

# KLASIFIKASI BERITA HOAKS BAHASA INDONESIA MENGUNAKAN INDOBERT *FINE-TUNING* DENGAN PENDEKATAN *FOCAL LOSS* PADA DATA TIDAK SEIMBANG

Aang Kunaefi<sup>\*1)</sup>, Zainal Abidin<sup>2)</sup>, Ririen Kusumawati<sup>3)</sup>

1. Magister Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Magister Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Magister Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

## Article Info

**Kata Kunci:** Berita Hoaks; *Fine Tuning*; *Focal Loss*; IndoBERT; Pemrosesan Bahasa Alami

**Keywords:** *Fine Tuning*; *Focal Loss*; Hoax News; IndoBERT; *Natural Language Processing*

## Article history:

Received 26 April 2025  
Revised 15 May 2025  
Accepted 18 May 2025  
Available online 22 May 2025

## DOI :

<https://doi.org/10.29100/jupi.v10i2.7811>

\* Corresponding author.  
Corresponding Author  
E-mail address:  
[arekbojonegoro@gmail.com](mailto:arekbojonegoro@gmail.com)

## ABSTRAK

Penyebaran berita hoaks di media online menjadi isu serius di tengah meningkatnya konsumsi informasi digital di kalangan masyarakat. Klasifikasi berita hoaks berbahasa Indonesia memiliki peran penting untuk menekan penyebaran informasi palsu. Salah satu tantangan utama dalam sistem klasifikasi ini adalah ketidakseimbangan distribusi data, di mana jumlah berita non-hoaks jauh lebih banyak dibandingkan berita hoaks. Penelitian ini mengusulkan pendekatan klasifikasi berita hoaks berbahasa Indonesia melalui teknologi *Natural Language Processing* (NLP) menggunakan *fine-tuning* model IndoBERT, yang merupakan *pre-trained language model* berbasis arsitektur BERT (*Bidirectional Encoder Representations from Transformers*) dan disesuaikan untuk Bahasa Indonesia. Ketidakseimbangan data diatasi menggunakan metode *Focal Loss*. Pendekatan *focal loss* dirancang untuk lebih menekankan pembelajaran pada sampel kelas minoritas yang sulit diklasifikasikan. Penelitian ini menggunakan dataset dari platform Kaggle, Huggingface dan Mendeley. Tataset mencakup berita Bahasa Indonesia dengan jumlah data berita hoaks jauh lebih kecil dari berita non hoaks. Hasil evaluasi menunjukkan bahwa kombinasi IndoBERT dan *Focal Loss* mampu meningkatkan performa model dengan akurasi sebesar 98.3% dibandingkan dengan pendekatan *Cross-Entropy Loss* yang mendapat akurasi 97%. Penelitian ini menunjukkan bahwa penggabungan model berbasis bahasa alami dengan strategi penanganan data tidak seimbang dapat memberikan hasil yang lebih akurat dalam mendeteksi berita hoaks.

## ABSTRACT

The spread of hoax news on online media has become a serious issue amidst the increasing consumption of digital information among the public. Classification of hoax news in Indonesian Language has an important role in reducing the spread of false information. One of the main challenges in this classification system is the imbalance in data distribution, where the number of non-hoax news is much greater than hoax news. This study proposes an approach to classifying hoax news in Indonesian through *Natural Language Processing* (NLP) technology using the IndoBERT *fine-tuning* model, which is a pre-trained language model based on the BERT (*Bidirectional Encoder Representations from Transformers*) architecture and is adapted for Indonesian Language. Data imbalance is overcome using the *Focal Loss* method. The *focal loss* approach is designed to emphasize learning more on minority class samples that are difficult to classify. This study uses datasets from the Kaggle, Huggingface and Mendeley platforms. The dataset includes Indonesian news with a much smaller amount of hoax news data than factual one. The evaluation results show that the combination of IndoBERT and *Focal Loss* is able to improve model performance with an accuracy of 98.3%, compared to the *Cross-Entropy Loss* approach, which only gets an accuracy of 97%. This study show that combining natural language-based models with imbalanced data handling strategies can provide more accurate in detecting hoax news.

## I. PENDAHULUAN

**P**ERKEMBANGAN teknologi digital telah membawa perubahan mendasar terhadap pola distribusi informasi, media online telah mengambil alih peran dominan dalam penyebaran informasi, melampaui media cetak. Kemajuan teknologi informasi telah mendorong meningkatnya konsumsi informasi digital di kalangan masyarakat. Pada tahun 2021, jumlah pengguna internet di Indonesia mengalami lonjakan dari 175 juta menjadi 202 juta pengguna [1]. Laporan terbaru dari *We Are Social* dan *Meltwater* (2024) menunjukkan bahwa hingga Januari 2024, terdapat 212,9 juta pengguna internet di Indonesia, setara dengan 77% dari total populasi. Dari jumlah tersebut, sekitar 167 juta pengguna (60,4%) merupakan pengguna aktif media sosial.

Kondisi ini menunjukkan bahwa media daring menjadi jalur utama dalam penyebaran informasi di era digital saat ini. Namun, kemudahan akses dan penyebaran informasi ini tidak disertai dengan mekanisme penyaringan atau verifikasi yang ketat. Berbeda dengan media cetak yang memiliki redaksi dan kontrol editorial, informasi di media daring dapat disebar oleh siapa saja tanpa proses validasi yang memadai. Hal ini menyebabkan meningkatnya potensi penyebaran informasi hoaks secara masif dan cepat [2]. Berdasarkan laporan Komite Litbang MAFINDO (Masyarakat Anti Fitnah Indonesia), pada semester pertama tahun 2024 tercatat temuan berita hoaks mencapai 2.119 kasus. Penyebaran berita hoaks memiliki dampak serius terhadap kepercayaan publik, stabilitas sosial, dan integritas informasi. Oleh karena itu, deteksi dan klasifikasi berita hoaks menjadi penting dalam upaya mengurangi dampak berita hoaks. Masyarakat harus bisa membedakan antara informasi yang benar dan yang informasi yang salah yang bersifat menyesatkan. Upaya klasifikasi berita hoaks secara otomatis menggunakan teknologi *Natural Language Processing* (NLP) menjadi solusi yang baik. Melalui pendekatan ini, sistem dapat membantu mengidentifikasi dan membedakan antara berita hoaks dari berita non-hoaks secara cepat dan akurat, sehingga dapat mendukung literasi digital masyarakat serta mencegah penyebarluasan informasi yang menyesatkan.

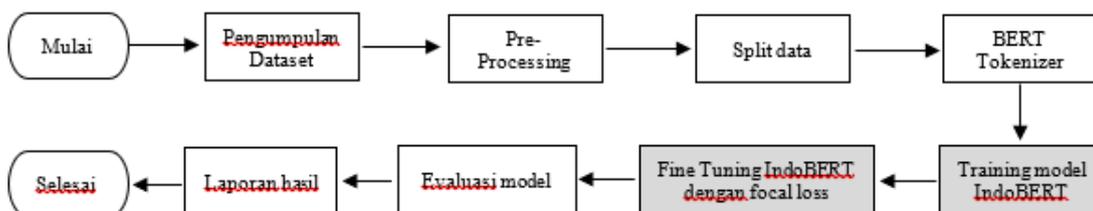
Penelitian terkait deteksi berita hoaks telah banyak dilakukan dengan memanfaatkan model BERT sebagai pendekatan klasifikasi. Salah satu masalah yang coba dipecahkan dalam penelitian klasifikasi ini adalah ketidakseimbangan jumlah data antara berita hoaks dan non-hoaks. Studi oleh Tobing *et al.* [3] menggunakan *Fine-Tuning* IndoBERT pada berita politik dan menerapkan *undersampling* untuk mengatasi ketimpangan data. Sementara itu, Ridho dan Yulianti [4] membandingkan IndoBERT dengan model lain seperti CNN-LSTM, *Logistic Regression*, dan *Naive Bayes*, serta menerapkan *oversampling* SMOTE untuk meningkatkan akurasi pada data tidak seimbang dari 92% menjadi 98%. Penelitian lain dilakukan oleh Fathin *et al* [5] yang melakukan klasifikasi berita hoaks dengan menerapkan *Random Oversampling* (ROS) dan *Random Undersampling* (RUS) untuk mengatasi ketidakseimbangan *dataset*. Dari hasil penerapan metode tersebut model mencapai akurasi 98,2%.

Tantangan utama dalam deteksi hoaks berbahasa Indonesia tidak hanya terletak pada distribusi data yang tidak seimbang, tetapi juga pada kompleksitas linguistik. Berita hoaks sering kali menyerupai berita non-hoaks dalam struktur dan gaya bahasa, serta cenderung menggunakan diksi yang emosional atau manipulatif, sehingga mempersulit proses klasifikasi. Model berbasis *transformer* seperti IndoBERT memiliki kemampuan representasi semantic secara menyeluruh, mendalam dan kontekstual, menjadikannya lebih unggul dalam menangani kompleksitas ini. Pendekatan ini digunakan sebagai respons terhadap keterbatasan metode klasik. Algoritma seperti *Naive Bayes* dan SVM yang mengandalkan representasi teks yang dangkal dan tidak kontekstual, sehingga kurang efektif dalam menangkap makna yang tergantung pada struktur kalimat [6]. Arsitektur CNN-LSTM memang menawarkan kemampuan yang lebih baik dalam mengenali urutan dan fitur lokal dan global dalam teks, namun tetap terbatas dalam memahami konteks semantik secara menyeluruh, terutama dalam teks dengan struktur kompleks [7]. Evaluasi empiris menunjukkan bahwa IndoBERT mampu memberikan performa yang lebih konsisten dalam memahami struktur dan makna dalam bahasa Indonesia. Namun performa model ini tetap menurun saat dihadapkan pada distribusi kelas yang tidak seimbang, untuk menangani ketidakseimbangan data tersebut, *Focal Loss* diterapkan untuk memfokuskan pembelajaran pada sampel yang sulit pada kelas minoritas (hoaks), sehingga meningkatkan sensitivitas model terhadap berita hoaks. pendekatan *Focal Loss* dipilih karena tidak memerlukan modifikasi data seperti yang dilakukan oleh metode *oversampling* maupun *undersampling*. Teknik *oversampling* berisiko menyebabkan *overfitting* karena melakukan duplikasi data pada kelas minoritas [8], sementara *undersampling* berpotensi mengurangi informasi penting dari kelas mayoritas [9]. *Focal Loss* menawarkan solusi yang lebih efisien dengan menangani ketidakseimbangan data tanpa mengubah distribusi data asli, sehingga model tetap dapat belajar dari keseluruhan data yang tersedia secara utuh. Oleh sebab itu, kombinasi IndoBERT dengan *Focal Loss* dipandang sebagai pendekatan yang lebih efektif dalam menangani tantangan klasifikasi berita hoaks, baik dari sisi representasi linguistik maupun dari sisi distribusi data yang tidak seimbang.

## II. METODE PENELITIAN

Penelitian ini menerapkan arsitektur IndoBERT *Fine Tuning* pada klasifikasi berita hoaks Bahasa Indonesia, dan menggunakan pendekatan *Focal Loss* untuk mengatasi ketidakseimbangan distribusi data. Penelitian dilakukan dengan memanfaatkan *cloud-base-service* dari *Google Colab* serta menggunakan *runtime* dari *Python 3*. Arsitektur model menggunakan *Transformer* dari *Hugging Fase* dan *framework deep learning PyTorch* untuk menjalankan arsitektur *Transformer*.

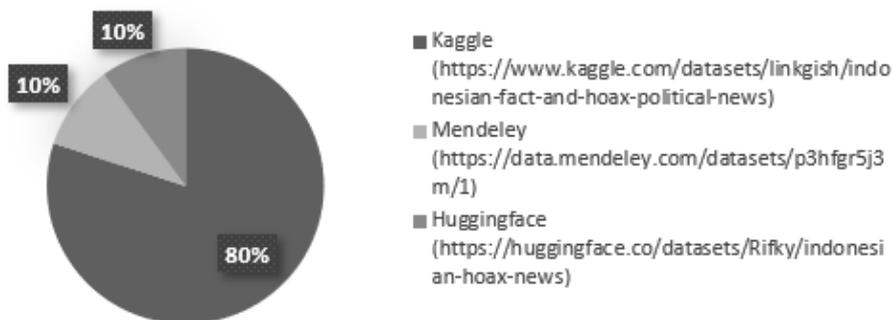
Tahapan dalam penelitian model klasifikasi berita dengan IndoBERT ini dimulai dengan pengumpulan data, *pre-processing* data, pembagian data, tokenisasi, pelatihan model IndoBERT, *fine-tuning* dengan *loss function focal loss*, evaluasi model, dan terakhir laporan hasil klasifikasi. Tahapan penelitian dapat dilihat dari pada gambar 1.



Gambar. 1. Tahapan Penelitian

### A. Pengumpulan Dataset

Penelitian ini memanfaatkan dataset sekunder yang berisi berita berbahasa Indonesia, yang telah diklasifikasikan ke dalam dua kategori utama: berita non-hoaks dan berita hoaks. Sumber data dikumpulkan dari tiga platform daring, yaitu Kaggle, Mendeley, dan Huggingface. Proporsi data menunjukkan adanya ketidakseimbangan kelas, dengan sekitar 75% merupakan berita non-hoaks dan 25% merupakan berita hoaks. Komposisi jumlah data dapat dilihat pada gambar 2. Perbedaan jumlah antara berita hoaks dan non-hoaks menunjukkan bahwa dataset bersifat tidak seimbang, yang berpotensi mempengaruhi kinerja model dalam mendeteksi kelas minoritas (berita hoaks). Ketidakseimbangan ini menjadi pertimbangan dalam pemilihan metode pelatihan model, khususnya dalam penggunaan *loss function* yang dapat mengatasi dominasi kelas mayoritas.



Gambar. 2. Komposisi dataset dari tiga platform

### B. Pre-processing Data

Tahap *pre-processing* data terdiri dari dua tahap yaitu tahap *cleaning* dan *encoding*, seperti yang ditunjukkan pada gambar 3. Proses *cleaning* dilakukan untuk menghilangkan berbagai simbol atau karakter yang tidak relevan dan tidak akurat dalam dataset [6]. Tahapan ini mencakup konversi seluruh teks menjadi huruf kecil, penghapusan URL, tanda baca, karakter berulang, serta penggantian baris baru dan spasi berlebih dengan spasi tunggal. Tujuan dari proses ini adalah untuk meningkatkan keakuratan dan kestabilan model dalam melakukan klasifikasi. Sementara itu, tahap *encoding* dalam penelitian ini dilakukan dengan menambahkan kolom baru bernama label, yang merepresentasikan kelas target dalam bentuk nilai numerik, di mana teks berlabel “non-hoaks” dikodekan sebagai 0 dan “hoaks” sebagai 1. Gambar 3 menunjukkan tahapan *pre-processing* dalam penelitian dapat.



Gambar. 3. Tahapan *pre-processing*

### C. Split data

Data yang telah melalui proses *preprocessing* akan dipisahkan menjadi tiga subset [10] yaitu data pelatihan, data validasi, dan data pengujian. Berbeda dengan beberapa model machine learning konvensional seperti *Logistic Regression* dan *Decision Tree* yang umumnya hanya membagi data menjadi dua bagian (pelatihan dan pengujian), model IndoBERT memerlukan skema pembagian menjadi tiga. Hal ini dikarenakan BERT memiliki kompleksitas dan fleksibilitas yang tinggi, sehingga keberadaan data validasi menjadi krusial untuk memantau performa model selama proses pelatihan dan mencegah terjadinya *overfitting* [11], yang sering kali sulit diatasi tanpa validasi yang memadai.

Penelitian ini menerapkan metode klasifikasi berita hoaks dengan menggunakan rasio pembagian data yang umum digunakan dalam studi-studi terdahulu yang melibatkan model BERT, yaitu 80% untuk data pelatihan, 10% untuk data validasi, dan 10% untuk data pengujian [12][13]. Rasio tersebut dipandang seimbang karena mampu menyediakan jumlah data yang memadai untuk proses pelatihan model, sekaligus tetap menyisihkan sebagian data yang cukup untuk keperluan validasi dan evaluasi. Pembagian ini bertujuan untuk mencegah *overfitting* dan memastikan kinerja model tetap optimal pada data yang belum pernah dilihat sebelumnya.

### D. BERT Tokenizer

Pada tahap ini, proses tokenisasi dilakukan untuk mengubah data teks menjadi format numerik yang dapat diproses oleh model BERT. Tokenisasi bertugas memecah setiap kalimat atau dokumen menjadi unit-unit yang lebih kecil [14] yang disebut token, yang dapat berupa kata utuh atau potongan sub-kata sesuai dengan aturan tokenisasi BERT. Setiap teks diawali dengan token khusus [CLS] yang berfungsi sebagai representasi keseluruhan kalimat, dan diakhiri dengan token [SEP] sebagai penanda akhir kalimat. Setiap token kemudian dikonversi menjadi representasi numerik berupa token ID, yang mencerminkan posisi token dalam kosakata BERT, dan *attention mask*, yang menandai token mana yang harus diproses oleh model dan mana yang harus diabaikan, seperti token hasil *padding*. Penggunaan *attention mask* sangat penting untuk memastikan bahwa model hanya fokus pada bagian input yang informatif, sehingga meningkatkan efektivitas pemahaman konteks dalam tugas-tugas pemrosesan bahasa alami. Oleh karena itu, tokenisasi memainkan peran penting dalam mentransformasikan data teks mentah menjadi format yang sesuai untuk proses komputasi dalam model BERT [11]

### E. Pelatihan Model IndoBERT

Pelatihan model dalam penelitian ini menggunakan IndoBERT, sebuah model berbasis *Transformer* yang telah melalui tahap *pre-trained* pada korpus besar berbahasa Indonesia. Korpus ini mencakup berbagai jenis teks dalam bahasa Indonesia, seperti artikel berita, media sosial, dokumen resmi, dan data publik lainnya, yang dirancang untuk mencerminkan struktur dan variasi penggunaan bahasa Indonesia secara luas. IndoBERT mengimplementasikan arsitektur BERT [15] yang memungkinkan pemrosesan konteks secara dua arah (*bidirectional*), yaitu dari sisi kiri dan kanan teks secara bersamaan [16]. Pendekatan ini memungkinkan model untuk memahami makna kata berdasarkan konteks penuh dalam kalimat.

Optimasi model dilakukan dengan menggunakan algoritma AdamW (*Adam with Weight Decay*), dengan nilai learning rate sebesar  $3e-6$ . AdamW dipilih karena kemampuannya dalam menggabungkan kecepatan konvergensi dari Adam dengan regularisasi eksplisit melalui *weight decay*, yang terbukti efektif dalam pelatihan model berskala besar seperti BERT. Selain itu AdamW mampu bekerja secara cepat dan efisien dalam mengoptimalkan sebuah model [17]. Optimizer ini menghitung dua estimasi statistik dari gradien, yaitu momen pertama menggunakan rumus (1) dan momen kedua menggunakan rumus (2)

$$m_t = \beta_1 \times m_{t-1} + (1 - \beta_1) g_t \quad (1)$$

$$v_t = \beta_2 \times v_{t-1} + (1 - \beta_2) g_t^2 \quad (2)$$

Dimana:

$m_t$  : momentum pertama (estimasi rata-rata gradien),

$v_t$  : momentum kedua (estimasi varians gradien),

$g_t$  : gradien dari fungsi loss pada waktu ke- $t$

$\beta_1$  dan  $\beta_2$  : koefisien penyusutan bobot eksponensial

Kombinasi dari proses tokenisasi yang sesuai format BERT, penggunaan model pre-trained IndoBERT, serta optimasi berbasis AdamW bertujuan untuk menghasilkan proses pelatihan yang stabil, efisien, dan akurat dalam

menyelesaikan tugas klasifikasi berita hoaks berbahasa Indonesia

#### F. *Fine-tuning IndoBERT dengan Focal Loss*

Pada tahap ini, model IndoBERT digunakan sebagai arsitektur dasar dalam proses *fine-tuning* untuk menyelesaikan tugas klasifikasi teks berita hoaks dalam bahasa Indonesia. *Fine-tuning* merujuk pada proses penyesuaian lanjutan terhadap parameter-parameter model yang telah dilatih sebelumnya (*pre-trained*), agar model dapat beradaptasi secara optimal terhadap tugas klasifikasi yang lebih spesifik. Dalam penelitian ini, klasifikasi dilakukan terhadap dua kelas, yaitu berita non-hoaks dan berita hoaks.

Pemanfaatan IndoBERT yang telah melalui tahap *pre-training* pada korpus besar berbahasa Indonesia memberikan dasar representasi bahasa yang kuat. Namun, untuk mencapai performa yang lebih akurat pada tugas klasifikasi tertentu, diperlukan proses *fine-tuning*. Proses ini memungkinkan model untuk mempelajari pola-pola khusus dari data yang lebih sempit dan relevan dengan domain yang dituju, seperti karakteristik linguistik dalam berita hoaks. Penyesuaian ini penting karena model *pre-trained* secara umum dilatih pada data yang luas dan bersifat umum, sehingga belum tentu optimal untuk tugas klasifikasi yang spesifik.

Menurut Rogers et al. [18] *fine-tuning* pada model berbasis *Transformer* seperti BERT secara signifikan meningkatkan performa karena mampu menyesuaikan representasi internal model terhadap distribusi data dan label yang baru. Oleh karena itu, pendekatan ini dipilih untuk memperoleh hasil klasifikasi yang lebih akurat dan relevan dalam mendeteksi hoaks berbahasa Indonesia. Dataset yang digunakan dalam penelitian ini bersifat tidak seimbang, dengan distribusi sebanyak 75% data untuk kelas non-hoaks dan 25% data untuk kelas hoaks. Ketidakseimbangan ini berpotensi menyebabkan bias model terhadap kelas mayoritas. Untuk mengatasi hal tersebut, digunakan *loss function* dengan *focal loss* sebagai pengganti dari *Cross-Entropy Loss*. Fungsi ini pertama kali diperkenalkan dalam konteks deteksi objek pada model RetinaNet [19]. *Focal Loss* dirancang untuk memberikan bobot lebih besar terhadap sampel yang sulit diklasifikasikan (dalam hal ini, berita hoaks yang berjumlah lebih sedikit), serta mengurangi kontribusi dari sampel yang lebih mudah dipelajari (berita non-hoaks yang lebih banyak). Formula *focal loss* dituliskan dalam rumus (3). Proses *fine-tuning* dilakukan dengan epoch sebanyak 10, batch-size 32 dan learning rate  $3e-6$

$$FL(p_t) = -a_t(1 - p_t)^\gamma \log(p_t) \quad (3)$$

Dimana:

$p_t$  : probabilitas prediksi untuk kelas yang benar

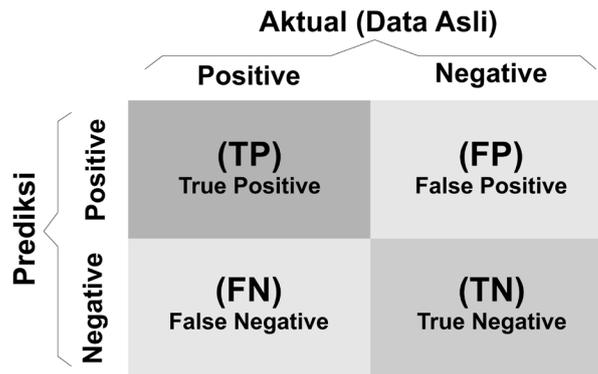
$\gamma$  : focus parameter

$a_t$  : balancing factor antara kelas

#### G. *Evaluasi Model*

Pada tahap evaluasi model, penelitian ini mengadopsi dua pendekatan utama, yaitu *confusion matrix* dan *classification report*, untuk menilai efektivitas model IndoBERT *fine tuning* dengan *focal loss* dalam melakukan klasifikasi terhadap berita hoaks bahasa Indonesia. *Confusion matrix* merupakan sebuah representasi tabular yang mengilustrasikan hubungan antara hasil prediksi model dan label aktual pada data uji. Matriks ini terdiri dari empat komponen utama, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN), yang masing-masing mencerminkan jenis-jenis prediksi yang benar maupun salah oleh model terhadap dua kelas yang dianalisis. Dengan menyediakan rincian informasi mengenai distribusi kesalahan dan keberhasilan prediksi, *confusion matrix* memungkinkan peneliti untuk melakukan analisis menyeluruh terhadap keunggulan serta keterbatasan model dalam mendeteksi masing-masing kelas.

Dalam klasifikasi berita hoaks, khususnya pada dataset yang tidak seimbang antara kelas hoaks dan non-hoaks, penggunaan *confusion matrix* menjadi penting untuk memahami sejauh mana model dapat membedakan kelas hoaks dan non-hoaks secara akurat. Visualisasi *confusion matrix* juga mempermudah proses identifikasi pola kesalahan klasifikasi yang mungkin terjadi akibat bias terhadap kelas mayoritas. *confusion matrix* tidak hanya berfungsi sebagai alat ukur performa secara numerik, tetapi juga sebagai sarana eksploratif dalam menganalisis perilaku model karena lebih informatif [20]. Visualisasi *confusion matrix* dapat dilihat pada gambar 4.



Gambar. 4. *Confusion Matrix*

Selain menggunakan *confusion matrix*, evaluasi lebih mendalam dilakukan melalui *classification report*, yang menyajikan metrik evaluasi seperti *precision*, *recall*, *F1-score*, dan *support*. *Precision* digunakan untuk mengukur proporsi prediksi positif yang benar terhadap seluruh prediksi positif yang dihasilkan oleh model, formula *precision* dituliskan dalam rumus (4)

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

dimana TP adalah jumlah prediksi benar untuk kelas positif, dan FP adalah jumlah prediksi salah yang diklasifikasikan sebagai positif. Nilai *precision* yang tinggi menunjukkan kemampuan model dalam meminimalkan kesalahan klasifikasi terhadap kelas hoaks.

Selain menilai *precision*, *classification report* juga menilai *Recall* atau sensitivitas, untuk menghitung sejauh mana model dapat mengidentifikasi seluruh instance dari kelas positif yang sebenarnya, dituliskan dalam rumus (5). dengan FN mengacu pada jumlah instance kelas positif yang gagal dikenali oleh model. *Recall* yang tinggi menunjukkan efektivitas model dalam mendeteksi seluruh berita hoaks dalam dataset

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

Setelah diperoleh nilai *precision* dan *recall*, dilakukan perhitungan *F1-score* untuk mengevaluasi keseimbangan antara kedua metrik tersebut. *F1-score* merupakan rata-rata harmonik dari *precision* dan *recall*, sebagaimana dirumuskan dalam Persamaan (6). Penggunaan *F1-score* menjadi krusial dalam konteks klasifikasi dengan distribusi kelas yang tidak seimbang, karena metrik tersebut mampu merepresentasikan performa model secara lebih menyeluruh dalam mengidentifikasi kedua kelas secara proporsional.

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (6)$$

Dalam konteks studi ini, pemilihan *F1-score* sebagai metrik evaluasi dianggap tepat karena peneliti mengharapkan pengembangan model yang tidak hanya memiliki kemampuan tinggi dalam mendeteksi berita hoaks, tetapi juga mampu mengklasifikasikan berita non-hoaks secara akurat dengan meminimalkan kesalahan prediksi. Dengan demikian, *F1-score* memberikan ukuran performa yang seimbang dalam mengakomodasi kebutuhan untuk mengurangi kesalahan klasifikasi pada kedua kategori tersebut..

### III. HASIL DAN PEMBAHASAN

Pada bagian ini disajikan hasil penelitian dan analisis performa model klasifikasi berita hoaks menggunakan IndoBERT yang telah melalui proses *fine-tuning*. Fokus utama penelitian ini adalah mengevaluasi efektivitas penggunaan *focal loss* dalam menangani ketidakseimbangan kelas antara berita non-hoaks dan berita hoaks. Dataset yang digunakan menunjukkan distribusi data yang tidak seimbang, Oleh karena itu, diperlukan pendekatan *loss function* yang adaptif terhadap distribusi kelas

### A. Hasil Evaluasi Model

Model dilatih selama 10 epoch dengan menggunakan *batch size* sebesar 32 dan *learning rate* sebesar 3e-6. Pada tahap pelatihan, digunakan fungsi *Focal Loss*, yang secara khusus dirancang untuk menangani masalah ketidakseimbangan data. *Focal Loss* bekerja dengan cara memberikan penalti yang lebih besar terhadap kesalahan prediksi pada kelas minoritas (berita hoaks), sehingga model dapat belajar lebih baik pada kategori yang kurang representatif. Sebagai pembandingan, penelitian ini juga menggunakan *Cross Entropy Loss*, yaitu fungsi *loss* standar yang sering digunakan dalam tugas klasifikasi.

Untuk mengevaluasi performa masing-masing metode yang digunakan dalam penelitian ini, dilakukan proses evaluasi *kuantitatif* dengan memanfaatkan *confusion matrix* dan *classification report*. Evaluasi ini mencakup sejumlah metrik performa yang digunakan dalam tugas klasifikasi, yaitu *precision*, *recall*, *F1-score*, dan *accuracy*. Keempat metrik ini memberikan gambaran yang komprehensif terhadap kemampuan model dalam mengklasifikasikan data, khususnya dalam konteks distribusi data yang tidak seimbang. Melalui analisis ini, efektivitas dari masing-masing fungsi *loss* dapat diukur secara objektif berdasarkan seberapa baik model mampu membedakan antara kelas mayoritas dan minoritas. Hasil evaluasi model dapat dilihat pada tabel 1.

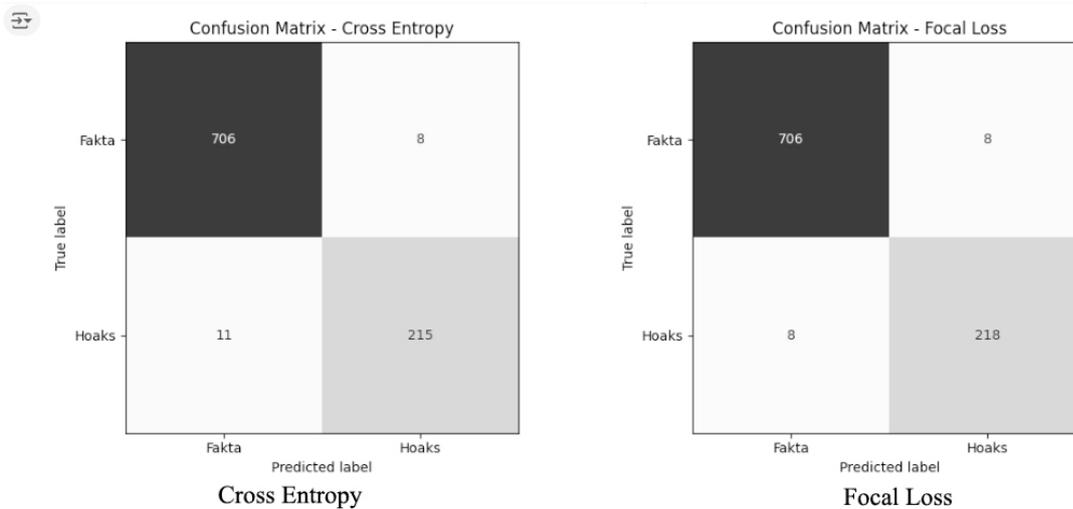
TABEL I  
 HASIL EVALUASI MODEL

Model	<i>Precision</i> (hoaks)	<i>Recall</i> (hoaks)	<i>F1-Score</i> (hoaks)	<i>Precision</i> (non-hoaks)	<i>Recall</i> (non-hoaks)	<i>F1-Score</i> (non-hoaks)	<i>Accuracy</i>
IndoBERT + <i>Cross Entropy</i>	0.9641	0.9513	0.9576	0.9846	0.9887	0.9867	0.9797
IndoBERT + <i>Focal Loss</i>	0.9646	0.9646	0.9646	0.9887	0.9887	0.9887	0.9829

Hasil evaluasi menunjukkan bahwa model yang dikembangkan mampu mencapai performa klasifikasi yang baik meskipun terjadi ketidakseimbangan distribusi data antar kelas. Salah satu perbedaan utama terlihat pada performa klasifikasi terhadap kelas hoaks sebagai kelas minoritas. Model yang dilatih menggunakan *Focal Loss* menunjukkan peningkatan *recall* pada kelas hoaks, dari 0,9513 pada model *Cross Entropy* menjadi 0,9646 pada model *Focal Loss*. Peningkatan ini menunjukkan bahwa model menjadi lebih sensitif dalam mengenali berita hoaks yang sesungguhnya, sehingga mampu mengurangi jumlah *false negatives* dalam konteks deteksi hoaks. Selain itu, peningkatan *recall* tersebut tidak disertai dengan penurunan *precision*. Sebaliknya, *precision* pada kelas hoaks mengalami meningkat meskipun tidak signifikan dari 0,9641 menjadi 0,9646, yang menunjukkan bahwa peningkatan sensitivitas model tidak menyebabkan lonjakan *false positive*. Secara keseluruhan, metrik evaluasi lainnya seperti akurasi dan *f1-score*, juga mengalami peningkatan. Hal ini mencerminkan bahwa penggunaan *Focal Loss* tidak hanya memperbaiki kinerja pada kelas minoritas, tetapi juga meningkatkan performa keseluruhan model.

Visualisasi menggunakan *confusion matrix* dapat dilihat pada gambar 5, yang menunjukkan perbandingan model *Cross Entropy* dan *Focal Loss* dengan *Confusion Matrix*. Visualisasi tersebut menunjukkan performa klasifikasi yang sangat baik pada kelas non-hoaks, di mana kedua model berhasil mengklasifikasikan 706 dari 714 data non-hoaks dengan benar. Namun, perbedaan yang lebih mencolok tampak pada performa terhadap kelas hoaks. Model dengan *Cross Entropy* menghasilkan 11 kesalahan klasifikasi hoaks sebagai non-hoaks, sementara model dengan *Focal Loss* hanya menghasilkan 8 kesalahan pada kasus serupa. Dengan kata lain, *Focal Loss* meningkatkan akurasi klasifikasi kelas hoaks, sebagaimana tercermin dari peningkatan *recall* yang telah disebutkan sebelumnya. Peningkatan ini konsisten dengan karakteristik utama *Focal Loss*, yaitu kemampuannya dalam menangani ketidakseimbangan kelas dengan memberikan penekanan lebih pada contoh yang sulit diklasifikasikan. Berdasarkan temuan ini, dapat disimpulkan bahwa penerapan *Focal Loss* merupakan pendekatan yang lebih efektif untuk klasifikasi berita hoaks dengan distribusi data yang tidak seimbang.

Gambar 5 menunjukkan model *Focal Loss* berhasil meningkatkan akurasi dalam klasifikasi berita hoaks, namun masih ditemukan delapan kasus di mana berita hoaks diklasifikasikan sebagai berita non-hoaks (fakta) pada model *focal loss*. Temuan ini mengindikasikan bahwa terdapat berita hoaks memiliki kemiripan yang tinggi dengan berita non-hoaks dalam hal struktur naratif dan gaya penyampaian. Karakteristik tersebut mencakup penggunaan bahasa yang netral, penyajian yang menyerupai laporan jurnalistik, serta minimnya elemen emosional atau provokatif. Efektivitas *Focal Loss* dalam menekan kesalahan pada kelas minoritas dan menjaga akurasi pada kelas mayoritas tetap menghadapi kendala pada kasus-kasus yang bersifat ambigu. Keberhasilan klasifikasi dalam konteks tersebut membutuhkan kemampuan model dalam memahami konteks yang lebih kompleks secara semantik maupun pragmatik.



Gambar. 5. Hasil perbandingan model *Cross Entropy* (kiri) dan *Focal Loss* (kanan)

## B. Pembahasan

Hasil yang diperoleh menunjukkan bahwa penerapan *fine-tuning* pada model IndoBERT terbukti efektif untuk tugas klasifikasi berita hoaks dalam bahasa Indonesia. Hal ini menunjukkan bahwa IndoBERT memiliki kemampuan yang baik dalam menangkap dan memahami kompleksitas linguistik yang terdapat dalam teks bahasa Indonesia. Penggunaan *Focal Loss* terbukti memberikan kontribusi positif terhadap peningkatan performa model pada data tidak seimbang, dengan meningkatkan *recall* pada kelas minoritas tanpa penurunan *precision*. Sebaliknya, *precision* pada kelas hoaks meningkat walaupun tidak secara signifikan, yang menunjukkan bahwa peningkatan sensitivitas model tidak menyebabkan lonjakan *false positive*. Dibandingkan studi sebelumnya yang menggunakan teknik *oversampling* dan *undersampling* untuk menangani ketidakseimbangan data, pendekatan *focal loss* dinilai lebih efisien karena tidak memerlukan manipulasi data, melainkan secara langsung menyesuaikan bobot kesalahan selama proses pelatihan. sehingga menghasilkan model yang lebih stabil dan dapat diimplementasikan tanpa mengurangi jumlah data asli.

Pendekatan lain seperti *re-weighting* juga menjadi bahan pertimbangan, namun metode *re-weighting* memerlukan penentuan bobot kelas secara manual, yang jika tidak ditentukan dengan tepat dapat menghasilkan performa yang tidak konsisten dan berisiko menyebabkan *overfitting*, terutama dalam kasus ketidakseimbangan data yang ekstrem [21]. Disisi lain, metode *ensemble learning* seperti *bagging* dan *boosting* cenderung meningkatkan kompleksitas model dan waktu pelatihan, serta membutuhkan sumber daya komputasi yang lebih besar [22]. Dengan demikian, pemanfaatan representasi semantik IndoBERT yang kuat dalam bahasa Indonesia, yang dikombinasikan dengan kemampuan adaptif *Focal Loss* dalam menangani distribusi data tidak seimbang, menunjukkan bahwa pendekatan ini menjadi solusi yang tepat dalam klasifikasi berita dengan distribusi data yang tidak seimbang.

## IV. KESIMPULAN

Penelitian ini menunjukkan bahwa model IndoBERT yang telah melalui proses *fine-tuning* dapat secara efektif digunakan untuk klasifikasi berita hoaks berbahasa Indonesia, bahkan dalam kondisi distribusi data yang tidak seimbang. Penggunaan *Focal Loss* terbukti meningkatkan performa model, khususnya dalam mendeteksi kelas minoritas, yaitu berita hoaks. Model yang diterapkan berhasil mencapai akurasi pada semua kelas sebesar 98.3%, dengan peningkatan signifikan pada metrik *F1-score* dan *Recall* untuk kelas hoaks dibandingkan dengan penggunaan *Cross Entropy Loss*. Pendekatan ini memungkinkan model untuk lebih sensitif terhadap kesalahan prediksi pada kelas yang kurang representatif, tanpa perlu melakukan manipulasi data seperti *oversampling* dan *undersampling* atau penentuan bobot secara manual seperti *re-weighting*. Dengan demikian, integrasi *Focal Loss* dalam pelatihan model berbasis IndoBERT menawarkan solusi yang lebih efisien dan stabil dalam menangani masalah ketidakseimbangan data pada tugas klasifikasi berita. Penelitian selanjutnya dapat mengeksplorasi kombinasi metode ini dengan strategi *augmentasi* data atau penggunaan model *multilingual* untuk meningkatkan kemampuan model terhadap data baru.

## DAFTAR PUSTAKA

- [1] B. Juarto and Yulianto, "Indonesian News Classification Using IndoBert," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 2, pp. 454–460, 2023.
- [2] S. Vosoughi, D. Roy, and S. Aral, "News On-line," *Science (80- )*, vol. 1151, no. March, pp. 1146–1151, 2018.
- [3] C. J. L. Tobing, I. G. N. L. Wijayakusuma, L. Putu, I. Harini, and U. Udayana, "Detection of Political Hoax News Using Fine-Tuning IndoBERT," vol. 9, no. 2, pp. 354–360, 2025.
- [4] M. Y. Ridho and E. Yulianti, "From Text to Truth : Leveraging IndoBERT and Machine Learning Models for Hoax Detection in Indonesian News," vol. 10, no. 3, pp. 544–555, 2024, doi: 10.26555/jiteki.v10i3.29450.
- [5] M. A. Fathin, Y. Sibaroni, and S. S. Prasetyowati, "Handling Imbalance Dataset on Hoax Indonesian Political News Classification using IndoBERT and Random Sampling," *J. Media Inform. Budidarma*, vol. 8, no. 1, p. 352, 2024, doi: 10.30865/mib.v8i1.7099.
- [6] B. Altmel and M. C. Ganiz, "Semantic text classification: A survey of past and recent advances," *Inf. Process. Manag.*, vol. 54, no. 6, pp. 1129–1153, 2018, doi: 10.1016/j.ipm.2018.08.001.
- [7] C. Li, C. Zhang, and Q. Fu, "Research on CNN + LSTM user intention classification based on multi-granularity features of texts," *J. Eng.*, vol. 2020, no. 13, pp. 486–490, 2020, doi: 10.1049/joe.2019.1175.
- [8] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Networks*, vol. 106, pp. 249–259, 2018, doi: 10.1016/j.neunet.2018.07.011.
- [9] C. J. Lin, J. G. Wang, and C. Y. Lee, "Pattern recognition using neural-fuzzy networks based on improved particle swarm optimization," *Expert Syst. Appl.*, vol. 36, no. 3 PART 1, pp. 5402–5410, 2009, doi: 10.1016/j.eswa.2008.06.110.
- [10] R. Yang *et al.*, "CNN-LSTM deep learning architecture for computer vision-based modal frequency detection," *Mech. Syst. Signal Process.*, vol. 144, p. 106885, 2020, doi: 10.1016/j.ymsp.2020.106885.
- [11] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Naacl-Hlt 2019*, no. Mlm, pp. 4171–4186, 2018, [Online]. Available: <https://aclanthology.org/N19-1423.pdf>
- [12] P. K. Verma, P. Agrawal, V. Madaan, and R. Prodan, "MCred: multi-modal message credibility for fake news detection using BERT and CNN," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 8, pp. 10617–10629, 2023, doi: 10.1007/s12652-022-04338-2.
- [13] S. Raza, D. Paulen-Patterson, and C. Ding, "Fake news detection: comparative evaluation of BERT-like models and large language models with generative AI-annotated data," *Knowl. Inf. Syst.*, pp. 1–30, 2025, doi: 10.1007/s10115-024-02321-1.
- [14] A. R. Hanum *et al.*, "Analisis Kinerja Algoritma Klasifikasi Teks BERT Dalam Mendeteksi Berita Hoaks," vol. 11, no. 3, pp. 537–546, 2024, doi: 10.25126/jtiik938093.
- [15] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 3, pp. 746–757, 2023, doi: 10.26555/jiteki.v9i3.26490.
- [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," *COLING 2020 - 28th Int. Conf. Comput. Linguist. Proc. Conf.*, pp. 757–770, 2020, doi: 10.18653/v1/2020.coling-main.66.
- [17] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.
- [18] A. Rogers, O. Kovaleva, and A. Rumshisky, "A primer in bertology: What we know about how bert works," *Trans. Assoc. Comput. Linguist.*, vol. 8, pp. 842–866, 2020, doi: 10.1162/tacl\_a\_00349.
- [19] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2017-Octob, pp. 2999–3007, 2017, doi: 10.1109/ICCV.2017.324.
- [20] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognit.*, vol. 91, pp. 216–231, 2019, doi: 10.1016/j.patcog.2019.02.023.
- [21] M. Ren, W. Zeng, B. Yang, and R. Urtasun, "Learning to reweight examples for robust deep learning," *35th Int. Conf. Mach. Learn. ICML 2018*, vol. 10, pp. 6900–6909, 2018.
- [22] E. Fersini, E. Messina, and F. A. Pozzi, "Sentiment analysis: Bayesian Ensemble Learning," *Decis. Support Syst.*, vol. 68, no. January 2015, pp. 26–38, 2014, doi: 10.1016/j.dss.2014.10.004.