Jumistik
E-ISSN: 2964-3953 P-ISSN: 2964-3888 Jurnal Manajemen Informatika, Sistem Informasi dan Teknologi Komputer

Vol.4 No.1 Juni 2025
Hal 450-461
Https://ojs.amiklps.ac.id

SISTEM DETEKSI BERITA HOAKS BERBASIS ALGORITMA *NATURAL LANGUAGE PROCESSING (NLP)* MENGGUNAKAN BERT

Azura Fardhina[1*], Ramadhan Mustaqim Siregar[2], Mika Ria Waty Br Sibarani[3], Irsya Chlara Br Ginting[4], Andre Pratama[5]

[1,2,3,4,5]Fakultas Teknologi dan Ilmu Komputer, Universitas Satya Terra Bhinneka
E-mail: azurafardina0405@gmail.com[*1], rmdhnmustaqim@gmail.com[2], sibaranimika183@gmail.com[3], irsyachlara26@gmail.com[4], andrepratama@satyaterrabhinneka.ac.id[5]

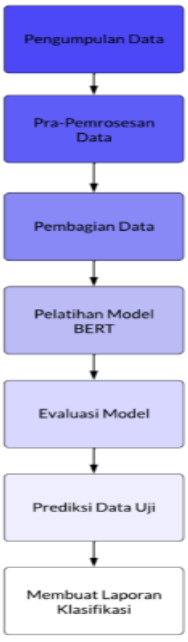
INFO ARTIKEL

Sejarah Artikel
Diterima : 25/05/2025
Direvisi : 28/06/2025
Diterbitkan : 30/06/2025

***Corresponding author**
azurafardina0405@gmail.com

DOI: 10.70247/jumistik.v4i1.156
Https://ojs.amiklps.ac.id

GRAPHICAL ABSTRACT



ABSTRACT

The advancement of digital technology brings both benefits and challenges, one of which is the increasing spread of hoaxes that can trigger conflicts in various sectors such as social, cultural, political, and economic. Hoaxes are unverifiable and often provocative information that spreads rapidly across digital platforms. Indonesian society remains vulnerable to unverified information. Therefore, an artificial intelligence (AI)-based system is needed to automatically detect hoaxes. This study employs the BERT model for its ability to understand word context and perform effective semantic classification through tokenization and transformer architecture. The dataset, sourced from Kaggle, consists of 730 articles: 425 labeled as hoax and 305 as non-hoax. After preprocessing and tokenization, the data was input into the model. BERT was chosen for its strong word representation capabilities trained on a large-scale corpus. Evaluation results show that BERT achieved 98% accuracy, outperforming the previous KNN model which reached 93.33%. These findings demonstrate the effectiveness of the BERT-based approach in detecting digital disinformation.

***Keywords**: Hoax Detection, BERT, Natural Language Processing, Text Classification*

ABSTRAK

Kemajuan teknologi digital membawa manfaat sekaligus tantangan, termasuk meningkatnya penyebaran hoaks yang dapat memicu konflik di berbagai bidang seperti sosial, budaya, politik, dan ekonomi. Hoaks merupakan informasi yang tidak dapat dipastikan kebenarannya, sering bersifat provokatif, dan menyebar cepat melalui platform digital. Masyarakat Indonesia masih rentan terhadap informasi yang belum diverifikasi. Oleh karena itu, diperlukan sistem deteksi otomatis berbasis kecerdasan buatan (AI). Penelitian ini menggunakan model BERT karena kemampuannya dalam memahami konteks kata dan melakukan klasifikasi semantik secara efektif melalui *tokenisasi* dan arsitektur transformer. Dataset yang digunakan berasal dari Kaggle dengan total 730 artikel, terdiri dari 425 hoaks dan 305 non-hoaks. Setelah tahap praproses dan *tokenisasi*, data dimasukkan ke dalam model. Pemilihan BERT didasarkan pada kemampuannya menghasilkan representasi kata yang akurat melalui pelatihan pada korpus berskala besar. Hasil evaluasi menunjukkan bahwa BERT mencapai akurasi sebesar 98%, melampaui model KNN sebelumnya yang hanya memperoleh 93,33%. Temuan ini membuktikan bahwa pendekatan berbasis BERT efektif dalam mendeteksi disinformasi di ruang digital.

Kata kunci: Deteksi Hoaks, BERT, Pemerosesan Bahasa Alami, Klasifikasi Teks

PENDAHULUAN

Perkembangan teknologi digital yang pesat telah menjadikan internet sebagai sarana utama dalam mengakses dan menyebarluaskan informasi. Meskipun memberikan kemudahan dan kecepatan dalam distribusi informasi, hal ini juga membuka celah bagi peredaran informasi yang tidak valid, salah satunya adalah berita bohong atau hoaks [1]. Hoaks merupakan informasi yang sengaja direkayasa untuk menyesatkan pembaca, umumnya ditandai dengan judul provokatif, narasi yang mengandung unsur SARA, ujaran kebencian, serta penggunaan sumber yang tidak kredibel atau bahkan palsu [2].

Fenomena ini menjadi tantangan serius di Indonesia. Tingkat literasi digital masyarakat yang masih rendah menyebabkan kecenderungan menyebarkan informasi tanpa proses verifikasi. Hal ini diperkuat oleh penelitian Rama *et* al. [3] yang menunjukkan bahwa penyebaran berita hoaks dapat memicu disinformasi massal dan berpotensi menimbulkan konflik sosial hingga disintegrasi bangsa. Data dari Dewan Pers juga mengungkapkan bahwa dari 61.800 situs yang mengklaim sebagai portal berita, hanya 1.700 di antaranya yang telah terverifikasi sebagai media resmi [4]. Kondisi ini menunjukkan tingginya potensi penyebaran informasi palsu yang perlu ditangani dengan pendekatan yang lebih canggih dan adaptif.

Dalam menjawab tantangan tersebut, dibutuhkan sistem cerdas yang mampu melakukan deteksi dan klasifikasi berita hoaks secara otomatis. Teknologi *Artificial Intelligence* (AI), khususnya pendekatan *machine learning* dan *deep learning*, menjadi salah satu solusi potensial. Teknologi ini memungkinkan sistem dapat memahami pola data secara mandiri tanpa perlu diprogram secara langsung, serta mampu meningkatkan performa seiring bertambahnya data pelatihan [5]. Ketika dikombinasikan dengan teknik *text mining*, sistem dapat menganalisis dan memahami makna semantik dalam teks guna mengidentifikasi indikasi hoaks [6].

Sejumlah penelitian sebelumnya telah mengimplementasikan berbagai algoritma *machine learning* dalam upaya deteksi hoaks. Ula [7] melakukan perbandingan terhadap empat algoritma, yakni *Multi-Layer Perceptron (MLP)*, *Naive Bayes, Support Vector Machine (SVM)*, dan *Random Forest*, di mana algoritma *Random Forest* memperoleh akurasi tertinggi sebesar 75,37%. Prasetya dan Ferdiansyah [8] menggunakan kombinasi *Naive Bayes* dan *Particle Swarm Optimization* (PSO) untuk mendeteksi berita hoaks terkait COVID-19 dan memperoleh akurasi 86,3%. Pendekatan *deep learning* juga telah diuji, seperti yang dilakukan oleh Kurniawan dan Mustikasari [9] dengan menggunakan kombinasi *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM), yang mencapai akurasi hingga 88%. Sementara itu, Awalina et al. [10] membuktikan bahwa model *Bidirectional Encoder Representations from Transformers* (BERT) memiliki performa terbaik dengan akurasi mencapai 90%, mengungguli model CNN, BiLSTM, dan pendekatan hibrida.

BERT merupakan model *pre-trained* yang dikembangkan oleh *Google Research*, dibangun dengan arsitektur *Transformer* dan dirancang untuk berbagai tugas *Natural Language Processing* (NLP) seperti prediksi kata, *Named Entity Recognition* (NER), serta pemahaman kalimat secara kontekstual [11]. Keunggulan utama BERT terletak pada kemampuannya melakukan representasi kontekstual dua arah (*bidirectional contextual representation*), memungkinkan pemahaman terhadap konteks kata dari sisi kiri dan kanan secara bersamaan. Proses pelatihan BERT terdiri dari dua tahap, yaitu *pre-training* dan *fine-tuning*. Tahap *pre-training* dilakukan secara *unsupervised* menggunakan tugas seperti *Masked Language Modeling* (MLM) dan *Next Sentence Prediction* (NSP) untuk membangun pemahaman terhadap struktur bahasa. Selanjutnya, pada tahap *fine-tuning*, model dilatih menggunakan data berlabel sesuai tugas tertentu seperti klasifikasi hoaks [12].

Penelitian ini bertujuan untuk menerapkan model BERT dalam mendeteksi berita hoaks berbahasa Indonesia secara otomatis. *Dataset* yang digunakan diperoleh dari platform *Kaggle*, terdiri dari 730 artikel yang terbagi atas 425 artikel hoaks dan 305 artikel non-hoaks. Selain mengevaluasi performa dari model BERT itu sendiri, penelitian ini juga membandingkan hasilnya dengan model klasik seperti *K-Nearest Neighbors* (KNN) dan *Naive Bayes* dari penelitian sebelumnya. Berdasarkan hasil studi terdahulu, model KNN mampu mencapai akurasi sebesar 93,33% dalam klasifikasi berita hoaks. Namun, pendekatan tersebut masih memiliki keterbatasan dalam menangkap konteks semantik secara mendalam antar-kalimat [13]. Dengan akurasi tinggi dan kemampuan memahami konteks semantik yang mendalam, pendekatan berbasis BERT diharapkan dapat menjadi dasar pengembangan sistem deteksi hoaks yang efisien dan dapat diintegrasikan pada berbagai platform digital. Semoga temuan dalam studi ini dapat turut andil dalam mendukung upaya penanggulangan disinformasi di Indonesia. Dengan akurasi tinggi dan kemampuan memahami konteks semantik yang mendalam, pendekatan berbasis BERT diharapkan dapat menjadi dasar pengembangan sistem deteksi hoaks yang efisien dan dapat diintegrasikan pada berbagai platform digital.

TINJAUAN PUSTAKA

Sistem Deteksi Berita Hoaks

Penyebaran berita bohong atau hoaks semakin marak di era digital dan menjadi tantangan serius dalam menjaga keakuratan informasi publik. Deteksi hoaks didefinisikan sebagai proses identifikasi dan klasifikasi informasi palsu atau menyesatkan yang tersebar, khususnya melalui media digital [14]. Berita hoaks memiliki potensi merusak opini publik, menimbulkan kepanikan, serta menciptakan konflik sosial. Oleh karena itu, deteksi otomatis terhadap

berita hoaks sangat penting sebagai bentuk perlindungan terhadap masyarakat dari misinformasi.

Dalam Membangun sistem deteksi hoaks, diperlukan metode yang mampu mengenali pola bahasa dan konteks informasi secara mendalam. Salah Satu Pendekatan yang berkembang pesat dalam beberapa tahun terakhir adalah penggunaan teknik pemrosesan bahasa alami (*Natural Language Processing)*.

Natural Language Processing (NLP)

Natural Language Processing (NLP), sebuah cabang dari kecerdasan buatan yang memfokuskan pada hubungan antara komputer dan bahasa manusia [15]. NLP memungkinkan mesin untuk menganalisis, menginterpretasikan, dan membentuk bahasa alami, baik tertulis maupun lisan. Dalam konteks deteksi hoaks, NLP digunakan untuk memahami struktur, konteks, dan makna kalimat dalam berita.

Salah satu pendekatan NLP yang digunakan adalah klasifikasi teks. Proses ini melibatkan pengelompokan dokumen atau kalimat ke dalam kategori tertentu berdasarkan kontennya. Metode ini banyak diaplikasikan untuk pengelompokan email spam, analisis sentimen, klasifikasi topik, hingga deteksi berita palsu [16]. Klasifikasi dilakukan melalui tahap *preprocessing*, ekstraksi fitur, pelatihan model, dan evaluasi, dengan algoritma seperti *Naive Bayes*, SVM, *Random Forest*, hingga model *deep learning* berbasis transformer.

Model BERT

Model BERT (*Bidirectional Encoder Representations from Transformers)* merupakan model berbasis *transformer* yang dikembangkan oleh *Google* dan dipublikasikan di tahun 2018 [17]. Keunggulan BERT dibandingkan model sebelumnya terletak pada kemampuannya memahami konteks kata dari dua arah sekaligus, yaitu dari kiri ke kanan dan sebaliknya. BERT dilatih menggunakan dua metode utama: *Masked Language Modeling* (MLM) dan *Next Sentence Prediction* (NSP), sehingga mampu mempelajari hubungan antar-kalimat secara mendalam [5].

Penelitian Terdahulu (State of the Art)

Penelitian mengenai deteksi berita hoaks dengan pendekatan *Natural Language Processing* (NLP) telah banyak dilakukan, khususnya dengan memanfaatkan metode berbasis *transformer* seperti BERT. Dalam studi oleh Awalina *et* al. [10], model BERT berhasil menunjukkan performa terbaik dalam klasifikasi berita hoaks berbahasa Indonesia, mengungguli model lain seperti CNN dan BiLSTM dengan akurasi mencapai 90%. Hasil tersebut menegaskan keunggulan representasi kontekstual dua arah yang dimiliki BERT dalam memahami makna mendalam dari teks.

Girinoto *et* al. [18] mengembangkan sistem deteksi *clickbait* pada judul berita menggunakan *Multilingual* BERT, yang secara arsitektural masih sama dengan BERT namun dilatih dalam banyak bahasa termasuk Bahasa Indonesia. Model ini diterapkan pada ekstensi *Chrome* dan diuji pada *dataset* berisi 6.632 headline dari berbagai kategori berita. Hasilnya, akurasi klasifikasi clickbait mencapai 92%, menandakan efektivitas model BERT dalam memahami konteks kalimat pendek dan ambigu seperti judul berita.

Penelitian lain oleh Rendragraha *et* al. [19] menggunakan BERT dan *Multilingual* BERT untuk deteksi ujaran kebencian dalam komentar berita *daring*. Meskipun bukan fokus pada hoaks, pendekatan ini menunjukkan kemampuan BERT dalam memproses data teks yang beragam gaya bahasanya, termasuk dalam bentuk komentar dan narasi pendek. Skor F1 makro yang diperoleh sebesar 54% membuktikan bahwa BERT cukup baik digunakan pada data sosial yang lebih dinamis.

Putri [20] dalam penelitiannya menggunakan *Multilingual* BERT untuk analisis sentimen pada ulasan film berbahasa Inggris. Penelitian ini menunjukkan bahwa BERT mampu dikustomisasi melalui *fine-tuning* pada berbagai domain teks, dan berhasil mencapai akurasi sebesar 73%. Meskipun domainnya berbeda, model dan metode yang digunakan serupa, dan menunjukkan bahwa BERT fleksibel untuk berbagai jenis klasifikasi teks.

Kurniawan dan Mustikasari [9] membandingkan dua pendekatan *deep learning*, yaitu CNN dan LSTM dalam klasifikasi berita hoaks berbahasa Indonesia. Dataset yang digunakan terdiri dari 1.786 artikel dan hasilnya menunjukkan CNN unggul dengan akurasi sebesar 88%, dibandingkan LSTM dengan 84%. Meskipun tidak menggunakan BERT, hasil ini menjadi pembanding awal terhadap pendekatan *sequential neural network*.

Di sisi lain, Penelitian oleh Rahmawati [21] memanfaatkan metode *Support Vector Machine* (SVM) serta *Random Forest* untuk mendeteksi berita hoaks dari situs *TurnbackHoax.id*. Dari 4.551 data, model SVM menghasilkan akurasi 83% dan *recall* sebesar 99% pada kelas hoaks. Penelitian ini membuktikan bahwa model konvensional masih dapat memberikan hasil kompetitif, namun belum mampu menangkap konteks bahasa secara mendalam sebagaimana dilakukan oleh BERT. BERT dilatih untuk memahami konteks secara dua arah, berbeda dari pendekatan sebelumnya yang satu arah. Arsitektur ini mampu menangkap hubungan semantik dan sintaksis antar-kalimat secara lebih akurat.

Dalam penerapannya untuk Bahasa Indonesia, penelitian oleh Delvin *et* al. [22] menunjukkan bahwa BERT yang dilatih melalui *fine-tuning* dengan dataset lokal mampu memberikan hasil yang mendekati model lokal seperti IndoBERT. Ini membuka kemungkinan bahwa BERT standar dapat diadaptasi tanpa perlu menggunakan model multibahasa ataupun turunan lokal.

Menurut studi oleh Sun *et* al. [23], BERT memiliki kemampuan generalisasi yang baik dan dapat digunakan dalam berbagai jenis klasifikasi teks seperti

entitas, relasi, sentimen, hingga pendeteksian fakta. Ini memperkuat argumen bahwa model ini cukup andal untuk tugas klasifikasi berita hoaks dengan bahasa non-Inggris, termasuk Bahasa Indonesia, asalkan dilakukan pelatihan (*fine-tuning*) dengan data yang sesuai.

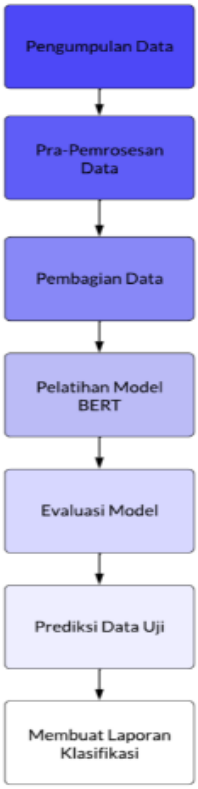
Dari beberapa penelitian di atas, terlihat bahwa model BERT memberikan keunggulan yang signifikan dalam memahami konteks semantik, baik untuk klasifikasi hoaks, clickbait, maupun sentimen. Berbeda dengan pendekatan tradisional atau model *deep learning* berbasis CNN/LSTM yang bekerja secara sekuensial, BERT mampu memahami kata berdasarkan konteks dua arah. Hal ini memungkinkan deteksi hoaks dilakukan dengan akurasi lebih tinggi dan fleksibilitas yang lebih besar terhadap berbagai jenis teks.

Kebaruan (*state of the art*) dari penelitian ini terletak pada penerapan model BERT murni (bukan IndoBERT atau versi multilingual), untuk mendeteksi berita hoaks berbahasa Indonesia. Fokus penelitian ini adalah mengadaptasi model BERT standar melalui proses *fine-tuning* terhadap *dataset* lokal, sehingga dapat dimanfaatkan sebagai sistem cerdas dalam menghadapi maraknya disinformasi secara real-time.

METODOLOGI PENELITIAN

Tahapan Penelitian

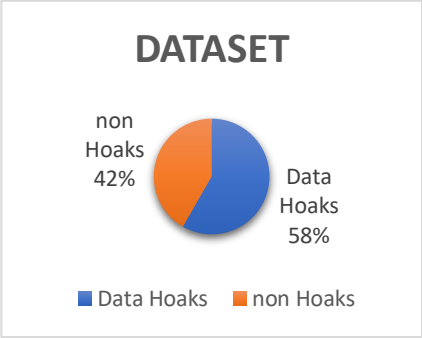
Penelitian ini menggunakan pendekatan eksperimental dengan menerapkan algoritma BERT (*Bidirectional Encoder Representations from Transformers*) dalam mendeteksi berita hoaks berbahasa Indonesia. BERT merupakan model bahasa berbasis *deep learning* yang memiliki keunggulan dalam memahami konteks dua arah, menjadikannya cocok untuk tugas klasifikasi teks seperti deteksi berita palsu. Metode penelitian ini terdiri dari beberapa tahapan utama yang saling terintegrasi: pengumpulan data, pra-pemrosesan data, *tokenisasi*, pelatihan model, evaluasi performa, serta prediksi dan pelaporan hasil klasifikasi seperti yang disajikan pada Gambar 1 berikut.



Gambar 1. Tahapan Metode Penelitian
Sumber: Hasil Olah Data

Berikut adalah penjelasan rinci dari setiap tahapan metode penelitian pada gambar 1.

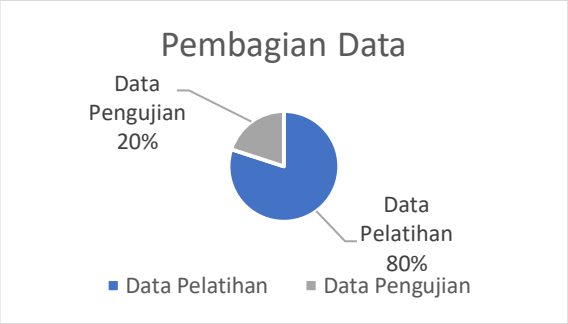
1. **Pengumpulan Data**: Data dikumpulkan dari satu dataset publik yang tersedia di platform *Kaggle.com*. Dataset ini berisi artikel berita dalam bahasa Indonesia yang telah diklasifikasikan sebagai hoaks dan non-hoaks. Komposisi data dapat dilihat pada Gambar 2 berikut.



Gambar 2. Komposisi Pembagian Data
Sumber: Hasil Olah Data

2. **Pra-Pemrosesan Data**: *Dataset* yang digunakan bernama (Data_Hoaks_2023.csv). Data ini diproses dengan penggabungan kolom judul dan konten, dilabeli dalam format biner, serta dilakukan *tokenisasi* menggunakan *tokenizer* dari model BERT.

3. **Pembagian Data**: Setelah dilakukan tokenisasi, data dibagi menjadi dua bagian, yaitu 80% untuk data pelatihan (*training set*) dan 20% untuk data pengujian (*test set*), dengan metode stratifikasi untuk menjaga keseimbangan distribusi kelas.



Gambar 3. Pembagian Data
Sumber: Hasil Olah Data

4. **Pelatihan Model BERT**: Model klasifikasi BERT diinisialisasi menggunakan arsitektur *bert-base-uncased*. Proses pelatihan dilakukan selama 4 *epoch* dengan menggunakan *optimizer* AdamW dan fungsi *loss CrossEntropyLoss* (*default* dari *BertForSequenceClassification*).
5. **Evaluasi Model**: Evaluasi dilakukan terhadap data validasi selama proses pelatihan untuk memantau performa dan menghindari *overfitting*. Evaluasi ini mencakup penghitungan metrik performa utama.
6. **Prediksi Data Uji**: Model yang telah dilatih digunakan untuk mengklasifikasikan berita pada data uji ke dalam dua kelas: hoaks dan non-hoaks.
7. **Pembuatan Laporan Klasifikasi**: Hasil prediksi dibandingkan dengan label asli dari data uji, dan dievaluasi menggunakan metrik klasifikasi seperti presisi, *recall*, *F1-score*, dan akurasi guna mengukur performa keseluruhan model.

Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari sumber publik dan telah diunggah ke *Google Drive*. *Dataset* tersebut terdiri dari 730 artikel dalam format CSV yang memuat tiga kolom: Judul, Konten, dan Label. Kolom Label mengindikasikan apakah suatu artikel tergolong berita hoaks (1) atau bukan hoaks (0). Dari total data, 425 artikel diklasifikasikan sebagai hoaks dan 305 sebagai non-hoaks. Beberapa entri hoaks mengandung judul provokatif dan konten yang tidak diverifikasi, seperti "JADI BURONAN KPK!! SRI MULYANI KABUR KE LUAR NEGERI TAKUT DI PENJARA" atau "Kuburan Ferdy Sambo Keluar Api dan Ular". *Dataset* ini dianggap sesuai karena mengandung variasi topik dan gaya bahasa yang relevan untuk pengujian kemampuan model dalam mendeteksi hoaks.

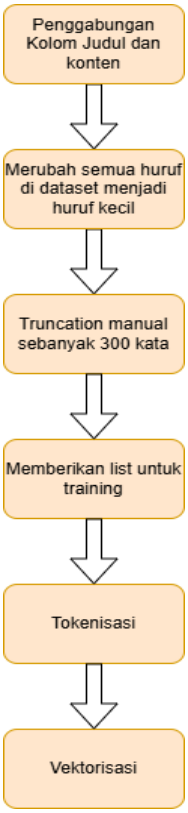
Pemilihan *dataset* dilakukan berdasarkan karakteristiknya yang :

a. Berasal dari kasus nyata dan kontekstual,
b. Ditulis dalam bahasa Indonesia,
c. Mewakili struktur umum berita *daring*,
d. Sudah memiliki label klasifikasi biner yang valid.

Menurut Agung *et* al. [12], struktur *dataset* yang kaya akan variasi konten dan panjang teks menjadi kunci keberhasilan pelatihan model berbasis NLP. Oleh karena itu, struktur *dataset* ini sangat sesuai dengan arsitektur BERT yang memerlukan pemahaman kontekstual dua arah.

Analisis Data

Sebagai bagian dari proses penelitian, dilakukan tahapan pra-pemrosesan data untuk mempersiapkan dataset sebelum digunakan pada model BERT. Gambar 4 berikut menunjukkan alur tahapan pra-pemrosesan data yang telah dirancang dalam penelitian ini.



Gambar 4. Tahapan Pra-Pemrosesan Data
Sumber: Hasil Olah Data

Berikut adalah penjelasan rinci dari gambar 4 tentang setiap tahap pra-pemrosesan data. Langkah pra-pemrosesan dilakukan untuk menyederhanakan dan mempersiapkan data agar sesuai dengan format *input* model BERT. Proses ini mencakup:

1. **Penggabungan Kolom Judul dan Konten** Tahap pertama dilakukan dengan menggabungkan kolom Judul dan Konten pada *dataset* menjadi satu *string* teks baru bernama *teks_input* untuk memberikan konteks yang lebih luas kepada model. Format penyatuannya ditulis sebagai (Judul: Konten).
2. **Konversi Seluruh Teks ke Huruf Kecil (*Lowercasing*)** Setelah penggabungan, seluruh teks dikonversi

menjadi huruf kecil untuk menjaga konsistensi *tokenisasi* dan mengurangi variasi kata akibat perbedaan kapitalisasi. Langkah ini dianggap sebagai praktik terbaik dalam pengolahan bahasa alami sebagaimana dikemukakan oleh Satya (2023) bahwa transformasi ke huruf kecil dapat membantu menyederhanakan struktur teks dan meningkatkan efisiensi model klasifikasl [24].

3. ***Truncation* Manual Sebanyak 300 Kata**
   Untuk menghindari *input* teks yang terlalu panjang dan melampaui batas maksimal *token* BERT (512 *token*), dilakukan pemangkasan manual konten sebanyak 300 kata pertama. Pemotongan ini diprioritaskan karena bagian awal dari artikel berita umumnya memuat informasi paling relevan. *Truncation* juga membantu mengoptimalkan waktu pemrosesan dan mencegah terjadinya error dalam *input* model.
4. **Pemberian Format List untuk Proses *Training***
   Selanjutnya, seluruh teks dan label dikonversi ke dalam bentuk list. Ini diperlukan karena fungsi *tokenizer* dan DataLoader dalam kerangka kerja *PyTorch* bekerja dengan format *iterable*.
5. **Tokenisasi dengan BERT *Tokenizer***
   Tokenisasi dilakukan menggunakan *BertTokenizer* dari pustaka Huggingface *Transformers* dengan model dasar *bert-base-uncased*. *Tokenisasi* melibatkan proses padding dan *truncation* secara otomatis, di mana teks dikonversi ke dalam *token-token* numerik (*input_ids*) dan *attention_mask*. Ini memungkinkan BERT membaca dan memahami teks dalam representasi numerik sesuai konteks dua arah (*bidirectional*).
6. ***Vektorisasi***
   Token hasil tokenisasi kemudian diubah menjadi vektor numerik yang dapat diteruskan ke model BERT dalam proses pelatihan. Vektorisasi ini merupakan representasi akhir dari teks dalam format tensor, yang mengandung informasi posisi, relasi kata, dan konteks kalimat, sesuai dengan mekanisme encoder Transformer.

Setiap tahapan pra-pemrosesan ini dirancang untuk memastikan bahwa model BERT menerima *input* yang konsisten, bersih, dan informatif. Menurut Zahratul *et* al. (2025), kualitas proses pra-pemrosesan memiliki kontribusi signifikan terhadap keberhasilan klasifikasi berbasis teks, karena tahap ini menentukan struktur representasi awal yang digunakan oleh model dalam membangun pemahaman semantik [25].

Tahapan Tokenisasi *dan Pembagian Data*

Tokenisasi dilakukan menggunakan *BertTokenizer* dari pustaka *Huggingface Transformers* dengan model dasar *bert-base-uncased*. Setiap teks diproses menjadi *token*, diberi *padding*, dan di-*truncate* hingga maksimum 512 *token*. *Token* ini kemudian dikonversi menjadi *input_ids* dan attention_mask sebagai *input* utama ke model BERT.

Setelah *tokenisasi*, data dibagi menjadi dua subset, yaitu 80% data untuk pelatihan dan 20% data untuk pengujian, dengan metode stratifikasi berdasarkan label agar distribusi kelas tetap proporsional.

Tahapan Pelatihan Model

Model yang digunakan adalah *BertForSequenceClassification* dengan dua kelas *output*. Proses pelatihan mencakup:

1. Inisialisasi model dengan bobot *pre-trained* dari *bert-base-uncased*,
2. Penyesuaian *optimizer* menggunakan Adam dengan *learning rate* sebesar 5e-5,
3. Pelatihan selama 4 *epoch* menggunakan *DataLoader* dengan *batch size* 8,
4. Evaluasi *loss* dilakukan setiap batch menggunakan fungsi *CrossEntropyLoss* bawaan dari model.

Pelatihan dijalankan di perangkat GPU jika tersedia, menggunakan torch.device("*cuda*"). Proses ini mengikuti metode pelatihan BERT seperti dijelaskan oleh Bayu et al. [26], yang menekankan pentingnya representasi konteks dua arah dalam memahami teks.

Tahapan Evaluasi dan Prediksi

Sebelum memasuki proses evaluasi, perlu dipahami bahwa salah satu alasan utama dari penggunaan model ini adalah untuk mensimulasikan klasifikasi konten sebagaimana yang terjadi di platform media sosial, di mana pengguna sering kali hanya membaca judul berita tanpa memeriksa keseluruhan kontennya. Oleh sebab itu, yang menjadi fokus utama adalah pada bagian judul berita. Variabel Label sebagai target klasifikasi ditambahkan secara manual, dan *dataset* dibagi menjadi bagian pelatihan (80%) dan pengujian (20%).

Setelah pelatihan selesai, data pengujian diperluas dengan menambahkan kolom hasil prediksi model. Evaluasi kemudian dilakukan dengan memanfaatkan *Confusion Matrix* sebuah tabel yang menggambarkan perbandingan antara hasil klasifikasi model dan label sebenarnya. Terdapat empat kemungkinan hasil:

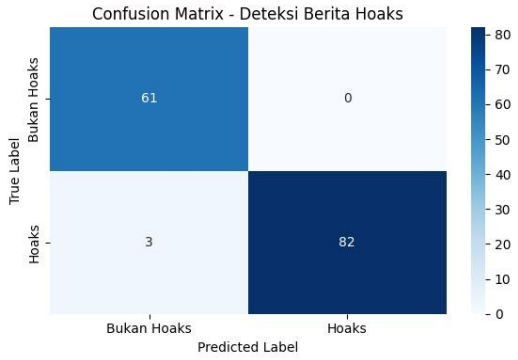
Tabel 1. *Visualisasi Confusion Matrix*

		Prediksi	
		Hoaks	Non–Hoaks
Aktual	Hoaks	*True Positive (TP)*	*False Negative (FN)*
	Non-Hoaks	*False Positif (FP)*	*True Negative (TN)*

Sumber: Assyfa et al. [27]

Confusion Matrix membantu memvisualisasikan kesalahan klasifikasi, sehingga kita bisa memahami apakah model cenderung mengklasifikasikan berita dengan benar, terlalu banyak *false positive*, atau justru banyak *false negative*. Gambar 5 berikut

merupakan hasil *Confusion Matrix* dari model yang telah dilatih:



Gambar 5. Hasil *Confusion Matrix*
Sumber: Hasil Olah Data

Evaluasi performa model bertujuan untuk mengukur dan menguji sejauh mana kemampuan model dalam mengklasifikasikan teks ke dalam kategori hoaks dan non-hoaks dengan benar. Evaluasi dilakukan terhadap data uji yang tidak pernah dilihat oleh model sebelumnya. Metrik yang digunakan dalam proses evaluasi adalah *Accuracy, Precision, Recall, F1-Score, Support, Macro Average, Weighted Average*.

Adapun metrik yang digunakan dalam proses evaluasi meliputi:

1. Akurasi (*Accuracy*)

Akurasi adalah salah satu ukuran evaluasi dalam *machine learning* yang digunakan untuk menilai sejauh mana model mampu dalam mengklasifikasikan data dengan benar [28]. Akurasi ini menghitung sebagai rasio antara jumlah prediksi yang benar (baik positif maupun negatif) terhadap total seluruh prediksi yang dilakukan. Secara matematis dapat dituliskan sebagai:

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

Keterangan:

- **TP (*True Positive*)**: jumlah data berlabel positif yang diklasifikasikan dengan benar oleh model.
- **TN (*True Negative*)**: jumlah data berlabel negatif yang diklasifikasikan dengan benar.
- **FP (*False Positive*)**: jumlah data negatif yang keliru diklasifikasikan sebagai positif.
- **FN (*False Negative*)**: jumlah data positif yang salah diklasifikasikan sebagai negatif.

Nilai akurasi yang tinggi menunjukkan bahwa model secara umum berperforma yang baik dalam mengklasifikasikan data secara umum. Namun, akurasi dapat menyesatkan jika digunakan pada *dataset* yang tidak seimbang karena tidak mencerminkan peforma model terhadap distribusi kelas yang tidak merata.

2. Presisi (*Precision*)

Rasio **presisi** merupakan ukuran evaluasi yang menunjukkan sejauh mana model mampu mengenali berita hoaks secara akurat. Presisi (P) dapat dirumuskan sebagai berikut:

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

Keterangan:

- **TP (*True Positive*)**: Model memprediksi data sebagai hoaks, dan memang hoaks.
- **FP (*False Positive*)**: Model memprediksi data sebagai hoaks, namun ternyata bukan.

Presisi menggambarkan proporsi data yang benar – benar hoaks dari seluruh data yang diprediksi sebagai hoaks (Jumlah TP dan FP). Nilai presisi yang tinggi menunjukkan bahwa model lebih akurat dalam mengidentifikasi berita hoaks dan menghasilkan sedikit kesalahan positif hoaks. Presisi tidak cukup untuk mengevaluasi kinerja model, karena tidak terdeteksi oleh model (*False Negative*). Presisi sering digunakan bersama dengan matriks lain untuk memberikan gambaran tentang kinerja model.

3. *Recall* (Sensitivitas)

Recall adalah metrik evaluasi yang digunakan untuk mengukur seberapa baik model dalam mengenali seluruh data yang memang termasuk dalam kelas positif, yaitu berita hoaks dalam konteks penelitian ini. *Recall* sering juga disebut sebagai *Sensitivity* atau *True Positive Rate (TPR)*.

Recall menjadi penting dalam kasus di mana konsekuensi dari tidak mendeteksi suatu kelas positif (*False Negative*) lebih serius daripada salah mendeteksi kelas positif (*False Positive*) [29]. Misalnya, dalam konteks deteksi berita hoaks, jika suatu berita hoaks tidak terdeteksi dan dianggap sebagai berita asli, maka informasi palsu tersebut dapat tersebar dan menimbulkan dampak sosial yang signifikan. *Recall* dapat dirumuskan dengan persamaan berikut:

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

Keterangan:

- ***Recall***: Perbandingan antara TP (*True Positive*) dengan banyaknya data yang sebenarnya positif.
- **TP (*True Positive*)**: Jumlah berita hoaks yang berhasil dideteksi dengan benar sebagai hoaks oleh model.
- **FN (*False Negative*)**: Jumlah berita hoaks yang gagal dikenali oleh model dan justru diprediksi sebagai berita asli.

Recall menunjukkan proporsi berita hoaks yang berhasil diidentifikasi dari seluruh berita yang memang sebenarnya adalah hoaks. Semakin tinggi nilai *recall*, maka semakin sedikit berita hoaks yang lolos dari deteksi (FN rendah) [30].

4. F1-Score

F1-Score adalah metrik yang menggabungkan presisi dan *recall* dalam satu nilai rata-rata harmonik [31]. Nilai F1-score tinggi menunjukkan bahwa model berhasil menjaga keseimbangan antara presisi dan *recall*. Ini sangat bermanfaat ketika kita berhadapan dengan *dataset* yang tidak seimbang. Dalam konteks deteksi hoaks, F1-score memberikan gambaran menyeluruh tentang kemampuan model dalam mendeteksi berita palsu dengan akurat dan menyeluruh.

F1-Score (f1) dapat dirumuskan dengan persamaan berikut:

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

Keterangan:

- **F1-Score**: perbandingan rata-rata *precision* dan *recall* yang dibobotkan.
- ***Recall***: perbandingan antara TP (*True Positive*) dengan banyaknya data yang sebenarnya positif
- ***Precision***: perbandingan antara TP (*True Positive*) dengan banyaknya data yang diprediksi positif.

Dengan demikian, F1-score merepresentasikan rata-rata harmonik antara nilai presisi dan *recall*. Metrik ini menjadi sangat bermanfaat dalam kondisi di mana diperlukan keseimbangan antara kemampuan model dalam mengidentifikasi kasus positif secara benar (presisi) dan mendeteksi seluruh kasus positif yang ada (*recall*). Pada konteks deteksi berita hoaks, F1-score membantu mengevaluasi apakah model mampu secara efektif menghindari kesalahan positif (mengklasifikasikan berita valid sebagai hoaks) sekaligus tidak melewatkan berita hoaks yang sebenarnya. Karena keduanya sama-sama penting, F1-score menjadi pilihan metrik evaluasi yang ideal untuk menilai kinerja model dalam tugas klasifikasi semacam ini.

5. *Support*

Support adalah jumlah sampel aktual dalam setiap kelas pada data pengujian. Dalam konteks klasifikasi biner seperti deteksi berita hoaks, *support* menunjukkan berapa banyak data yang benar-benar termasuk dalam kelas *hoaks* maupun *non-hoaks*. Nilai ini penting untuk mengetahui proporsi kelas dalam *dataset*, yang berpengaruh dalam interpretasi metrik evaluasi lain seperti *recall* dan *precision*.

6. *Macro Average*

Macro Average adalah teknik untuk menghitung rata-rata dari metrik evaluasi seperti *precision*, *recall*, dan F1-score dengan memberikan bobot yang sama kepada setiap kelas. Metode ini pertama-tama menghitung metrik evaluasi untuk masing-masing kelas [32], lalu menjumlahkan dan membaginya dengan jumlah kelas. Tidak memperhitungkan jumlah data (*support*) dalam setiap kelas, sehingga cocok digunakan jika distribusi kelas seimbang.

Macro Avg dapat dirumuskan dengan persamaan berikut:

$$Macro\ Avg = \frac{\sum_{i=1}^{n} M_i}{n}$$

Keterangan:

- ***Macro avg***: Rata-rata sederhana dari matrik evaluasi
- **Σ Nilai matrik evaluasi semua kelas**: Jumlah total nilai matrik evaluasi semua kelas
- **M_i**: Nilai metrik (*precision*, *recall*, atau F1-score) pada kelas ke-i
- ***n***: Jumlah kelas

7. Weighted Average

Weighted average menghitung metrik evaluasi dengan mempertimbangkan proporsi data (*support*) di masing-masing kelas. Setiap metrik dikalikan dengan jumlah data dalam kelas tersebut, lalu dijumlahkan dan dibagi dengan total seluruh data. Ini berguna ketika distribusi kelas tidak seimbang, karena menghasilkan evaluasi yang mencerminkan performa model secara keseluruhan.

Weighted average dapat dirumuskan dengan persamaan berikut:

$$Weighted\ Avg = \frac{\sum_{i=1}^{n}(w_i \times M_i}{\sum_{i=1}^{n} w_i}$$

Keterangan:

- ***Weighted Avg***: Rata-rata tertimbang dari matrik evaluasi.
- **w_i**: Bobot untuk kelas
- **M_i**: nilai metrik (seperti presisi, *recall*, F1-Score) untuk kelas i.

Penggunaan *weighted average* dalam kode, memastikan bahwa performa model dievaluasi secara adil berdasarkan distribusi aktual data hoaks dan non-hoaks yang telah disiapkan [33].

HASIL DAN PEMBAHASAN

Hasil Evaluasi Model

Dalam penelitian ini, model *BERT* dengan varian *bert-base-uncased* diterapkan untuk melakukan klasifikasi teks dalam mendeteksi konten berita hoaks berbahasa Indonesia. Proses pelatihan dijalankan selama empat *epoch* menggunakan algoritma optimasi *Adam* dengan nilai *learning rate* sebesar 5e-5. *Dataset* yang digunakan mencakup total 730 artikel, yang terdiri dari 425 berita terindikasi hoaks dan 305 berita yang valid atau non-hoaks. Untuk menjaga representasi kelas yang seimbang, data dibagi secara stratifikasi menjadi dua bagian: 80% dialokasikan untuk pelatihan dan 20% untuk pengujian.

Tahapan *tokenisasi* dilakukan menggunakan *BertTokenizer* dari pustaka *Transformers* milik

Huggingface. Dalam proses ini, setiap entri teks dikonversi menjadi *token* hingga batas maksimum 512 *token* setiap sampel. Kemudian, pelatihan dilakukan menggunakan *BertForSequenceClassification*, dan model hasil pelatihan disimpan dalam direktori saved_model/bert-hoax untuk keperluan prediksi dan inferensi di tahap selanjutnya.

Evaluasi performa model dilakukan menggunakan data latih sebesar 80% dari *dataset*, di mana data ini digunakan selama pelatihan agar hasil evaluasi benar-benar mencerminkan kemampuan generalisasi model. Model yang telah dilatih selama 4 *epoch* ini kemudian diuji untuk melihat kemampuannya dalam mengklasifikasikan teks ke dalam dua kategori: hoaks dan non-hoaks.

Selama proses pelatihan, performa model terus menunjukkan perbaikan dengan menurunnya nilai *loss* di setiap *epoch*, yang mengindikasikan bahwa model semakin memahami pola data. Adapun hasil *loss* pada setiap *epoch* dapat dilihat pada Tabel 2 berikut.

Tabel 2. Hasil *Loss* setiap *Epoch*

Epoch	Jumlah Batch	*Loss*
0	73/73	0.247
1	73/73	0.032
2	73/73	0.0069
3	73/73	0.111

Hasil ini menunjukkan bahwa model mencapai konvergensi yang baik dan stabil pada akhir pelatihan. Evaluasi selanjutnya difokuskan pada pengukuran metrik-metrik seperti akurasi, presisi, *recall*, dan F1-score untuk menilai efektivitas model dalam mendeteksi hoaks secara akurat dan andal. Hasil evaluasi berdasarkan klasifikasi dua kelas: hoaks dan non-hoaks disajikan pada Tabel 3 berikut.

Tabel 3. Hasil Evaluasi Model BERT

Label	*Precision*	*Recall*	*F1-Score*	*Support*
Non-Hoaks (0)	0.96	0.96	0.96	61
Hoaks (1)	0.98	0.98	0.98	85
Accuracy			**0.9863**	146
Macro Avg	0.98	0.98	0.98	146
Weighted Avg	0.98	0.98	0.98	146

Setelah keseluruhan proses pelatihan selesai, model dievaluasi menggunakan data uji. Model berhasil memperoleh skor akurasi yang sangat tinggi, yakni 98%, yang menunjukkan bahwa dari seluruh data uji yang tersedia, 98% berhasil diklasifikasikan dengan benar. Performa tersebut menunjukkan efisiensi dan keakuratan model dalam mendeteksi informasi yang valid maupun hoaks.

Performa unggul ini didukung oleh kemampuan BERT dalam memahami konteks kalimat secara dua arah (*bidirectional*), berbeda dengan pendekatan satu arah seperti pada LSTM atau SVM. Arsitektur *transformer* yang digunakan oleh BERT memungkinkan model menangkap hubungan semantik dan sintaksis antar kata secara lebih akurat, yang menjadi keunggulan utama dalam tugas deteksi hoaks berbasis teks.

Pendekatan ini juga sejalan dengan penelitian yang dilakukan oleh Putri [20], yang menggunakan Multilingual BERT untuk analisis sentimen pada ulasan film berbahasa Inggris. Meskipun domain aplikasinya berbeda, yakni pada klasifikasi sentimen, model dan metode yang digunakan memiliki kesamaan, yaitu *fine-tuning* BERT pada data spesifik. Hasil penelitian tersebut menunjukkan bahwa BERT tidak hanya efektif dalam tugas klasifikasi tertentu, tetapi juga fleksibel untuk berbagai jenis pemrosesan teks, termasuk deteksi hoaks seperti yang dilakukan dalam penelitian ini.

Pembahasan

Hasil penelitian menunjukkan bahwa model BERT sangat efektif dalam mengklasifikasikan berita hoaks berbahasa Indonesia, dengan akurasi mencapai 98% serta presisi dan *recall* yang stabil di kedua kelas. Hal ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik dan tidak bias terhadap kelas tertentu. Meskipun hanya menggunakan gabungan teks dari judul dan isi berita sebagai *input*, model mampu memahami konteks semantik secara mendalam berkat proses *tokenisasi* dan embedding dari arsitektur bert-base-uncased. Kemampuan ini mendukung deteksi pola hoaks secara akurat.

Sebagai pembanding performa, penelitian oleh Hendriyanto (2022) menggunakan algoritma *K-Nearest Neighbor* (KNN) pada judul berita dengan hasil akurasi maksimal sebesar 93,33%, *precision* 100%, *recall* 80%, dan f1-score 88,89%. Hasil tersebut digunakan sebagai acuan awal dalam menilai sejauh mana model lain dapat meningkatkan performa klasifikasi. Pendekatan KNN pada penelitian tersebut hanya memanfaatkan representasi vektor berbasis *TF-IDF*, sehingga kemampuan dalam menangkap relasi semantik antar kata masih terbatas dan lebih bergantung pada kemiripan kata secara permukaan.

Berbeda dengan penelitian tersebut, penelitian ini memanfaatkan gabungan judul dan isi berita, serta tokenisasi kontekstual dengan pretrained BERT. Pendekatan ini memungkinkan model untuk menangkap relasi kata dan makna kalimat secara lebih mendalam dibandingkan perhitungan jarak vektor pada KNN. Integrasi fungsi prediksi ke dalam antarmuka pengguna menjadikan model ini sangat aplikatif untuk mendukung literasi digital dan memerangi penyebaran informasi palsu secara real-time di masyarakat. Keunggulan model ini juga didukung oleh studi Devlin *et* al. [34] yang menegaskan performa superior BERT dalam klasifikasi teks, serta pendapat Tharwat [35] yang menyarankan

penggunaan kombinasi metrik evaluasi untuk penilaian yang menyeluruh. Jika dibandingkan dengan metode lain seperti *Random Forest*, *Naive Bayes*, atau KNN (Ula, 2020; Prasetya & Ferdiansyah, 2022), BERT menunjukkan performa yang lebih unggul. Penelitian oleh Awalina *et* al. (2021) dan Rohmah & Maharani (2022) juga memperkuat bahwa BERT, termasuk IndoBERT, memiliki kemampuan representasi bahasa Indonesia yang kuat.

Selain itu, BERT mampu mengatasi ketidakseimbangan data tanpa perlakuan khusus seperti *oversampling*, karena tokenizer-nya sudah dilatih pada korpus yang sangat besar. Dengan performa dan fleksibilitas ini, model BERT sangat potensial untuk diterapkan pada sistem verifikasi otomatis, aplikasi *mobile* anti-hoaks, hingga integrasi ke media sosial.

KESIMPULAN DAN SARAN

Berdasarkan rangkaian tahapan yang telah dilakukan dalam penelitian ini, mulai dari pengumpulan data, pelatihan model, hingga evaluasi, dapat disimpulkan beberapa hal penting sebagai berikut.

1. Model BERT terbukti memiliki performa yang sangat baik dalam mengklasifikasikan berita hoaks dan non-hoaks, dengan skor akurasi mencapai 98%, *precision* 97,6%, *recall* 98,2%, dan *F1-score* 97,9%.
2. Dataset yang digunakan sebanyak 730 data (425 hoaks dan 305 non-hoaks), dibagi secara stratifikasi dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian guna menjaga keseimbangan kelas.
3. Model dikembangkan menggunakan tokenizer *bert-base-uncased* dan arsitektur *BertForSequenceClassification*, serta dilatih selama 4 epoch menggunakan *optimizer Adam* dan *learning rate 5e-5*.
4. Model mampu melakukan prediksi terhadap data baru melalui fungsi prediksi_berita(teks) yang menyajikan hasil klasifikasi disertai tingkat kepercayaan, sehingga dapat dipahami dengan mudah oleh pengguna.
5. Hasil penelitian ini menunjukkan bahwa pendekatan berbasis BERT siap untuk diimplementasikan, dan berpotensi menjadi solusi praktis dalam menangkal penyebaran berita hoaks di masyarakat.

Adapun saran yang dapat diberikan untuk pengembangan lebih lanjut adalah sebagai berikut:

1. Eksplorasi model pralatih yang lebih relevan untuk bahasa Indonesia, seperti IndoBERT, agar pemahaman konteks linguistik lokal dapat ditingkatkan.
2. Perluasan dataset baik dari sisi jumlah, sumber, periode waktu, maupun jenis media sangat diperlukan guna meningkatkan kemampuan generalisasi model.
3. Penggabungan fitur tambahan seperti analisis sentimen dan ekstraksi entitas dapat memperkaya konteks semantik dan mendukung klasifikasi yang lebih akurat.
4. Implementasi model dalam bentuk aplikasi nyata berbasis web atau mobile sangat disarankan agar hasil penelitian dapat dimanfaatkan langsung oleh masyarakat luas.
5. Pengujian model terhadap data dunia nyata, terutama pada berita-berita viral dan tren terkini, penting dilakukan untuk menguji ketahanan model terhadap dinamika informasi digital yang cepat berubah.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan, bantuan, dan kontribusi dalam penyelesaian penelitian ini. Ucapan terima kasih khusus disampaikan kepada dosen pembimbing yang telah memberikan arahan, masukan, serta motivasi selama proses penyusunan.

DAFTAR PUSTAKA

[1] N. P. Raharjo dan B. Winarko, "Analisis Tingkat Literasi Digital Generasi Milenial Kota Surabaya dalam Menanggulangi Penyebaran Hoaks," *J. Komunika J. Komun. Media Dan Inform.*, vol. 10, no. 1, hlm. 33, Sep 2021, doi: 10.31504/komunika.v10i1.3795.

[2] Y. R. Silitonga, "SISTEM PENDETEKSI BERITA *HOAX* DI MEDIA SOSIAL DENGAN TEKNIK DATA MINING SCIKIT LEARN," *J. Ilmu Komput.*, vol. 4, no. 2, 2019.

[3] M. R. D. Sulistyo dan F. U. Najicha, "PENGARUH BERITA *HOAX* TERHADAP KESATUAN DAN PERSATUAN BANGSA INDONESIA," vol. 6, no. 1, 2022.

[4] B. Mufarida, "Dewan Pers: Baru 1.700 Media yang Sudah Terverifikasi," SINDOnews Nasional. Diakses: 12 Juni 2025. [*Daring*]. Tersedia pada: https://nasional.sindonews.com/read/1331869/15/dewan-pers-baru-1700-media-yang-sudah-terverifikasi-1709283799

[5] "593576-machine-learning-teori-program-dan-studi-5c99fd07.pdf." Diakses: 12 Juni 2025. [*Daring*]. Tersedia pada: https://repository.deepublish.com/media/publications/593576-machine-learning-teori-program-dan-studi-5c99fd07.pdf

[6] T. Silwattananusarn dan P. Kulkanjanapiban, *(PDF) A text mining and topic modeling based bibliometric exploration of information science research*. Diakses: 12 Juni 2025. [*Daring*]. Tersedia pada: https://www.researchgate.net/publication/361864566_A_text_mining_and_topic_modeling_based_bibliometric_exploration_of_information_science_research

[7] M. Ula, "ANALISA DAN DETEKSI KONTEN *HOAX* PADA MEDIA BERITA INDONESIA MENGGUNAKAN MACHINE LEARNING," *J. Teknol. Terap. Sains 40*, vol. 1, no. 2, hlm. 229, Des 2020, doi: 10.29103/tts.v1i2.3263.

[8] F. Prasetya dan F. Ferdiansyah, "Analisis Data Mining Klasifikasi Berita *Hoax* COVID 19 Menggunakan Algoritma Naive Bayes," *J. Sist. Komput. Dan Inform. JSON*, vol. 4, no. 1, hlm. 132, Sep 2022, doi: 10.30865/json.v4i1.4852.

[9] A. A. Kurniawan dan M. Mustikasari, "Implementasi *Deep Learning* Menggunakan Metode CNN dan LSTM untuk Menentukan Berita Palsu dalam Bahasa Indonesia," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, hlm. 544, Des 2021, doi: 10.32493/informatika.v5i4.6760.

[10] J. Fawaid, A. Awalina, R. Y. Krisnabayu, dan N. Yudistira, "Indonesia's Fake News Detection using Transformer Network," dalam *Proceedings of the 6th International Conference on Sustainable Information Engineering and Technology*, dalam SIET '21. New York, NY, USA: Association for Computing Machinery, Nov 2021, hlm. 247–251. doi: 10.1145/3479645.3479666.

[11] B. Ghojogh dan A. Ghodsi, "Neural Network Compression and Knowledge Distillation: Tutorial and Survey," 15 Oktober 2024, *Open Science Framework*. doi: 10.31219/osf.io/4n2cb.

[12] A. Putra Pratama, Ismarmiaty, dan Apriani, "Pengukuran Tingkat Akurasi Pada Ulasan E-Commerce Menggunakan Metode INDOBERTDengan *Optimizer* Adam," *J. Komput. Inf. Dan Teknol.*, vol. 5, no. 1.

[13] M. D. Hendriyanto dan B. N. Sari, "PENERAPAN ALGORITMA K-NEAREST NEIGHBOR DALAM KLASIFIKASI JUDUL BERITA *HOAX*," *J. Ilm. Inform.*, vol. 10, no. 02, hlm. 80–84, Sep 2022, doi: 10.33884/jif.v10i02.5477.

[14] H. Allcott dan M. Gentzkow, "Social Media and Fake News in the 2016 Election," *J. Econ. Perspect.*, vol. 31, no. 2, hlm. 211–236, Mei 2017, doi: 10.1257/jep.31.2.211.

[15] L. A. Kumar dan D. K. Renuka, *Deep Learning Approach for Natural Language Processing, Speech, and Computer Vision: Techniques and Use Cases*, 1 ed. Boca Raton: CRC Press, 2023. doi: 10.1201/9781003348689.

[16] W. R. Pratiwi dan R. E. Putra, "Perbandingan Performa Algoritma GA-SVM dan BOA-SVM dalam Mengklasifikasi Artikel Berita Berbahasa Indonesia," *J. Inform. Comput. Sci. JINACS*, vol. 2, no. 04, hlm. 252–258, Jun 2021, doi: 10.26740/jinacs.v2n04.p252-258.

[17] R. D. Rahman, N. Y. Setiawan, dan F. A. Bachtiar, "Analisis Sentimen Pengguna Aplikasi *Mobile* Berbasis Review Pada Platform Blibli Menggunakan Metode Bidirectional Encoder Representations from Transformers (BERT)".

[18] G. Girinoto, D. A. Alwan, G. A. N. Gde K.T. D, O. G. Nabila, A. Arizal, dan D. F. Priambodo, "Implementasi Deteksi Judul Berita Clickbait Berbahasa Indonesia dengan pre-trained model Multilingual BERT Pada Aplikasi Berbasis Chrome Extension," *J. Ilm. SINUS*, vol. 20, no. 2, hlm. 25, Jul 2022, doi: 10.30646/sinus.v20i2.624.

[19] A. D. Rendragraha, M. A. Bijaksana, dan A. Romadhony, "Pendekatan Metode Transformers untuk Deteksi Bahasa Kasar dalam Komentar Berita Online Indonesia," vol. 8, no. 2, hlm. 3385–3395, 2021.

[20] C. A. Putri, "Analisis Sentimen Review Film Berbahasa Inggris Dengan Pendekatan Bidirectional Encoder Representations from Transformers," *JATISI J. Tek. Inform. Dan Sist. Inf.*, vol. 6, no. 2, hlm. 181–193, Jan 2020, doi: 10.35957/jatisi.v6i2.206.

[21] "DETEKSI BERITA *HOAX* PADA WEBSITE TURNBACK *HOAX* DENGAN MENGGUNAKAN MACHINE LEARNING UIN SYARIF HIDAYATULLAH JAKARTA 2021 M / 1442 M."

[22] D. Nuryadi *dkk.*, "FINE TUNING INDOBERT UNTUK ANALISIS SENTIMEN PADA ULASAN PENGGUNA APLIKASI TIKET.COM DI GOOGLE PLAY STORE," vol. 9, no. 2, 2025.

[23] C. Sun, X. Qiu, Y. Xu, dan X. Huang, "How to Fine-Tune BERT for Text Classification?," 2019, *arXiv*. doi: 10.48550/ARXIV.1905.05583.

[24] S. A. H. Bahtiar, "PERBANDINGAN NAÏVE BAYES DAN LOGISTIC REGRESSION DALAM SENTIMENT ANALYSIS PADA REVIEW MARKETPLACE MENGGUNAKAN RATING- BASED LABELING".

[25] Z. Jannah, R. Kurniawan, dan S. Anwar, "STUDI ALGORITMA NEURAL NETWORK DALAM KLASIFIKASI SENTIMEN PENGGUNA SHOPEE: PENINGKATAN AKURASI MODEL," *J. Inform. Dan Tek. Elektro Terap.*, vol. 13, no. 2, Apr 2025, doi: 10.23960/jitet.v13i2.6113.

[26] B. Firmanto, A. S. Aziz, dan J. Sesoca, "Tinjauan Perkembangan Kecerdasan Buatan Berbasis Arsitektur Transformer □," . *p*, no. 1.

[27] A. R. Hanum *dkk.*, "Analisis Kinerja Algoritma Klasifikasi Teks Bert dalam Mendeteksi Berita Hoaks," *J. Teknol. Inf. Dan Ilmu Komput.*, vol. 11, no. 3, hlm. 537–546, Jul 2024, doi: 10.25126/jtiik.938093.

[28] W. Musu dan A. Ibrahim, "Pengaruh Komposisi Data *Training* dan Testing terhadap Akurasi Algoritma C4.5," no. 1, 2021.

[29] P. Khakhar dan R. Kumar Dubey, "Recall Score - an overview | ScienceDirect Topics." Diakses: 15 Juni 2025. [Daring]. Tersedia pada: https://www.sciencedirect.com/topics/engineering/recall-score?utm_source=chatgpt.com

[30] C. Destitus, W. Wella, dan S. Suryasari, "Support Vector Machine VS Information Gain: Analisis Sentimen Cyberbullying di Twitter Indonesia," *Ultima InfoSys J. Ilmu Sist. Inf.*, vol. 11, no. 2, hlm. 107–111, Des 2020, doi: 10.31937/si.v11i2.1740.

[31] Ernianti Hasibuan dan Elmo Allistair Heriyanto, "ANALISIS SENTIMEN PADA ULASAN APLIKASI AMAZON SHOPPING DI GOOGLE PLAY STORE MENGGUNAKAN NAIVE BAYES CLASSIFIER," *J. Tek. Dan Sci.*, vol. 1, no. 3, hlm. 13–24, Okt 2022, doi: 10.56127/jts.v1i3.434.

[32] T. Gowda dan W. You, "Macro-Average: Rare Types Are Important Too".
[33] T. Hesterberg, "Weighted Average Importance Sampling and Defensive Mixture Distributions," vol. 37, no. 2, 1995.
[34] J. Devlin, M.-W. Chang, K. Lee, dan K. Toutanova, "BERT: *Pre-training* of Deep Bidirectional Transformers for Language Understanding".
[35] A. Tharwat, "Classification assessment methods," *Appl. Comput. Inform.*, vol. 17, no. 1, hlm. 168–192, Jan 2021, doi: 10.1016/j.aci.2018.08.003.