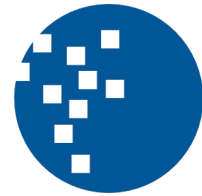


**IMPLEMENTASI MODEL INDOBERT DALAM MENDETEKSI BERITA  
HOAKS BERBAHASA INDONESIA**



**UMN**  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

**LAPORAN MBKM PENELITIAN**

**RIVO JUICER WOWOR  
00000059635**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA  
TANGERANG  
2025**

**IMPLEMENTASI MODEL INDOBERT DALAM MENDETEKSI BERITA  
HOAKS BERBAHASA INDONESIA**



**UMMN**

**UNIVERSITAS  
MULTIMEDIA  
NUSANTARA**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA**

**TANGERANG**

**2025**

## HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Rivo Juicer Wowor  
NIM : 00000059635  
Program Studi : Informatika

Menyatakan dengan sesungguhnya bahwa Tesis/Skripsi/Tugas Akhir/Laporan Magang/MBKM saya yang berjudul:

**Implementasi Model *IndoBERT* dalam Mendeteksi Berita Hoaks Berbahasa Indonesia**

merupakan hasil karya saya sendiri, bukan merupakan hasil plagiat, dan tidak pula dituliskan oleh orang lain; Semua sumber, baik yang dikutip maupun dirujuk, telah saya cantumkan dan nyatakan dengan benar pada bagian Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan karya ilmiah, saya bersedia menerima konsekuensi untuk dinyatakan TIDAK LULUS. Saya juga bersedia menanggung segala konsekuensi hukum yang berkaitan dengan tindak plagiarisme ini sebagai kesalahan saya pribadi dan bukan tanggung jawab Universitas Multimedia Nusantara.

Tangerang, 03 Januari 2025



(Rivo Juicer Wowor)

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK  
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Rivo Juicer Wowor  
NIM : 00000059635  
Program Studi : Informatika  
Jenjang : S1  
Jenis Karya : Laporan MBKM Penelitian

Menyatakan dengan sesungguhnya bahwa:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia karena dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)\*\*.

Tangerang, 03 Januari 2025

Yang menyatakan



Rivo Juicer Wowor

UMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

\*\* Jika tidak bisa membuktikan LoA jurnal/HKI selama enam bulan ke depan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

## Halaman Persembahan / Motto

”Mimpi saya yang terbesar, yang ingin saya laksanakan adalah, agar mahasiswa Indonesia berkembang menjadi *'manusia-manusia yang biasa'*. Menjadi pemuda-pemuda dan pemudi-pemudi yang bertingkah laku sebagai seorang manusia yang normal, sebagai seorang manusia yang tidak mengingkari eksistensi hidupnya sebagai seorang mahasiswa, sebagai seorang pemuda dan sebagai seorang manusia”

Soe Hok Gie  
(Catatan Seorang Demonstran, 1964)

UMMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

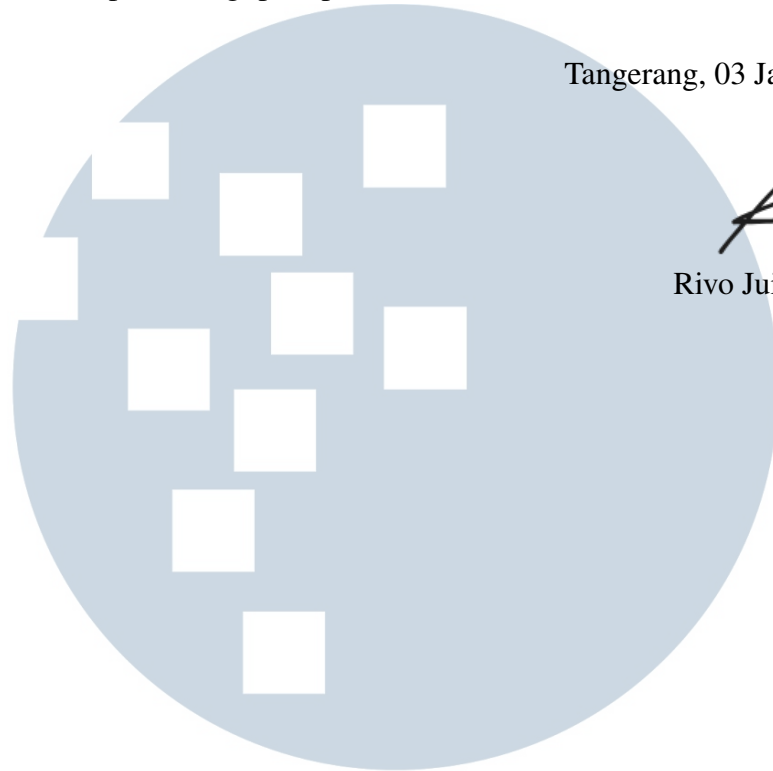
## KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Laporan MBKM Penelitian ini dengan judul: Implementasi Model IndoBERT dalam Mendeteksi Berita Hoaks Berbahasa Indonesia dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan laporan penelitian ini, sangatlah sulit bagi saya untuk menyelesaikan laporan penelitian ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Assoc. Prof. Arya Wicaksana, S.Kom.,M.Eng.Sc., OCA., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Dennis Gunawan, S.Kom., M.Sc., sebagai Pembimbing penelitian yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya laporan penelitian ini.
5. Ibu Dr. Sy. Yuliani Yakub, S.Kom., M.T, selaku Kepala Tim Peneliti UNIIC UMN.
6. Christian Ivan Wibowo, Muhammad Alfarizky Ramadhani Oscandar, serta Fadhil Dzaky Muhammad selaku rekan-rekan tim Peneliti UNIIC UMN.
7. Alm. Ayah beserta Ibu, keluarga inti Wowor-Badilo, keluarga besar GPdI Hebron Gading Serpong, keluarga besar Pengurus Laboratorium Cybersecurity, serta teman-teman saya dari Himpunan Mahasiswa Informatika UMN, Indobest Artha Kreasi dan lainnya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan laporan penelitian ini.

Semoga laporan penelitian ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 03 Januari 2025



Rivo Juicer Wowor

UMMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

# **IMPLEMENTASI MODEL INDOBERT DALAM MENDETEKSI BERITA HOAKS BERBAHASA INDONESIA**

Rivo Juicer Wowor

## **ABSTRAK**

Peningkatan jumlah hoaks yang semakin canggih menjadi ancaman serius bagi keamanan siber saat ini, mengurangi kepercayaan dan menyebarkan informasi yang menyesatkan di kalangan masyarakat. Hoaks-hoaks ini sering muncul sebagai berita palsu dan penipuan online, yang dapat berdampak pada kerugian ekonomi dan reputasi bagi individu maupun organisasi. Di Indonesia, survei menunjukkan bahwa lebih dari 60% orang yang terpapar berita hoaks percaya bahwa informasi tersebut benar, sehingga menunjukkan perlunya metode deteksi yang lebih baik. Pendekatan keamanan siber tradisional sulit mengikuti perkembangan dan kecanggihan serangan-serangan ini. Untuk mengatasi hal ini, penelitian ini melihat potensi penggunaan teknik pembelajaran mesin, khususnya dalam mendeteksi hoaks berbasis teks dalam bahasa Indonesia. Dengan menggunakan IndoBERT, model kecerdasan buatan yang sudah dilatih khusus untuk bahasa Indonesia, penelitian ini bertujuan meningkatkan akurasi deteksi hoaks. Model ini disesuaikan dengan menggunakan data berisi hoaks yang sudah diverifikasi dan konten asli. Hasil penelitian menunjukkan bahwa IndoBERT berpotensi menjadi alat yang baik untuk mendeteksi konten penipuan secara otomatis, sehingga dapat menjadi solusi untuk memperkuat pertahanan keamanan siber. Penelitian ini berkontribusi pada upaya untuk menggunakan teknologi pembelajaran mesin dalam keamanan siber untuk menghadapi tantangan yang terus berkembang.

**Kata kunci:** Deep learning, Deteksi Hoaks, IndoBERT, Keamanan Siber

U M N  
U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A



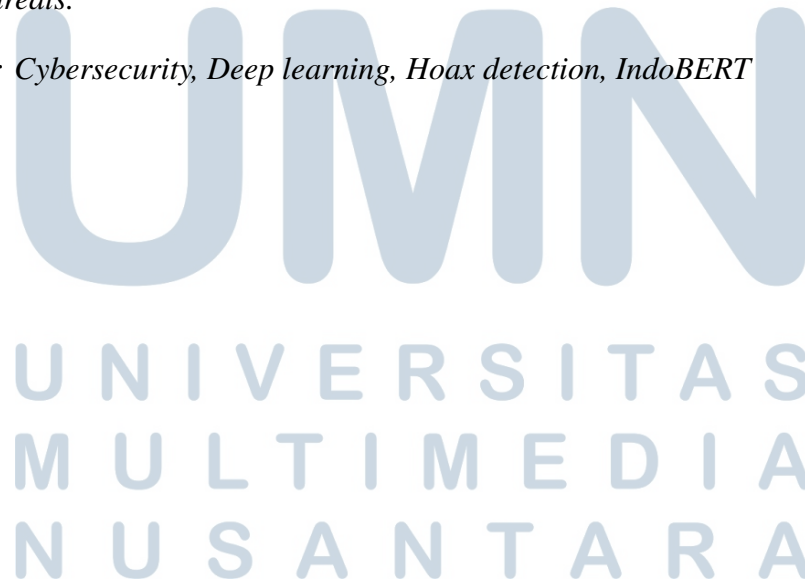
**INDOBERT MODEL IMPLEMENTATION FOR DETECTING FAKE NEWS  
IN INDONESIAN**

Rivo Juicer Wowor

**ABSTRACT**

*The increasing prevalence of deep hoaxes represents a major threat to modern cybersecurity, undermining trust and contributing to the spread of misinformation. These hoaxes, which often manifest as fake news, phishing schemes, or deceptive social engineering tactics, can inflict serious economic and reputational harm on both individuals and organizations. In Indonesia, surveys show that over 60% of people exposed to hoax news believe it to be true, highlighting the critical need for improved detection methods. Traditional cybersecurity approaches struggle to keep pace with the growing scale and sophistication of these attacks. In response, this research explores the potential of deep learning techniques, specifically focusing on text-based hoaxes in the Indonesian language. By leveraging IndoBERT, a pre-trained deep learning model optimized for Indonesian, the study aims to enhance the accuracy of hoax detection. The model was fine-tuned using supervised learning on a dataset comprising verified hoaxes and legitimate content. The results indicate that IndoBERT is a promising tool for automating the detection of deceptive content, offering a scalable solution for bolstering cybersecurity defenses. This research contributes to ongoing efforts to integrate advanced machine learning models into cybersecurity systems, addressing the challenges posed by the evolving landscape of cyber threats.*

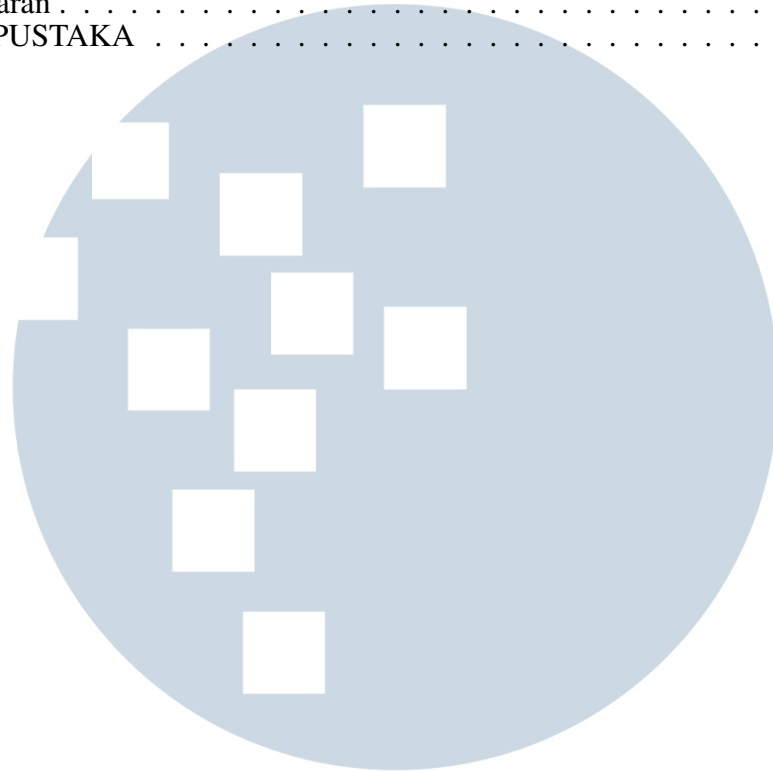
**Keywords:** *Cybersecurity, Deep learning, Hoax detection, IndoBERT*



## DAFTAR ISI

HALAMAN JUDUL . . . . .	i
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH . . . . .	ii
HALAMAN PERSEMBAHAN/MOTO . . . . .	iv
KATA PENGANTAR . . . . .	v
ABSTRAK . . . . .	vii
ABSTRACT . . . . .	viii
DAFTAR ISI . . . . .	ix
DAFTAR GAMBAR . . . . .	xi
DAFTAR TABEL . . . . .	xii
DAFTAR KODE . . . . .	xiii
DAFTAR LAMPIRAN . . . . .	xiv
BAB 1 PENDAHULUAN . . . . .	1
1.1 Latar Belakang . . . . .	1
1.2 Rumusan Masalah . . . . .	3
1.3 Tujuan Penelitian . . . . .	3
1.4 Urgensi Penelitian . . . . .	3
1.5 Luaran . . . . .	3
1.6 Manfaat . . . . .	4
BAB 2 TINJAUAN PUSTAKA . . . . .	5
2.1 Hoaks dalam Kaidah Bahasa . . . . .	5
2.2 BERT dan IndoBERT . . . . .	6
2.2.1 BERT ( <i>Bidirectional Encoder Representations from Transformers</i> ) . . . . .	6
2.2.2 IndoBERT . . . . .	9
2.3 <i>Evaluation Metrics</i> . . . . .	9
BAB 3 METODOLOGI . . . . .	12
3.1 Pengumpulan Data . . . . .	12
3.1.1 Dataset Hoaks . . . . .	13
3.1.2 Dataset Berita Asli . . . . .	14
3.2 Pembersihan Data . . . . .	15
3.2.1 Dataset Hoaks . . . . .	15
3.2.2 Dataset Berita Asli . . . . .	17
3.3 Penggabungan dan Penyeimbangan Data . . . . .	18
3.4 Tokenisasi . . . . .	19
3.5 Konfigurasi Model . . . . .	20
3.5.1 Arsitektur . . . . .	20
3.5.2 Konfigurasi . . . . .	20
3.6 Pengukuran <i>Evaluation Metrics</i> . . . . .	22
BAB 4 HASIL DAN PEMBAHASAN . . . . .	24
4.1 Hasil . . . . .	24
4.1.1 Perbandingan Performa Model terhadap Learning Rate dengan menggunakan Epoch 2 . . . . .	24
4.1.2 Perbandingan Performa Model terhadap Learning Rate dengan menggunakan Epoch 3 . . . . .	25
4.1.3 <i>Confusion Matrix</i> dan <i>evaluation metrics</i> . . . . .	27
4.2 Diskusi . . . . .	30
4.2.1 Dampak Hyperparameter terhadap Performa Model . . . . .	30
4.2.2 Analisis Performa Klasifikasi . . . . .	31

BAB 5	SIMPULAN DAN SARAN . . . . .	33
5.1	Simpulan . . . . .	33
5.2	Saran . . . . .	33
DAFTAR PUSTAKA	. . . . .	34



**UMMN**  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## DAFTAR GAMBAR

Gambar 2.1	Ilustrasi BERT dalam proses <i>fine-tuning</i> untuk klasifikasi kalimat. Diambil di [13] . . . . .	7
Gambar 2.2	Contoh proses tokenisasi pada BERT. Diambil di [19] . . . . .	8
Gambar 3.1	<i>Flowchart</i> proses penelitian . . . . .	12
Gambar 4.1	Perbandingan metrik akurasi, $F_1$ -score, <i>recall</i> , presisi, <i>training loss</i> , dan <i>validation loss</i> pada Epoch 2 dengan variasi <i>learning rate</i> . . . . .	25
Gambar 4.2	Perbandingan metrik akurasi, $F_1$ -score, <i>recall</i> , presisi, <i>training loss</i> , dan <i>validation loss</i> pada Epoch 3 dengan variasi <i>learning rate</i> . . . . .	26
Gambar 4.3	<i>Confusion matrix</i> pada data <i>training</i> . . . . .	28
Gambar 4.4	<i>Confusion matrix</i> pada data validasi. . . . .	29
Gambar 4.5	<i>Confusion matrix</i> pada data <i>testing</i> . . . . .	30



## DAFTAR TABEL

Tabel 2.1	Struktur Confusion Matrix untuk Klasifikasi Berita Hoaks .	10
Tabel 3.1	Sumber dan Karakteristik Pengumpulan Data . . . . .	14
Tabel 3.2	Jumlah serta Rasio Pembagian Dataset yang digunakan . .	18
Tabel 3.3	Parameter Konfigurasi Model . . . . .	20
Tabel 3.4	Hyperparameter Value . . . . .	21
Tabel 4.1	<i>Evaluation Metrics</i> untuk Model IndoBERT pada Data <i>Training, Validasi, dan Uji</i> . . . . .	27



## DAFTAR KODE

3.1	Hoax Data Fetching . . . . .	13
3.2	Hoax Data Cleaning . . . . .	15
3.3	Real News Data Cleaning . . . . .	17
3.4	Train, test, and validation dataset splitting process . . . . .	18
3.5	Tokenizer Function . . . . .	19
3.6	Hyperparameter Function . . . . .	21
3.7	Model Evaluation Process . . . . .	22



## DAFTAR LAMPIRAN

Lampiran 1	MBKM-01 Cover Letter MBKM Research . . . . .	38
Lampiran 2	MBKM-02 MBKM Research Card . . . . .	39
Lampiran 3	MBKM-03 Daily Task - Research . . . . .	40
Lampiran 4	MBKM-04 Verification Form of Report MBKM Research . . .	49
Lampiran 5	Form Bimbingan . . . . .	50
Lampiran 6	Laporan Hasil Turnitin . . . . .	51
Lampiran 7	Kontrak Kerjasama Penelitian . . . . .	52
Lampiran 8	Draft Paper . . . . .	53

