value of time. $(D_{iq}(T)'$ illustrates the Euclidian gap amongst glowworms "i" with "q" at the period "T". "$(LE_q(T)$" shows luciferin stage by glowworm "q" at period "T". "$(R_d^i(T)$" means uneven local-result variation by glowworm 'i' at the time 'T'. "R" represents the range of radial from the luciferin sensor. All the glowworms update its position by

$$X_i(T+1) = X_i(T) + S_s\left(\frac{X_q(T) - X_i(T)}{X_q(T) - X_i(T)}\right). \tag{23}$$

where, "S_s" means the step size and $(|X_q(T) - X_i(T)|$" specifies Euclidean uses glowworm. Local decision range updating rule:- The dynamic glowworm of decision spaces depends on the linked sensor of luciferin of radial range and decision space in the present radius. To determine the location of glowworm that based on local details, it is evaluated by the stronger function in the radial sensor selection as follows

$$R_d^i(T+1) = \min\left\{R, \max\left\{0, R_d^i(T) + \beta N_T - |N_i(T)|\right\}\right\}. \tag{24}$$

Now, β" denotes a constant parameter. "N_t" indicates the clear threshold parameter. Thereby, the better glowworm can be recognized by the TSGSO effectively.

The TSGSO system develops a fitness function (FF) for obtaining enriched efficacy of the classification. It describes the positive integer to characterize the improved performance of the candidate results. In the presented method, the minimizing of the classifier rate of errors is regarded as the fitness function.

$$\begin{aligned} fitness(x_i) &= ClassifierErrorRate(x_i) \\ &= \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} * 100 \end{aligned} \tag{25}$$

IV. RESULTS AND DISCUSSION

The proposed model is simulated using Python 3.6.5 tool. The cyberbullying detection performance of the EDL-TSGSO technique was tested utilizing the Twitter dataset from the Kaggle repository [22]. The dataset includes 31353 offensive and 24435 non-offensive tweets as defined in Table 1.

In Fig. 3, the confusion matrices of the EDL-TSGSO technique are studied under 80:20 of TRS/TSS. The results inferred that the EDL-TSGSO technique has properly categorized the offensive and non-offensive tweets.

TABLE 1. Details of the dataset.

Category	Number of Instances
Offensive	31353
Non-offensive	24435

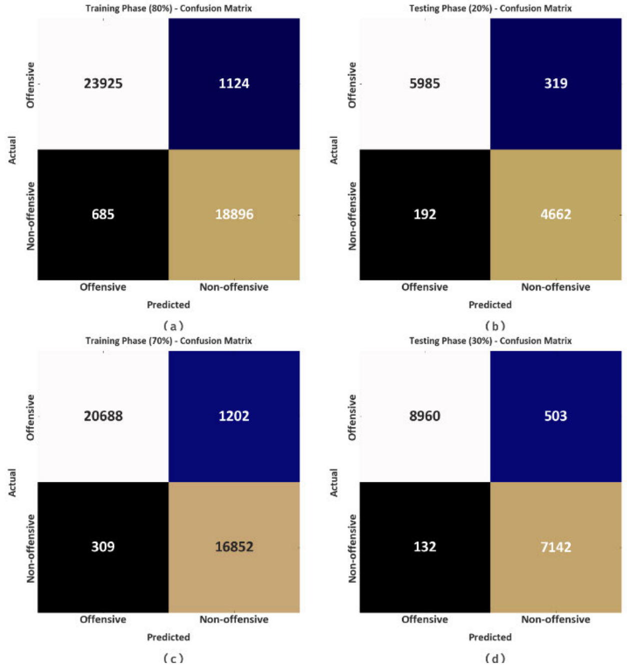

FIGURE 3. Confusion matrices of EDL-TSGSO system (a-b) TRS/TSS of 80:20 and (c-d) TRS/TSS of 70:30.

Table 2 and Fig. 4 report the overall cyberbullying detection results of the EDL-TSGSO technique with 80:20 of TRS/TSS. The results indicated that the EDL-TSGSO technique has properly recognized the tweet classes. For instance, with 80% of TRS, the EDL-TSGSO technique achieves an average $accu_{bal}$ of 96.01%, $sens_y$ of 96.01%, $spec_y$ of 96.01%, F_{score} of 95.89%, and MCC of 91.81%. In addition, with 20% of TSS, the EDL-TSGSO method attains an average $accu_{bal}$ of 95.49%, $sens_y$ of 95.49%, $spec_y$ of 95.49%, F_{score} of 95.36%, and MCC of 90.74%.

TABLE 2. Cyberbullying detection outcome of EDL-TSGSO approach on 80:20 of TRS/TSS.

Class	Accuracybal	Sensitivity	Specificity	F-Score	MCC
Training Phase (80%)					
Offensive	95.51	95.51	96.50	96.36	91.81
Non-offensive	96.50	96.50	95.51	95.43	91.81
Average	96.01	96.01	96.01	95.89	91.81
Testing Phase (20%)					
Offensive	94.94	94.94	96.04	95.91	90.74
Non-offensive	96.04	96.04	94.94	94.80	90.74
Average	95.49	95.49	95.49	95.36	90.74

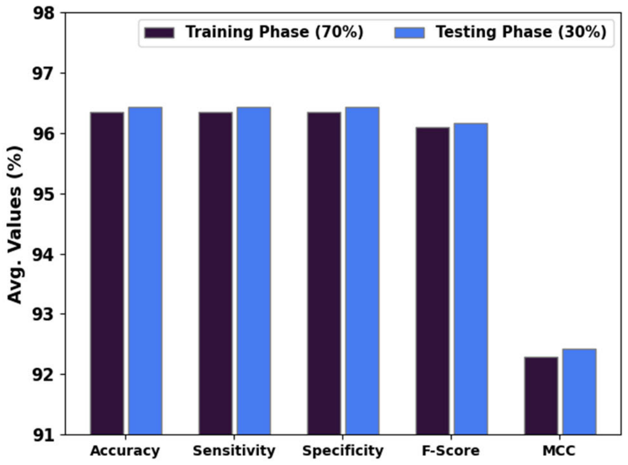

FIGURE 4. Average outcome of EDL-TSGSO approach on 80:20 of TRS/TSS.

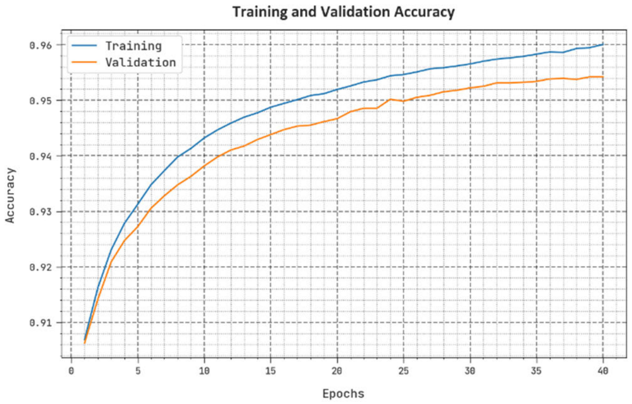

FIGURE 5. TACY and VACY outcome of EDL-TSGSO approach on 80:20 of TRS/TSS.

The TACY and VACY of the EDL-TSGSO approach on 80:20 of TRS/TSS have been investigated in Fig. 5. The figure implied that the EDL-TSGSO method has shown improved performance with increased values of TACY and VACY. Notably, the EDL-TSGSO approach has reached maximum TACY outcomes.

The TLOS and VLOS of the EDL-TSGSO method on 80:20 of TRS/TSS are signified in Fig. 6. The figure shows the EDL-TSGSO approach has better performance with minimal values of TLOS and VLOS. Visibly the EDL-TSGSO technique has reduced VLOS outcomes.

Table 3 and Fig. 7 report the overall cyberbullying detection results of the EDL-TSGSO method with 70:30 of TRS/TSS. The fallouts exhibited by the EDL-TSGSO method have properly recognized the tweet classes. For example, with 70% of TRS, the EDL-TSGSO system attains an average $accu_{bal}$ of 96.35%, $sens_y$ of 96.35%, $spec_y$ of 96.35%, F_{score} of 96.09%, and MCC of 92.29%. Also, with 30% of TSS, the EDL-TSGSO method attains an average $accu_{bal}$ of 96.43%, $sens_y$ of 96.43%, $spec_y$ of 96.43%, F_{score} of 96.16%, and MCC of 92.42%.

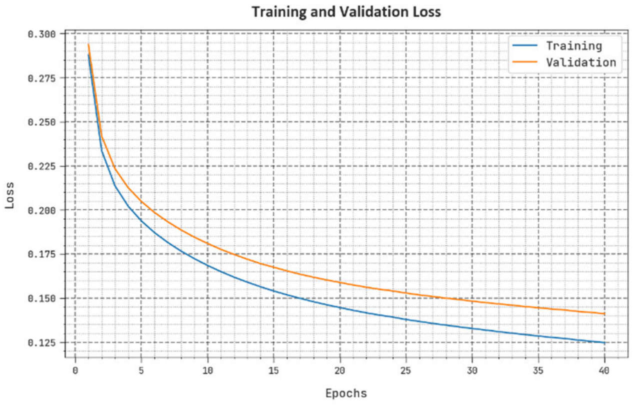

FIGURE 6. TLOS and VLOS outcome of EDL-TSGSO approach on 80:20 of TRS/TSS.

TABLE 3. Cyberbullying detection outcome of EDL-TSGSO approach on 70:30 of TRS/TSS.

Class	Accuracybal	Sensitivity	Specificity	F-Score	MCC
Training Phase (70%)					
Offensive	94.51	94.51	98.20	96.48	92.29
Non-offensive	98.20	98.20	94.51	95.71	92.29
Average	96.35	96.35	96.35	96.09	92.29
Testing Phase (30%)					
Offensive	94.68	94.68	98.19	96.58	92.42
Non-offensive	98.19	98.19	94.68	95.74	92.42
Average	96.43	96.43	96.43	96.16	92.42

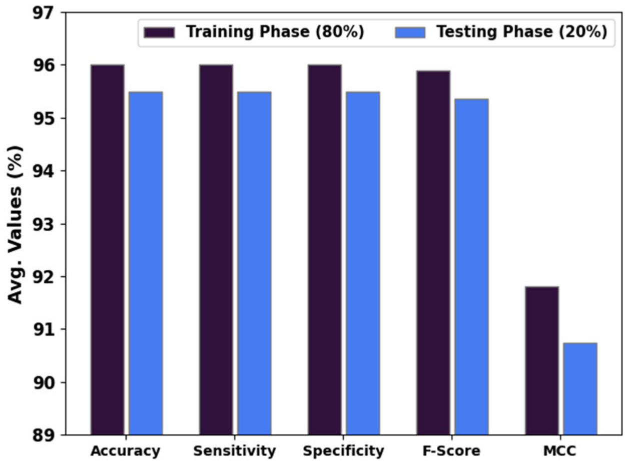

FIGURE 7. Average outcome of EDL-TSGSO approach on 70:30 of TRS/TSS.

The TACY and VACY of the EDL-TSGSO approach on 70:30 of TRS/TSS have been investigated in Fig. 8. The figure implied that the EDL-TSGSO technique has shown improved performance with increased values of TACY and VACY. Notably, the EDL-TSGSO technique has reached maximum TACY outcomes.

The TLOS and VLOS of the EDL-TSGSO method on 70:30 of TRS/TSS are shown in Fig. 9. The figure shows that

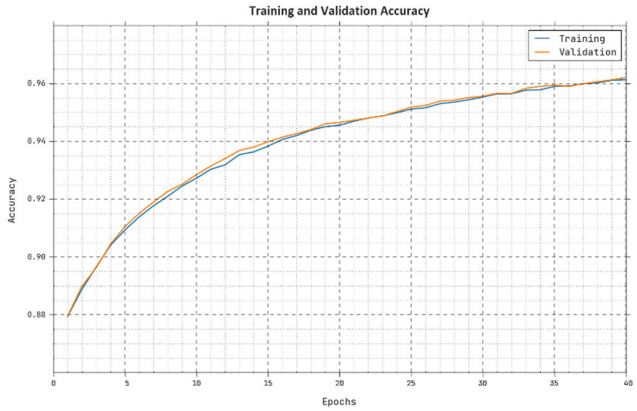

FIGURE 8. TACY and VACY outcome of EDL-TSGSO approach on 70:30 of TRS/TSS.

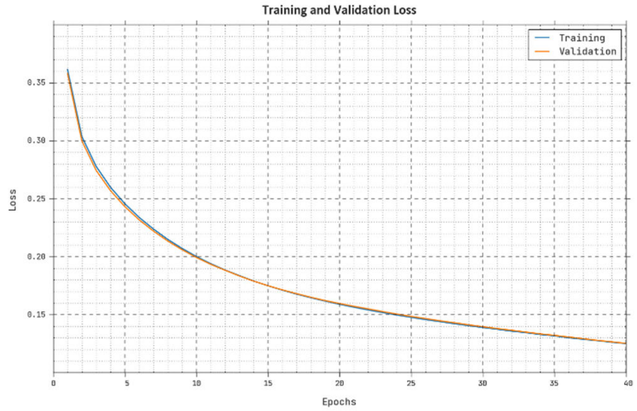

FIGURE 9. TLOS and VLOS outcome of EDL-TSGSO approach on 70:30 of TRS/TSS.

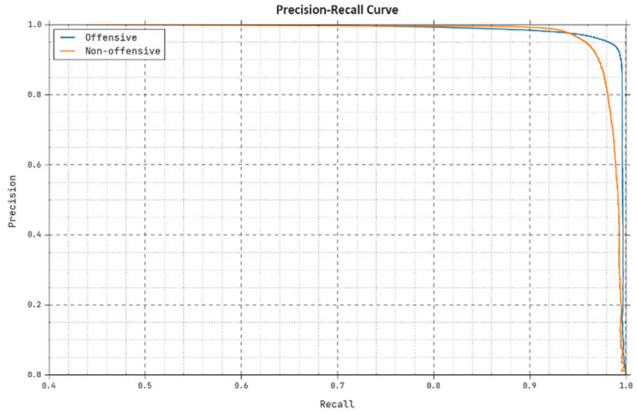

FIGURE 10. Precision recall outcome of EDL-TSGSO approach.

the EDL-TSGSO technique has revealed better performance with least values of TLOS and VLOS. Visibly the EDL-TSGSO approach has reduced VLOS outcomes.

A brief precision-recall inspection of the EDL-TSGSO technique under the test database is shown in Fig. 10. The figure designated the EDL-TSGSO technique has greater values of precision-recall values under all classes.

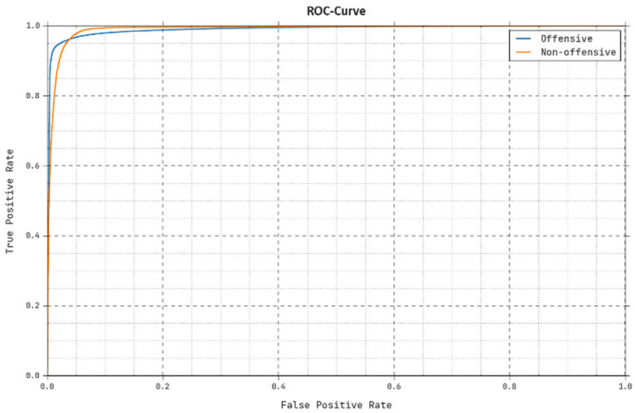

FIGURE 11. ROC outcome of EDL-TSGSO approach.

The detailed ROC analysis of the EDL-TSGSO method under the test database is shown in Fig. 11. The outcome implied the EDL-TSGSO approach has displayed its capability in classifying different class labels.

In Table 4, a comparative $accu_y$ examination of the EDL-TSGSO technique is provided briefly [12]. The experimental values highlighted that the EDL-TSGSO technique has obtained superior outcomes over other models. In addition, it is noticed that the EDL-TSGSO technique reaches $anaccu_y$ of 96.43%. Contrastingly, the EDL, MC-DL, linear SVC, TC, bagging, CNN, and BiGRU models result in $accu_y$ of 94.86%, 89.13%, 50.72%, 87.33%, 68.74%, 88.14%, and 88.14% respectively.

TABLE 4. Accuracy analysis of EDL-TSGSO approach with other recent algorithms.

Method	Accuracy (%)
EDL-TSGSO	96.43
EDL	94.86
MC-DL Model	89.13
Linear SVC	50.72
Transformer block	87.33
Bagging Classifier	68.74
CNN Model	88.14
BiGRU Model	88.14

In Table 5, a comparative CT inspection of the EDL-TSGSO method is provided briefly. The experimental values emphasized that the EDL-TSGSO method has gained lesser time taken over other methods. Moreover, it is noted that the EDL-TSGSO method reaches a CT of 0.32s. Contrastingly, the EDL, MC-DL, linear SVC, TC, bagging, CNN, and BiGRU methods result in CT of 0.45s, 1.40s, 0.73s, 0.76s, 1.10s, 0.63s, and 0.40s correspondingly.

These results demonstrated the betterment of the EDL-TSGSO technique over other approaches.

TABLE 5. CT analysis of EDL-TSGSO approach with other recent algorithms.

Method	Computational Time (sec)
EDL-TSGSO	0.32
EDL	0.45
MC-DL Model	1.40
Linear SVC	0.73
Transformer block	0.76
Bagging Classifier	1.10
CNN Model	0.63
BiGRU Model	0.40

V. CONCLUSION

In this study, we have developed a novel EDL-TSGSO algorithm for the accurate cyberbullying detection and classification on Twitter data by the use of NLP and ensemble learning process. The presented EDL-TSGSO technique follows a series of processes namely pre-processing, Glove word embedding, ELSTM-AB classification, and TSGSO-based hyperparameter tuning. The presented EDL-TSGSO technique integrates the LSTM and Adaboost classifiers for effective cyberbullying detection and classification. The ensemble ELSTM-AB classifier integrates the prediction of LSTM and Adaboost models to enhance the overall classification performance. To further develop the cyberbullying detection performance of the EDL-TSGSO algorithm, the TSGSO algorithm is applied as a hyperparameter optimizer. The experimental validation of the EDL-TSGSO algorithm on the Twitter dataset demonstrates its promising performance over other state of art approaches in terms of different measures. Future work can focus on the design of advanced hybrid metaheuristic optimization algorithm for hyperparameter tuning process. In addition, the proposed model can be tested on large scale real time dataasets. Future work should explore privacy-preserving artificial intelligence techniques that can detect cyberbullying without compromising individuals' personal information.

ACKNOWLEDGMENT

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R193), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

REFERENCES

[1] J. Batani, E. Mbunge, B. Muchemwa, G. Gaobotse, C. Gurajena, S. Fashoto, T. Kavu, and K. Dandajena, "A review of deep learning models for detecting cyberbullying on social media networks," in *Proc. Comput. Sci. Online Conf.* Cham, Switzerland: Springer, 2022, pp. 528–550.

[2] S. T. Laxmi, R. Rismala, and H. Nurrahmi, "Cyberbullying detection on Indonesian Twitter using Doc2Vec and convolutional neural network," in *Proc. 9th Int. Conf. Inf. Commun. Technol. (ICoICT)*, Aug. 2021, pp. 82–86.

[3] V. Balakrishnan, S. Khan, and H. R. Arabnia, "Improving cyberbullying detection using Twitter users' psychological features and machine learning," *Comput. Secur.*, vol. 90, Mar. 2020, Art. no. 101710.

[4] A. Mangaonkar, R. Pawar, N. S. Chowdhury, and R. R. Raje, "Enhancing collaborative detection of cyberbullying behavior in Twitter data," *Cluster Comput.*, vol. 25, no. 2, pp. 1263–1277, Apr. 2022.

[5] A. M. Alduailaj and A. Belghith, "Detecting Arabic cyberbullying tweets using machine learning," *Mach. Learn. Knowl. Extraction*, vol. 5, no. 1, pp. 29–42, Jan. 2023.

[6] E. Bashir and M. Bouguessa, "Data mining for cyberbullying and harassment detection in Arabic texts," *Int. J. Inf. Technol. Comput. Sci.*, vol. 13, no. 5, pp. 41–50, Oct. 2021.

[7] R. Ying, Y. Shou, and C. Liu, "Prediction model of Dow Jones index based on LSTM-AdaBoost," in *Proc. Int. Conf. Commun., Inf. Syst. Comput. Eng. (CISCE)*, May 2021, pp. 808–812.

[8] Y. Bai, J. Xie, D. Wang, W. Zhang, and C. Li, "A manufacturing quality prediction model based on AdaBoost-LSTM with rough knowledge," *Comput. Ind. Eng.*, vol. 155, May 2021, Art. no. 107227.

[9] S. Sun, Y. Wei, and S. Wang, "AdaBoost-LSTM ensemble learning for financial time series forecasting," in *Proc. 18th Int. Conf. Comput. Sci. (ICCS)*. Wuxi, China: Springer, Jun. 2018, pp. 590–597.

[10] A. Muneer and S. M. Fati, "A comparative analysis of machine learning techniques for cyberbullying detection on Twitter," *Future Internet*, vol. 12, no. 11, p. 187, Oct. 2020.

[11] B. A. H. Murshed, J. Abawajy, S. Mallappa, M. A. N. Saif, and H. D. E. Al-Ariki, "DEA-RNN: A hybrid deep learning approach for cyberbullying detection in Twitter social media platform," *IEEE Access*, vol. 10, pp. 25857–25871, 2022.

[12] M. Alotaibi, B. Alotaibi, and A. Razaque, "A multichannel deep learning framework for cyberbullying detection on social media," *Electronics*, vol. 10, no. 21, p. 2664, Oct. 2021.

[13] S. Bharti, A. K. Yadav, M. Kumar, and D. Yadav, "Cyberbullying detection from tweets using deep learning," *Kybernetes*, vol. 51, no. 9, pp. 2695–2711, Sep. 2022.

[14] T. Mahlangu and C. Tu, "Deep learning cyberbullying detection using stacked embbedings approach," in *Proc. 6th Int. Conf. Soft Comput. Mach. Intell. (ISCMI)*, Nov. 2019, pp. 45–49.

[15] M. A. Al-Ajlan and M. Ykhlef, "Optimized Twitter cyberbullying detection based on deep learning," in *Proc. 21st Saudi Comput. Soc. Nat. Comput. Conf. (NCC)*, Apr. 2018, pp. 1–5.

[16] Y. Fang, S. Yang, B. Zhao, and C. Huang, "Cyberbullying detection in social networks using bi-GRU with self-attention mechanism," *Information*, vol. 12, no. 4, p. 171, Apr. 2021.

[17] M. Agbaje and O. Afolabi, "Neural network-based cyber-bullying and cyber-aggression detection using Twitter text," *Res. Square*, Jul. 2020, doi: 10.21203/rs.3.rs-1878604/v1.

[18] R. R. Dalvi, S. Baliram Chavan, and A. Halbe, "Detecting a Twitter cyberbullying using machine learning," in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2020, pp. 297–301.

[19] A. Setyanto, A. Laksito, F. Alarfaj, M. Alreshoodi, Kusrini, I. Oyong, M. Hayaty, A. Alomair, N. Almusallam, and L. Kurniasari, "Arabic language opinion mining based on long short-term memory (LSTM)," *Appl. Sci.*, vol. 12, no. 9, p. 4140, Apr. 2022.

[20] F. Shen, X. Zhao, G. Kou, and F. E. Alsaadi, "A new deep learning ensemble credit risk evaluation model with an improved synthetic minority oversampling technique," *Appl. Soft Comput.*, vol. 98, Jan. 2021, Art. no. 106852.

[21] U. Chandran, S. Kumarasamy, R. Samikannu, M. P. E. Rajamani, V. Krishnamoorthy, and S. Murugesan, "Tournament selected glowworm swarm optimization based measurement of selective harmonic elimination in multilevel inverter for enhancing output voltage and current," *Math. Problems Eng.*, vol. 2022, Mar. 2022, Art. no. 5845249.

[22] DataTurks. (2018). *Tweets Dataset for Detection of Cyber-Trolls*. Kaggle. [Online]. Available: https://www.kaggle.com/datatruks/dataset-for-detection-of-cybertrolls