JSAI : Journal Scientific and Applied Informatics
Vol. 08, No. 3, November 2025, hal. 626~631
E-ISSN: 2614-3054; P-ISSN: 2614-3062, accredited by Kemenristekdikti, Sinta 4
DOI: 10.36085

Model Deteksi Berita Palsu Menggunakan BERT dan Bi-LSTM Berbasis *Discriminative Approach*

Dwi Fitri Brianna[a], Paisal[b], M. Apreza Saputra[c], Muhammad Al Hapiz[d]
Fakultas Ilmu Komputer, Universitas Sjakhyakirti, Palembang, Indonesia
[a]*dwiftribrianna@unisti.ac.id*, [b]*paisal@unisti.ac.id*, [c]*21610009@student.unisti.ac.id*, [d]*21610014@student.unisti.ac.id*

Article Info

Article history:

Received, 2025-10-28
Revised, 2025-11-08
Accepted, 2025-11-11

Kata Kunci:

Penambangan teks,
Pembelajaran mesin,
Pembelajaran mendalam,
Word embedding,
Berita palsu

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan model klasifikasi teks dalam mendeteksi hoaks dengan menggunakan pendekatan berbasis pembelajaran mendalam dan representasi teks. Data teks yang telah melalui tahap pra-pemrosesan kemudian diekstraksi fiturnya dengan menggunakan tiga pendekatan, yaitu Word2Vec, Doc2Vec, dan Bidirectional Encoder Representations from Transformers (BERT). Dataset penelitian terdiri atas berita asli (label 0) yaitu 2.325 data dan berita palsu (label 1) sebanyak 2.287 data. Dalam penelitian ini, vektor fitur BERT dengan dimensi 768 dikombinasikan dengan algoritma *Bidirectional Long Short-Term Memory* (Bi-LSTM) untuk menangkap ketergantungan sekuensial dalam teks serta algoritma *Support Vector Machine* (SVM) sebagai klasifikator akhir. Proses pelatihan dilakukan pada perangkat keras Dell Precision 7750 menggunakan parameter dimensi *embedding* sebesar 128, jumlah *hidden unit* sebanyak 64, tingkat *dropout* 0,3, serta *learning rate* sebesar 0,001. Pengujian dan pelatihan dilakukan selama 10 *epoch* dengan ukuran batch 32. Hasil penelitian menunjukkan bahwa model Word2Vec dan Bi-LSTM menghasilkan akurasi 87,4% dengan F1-Score 87,0%, sedangkan Doc2Vec dan Bi-LSTM sedikit lebih rendah dengan akurasi 85,6% dan F1-Score 85,4%. Model terbaik diperoleh dari BERT, Bi-LSTM dan SVM dengan akurasi 93,8%, precision 94,1%, recall 93,5%, dan F1-Score 93,7%.

Keywords:

Text mining,
Image processing,
Deep learning,
Word embedding,
Hoax news

ABSTRACT

This study aimed to develop a text classification model for detecting hoaxes using a deep learning approach and text representation methods. The text data that had undergone preprocessing were then extracted using three approaches: Word2Vec, Doc2Vec, and Bidirectional Encoder Representations from Transformers (BERT). The research dataset consisted of 2,325 genuine news articles (label 0) and 2,287 fake news articles (label 1). In this study, BERT feature vectors with a dimension of 768 were combined with the Bidirectional Long Short-Term Memory (Bi-LSTM) algorithm to capture sequential dependencies in the text, along with the Support Vector Machine (SVM) algorithm as the final classifier. The training process was carried out on Dell Precision 7750 hardware using parameters of embedding dimension 128, 64 hidden units, a dropout rate of 0.3, and a learning rate of 0.001. Training and testing were conducted for 10 epochs with a batch size of 32. The results indicated that the Word2Vec and Bi-LSTM model achieved an accuracy of 87.4% with an F1-Score of 87.0%, while the Doc2Vec and Bi-LSTM model performed slightly lower with an accuracy of 85.6% and an F1-Score of 85.4%. The best performance was obtained by the BERT, Bi-LSTM, and SVM model, which achieved an accuracy of 93.8%, precision of 94.1%, recall of 93.5%, and an F1-Score of 93.7%.

Penulis Korespondensi:

Dwi Fitri Brianna,
Fakultas Ilmu Komputer
Universitas Sjakhyakirti, Palembang, Indonesia
Email: dwiftribrianna@unisti.ac.id

JSAI : Journal Scientific and Applied Informatics
Vol. 08, No. 3, November 2025, hal. 626~631
E-ISSN: 2614-3054; P-ISSN: 2614-3062, accredited by Kemenristekdikti, Sinta 4
DOI: 10.36085

1. PENDAHULUAN

Perkembangan teknologi digital dan pesatnya penggunaan media sosial telah memberikan kemudahan dalam berbagi informasi secara cepat dan luas. Namun, perkembangan ini juga diiringi dengan meningkatnya penyebaran informasi yang tidak akurat atau bahkan sengaja dimanipulasi, dikenal sebagai berita palsu (fake news) [1], [2]. Urgensi penelitian berkaitan fenomena penyebaran berita palsu yang semakin meluas di platform media sosial, yang dapat memengaruhi opini publik dan stabilitas sosial serta politik. Fenomena ini menjadi tantangan serius, khususnya dalam konteks Indonesia, yang memiliki populasi pengguna internet yang besar dan beragam latar belakang literasi digital. Untuk mengatasi permasalahan ini, diperlukan sistem deteksi otomatis yang mampu mengidentifikasi berita palsu [3], [4].

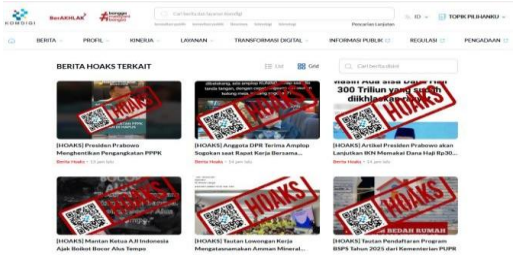
Pendekatan yang banyak digunakan pada penelitian sebelumnya adalah *discriminative modes*, yang berfokus pada pemetaan langsung antara data input dan label keluaran. Qu et al. (2024) menggunakan kombinasi *convolutional neural network* (CNN) untuk fitur ekstraksi pada teks dan citra dalam bahasa Inggris, sementara Luvembe et al. (2024) mengaplikasikan *deep neural networks* (DNN) pada kedua modalitas tersebut untuk dataset dalam bahasa Inggris [5], [6]. Riset oleh Peng et al. (2024) menggabungkan BERT untuk teks dataset berbahasa Inggris dan VGG16 untuk analisis data citra [7]. Penelitian oleh Fang et al. (2024) memanfaatkan Bidirectional Encoder Representations from Transformers (BERT) dan Bidirectional Long Short-Term Memory (Bi-LSTM) untuk ekstraksi fitur pada teks untuk analisis dataset dalam bahasa Inggris serta China [8]. Penelitian oleh Hashmi et al. (2024) menggabungkan CNN dan LSTM untuk ekstraksi fitur pada teks dalam bahasa Inggris [9]. Riset oleh Alghamdi et al. (2024) menggunakan mBERT untuk analsis teks bahasa Inggris, Hindi, Indonesia, Swahili, dan Vietnam, tanpa melibatkan citra [10]. Penelitan yang dilakukan oleh Jiang et al. (2024) menggabungkan Albert untuk analisis teks dan ResNet50 untuk analisis citra dari dataset multimodal berbahasa Inggris [11].

Meskipun penelitian sebelumnya telah berhasil meningkatkan akurasi deteksi teks, masih terdapat gap penelitian terkait keterbatasan riset penggunaan dataset bahasa Indonesia dan masalah keanekaragaman semantik antara modalitas. Penelitian ini akan menggunakan BERT untuk analisis teks berita palsu dalam Bahasa Indonesia memiliki keunggulan dibandingkan model monolingual seperti IndoBERT atau model global seperti BERT-Base, karena BERT mampu menangani teks multibahasa, termasuk Bahasa Indonesia, tanpa pelatihan tambahan [10], [12], [13]. Selain itu, BERT lebih efektif dalam mengatasi morfologi kompleks dan struktur kalimat khas Bahasa Indonesia, terutama ketika dipasangkan dengan Bi-LSTM yang menangkap konteks dua arah secara forward dan backward [14], [15]. Penelitian ini bertujuan untuk mengembangkan model deteksi berita palsu dalam bahasa Indonesia berbasis fitur multimodal menggunakan pendekatan discriminative. Pendekatan pemecahan masalah dalam penelitian ini akan dilakukan dengan mengembangkan model deteksi berita palsu pendekatan *discriminative*. Penelitian ini akan memanfaatkan BERT dan Bi-LSTM untuk analisis teks berita palsu dalam bahasa Indonesia.

Penelitian ini mengusulkan kebaruan (*state-of-the-art*) terutama dalam henggunaan BERT untuk analisis teks berita palsu dalam Bahasa Indonesia. Berbeda dengan model monolingual seperti IndoBERT atau model global seperti BERT-Base, BERT dapat menangani teks dalam berbagai bahasa tanpa pelatihan tambahan, memberikan fleksibilitas yang lebih besar untuk mendeteksi berita palsu dalam konteks multibahasa. Selain itu, dengan memadukan BERT dengan Bi-LSTM, penelitian ini mampu menangani morfologi kompleks dan struktur kalimat khas Bahasa Indonesia dengan lebih baik, yang sebelumnya tidak banyak diperhatikan dalam penelitian sebelumnya.

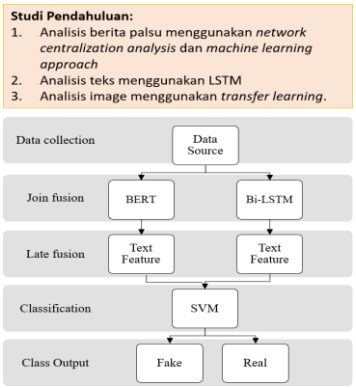
2. METODE PENELITIAN

Penelitian ini akan menggunakan perangkat keras Dell Precision 7750 dengan GPU NVIDIA Quadro RTX 5000 serta layanan cloud Amazon EC2 P3.2xlarge yang dilengkapi dengan GPU NVIDIA Tesla V100, 8 vCPUs, 61 GB RAM, dan 1 GPU Tesla V100 serta Elastic Block Storage (EBS) untuk penyimpanan data yang diperlukan selama penelitian. Dataset yang digunakan dalam penelitian ini diperoleh dari Komdigi dan terdiri dari 4.650 sampel data teks. Data ini dibagi menjadi dua kategori, yaitu berita asli (dengan label 0) dan berita palsu (dengan label 1). Dataset tersebut memiliki berbagai fitur utama, seperti teks yang terdiri dari judul dan isi berita yang menyertai artikel berita seperti pada **Gambar 1**.



Gambar 1 Dataset Penelitian

Studi pendahuluan terkait analisis berita palsu telah dilakukan oleh tim peneliti dengan menggunakan berbagai pendekatan antara lain analisis teks berita palsu menggunakan *network centralization analysis* [16], LSTM [17], *machine learning* [18] dan analisis citra menggunakan transfer learning [19] dalam kurun waktu 2022 hingga 2024. Adapun metode penelitian yang diusulkan dapat dilihat pada Gambar 2.



Gambar 2 Tahap Penelitian

Tahap *data collection* dimulai dengan melakukan pembersihan data yang tidak relevan. Untuk data teks dilakukan tokenisasi, penghapusan tanda baca, stiker, dan hyperlink, stop word dan pemeriksaan ejaan untuk memperbaiki kata-kata bahasa yang salah eja dan menggantinya dengan kata yang benar. Penelitian ini menggunakan dua metode untuk ekstraksi fitur teks Word2vec (*word embeddings, word vectors*) dan Doc2vec (*sentence embeddings, paragraph vectors, Paragraph2vec*).

Model Word2vec digunakan untuk memahami konteks dan hubungan antar kata, sehingga fitur numerik yang dihasilkan dapat menangkap arti dan asosiasi antar kata dalam kalimat atau dokumen. Misalnya, kata-kata seperti "berita", "informasi", dan "berita palsu" akan memiliki representasi vektor yang saling berdekatan karena memiliki hubungan makna yang relevan dalam konteks tertentu. Model Doc2Vec bekerja dengan cara memetakan setiap kalimat atau paragraf ke dalam ruang vektor yang mempertimbangkan seluruh konteks kalimat atau paragraf tersebut, bukan hanya kata-kata terpisah.

Pada penelitian ini, vektor fitur BERT yang digunakan memiliki dimensi 768, yang kemudian dikombinasikan dengan algoritma Bidirectional Long Short-Term Memory (Bi-LSTM). Metode BERT digunakan untuk menganalisis data teks, sedangkan Bi-LSTM membantu menangkap ketergantungan sekuensial yang lebih baik dalam teks, baik secara maju maupun mundur, sehingga memberikan pemahaman yang lebih dalam terhadap konteks kalimat.

Untuk data teks, parameter dimensi *embedding* diatur pada nilai 128, sehinggga model dapat menerima informasi semantik yang lebih baik dalam data teks. Parameter *hidden unit* diterapkan sebanyak 64 untuk memberikan keseimbangan antara kapasitas model dan kinerja komputasi. Tingkat *dropout* diatur pada 0,3 untuk mengurangi risiko *overfitting* dan memastikan model tidak terlalu mengandalkan data latih tertentu. Parameter *learning rate* diatur pada 0,001 untuk memastikan proses pelatihan tetap stabil. Proses pelatihan dilakukan selama 10 *epoch*, dengan ukuran *batch* sebanyak 32 untuk mengoptimalkan pembelajaran model dalam mendeteksi pola dalam data. Ringkasan penelitian terkait yang membahas berbagai pendekatan dalam deteksi berita palsu menggunakan teknik ekstraksi fitur pada teks dan citra, serta dataset yang digunakan dapat dilihat pada **Tabel 1**.

Tabel 1 Penelitian Terkait

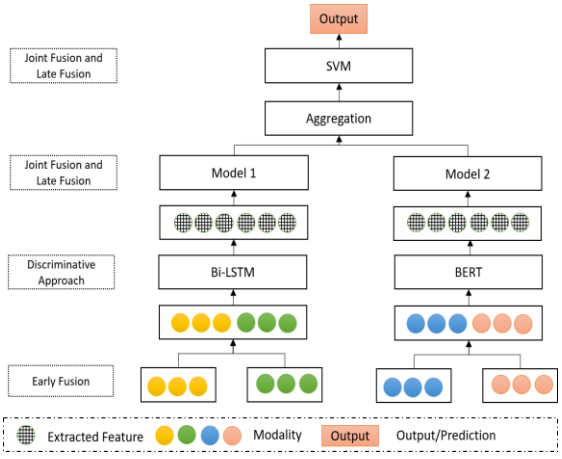
No	Sumber	Metode	Dataset
1	[5]	CNN	*English*
2	[6]	DNN	*English*
3	[7]	BERT	*English*
4	[8]	BERT,Bi-LSTM	*English, Chinese*
5	[9]	CNN-LSTM	*English*
6	[10]	mBERT	*English, Hindi, Indonesian, Swahili* dan *Vietnamese*
7	[11]	ALBERT	*English*
8	[20]	BERT	*English*`
9	Usulan	BERT, Bi-LSTM, SVM	Indonesia

3. HASIL DAN ANALISIS

Tahap awal penelitian ini adalah pembersihan dan normalisasi data. Dari total 4.650 sampel data teks yang diperoleh dari Komdigi, setelah dilakukan data cleaning (tokenisasi, penghapusan tanda baca, hyperlink,

stop word, dan koreksi ejaan), jumlah data dalam penelitian yaitu 4.612 sampel. Data tersebut terdiri atas berita asli (label 0) yaitu 2.325 data dan berita palsu (label 1) sebanyak 2.287 data. Penelitian ini menggunakan tiga teknik ekstraksi fitur, yaitu Word2Vec, Doc2Vec, dan BERT. Metode Word2Vec menghasilkan representasi kata dengan dimensi vektor 128. Kata-kata yang memiliki makna semantik mirip, seperti "hoaks" dan "berita palsu", ditampilkan dengan vektor yang saling berdekatan dalam ruang vektor. Untuk metode Doc2Vec merepresentasikan dokumen teks (judul + isi berita) ke dalam vektor berdimensi 128, sehingga memperhitungkan konteks kalimat secara keseluruhan. Metode BERT menghasilkan representasi vektor berdimensi 768 yang kemudian diproses dengan model Bi-LSTM.

Penelitian ini menggabungkan dua jenis fitur, BERT dan Bi-LSTM untuk meningkatkan akurasi. Ada tiga metode fusi yang digunakan: *early fusion, joint fusion,* dan *late fusion.* Ketiga metode ini digunakan untuk menggabungkan informasi dari teks, sehingga meningkatkan kinerja deteksi berita palsu seperti yang terlihat pada **Gambar 3**.



Gambar 3 Arsitektur Model Usulan

Berdasarkan arsitektur model usulan, pendekatan fusi digunakan untuk deteksi berita palsu dengan menggabungkan informasi dari teks yang dimulai dengan early fusion, fitur yang diekstrak dari teks dan gambar digabungkan pada tahap awal. Fitur teks dari masing-masin model diberi warna kuning, dan diberi warna biru, diproses secara terpisah dan kemudian digabungkan untuk membentuk satu set fitur gabungan. Pada bagian discriminative approach, model menggunakan BERT untuk memproses data teks, sementara Bi-LSTM (Bidirectional Long Short-Term Memory) digunakan untuk menangkap ketergantungan sekuensial dalam teks dan memahami konteks secara lebih mendalam. Fitur teks yang telah diproses kemudian diteruskan melalui mBERT dan Bi-LSTM untuk ekstraksi informasi kontekstual yang relevan. Selanjutnya, pada tahap joint fusion dan late fusion, fitur teks yang telah diproses secara terpisah digabungkan.

Joint fusion menggabungkan fitur pada level yang lebih tinggi, sementara late fusion menggabungkan hasil prediksi dari kedua fitur untuk menghasilkan keputusan akhir. Penerapan *aggregation* dilakukan dengan menggabungkan output dari kedua model (model 1 dan model 2) untuk prediksi akhir. Model 1 memproses teks menggunakan BERT, sementara model 2 memproses teks menggunakan Bi-LSTM. Kedua model ini bekerja secara paralel namun menangani data dari fitur yang berbeda. Output dari kedua model ini kemudian diproses melalui tahap fusi dan agregasi. Untuk klasifikasi akhir, digunakan Support Vector Machine (SVM) yang memanfaatkan fitur yang telah diekstraksi dan digabungkan melalui teknik fusi. Prediksi akhir menentukan apakah berita yang dianalisis termasuk berita asli atau palsu.

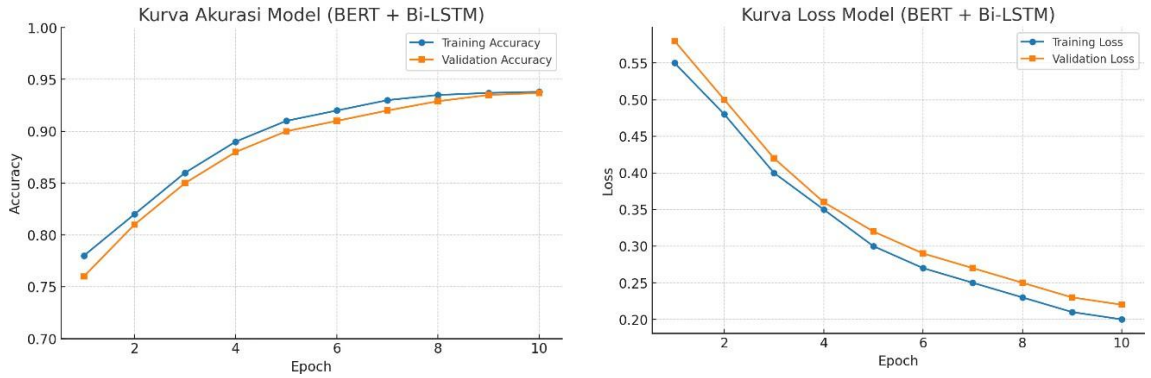
Proses pelatihan dilakukan menggunakan perangkat keras Dell Precision 7750 yang dilengkapi dengan GPU Quadro RTX 5000 sehingga mampu mendukung komputasi intensif pada model yang digunakan. Parameter pelatihan ditetapkan dengan dimensi embedding sebesar 128, serta jumlah hidden units sebanyak 64 untuk mengoptimalkan representasi data. Untuk mencegah overfitting, diterapkan dropout sebesar 0,3. Selain itu, learning rate ditetapkan pada 0,001 agar proses pembelajaran berlangsung stabil, dengan jumlah epoch sebanyak 10 dan batch size sebesar 32 guna menyeimbangkan efisiensi komputasi dan kualitas hasil pelatihan. Setelah dilakukan pengujian, diperoleh hasil performa model seperti ditunjukkan pada **Tabel 2**.

Tabel 2 Hasil Penelitian

Model	Akurasi	Precision	Recall	F1-Score
Word2Vec + Bi-LSTM	87,4%	86,9%	87,2%	87,0%
Doc2Vec + Bi-LSTM	85,6%	85,0%	85,8%	85,4%
BERT + Bi-LSTM + SVM	**93,8%**	**94,1%**	**93,5%**	**93,7%**

Berdasarkan hasil eksperimen, model Word2Vec + Bi-LSTM menghasilkan akurasi sebesar 87,4% dengan nilai precision 86,9%, recall 87,2%, dan F1-Score 87,0%. Model Doc2Vec + Bi-LSTM sedikit lebih

rendah dengan akurasi 85,6% dan F1-Score 85,4%, menunjukkan performa yang stabil namun tidak sebaik Word2Vec. Untuk model BERT + Bi-LSTM + SVM memberikan hasil terbaik dengan akurasi 93,8%, precision 94,1%, recall 93,5%, serta F1-Score 93,7%. Hasil akurasi dan loss untuk setiap epoch eksperimen pada model BERT + Bi-LSTM + SVM dapat dilihat pada **Gambar 4**.



Gambar 4 Hasil Grafik Akurasi dan Loss Untuk Model Usulan

Model BERT + Bi-LSTM yang dilatih selama 10 epoch menunjukkan performa yang stabil dan konsisten. Kurva akurasi meningkat sejak awal pelatihan dan mencapai kestabilan pada sekitar 93,8% mulai epoch ke-7 hingga epoch ke-10, menandakan bahwa model mampu belajar secara efektif. Di sisi lain, kurva loss mengalami penurunan signifikan dan kemudian stabil dengan gap yang kecil antara training dan validation loss. Kondisi ini menunjukkan bahwa model tidak mengalami overfitting, sehingga generalisasi terhadap data uji tetap terjaga dengan baik.

4. KESIMPULAN

Hasil pelatihan model BERT + Bi-LSTM selama 10 epoch ditunjukkan pada Gambar 3 (akurasi) dan Gambar 4 (loss). Dari kurva akurasi terlihat bahwa performa model mengalami peningkatan signifikan sejak epoch ke-1 hingga epoch ke-3, kemudian stabil pada nilai di atas 90% mulai epoch ke-5. Pada akhir pelatihan, akurasi validasi mencapai 93,7%, sedangkan akurasi data latih mencapai 93,8%. Perbedaan yang relatif kecil antara keduanya menunjukkan bahwa model memiliki kemampuan generalisasi yang baik pada dataset penelitian. Sementara itu, kurva loss memperlihatkan tren penurunan yang konsisten pada data latih maupun data validasi. Nilai loss awal yang berada pada kisaran 0,55–0,58 menurun secara stabil hingga mencapai sekitar 0,20–0,22 pada epoch ke-10. Berdasarkan hasil penelitian, kombinasi BERT embeddings dengan arsitektur Bi-LSTM dapat menghasilkan representasi teks yang lebih kontekstual, sehingga menghasilkan performa klasifikasi yang lebih baik dalam mendeteksi berita palsu dibandingkan dengan metode berbasis Word2Vec maupun Doc2Vec.

UCAPAN TERIMA KASIH

Tim peneliiti menyampaikan terima kasih kepada Kementerian Pendidikan Tinggi, Sains, dan Teknologi Republik Indonesia yang telah memberikan hibah penelitian tahun 2025 dengan nomor kontrak turunan 175/LL2/DT.05.00/PL/2025 dan 03/VII/PG.2/VI/2025 dan Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Sjakhyakirti atas dukungan dan fasilitasi dalam pelaksanaan penelitian ini.

REFERENSI

[1] F. Olan, U. Jayawickrama, E. O. Arakpogun, J. Suklan, and S. Liu, “Fake news on social media: the impact on society,” *Inf. Syst. Front.*, vol. 26, no. 2, pp. 443–458, 2024.
[2] A. M. French, V. C. Storey, and L. Wallace, “The impact of cognitive biases on the believability of fake news,” *Eur. J. Inf. Syst.*, vol. 34, no. 1, pp. 72–93, 2025.
[3] S. Arora, V. Agrawal, D. Kumar, S. Arora, and S. K. Banshal, “Sentimental impact of fake news on social media using an integrated ensemble framework,” *Soc. Netw. Anal. Min.*, vol. 14, no. 1, p. 185, 2024.
[4] B. Hu, Z. Mao, and Y. Zhang, “An overview of fake news detection: From a new perspective,” *Fundam. Res.*, vol. 5, no. 1, pp. 332–346, 2025.
[5] Z. Qu, Y. Meng, G. Muhammad, and P. Tiwari, “QMFND: A quantum multimodal fusion-based fake news detection model for social media,” *Inf. Fusion*, vol. 104, p. 102172, 2024.
[6] A. M. Luvembe, W. Li, S. Li, F. Liu, and X. Wu, “CAF-ODNN: Complementary attention fusion with

optimized deep neural network for multimodal fake news detection," *Inf. Process. Manag.*, vol. 61, no. 3, p. 103653, 2024.

[7] L. Peng, S. Jian, Z. Kan, L. Qiao, and D. Li, "Not all fake news is semantically similar: Contextual semantic representation learning for multimodal fake news detection," *Inf. Process. Manag.*, vol. 61, no. 1, p. 103564, 2024.

[8] X. Fang *et al.*, "NSEP: Early fake news detection via news semantic environment perception," *Inf. Process. Manag.*, vol. 61, no. 2, p. 103594, 2024.

[9] E. Hashmi, S. Y. Yayilgan, M. M. Yamin, S. Ali, and M. Abomhara, "Advancing fake news detection: Hybrid deep learning with fasttext and explainable ai," *IEEE Access*, 2024.

[10] J. Alghamdi, Y. Lin, and S. Luo, "Fake news detection in low-resource languages: A novel hybrid summarization approach," *Knowledge-Based Syst.*, vol. 296, p. 111884, 2024.

[11] M. Jiang, C. Jing, L. Chen, Y. Wang, and S. Liu, "An application study on multimodal fake news detection based on Albert-ResNet50 Model," *Multimed. Tools Appl.*, vol. 83, no. 3, pp. 8689–8706, 2024.

[12] E. Shushkevich, M. Alexandrov, and J. Cardiff, "Improving multiclass classification of fake news using Bert-based models and CHATGPT-augmented data," *Inventions*, vol. 8, no. 5, p. 112, 2023.

[13] R. Sivanaiah, S. Suresh, S. Pandian, and A. D. Suseelan, "Bridging the Language Gap: Transformer-Based BERT for Fake News Detection in Low-Resource Settings," in *International Conference on Speech and Language Technologies for Low-resource Languages*, Springer, 2023, pp. 398–411.

[14] X. Jiang, C. Song, Y. Xu, Y. Li, and Y. Peng, "Research on sentiment classification for netizens based on the BERT-BiLSTM-TextCNN model," *PeerJ Comput. Sci.*, vol. 8, p. e1005, 2022.

[15] F. Meng, S. Yang, J. Wang, L. Xia, and H. Liu, "Creating knowledge graph of electric power equipment faults based on BERT–BiLSTM–CRF model," *J. Electr. Eng. Technol.*, vol. 17, no. 4, pp. 2507–2516, 2022.

[16] D. F. Brianna, E. S. Negara, and Y. N. Kunang, "Network centralization analysis approach in the spread of hoax news on social media," in *International Conference on Electrical Engineering and Computer Science (ICECOS)*, IEEE, 2022, pp. 303–308.

[17] M. Purba, P. Paisal, C. P. Darmo, H. Noprisson, and V. Ayumi, "Model of Indonesian Cyberbullying Text Detection Using Modified Long Short-Term Memory," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 10, no. 1, pp. 9–14, 2023.

[18] P. Paisal, "Analisis Sentimen Masyarakat Berdasarkan Opini dari Sosial Media Menggunakan Metode Naive Bayes Classifier (Study Kasus: Universitas Sjakhyakirti)," *J. Ilm. Inform. Glob.*, vol. 11, no. 1, 2022.

[19] D. F. Brianna, L. I. Kesuma, D. Geovani, and P. Sari, "Combination of Image Enhancement and Double U-Net Architecture for Liver Segmentation in CT-Scan Images," *J. Electron. Electromed. Eng. Med. Informatics*, vol. 7, no. 1, pp. 208–219, 2024.

[20] K. Yu, S. Jiao, and Z. Ma, "Fake news detection based on BERT multi-domain and multi-modal fusion network," *Comput. Vis. Image Underst.*, p. 104301, 2025.