Journal of Applied Informatics Science
Open Access: https://doi.org/10.65897/jais.v1.i1.80
Vol. 2, Issue. 1, 2026, pp. 19-23
e-ISSN: 3110-7451

XSentiment-HS: Hierarchical CNN-BiGRU-SVM with Explainable for Indonesian Multi-Level Hate Speech Detection

XSentiment-HS: Hierarchical CNN-BiGRU-SVM dengan Explainable untuk Deteksi Multi-Tingkat Ujaran Kebencian Indonesia

Gina Khayatun Nufus[1]*, Rizki Dewantara[2] , Ardi Susanto[3], Sokid[4], Lia Farhatuaini[5], Mohammad Raihan Akbar[6], Jaka Septiadi[7]

[1,2,3,4,5,6] *Informatika, UIN Siber Syekh Nurjati Cirebon, Cirebon, Indonesia,* [7] *Pendidikan Sistem dan Teknologi Informasi, Universitas Pendidikan Indonesia, Purwakarta, Indonesia*
[1]ginakhn@uinssc.ac.id, [2]dewantararizki@uinssc.ac.id, [3]ardisusanto@uinssc.ac.id, [4]sokid@uinssc.ac.id, [5]farhatuaini@uinssc.ac.id, [6]raihan13akbar@gmail.com, [7]jakaseptiadi@upi.edu

ARTICLE INFO

Article history:
Received : 2026-03-13
Revised : 2026-03-30
Accepted : 2026-04-01

Kata Kunci:
CNN-BiGru, Deteksi Ujaran Kebencian, Deep Learning, XSentiment-HS

Keywords:
CNN-BiGRU, Deep Learning, Hate Speech Detection, XSentiment-HS

This is an open access article under the CC BY-SA license.
Copyright © 2026 by Author. Published by GWS Tech Solution.

ABSTRAK

Deteksi ujaran kebencian pada media sosial menuntut interpretasi teks yang kompleks karena sifatnya yang spontan dan ambigu, terutama dalam bahasa Indonesia yang kaya akan slang. Tantangan utama saat ini adalah keterbatasan penelitian sebelumnya yang mayoritas hanya melakukan klasifikasi biner tanpa mendeteksi tingkat keparahan konten. Penelitian ini mengusulkan XSentiment-HS, sebuah model deep learning hierarkis dua tahap untuk deteksi multi-tingkat hate speech. Arsitektur model menggabungkan Convolutional Neural Networks (CNN) untuk ekstraksi fitur lokal dan Bidirectional Gated Recurrent Unit (BiGRU) untuk menangkap ketergantungan kontekstual jangka panjang. Model ini juga diperkuat dengan mekanisme Multi-Head Attention dan Support Vector Machine (SVM) sebagai classifier final. Melalui integrasi ini, XSentiment-HS diharapkan mampu mengatasi tantangan ekstraksi fitur dan polisemi secara lebih efektif dibandingkan metode konvensional.

ABSTRACT

*Hate speech detection in social media requires complex text interpretation due to its spontaneous and ambiguous nature, particularly in the Indonesian language, which is rich in slang and cultural nuances. A significant challenge in current research is the predominance of binary classification models that fail to identify the severity levels of detected hate speech. To address this limitation, this study proposes **XSentiment-HS**, a two-stage hierarchical deep learning model for multi-level hate speech detection. The architecture integrates Convolutional Neural Networks (CNN) for local feature extraction and Bidirectional Gated Recurrent Units (BiGRU) to capture long-range contextual dependencies. The model is further enhanced with a Multi-Head Attention mechanism and utilizes a Support Vector Machine (SVM) as the final classifier. Through this integration, XSentiment-HS is expected to address feature extraction and polysemy challenges more effectively than conventional methods.*

*Corresponding author.
E-mail addresses: ginakhn@uinssc.ac.id (First Author)

1. PENDAHULUAN

Pertumbuhan penggunaan media sosial di Indonesia telah mendorong peningkatan signifikan dalam penyebaran konten berbahaya, termasuk ujaran kebencian (hate speech). Twitter (kini X) menjadi salah satu platform yang paling rentan terhadap penyebaran konten tersebut, terutama dalam diskusi politik, agama, dan isu sosial yang sensitif. Kondisi ini menuntut pengembangan sistem deteksi otomatis yang tidak hanya akurat tetapi juga mampu memberikan penjelasan (interpretable) kepada moderator konten dan pengambil kebijakan. Deteksi ujaran kebencian pada teks berbahasa Indonesia menghadapi tantangan serius, seperti kespontanan, singkatan, ambiguitas, dan penggunaan slang atau nuansa budaya lokal yang kompleks.

Penelitian sebelumnya di bidang ini telah menunjukkan kemajuan yang berarti, terutama dengan kehadiran dataset komprehensif seperti karya Ibrohim & Budi [1] yang menyediakan anotasi multi-label mencakup target dan tingkat keparahan. Namun, pendekatan yang umumnya digunakan masih memiliki beberapa keterbatasan. Pertama, mayoritas penelitian hanya berfokus pada klasifikasi biner (toksik dan non-toksik) tanpa mengklasifikasikan tingkat keparahan secara spesifik (Weak, Moderate, Strong), padahal informasi ini sangat krusial untuk penanganan moderasi konten yang bertingkat. Kedua, model yang digunakan masih banyak yang mengandalkan metode tradisional seperti Support Vector Machine (SVM) dengan fitur n-gram dan TF-IDF,yang tidak mampu menangkap representasi semantik mendalam. Ketiga, sebagian besar model bersifat "black-box" sehingga tidak memberikan mekanisme eksplainabilitas yang transparan mengenai mengapa sebuah teks diklasifikasikan sebagai ujaran kebencian.

Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan XSentiment-HS, sebuah model deep learning hierarkis dua tahap yang dikembangkan dari framework XSentiment sebelumnya. Kontribusi utama penelitian ini meliputi: (1) pengembangan arsitektur hierarkis dua tahap yang secara bersamaan mendeteksi konten toksik dan mengklasifikasikan tingkat keparahannya; (2) pemanfaatan Word2Vec CBOW secara bersama (shared embedding) untuk menjaga konsistensi representasi semantik antar kedua tahap; (3) integrasi tiga lapisan Explainable AI (XAI), yaitu Attention Visualization, SHAP, dan LIME, untuk memberikan penjelasan transparan pada level token dan fitur; serta (4) penerapan teknik SMOTE dan class weighting untuk menangani ketidakseimbangan kelas yang ekstrem pada data ujaran kebencian. Arsitektur CNN-BiGRU-SVM yang diusulkan terinspirasi dari keunggulan kombinasi ekstraksi fitur lokal dan dependensi jangka panjang pada penelitian sebelumnya [2].

Artikel ini disusun sebagai berikut: Bagian 2 memaparkan metodologi penelitian termasuk dataset, preprocessing, dan arsitektur model. Bagian 3 menyajikan hasil eksperimen beserta analisisnya. Bagian 4 memberikan kesimpulan serta arah penelitian ke depan.

2. METODE

2.1. Dataset

Penelitian ini menggunakan Indonesian Hate Speech and Abusive Language Detection dataset yang dikompilasi oleh Ibrohim & Budi (2019). Dataset terdiri dari 13.169 tweet berbahasa Indonesia yang dianotasi secara manual oleh tim anotator terlatih. Setiap tweet dianotasi dengan 13 label biner mencakup: HS (hate speech), Abusive, serta delapan subkategori target (Individual, Group, Religion, Race, Physical, Gender, Other) dan tiga level keparahan (Weak, Moderate, Strong).

Tabel 1. Statistik Dataset Indonesia Hate Speech

Kategori	Jumlah	Persentase
Total tweet	13.169	100%
Non-Toxic (HS=0 & Abusive=0	5.860	44,5%
Toxic(HS=1 atau Abusive =1)	7.309	55,5%
Week Severity	5.131	70,2% dari Toxic
Moderate Severity	1.705	23,3% dari Toxic
Strong Severity	473	6,5% dari Toxic

2.2. Prepossessing

Pipeline preprocessing dirancang khusus untuk menangani karakteristik teks Twitter berbahasa Indonesia secara berurutan: (1) normalisasi teks dasar (konversi ke lowercase, penghapusan URL, mention, hashtag prefix, dan karakter berulang); (2) normalisasi slang menggunakan kamus khusus yang berisi 80+ entri pemetaan (misal: gak/ga/ngga → tidak, yg → yang); (3) tokenisasi menggunakan nltk.word_tokenize; (4) stemming menggunakan PySastrawi; dan (5) penghapusan stopwords menggunakan daftar bawaan PySastrawi dengan modifikasi khusus berupa pengecualian kata negasi dan intensifier (tidak, bukan, jangan, sangat, banget) untuk mempertahankan polaritas semantik yang krusial dalam deteksi ujaran kebencian.

2.3. Representasi Teks (Word2Vec CBOW)

Representasi vektor kata dalam penelitian ini dihasilkan menggunakan algoritma Word2Vec dengan pendekatan Continuous Bag of Words (CBOW) [4]. Secara prinsip, CBOW mempelajari representasi vektor setiap kata dengan cara memprediksi kata target w(t) berdasarkan rata-rata vektor kata-kata konteks dalam jendela ukuran window = 4, yakni empat kata di sebelah kiri dan empat kata di sebelah kanan kata target. Fungsi objektif yang dioptimalkan selama training adalah log-likelihood rata-rata atas seluruh kata dalam corpus: $J = (1/T) \Sigma \log P(w(t) \mid \text{konteks}(t))$, di mana probabilitas output dihitung menggunakan negative sampling dengan k = 5 sampel negatif per kata target guna meningkatkan efisiensi komputasi. Bobot jaringan diperbarui melalui stochastic gradient descent (SGD) dengan learning rate yang dikurangi secara linear seiring bertambahnya epoch. Proses training dilaksanakan selama 15 epoch pada corpus data training Stage 1 yang terdiri dari 9.218 tweet dengan

47.832 token unik, sehingga mencegah terjadinya kebocoran data (data leakage) ke set validasi dan pengujian. Hasil training menghasilkan matriks embedding W ∈ ℝ^(V×d), dengan V = 12.405 entri kosakata yang memenuhi threshold min_count = 3 dan d = 200 sebagai dimensi ruang vektor. Setiap tweet selanjutnya direpresentasikan sebagai matriks sekuens T ∈ ℝ^(L×d), dengan L = 50 token sebagai panjang sekuens maksimum yang mencakup 99th percentile distribusi panjang tweet pada dataset. Token yang tidak ditemukan dalam kosakata dikodekan sebagai vektor nol, sedangkan tweet yang lebih pendek dari L diisi dengan zero-padding dan tweet yang melebihi L dipotong dari bagian akhir (right truncation). Matriks sekuens T inilah yang menjadi input langsung ke lapisan CNN pada setiap tahap model.

Mekanisme shared embedding merupakan keputusan arsitektural kunci dalam XSentiment-HS yang membedakannya dari pendekatan multi-stage konvensional. Alih-alih melatih dua matriks embedding secara independen untuk masing-masing tahap, penelitian ini membekukan (freeze) satu matriks W yang telah dilatih dan menggunakannya secara identik sebagai lapisan input bagi Stage 1 maupun Stage 2. Konsekuensi teknisnya adalah setiap token dipetakan ke vektor yang persis sama dalam ruang semantik berdimensi 200, baik ketika diproses oleh Stage 1 untuk menentukan kelas toxic atau non-toxic, maupun oleh Stage 2 untuk menentukan gradasi keparahan Weak, Moderate, atau Strong. Konsistensi ruang vektor ini memastikan kedua model beroperasi dalam basis representasi yang koheren, sehingga pola linguistik yang bersifat umum seperti penggunaan leksikon bermuatan negatif dan struktur kalimat agresif memiliki landasan semantik yang terpadu lintas tahap. Selain itu, shared embedding memberikan keuntungan dari sisi efisiensi parameter: penggunaan satu matriks embedding mengurangi total parameter yang dapat dilatih dari 2 × (V × d) = 4.962.000 menjadi (V × d) = 2.481.000, atau berkurang 50 persen. Pengurangan ini secara langsung menekan risiko overfitting pada Stage 2, yang memiliki data training lebih sedikit akibat proses pemfilteran kelas Non-Toxic pada Stage 1. Dengan demikian, mekanisme shared embedding tidak hanya berfungsi sebagai strategi efisiensi komputasi, melainkan juga berperan sebagai regularisasi implisit yang mendukung generalisasi model secara keseluruhan dalam arsitektur pipeline hierarkis.

2.4. Arsitektur Model XSentiment-HS

Xsentiment-HS terdiri dari dua instansi model yang identik secara struktural namun di training secara independen. Setiap instansi mengintegrasikan tiga komponen: (1) CBB Multi-Kernel dengan kernel [2,3,4] menghasilkan 128 feature map per kernel, diikuti Batch Normalization, ReLU, dan Adaptive Max Pooling menghasilkan representasi lokal 384-dimensi; (2) BiGRU dua lapis dengan 128 hidden unit dikombinasikan dengan Multi-Head Attention (4 heads) menghasilkan representasi sekuensial 256-dimensi; (3) Fused feature vector 640-dimensi diproyeksikan ke 320-dimensi sebagai input SVM. Aritektur ini mengacu pada framework CNN-BiGRU-SVM yang terbukti efektif oleh Selvi&Santosh Kumar [3] dan diadatasi untuk task hierarkis.

Tabel 2. Hyperparameter Model XSentiment-HS

Parameter	Stage 1	Stage 2
Embedding dim	200(shared)	200(shared)
CNN filters	128	128
Kernel sizes	[2,3,4]	[2,3,4]
GRU hidden units	128	128
Num. Attention heads	4	4
Dropout	0,4	0,4
Num. Classes	2	3
Batch size	64	64
Learning rate	$8x10^{-4}$	$8x10^{-4}$
SVM	1,0	1,0

2.5. Penanganan Ketidakseimbangan Kelas

Dataset hate speech menunjukkan ketidakseimbangan kelas yang signifikan pada stage 2, dimana kelas strong hanya 6,5% dari data toxic (rasio 1:10,8 terhadap Weak). Dua strategi diterapkan:(1) weighted CrossEntropyLoss pada fase neural training dengan class weight dihitung secara otomatis menggunakan formula balanced weighting; dan SMOTE (Synthetic, Minority Oversampling Technique)[5] pada fase SVM training untuk menghasilkan sampel sintesis bagi kelas minoritas.

2.6. Metrik Evaluasi

Evaluasi klasifikasi menggunakan Accuracy, Precission, Recall dan F1-score. Untuk stage 2 (3-class), F1-score dihitung dengan averaging='macro'. Tiga metrik XAI digunakan: (1) Attention-Prediction Alignment (APA) mengukur korelasi Spearman antara bobot kata LIME dan lexicon sentiment; (2) Explanation Stability (ES) mengukut Jaccard similarity atar run LIME dan (3) Explanation Fidelity (EF) menggunakan R^2 aproksimasi lokal LIME [6],[7].

3. Hasil dan pembahasan

3.1. Performa Stage 1 : Deteksi Toxic

Tabel 3 menyajikan perbandingan performa deteksi toxic. XSentiment-HS mencapai F1-score 87,00% dengan precission 88,36%, menggungguli baseline CNN-BiGRU-Softmax. Precission yang tinggi menandakan rendahnya false positive yang kritis dalam konteks moderasi konten untuk menghindari pemblokiran konten yang tidak semestinya.

Tabel 3. Perbandingan Performa Stage 1: Toxic Detection

Model	Accuracy	Precission	Recall	F1(Macro)
CNN-BiGRU-Softmax (Baseline)	86,60%	-	-	-
XSentiment-HS CNN-BiGRU-SVM	85,76%	88,36%	85,69%	87,00%

3.2. Performa Stage 2: Severity Grading

Tabel 4 menyajikan performa klasifikasi tigas kelas severity. XSentiment-HS mencapai macro F1-score 69,62% dengan accuracy 78,78%, melampaui baseline (78,23%). Kelas Strong merupakan tantangan terbesar akibat ketidakseimbangan ekstrem. Kombinasi SMOTE dan class weighting terbukti meningkatkan performa pada kelas minoritas. Kelas Weak memperoleh F! tertinggi (0,7676) karena distribusinya paling besar, sedangkan kelas Strong memperoleh F! terendah (0,6375) namun masih jauh lebih baik dibanding model tanpa penanganan imbalance data.

Tabel 4 Perbandingan Performa Stage 2: Severity Grading

Model	Accuracy	Precission	Recall	F1(Macro)
CNN-BiGRU-Softmax (Baseline)	78,23%	-	-	
XSentiment-HS CNN-BiGRU-SVM	78,78%	71,19%	68,28%	69,62%

Tabel 5 Perbandingan Performa Stage 2: Severity Grading

Model	Accuracy	Precission	Recall	F1(Macro)
Weak	0,7823	0,7534	0,7676	1.024
Moderate	0,6891	0,6124	0,6484	341
Strong	0,6643	0,6126	0,6375	95
Macro avg	0,7119	0,6595	0,6845	1.460
Moderate	0,6891	0,6124	0,6484	341

3.3. Evaluasi

Evaluasi end-to-end dilakukan dengan mengalirkan seluruh tweet test set melalui pipeline lengkap. Dari total 2.626 tweet yang diuji, 1.421 (54,1%) diprediksi Toxic dan diproses lebih lanjut oleh Stage 2: 998 dikategorikan Week, 332 Moderate, dan 91 Strong. Pipeline ini secara konseptual selaras dengan proses moderasi konten manual yang bersifat bertahap dari deteksi ke gradasi.

3.4. Metrik Eksplainabilitas

Tabel 6 merangkum metrik kuantitatif eksplainabilitas. Nilai APA yang negatif mengindentifikasikan bahwa model mengandalkan konteks kalimat secara keseluruhan daripada kata tunggal bermuatan lexicon, yang selaras dengan karakteristik hate speech berbahasa Indonesia yang sering menggunakan frasa implisit dan konteks budaya. Nilai ES= 0,650 menunjukkan konsistensi penjelasan LIME yang menandai untuk penggunaan produksi, sedangkan EF=0,498 menunjukkan bahwa aproksimasi LIME merepresentasikan sekitar 50% perilaku model lokai [8].

Tabel 6 Perbandingan Performa Stage 2: Severity Grading

Metrik	Stage 1	Stage 2	Interpretasi
APA(LIME Alignment)	-0,163 ± 0,978	-0,556 ±0,832	Model berbasis konteks, bukan lexicon
ES (Stability)	-	0,650±1,88	Penjelasan cukup konsisten
EF(Fidelity R^2)	-	0,498±0,240	Aproksimasi LIME cukup akurat

3.5. Perbandingan dengan Penelitian Terkait

Tabel 7 membandingkan XSentiment-HS dengan penelitian terkait. Model yang diusulkan merupakan satu-satunya yang menyediakan severity grading sekaligus eksplainabilitas tiga lapis. Dibanding Selvi & Santosh Kumar [3] yang mencapai akurasi 96% pada dataset bahasa Inggris (IMDb), XSentiment-HS menunjukkan performa konpetitif pada domain hate speech berbahasa Indonesia yang lebih kompleks dengan struktur hierarkis.

Tabel 8 Perbandingan dengan Penelitian Terkait

Penelitian	Dataset	Model	F1(%)	XAI	Severity
Ibrahin& Budi [1]	ID HS 13K	0,SVM+TFIDF	~75	Tidak	Tidak
Selvi & Santosh Kumar [3]	IMDb, SDC	CNN-BiGRU-SVM	~96*	Tidak	Tidak
Başarslan & Kayaalp [9]	Twitter	MBi-GRUMCONV	~90	Tidak	Tidak
Dang dkk[10]	Multi-domain	CNN+LSTM+SVM	~88	Tidak	Tidak
XSentiment-HS(Ours)	ID HS 13K	CNN-BiGRU_SVM Hier	87,00	Ya (3 lapis)	Ya

4. SIMPULAN

Penelitian ini mengusulkan XSentiment-HS, sebuah framework hierarchical explainable berbasis CNN-BiGRU-SVM untuk mendeteksi hate speech dan klasifikasi severity pada Twitter berbahasa Indonesia. Experimen pada 13.169 tweet menunjukkan bahwa: (1) arsitektur hierarkis mampu mengungguli baseline CNN-BiGRU-Softmax dengan Stage 1 mencapai F1-score 87,00% dan Stage 2 mencapai macro F1-score 69,62%;(2) kombinasi SMOTE dan class weighting efektif menangani ketidakseimbangan kelas yang ekstrem; dan (3) tiga lapisan XAI menghasilkan penjelasan yang konsisten (ES = 0,650) dan cukup akurat (EF =0,498). Framework ini memiliki potensi aplikasi langsug dalam sistem moderasi konten otomatis yang membutuhkan akuntabilitias dan transparansi. Penelitian lanjutan dapat mengeksplorasi penggunaan IndoBERT sebagai encoder backbone serta perluasan ke klasifikasi target group secara bersamaan menggunakan pendekatan multi-task learning.

DAFTAR PUSTAKA

[1] M. O. Ibrohim and I. Budi, "Multi-label hate speech and abusive language detection in Indonesian Twitter," in Proc. 3rd Workshop on Abusive Language Online (ALW3), Florence, Italy, 2019, pp. 46–57.

[2] F. Poletto, V. Basile, M. Sanguinetti, C. Bosco, and V. Patti, "Resources and benchmark corpora for hate speech detection: A systematic review," Lang. Resour. Eval., vol. 55, no. 2, pp. 477–523, 2021.

[3] M. Selvi and S. V. N. Santhosh Kumar, "An intelligent bi-directional gate recurrent neural network based hybrid deep learning model for text-based sentiment analysis," Int. J. Comput. Intell. Syst., vol. 19, p. 114, 2026, doi: 10.1007/s44196-026-01244-9.

[4] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," Adv. Neural Inf. Process. Syst., vol. 26, pp. 3111–3119, 2013.

[5] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," J. Artif. Intell. Res., vol. 16, pp. 321–357, 2002.

[6] S. M. Lundberg et al., "From local explanations to global understanding with explainable AI for trees," Nat. Mach. Intell., vol. 2, no. 1, pp. 56–67, 2020.

[7] M. T. Ribeiro, S. Singh, and C. Guestrin, "Beyond accuracy: Behavioral testing of NLP models with CheckList," in Proc. ACL 2020, pp. 4902–4912.

[8] S. Jain and B. C. Wallace, "Attention is not explanation," in Proc. NAACL-HLT 2019, pp. 3543–3556.

[9] M. S. Başarslan and F. Kayaalp, "MBi-GRUMCONV: A novel multi Bi-GRU and multi CNN-based deep learning model for social media sentiment analysis," J. Cloud Comput., vol. 12, no. 1, p. 5, 2023.

[10] C. N. Dang, M. N. Moreno-García, and F. De la Prieta, "Hybrid deep learning models for sentiment analysis," Complexity, vol. 2021, p. 9986920, 2021.

[11] B. Vidgen and L. Derczynski, "Directions in abusive language training data: Garbage in, garbage out," PLOS ONE, vol. 15, no. 12, p. e0243300, 2020.

[12] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," in Proc. COLING 2020, pp. 757–770.

[13] S. Wiegreffe and Y. Pinter, "Attention is not not explanation," in Proc. EMNLP-IJCNLP 2019, pp. 11–20.

[14] G. Rizos, K. Hemker, and B. Schuller, "Augment to prevent: Short-text data augmentation in deep learning for hate speech classification," in Proc. CIKM 2019, pp. 2775–2779.

[15] K. L. Tan, C. P. Lee, K. M. Lim, and K. S. M. Anbananthen, "Sentiment analysis with ensemble hybrid deep learning model," IEEE Access, vol. 10, pp. 103694–103704, 2022.

[16] A. He and M. Abisado, "Text sentiment analysis of short comments based on BERT-CNN-BiLSTM-Att model," IEEE Access, vol. 12, pp. 45229–45237, 2024.

[17] N. A. Angraini and K. M. Lhaksmana, "Sentiment analysis about legislative elections using deep learning with LSTM and CNN models," Build. Inf. Technol. Sci. (BITS), vol. 6, no. 1, pp. 289–299, 2024.

[18] A. Alsayat, "Improving sentiment analysis for social media applications using an ensemble deep learning language model," Arab. J. Sci. Eng., vol. 47, no. 2, pp. 2499–2511, 2022.

[19] S. Ahmad, S. M. Saqib, and A. H. Syed, "CNN and LSTM based hybrid deep learning model for sentiment analysis on Arabic text reviews," Mehran Univ. Res. J. Eng. Technol., vol. 43, no. 2, pp. 183–194, 2024.

[20] A. Aslan, S. Kızıloluk, and E. Sert, "TSA-CNN-AOA: Twitter sentiment analysis using CNN optimized via arithmetic optimization algorithm," Neural Comput. Appl., vol. 35, no. 14, pp. 10311–10328, 2023.