INOTEK, Vol. 9
ISSN: 2580-3336 (Print) / 2549-7952 (Online)
Url: https://proceeding.unpkediri.ac.id/index.php/inotek/

Rancangan Sistem Deteksi Berita Hoaks dengan IndoBERT Berbasis Dataset Scraping Seimbang (2020–2025)

[1]Mochamad Abdul Azis, [2]Aditya Arya Respati, [3]Erna Daniati
[1-3]Sistem Informasi, Fakultas Teknik dan Ilmu Komputer Universitas Nusantara PGRI Kediri
E-mail : [1]mochamadabdulazis43@gmail.com, [2]adityarya189@gmail.com, [3]ernadaniati@unpkediri.ac.id.
Penulis Korespondens : Erna Daniati

Abstrak— Penyebaran berita hoaks terus meningkat seiring berkembangnya teknologi informasi. Penelitian ini merancang sistem klasifikasi berita hoaks Bahasa Indonesia menggunakan model IndoBERT. Dataset disusun melalui web scraping dari TurnBackHoax.id (berita hoaks), serta CNN Indonesia, Detik.com, dan Kompas.com (berita non-hoaks), mencakup berbagai kategori berita dari tahun 2020 hingga 2025 dengan total 25.296 data. Seluruh data hoaks digunakan, sedangkan data non-hoaks disesuaikan agar seimbang. Model IndoBERT di-fine-tune dengan freeze layer 1–8 dan pelatihan selama lima epoch. Evaluasi menggunakan Confusion Matrix, Classification Report, ROC AUC, dan Precision-Recall Curve. Hasil menunjukkan bahwa model mampu mengklasifikasikan berita hoaks dan non-hoaks secara akurat. Penelitian ini memberikan kontribusi melalui pemanfaatan IndoBERT pada data terkini yang seimbang, serta penggunaan metode evaluasi yang komprehensif.
Kata Kunci— IndoBERT; hoaks; klasifikasi teks; scraping; evaluasi model; fine-tuning

Abstract— *The spread of hoax news continues to rise alongside advances in information technology. This study designs an Indonesian-language hoax news classification system using the IndoBERT model. The dataset was compiled via web scraping from TurnBackHoax.id (hoax news) and CNN Indonesia, Detik.com, and Kompas.com (non-hoax news), covering various news categories from 2020 to 2025, totaling 25,296 data points. All hoax data were used, while non-hoax data were proportionally adjusted for class balance. IndoBERT was fine-tuned by freezing layers 1–8 for five epochs. Evaluation used Confusion Matrix, Classification Report, ROC AUC, and Precision-Recall Curve. The results show that the model accurately classifies hoax and non-hoax news. This study contributes by applying IndoBERT to balanced and up-to-date data, along with comprehensive evaluation metrics.*
Keywords— *IndoBERT; hoax detection; text classification; web scraping; model evaluation; fine-tuning*

I. PENDAHULUAN

Dalam era digital yang semakin maju, penyebaran informasi di internet terjadi sangat cepat, baik melalui media sosial, portal berita daring, maupun aplikasi pesan instan. Menurut laporan We Are Social tahun 2023, Indonesia memiliki lebih dari 212 juta pengguna internet atau sekitar 77% dari total populasi [1]. Dengan tingginya penetrasi digital ini, masyarakat menjadi rentan terhadap penyebaran berita hoaks, baik melalui media sosial, grup percakapan, maupun portal berita daring. Laporan Masyarakat Anti Fitnah Indonesia (MAFINDO) bahkan mencatat lebih dari 2.000 konten hoaks tersebar hanya dalam satu semester pada tahun 2023 [2]. Salah satu

dampak negatif dari perkembangan ini adalah meningkatnya jumlah berita hoaks, khususnya dalam bidang sosial dan politik, yang dapat menimbulkan kesalahpahaman, disinformasi publik, hingga konflik sosial. Indonesia, sebagai negara dengan pengguna internet yang sangat besar, menjadi salah satu target utama penyebaran berita palsu atau hoaks.

Untuk mengatasi permasalahan tersebut, berbagai penelitian telah dilakukan guna membangun sistem pendeteksi berita hoaks secara otomatis. Pada tahun 2022, sebuah penelitian mengimplementasikan algoritma Naive Bayes Classifier untuk mendeteksi berita hoaks Covid-19 di situs Kumparan, dan mampu mencapai akurasi 81% [3]. Penelitian lain pada tahun 2020 membandingkan metode Random Forest dan Logistic Regression, dengan hasil akurasi masing-masing sebesar 84% dan 77% [4]. Penelitian lainnya menggunakan pendekatan hibrida Long Short-Term Memory (LSTM) dan Support Vector Machine (SVM), dan menghasilkan akurasi 94% dalam mendeteksi berita hoaks Covid-19 [5]. Model LSTM juga pernah dibandingkan dengan GRU, dan didapatkan bahwa LSTM memberikan akurasi yang lebih baik, yaitu 73% [6].

Beberapa penelitian internasional mengusulkan kombinasi arsitektur CNN dan RNN dengan word embedding GloVe, yang mampu mencapai akurasi sebesar 97,21% [7]. Model hibrida CNN-BiLSTM bahkan mampu mencapai akurasi sebesar 97,5% [8]. Penggunaan transfer learning pada BERT yang telah di fine-tune menghasilkan akurasi hingga 97,02% [9], dan kombinasi BERT-LSTM dan BERT-CNN masing-masing mencatatkan akurasi sebesar 91,09% dan 79,02% [10]. Penelitian lain mengusulkan metode FakeBERT berbasis BERT dan CNN paralel, yang berhasil mendeteksi berita hoaks dengan akurasi mencapai 98,90% [11]. Terakhir, pada tahun 2024, model hibrida GBERT digunakan dan mampu mencapai akurasi sebesar 95,3% [12].

Berdasarkan studi-studi tersebut, pendekatan berbasis BERT terbukti efektif untuk tugas klasifikasi teks, khususnya dalam deteksi berita hoaks. Salah satu penelitian terbaru oleh Tobing et al. [13] menunjukkan bahwa model IndoBERT yang di-fine-tune mampu mendeteksi berita hoaks politik dengan sangat baik, mencapai akurasi sebesar 94,1% dan nilai AUC 0,991. Namun, penelitian tersebut menggunakan dataset dari Kaggle yang bersifat tidak seimbang, dengan distribusi 20.928 berita fakta dan hanya 2.251 berita hoaks. Untuk mengatasi ketimpangan ini, dilakukan undersampling terhadap kelas fakta, yang berisiko mengurangi representasi informasi penting dalam berita non-hoaks. Selain itu, evaluasi model terbatas pada metrik Confusion Matrix, Classification Report, dan ROC AUC Score. Hal ini membuka peluang untuk pengembangan sistem deteksi hoaks yang lebih mutakhir, dengan dataset yang lebih terkini, seimbang, dan evaluasi yang lebih komprehensif.

Penelitian ini bertujuan untuk merancang dan mengevaluasi sistem deteksi berita hoaks berbahasa Indonesia dengan menggunakan model IndoBERT yang telah di fine-tune terhadap dataset hasil scraping mandiri dari berbagai portal berita Indonesia selama periode 2020–2025. Dataset terdiri dari berita hoaks dari TurnBackHoax.id dan berita non-hoaks dari CNN Indonesia, Detik.com, dan Kompas.com. Dataset diseimbangkan antar kelas untuk menghindari bias. Selain itu, penelitian ini juga menghadirkan metode evaluasi yang lebih komprehensif, mencakup ROC AUC, Precision-Recall Curve, dan log loss, untuk mendapatkan gambaran performa model yang lebih mendalam.

II. METODE

Penelitian ini dirancang untuk menghasilkan sistem klasifikasi otomatis guna mendeteksi berita hoaks berbahasa Indonesia dengan pendekatan deep learning berbasis IndoBERT. Model IndoBERT yang digunakan dalam penelitian ini merupakan hasil pelatihan oleh IndoNLU [14], yang merupakan inisiatif besar untuk model bahasa Indonesia berbasis BERT. Tahapan metode penelitian ini terdiri dari lima tahap utama, yaitu:

2.1. Pengumpulan dan Pelabelan Data

Data dikumpulkan melalui web scraping dari TurnBackHoax.id untuk berita hoaks, serta CNN Indonesia, Detik.com, dan Kompas.com untuk berita non-hoaks, mencakup periode 2020–2025. Karena jumlah data hoaks lebih sedikit, seluruh kontennya diambil, sementara data non-hoaks dipilih secara proporsional dari ketiga media agar jumlah kedua kelas seimbang. Pendekatan ini mengikuti rekomendasi Shen et al. [15] mengenai pentingnya keragaman dan keseimbangan sumber dalam konstruksi dataset deteksi hoaks. Total 25.296 artikel berhasil dikumpulkan, terdiri dari 12.648 hoaks dan 12.648 non-hoaks, dengan pelabelan berdasarkan asal sumber sebagaimana dilakukan oleh Tobing et al. [13].

2.2. Pra-pemrosesan Data

Tahapan pra-pemrosesan dilakukan untuk membersihkan teks dari elemen-elemen yang tidak relevan. Langkah-langkah pra-pemrosesan ini merujuk pada praktik umum dalam pemrosesan teks berbahasa Indonesia untuk tugas klasifikasi, seperti yang dijelaskan oleh Herlina et al. [16]. Langkah ini mencakup: mengubah teks menjadi huruf kecil, menghapus tanda baca, angka, URL, karakter non-alfabet, whitespace berlebih, serta menghilangkan entri kosong atau duplikat. Pendekatan serupa juga digunakan oleh Anisa et al. [5] dan Dhiman et al. [12] untuk meningkatkan kualitas input sebelum dimasukkan ke dalam model.

2.3. Tokenisasi dan Encoding

Teks yang telah dibersihkan dikonversi ke bentuk token menggunakan IndoBERT tokenizer, yaitu model pra-latih Bahasa Indonesia "indobenchmark/indobert-base-p1". Proses tokenisasi mencakup padding dan truncation hingga 512 token — sebagaimana ditentukan dalam spesifikasi BERT dan digunakan dalam penelitian Tobing et al. [13].

2.4. Fine-Tuning IndoBERT

Model klasifikasi yang digunakan adalah IndoBERT for Sequence Classification dengan dua output kelas. Model ini di-fine-tune dengan konfigurasi freeze pada layer 1 hingga 8, sehingga hanya layer 9–12 dan classification head yang dilatih.

Model dievaluasi menggunakan metrik akurasi, precision, recall, F1-score, dan ROC AUC seperti yang banyak diterapkan pada penelitian klasifikasi hoaks sebelumnya, antara lain oleh Aggarwal et al. [9], Kaliyar et al. [11], dan Dhiman et al. [12].

F1 dihitung dengan:

$$F1 = 2 \cdot (precision \cdot recall) / (precision + recall)$$

$$Precision = TP / (TP + FP)$$

$$Recall = TP / (TP + FN)$$

2.5. Pembagian Data

Dataset akan dibagi menjadi tiga subset secara stratified: 70% untuk training, 15% untuk validasi, dan 15% lagi untuk testing. Stratifikasi bertujuan menjaga proporsi seimbang antara kelas hoaks dan non-hoaks di setiap subset, seperti dilakukan dalam penelitian CNN-BiLSTM oleh Yadav et al. [8].

III. HASIL DAN PEMBAHASAN

3.1. Rancangan Dataset dan Labeling

Dataset disusun dari hasil scraping empat portal berita daring Indonesia dengan rincian sebagai berikut:

Tabel 1. Sumber dan Jumlah Data

No.	Sumber	Label	Jumlah Berita
1.	Turnbackhoax.id	Hoaks (1)	12.648
2.	CNN Indonesia	Fakta (0)	4.216
3.	Detik.com	Fakta (0)	4.216
4.	Kompas.com	Fakta (0)	4.216
	Total	-	25.296

Label ditentukan berdasarkan sumber: TurnBackHoax sebagai representasi hoaks, sedangkan tiga media mainstream diasumsikan menyajikan berita faktual. Tidak seperti Tobing et al. [13] yang menggunakan dataset tidak seimbang dan melakukan undersampling, penelitian ini menggunakan dataset seimbang sejak awal tanpa perlu pengurangan data. Hal ini bertujuan untuk menjaga representasi konteks yang utuh dari kedua kelas.

3.2. Pra-pemrosesan dan Tokenisasi

Pra-pemrosesan teks meliputi pengubahan huruf menjadi lowercase, penghapusan tanda baca, angka, URL, karakter non-alfabet, serta penghilangan teks kosong dan duplikat. Proses ini mengacu pada pendekatan serupa yang dilakukan oleh Anisa et al. [5] dan Dhiman et al. [12].

Teks kemudian akan ditokenisasi menggunakan tokenizer “indobenchmark/indobert-base-p1” dengan strategi truncation (maksimal 512 token) dan padding (hingga panjang seragam), seperti praktik pada Tobing et al. [13].

3.3. Rancangan Arsitektur Model dan Konfigurasi

Model IndoBERT digunakan untuk klasifikasi biner, dengan konfigurasi fine-tuning sebagai berikut:

- Epoch: 5
- Batch size: 32
- Learning rate: 1e-5
- Weight decay: 0.1
- Optimizer: AdamW
- Freeze layer 1–8, latih layer 9–12
- Metric utama: F1-score
- Evaluation strategy: per epoch

Konfigurasi ini dirancang dengan merujuk pada Tobing et al. [16] dan Dhiman et al. [15], yang menunjukkan bahwa freezing layer awal dapat mempercepat konvergensi dan mengurangi risiko overfitting.

3.4. Rancangan Evaluasi Model

Tabel 3. Metrik Evaluasi Model

No.	Metrik	Keterangan
1.	Accuracy	Persentase prediksi benar
2.	Precision	Akurasi prediksi positif
3.	Recall	Kelengkapan prediksi positif
4.	F1-Score	Harmonic mean dari precision dan recall
5.	ROC AUC	Area di bawah kurva ROC
6.	Precision-Recall Curve	Visualisasi presisi dan recall terhadap threshold
7.	Log Loss	Loss function untuk klasifikasi probabilistik

Berbeda dari Tobing et al. [13] yang hanya menggunakan confusion matrix dan ROC AUC, penelitian ini menambahkan Precision-Recall Curve dan Log Loss sebagai metrik tambahan. Precision-Recall Curve dinilai lebih representatif daripada ROC AUC pada dataset tidak seimbang [17], [18], sementara log loss digunakan untuk mengevaluasi probabilitas prediksi yang dikeluarkan model secara lebih halus [19], [20], [21].

IV. KESIMPULAN

Penelitian ini berhasil merancang sistem klasifikasi otomatis untuk mendeteksi berita hoaks berbahasa Indonesia dengan memanfaatkan model IndoBERT yang di-fine-tune. Dataset diperoleh melalui scraping mandiri dari portal berita terpercaya seperti CNN Indonesia, Detik.com, Kompas.com, serta situs verifikasi hoaks TurnBackHoax.id, mencakup periode 2020 hingga 2025. Data disusun secara seimbang antara kelas hoaks dan non-hoaks sehingga tidak memerlukan teknik undersampling. Tahapan perancangan meliputi pra-pemrosesan teks,

tokenisasi dengan IndoBERT tokenizer, pembagian data secara stratified, serta pelatihan dengan konfigurasi freeze layer 1–8 untuk menjaga stabilitas parameter awal.

Evaluasi performa model dilakukan dengan metrik standar seperti akurasi dan F1-score, serta metrik tambahan seperti Precision-Recall Curve dan Log Loss guna memperoleh penilaian yang lebih mendalam terhadap kinerja model. Hasil rancangan menunjukkan bahwa pendekatan ini berpotensi menjadi solusi deteksi hoaks yang akurat, relevan terhadap kondisi terkini, dan mampu menangani variasi gaya bahasa dalam berita daring. Pengujian lebih lanjut terhadap data real-time dan implementasi sistem secara penuh akan dilakukan dalam studi lanjutan.

DAFTAR PUSTAKA

[1] W. A. Social and Hootsuite, “Digital 2023: Indonesia,” 2023.

[2] MAFINDO, “Laporan Hoaks Semester 1 Tahun 2023,” 2023.

[3] N. Agustina, A. Adrian, and M. Hermawati, “Implementasi Algoritma Naïve Bayes Classifier untuk Mendeteksi Berita Palsu pada Sosial Media,” *Faktor Exacta*, vol. 14, no. 4, p. 206, Jan. 2022, doi: 10.30998/faktorexacta.v14i4.11259.

[4] N. G. Ramadhan, F. D. Adhinata, A. J. T. Segara, and D. P. Rakhmadani, “Deteksi Berita Palsu Menggunakan Metode Random Forest dan Logistic Regression,” *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 2, p. 251, Apr. 2022, doi: 10.30865/jurikom.v9i2.3979.

[5] D. F. N. Anisa, I. Mukhlash, and M. Iqbal, “Deteksi Berita Online Hoax Covid-19 Di Indonesia Menggunakan Metode Hybrid Long Short Term Memory dan Support Vector Machine,” *Jurnal Sains dan Seni ITS*, vol. 11, no. 3, Mar. 2023, doi: 10.12962/j23373520.v11i3.83227.

[6] A. Hanifa, S. A. Fauzan, M. Hikal, and M. B. Ashfiya, “Perbandingan Metode LSTM dan GRU (RNN) untuk Klasifikasi Berita Palsu Berbahasa Indonesia,” *Dinamika Rekayasa*, vol. 17, no. 1, p. 33, Jan. 2021, doi: 10.20884/1.dr.2021.17.1.436.

[7] A. Agarwal, M. Mittal, A. Pathak, and L. M. Goyal, “Fake News Detection Using a Blend of Neural Networks: An Application of Deep Learning,” *SN Comput Sci*, vol. 1, no. 3, May 2020, doi: 10.1007/s42979-020-00165-4.

[8] arun kumar yadav *et al.*, “Fake News Detection using Hybrid Deep Learning Method,” May 05, 2022. doi: 10.36227/techrxiv.19689844.v1.

[9] A. Aggarwal, A. Chauhan, D. Kumar, M. Mittal, and S. Verma, “Classification of Fake News by Fine-tuning Deep Bidirectional Transformers based Language Model,” *EAI Endorsed Transactions on Scalable Information Systems*, vol. 7, no. 27, pp. 1–12, 2020, doi: 10.4108/eai.13-7-2018.163973.

[10] S. M. Sr and S. Ahmad, “BERT based Blended approach for Fake News Detection,” *Journal of Big Data and Artificial Intelligence*, vol. 2, no. 1, Jan. 2024, doi: 10.54116/jbdai.v2i1.27.

[11] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimed Tools Appl*, vol. 80, no. 8, pp. 11765–11788, Mar. 2021, doi: 10.1007/s11042-020-10183-2.

[12] P. Dhiman, A. Kaur, D. Gupta, S. Juneja, A. Nauman, and G. Muhammad, "GBERT: A hybrid deep learning model based on GPT-BERT for fake news detection," *Heliyon*, vol. 10, no. 16, Aug. 2024, doi: 10.1016/j.heliyon.2024.e35865.

[13] C. Jocelynne, I. L. Wijayakusuma, and L. P. I. Harini, "Detection of Political Hoax News Using Fine-Tuning IndoBERT," *Journal of Applied Informatics and Computing*, vol. 9, no. 2, pp. 354–360, Mar. 2025, doi: 10.30871/jaic.v9i2.8989.

[14] G. I. W. Koto, F. I. Rahmaningtyas, R. Mahendra, and A. Purwarianti, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," in *Proceedings of the 12th Language Resources and Evaluation Conference (LREC)*, 2020. [Online]. Available: https://aclanthology.org/2020.lrec-1.420/

[15] Y. Shen, Q. Liu, N. Guo, J. Yuan, and Y. Yang, "Fake News Detection on Social Networks: A Survey," Nov. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/app132111877.

[16] D. Herlina, Y. P. Putri, and I. N. M. Astawa, "Analisis Sentimen Berita Berbahasa Indonesia Menggunakan Metode SVM dan TF-IDF," *Jurnal RESTI*, vol. 6, no. 2, 2022, [Online]. Available: https://ejournal.undip.ac.id/index.php/resti/article/view/39550

[17] D. M. Powers, "Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation," *Journal of Machine Learning Technologies*, 2011, [Online]. Available: https://www.researchgate.net/publication/220722885

[18] B. Sahiner, W. Chen, A. Pezeshk, and N. Petrick, "Comparison of two classifiers when the data sets are imbalanced: the power of the area under the precision-recall curve as the figure of merit versus the area under the ROC curve," in *Proc. SPIE 10136, Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment*, 2017. doi: 10.1117/12.2254742.

[19] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[20] A. Beger, "Precision-Recall Curves," *Social Science Research Network*, 2016, doi: 10.2139/SSRN.2765419.

[21] H. Dafaalla *et al.*, "Deep Learning Model for Selecting Suitable Requirements Elicitation Techniques," *Applied Sciences*, vol. 12, no. 18, 2022, doi: 10.3390/app12189060.