Adaptive Ensemble Learning with a Fine-Tuned Framework for Cyberbullying Detection in Cross-Platform Social Media Environments

I. Anand Raj
PG & Research Department of Computer Science and Artificial Intelligence, St. Joseph's College of Arts & Science (Autonomous), Affiliated to Annamalai University, Cuddalore, Tamilnadu, India
ammayesuraja75@gmail.com (corresponding author)

R. Vidya
PG & Research Department of Computer Science and Artificial Intelligence, St. Joseph's College of Arts & Science (Autonomous), Affiliated to Annamalai University, Cuddalore, Tamilnadu, India
vidya.sjc@gmail.com

Received: 27 September 2025 | Revised: 3 November 2025, 26 November 2025, 1 December 2025, and 4 December 2025 | Accepted: 5 December 2025

 DOI: https://doi.org/10.48084/etasr.15164

ABSTRACT

Recently, on social media platforms, cyberbullying has been a significant concern for organizations, individuals, and society, and its detection is gaining increasing attention. The simplicity of access to social media networks such as Instagram, Twitter, and Facebook has led to an exponential upsurge in the exploitation of individuals in the form of bullying, toxic comments, sexism, hateful messages, racism, harassment, aggressive content, etc. Machine learning models are trained to identify and flag latent cyberbullying content, in addition to recognizing behavioral patterns that are suggestive of cyberbullying. This study presents an Ensemble Learning for Cyberbullying Detection across Social Media Platforms Using Word Vector Representations (ELCDSMP-WVR) approach. Initially, text preprocessing is performed in three-levels. The GloVe method is employed for the word embedding process. An Ensemble Voting Classifier (EVC) integrates three advanced DL techniques, a Temporal Convolutional Network (TCN), a Graph Wasserstein Autoencoder (GWAE), and a Deep Belief Network (DBN), to improve the robustness of the classification process. Finally, the Black-Winged Kite Optimization Algorithm (BKA) is employed to improve overall performance. A comparison study of the ELCDSMP-WVR technique showed superior accuracies of 95.27% and 97.88% over existing approaches on two cyberbullying datasets.

Keywords-cyberbullying detection; social media platforms; ensemble learning; black-winged kite optimisation algorithm; word vector representation

I. INTRODUCTION

Social networks have become an integral part of life, and the increased use of these platforms has created opportunities for connection but also misuse [1]. Cyberbullying has adverse psychological effects, and manual detection is impractical given the scale and speed of online interactions [2], making automated solutions based on Machine Learning (ML) increasingly crucial [3]. Conventional ML methods can detect inappropriate and abusive patterns [4], but single models encounter difficulties in accuracy and generalization [5]. Ensemble methods, such as stacking [6], integrate diverse models to enhance performance and provide more robust predictions [7]. By using multiple models, detection accuracy can be improved, leading to safer online environments. With the continuous growth of online communication, the nature of cyberbullying becomes more intrinsic [8], threatening detection systems. This can be handled by developing advanced and adaptive models for creating safer online environments [9], finally mitigating online abuse and its psychological and social impacts [10].

In [11], supervised ML and NLP approaches were used to automatically detect cyberbullying. In [12], Sentiment Analysis (SA) was performed using ML models to detect cyberbullying. In [13], a Heterogeneous GNN (HGNN) model was used for session-based cyberbullying detection. In [14], DL and ML techniques were utilized for feature extraction and cyberbullying detection. The CDNN architecture in [15] integrated numerous convolutional layers. In [16], the Fuzzy Adaptive Equilibrium Optimizer (FAEO) was integrated with an Extended CNN (ECNN) to automatically discover hidden topics. In [3], an ensemble stacking technique was proposed, using Word2Vec with CBOW for feature extraction, and BERT-M, an adapted version of BERT.

This study presents an Ensemble Learning for Cyberbullying Detection across Social Media Platforms Using Word Vector Representations (ELCDSMP-WVR) approach with the following key contributions:

- The preprocessing phase involves a three-level process, namely noise removal, stop-word elimination, and content transformation. This step improves data quality and consistency for more accurate detection. It also standardizes diverse social media inputs and significantly improves the performance of the downstream models.
- GloVe embedding is used for efficiently capturing semantic relationships between words in a dense vector format. DL is also utilized for comprehending context and variances in social media posts, improving the accuracy and robustness of cyberbullying detection.
- The TCN, GWAE, and DBN models are integrated in an ensemble voting classifier optimized by BKA to enhance robustness. This method fine-tunes training and hyperparameters, improving convergence and reducing overfitting. It also improves accuracy, efficiency, and reliability in cross-platform cyberbullying detection.
- A novel adaptive ensemble learning framework incorporates TCN, GWAE, and DBN models for capturing temporal, graph-based, and probabilistic features. The novelty is in the incorporation of BKA for improved performance. This fusion is specifically designed for accurate and robust cross-platform cyberbullying detection.

II. METHODOLOGY

Figure 1 represents the entire flow of the ELCDSMP-WVR model.

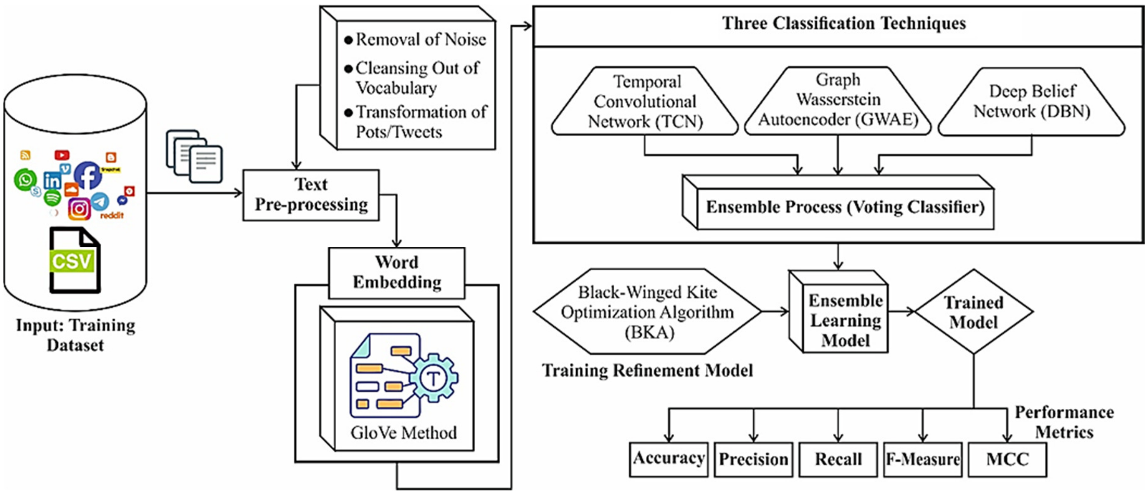

Fig. 1. Entire workflow of the ELCDSMP-WVR model.

A. *Text Preprocessing*

The text pre-processing step involves three levels. In general, social media data, such as tweets, Instagram comments, and posts, are concise and noisy [16], including a wide variety of inconsistencies, such as acronyms, emotions, and typos. Original data also suffer from elongated words, spelling mistakes, slang, word boundary errors, and concatenated words. First, noise, such as hashtags, mentions, punctuation, symbols, URLs, and emoticons, is removed. Next, OoV issues are addressed through acronym expansion, slang correction, spell checking, and repeated character reduction. Finally, text is transformed through lowercase, tokenization, stemming, and stop-word removal.

Text preprocessing begins with the removal of noise by converting emoticons, removing extra symbols and punctuation, replacing URLs, and deleting unnecessary characters. This is followed by out-of-vocabulary cleansing, which addresses issues like elongated words (e.g., converting "looooooove" to "love"), splitting concatenated words (e.g., "IHateBlackPeople" into "I", "Hate", "Black", "People"), and modifying slang or acronyms using a slang dictionary (e.g., converting "luv" to "love" or "gr8" to "great"). The text is standardized by transforming tweets through lowercase conversion, tokenization (breaking tweets into individual words or tokens), stop-word removal (common words like "is", "the", "with" that add little semantic value), and stemming (reducing words to their base or root form). Collectively, these preprocessing steps are crucial for converting noisy, informal social media text into structured and analyzable data.

B. *GloVe-Based Word Embedding*

The GloVe method is then utilized for the numeric representation of text [17]. This model was chosen for capturing global word co-occurrence statistics, giving richer semantic relationships than local context-based methods like Word2Vec. Performance is improved by comprehending complex language in social media text, and the pre-trained embeddings present an efficient and accurate representation. GloVe is a word embedding technique that combines global statistics with local context features. As a regression-based, unsupervised model, it captures word co-occurrence patterns and learns context-aware vectors using gradient descent. GloVe embeddings, learned through random gradient descent after error convergence, improve attention by mitigating the effect of rare word pairs.

C. EVC Model

This ensemble model integrates multiple methods to improve classification performance and mitigate the drawbacks of individual models. A fusion model is used to capture temporal dependencies. GWAE is efficient for modelling complex graph structures, and DBN extracts deep hierarchical features. This incorporation improves robustness and accuracy compared to single-model approaches, making it appropriate for the diverse and intrinsic behavior of social media data. EVC improves classification by combining the outputs of multiple classifiers, using hard or soft voting. In hard voting, the final class prediction is made based on the majority vote as:

$$C_{final}(x) = \text{argmax} \sum_{i-1}^{n} 1\,[C_i(x) = c] \tag{2}$$

where c denotes a class label, $C_i(x)$ represents the prediction of the i-th classifier, and $1[\cdot]$ signifies an indicator function that matches 1 when the condition is true and 0 otherwise. The class receiving the most votes is selected. In soft voting, the final prediction is based on averaging the predicted probabilities as:

$$C_{\text{final}}(x) = \text{argmax} \sum_{i-1}^{n} w_i\, P[C_i = c|x] \tag{3}$$

where $P[C_i = c|x]$ shows the predicted probability of class c from the classifier C_i, and w_i denotes the weight [18].

TCN is a DL model for sequential data that captures temporal dependencies using causal and dilated convolutions, enabling parallel processing while preserving input order. TCN processes sequential data using causal and dilated convolutions. The basic output at time t is calculated using:

$$y_t = \sum_{i=0}^{k-1} W_i\, x_{t-i} \tag{4}$$

where W_i indicates a weighted convolutional filter, y_t depicts output at time t, k signifies the filter size, and x_{t-i} depicts the previous inputs. The receptive fields are expanded by dilated convolutions without deepening the network, enabling long-range dependency learning:

$$y_t = \sum_{i=0}^{k-1} W_i\, x_{t-di} \tag{5}$$

Residual connections stabilize training by integrating layer outputs with earlier representations, preventing gradient issues [19-20]. GWAE utilizes hierarchical graph convolutions to update node features by aggregating both local and global neighborhood data. The encoder learns a latent distribution $Q(Z|X)$, from which the decoder reconstructs the original graph. The training objective combines reconstruction loss $c\big(X, G(Z)\big)$ to evaluate decoder accuracy with a regularization term $D\big(Q(Z|X), P(Z)\big)$ that aligns the latent space distribution to a prior $P(Z)$. The total loss is given by:

$$D_{WAE}\big(P(X), P(X|Z)\big) = \inf \lambda D\big(Q(Z|X), P(Z)\big) + \mathbb{E}_{P_X}\mathbb{E}_{Q(Z|X)}\big[c\big(X, G(Z)\big)\big] \tag{6}$$

The reconstruction and regularization terms are balanced by λ, ensuring accurate graph reconstruction and a structured latent space. Additionally, the DBN uses stacked RBMs to learn hierarchical features unsupervised, aiding pattern recognition before classification.

D. Fine-Tuned Optimization Using BKA

Finally, the fine-tuning process is performed by using BKA [21]. This is a nature-inspired population-intelligence optimization model that is motivated by the fact that Black-Winged Kites (BKs) show intelligent behaviors and are highly adaptive during attack and migration. This method highlights robust global search capability and a balanced exploration-exploitation mechanism. The model also outperforms conventional optimizers in avoiding local minima, making it ideal for fine-tuning intrinsic DL techniques.

1) Population Initialization

In BKA, generating a set of arbitrary solutions is the primary step for initializing the population. The succeeding matrix is employed to depict the position of every BK:

$$BK = \begin{bmatrix} BK_{1,1} & BK_{2,2} & \dots & \dots & BK_{1,dim} \\ BK_{2,1} & BK_{2,2} & \dots & \dots & BK_{2,dim} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ BK_{m,1} & BK_{m,2} & \dots & \dots & BK_{m,dim} \end{bmatrix} \tag{7}$$

where dim denotes the dimension of the specified problem, BK_{ij} indicates the j-th dimension of the i-th BK, and m depicts the potential solution counts. Every BK is allocated a location:

$$X_i = BK_{lb} + rand(BK_{ub} - BK_{lb}) \tag{8}$$

where BK_{ub} and BK_{lb} refer to the upper and lower bounds of i-th BK in the j-th dimension, i signifies an integer between one and m, and $rand$ is a random value among [0, 1].

2) Attack Behavior

BKs exhibit dual attack approaches, hovering in the air, waiting for an attack, and looking for prey. This approach comprises various attack behaviors employed for global searching and exploration. The mathematical expression of BK's attack behavior is given by:

$$n = 0.05 \times e^{-2\times\left(\frac{t}{T}\right)^2} \tag{9}$$

$$y_{t+1}^{i,j} = \begin{cases} y_t^{i,j} + n\big(1+\sin(r)\big) \times y_t^{i,j}, & p < r \\ y_t^{i,j} + n \times (2r-1) \times y_t^{i,j}, & \text{else} \end{cases} \tag{10}$$

where $y_t^{i,j}$ and $y_{t+1}^{i,j}$ represent the position of the i-th BK in the j-th dimension in the t-th and $t+1$-th iterations, respectively, p is a constant value of 0.9, r is a random number in [0, 1], t is the iteration count, n is a balance parameter, and T indicates the entire iteration count.

3) Migration Behavior

Migration is guided by a leader who steps down if outperformed by a random individual; otherwise, they lead the group. This ensures adaptive and effective migration.

$$m = 2 \times \sin(r + \pi/2) \quad (11)$$

$$y_{t+1}^{i,j} = \begin{cases} y_t^{i,j} + C(0,1) \times (y_t^{i,j} - L_t^j), & F_i < F_{ri} \\ y_t^{i,j} + C(0,1) \times (L_t^j - m \times y_t^{i,j}), & \text{else} \end{cases} \quad (12)$$

Here, L_t^j signifies the leader of BK in dimension j in the t-th iteration, $y_t^{i,j}$ and $y_{t+1}^{i,j}$ refer to the position of the i-th BK in dimension j at iteration t and $t+1$, respectively, F_j is the current position in dimension j at the t-th iteration, F_{ri} is the fitness of a random position in dimension j at the same iteration, and $C(0,1)$ denotes a Cauchy mutation.

Fitness selection is key to BKA to evaluate candidate efficiency. Precision is used to estimate the fitness function, as:

$$Fitness = \max(P) \quad (13)$$

$$P = \frac{TP}{TP+FP} \quad (14)$$

where TP and FP depict the true and false positive values.

III. PERFORMANCE ANALYSIS AND DISCUSSION

The experimental analysis of the ELCDSMP-WVR method was investigated on [22] (Dataset 1) and [23] (Dataset 2). These datasets were gathered from various sources associated with the automated detection of cyberbullying. This data consists of diverse kinds of cyberbullying, such as insults, hate speech, toxicity, and aggression.

The proposed method was run using Python 3.6.5 on an i5-8600K CPU, 4 GB GPU, 16 GB RAM, 250 GB SSD, and 1 TB HDD, using a learning rate of 0.01, ReLU, 50 epochs, 0.5 dropout, and a batch size of 5. Table I describes the datasets, and Table II shows sample texts. The dataset already includes annotations, and although some of them may vary in interpretation. Table III compares the ELCDSMP-WVR model on both datasets, using results from the literature verified through the experimental setup to ensure accuracy [24-26]. The proposed ELCDSMP-WVR model attained the highest accuracy of 95.27%. In contrast, the LR, SVM, BiLSTM, RNN, Multinomial NB, SOSNet+SBERT, and Naïve Bayes (NB) models achieved lower accuracies at 88.61%, 86.00%, 95.08%, 87.98%, 87.08%, 95.02%, and 91.12%, respectively.

TABLE I. DETAILS OF DATASETS

CBLabel	Dataset 1 [22]	Dataset 2 [23]
Cyberbullying_No (0)	5000	11501
Cyberbullying _Yes (1)	126	5347
Total	**5126**	**16848**

TABLE II. SAMPLE TEXTS

Cyberbullying	Label	Sample texts: Dataset 1 [22]	Sample texts: Dataset 2 [23]
Cyberbullying_No	0	Thanks. Anyway; the answer is yes.	I'm thinking #MKR will go past Christmas Eve and get into 2016 at this rate!
Cyberbullying_No	0	Pictured NeededWhat do these people look like? There should be a picture.	I wish they went to death row instead of sudden death #Mkr
Cyberbullying_yes	1	You must be a terrosist yourself to revert thta. But what could i expect from someone of your ethnicity	I am so glad those annoying bitchy blonde thots got kicked off #MKR
Cyberbullying_yes	1	You either know that is a lie or you have not checked.	Its not the only thing she's done a lot of I'm guessing #mkr

TABLE III. COMPARATIVE ANALYSIS OF ELCDSMP-WVR

Method	$Accu_y$	$Prec_n$	$Reca_l$	$F_{Measure}$
Dataset 1 [22]				
LR Method [24]	88.61	86.62	93.41	85.02
SVM [24]	86.00	80.38	89.75	85.04
BiLSTM [25]	95.08	80.08	93.31	89.94
RNN [25]	87.98	84.52	85.69	91.50
Multinominal NB [25]	87.08	84.08	87.25	91.98
SOSNet+SBERT [26]	95.02	81.79	94.47	90.55
NB [26]	91.12	83.64	88.29	91.77
ELCDSMP-WVR (Proposed)	95.27	90.08	95.27	92.51
Dataset 2 [23]				
Word2vec [25]	95.08	89.38	86.38	97.40
BERT [25]	95.92	97.58	93.34	93.64
MLPNN [25]	87.70	95.50	88.94	96.26
GRU [25]	90.96	88.22	91.72	89.28
SGD [25]	89.82	96.96	91.75	93.18
Adaboost [26]	91.68	96.17	95.07	88.47
SA+ML+SMOTE [26]	88.39	87.52	87.34	86.02
ELCDSMP-WVR (Proposed)	97.88	98.04	97.04	97.52

Based on $F_{Measure}$, the ELCDSMP-WVR technique achieved 92.51%. Likewise, the LR, SVM, BiLSTM, RNN, Multinomial NB, SOSNet+SBERT, and NB models achieved lower $F_{Measure}$ of 85.02%, 85.04%, 89.94%, 91.50%, 91.98%, 90.55%, and 91.77%, respectively.

IV. CONCLUSION

This study presented the ELCDSMP-WVR approach, which involves text pre-processing, GloVe-based word embedding, fusion of TCN, GWAE, and DBN-based classification, and BKA-based tuning. The comparison study of the ELCDSMP-WVR technique showed a superior accuracy of 95.27% and 97.88% over existing approaches on two cyberbullying datasets. The limitations comprise reliance on static datasets, which may not fully capture cyberbullying or handle sarcasm. Also, difficulties with scalability can be seen across new or less popular social media platforms. Future work may explore real-time detection, multilingual support, and continuous learning for dynamic content environments.

REFERENCES

[1] T. Ahmed, M. Kabir, S. Ivan, H. Mahmud, and K. Hasan, "Am I Being Bullied on Social Media? An Ensemble Approach to Categorize

Cyberbullying," in *2021 IEEE International Conference on Big Data (Big Data)*, Orlando, FL, USA, Sept. 2021, pp. 2442–2453, https://doi.org/10.1109/BigData52589.2021.9671594.

[2] K. S. Alam, S. Bhowmik, and P. R. K. Prosun, "Cyberbullying Detection: An Ensemble Based Machine Learning Approach," in *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Tirunelveli, India, Oct. 2021, pp. 710–715, https://doi.org/10.1109/ICICV50876.2021.9388499.

[3] A. Muneer, A. Alwadain, M. G. Ragab, and A. Alqushaibi, "Cyberbullying Detection on Social Media Using Stacking Ensemble Learning and Enhanced BERT," *Information*, vol. 14, no. 8, Aug. 2023, Art. no. 467, https://doi.org/10.3390/info14080467.

[4] A. F. Alqahtani and M. Ilyas, "A Machine Learning Ensemble Model for the Detection of Cyberbullying." arXiv, Feb. 19, 2024, https://doi.org/10.48550/arXiv.2402.12538.

[5] Y. J. N. Kumar *et al.*, "Detecting cyberbullying in social media using text analysis and ensemble techniques," *E3S Web of Conferences*, vol. 507, Nagpur, India, 2024, Art. no. 01069, https://doi.org/10.1051/e3sconf/202450701069.

[6] D. Van Bruwaene, Q. Huang, and D. Inkpen, "A multi-platform dataset for detecting cyberbullying in social media," *Language Resources and Evaluation*, vol. 54, no. 4, pp. 851–874, Dec. 2020, https://doi.org/10.1007/s10579-020-09488-3.

[7] S. Abarna, J. I. Sheeba, and S. P. Devaneyan, "A novel ensemble model for identification and classification of cyber harassment on social media platform," *Journal of Intelligent & Fuzzy Systems*, vol. 45, no. 1, pp. 13–36, July 2023, https://doi.org/10.3233/JIFS-230346.

[8] S. Unnava and S. R. Parasana, "A Study of Cyberbullying Detection and Classification Techniques: A Machine Learning Approach," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15607–15613, Aug. 2024, https://doi.org/10.48084/etasr.7621.

[9] A. T. Azar, H. M. Noori, A. R. Mahlous, A. Al-Khayyat, and I. K. Ibraheem, "Quasi-Reflection Learning Arithmetic Firefly Search Optimization with Deep Learning-based Cyberbullying Detection on Social Networking," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 17162–17169, Oct. 2024, https://doi.org/10.48084/etasr.8314.

[10] S. S. Alzahrani, "Data Mining Regarding Cyberbullying in the Arabic Language on Instagram Using KNIME and Orange Tools," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9364–9371, Oct. 2022, https://doi.org/10.48084/etasr.5184.

[11] B. Joshi, B. K. Joshi, S. Pant, A. Kumar, and H. K. Sharma, "An Efficient Method for Detecting Cyberbullying Using Supervised Machine Learning Techniques," *Procedia Computer Science*, vol. 258, pp. 1254–1261, Jan. 2025, https://doi.org/10.1016/j.procs.2025.04.359.

[12] M. F. Almufareh, N. Z. Jhanjhi, M. Humayun, G. Alwakid, D. Javed, and S. N. Almuayqil, "Integrating Sentiment Analysis With Machine Learning for Cyberbullying Detection on Social Media," *IEEE Access*, vol. 13, pp. 78348–78359, 2025, https://doi.org/10.1109/ACCESS.2025.3558843.

[13] M. Buyankhishig, T. Shwe, I. Mendonça, and M. Aritsugi, "Heterogeneous Graph Neural Network Framework for Session-Based Cyberbullying Detection," *IEEE Access*, vol. 13, pp. 110926–110940, 2025, https://doi.org/10.1109/ACCESS.2025.3583332.

[14] A. Almomani, K. Nahar, M. Alauthman, M. A. Al-Betar, Q. Yaseen, and B. B. Gupta, "Image cyberbullying detection and recognition using transfer deep machine learning," *International Journal of Cognitive Computing in Engineering*, vol. 5, pp. 14–26, Jan. 2024, https://doi.org/10.1016/j.ijcce.2023.11.002.

[15] S. Niranjana, D. V. Babu, K. Malathy, S. D. Devi, K. K. Babu, and C. Karthikeyan., "CDNN: A Novel Methodology Development to Detect Online Social Network Cyberbullying Threats Using Convoluted Deep Neural Network Principle," in *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India, Sept. 2024, pp. 1–5, https://doi.org/10.1109/ICSES63760.2024.10910787.

[16] B. A. H. Murshed, Suresha, J. Abawajy, M. A. N. Saif, H. M. Abdulwahab, and F. A. Ghanem, "FAEO-ECNN: cyberbullying detection in social media platforms using topic modelling and deep learning," *Multimedia Tools and Applications*, vol. 82, no. 30, pp. 46611–46650, Dec. 2023, https://doi.org/10.1007/s11042-023-15372-3.

[17] H. Çetiner and H. Yüksel, "CBLTwitter: Twitter disaster detection analysis using CNN-BiLSTM deep learning methods," *Gümüşhane Üniversitesi Fen Bilimleri Dergisi*, vol. 15, no. 2, pp. 563–576, June 2025, https://doi.org/10.17714/gumusfenbil.1653072.

[18] S. Al-Dahidi *et al.*, "Multistep PV power forecasting using deep learning models and the reptile search algorithm," *Results in Engineering*, vol. 27, Sept. 2025, Art. no. 106265, https://doi.org/10.1016/j.rineng.2025.106265.

[19] G. Shishaev, V. Demyanov, and D. Arnold, "History Matching under Uncertainty of Geological Scenarios with Implicit Geological Realism Control with Generative Deep Learning and Graph Convolutions." arXiv, July 14, 2025, https://doi.org/10.48550/arXiv.2507.10201.

[20] H. S. Kauser and M. V. Anu, "Hybrid deep learning model for accurate and efficient android malware detection using DBN-GRU," *PLOS ONE*, vol. 20, no. 5, 2025, Art. no. e0310230, https://doi.org/10.1371/journal.pone.0310230.

[21] J. Gao *et al.*, "A Short-Term Wind Power Forecasting Method Based on Adaptive Bka-Tcn-Bilstm Hybrid Model with AP Clustering." Social Science Research Network, July 14, 2025, https://doi.org/10.2139/ssrn.5350915.

[22] N. Ejaz, S. Choudhury, and F. Razi, "Towards Comprehensive Cyberbullying Detection: A Dataset Incorporating Aggressive Texts, Repetition, Peerness, and Intent to Harm," Nov. 2023, https://doi.org/10.17632/wmx9jj2htd.1.

[23] "Cyberbullying Dataset." Kaggle, Available: https://www.kaggle.com/datasets/saurabhshahane/cyberbullying-dataset.

[24] Y. M. Ibrahim, R. Essameldin, and S. M. Saad, "Social Media Forensics: An Adaptive Cyberbullying-Related Hate Speech Detection Approach Based on Neural Networks With Uncertainty," *IEEE Access*, vol. 12, pp. 59474–59484, 2024, https://doi.org/10.1109/ACCESS.2024.3393295.

[25] E. Y. Daraghmi, S. Qadan, Y. A. Daraghmi, R. Yousuf, O. Cheikhrouhou, and M. Baz, "From Text to Insight: An Integrated CNN-BiLSTM-GRU Model for Arabic Cyberbullying Detection," *IEEE Access*, vol. 12, pp. 103504–103519, 2024, https://doi.org/10.1109/ACCESS.2024.3431939.

[26] M. F. Almufareh, N. Z. Jhanjhi, M. Humayun, G. N. Alwakid, D. Javed, and S. N. Almuayqil, "Integrating Sentiment Analysis With Machine Learning for Cyberbullying Detection on Social Media," *IEEE Access*, vol. 13, pp. 78348–78359, 2025, https://doi.org/10.1109/ACCESS.2025.3558843.