IMPLEMENTASI MODEL INDOBERT DAN MBERT UNTUK DETEKSI BERITA HOAKS BERBASIS WEB

Ananda Nuril Wahyuni, Tri Listyorini, Endang Supriyati
Teknik Informatika, Universitas Muria Kudus
Jl. Lkr. Utara, Kayuapu Kulon, Gondangmanis, Kec. Bae, Kabupaten Kudus, Jawa Tengah
anandanurilw@gmail.com

ABSTRAK

Penyebaran berita hoaks di media digital menjadi permasalahan serius karena dapat mempengaruhi opini publik dan menimbulkan dampak sosial yang luas. Tingginya volume informasi daring menyebabkan proses verifikasi manual menjadi tidak efektif dan sulit dilakukan secara *real-time*. Penelitian ini bertujuan untuk mengimplementasikan serta membandingkan kinerja model IndoBERT dan Multilingual BERT (mBERT) dalam mendeteksi berita hoaks berbahasa Indonesia berbasis web. Metode yang digunakan meliputi pengumpulan dataset sebanyak 21.373 data, pra-pemrosesan teks, pembagian data menggunakan *stratified split*, proses *fine-tuning* model *transformer*, dan evaluasi menggunakan metrik *accuracy, precision, recall*, dan *F1-score*, serta implementasi dalam sistem berbasis web menggunakan Flask. Hasil penelitian menunjukkan bahwa IndoBERT memperoleh performa terbaik dengan akurasi 94,52% dan *F1-score* 94,53%, sedangkan mBERT mencapai akurasi 91,58% dan *F1-score* 91,64%. Hasil ini menunjukkan bahwa model monolingual IndoBERT lebih efektif dalam memahami konteks bahasa Indonesia dibandingkan model multibahasa. Sistem yang dikembangkan dapat digunakan sebagai alat bantu verifikasi informasi secara cepat dan praktis.

Kata kunci : *deteksi hoaks, IndoBERT, mBERT, klasifikasi teks, sistem web*

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah mengubah cara masyarakat dalam memperoleh serta menyebarkan informasi. Media digital seperti portal berita daring, media sosial, serta berbagai platform berbasis *web* kini menjadi sumber utama informasi. Namun, kemudahan akses tersebut tidak selalu diimbangi dengan kualitas dan validitas informasi yang beredar. Salah satu permasalahan yang muncul secara signifikan adalah penyebaran berita bohong (hoaks), yang berpotensi menyesatkan masyarakat memengaruhi opini publik, serta menimbulkan dampak sosial yang merugikan [1]. Di Indonesia, penyebaran hoaks menjadi tantangan serius karena tingginya konsumsi informasi digital serta karakteristik bahasa yang beragam dan dinamis.

Tingginya volume berita yang diproduksi setiap hari menyebabkan proses verifikasi secara manual menjadi tidak efektif dan sulit dilakukan secara *real-time*. Hoaks tidak hanya berkaitan dengan isu politik, tetapi juga mencakup bidang kesehatan, ekonomi, dan sosial yang dapat memicu kepanikan publik serta menurunkan tingkat kepercayaan terhadap sumber informasi resmi [2]. Kondisi ini menunjukkan perlunya sistem otomatis yang mampu mendeteksi berita hoaks secara cepat dan akurat, khususnya untuk teks berbahasa Indonesia yang memiliki karakteristik linguistik tersendiri.

Perkembangan *Artificial Intelligence* pada bidang *Natural Language Processing* (NLP) memberikan pendekatan yang lebih efektif dalam melakukan klasifikasi teks. Model berbasis *transformer* seperti *Bidirectional Encoder Representations from Transformers* (BERT) memiliki kemampuan dalam memahami konteks semantik teks secara dua arah sehingga lebih unggul dibandingkan metode konvensional [3]. Untuk bahasa Indonesia, IndoBERT dikembangkan menggunakan korpus berbahasa Indonesia sehingga mampu menangkap struktur linguistik dan makna kontekstual secara lebih mendalam [4]. Di sisi lain, Multilingual BERT (mBERT) merupakan model multibahasa yang dilatih menggunakan berbagai bahasa termasuk bahasa Indonesia dan memiliki kemampuan representasi lintas bahasa [5]. Beberapa penelitian menunjukkan bahwa mBERT tetap dapat digunakan secara efektif untuk klasifikasi teks berbahasa Indonesia, meskipun dalam beberapa kasus performanya masih berada di bawah model monolingual seperti IndoBERT [6].

Berbagai penelitian telah membahas penggunaan model *transformer* untuk deteksi hoaks, namun sebagian besar masih berfokus pada evaluasi performa model secara eksperimental tanpa mengintegrasikannya ke dalam sistem yang dapat diakses langsung oleh pengguna [2], [6]. Padahal, implementasi berbasis *web* sangat penting untuk meningkatkan pemanfaatan teknologi deteksi hoaks oleh masyarakat serta mendukung literasi digital. Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengimplementasikan dan membandingkan performa IndoBERT dan mBERT dalam mendeteksi berita hoaks berbahasa Indonesia serta mengintegrasikannya ke dalam sistem deteksi hoaks berbasis *web*. Penelitian ini diharapkan dapat memberikan kontribusi akademis melalui evaluasi komparatif kedua model serta kontribusi praktis berupa sistem yang dapat digunakan secara langsung oleh pengguna.

2. TINJAUAN PUSTAKA

2.1. Hoaks

Hoaks merupakan informasi palsu atau menyesatkan yang disebarkan secara sengaja dengan tujuan tertentu, seperti memengaruhi opini publik, menciptakan keresahan sosial, atau memperoleh keuntungan pribadi. Penyebaran hoaks di Indonesia meningkat seiring dengan masifnya penggunaan internet dan media sosial sebagai sumber informasi utama masyarakat [7]. Konten hoaks umumnya dikemas menyerupai berita faktual sehingga sulit dibedakan oleh pembaca awam.

Hoaks di media online Indonesia cenderung sulit dibedakan karena menggunakan gaya bahasa persuasif dan struktur narasi yang menyerupai berita faktual [4]. Karakteristik teks hoaks umumnya meliputi penggunaan judul sensasional, kata-kata emosional, minimnya sumber terpercaya, serta adanya ajakan untuk segera menyebarkan informasi [8]. Dampak dari penyebaran hoaks tidak hanya merugikan individu, tetapi juga dapat mengganggu stabilitas sosial dan kepercayaan publik terhadap media. Untuk mengatasi permasalahan tersebut, berbagai pendekatan deteksi hoaks berbasis teks telah dikembangkan, mulai dari metode *rule-based*, *machine learning*, hingga *deep learning* yang mampu menganalisis pola linguistik secara otomatis.

2.2. Natural Language Processing

Natural Language Processing atau NLP merupakan cabang kecerdasan buatan yang fokus pada pemrosesan dan pemahaman bahasa alami oleh komputer. NLP memungkinkan sistem untuk memproses teks, memahami konteks, serta mengekstraksi informasi penting dari data berbentuk bahasa alami [9].

Pada tahap awal, pengembangan NLP banyak mengandalkan pendekatan *machine learning* konvensional yang menggunakan fitur buatan seperti TF-IDF untuk melakukan klasifikasi teks. Pendekatan ini cukup efektif, namun memiliki keterbatasan dalam memahami hubungan semantik dan konteks kalimat secara mendalam [4]. Seiring meningkatnya kompleksitas data teks, pendekatan *deep learning* mulai digunakan karena kemampuannya dalam mempelajari representasi fitur secara otomatis tanpa rekayasa fitur manual. Pendekatan *machine learning* dan *deep learning* telah terbukti efektif pada berbagai domain, termasuk kesehatan dan sistem pendukung keputusan, dengan tingkat akurasi klasifikasi yang tinggi [13].

Perkembangan *deep learning* dalam NLP kemudian ditandai dengan munculnya arsitektur *transformer* yang mampu menangkap dependensi jangka panjang dan konteks semantik secara lebih baik dibandingkan model sebelumnya seperti RNN. Penelitian yang membandingkan kedua pendekatan tersebut pada pemrosesan bahasa Indonesia dan bahasa daerah menunjukkan bahwa *transformer* memiliki performa yang lebih unggul dalam hal stabilitas pelatihan dan pemahaman konteks [15]. Keunggulan arsitektur *transformer* tersebut menjadikannya sangat relevan untuk tugas analisis teks yang kompleks, termasuk deteksi berita hoaks, karena mampu memahami konteks kalimat secara menyeluruh dan mengurangi kesalahan klasifikasi akibat ambiguitas bahasa [1], [10].

2.3. IndoBERT

IndoBERT merupakan model bahasa berbasis transformer yang dikembangkan khusus untuk bahasa Indonesia. Model ini dilatih menggunakan korpus teks berbahasa Indonesia dalam jumlah besar sehingga mampu memahami struktur linguistik dan konteks semantik secara lebih baik dibandingkan model konvensional [12].

Penggunaan IndoBERT dalam berbagai penelitian menunjukkan performa yang unggul dalam tugas klasifikasi teks, termasuk deteksi berita hoaks, karena kemampuannya dalam merepresentasikan konteks kalimat secara dua arah (*bidirectional*) [1], [4]. Selain itu, IndoBERT juga telah banyak dimanfaatkan pada tugas NLP lainnya, seperti analisis sentimen dan klasifikasi teks.

2.4. mBERT

Multilingual BERT (mBERT) merupakan model *transformer* yang dilatih menggunakan data multibahasa dan dirancang untuk menangani berbagai bahasa dalam satu model. Keunggulan mBERT terletak pada kemampuannya dalam mempelajari pola lintas bahasa (*cross-lingual representation*), sehingga dapat digunakan pada bahasa dengan sumber daya terbatas [11].

Meskipun tidak dilatih secara khusus untuk bahasa Indonesia, mBERT tetap menunjukkan performa yang kompetitif dalam berbagai tugas pemrosesan bahasa alami. Hal ini disebabkan oleh kemampuan model *transformer* dalam membangun representasi konteks yang kaya dan bersifat kontekstual (*contextualized representation*), sehingga mampu menangkap makna kata berdasarkan konteks kalimat secara efektif [10], [11]. Oleh karena itu, mBERT layak dibandingkan dengan IndoBERT untuk mengevaluasi efektivitas model multibahasa dalam deteksi berita hoaks berbahasa Indonesia.

2.5. Sistem Berbasis Web

Sistem berbasis *web* merupakan aplikasi perangkat lunak yang diakses dan dijalankan melalui browser *web*, memanfaatkan teknologi internet untuk menyediakan layanan atau informasi secara online. Sistem berbasis *web* memiliki keunggulan dalam hal aksesibilitas, fleksibilitas, dan kemudahan pemeliharaan, karena pengguna dapat mengaksesnya dari berbagai perangkat tanpa perlu instalasi khusus [11].

Beberapa penelitian mengintegrasikan model NLP ke dalam sistem *web* menggunakan *framework* seperti Flask atau Streamlit [2]. Pendekatan ini dinilai

efektif dalam meningkatkan aksesibilitas dan pemanfaatan sistem deteksi hoaks oleh masyarakat umum.

2.6. Penelitian Terdahulu

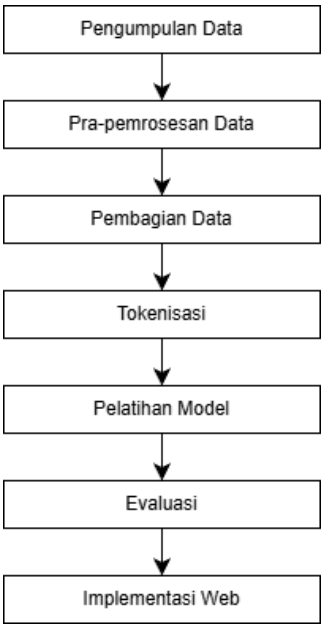
Penelitian deteksi berita hoaks di Indonesia pada tahap awal banyak menggunakan algoritma *machine learning* seperti Naïve Bayes, Support Vector Machine, dan Random Forest dengan representasi fitur berbasis TF-IDF [4], [7]. Namun, pendekatan ini memiliki keterbatasan dalam memahami konteks dan hubungan antar kata secara mendalam.

Seiring berkembangnya *deep learning*, penelitian mulai beralih ke model neural seperti Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), dan Long Short-Term Memory (LSTM) yang menunjukkan peningkatan performa pada klasifikasi teks [11]. Meskipun demikian, model-model tersebut masih belum optimal dalam merepresentasikan konteks kalimat secara menyeluruh.

Perkembangan selanjutnya ditandai dengan munculnya model berbasis *transformer* dan BERT yang mampu membangun representasi semantik teks secara kontekstual dan dua arah, sehingga menunjukkan peningkatan performa yang signifikan dalam mendeteksi hoaks [10], [12], serta pada berbagai tugas pemrosesan bahasa Indonesia berbasis *transformer* [15].

Meski demikian, sebagian besar penelitian masih berfokus pada evaluasi model secara *offline* tanpa integrasi sistem berbasis *web* [1]. Oleh karena itu, penelitian ini bertujuan mengisi celah tersebut dengan mengimplementasikan dan membandingkan IndoBERT serta mBERT dalam sistem deteksi hoaks berbasis *web* secara *end-to-end*, sehingga hasil yang diperoleh lebih aplikatif dan relevan bagi pengguna.

3. METODE PENELITIAN



Gambar 1. Alur penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimental untuk membandingkan kinerja dua model berbasis transformer, yakni IndoBERT dan mBERT untuk deteksi hoaks berita berbahasa Indonesia dan mengimplementasikannya dalam sistem berbasis *web*. Metode penelitian ini terdiri atas beberapa tahapan yang saling terintegrasi, sebagaimana ditunjukkan pada Gambar 1.

3.1. Pengumpulan Data

Dataset diperoleh dari platform Kaggle, dengan nama *dataset* Deteksi Berita Hoaks Indonesia Dataset (https://www.kaggle.com/datasets/mochamadabdulazis/deteksi-berita-hoaks-indo-dataset). *Dataset* berisi berita berbahasa Indonesia yang telah diberi label menjadi dua kelas, yaitu hoaks dan fakta. Data hoaks bersumber dari situs TurnBackHoax, sedangkan data fakta berasal dari media berita kredibel yaitu Antara, Kompas, dan Detik.

Distribusi data sebelum proses pembersihan ditunjukkan pada Tabel 1.

Tabel 1. Distribusi data awal

Kategori	Jumlah Data
Hoaks (TurnBackHoax)	12.746
Fakta (Antara, Kompas, Detik)	11.197

Setelah dilakukan proses pembersihan data (*data cleaning*), total data yang digunakan dalam penelitian ini berjumlah 21.373 data, dengan distribusi label seperti ditunjukkan pada Tabel 2.

Tabel 2. Distribusi data setelah cleaning

Label	Keterangan	Jumlah
0	Fakta	10.911
1	Hoaks	10.462

3.2. Pra-pemrosesan Data

Pra-pemrosesan data dilakukan untuk membersihkan teks berita dari elemen yang tidak relevan serta memastikan kualitas data sebelum digunakan dalam proses pelatihan model. Mengingat model IndoBERT dan mBERT merupakan model berbasis *transformer* yang telah dilatih pada korpus teks mentah, proses pra-pemrosesan dilakukan secara minimal namun tetap menjaga keamanan linguistik (*linguistically safe cleaning*).

Tahapan pra-pemrosesan yang diterapkan dalam penelitian ini meliputi:

a. Case folding, yaitu mengubah seluruh teks menjadi huruf kecil.
b. Penghapusan anotasi tambahan, seperti teks yang berada di dalam tanda kurung siku.
c. Penghapusan URL, baik yang diawali dengan protokol HTTP maupun WWW.
d. Penyaringan karakter, dengan mempertahankan huruf alfabet, angka, spasi, serta tanda baca penting seperti tanda tanya dan tanda seru.

e. Normalisasi spasi, dengan menghapus spasi berlebih agar teks menjadi lebih konsisten.

Selain proses pembersihan teks, dilakukan pula tahapan seleksi data untuk meningkatkan kualitas *dataset*, yaitu:

a. Menghapus teks berita yang terlalu pendek (kurang dari 10 karakter).
b. Menghapus data duplikat berdasarkan kesamaan teks berita.

Pendekatan ini bertujuan untuk mengurangi noise, mencegah *data leakage*, serta meningkatkan stabilitas proses pelatihan model tanpa menghilangkan informasi semantik penting dalam teks.

3.3. Pembagian Data

Dataset yang telah melalui tahap pra-pemrosesan selanjutnya dibagi menjadi tiga subset, yaitu data latih (*training*), data validasi (*validation*), dan data uji (*testing*). Pembagian data dilakukan menggunakan metode *stratified split* agar distribusi kelas tetap seimbang.

Proporsi pembagian data yang digunakan adalah 80% untuk data latih, serta masing-masing 10% untuk data validasi dan data uji. Distribusi pembagian *dataset* ditunjukkan pada Tabel 3.

Tabel 3. Pembagian dataset

Subset	Jumlah Data
Data latih	17.098
Data validasi	2.137
Data uji	2.138

Pendekatan *stratified split* ini bertujuan untuk memastikan bahwa proses pelatihan dan evaluasi model dilakukan secara adil dan representatif terhadap distribusi kelas yang ada pada *dataset*. Data latih digunakan untuk melatih model, data validasi untuk memilih model terbaik selama proses pelatihan, dan data uji untuk mengevaluasi performa akhir model.

3.4. Tokenisasi Data

Tokenisasi merupakan proses konversi teks berita menjadi representasi numerik yang dapat diproses oleh model *transformer*. Penelitian ini menggunakan tokenizer bawaan dari masing-masing model, yaitu tokenizer IndoBERT dan tokenizer mBERT, yang berbasis *WordPiece*.

Pada tahap tokenisasi diterapkan *truncation* dengan panjang maksimum 96 token untuk menjaga efisiensi komputasi tanpa menghilangkan konteks utama teks. Proses *padding* dilakukan secara dinamis menggunakan DataCollatorWithPadding, sehingga panjang input dalam setiap *batch* disesuaikan secara otomatis.

Tokenisasi menghasilkan dua komponen utama, yaitu *input IDs* dan *attention mask*, yang selanjutnya digunakan sebagai masukan pada proses pelatihan dan evaluasi model.

3.5. Pelatihan Model

Pelatihan model dilakukan menggunakan pendekatan *supervised learning* dengan memanfaatkan dua model *pre-trained transformer*, yaitu IndoBERT dan Multilingual BERT (mBERT). Kedua model dilakukan proses *fine-tuning* untuk tugas klasifikasi biner deteksi berita hoaks berbahasa Indonesia. Sebelum proses pelatihan, dilakukan pengaturan *seed* secara menyeluruh pada pustaka Python, NumPy, dan PyTorch untuk memastikan keterulangan (*reproducibility*) hasil eksperimen.

Proses pelatihan diimplementasikan menggunakan *framework* PyTorch dan pustaka HuggingFace *transformers* dengan memanfaatkan kelas Trainer. Tokenisasi data dilakukan secara dinamis (*dynamic padding*) menggunakan DataCollatorWithPadding, sehingga panjang input pada setiap *batch* dapat disesuaikan secara otomatis dan efisien.

Pada penelitian ini diterapkan strategi partial *fine-tuning*, di mana enam lapisan encoder terbawah pada masing-masing model dibekukan (*layer freezing*), sedangkan lapisan di atasnya serta *classification head* dilatih ulang. Pendekatan ini bertujuan untuk mengurangi risiko *overfitting*, mempercepat proses pelatihan, serta mempertahankan representasi linguistik dasar yang telah dipelajari oleh model *pre-trained*. Parameter utama pelatihan model ditunjukkan pada Tabel 4.

Tabel 4. Parameter pelatihan model

Parameter	Nilai
Learning rate	1×10^{-5}
Batch size	16
Epoch maksimum	5
Optimizer	AdamW
Panjang maksimum token	96
Weight decay	0.05

Pemilihan model terbaik dilakukan secara otomatis berdasarkan nilai *validation loss* terendah. Selain itu, diterapkan mekanisme *early stopping* dengan *patience* sebesar satu epoch untuk menghentikan pelatihan secara dini apabila tidak terjadi peningkatan kinerja model pada data validasi. Pendekatan ini bertujuan untuk meningkatkan stabilitas pelatihan dan mencegah *overfitting*.

3.6. Evaluasi Model

Evaluasi performa model dilakukan menggunakan data uji (*test set*) yang tidak pernah dilibatkan selama proses pelatihan maupun validasi. Evaluasi bertujuan untuk mengukur kemampuan generalisasi model dalam mendeteksi berita hoaks.
Metrik evaluasi yang digunakan dalam penelitian ini meliputi:

a. *Accuracy*, untuk mengukur tingkat ketepatan prediksi secara keseluruhan.

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

b. *Precision*, untuk mengukur ketepatan prediksi kelas hoaks.

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

c. *Recall*, untuk mengukur kemampuan model dalam mendeteksi seluruh berita hoaks.

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

d. *F1-Score*, sebagai rata-rata harmonik antara *precision* dan *recall*.

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

Keterangan:

- TP (*True Positive*): jumlah berita hoaks yang terklasifikasi benar.
- TN (*True Negative*): jumlah berita fakta yang terklasifikasi benar.
- FP (*False Positive*): jumlah berita fakta yang salah diklasifikasi sebagai hoaks.
- FN (*False Negative*): jumlah berita hoaks yang salah diklasifikasi sebagai fakta. Hasil evaluasi digunakan untuk menilai efektivitas model kombinasi dibandingkan model tunggal, serta menentukan kelayakan model untuk diimplementasikan ke dalam sistem *web*.

Nilai *F1-Score* digunakan sebagai metrik utama dalam membandingkan performa model IndoBERT dan mBERT, karena mampu memberikan gambaran kinerja model secara lebih seimbang pada tugas klasifikasi biner.

3.7. Implementasi Web

Model dengan performa terbaik selanjutnya diimplementasikan ke dalam sistem berbasis *web* yang dapat digunakan oleh publik. Sistem ini memungkinkan pengguna untuk memasukkan teks berita dan memperoleh hasil klasifikasi secara langsung. Arsitektur sistem terdiri atas dua komponen utama:

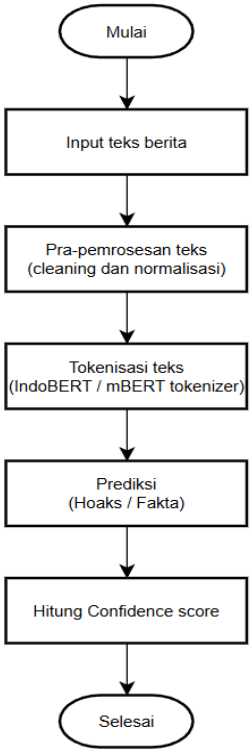
a. *Frontend* (Antarmuka Pengguna)
Dibangun menggunakan HTML, CSS, dan JavaScript untuk menerima input teks berita dan menampilkan hasil deteksi (hoaks atau valid).

b. *Backend* (Pemrosesan Pengguna)
Menggunakan bahasa pemrograman Python dengan *framework* Flask untuk menjalankan model IndoBERT dan mBERT, memproses teks input, dan mengembalikan hasil prediksi ke antarmuka pengguna.

Fitur utama pada sistem mencakup:

a. Input teks berita
b. Output klasifikasi berita (Hoaks atau Fakta)
c. *Confidence score* prediksi

Alur sistem dapat dilihat pada Gambar 2.

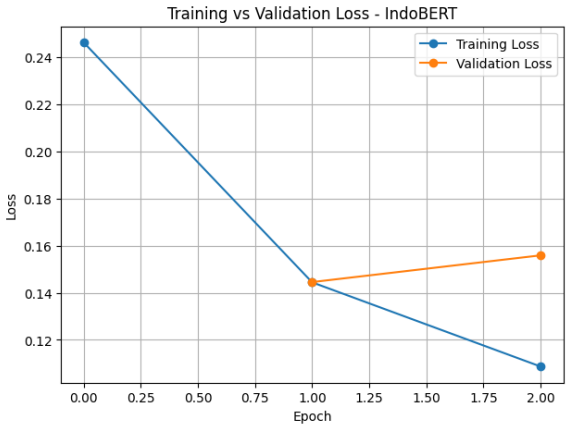


Gambar 2. Alur sistem web

4. HASIL DAN PEMBAHASAN

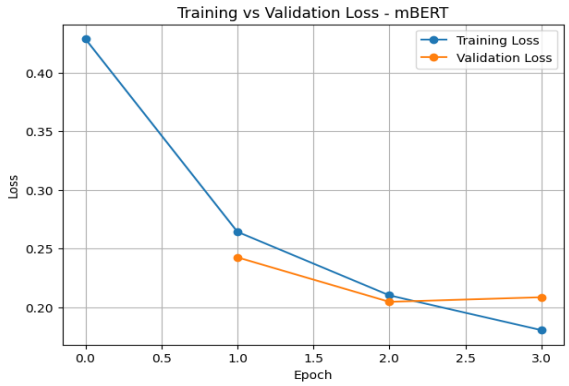
4.1. Hasil Pelatihan Model

Proses pelatihan model IndoBERT dan mBERT dilakukan menggunakan metode *fine-tuning* dengan mekanisme *early stopping* untuk mencegah terjadinya *overfitting*. Selama pelatihan, model dievaluasi pada setiap epoch menggunakan data validasi dengan mengamati perubahan nilai *training loss* dan *validation loss*. Berdasarkan hasil pelatihan, IndoBERT mencapai konvergensi lebih cepat dan dihentikan pada epoch ke-2, sedangkan mBERT mencapai titik optimal pada epoch ke-3. Perbedaan jumlah epoch optimal ini menunjukkan perbedaan kecepatan konvergensi antara model yang dilatih secara spesifik untuk bahasa Indonesia dan model multibahasa.



Gambar 3. Grafik training loss dan validation loss model IndoBERT

Berdasarkan grafik pada Gambar 3, dapat dilihat bahwa nilai *training loss* IndoBERT menurun secara signifikan dari epoch awal hingga epoch ke-2. Sementara itu, *loss* mencapai nilai terendah pada epoch ke-1 dan meningkat sedikit pada epoch ke-2. Pola ini menunjukkan bahwa model mulai mengalami *overfitting* ringan setelah epoch ke-1, sehingga pelatihan dihentikan pada epoch ke-2 menggunakan mekanisme *early stopping* untuk menjaga kemampuan generalisasi model.



Gambar 4. Grafik training loss dan validation loss model mBERT

Dari grafik pada Gambar 4, terlihat bahwa *training loss* mBERT menurun secara bertahap hingga epoch ke-3. Nilai *validation loss* mencapai titik minimum pada epoch ke-2 dan mengalami kenaikan pada epoch ke-3. Hal ini menunjukkan bahwa mBERT memiliki kecepatan konvergensi yang lebih lambat dibandingkan IndoBERT dan membutuhkan waktu pelatihan lebih lama untuk menyesuaikan diri dengan data berbahasa Indonesia.

4.2. Hasil Evaluasi

Evaluasi performa model dilakukan menggunakan data uji (*test set*) yang tidak digunakan dalam proses pelatihan. Metrik evaluasi yang digunakan meliputi *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur kemampuan model dalam melakukan klasifikasi hoaks dan fakta.

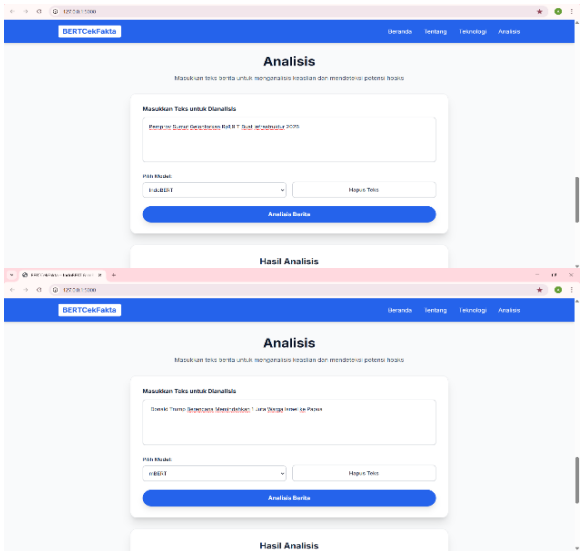
Tabel 5. Hasil evaluasi performa IndoBERT dan mBERT

Model	*Accuracy*	*Precision*	*Recall*	*F1-Score*
IndoBERT	0.945276	0.925824	0.965616	0.945302
mBERT	0.915809	0.891599	0.942693	0.916435

Berdasarkan Tabel 5, dapat dilihat bahwa model IndoBERT memperoleh performa terbaik pada seluruh metrik evaluasi dengan nilai akurasi sebesar 94,52% dan *F1-score* sebesar 94,53%. Sementara itu, mBERT mencapai nilai akurasi sebesar 91,58% dan *F1-score* sebesar 91,64%. Perbedaan ini menunjukkan bahwa IndoBERT lebih efektif dalam memahami konteks berita berbahasa Indonesia dibandingkan mBERT yang bersifat multibahasa.

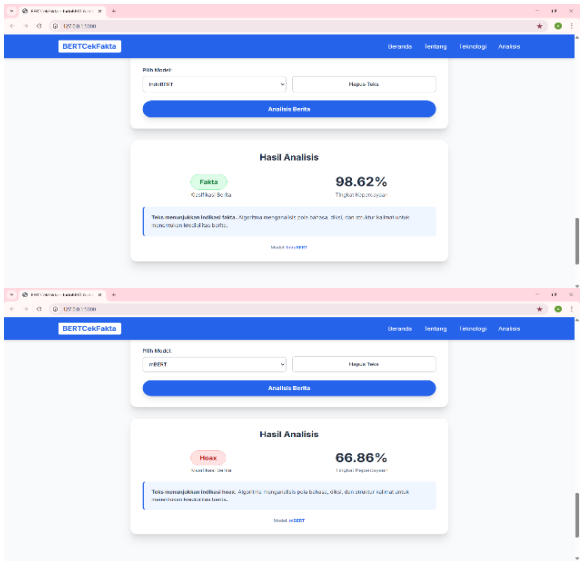
4.3. Implementasi Sistem Berbasis Web

Model IndoBERT dan mBERT yang telah dilatih diimplementasikan dalam sistem deteksi hoaks berbasis *web* untuk memudahkan pengguna dalam melakukan pengecekan informasi. Sistem dirancang agar pengguna dapat memilih model klasifikasi yang akan digunakan sebelum melakukan prediksi, sehingga memungkinkan perbandingan hasil deteksi antar model dalam satu platform.



Gambar 5. Tampilan analisis sistem web

Berdasarkan Gambar 5, sistem *web* menyediakan antarmuka yang memungkinkan pengguna melakukan analisis berita secara interaktif. Pengguna dapat memilih model klasifikasi yang akan digunakan, yaitu IndoBERT atau mBERT, melalui menu pilihan model sebelum melakukan prediksi. Fitur ini memberikan fleksibilitas kepada pengguna dalam menentukan model yang digunakan serta mendukung transparansi hasil prediksi.



Gambar 6. Contoh hasil prediksi hoaks/fakta pada sistem web

Pada Gambar 6, setelah teks berita dimasukkan, sistem menampilkan hasil analisis berupa label klasifikasi (fakta atau hoaks) serta confidence score (tingkat kepercayaan) prediksi dalam bentuk persentase. Tampilan hasil juga dilengkapi dengan penjelasan singkat yang membantu pengguna memahami alasan klasifikasi yang diberikan oleh model. Antarmuka ini dirancang sederhana dan responsif sehingga mudah digunakan oleh pengguna umum dalam melakukan pengecekan informasi secara mandiri.

4.4. Pembahasan

Berdasarkan seluruh hasil yang diperoleh, dapat disimpulkan bahwa IndoBERT menunjukkan performa yang lebih unggul dibandingkan mBERT dalam mendeteksi berita hoaks berbahasa Indonesia. Hal ini disebabkan oleh proses *pre-training* IndoBERT yang menggunakan korpus bahasa Indonesia, sehingga representasi semantik yang dihasilkan lebih sesuai dengan konteks lokal. Sebaliknya, mBERT memiliki keunggulan dalam generalisasi lintas bahasa, namun kurang optimal dalam menangkap nuansa spesifik bahasa Indonesia.

Kesalahan prediksi yang terjadi umumnya disebabkan oleh judul berita yang ambigu, sensasional, atau menggunakan bahasa persuasif yang menyerupai karakteristik berita hoaks. Temuan ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa klasifikasi hoaks berbasis judul berita masih memiliki tantangan pada teks yang bersifat subjektif dan provokatif. Implementasi kedua model dalam sistem berbasis *web* menunjukkan bahwa pendekatan ini dapat digunakan sebagai solusi praktis dalam membantu masyarakat memverifikasi informasi.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa model IndoBERT dan mBERT mampu melakukan klasifikasi berita hoaks dan fakta dengan baik. IndoBERT menunjukkan kinerja yang lebih optimal dalam konteks bahasa Indonesia karena dilatih secara khusus pada korpus berbahasa Indonesia, sehingga lebih mampu menangkap konteks semantik dan karakteristik linguistik lokal. Sementara itu, mBERT memiliki keunggulan dalam fleksibilitas multibahasa, namun performanya sedikit di bawah IndoBERT untuk kasus deteksi hoaks berbahasa Indonesia. Implementasi sistem berbasis *web* membuktikan bahwa model dapat digunakan secara praktis untuk membantu pengguna memverifikasi informasi secara cepat dan efisien. Penelitian selanjutnya disarankan untuk memperluas *dataset*, menguji model *transformer* terbaru, serta mengembangkan deteksi hoaks multimodal (teks, gambar, dan video).

DAFTAR PUSTAKA

[1] C. J. L. Tobing, IGN Lanang Wijayakusuma, and Luh Putu Ida Harini, "Perbandingan Kinerja IndoBERT dan MBERT untuk Deteksi Berita Hoaks Politik dalam Bahasa Indonesia", *j. sains. teknologi.*, vol. 14, no. 1, pp. 114–123, Apr. 2025.

[2] I. Bagaskara, E. Purwanto, dan J. Maulindar, "Sistem Cerdas Deteksi Berita Hoaks Berbasis IndoBERT dengan Antarmuka Web," *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB)*, pp. 197-206, Jul. 2025.

[3] V. Prisscilya and A. S. . Girsang, "Classification of Indonesia False News Detection Using Bertopic and Indobert", *jist*, vol. 5, no. 8, pp. 3061–3079, Aug. 2024.

[4] A. M. Kamal, Y. H. Chrisnanto, and R. Yuniarti, "Identifikasi Berita Palsu di Portal Media Online Menggunakan Model IndoBERT dan LSTM", *Jur. Ris. Kom.*, vol. 12, no. 3, pp. 287–297, Jun. 2025.

[5] M. Y. Ridho dan E. Yulianti, "From Text to Truth: Leveraging IndoBERT and Machine Learning Models for Hoax Detection in Indonesian News," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, vol. 10, no. 3, pp. 544-555, Sep. 2024.

[6] M. A. Fathin, Y. Sibaroni, dan S. S. Prasetyowati, "Handling Imbalance Dataset on Hoax Indonesian Political News Classification Using IndoBERT and Random Sampling," *Jurnal Media Informatika Budidarma*, vol. 8, no. 1, pp. 352-360, Jan. 2024.

[7] A. Sarjito, "Hoaks, Disinformasi, dan Ketahanan Nasional: Ancaman di Era Digital," *Journal of Governance and Local Politics (JGLP)*, vol. 6, no. 2, pp. 175-185, Nov. 2024.

[8] N. Nuryanto, "Deteksi Berita Hoaks Berbahasa Indonesia Menggunakan Natural Language Processing (NLP) dengan Model IndoBERT dan Implementasi Berbasis Web," *Prosiding Seminar Nasional Sistem Informasi dan Teknik Informatika*, pp.134-140, Sep. 2025.

[9] A. Fardhina, "Sistem Deteksi Berita Hoaks Berbasis Algoritma Natural Language Processing (NLP) Menggunakan BERT," *Jurnal Manajemen Informatika, Sistem Informasi dan Teknologi Komputer (JUMISTIK)*, vol. 4, no. 1, pp. 450-461, Jun. 2025.

[10] A. R. Hakim and A. H. Muhammad, "PERBANDINGAN MODEL TRANSFORMER, DEEP LEARNING, DAN MACHINE LEARNING UNTUK DETEKSI BERITA PALSU: STUDI KASUS PADA TEKS BERBAHASA INDONESIA," *Jurnal Manajemen Informatika dan Sistem Informasi (MISI)*, vol. 8, no. 2, pp. 188–197, Jun. 2025.

[11] W. Widodo and M. Nugraheni, "Indonesian Fake News Classification Using Transfer Learning in CNN and LSTM," *JOIV: International Journal*

on Informatics Visualization, vol. 8, no. 3, pp. 1213–1221, Sep. 2024.

[12] R. A. Pratama and I. K. G. Darma Putra, "Klasifikasi Teks Berita Menggunakan Bidirectional Encoder Representations from Transformers (BERT)," *Jurnal Ilmu Komputer dan Informasi*, vol. 13, no. 2, 2022.

[13] I. Akbar, F. Supriadi, and D. I. Junaedi, "Pemanfaatan Machine Learning di Bidang Kesehatan," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 1, pp. 1744-1749, Feb. 2025.

[14] L. Titania, "Penerapan Aplikasi Fuzzy Logic dalam Pemilihan Minuman Kaleng," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 3, pp. 4749-4755, Jun. 2025.

[15] T. Listyorini, Muljono, Purwanto and R. S. Basuki, "Neural Approaches for Indonesian–Javanese Translation: A Comparison of RNN and Transformer Models," *2025 International Seminar on Application for Technology of Information and Communication (iSemantic)*, pp. 575-584, 2025.