

Detection of Political Hoax News Using Fine-Tuning IndoBERT

Charlotte Jocelynne L Tobing ¹, IGN Lanang Wijayakusuma ², Luh Putu Ida Harini ³

Matematika, Universitas Udayana

charlottejocelynne@gmail.com ¹, lanang_wijaya@unud.ac.id ², ballidah@unud.ac.id ³

Article Info

Article history:

Received 2024-12-02

Revised 2024-12-09

Accepted 2025-01-21

Keyword:

*Hoax
News Detection,
IndoBERT,
Political Hoax News,
Machine Learning,
Indonesian News*

ABSTRACT

Indonesia has experienced a surge in the spread of political hoax news, posing a potential threat to democratic and social stability. This study aims to develop a model for detecting political hoax news in the Indonesian language using IndoBERT, a language model optimized for Indonesian text. The dataset was sourced from Kaggle and comprises 20,928 factual news articles and 2,251 hoax news articles from major Indonesian media outlets, including CNN, Kompas, Tempo, and Turnbackhoax. The imbalance between factual and hoax news articles was addressed through undersampling, resulting in 1,302 samples for each class. The research stages include data collection, preprocessing, IndoBERT model training, and model evaluation. Results indicate that fine-tuning IndoBERT can detect political hoax news with an accuracy of 94.1% and an ROC AUC of 0.991, demonstrating high performance in accuracy and generalization capability. This research is expected to contribute to minimizing the spread of political hoax news in Indonesia and enhance media literacy among the public.



This is an open access article under the [CC-BY-SA](#) license.

I. PENDAHULUAN

Sebagai negara dengan salah satu tingkat pertumbuhan pengguna internet tertinggi di dunia, Indonesia mengalami peningkatan yang cepat dalam pemakaian media sosial. Menurut hasil penelitian Masyarakat Telematika Indonesia, ditemukan bahwa 44,3% dari responden mengakui sering menerima berita bohong atau misinformasi setiap harinya [1]. Selama pemilihan presiden di Indonesia dari tahun 2017 hingga 2019, berita bohong dan kampanye disinformasi marak terjadi, yang menyebabkan meningkatnya polarisasi politik di negara ini [2]. Penyebaran berita bohong sangat cepat di dunia maya karena masyarakat Indonesia menyukai hal-hal yang viral dan membuat mereka penasaran [3]. Selain itu, disinformasi dalam arti luas mengacu pada penyebaran informasi yang tidak benar dan dalam arti khusus adalah penyebaran informasi legit yang dikemas dengan cara tertentu untuk mengaburkan fondasi kontekstual awal informasi tersebut menjadi kontekstual baru yang didasarkan pada kepentingan politik penyebar informasi[1].

Beberapa isu terkait berita hoaks politik di Indonesia antara lain dampaknya yang dapat mengganggu proses demokrasi dan pemilihan umum yang adil, fungsinya untuk menjelaskan lawan politik, serta penyebarannya yang cepat di

media sosial sehingga sulit dicegah. Karena itu, sebuah sistem yang efektif diperlukan untuk mendeteksi dan memberantas berita hoaks politik. Salah satu langkah untuk mengurangi penyebaran informasi yang menyesatkan adalah dengan melakukan deteksi untuk memastikan keaslian berita, sehingga pembaca mendapatkan penjelasan yang jelas.

Pada beberapa tahun belakangan ini, telah dilakukan banyak penelitian mengenai deteksi berita hoaks yang dimotivasi oleh kombinasi antara kemajuan teknologi dan meningkatnya kesadaran akan dampak negatif dari berita yang menyesatkan [4]. Kondisi yang selalu berubah ini membuat tantangan baru bagi para pencipta berita bohong untuk menyempurnakan cara mereka dalam mengelabui mekanisme deteksi [5]. Mulai dari konten buatan AI yang canggih hingga gambar atau video yang dimanipulasi dengan cermat, sarana pelaku hoaks semakin beragam, sehingga dibutuhkan teknik deteksi yang lebih canggih dan fleksibel. Tentunya penelitian-penelitian tersebut sangat membantu pembaca dalam menghindari berita hoaks.

Penelitian sebelumnya telah melakukan beberapa penelitian untuk mendeteksi berita hoaks. Pada tahun 2022, penelitian mengimplementasikan Algoritma Naive Bayes Classifier untuk membangun sistem deteksi berita hoaks Covid-19 di situs kumparan dan mampu mencapai akurasi

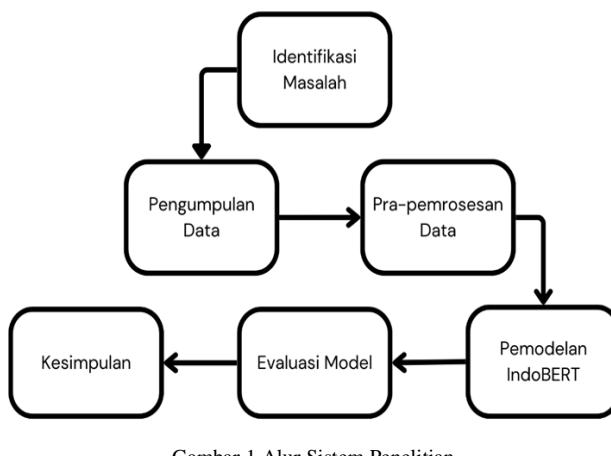
81% [6]. Penelitian lain pada tahun 2020 membahas deteksi berita hoaks menggunakan metode Random Forest dan Logistic Regression yang mencapai akurasi 84% dan 77% secara berurutan [7]. Penelitian lain melakukan deteksi berita hoax Covid-19 di Indonesia dengan metode hibrida LSTM dan SVM yang menghasilkan akurasi 94% [8]. Penelitian lain mengusulkan perbandingan model LSTM dan GRU yang mencapai akurasi terbaik 73% dengan model LSTM [9].

Penelitian internasional mengusulkan kombinasi arsitektur CNN dan RNN dengan pemrosesan teks menggunakan word embedding GloVe, berhasil mencapai akurasi prediksi berita hoaks sebesar 97,21% [10], kemudian ada penelitian lain yang mengusulkan model hibrida CNN-BiLSTM dengan akurasi 97,5% [11]. Ada penelitian yang mengusulkan penggunaan transfer learning pada BERT yang telah di fine-tune mencapai akurasi 97,02% [12]. Kemudian pada penelitian lain mengatakan bahwa kombinasi model BERT-LSTM mencapai akurasi 91,09% dan kombinasi model BERT-CNN mencapai akurasi 79,02% [13]. Kemudian penelitian lain menggunakan Metode FakeBERT yang berbasis BERT dikombinasikan dengan blok paralel dari single-layer CNN dengan ukuran kernel dan filter yang berbeda berhasil mencapai akurasi 98,90% dalam mendekripsi berita hoaks [14]. Penelitian tahun 2024 menguji model hibrida GBERT yang mampu mendekripsi berita hoaks dengan akurasi 95,3% [15].

Kemampuan model BERT efektif untuk mendekripsi berita hoaks berbagai bahasa namun kurang optimal untuk bahasa Indonesia. Berdasarkan permasalahan dan solusi yang telah dijelaskan, penelitian ini melakukan pengelompokan berita hoaks dalam bahasa Indonesia dengan menerapkan Fine-Tune pada model IndoBERT. Penggunaan IndoBERT yang di pre-training khusus untuk bahasa Indonesia dianggap paling relevan untuk meningkatkan akurasi deteksi berita hoaks politik di Indonesia.

II. METODE

Pada penelitian ini, diusulkan mesin pembelajaran IndoBERT dalam melakukan deteksi berita politik hoaks di Indonesia.



A. Pengumpulan Data

Dataset diperoleh dari (<https://www.kaggle.com/datasets/linkgish/indonesian-fact-and-hoax-political-news>) pada situs kaggle, yang berisikan data scrapping dari 4 situs berita yaitu CNN, Kompas, Tempo, dan Turnbackhoax. Setiap data berita sudah diberi label sesuai tipe beritanya. Jumlah dataset yang didapatkan adalah 20928 data berita fakta dan 2251 data berita fakta.

B. Pra-pemrosesan Data

Pada tahap pra-pemrosesan data, data yang sudah didapatkan akan diproses dan dipersiapkan untuk digunakan dalam model. Dalam pra-pemrosesan data dilakukan pembersihan data yang mencakup konversi teks menjadi huruf kecil (lowercase), menghapus tanda baca dan symbol khusus, membersihkan data karakter non-alfabet, menghapus teks kosong, menghapus URL dan tautan, serta menghapus data yang memiliki nilai kosong (NaN).

Untuk mengatasi ketidakseimbangan data pada dataset, penelitian ini menggunakan dua teknik: undersampling dan class weighting. Ketidakseimbangan ini dapat menyebabkan model lebih cenderung memprediksi kelas mayoritas (fakta), sehingga mengurangi kemampuan untuk mengenali pola berita hoaks. Teknik pertama, undersampling, dilakukan dengan mengurangi jumlah data pada kelas mayoritas (fakta) hingga sebanding dengan kelas minoritas (hoaks). Teknik ini dipilih untuk memastikan model dapat mempelajari pola secara merata antara kedua kelas, meskipun dengan risiko kehilangan informasi dari data yang dihapus. Pendekatan kedua, class weighting, memberikan bobot lebih besar pada kelas minoritas (hoaks) dan bobot lebih kecil pada kelas mayoritas (fakta) selama proses pelatihan. Teknik ini memungkinkan seluruh dataset digunakan tanpa perlu mengurangi data dari kelas mayoritas. Kemudian, dataset dibagi menjadi 3 bagian yaitu data training, data validation, dan data test. Distribusi pembagian dataset ditunjukkan pada gambar 2.



Gambar 2. Pembagian Data

Setelah dilakukan proses-proses tersebut, kemudian proses tokenisasi dilakukan untuk memecah teks menjadi token yang lebih kecil. Semua langkah ini bertujuan untuk memaksimalkan kinerja model IndoBERT yang akan digunakan untuk deteksi hoaks.

C. IndoBERT

BERT adalah model berbasis transformer yang efektif dalam menangkap sintaksis dan semantik kalimat melalui mekanisme attention penuh, memungkinkan model ini memahami konteks kata dalam dua arah. IndoBERT, sebagai

adaptasi BERT untuk bahasa Indonesia, dilatih secara khusus menggunakan data dari Wikipedia Bahasa Indonesia, artikel berita dari Kompas, Tempo, dan Liputan6, serta korpus web Indonesia. Dengan 12 lapisan tersembunyi dan 12 kepala attention, IndoBERT telah menunjukkan performa tinggi dalam berbagai tugas NLP berbahasa Indonesia[16].

D. Pemodelan IndoBERT

Pada penelitian ini, model IndoBERT digunakan untuk melakukan klasifikasi berita politik sebagai upaya mendeteksi berita hoaks di Indonesia. Pemodelan dilakukan dengan menggunakan transformer-based model IndoBERT, yang telah di-fine-tune pada dataset berita yang telah ditokenisasi. Proses pelatihan dilakukan dengan mengatur agar layer IndoBERT ke-9 hingga layer terakhir unfreeze (tidak dibekukan), sementara layer ke-8 ke bawah freeze (dibekukan). Pendekatan ini bertujuan untuk menjaga informasi dasar dalam embedding kata yang telah dilatih, sekaligus mengizinkan lapisan atas untuk belajar pola yang lebih spesifik terkait berita politik hoaks. Teknik freezing dan unfreezing ini efektif untuk menghemat sumber daya komputasi, sambil tetap mempertahankan performa tinggi dalam klasifikasi [17].

Selama proses pelatihan, argumen TrainingArguments digunakan untuk mengatur strategi logging dan evaluasi yang dilakukan di setiap epoch. Selain itu, opsi save_best_model diaktifkan agar model dengan performa terbaik otomatis disimpan di akhir pelatihan, yang memungkinkan pemilihan model terbaik berdasarkan metrik evaluasi (misalnya, F1-score).

Untuk melatih model, Trainer dari pustaka transformers digunakan, dengan konfigurasi dataset pelatihan (train) dan dataset evaluasi (test). Data tersebut telah dipersiapkan menggunakan tokenisasi berbasis IndoBERT. Untuk melatih model, Trainer dari pustaka transformers digunakan, dengan konfigurasi dataset pelatihan (train) dan dataset evaluasi (test). Data tersebut telah dipersiapkan menggunakan tokenisasi berbasis IndoBERT.

E. Evaluasi Model

Pada tahap ini, model dievaluasi untuk mengukur kinerjanya dalam mendeteksi berita hoaks dan fakta. Proses evaluasi ini dilakukan pada data validasi untuk memberikan gambaran tentang performa model sebelum diujikan pada data yang benar-benar baru. Dengan evaluasi ini, performa model dapat direfleksikan lebih akurat tanpa bias dari data yang digunakan untuk pelatihan atau penyesuaian model. Metrik evaluasi yang digunakan mencakup:

1. *Confusion Matrix*: Matriks ini memberikan total prediksi yang benar dan salah pada masing-masing kelas (fakta dan hoaks). Matriks ini memberikan gambaran tentang jumlah True Positives (TP), False Positives (FP), True Negatives (TN), dan False Negatives (FN), yang merupakan dasar untuk menghitung metrik evaluasi lainnya. Confusion matrix membantu dalam memahami distribusi kesalahan dan bagaimana model menangani setiap kelas.

2. *Classification Report*: Classification report memberikan metrik seperti presisi, recall, F1-score, dan akurasi untuk setiap kelas.

- Presisi menunjukkan seberapa banyak prediksi "hoaks" yang benar-benar merupakan hoaks, sehingga membantu menilai keakuratan model dalam menangani prediksi positif.
- Recall menunjukkan seberapa baik model menangkap semua sampel "hoaks" yang sebenarnya, yang penting dalam mendeteksi sebanyak mungkin berita hoaks.
- F1-score adalah gabungan presisi dan recall yang memberikan pandangan keseimbangan antara keduanya, terutama berguna pada dataset yang tidak seimbang.

3. *ROC AUC Score*: Area under the ROC curve (ROC AUC) adalah metrik yang mengukur kemampuan model dalam memisahkan kelas positif dan negatif pada berbagai ambang batas. Nilai AUC yang mendekati 1 menunjukkan bahwa model memiliki performa yang baik dalam membedakan antara kelas fakta dan hoaks. ROC AUC digunakan karena mampu menggambarkan performa model secara lebih umum tanpa bergantung pada ambang batas tertentu.

Dengan metrik evaluasi ini, kita dapat menilai performa keseluruhan model dan menganalisis kekuatannya dalam mendeteksi berita hoaks serta kelemahannya dalam kesalahan prediksi. Evaluasi ini akan menjadi dasar dalam tahap pembahasan untuk interpretasi lebih lanjut mengenai efektivitas model dan potensinya untuk diterapkan pada kasus nyata di Indonesia.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari situs Kaggle, tepatnya dari kumpulan data berjudul Indonesian Fact and Hoax Political News. Dataset ini berisi berita yang diperoleh melalui scraping dari empat sumber berita utama di Indonesia: CNN, Kompas, Tempo, dan Turnbackhoax. Setiap berita telah dilabeli sebagai "fakta" atau "hoaks" sesuai tipe beritanya.

Jumlah data awal terdiri dari 20.928 berita fakta dan 2.251 berita hoaks. Ketidakseimbangan ini memerlukan perhatian khusus karena model dapat lebih cenderung memprediksi kelas mayoritas (fakta) dibandingkan kelas minoritas (hoaks). Oleh karena itu, setelah melalui tahap cleaning, dilakukan teknik undersampling pada kelas mayoritas sehingga dataset menjadi seimbang, dengan 1.302 data pada setiap kelas (fakta dan hoaks).

Pada tahap cleaning, selain normalisasi teks, penghapusan tanda baca, stopword, dan karakter tidak relevan, data judul dan teks berita digabungkan untuk menyediakan konteks yang lebih lengkap bagi model. Selain itu, penanda di awal judul berita hoaks, seperti "[HOAKS]" dan "[SALAH]", dihapus agar model tidak bergantung pada tanda ini dalam mendeteksi

hoaks. Dengan data yang telah dibersihkan dan seimbang, model diharapkan dapat mempelajari pola kedua kelas secara adil dan meningkatkan akurasi dalam mendeteksi berita hoaks dan fakta.

B. Pra-pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk menyiapkan data yang bersih dan seimbang sebagai input bagi model IndoBERT. Berikut tahapan pra-pemrosesan data yang dilakukan:

1. *Pembersihan Data*: Data terlebih dahulu dibersihkan untuk memastikan kualitas teks dan keseragaman format data. Tahapan pembersihan ini meliputi:

- Konversi teks menjadi huruf kecil (lowercasing) untuk menghindari perbedaan akibat huruf besar dan kecil.
- Penghapusan tanda baca dan simbol khusus yang tidak relevan bagi pemahaman teks.
- Penghapusan karakter non-alfabet untuk mempertahankan hanya huruf dan kata yang relevan.
- Penghapusan teks kosong dan data yang memiliki nilai NaN untuk menghindari data yang tidak valid.
- Penghapusan URL atau tautan yang mungkin ada dalam teks, karena informasi ini tidak relevan untuk analisis konten berita.

2. *Pembagian Dataset*: Dataset yang telah dibersihkan dan seimbang kemudian dibagi menjadi tiga bagian:

- Data train (70% dari total data), berjumlah 1.822 data, digunakan untuk melatih model.
- Data validasi (15% dari total data), berjumlah 390 data, digunakan untuk mengukur performa model selama pelatihan dan mencegah overfitting.
- Data test (15% dari total data), berjumlah 392 data, yang akan digunakan sebagai data uji akhir untuk mengevaluasi model setelah pelatihan selesai.

3. *Tokenisasi*: Proses tokenisasi dilakukan untuk mempersiapkan teks dalam format yang dipahami oleh model IndoBERT dari Hugging Face. Tokenisasi ini mengubah teks menjadi token numerik dengan mempertahankan struktur kata dan sub-kata yang khas dalam bahasa Indonesia. Setiap teks ditokenisasi dengan pengaturan berikut:

- Truncation hingga batas 512 token untuk membatasi panjang input yang terlalu besar.
- Padding hingga panjang maksimum agar setiap input memiliki panjang yang konsisten dalam setiap batch pelatihan.

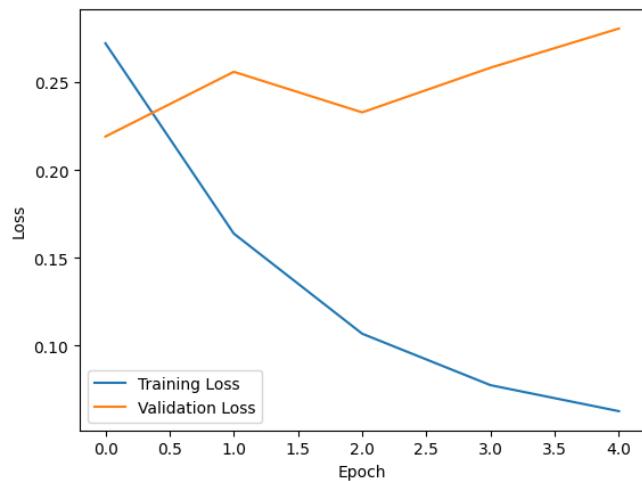
Tahapan ini memastikan data bersih, konsisten, dan siap untuk dilatih pada model IndoBERT, sehingga diharapkan dapat meningkatkan performa model dalam mendeteksi berita hoaks dan fakta.

C. Pemodelan IndoBERT

Pemodelan dilakukan dengan menggunakan transformer-based model IndoBERT yang telah di-finetune pada dataset.

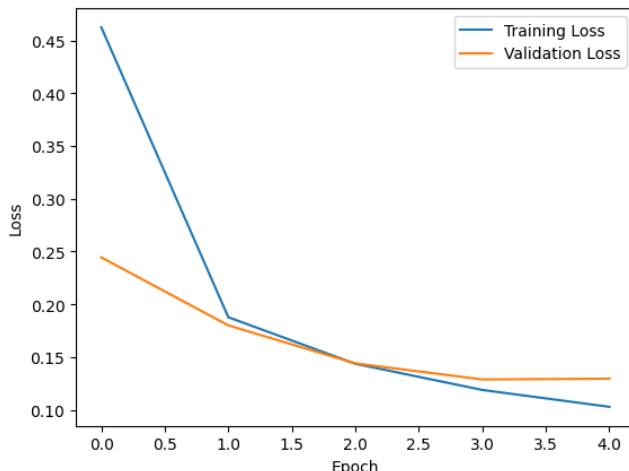
Dalam pemodelan ini diawali dengan mengatur hyperparameter yang akan digunakan. Hyperparameter tersebut mencakup learning rate sebesar 1e-5, batch size sebanyak 32, dan jumlah epoch sebanyak 5, dan beberapa lapisan awal model IndoBERT dibekukan hingga lapisan ke-8 untuk mencegah overfitting. Selain itu, weight decay sebesar 0.1 digunakan untuk mencegah overfitting pada model. Selama proses pelatihan, argumen TrainingArguments mengatur strategi logging dan evaluasi untuk dilakukan di setiap epoch, dengan opsi untuk menyimpan model terbaik di akhir pelatihan. Dataset train digunakan selama pelatihan model untuk mempelajari pola yang membedakan berita fakta dan hoaks dengan melakukan penyesuaian bobot berdasarkan Training Loss. Dataset validation, yang dievaluasi pada akhir setiap epoch, digunakan untuk memantau performa model selama pelatihan, membantu mendeteksi potensi masalah seperti overfitting dan memastikan model tidak hanya menghafal data train.

Dalam penelitian ini, dilakukan perbandingan antara dua teknik untuk menangani ketidakseimbangan data, yaitu undersampling dan class weighting. Hasil pelatihan dari metode yang memakai data dengan teknik undersampling dan class weightring dapat dilihat pada plot di Gambar 3 dan Gambar 4.



Gambar 3. Riwayat Training Data Dengan Teknik Weighting Class

Pada Gambar 3 terlihat bahwa model yang menggunakan teknik class weighting menunjukkan indikasi overfitting. Hal ini diduga terjadi karena ketidakseimbangan data yang sangat ekstrem, di mana jumlah sampel berita fakta jauh lebih besar dibandingkan berita hoaks. Ketidakseimbangan ini mengakibatkan model kesulitan mempelajari pola secara optimal, terutama pada kelas minoritas, sehingga performa pada data validasi menurun meskipun akurasi pada data pelatihan tinggi.



Gambar 4. Riwayat Training Data Dengan Teknik Undersampling

Pada Gambar 4, terlihat bahwa model yang dilatih menggunakan teknik undersampling menunjukkan performa yang lebih stabil dan memuaskan dibandingkan dengan model yang menggunakan class weighting. Pada awal pelatihan (epoch pertama), training loss sebesar 0,4625 dan validation loss sebesar 0,2444 menunjukkan bahwa model sudah memiliki kemampuan awal yang baik untuk mempelajari pola data. Dengan akurasi sebesar 89,5% dan AUC 0,975, model menunjukkan kinerja awal yang menjanjikan dalam membedakan berita fakta dan hoaks di domain politik.

Seiring berjalannya pelatihan, training loss terus menurun secara signifikan hingga mencapai 0,1028 pada epoch kelima, mencerminkan bahwa model dapat belajar pola secara efektif dari data pelatihan. Penurunan validation loss hingga titik terendah pada epoch keempat (0,1287) menunjukkan bahwa model dapat mempelajari pola yang relevan dari data validasi.

Akurasi (95,2%) dan AUC (0,991) yang tetap tinggi pada epoch kelima mengindikasikan bahwa model mampu menjaga performa generalisasi yang baik. Nilai AUC yang tinggi khususnya menguatkan bahwa model dapat membedakan antara kelas berita fakta dan berita hoaks secara konsisten, bahkan pada data validasi yang tidak terlihat selama pelatihan.

Keunggulan teknik undersampling dibandingkan class weighting dalam konteks ini dapat dijelaskan melalui distribusi data yang lebih seimbang setelah undersampling. Dengan mengurangi dominasi kelas mayoritas, model dapat lebih fokus mempelajari pola yang relevan dari kedua kelas. Hal ini berbeda dengan teknik class weighting, yang meskipun memberikan bobot lebih besar pada kelas minoritas, sering kali masih terpengaruh oleh jumlah sampel kelas mayoritas yang jauh lebih besar, sehingga meningkatkan risiko overfitting pada kelas tertentu.

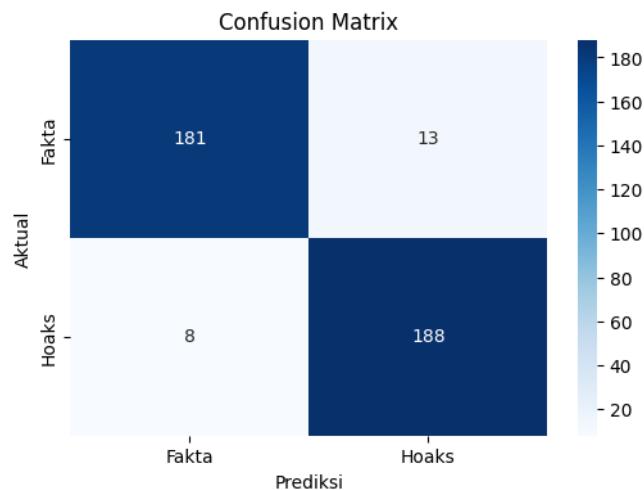
Berdasarkan hasil ini, teknik undersampling dipilih untuk evaluasi akhir pada dataset uji. Pilihan ini didasarkan pada konsistensi performa model dalam menjaga keseimbangan antara training loss dan validation loss, serta akurasi dan AUC yang tinggi pada data validasi. Teknik ini terbukti lebih efektif

dalam menangani ketidakseimbangan data yang ekstrem, memberikan model kemampuan yang lebih baik untuk generalisasi pada data baru.

D. Evaluasi Model Pada Data Uji

Pada tahap prapemrosesan data sebelumnya dataset sudah dipisah menjadi tiga subset yaitu train, test, dan validation. Pembagian dataset sejak awal bertujuan agar evaluasi pada data uji benar-benar merefleksikan performa model dalam kondisi sebenarnya, tanpa bias dari data yang digunakan untuk melatih atau menyesuaikan model selama pelatihan. Dalam evaluasi model ini dilakukan menggunakan dataset validation yang berisikan 392 teks berita. Model yang telah dilatih pada tahap sebelumnya digunakan untuk mengevaluasi dan menguji performanya pada dataset validation ini. Model menghasilkan hasil yang memuaskan dengan accuracy 94,1% dan AUC 0,991 pada data uji, menunjukkan bahwa model mampu mengklasifikasikan berita politik secara akurat dan efektif. Ada beberapa metrik evaluasi utama yaitu:

1. *Confusion Matrix*: Pada confusion matrix hasil evaluasi, model memiliki performa yang baik dalam mendeteksi berita hoaks dan faktual. True Positives (TP) adalah kasus ketika model berhasil memprediksi berita Hoaks secara benar, yang berjumlah 181 contoh. Sebaliknya, True Negatives (TN) adalah kasus ketika model memprediksi berita Fakta secara tepat, yang berjumlah 188 contoh seperti yang ditunjukkan pada Gambar 5.

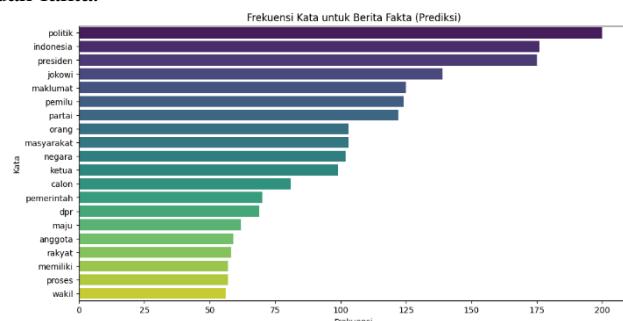


Gambar 5. Confusion Matrix

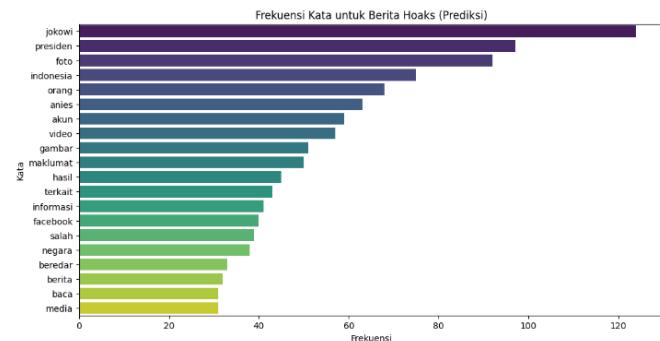
2. *Classification Report*: Classification Report menunjukkan precision, recall, dan f1-score yang tinggi untuk kedua kelas. Nilai precision dan recall untuk kelas "Fakta" dan "Hoaks" masing-masing berada diatas 0.94 yang menunjukkan kemampuan model dalam mengklasifikasikan berita dengan tingkat kesalahan rendah. Nilai F1-score yaitu 0.95 untuk kedua kelas menunjukkan keseimbangan yang baik antara precision dan recall, mengindikasikan bahwa model ini andal untuk mendeteksi baik fakta maupun hoaks.

3. ROC AUC Score: Model mencapai ROC AUC score sebesar 0.946 pada data uji, menunjukkan kemampuan model untuk membedakan antara kelas "Fakta" dan "Hoaks" dengan baik, yang penting untuk aplikasi deteksi berita hoaks.

Secara keseluruhan, model IndoBERT menunjukkan performa yang kuat dalam klasifikasi berita hoaks politik pada data uji. Gambar 6 dan Gambar 7 adalah grafik yang memperlihatkan kata-kata yang sering muncul di berita hoaks dan fakta.



Gambar 6. Frekuensi Kata untuk Berita Fakta



Gambar 7. Frekuensi Kata untuk Berita Hoaks

Grafik analisis kata pada berita fakta pada gambar 5 menunjukkan bahwa kata-kata seperti "politik," "Indonesia," "presiden," dan "pemilu" sering muncul, mencerminkan fokus berita pada isu-isu politik nasional yang penting dan relevan bagi masyarakat. Kehadiran kata-kata seperti "pemerintah," "negara," dan "masyarakat" menunjukkan bahwa berita fakta cenderung membahas topik yang berdampak luas dan signifikan bagi publik, serta mengedukasi pembaca tentang proses dan peran dalam struktur pemerintahan. Berita fakta umumnya memberikan informasi yang terverifikasi, dengan konteks yang membantu masyarakat memahami perkembangan politik dan sosial yang terjadi.

Pada gambar 6, terlihat bahwa kata-kata seperti "Jokowi" dan "presiden" mendominasi, menunjukkan bahwa berita hoaks sering kali menjadikan tokoh-tokoh politik utama sebagai subjek utama. Hal ini bisa jadi karena topik politik dan sosok pemimpin seperti presiden cenderung menarik perhatian dan memicu emosi pembaca, sehingga lebih mudah menyebar. Selain itu, kata-kata seperti "foto," "video," "akun," dan "gambar" muncul dengan frekuensi tinggi, menandakan bahwa banyak berita hoaks memanfaatkan media visual untuk memperkuat atau menambahkan

kredibilitas pada klaim hoaks. Media visual ini sering kali berupa foto atau video yang sudah dimanipulasi atau diambil di luar konteks untuk memberikan kesan yang salah. Kata-kata lain seperti "maklumat," "hasil," dan "terkait" juga menunjukkan bahwa berita hoaks kerap kali mencoba menciptakan kesan seolah-olah ada informasi resmi atau hasil tertentu yang mendukung narasi yang disampaikan.

Model IndoBERT yang digunakan untuk mendeteksi berita palsu politik pada penelitian ini memiliki keterbatasan yang perlu dipertimbangkan. Salah satu kendala utama adalah adanya bias pada dataset, seperti distribusi data yang tidak seimbang atau representasi yang terlalu dominan dari sumber tertentu, sehingga dapat memengaruhi kemampuan model dalam mengenali pola pada berita dari sumber lain. Selain itu, model ini membutuhkan dataset yang lebih besar dan beragam, mengingat domain politik memiliki narasi yang kompleks dan dinamis yang sering kali melibatkan isu lintas sektor. Kelemahan lainnya adalah IndoBERT belum sepenuhnya optimal untuk memahami konteks rumit dalam berita politik, terutama ketika menghadapi narasi panjang, sindiran, atau perubahan gaya bahasa yang tidak terduga. Hal ini menunjukkan perlunya perbaikan lebih lanjut, termasuk peningkatan dataset dan strategi fine-tuning yang lebih spesifik untuk domain politik.

Namun, model yang dikembangkan dalam penelitian ini menunjukkan keunggulan dalam mengklasifikasikan berita politik sebagai fakta atau hoaks dengan tingkat akurasi yang sangat tinggi, mencapai 95% pada data uji serta nilai AUC sebesar 0.946. Keberhasilan ini menempatkan model pada posisi yang kompetitif dibandingkan dengan berbagai pendekatan lain yang telah diterapkan sebelumnya. Misalnya, Algoritma Naive Bayes Classifier yang digunakan untuk mendeteksi berita hoaks Covid-19 mencapai akurasi 81%, sedangkan Random Forest dan Logistic Regression menghasilkan akurasi masing-masing sebesar 84% dan 77%. Pendekatan hibrida LSTM dan SVM telah mencapai akurasi 94%, sementara perbandingan model LSTM dan GRU hanya menghasilkan akurasi terbaik sebesar 73% dengan model LSTM.

Berdasarkan tingkat keakuratan yang tinggi yang telah dicapai oleh model IndoBERT, model ini dapat diterapkan secara nyata untuk mendukung literasi digital dan mencegah penyebaran berita hoaks melalui berbagai cara. Salah satu skenario implementasinya adalah melalui integrasi dengan platform pengecekan fakta online. Dengan akurasi 95% dan nilai AUC sebesar 0.946, model ini mampu secara otomatis menganalisis teks berita atau klaim yang diunggah pengguna, memberikan klasifikasi antara fakta dan hoaks, serta menyajikan skor kepercayaan sebagai panduan yang andal. Selain itu, model ini dapat diimplementasikan dalam aplikasi media literasi, di mana pengguna diajak untuk mempelajari cara mengenali ciri-ciri berita hoaks melalui simulasi atau modul pembelajaran berbasis kasus nyata yang didukung oleh kecerdasan buatan. Kemudian, dengan kemampuan klasifikasi yang unggul, model ini dapat digunakan dalam sistem monitoring berita pada media sosial untuk mendeteksi

dan menandai konten yang berpotensi hoaks secara real-time. Sistem ini dapat memberikan peringatan kepada pengguna atau mengarahkan mereka ke sumber informasi yang lebih kredibel. Selain itu, penerapan model ini dapat diwujudkan dalam aplikasi chatbot interaktif untuk membantu masyarakat memverifikasi informasi dengan cepat dan akurat. Chatbot ini dapat memberikan respons berdasarkan analisis mendalam model IndoBERT, membantu pengguna memahami keabsahan berita serta memberikan konteks tambahan dari sumber yang relevan. Akhirnya, model ini juga berpotensi besar untuk digunakan oleh institusi pendidikan atau pemerintah dalam kampanye literasi digital yang bertujuan meningkatkan kemampuan masyarakat dalam membedakan berita yang kredibel dari berita palsu, sekaligus mengurangi dampak negatif dari penyebaran hoaks di masyarakat.

IV. KESIMPULAN

Berdasarkan hasil penelitian mengenai implementasi fine-tune menggunakan model IndoBERT untuk deteksi berita hoaks politik, diperoleh beberapa kesimpulan. Model IndoBERT yang dikembangkan mampu mengklasifikasikan berita politik sebagai fakta atau hoaks dengan tingkat akurasi yang tinggi, mencapai 95% pada data uji dengan nilai AUC sebesar 0.946. Keunggulan utama model ini terletak pada kemampuannya dalam menangani bahasa Indonesia secara efektif, khususnya dalam memahami konteks berita politik, sehingga menghasilkan prediksi yang lebih akurat dan relevan. Namun, model ini memiliki keterbatasan, terutama sensitivitas terhadap bias dataset, terutama ketika menghadapi distribusi data yang tidak seimbang antara berita fakta dan hoaks. Meskipun teknik undersampling diterapkan untuk meningkatkan performa model, pendekatan ini secara tidak langsung mengurangi jumlah data pelatihan yang tersedia, yang dapat memengaruhi kemampuan generalisasi model. Selain itu, cakupan penerapan model IndoBERT dalam penelitian ini terbatas pada domain berita politik berbahasa Indonesia, sehingga performanya mungkin tidak optimal jika diterapkan pada jenis berita atau bahasa lain.

DAFTAR PUSTAKA

- [1] A. Lee, "Online Hoaxes, Existential Threat, and Internet Shutdown: A Case Study of Securitization Dynamics in Indonesia," *Journal of Indonesian Social Sciences and Humanities*, vol. 10, no. 1, pp. 17–34, Jun. 2020, doi: 10.14203/jissh.v10i1.156.
- [2] D. R. Fatmala, A. Amelia, and F. A. Trianingsih, "The Use of Social Media Bot Accounts on Influencing Public Opinion: A Legal Review in Indonesia," *Legality: Jurnal Ilmiah Hukum*, vol. 28, no. 2, pp. 169–182, Sep. 2020, doi: 10.22219/jih.v28i2.12148.
- [3] I. B. K. A. Dwipayana, M. I. S. Abenk, and N. H. Lukman, "Public Awareness in Efforts to Defend The Country in The Digital Era in Order to Fight Hoaxes to Maintain State Resilience," *Journal of Digital Law and Policy*, vol. 2, no. 1, pp. 1–10, Sep. 2022, doi: 10.58982/jdlp.v2i1.197.
- [4] X. Zhou and R. Zafarani, "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," *ACM Comput Surv*, vol. 53, no. 5, Sep. 2020, doi: 10.1145/3395046.
- [5] A. Orhan, "Fake news detection on social media: the predictive role of university students' critical thinking dispositions and new media literacy," *Smart Learning Environments*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40561-023-00248-8.
- [6] 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.
- [7] N. G. Ramadhan, F. D. Adhina, A. J. T. Segara, and D. P. Rakhamdani, "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.
- [8] D. F. N. Anisa, I. Mukhlas, 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.
- [9] 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.
- [10] 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.
- [11] A. K. Yadav *et al.*, "Fake News Detection Using Hybrid Deep Learning Method," *SN Comput Sci*, vol. 4, no. 6, Nov. 2023, doi: 10.1007/s42979-023-02296-w.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," Nov. 2020, doi: <https://doi.org/10.48550/arXiv.2011.00677>.
- [17] Y. Liu, S. Agarwal, and S. Venkataraman, "AutoFreeze: Automatically Freezing Model Blocks to Accelerate Fine-tuning," Feb. 2021, doi: <https://doi.org/10.48550/arXiv.2102.01386>.