

**DETEKSI BERITA DEEPHOAKS SEBAGAI UPAYA  
KEAMANAN SIBER BERBASIS INDOBERT UNTUK  
BAHASA INDONESIA**



**SKRIPSI**

**VILLE JASON KANNISTO  
00000063305**

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

**DETEKSI BERITA DEEPHOAKS SEBAGAI UPAYA  
KEAMANAN SIBER BERBASIS INDOBERT UNTUK  
BAHASA INDONESIA**



**SKRIPSI**

Diajukan sebagai salah satu syarat untuk memperoleh  
Gelar Sarjana Komputer (S.Kom.)

**VILLE JASON KANNISTO  
00000063305**

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

## HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Ville Jason Kannisto

Nomor Induk Mahasiswa : 00000063305

Program Studi : Informatika

Skripsi dengan judul:

**Deteksi Berita Deephoaks sebagai Upaya Keamanan Siber Berbasis IndoBERT untuk Bahasa Indonesia**

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.



Tangerang, 03 Juli 2025

(Ville Jason Kannisto)

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

## HALAMAN PENGESAHAN

Skripsi dengan judul

### **DETEKSI BERITA DEEPHOAKS SEBAGAI UPAYA KEAMANAN SIBER BERBASIS INDOBERT UNTUK BAHASA INDONESIA**

oleh

Nama : Ville Jason Kannisto  
NIM : 00000063305  
Program Studi : Informatika  
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Selasa, 15 Juli 2025

Pukul 13.00 s/d 15.00 dan dinyatakan

**LULUS**

Dengan susunan penguji sebagai berikut

Ketua Sidang

(Wirawan Istiono, S.Kom., M.Kom.)

NIDN: 0313048304

Penguji

(Alexander Waworuntu, S.Kom., M.T.I.)

NIDN: 0309068503

Pembimbing

(SY.Yuliani, S.Kom., MT., PhD)

NIDN: 0411037904

Ketua Program Studi Informatika,

(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

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

Yang bertanda tangan di bawah ini:

Nama : Ville Jason Kannisto  
NIM : 00000063305  
Program Studi : Informatika  
Jenjang : S1  
Judul Karya Ilmiah : Deteksi Berita Deephoaks sebagai  
Upaya Keamanan Siber Berbasis  
IndoBERT untuk Bahasa Indonesia

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) \*\*.
- ☐ Lainnya, pilih salah satu:
  - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
  - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 03 Juli 2025

Yang menyatakan



Ville Jason Kannisto

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

**HALAMAN PERSEMBAHAN / MOTTO**

"Fortune favors the bold"

Virgil, "Aeneid"



## KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan skripsi ini dengan judul: Deteksi Berita Deephoaks sebagai Upaya Keamanan Siber Berbasis IndoBERT untuk Bahasa 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 magang ini, sangatlah sulit bagi saya untuk menyelesaikan laporan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu SY.Yuliani.,S.Kom.,MT .,PhD, sebagai Pembimbing skripsi yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya penelitian skripsi ini.
5. Orang Tua dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan laporan magang ini.

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

Tangerang, 03 Juli 2025



Ville Jason Kannisto



# DETEKSI BERITA DEEPHOAKS SEBAGAI UPAYA KEAMANAN SIBER BERBASIS INDOBERT UNTUK BAHASA INDONESIA

Ville Jason Kannisto

## ABSTRAK

Perkembangan pesat *Generative AI* telah menimbulkan ancaman keamanan siber yang signifikan. Teks hoaks yang dihasilkan AI, atau *deephoaks*, dapat digunakan untuk memanipulasi opini publik, melakukan penipuan, dan mengikis kepercayaan pada informasi digital, sehingga deteksi yang andal menjadi sangat krusial. Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model deteksi berita hoaks yang efektif untuk konten berbahasa Indonesia. Metode yang digunakan adalah *fine-tuning* model Transformer berbasis arsitektur BERT, yaitu IndoBERT, pada tugas klasifikasi biner. Dataset penelitian dibangun dengan menggabungkan 12.600 berita fakta dari dataset publik Kaggle dan 12.542 berita hoaks yang dikumpulkan melalui *web scraping* dari situs TurnBackHoax.id, menghasilkan total 25.142 data yang seimbang. Setelah melalui tahap *pre-processing* yang ekstensif, model dilatih dengan strategi *feature extraction* di mana hanya lapisan klasifikasi yang diperbarui. Hasil evaluasi akhir pada data uji menunjukkan performa yang sangat baik, dengan akurasi mencapai 95%, F1-Score 95%, dan ROC-AUC 99%. Hasil ini membuktikan bahwa model IndoBERT yang di-*fine-tuning* sangat efektif dan memiliki potensi besar untuk diimplementasikan sebagai sistem pertahanan otomatis terhadap ancaman disinformasi dalam ekosistem keamanan siber Indonesia.

**Kata kunci:** *Deep Learning*, Deteksi Hoaks, IndoBERT, Keamanan Siber, *Natural Language Processing*

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



**DEEPHOAX NEWS DETECTION AS A CYBERSECURITY EFFORT  
INDOBERT-BASED FOR INDONESIAN LANGUAGE**

Ville Jason Kannisto

**ABSTRACT**

*The rapid development of Generative AI poses a significant cybersecurity threat. AI-generated hoax texts, or deephoaks, can be used to manipulate public opinion, commit fraud, and erode trust in digital information, making reliable detection crucial. This research aims to develop and evaluate an effective hoax news detection model for Indonesian language content. The method employed is the fine-tuning of a BERT-based Transformer model, specifically IndoBERT, for a binary classification task. The research dataset was constructed by combining 12,600 real news articles from a public Kaggle dataset and 12,542 hoax news articles collected via web scraping from the TurnBackHoax.id website, resulting in a balanced dataset of 25,142 entries. Following an extensive preprocessing phase, the model was trained using a feature extraction strategy, where only the classification layer was updated. The final evaluation on the test set demonstrated excellent performance, achieving an accuracy of 95%, an F1-Score of 95%, and a ROC-AUC of 99%. These results prove that the fine-tuned IndoBERT model is highly effective and holds significant potential for implementation as an automated defense system against disinformation threats within Indonesia's cybersecurity ecosystem.*

**Keywords:** Cybersecurity, Deep Learning, Hoax Detection, IndoBERT, Natural Language Processing.



## DAFTAR ISI

HALAMAN JUDUL . . . . .	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT . . . . .	ii
HALAMAN PENGESAHAN . . . . .	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH . . . . .	iv
HALAMAN PERSEMBAHAN/MOTO . . . . .	v
KATA PENGANTAR . . . . .	vi
ABSTRAK . . . . .	vii
ABSTRACT . . . . .	viii
DAFTAR ISI . . . . .	ix
DAFTAR TABEL . . . . .	xi
DAFTAR GAMBAR . . . . .	xii
DAFTAR KODE . . . . .	xiii
DAFTAR RUMUS . . . . .	xiv
DAFTAR LAMPIRAN . . . . .	xv
BAB 1 PENDAHULUAN . . . . .	1
1.1 Latar Belakang Masalah . . . . .	1
1.2 Rumusan Masalah . . . . .	3
1.3 Batasan Permasalahan . . . . .	4
1.4 Tujuan Penelitian . . . . .	4
1.5 Manfaat Penelitian . . . . .	4
1.6 Sistematika Penulisan . . . . .	5
BAB 2 LANDASAN TEORI . . . . .	6
2.1 Keamanan Siber (Cyber Security) . . . . .	6
2.1.1 Definisi Keamanan Siber . . . . .	6
2.1.2 Ancaman dan Serangan dalam Keamanan Siber . . . . .	6
2.1.3 Peran AI dan Machine Learning dalam Keamanan Siber . . . . .	7
2.2 Konsep Dasar Deepfakes dan Dampaknya . . . . .	8
2.3 Model BERT dan IndoBERT (Bidirectional Encoder Representations from Transformers) . . . . .	10
2.3.1 Pengertian dan Prinsip Kerja BERT . . . . .	10
2.3.2 Arsitektur Transformer Encoder pada BERT . . . . .	11
2.3.3 Representasi Input dan Tokenisasi pada BERT . . . . .	13
2.3.4 IndoBERT: Adaptasi BERT untuk Bahasa Indonesia . . . . .	14
2.4 Pendekatan Deteksi Berita Hoaks Berbasis Teks . . . . .	14
2.5 Dataset untuk Deteksi Berita Hoaks Teks . . . . .	15
2.6 Studi Terkait . . . . .	16
BAB 3 METODOLOGI PENELITIAN . . . . .	17
3.1 Desain Penelitian Deteksi Berita Hoaks . . . . .	17
3.2 Pengumpulan Data . . . . .	18
3.2.1 Web Scraping . . . . .	20
3.3 Implementasi Model . . . . .	22
3.4 Evaluasi Model . . . . .	23
BAB 4 HASIL DAN DISKUSI . . . . .	26
4.1 Spesifikasi Sistem . . . . .	26
4.1.1 Perangkat Keras . . . . .	26
4.1.2 Perangkat Lunak . . . . .	26
4.2 Deskripsi Dataset dan Preprocessing . . . . .	27
4.2.1 Dataset . . . . .	27

4.2.2	Proses dan Hasil Preprocessing . . . . .	28
4.3	Implementasi dan Evaluasi Model . . . . .	41
4.3.1	Implementasi Model Diteksi Berita Bahasa Indonesia dengan IndoBERT . . . . .	41
4.4	Analisa Hasil Evaluasi Model . . . . .	51
4.4.1	Analisis Performa Model . . . . .	51
4.4.2	Perbandingan dengan Studi Terkait . . . . .	56
BAB 5	KESIMPULAN DAN SARAN . . . . .	58
5.1	Kesimpulan . . . . .	58
5.2	Saran . . . . .	58
DAFTAR PUSTAKA	. . . . .	60



## DAFTAR TABEL

Tabel 2.1	Perkembangan Pendekatan Deteksi Berita Hoaks Berbasis Teks . . . . .	15
Tabel 2.2	Perbandingan Hasil Akurasi dengan Penelitian Sebelumnya	16
Tabel 4.1	Perbandingan Rata-rata Panjang Kata per Kategori Berita .	30
Tabel 4.2	Format Tabel sebelum Unifikasi . . . . .	31
Tabel 4.3	Contoh Format Tabel Setelah Unifikasi untuk Dataset Fakta dan Hoaks . . . . .	31
Tabel 4.4	Perbandingan Contoh Teks Sebelum dan Sesudah Penghapusan Tag Sumber . . . . .	33
Tabel 4.5	Perbandingan Berita Hoaks Sebelum dan Sesudah Ekstraksi	35
Tabel 4.6	Perbandingan Teks Sebelum dan Sesudah Proses Pembersihan Global . . . . .	37
Tabel 4.7	Perbandingan Teks Sebelum dan Sesudah Tokenisasi . . .	46
Tabel 4.8	Hasil Metrik Evaluasi Final pada Data Uji . . . . .	52
Tabel 4.9	Hasil Uji Performa Model pada Berita Fakta dengan Panjang Bervariasi . . . . .	56
Tabel 4.10	Perbandingan Hasil Akurasi dengan Penelitian Sebelumnya	57



## DAFTAR GAMBAR

Gambar 1.1	Contoh Berita Hoaks . . . . .	2
Gambar 2.1	Penipuan Deepfake Melonjak di Indonesia Sebesar 1550% . . . . .	9
Gambar 2.2	Isu penyebaran hoaks di Indonesia . . . . .	10
Gambar 2.3	Arsitektur BERT dalam prosedur pre-training dan fine-tuning . . . . .	11
Gambar 2.4	Representasi Input BERT . . . . .	14
Gambar 3.1	Desain penelitian deteksi deephoaks . . . . .	17
Gambar 3.2	Pengumpulan Data . . . . .	18
Gambar 3.3	Contoh dari Dataset Fakta - diambil dari Kaggle . . . . .	19
Gambar 3.4	Contoh dari Dataset Hoaks - diambil dari TurnBackHoax.id . . . . .	19
Gambar 3.5	Proses Web-Scraping . . . . .	20
Gambar 3.6	<i>pre-processing</i> Data . . . . .	21
Gambar 4.1	Perbandingan Data Awal . . . . .	28
Gambar 4.2	Distribusi Sumber Berita Fakta (Sebelum balancing) . . . . .	38
Gambar 4.3	Distribusi Sumber Berita Fakta (Sesudah balancing) . . . . .	39
Gambar 4.4	Komposisi Dataset Final Setelah merging . . . . .	40
Gambar 4.5	Proses <code>Trainer.train()</code> . . . . .	50
Gambar 4.6	Kurva Training dan Validation Loss Selama Pelatihan . . . . .	53
Gambar 4.7	Kurva Peningkatan Metrik pada Validation Set per Epoch . . . . .	53
Gambar 4.8	Confusion Matrix Hasil Prediksi pada Data Uji . . . . .	54
Gambar 4.9	Kurva ROC dengan Nilai AUC . . . . .	55



## DAFTAR KODE

Kode 4.1	Pemeriksaan dan Pembersihan Data Tidak Valid . . . . .	28
Kode 4.2	Skrip untuk Unifikasi Struktur Dataset . . . . .	30
Kode 4.3	Function untuk Penghapusan Tag Sumber Berita . . . . .	32
Kode 4.4	Pipeline Fungsi untuk Pembersihan dan Ekstraksi Teks Hoaks . .	33
Kode 4.5	Fungsi untuk Pembersihan dan Normalisasi Teks Global . . . . .	35
Kode 4.6	Penyamaan Value Berita Fakta . . . . .	37
Kode 4.7	Mengabungkan Dataset . . . . .	39
Kode 4.8	Proses Pembagian Dataset . . . . .	40
Kode 4.9	Pemuatan dan Persiapan Dataset untuk Pelatihan . . . . .	42
Kode 4.10	Inisialisasi Model IndoBERT dan Pembekuan Lapisan Dasar . .	43
Kode 4.11	Proses Tokenisasi Dataset . . . . .	45
Kode 4.12	Strukturisasi Data dengan Kelas Dataset dan Data Collator . . . .	46
Kode 4.13	Fungsi untuk Kalkulasi Metrik Evaluasi . . . . .	47
Kode 4.14	Konfigurasi Training . . . . .	49
Kode 4.15	Pelatihan Model dengan Trainer . . . . .	50
Kode 4.16	Evaluasi Final Model pada Data Uji . . . . .	51



## DAFTAR RUMUS

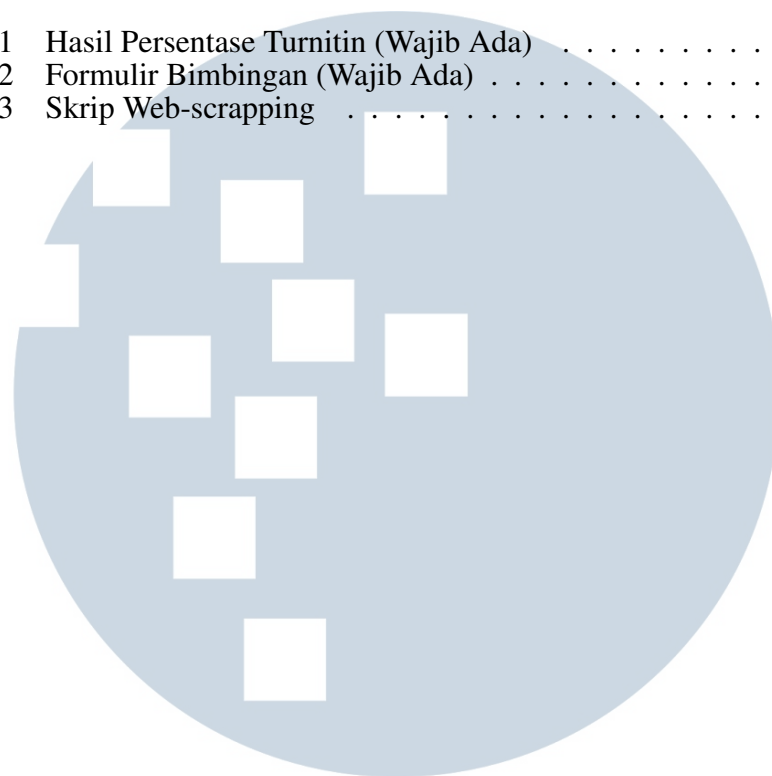
Rumus 2.1	<i>self-attention</i>	12
Rumus 2.2	<i>multi-head attention</i>	12
Rumus 3.1	<i>Accuracy</i>	24
Rumus 3.2	<i>Precision</i>	24
Rumus 3.3	<i>Recall</i>	24
Rumus 3.4	<i>F1-Score</i>	24





## DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin (Wajib Ada) . . . . .	62
Lampiran 2	Formulir Bimbingan (Wajib Ada) . . . . .	70
Lampiran 3	Skrip Web-scraping . . . . .	71



UMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA