

# Sistem Cerdas Deteksi Berita Hoaks Berbasis IndoBERT dengan Antarmuka Web Interaktif

Ikrar Bagaskara<sup>1\*</sup>, Eko Purwanto<sup>2</sup>, Joni Maulindar<sup>3</sup>

<sup>1</sup>Fakultas Ilmu Komputer  
Universitas Duta Bangsa Surakarta  
1\*210103101@mhs.udb.ac.id

<sup>2</sup>Fakultas Ilmu Komputer  
Universitas Duta Bangsa Surakarta  
2eko\_purwanto@udb.ac.id

<sup>3</sup>Fakultas Ilmu Komputer  
Universitas Duta Bangsa Surakarta  
3joni\_maulindar@udb.ac.id

**Abstrak**— Penelitian ini bertujuan untuk merancang dan membangun sebuah sistem cerdas yang mampu mendeteksi berita hoaks berbahasa Indonesia secara akurat dan mudah diakses oleh publik. Tujuan ini untuk mengatasi tantangan penyebaran disinformasi serta kurangnya alat deteksi yang dioptimalkan untuk Bahasa Indonesia dan ramah pengguna. Metode yang digunakan adalah *fine-tuning* model Transformer *pre-trained*, IndoBERT, pada dataset berita lokal yang telah melalui pra-pemrosesan dan menghasilkan 27.431 artikel bersih. Model yang telah dilatih kemudian diimplementasikan ke dalam sebuah prototipe aplikasi web interaktif menggunakan *framework* Streamlit untuk pengujian fungsional. Hasil evaluasi pada data uji menunjukkan performa model yang sangat tinggi, dengan F1-Score mencapai 0.99 untuk kelas hoaks dan akurasi rata-rata tertimbang 100%. Meskipun sangat akurat, hasil penelitian juga mengidentifikasi batasan pada kemampuan model dalam mendeteksi hoaks yang ditulis dengan gaya jurnalistik profesional. Prototipe aplikasi web terbukti fungsional dalam skenario pengujian praktis

**Kata kunci**— Deteksi Hoaks, IndoBERT, *Natural Language Processing* (NLP), Klasifikasi Teks, Streamlit

**Abstract**— This research aims to design and build an intelligent system capable of accurately and accessibly detecting Indonesian-language hoax news, addressing the challenges of disinformation and the lack of user-friendly, Indonesian-optimized detection tools. The implemented method is the fine-tuning of a pre-trained Transformer model, IndoBERT, on a local news dataset that was pre-processed to yield 27,431 clean articles. The trained model was then implemented as an interactive web application prototype using the Streamlit framework for functional testing. The evaluation results on the test data show exceptionally high model performance, achieving an F1-Score of 0.99 for the hoax class and a weighted average accuracy of 100%. Despite its high accuracy, the results also identified a limitation in the model's ability to detect hoaxes written in a professional journalistic style. The web application prototype was proven functional in practical testing scenarios.

**Keywords**— Hoax Detection, IndoBERT, *Natural Language Processing* (NLP), Text Classification, Streamlit

## I. PENDAHULUAN

Penyebaran berita hoaks di Indonesia telah menjadi masalah serius yang mengancam stabilitas sosial, keamanan, dan kesehatan publik. Maraknya penggunaan media sosial dan platform digital sebagai sumber informasi utama masyarakat turut mempercepat distribusi konten menyesatkan. Pada awal tahun 2024, Kementerian Komunikasi dan Digital (Komdigi) melalui Siaran Pers No. 02/HM/KOMINFO/01/2024 mengungkapkan bahwa sebanyak 12.547 konten hoaks telah ditangani sejak Agustus 2018 hingga 31 Desember 2023. Mayoritas hoaks tersebut berkaitan dengan isu Covid-19, diikuti oleh politik dan pemilu 2024. Meluasnya penyebaran hoaks ini menjadi ancaman serius karena dapat memicu disinformasi, menurunkan kepercayaan publik terhadap informasi, serta mengganggu stabilitas sosial menjelang pemilu [1]. Penelitian yang dilakukan [2] mengungkapkan bahwa hoaks berdampak merugikan bagi masyarakat

Indonesia. Dampak yang terasa adalah menurunnya kepercayaan masyarakat terhadap informasi yang diterima, hoaks juga menimbulkan kebingungan dan ketidakpastian, sehingga masyarakat kesulitan membedakan antara fakta dan informasi yang dimanipulasi.

Secara khusus, solusi deteksi hoaks berbasis *artificial intelligence* terutama dalam bahasa Indonesia saat ini masih menghadapi sejumlah tantangan mendasar. Pertama, mayoritas model *Natural Language Processing* (NLP) yang digunakan tidak dioptimalkan untuk bahasa Indonesia, seperti model *Birectditional Encoder Representational Transform* (BERT) yang masih berbahasa Inggris, sehingga gagal menangkap konteks linguistik bahasa Indonesia [3]. Kedua, kurangnya integrasi antara model *artificial intelligence* dengan antarmuka yang ramah pengguna menyebabkan teknologi ini belum dimanfaatkan secara maksimal oleh masyarakat

umum. Hal ini sejalan dengan temuan [4], yang menyatakan bahwa meskipun terdapat berbagai alat *artificial intelligence* yang dirancang untuk pengguna non-teknis, banyak diantaranya masih memerlukan pemahaman dasar tentang konsep *artificial intelligence*, yang menjadi hambatan bagi adopsi yang lebih luas terutama bagi masyarakat awam teknologi.

Untuk mengatasi masalah tersebut, penelitian ini mengusulkan implementasi sistem cerdas berbasis *IndoBERT*. Varian dari model *BERT* yang sudah dilatih dengan berbagai korpus Bahasa Indonesia, seperti Wikipedia Bahasa Indonesia, kamus besar Bahasa Indonesia, berita, sosial media, dan lainnya. Solusi ini tidak hanya meningkatkan presisi deteksi melalui pemahaman semantik yang lebih baik terhadap teks Indonesia, tetapi juga memperluas aksesibilitas melalui antarmuka web yang *user-friendly*. Dengan mengintegrasikan model *NLP* dan desain antarmuka interaktif, sistem ini diharapkan menjadi alat edukasi publik yang efektif sekaligus mempercepat respon terhadap penyebaran disinformasi.

### 1.1. Penelitian Terkait

Penelitian terkait klasifikasi telah menjadi area penelitian yang aktif, seperti klasifikasi gambar yang dapat membantu perusahaan atau individu menentukan suatu objek, selain itu terdapat klasifikasi text berdasarkan kata yang dipelajari oleh model untuk menentukan suatu kalimat. Seperti yang dilakukan oleh [5] dalam konteks deteksi hoaks seputar Covid-19 menggunakan berbagai pendekatan model *machine learning* klasik dan menghasilkan model terbaik *random forest* dengan menambahkan *feature engineering* yang mampu mendapatkan skor akurasi hingga 96.05%. Adapun juga pendekatan yang berbeda dilakukan oleh [6] dimana memilih menggunakan *deep learning* dengan model LSTM dengan penambahan fitur *word embedding Glove* untuk memprediksi hoaks pada berita politik dan mendapatkan akurasi akhir di angka 99.8%. Penelitian lain yang lebih berfokus dalam pembangunan fitur untuk identifikasi cerdas hoaks menggunakan *decision tree classifier* juga berhasil mendapatkan hasil yang memuaskan dengan nilai akurasi 93.5% dan *F1-Score* 0.935 [7].

Adapun penelitian yang menggunakan pendekatan yang sama, yaitu menggunakan model *IndoBERT* mampu mencapai 95% pada data uji dengan nilai AUC sebesar 0.946 [8].

Meskipun penelitian-penelitian tersebut berhasil menunjukkan kapabilitas untuk klasifikasi berita hoaks, terdapat celah utama yang coba diatasi oleh penelitian ini. Pertama, seluruh studi kasus diatas hanya berfokus pada pembangunan model tanpa adanya implementasi nyata, sehingga penilaian kualitas model hanya didasarkan pada evaluasi yang sudah ditentukan. Nyatanya pada praktik lapangan, pembuat hoaks makin pintar dan mampu menarasikan beritanya seolah-olah berita asli, akibatnya tidak bisa dipastikan apakah model yang dibangun mampu menangani masalah tersebut dan akhirnya hanya bersifat konseptual. Kedua, mayoritas penelitian sebelumnya tidak mempertimbangkan aspek keterlibatan pengguna melalui antarmuka yang interaktif sebagai bagian dari sistem deteksi hoaks. Padahal, dalam penerapannya, sistem deteksi yang baik tidak hanya mengandalkan akurasi model, tetapi juga bagaimana informasi hasil klasifikasi disampaikan kepada pengguna secara transparan dan mudah dipahami. Kurangnya integrasi antara model deteksi dengan sistem antarmuka yang dirancang secara *human-centered* dapat mengurangi efektivitas sistem dalam mendorong tindakan preventif terhadap penyebaran hoaks, terutama pada pengguna non-teknis.

Penelitian ini memiliki perbedaan mendasar dibandingkan dengan studi-studi sebelumnya, terutama pada aspek penerapan model dalam skenario nyata. Model yang telah dilatih tidak hanya dievaluasi secara *offline*, tetapi juga diuji secara langsung terhadap teks-teks baru yang tidak ada pada dataset, termasuk berita aktual. Selain itu, model diintegrasikan ke dalam sebuah sistem interaktif menggunakan *framework* Streamlit, sehingga memungkinkan pengguna untuk memeriksa keaslian suatu berita secara langsung. Dengan demikian, penelitian ini menunjukkan bahwa model yang dikembangkan memiliki potensi operasional dan tidak hanya terbatas pada validasi konseptual di atas data uji.

## 1.2. Landasan Teori

Penelitian ini menggabungkan domain *natural language processing* yang dijelaskan oleh [9] sebagai cabang dari Artificial Intelligence (AI) yang berfokus pada interaksi antara manusia dan komputer melalui bahasa alami. NLP mempelajari berbagai aspek penggunaan bahasa, mulai dari kata, kalimat, paragraf, hingga keseluruhan dokumen. Dan domain klasifikasi yang dijabarkan oleh [10] sebagai pendekatan untuk mengelompokkan target ke dalam kelas yang sudah ditentukan. Dengan memanfaatkan kedua domain ini, maka prediksi berita hoaks dapat diintegrasikan menjadi sistem cerdas yang mampu meniru kecerdasan manusia untuk berinteraksi dengan lingkungan luar [11]. Hadirnya model unggul seperti BERT yang dimaksudkan untuk pembelajaran awal yang mendalam dari representasi teks dua arah dan kemudahan dalam penggunaannya [12] ditambah dengan hadirnya model IndoBERT varian dari BERT yang di kembangkan khusus menggunakan Bahasa Indonesia yang dilatih murni sebagai *masked language model* dengan konfigurasi bawaan BERT-Base sehingga mampu menangkap suasana linguistik Bahasa Indonesia secara tepat [3].

## II. METODOLOGI PENELITIAN

### 2.1 Data Collecting

Pada tahap ini, pengumpulan data akan dilakukan dengan mengumpulkan dataset berita hoaks dan non-hoaks dari sumber yang terpercaya. Dataset utama akan diperoleh dari Turnbackhoax.id dipilih karena merupakan platform *fact-checking* terverifikasi di Indonesia. Ditambah dengan sumber lain seperti CNN, Tempo, dan Kompas. Dataset dipilih sebagai representasi sumber berita non-hoaks yang kredibel dan memiliki standar jurnalisme yang diakui. Data ini kemudian akan disimpan dalam format *XLSX* untuk memudahkan proses selanjutnya.

### 2.2 Data Understanding

Pada tahapan ini, *dataset* hasil *scraping* akan dilakukan eksplorasi mendalam terlebih dahulu, termasuk mengetahui jumlah *missing values* atau jumlah duplikasi pada *dataset*. Selain itu tahapan *Exploratory Data Analysis (EDA)*, seperti melakukan analisis deskriptif untuk mengenal

*dataset* lebih dalam. Tahapan *data understanding* diperlukan untuk menentukan tindakan apa saja yang akan diambil pada tahapan berikutnya, yaitu *data preparation*.

### 2.3 Data Preparation

Setelah data dikumpulkan, langkah selanjutnya adalah melakukan *data preparation* atau menyiapkan data sebelum dilakukan *modeling* dan *evaluation*, tahapan *data preparation* sendiri meliputi, penghapusan kolom data yang memiliki nilai kosong, memecah data dengan skenario pembagian 60% untuk data latih, 20% untuk data validasi, dan sisa 20% untuk data uji untuk memastikan model dapat belajar dan menguji data secara tepat. Setelah itu, teks akan dipecah menjadi kata-kata individu melalui proses *feature extraction* menggunakan bantuan *tokenizer* yang khusus digunakan untuk model IndoBERT.

### 2.4 Modelling

Proses pemodelan akan dilakukan dengan menggunakan *IndoBERT*, sebuah model berbasis *transformer* yang dioptimalkan untuk bahasa Indonesia. Langkah pertama adalah *fine-tuning*, di mana model *IndoBERT* akan disesuaikan menggunakan dataset yang telah di-*preprocessing*. *Fine-tuning* akan dilakukan dengan menambahkan layer klasifikasi di atas model *pre-trained*. Selanjutnya, akan dilakukan pemilihan *hyperparameter* seperti *learning rate*, *batch size*, dan jumlah *epoch* untuk mendapatkan performa terbaik. Setelah itu, model akan dilatih menggunakan data latih dan divalidasi menggunakan data validasi. Proses pelatihan akan dilakukan dengan memantau *loss* dan akurasi pada *setiap epoch* untuk menghindari *overfitting*. Setelah pelatihan selesai, model akan siap digunakan untuk deteksi hoaks.

### 2.5. Evaluation

Evaluasi dilakukan untuk mengukur performa model dalam mendeteksi berita hoaks. Untuk mengukur performa model klasifikasi multi-kelas secara komprehensif, beberapa matrik evaluasi digunakan. Landasan dari evaluasi ini adalah *Confusion Matrix*, yang kemudian menjadi dasar untuk menghitung matrik kuantitatif lainnya seperti *Accuracy*, *Precision*, *Recall*, dan *F1-Score*

#### 2.5.1 Confusion Matrix

Merupakan sebuah tabel yang memvisualisasikan kinerja model dengan membandingkan hasil prediksi dengan label aktual. Dari matriks ini, dapat melihat secara detail jenis-jenis kesalahan yang dibuat oleh model. Komponen utamanya adalah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

### 2.5.2 Matriks Kinerja

a. **Accuracy (Akurasi):** Mengukur persentase prediksi yang benar dari keseluruhan data uji. Meskipun umum digunakan, akurasi bisa menjadi matrik yang kurang representatif jika dataset tidak seimbang (*imbalanced*). Dengan formula matematis sebagai berikut

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

b. **Precision (Presisi):** Mengukur tingkat ketepatan dari prediksi positif. Dari semua ulasan yang diprediksi sebagai kelas tertentu (misal: 'positif'), berapa persen yang benar-benar 'positif'. Dengan formula matematis sebagai berikut

$$Precision = \frac{TP}{TP + FP}$$

c. **Recall (Daya Ingat):** Mengukur tingkat keberhasilan model dalam menemukan kembali sebuah kelas. Dari semua ulasan yang seharusnya 'positif', berapa persen yang berhasil diprediksi dengan benar oleh model. Dengan formula matematis sebagai berikut

$$Recall = \frac{TP}{TP + FN}$$

d. **F1-Score:** Merupakan rata-rata harmonik dari *Precision* dan *Recall*. Metrik ini memberikan gambaran yang lebih seimbang mengenai kinerja model, terutama pada dataset yang tidak seimbang, karena mempertimbangkan baik *False Positive* maupun *False Negative*. [13] Dengan formula matematis sebagai berikut

$$F1 - score = 2x \frac{Precision \times Recall}{Precision + Recall}$$

### 2.6. Deployment

Tahap *deployment* diimplementasikan menggunakan metode *prototyping*. [14] menjelaskan pendekatan *prototyping* adalah mewujudkan konsep dengan maksud untuk menganalisis, dengan prototipe sebagai contoh desain yang lebih ringkas

daripada produk akhir. Prototipe sering menjadi metode andalan karena memungkinkan pengiriman produk lebih awal kepada pengguna untuk mendapatkan umpan balik secara cepat [15]. Pada pengembangan sistem *machine learning* menggunakan *framework Streamlit* sering menjadi rujukan karena memungkinkan untuk mengembangkan sistem interaktif secara cepat dan menggunakan sedikit kode. Streamlit sendiri dirancang untuk mempermudah konversi skrip data menjadi aplikasi web yang dapat dibagikan dengan mudah [16].

## III. HASIL DAN PEMBAHASAN

### 3.1. Hasil Pra-Pemrosesan

Tahap awal penelitian berfokus pada persiapan data. Dari total 31.332 artikel berita mentah yang dikumpulkan, dilakukan proses pra-pemrosesan untuk menangani nilai yang hilang dan data duplikat. Proses ini menghasilkan dataset final yang bersih dan siap digunakan untuk pemodelan sebanyak 27.431 artikel. Analisis pada dataset final ini menunjukkan adanya ketidakseimbangan kelas (*class imbalance*), dengan distribusi yang terdiri dari 20.945 artikel valid (label 0) dan 6.486 artikel hoaks (label 1). Untuk kebutuhan eksperimen, dataset tersebut kemudian dibagi menjadi tiga bagian dengan proporsi 60% data latih (16.458 artikel), 20% data validasi (5.486 artikel), dan 20% data uji (5.487 artikel).

Tabel 1. Informasi Deskriptif Dataset

Deskripsi	Jumlah
Data Mentah	31.332 baris data
Data Bersih	27.431 baris data
Artikel Valid	20.945 baris data
Artikel Hoaks	6.486 baris data

### 3.2. Hasil Eksperimen dan Evaluasi Kinerja Model

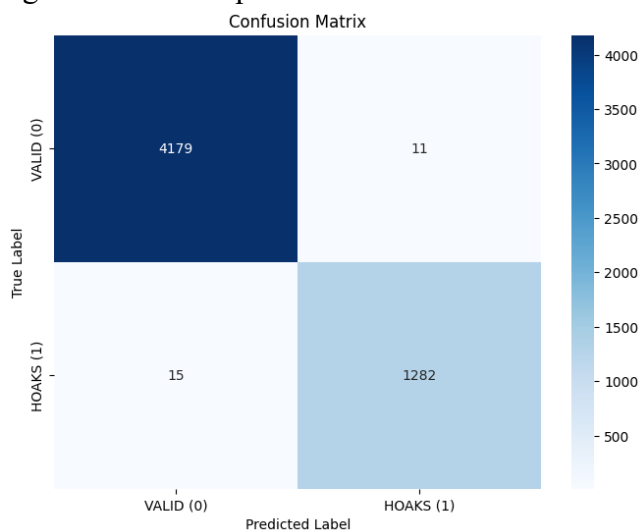
Pada tahap evaluasi, model diuji menggunakan 5.487 data uji yang belum pernah dilihat sebelumnya untuk mengukur performa generalisasinya. Hasil kuantitatif, yang disajikan dalam Laporan Klasifikasi (Tabel 2), menunjukkan kinerja yang sangat tinggi. Model berhasil mencapai akurasi rata-rata sebesar 100%, mengindikasikan kemampuan

model secara keseluruhan dalam membedakan antara berita valid dan hoaks sangat baik.

Tabel 2. Laporan Evaluasi Klasifikasi

	precision	recall	f1-score	support
valid (0)	1.00	1.00	1.00	4190
hoaks(1)	0.99	0.99	0.99	1297
accuracy			1.00	5487
macro avg	0.99	0.99	0.99	5487
weighted avg	1.00	1.00	1.00	5487

Untuk analisis yang lebih granular, metrik *Precision*, *Recall*, dan *F1-Score* memberikan wawasan yang lebih dalam. Untuk kelas hoaks (label 1), model mencatatkan *Precision* sebesar 0.99, yang berarti dari semua berita yang diprediksi sebagai hoaks oleh model, 99% di antaranya memang benar-benar hoaks. *Recall* sebesar 0.99 juga menunjukkan bahwa model mampu mengidentifikasi 99% dari seluruh berita hoaks yang ada di dalam data uji. Nilai *F1-Score* 0.99 untuk kelas ini, yang merupakan rata-rata harmonik dari *Precision* dan *Recall*, mengonfirmasi bahwa model memiliki keseimbangan yang sangat baik antara tidak salah menuduh berita valid sebagai hoaks dan tidak melewatkan berita hoaks yang sebenarnya. Performa pada kelas valid bahkan lebih sempurna, dengan ketiga metrik mencapai skor 1.00.

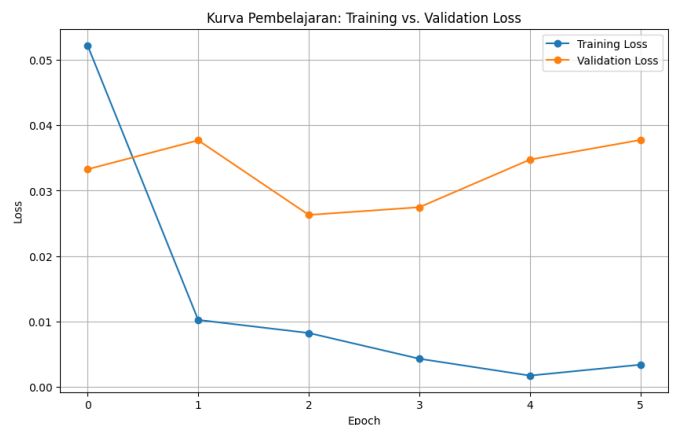


Gambar 1. Confusion Matrix untuk memberikan visualisasi detail.

Analisis lebih lanjut pada *Confusion Matrix* (Gambar 1) memberikan visualisasi detail mengenai jenis kesalahan yang dibuat model. Dari total 1.297

berita hoaks, model hanya melakukan 15 kesalahan klasifikasi sebagai 'valid' (*False Negative*). Jenis kesalahan ini tergolong paling kritis dalam skenario dunia nyata, karena berpotensi membiarkan disinformasi terlewatkan. Di sisi lain, dari 4.179 berita valid, hanya ada 11 kesalahan klasifikasi sebagai 'hoaks' (*False Positive*). Jumlah kedua jenis kesalahan yang sangat minim ini menunjukkan bahwa model tidak hanya akurat, tetapi juga andal dan dapat dipercaya dalam tugas klasifikasinya.

Untuk memastikan model tidak mengalami *overfitting*, learning curve selama proses pelatihan dianalisis (Gambar 2). Grafik menunjukkan bahwa nilai *training loss* dan *validation loss* sama-sama menurun secara konsisten pada *epoch-epoch* awal, menandakan model belajar secara efektif. Proses pelatihan juga menerapkan mekanisme *early stopping* yang secara otomatis menghentikan pelatihan pada *epoch* ke-5 saat *validation loss* tidak lagi menunjukkan perbaikan, dan menyimpan bobot model terbaik dari *epoch* ke-2 di mana *validation loss* berada di titik terendah (0.0255). Pola kurva ini mengonfirmasi bahwa model memiliki kemampuan generalisasi yang baik dan tidak hanya 'menghafal' data latih.



Gambar 2. Grafik Training dan Validation Loss untuk menguji pelatihan

### 3.3. Pembahasan dan Analisis Hasil

#### 1. Interpretasi Kinerja Model yang Sangat Tinggi

Kinerja klasifikasi yang sangat tinggi, khususnya F1-Score 0.99 untuk kelas hoaks, dapat diatribusikan pada dua faktor utama dalam metodologi penelitian. Pertama adalah pemilihan model IndoBERT. Sebagai model yang telah dilatih secara khusus pada korpus besar Bahasa Indonesia, IndoBERT memiliki

pemahaman mendalam mengenai konteks, semantik, dan nuansa linguistik yang khas dalam teks berita berbahasa Indonesia. Kedua, proses *fine-tuning* yang dilakukan memungkinkan model dasar tersebut untuk mengadaptasi pengetahuannya secara spesifik pada tugas klasifikasi berita hoaks, dengan menambahkan *layer* klasifikasi di atas model *pre-trained*. Kombinasi dari pemahaman bahasa yang kuat dan spesialisasi tugas inilah yang memungkinkan model untuk mencapai tingkat akurasi yang tinggi.

## 2. Analisis Kesalahan (*Error Analysis*)

Meskipun kinerjanya tinggi, analisis pada 26 kasus kesalahan memberikan wawasan penting mengenai batasan model. Ditemukan bahwa 15 kasus *False Negative* cenderung terjadi pada berita hoaks yang ditulis dengan narasi seolah-olah seperti berita asli. Hal ini dikarenakan salah satu cara model belajar untuk mengklasifikasikan suatu berita adalah berdasarkan gaya penulisan pada artikel, meliputi penggunaan huruf besar kecil, karakter pada berita, dan narasi yang tidak dibuat secara profesional. Oleh karena itu jika model dihadapkan dengan narasi yang dibuat dengan cara jurnalistik, maka model akan sulit untuk mengklasifikasikan dan cenderung salah memberikan output.

Sebagai contoh spesifik, salah satu berita hoaks yang keliru diklasifikasikan sebagai 'valid' adalah artikel mengenai 'program nasional melawan diabetes' (Tabel 3). Teks ini berhasil menipu model karena narasinya meniru struktur siaran pers resmi, menggunakan kutipan fiktif dari tokoh publik (Raffi Ahmad), dan menyajikan klaimnya dengan gaya bahasa yang terstruktur dan meyakinkan, pola yang dominan ditemukan pada berita valid dalam data latih.

## 3. Perbandingan dengan Penelitian Terdahulu

Ketika dibandingkan dengan penelitian terkait yang telah dibahas, kinerja model IndoBERT dalam penelitian ini menunjukkan hasil yang sangat kompetitif. Model ini berhasil melampaui *F1-Score* yang dicapai oleh model *Decision Tree Classifier* (0.935) dan akurasi yang dihasilkan oleh pendekatan *machine learning* klasik seperti *Random Forest* (96.05%). Bahkan jika dibandingkan dengan pendekatan *deep learning* lain yang menggunakan

IndoBERT (akurasi 95%), model saat ini menunjukkan keunggulan. Yang lebih penting, penelitian ini tidak berhenti pada validasi konseptual, melainkan berhasil mengatasi celah utama dari studi-studi sebelumnya dengan mengintegrasikan model ke dalam sebuah sistem antarmuka web interaktif, yang membuktikan potensi operasionalnya

## 3.4. Uji Coba pada Skenario Praktis.

Untuk mendemonstrasikan kapabilitas operasional model di dunia nyata, dilakukan uji coba inferensi langsung pada teks berita baru yang tidak termasuk dalam dataset pelatihan maupun pengujian. Pengujian ini dirancang untuk mensimulasikan penggunaan sistem oleh pengguna akhir yang memasukkan teks secara dinamis. Sebagai contoh pertama, ketika model diberi masukan teks berita valid mengenai peresmian tol oleh Presiden Joko Widodo, model berhasil mengklasifikasikannya dengan benar sebagai 'valid'. Tingkat kepercayaan (*confidence score*) yang dihasilkan untuk prediksi ini pun sangat meyakinkan, yaitu sebesar 97.45%.

Teks: 'Presiden Joko Widodo meresmikan tol baru yang meng...'  
Hasil Prediksi: VALID (Kepercayaan: 97.45%)

Gambar 3. Hasil inferensi model pada contoh berita valid tentang peresmian tol, menunjukkan prediksi 'VALID' dengan tingkat kepercayaan 97.45%

Pada skenario kedua, sebuah teks berisi disinformasi populer mengenai efek minum air kelapa setelah vaksinasi diujikan pada model. Hasilnya, model kembali menunjukkan ketepatannya dengan memberikan klasifikasi 'hoaks', dengan tingkat kepercayaan yang nyaris sempurna mencapai 99.99%. Kedua hasil uji coba praktis ini mengonfirmasi bahwa model yang dikembangkan tidak hanya akurat pada data terstruktur, tetapi juga fungsional dan dapat diandalkan saat dihadapkan pada data baru. Kemampuan model untuk memberikan label beserta tingkat kepercayaannya juga meningkatkan aspek transparansi dan kegunaan sistem bagi pengguna.

Teks: 'Beredar kabar bahwa minum air kelapa setelah vaksi...'  
Hasil Prediksi: HOAKS (Kepercayaan: 99.99%)

Gambar 4. Hasil inferensi model pada contoh berita hoaks tentang vaksinasi, menunjukkan prediksi 'HOAKS' dengan tingkat kepercayaan 99.99%

## 3.5. Implementasi Prototipe Interaktif

Sebagai realisasi dari tujuan penelitian untuk menciptakan alat yang aplikatif, model deteksi hoaks yang telah divalidasi kemudian diimplementasikan ke dalam sebuah prototipe sistem cerdas berbasis web. Sesuai dengan rancangan pada bab metodologi, sistem ini dibangun menggunakan framework *Streamlit*, yang dipilih karena kemampuannya dalam mengonversi skrip analisis data menjadi aplikasi web interaktif secara efisien.



Gambar 5. Interface dari aplikasi interaktif untuk memeriksa berita hoaks

Antarmuka pengguna dirancang dengan mengutamakan kemudahan penggunaan, seperti

yang diperlihatkan pada Gambar 5 Pengguna disediakan sebuah area input teks di mana mereka dapat langsung mengetik atau menempelkan keseluruhan isi berita yang ingin diperiksa. Setelah menekan tombol 'Prediksi', sistem akan memanggil fungsi inferensi untuk memproses teks tersebut menggunakan model IndoBERT yang telah di-fine-tune. Hasilnya kemudian ditampilkan secara langsung pada antarmuka. Output yang diberikan tidak hanya berupa label klasifikasi akhir ('valid' atau 'hoaks'), tetapi juga disertai dengan skor kepercayaan (confidence score) dari prediksi tersebut.

Penyajian skor kepercayaan ini bertujuan untuk memberikan transparansi kepada pengguna mengenai seberapa yakin model terhadap keputusannya. Keberhasilan implementasi prototipe ini menjadi bukti konkret bahwa penelitian ini tidak hanya menghasilkan sebuah model, tetapi juga sebuah sistem fungsional yang menjembatani antara kompleksitas teknologi AI dengan kebutuhan praktis pengguna akhir. Berikut adalah hasil pengujian dengan berbagai skenario yang disajikan pada Tabel 3.

Tabel 3. Hasil Pengujian Model

Potongan Kalimat	Skor Kepercayaan	Hasil	Tautan Artikel
Saat ini, tambah Ferry seluruh penumpang telah dievakuasi ke ruang tunggu internasional KNIA untuk keamanan. Selain itu, Tim Jibom melakukan skrining dan pengecekan menyeluruh atas ancaman tersebut ....	100%	Valid	<a href="https://www.cnnindonesia.com/nasional/20250617131038-20-1240609/saudia-airlines-mendarat-darurat-di-kualanamu-karena-ada-ancaman-bom">https://www.cnnindonesia.com/nasional/20250617131038-20-1240609/saudia-airlines-mendarat-darurat-di-kualanamu-karena-ada-ancaman-bom</a>
Kekhawatiran akan eskalasi letusan Gunung Lewotobi Laki-laki membuat sebagian warga memilih meninggalkan rumah mereka. Sejumlah warga Desa Waiula memutuskan mengungsi ke Desa Riang Baring, Kecamatan Ile Bura, yang dianggap cukup aman.....	99.99%	Valid	<a href="https://www.kompas.com/sulawesi-selatan/read/2025/06/17/210025588/gunung-lewotobi-laki-laki-meletus-dahsyat-warga-mengungsi-usai">https://www.kompas.com/sulawesi-selatan/read/2025/06/17/210025588/gunung-lewotobi-laki-laki-meletus-dahsyat-warga-mengungsi-usai</a>
PARAH! Candi Borobudur Akan Dipasang Eskalator Agar Prabowo Bisa Naik ke Atas Gak cukup sejarah yang direvisi, bangunan sejarah sehebat candi borobudur pun juga direnovasi menurut kemauan rezim	100%	Hoaks	<a href="https://turnbackhoax.id/2025/06/06/salah-video-candi-borobudur-bakal-dipasang-eskalator-agar-prabowo-bisa-naik-ke-atas/">https://turnbackhoax.id/2025/06/06/salah-video-candi-borobudur-bakal-dipasang-eskalator-agar-prabowo-bisa-naik-ke-atas/</a>
TOLAK Program Membagi <sup>2</sup> OBAT CACING/IMUNISASI PLUS VACSIN dari Kementerian Kesehatan Sebab ini Bagian dari Misi Global Dalam Rangka Penyebaran Virus Baru. Target Utama Mereka Sekolah <sup>2</sup> dan Keluarga yang Awam Terhadap Kesehatan.	99.83%	Hoaks	<a href="https://turnbackhoax.id/2025/05/30/salah-virus-baru-disebarkan-ke-anak-sekolah-melalui-pembagian-obat-cacing/">https://turnbackhoax.id/2025/05/30/salah-virus-baru-disebarkan-ke-anak-sekolah-melalui-pembagian-obat-cacing/</a>

#### Pengujian Berita Hoaks dengan Gaya Penulisan Jurnalistik

Dalam adegan yang mengejutkan dan memilukan, pendudukan Zionis meluncurkan roket ke Gaza dengan tulisan “Made in Türkiye – 2024”, pada saat Türkiye dikenal karena sikap mendukungnya terhadap Gaza. Ironi yang menyakitkan adalah bahwa senjata dari negara yang menyatakan solidaritasnya digunakan saat ini untuk mengebom orang-orang ....	99.01%	Valid	<a href="https://turnbackhoax.id/2025/05/16/salah-israel-pakai-bom-buatan-turkiye-untuk-meny Serang-gaza/">https://turnbackhoax.id/2025/05/16/salah-israel-pakai-bom-buatan-turkiye-untuk-meny Serang-gaza/</a>
Hari ini, pemerintah Indonesia secara resmi memasukkan metode ini ke dalam program nasional untuk melawan diabetes sesuai jadwal hingga akhir musim dingin 2025 seluruh penduduk Indonesia akan bebas dari diabetes metode yang dikembangkan oleh dokter Tony Setiobudi ini sangat sederhana...	99.96%	Hoaks	<a href="https://turnbackhoax.id/2025/03/27/penipuan-raffi-ahmad-metode-air-garam-jadi-program-nasional-perangi-diabetes/">https://turnbackhoax.id/2025/03/27/penipuan-raffi-ahmad-metode-air-garam-jadi-program-nasional-perangi-diabetes/</a>

Hasil pengujian kualitatif yang dirangkum pada Tabel 3 memperlihatkan performa model dalam berbagai skenario penulisan. Teridentifikasi bahwa model memiliki keunggulan pada dua skenario umum: pertama, berita valid yang ditulis sesuai kaidah jurnalistik berhasil dikenali dengan akurasi sangat tinggi. Kedua, model juga efektif mendeteksi berita hoaks dengan narasi yang kurang terstruktur atau cenderung provokatif.

Namun, temuan paling krusial dari pengujian ini adalah teridentifikasinya kelemahan model saat dihadapkan pada berita hoaks yang disusun secara cermat untuk meniru gaya jurnalistik standar. Pada skenario yang paling menantang ini, model menunjukkan kecenderungan untuk keliru mengklasifikasikan disinformasi tersebut sebagai berita valid, yang merupakan jenis kesalahan prediksi paling berisiko (*False Negative*).

#### IV. KESIMPULAN

Penelitian ini dilaksanakan untuk menjawab tantangan serius penyebaran berita hoaks di Indonesia, yang diperparah oleh kurangnya alat deteksi yang dioptimalkan untuk Bahasa Indonesia dan dapat diakses oleh masyarakat umum. Tujuan utama dari penelitian ini adalah merancang, membangun, dan mengevaluasi sebuah sistem cerdas yang mampu mengklasifikasikan berita hoaks berbahasa Indonesia secara akurat. Untuk mencapai tujuan tersebut, penelitian ini mengimplementasikan model *Natural Language Processing* (NLP) mutakhir, IndoBERT, yang kemudian di-*fine-tune* menggunakan dataset berita lokal dan diintegrasikan

ke dalam sebuah prototipe aplikasi web interaktif menggunakan *framework* Streamlit.

Temuan utama dari penelitian ini menunjukkan bahwa pendekatan yang diusulkan sangat efektif. Secara kuantitatif, model yang telah di-*fine-tune* berhasil mencapai kinerja yang sangat tinggi pada data uji, dengan F1-Score mencapai 0.99 untuk kelas hoaks dan akurasi keseluruhan 100%, yang mengindikasikan kemampuan generalisasi yang kuat. Namun, analisis kualitatif juga mengidentifikasi batasan utama model, yaitu kesulitan dalam mendeteksi berita hoaks yang ditulis secara cermat menyerupai gaya jurnalistik profesional. Di sisi implementasi, penelitian ini berhasil mewujudkan sebuah prototipe aplikasi web yang fungsional, di mana pengguna dapat secara langsung menguji keaslian berita dan menerima umpan balik instan beserta skor kepercayaan.

Kontribusi utama dari penelitian ini ada pada dua level. Secara teoritis, penelitian ini mengonfirmasi bahwa arsitektur berbasis *Transformer* seperti IndoBERT merupakan solusi yang sangat andal untuk tugas klasifikasi teks dalam Bahasa Indonesia. Secara praktis, dan ini yang paling signifikan, penelitian ini berhasil menjembatani celah antara riset konseptual dan aplikasi dunia nyata dengan menghasilkan sistem fungsional pertama yang mudah diakses, menjawab kritik terhadap penelitian sebelumnya yang seringkali berhenti pada level model.

Meskipun hasil yang dicapai sangat memuaskan, penelitian ini memiliki beberapa keterbatasan yang dapat menjadi landasan untuk penelitian di masa depan. Keterbatasan utama adalah kerentanan model

terhadap hoaks yang ditulis dengan gaya jurnalistik rapi. Oleh karena itu, saran untuk penelitian mendatang adalah:

1. Mengembangkan model *hybrid* yang tidak hanya menganalisis isi teks, tetapi juga mempertimbangkan fitur-fitur eksternal seperti reputasi sumber berita atau analisis pola penyebaran tautan.
2. Memperkaya dataset pelatihan dengan lebih banyak contoh hoaks canggih (*sophisticated hoaxes*) untuk meningkatkan ketangguhan model.
3. Melakukan studi pengalaman pengguna (UX research) pada prototipe yang ada untuk mengukur efektivitasnya dalam meningkatkan literasi digital masyarakat dan mengembangkannya menjadi aplikasi skala penuh.

#### UCAPAN TERIMA KASIH

Penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada semua individu dan pihak yang telah memberikan kontribusi dan dukungan selama penelitian ini. Tanpa kerja keras, panduan, dan bantuan mereka, penelitian ini tidak akan terwujud. Penulis ingin mengucapkan terima kasih kepada:

1. Pembimbing penulis, Bapak Eko Purwanto, S.Kom., M.Kom, atas arahan, saran, dan pengawasannya yang berharga sepanjang penelitian ini.
2. Pembimbing penulis, Bapak Joni Maulindar, S.Kom., M.Eng, atas arahan, saran, dan pengawasannya yang berharga sepanjang penelitian ini.
3. Keluarga dan teman-teman penulis atas dukungan, motivasi, dan pengertian yang mereka berikan selama proses penelitian ini.
4. Institusi dan lembaga yang telah memberikan akses dan izin untuk mengumpulkan data yang diperlukan.

Terima kasih atas semua dukungan dan kontribusi yang telah diberikan. Semua bantuan ini sangat berarti bagi kelancaran dan kesuksesan penelitian penulis.

#### REFERENSI

- [1] Kementerian Komunikasi dan Digital, "Siaran Pers No. 02/HM/KOMINFO/01/2024 tentang Hingga Akhir Tahun 2023, Kominfo Tangani 12.547 Isu Hoaks," 2024. Diakses: 1 Mei 2025. [Daring]. Tersedia pada: <https://www.komdigi.go.id/berita/pengumuman/detail/siaran-pers-no-02-hm-kominfo-01-2024-tentang-hingga-akhir-tahun-2023-kominfo-tangani-12-547-isu-hoaks>
- [2] D. Rahmawati, R. S. P. Robawa, M. F. A. Abiyi, P. D. N. Rf, R. I. Nugraha, dan F. P. Margono, "Analisis Hoaks dalam Konteks Digital: Implikasi dan Pencegahannya di Indonesia," *INNOVATIVE: Journal Of Social Science Research*, vol. 3, hlm. 10819–10829, 2023.
- [3] F. Koto, A. Rahimi, J. H. Lau, dan T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," 2020, *arXiv*. doi: 10.48550/ARXIV.2011.00677.
- [4] A. R. Jakkula, "Building AI Tools for Non-Technical Users: A Case Study Approach," 2023.
- [5] A. T. B. Panjaitan dan I. Santoso, "Deteksi Hoaks Pada Berita Berbahasa Indonesia Seputar COVID-19," *FORMAT*, vol. 10, no. 1, hlm. 76, Feb 2021, doi: 10.22441/format.2021.v10.i1.007.
- [6] R. A. Sunan, H. F. E. K., dan C. S. K. Aditya, "Klasifikasi Hoax Berita Politik Menggunakan Algoritma Long Short-Term Memory (LSTM) dengan Penambahan Fitur Embedding Global Vector (GloVe)," *JEPIN*, vol. 10, no. 2, hlm. 287, Agu 2024, doi: 10.26418/jp.v10i2.76042.
- [7] M. U. Shalih dan T. E. E. Tju, "Pembangunan Fitur dalam Identifikasi Cerdas Hoaks dengan Naïve Bayes dan Klasifikasi Decision Tree," *Jutisi J. Tek. Sis. Info*, vol. 13, no. 1, hlm. 142, Apr 2024, doi: 10.35889/jutisi.v13i1.1731.
- [8] C. Jocelyne, I. L. Wijayakusuma, dan L. P. I. Harini, "Detection of Political Hoax News Using Fine-Tuning IndoBERT," *JAIC*, vol. 9, no. 2, hlm. 354–360, Mar 2025, doi: 10.30871/jaic.v9i2.8989.
- [9] M. Zhou, N. Duan, S. Liu, dan H.-Y. Shum, "Progress in Neural NLP: Modeling, Learning, and Reasoning," *Engineering*, vol. 6, no. 3, hlm. 275–290, Mar 2020, doi: 10.1016/j.eng.2019.12.014.
- [10] J. E. Black, J. K. Kueper, dan T. S. Williamson, "An introduction to machine learning for classification and prediction," *Family Practice*, vol. 40, no. 1, hlm. 200–204, Feb 2023, doi: 10.1093/fampra/cm104.
- [11] M. Amin, "Sistem Cerdas Kontrol Kran Air Menggunakan Mikrokontroler Arduino dan Sensor Ultrasonic," *InfoTekJar : J. nas. inform. dan teknol.*, 2020, doi: <https://doi.org/10.30743/infotekjar.v4i2.2386>.
- [12] M. V. Koroteev, "BERT: A Review of Applications in Natural Language Processing and Understanding," 2021, *arXiv*. doi: 10.48550/ARXIV.2103.11943.
- [13] S. A. Hicks *dkk.*, "On evaluation metrics for medical applications of artificial intelligence," *Sci Rep*, vol. 12, no. 1, hlm. 5979, Apr 2022, doi: 10.1038/s41598-022-09954-8.
- [14] L. Kent, C. Snider, J. Gopsill, dan B. Hicks, "Mixed reality in design prototyping: A systematic review," *Design Studies*, vol. 77, hlm. 101046, Nov 2021, doi: 10.1016/j.destud.2021.101046.
- [15] R. Wasik, "The Advantages & Disadvantages of Prototyping - Rapids Reproductions," *The Advantages & Disadvantages of Prototyping*. Diakses: 6 Juni 2025. [Daring]. Tersedia pada:

- <https://rapidsrepro.com/advantages-disadvantages-prototyping/>
- [16] Y. Akkem, B. S. Kumar, dan A. Varanasi, “Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation,” *IJST*, vol. 16, no. 48, hlm. 4688–4702, Des 2023, doi: 10.17485/IJST/v16i48.2850.