

**PERBANDINGAN MODEL KLASIFIKASI UJARAN
KEBENCIAN DI MEDIA SOSIAL BERBASIS BERT,
XGBOOST, DAN HYBRID BERT-XGBOOST**



SKRIPSI

**ARRAFI AJI PAMUNGKAS
00000064717**

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

**PERBANDINGAN MODEL KLASIFIKASI UJARAN
KEBENCIAN DI MEDIA SOSIAL BERBASIS BERT,
XGBOOST, DAN HYBRID BERT-XGBOOST**



SKRIPSI

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

**ARRAFI AJI PAMUNGKAS
00000064717**

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

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Arrafi Aji Pamungkas
Nomor Induk Mahasiswa : 00000064717
Program Studi : Informatika

Skripsi dengan judul:

Perbandingan Model Klasifikasi Ujaran Kebencian di Media Sosial Berbasis BERT, XGBoost, dan Hybrid BERT-XGBoost serta Interpretasi Naratif Hasil Menggunakan Generative AI

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, 16 Juni 2025



(Arrafi Aji Pamungkas)

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

PERBANDINGAN MODEL KLASIFIKASI UJARAN KEBENCIAN DI MEDIA SOSIAL BERBASIS BERT, XGBOOST, DAN HYBRID BERT-XGBOOST

oleh

Nama : Arrafi Aji Pamungkas
NIM : 00000064717
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Senin, 14 Juli 2025

Pukul 13.00 s/d 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang

Penguji

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

NIDN: 0315109103

(Marlinda Vasty Overbeek, S.Kom.,
M.Kom.)

NIDN: 0818038501

Pembimbing

(Aditiyawan, S.Kom., M.Si.)

NIDK: 8994550022

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 : Arrafi Aji Pamungkas
NIM : 00000064717
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Perbandingan Model Klasifikasi
Ujaran Kebencian di Media Sosial
Berdasarkan BERT, XGBoost, dan Hybrid
BERT-XGBoost

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.

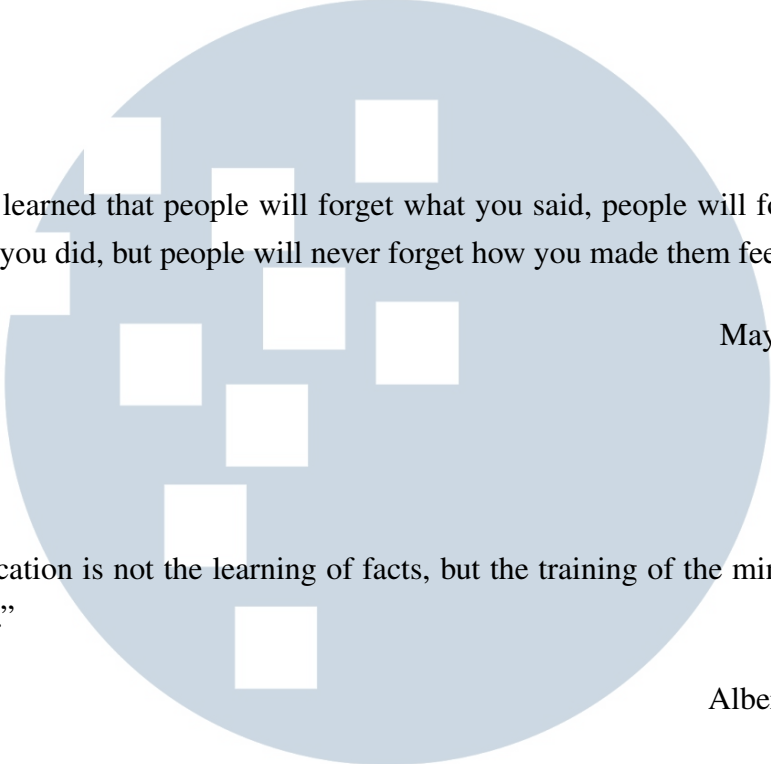
Tangerang, 16 Juni 2025

Yang menyatakan



Arrafi Aji Pamungkas

HALAMAN PERSEMBAHAN / MOTTO



"I've learned that people will forget what you said, people will forget what you did, but people will never forget how you made them feel."

Maya Angelou

"Education is not the learning of facts, but the training of the mind to think."

Albert Einstein

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala limpahan rahmat, karunia, petunjuk, dan kekuatan yang senantiasa diberikan sehingga penulis dapat menyelesaikan skripsi yang berjudul “*Perbandingan Model Klasifikasi Ujaran Kebencian di Media Sosial Berbasis BERT, XGBoost, dan Hybrid BERT-XGBoost*” dengan baik dan lancar. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika di Universitas Multimedia Nusantara. Penulis menyadari bahwa proses penyusunan skripsi ini bukanlah hal yang mudah, karena di dalamnya terdapat berbagai tantangan yang menguji kemampuan dalam mengatur waktu, mengendalikan emosi, menjaga fokus, dan mempertahankan motivasi ditengah berbagai keterbatasan. Oleh sebab itu, penulis ingin menyampaikan rasa terima kasih yang sedalam-dalamnya kepada semua pihak yang telah membantu, membimbing, dan memberikan dorongan moral maupun materiil selama proses penyusunan skripsi ini 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. Bapak Aditiyawan, S.Kom., M.Si., sebagai Pembimbing pertama yang telah meluangkan waktu serta memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Dosen dan staf pengajar di Teknik Informatika, yang telah memberikan ilmu dan wawasan selama masa perkuliahan.
6. Keluarga saya yang telah memberikan dukungan penuh serta selalu mendukung apapun keputusan yang penulis ambil hingga dapat menyelesaikan penulisan skripsi.

7. Teman-teman selama perkuliahan penulis yang selalu menemani, *shout out* buat kalian **Fadhil Rahman, Cornellius Baros, Ville Jason, William Reyhan** secara tidak langsung dalam penulisan skripsi ini.
8. Seluruh pihak yang telah menemani selama proses penulisan skripsi ini dan memberikan kontribusi, dukungan, serta doa kepada penulis

Semoga karya ilmiah ini dapat bermanfaat bagi pembaca dan menjadi referensi berguna untuk mengembangkan keilmuan dalam bidang teknologi.

Tangerang, 16 Juni 2025



Arrafi Aji Pamungkas



**PERBANDINGAN MODEL KLASIFIKASI UJARAN KEBENCIAN DI
MEDIA SOSIAL BERBASIS BERT, XGBOOST, DAN HYBRID
BERT-XGBOOST**

Arrafi Aji Pamungkas

ABSTRAK

Perkembangan media sosial telah memicu peningkatan signifikan dalam penyebaran ujaran kebencian daring, khususnya pada platform X (Twitter). Penelitian ini mengusulkan pendekatan *hybrid* berbasis *Bidirectional Encoder Representations from Transformers* (BERT) dan algoritma *Extreme Gradient Boosting* (XGBoost) untuk mendeteksi konten bermuatan ujaran kebencian secara lebih akurat. Dataset dikumpulkan melalui scraping menggunakan Twitter Search API dan anotasi berbasis *crowdsourcing*, mencakup kategori, target, dan intensitas ujaran kebencian, dengan total sebanyak 13.014 data. Tiga model dikembangkan dan dibandingkan: BERT + XGBoost, XGBoost berbasis TF-IDF, dan BERT *finetuned*. Hasil evaluasi menunjukkan bahwa model *hybrid* mencapai akurasi 81%, namun tidak melampaui performa model BERT *finetuned*, yang memperoleh akurasi 88,99% dan F1-score 0,8893. Temuan ini mengindikasikan bahwa pendekatan *hybrid* tidak selalu menjamin peningkatan kinerja dalam konteks klasifikasi teks berbahasa Indonesia.

Kata kunci: klasifikasi teks, ujaran kebencian, BERT, XGBoost, *hybrid model*.



**COMPARATIVE ANALYSIS OF HATE SPEECH CLASSIFICATION
MODELS ON SOCIAL MEDIA: BERT, XGBOOST, AND HYBRID
BERT-XGBOOST**

Arrafi Aji Pamungkas

ABSTRACT

The rise of social media has significantly accelerated the spread of online hate speech, particularly on Platform X (Twitter). This study proposes a hybrid approach that combines Bidirectional Encoder Representations from Transformers (BERT) and Extreme Gradient Boosting (XGBoost) to improve the accuracy of hate speech detection. The dataset was collected via scraping using the Twitter Search API and annotated through crowdsourcing, covering the categories, targets, and intensity levels of hate speech, resulting in a total of 13,014 entries. Three models were developed and compared: a hybrid BERT + XGBoost model, an XGBoost model with TF-IDF features, and a fine-tuned BERT model. Evaluation results show that the hybrid model achieved an accuracy of 81%, but did not outperform the fine-tuned BERT model, which attained an accuracy of 88.99% and an F1-score of 0.8893. These findings indicate that hybrid approaches do not always guarantee performance improvements, especially in the context of Indonesian-language text classification.

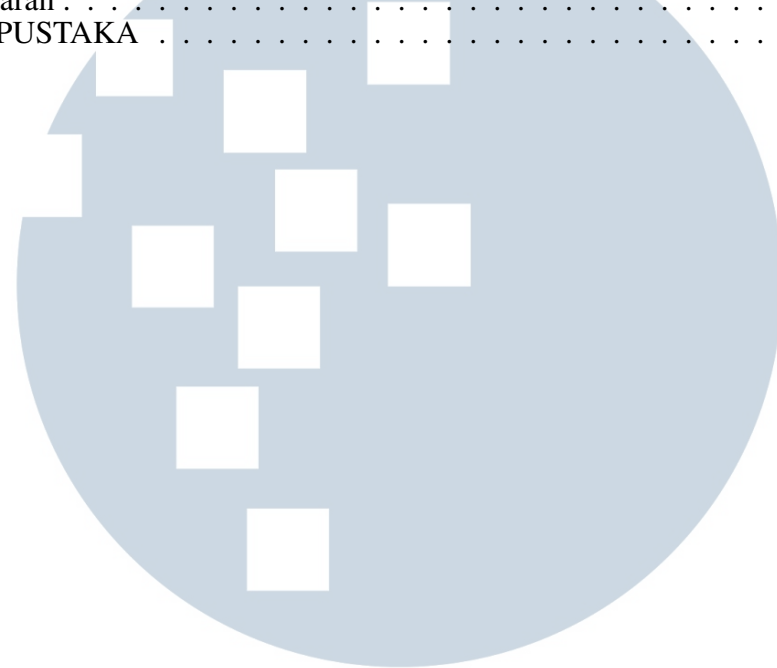
Keywords: text classification, hate speech, BERT, XGBoost, *hybrid model*.



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	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR KODE	xiv
DAFTAR RUMUS	xv
DAFTAR LAMPIRAN	xvi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	5
2.1 Ujaran Kebencian	5
2.2 Media Sosial X	5
2.3 TF-IDF	6
2.4 Transformer	7
2.4.1 Bidirectional Encoder Representations from Transformers (BERT)	8
2.4.2 indoBERT	10
2.5 Ensemble Learning	10
2.5.1 Gradient Boosting	11
2.5.2 XGBoost	12
2.6 Confussion Matrix	15
BAB 3 METODOLOGI PENELITIAN	17
3.1 Studi Literatur	17
3.2 Pengumpulan Data	18
3.3 Preprocessing Data	18
3.4 Perancangan Model	20
3.4.1 Split Dataset	21
3.4.2 Feature Extraction	22
3.4.3 Hyperparameter Tuning	22
3.5 Pengujian dan Evaluasi	23
BAB 4 HASIL DAN DISKUSI	25
4.1 Spesifikasi Sistem	25
4.2 Deskripsi Dataset	26
4.3 Implementasi Sistem	27
4.3.1 Import Libraries	27
4.3.2 Preprocessing Data	28

4.3.3	Perancangan Model	36
4.4	Uji Coba dan Evaluasi Model	47
4.4.1	Rangkuman dan Evaluasi Akhir Antar Model	54
BAB 5	SIMPULAN DAN SARAN	56
5.1	Simpulan	56
5.2	Saran	57
DAFTAR PUSTAKA	58



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

Tabel 4.1	Contoh lima data pertama	26
Tabel 4.2	Hasil <i>remove</i> kolom	29
Tabel 4.3	Data sebelum dan sesudah <i>cleaning</i>	33
Tabel 4.4	Ringkasan jumlah data dan hasil audit label	34
Tabel 4.5	Distribusi label <i>hate speech</i> sebelum dan sesudah <i>relabeling</i>	35
Tabel 4.6	Distribusi dataset	37
Tabel 4.7	Hasil evaluasi model pada data uji	48
Tabel 4.8	Hasil evaluasi model pada data uji	50
Tabel 4.9	Hasil evaluasi model pada data uji	53
Tabel 4.10	Evaluasi model untuk klasifikasi <i>hate-speech</i> dan <i>non-hate</i>	54



DAFTAR GAMBAR

Gambar 2.1	Arsitektur <i>transformer</i>	8
Gambar 2.2	Arsitektur proses <i>pretraining</i> dan <i>finetuning</i> pada <i>bidirectional encoder representations from transformers</i> (bert)	9
Gambar 2.3	Arsitektur Gradient Boosting	12
Gambar 2.4	Arsitektur XGBoost	15
Gambar 3.1	Alur metodologi penelitian	17
Gambar 3.2	Alur Preprocessing Dataset	19
Gambar 3.3	Alur Perancangan Model	20
Gambar 4.1	Grafik <i>loss</i> dan akurasi <i>training</i> dan <i>validation</i> model bert+xgboost	47
Gambar 4.2	<i>Confusion matrix</i> model bert+xgboost	48
Gambar 4.3	Grafik <i>loss</i> dan akurasi <i>training</i> dan <i>validation</i> model xgboost tf-idf	50
Gambar 4.4	<i>Confusion matrix</i> model xgboost tf-idf	51
Gambar 4.5	Grafik <i>loss</i> dan akurasi <i>training</i> dan <i>validation</i> model bert finetuned	52
Gambar 4.6	<i>Confusion matrix</i> model bert finetuned	53



DAFTAR KODE

Kode 4.1	Import libraries	27
Kode 4.2	Pemeriksaan dataset	28
Kode 4.3	Potongan kode <i>remove</i> kolom	29
Kode 4.4	Cek <i>null</i> dan hapus duplikasi	30
Kode 4.5	Fungsi-fungsi <i>cleaning text</i>	30
Kode 4.6	Fungsi <i>preprocessing</i>	32
Kode 4.7	Deteksi <i>mislabeled</i>	34
Kode 4.8	<i>Update label</i>	35
Kode 4.9	Split dataset	36
Kode 4.10	Menampilkan distribusi dataset setelah split	37
Kode 4.11	Fungsi ekstraksi <i>embedding</i> dengan <i>mean pooling</i>	38
Kode 4.12	Proses <i>embedding</i> untuk seluruh dataset	40
Kode 4.13	<i>Hyperparameter</i> dengan grid search	41
Kode 4.14	Inisialisasi <i>dmatrix</i> dan <i>best parameters</i>	42
Kode 4.15	<i>Callback logging metrics</i> evaluasi	43
Kode 4.16	Pelatihan model <i>xgboost</i>	45
Kode 4.17	Evaluasi <i>metrics</i> pada data uji	46



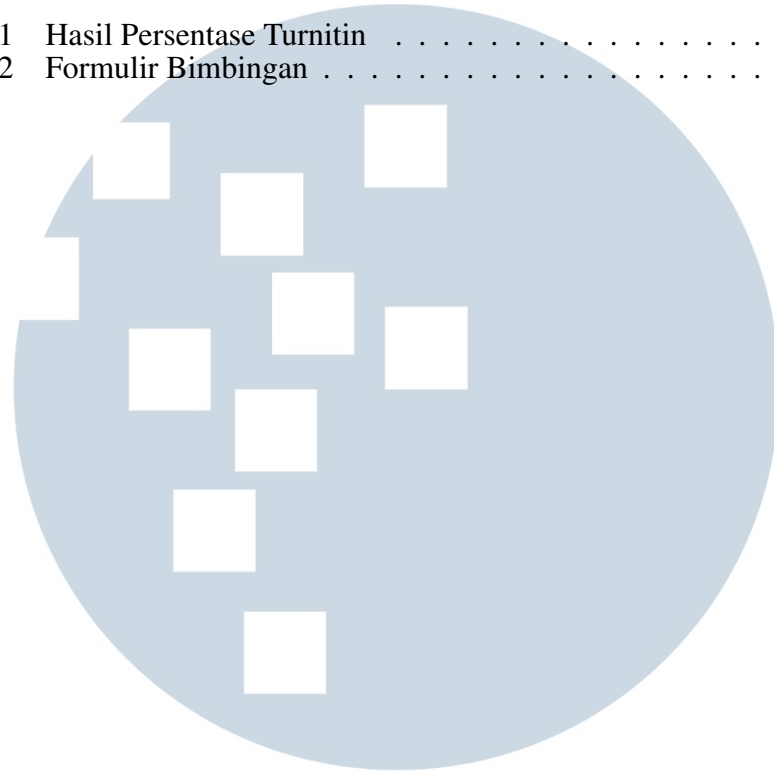
DAFTAR RUMUS

Rumus 2.1	<i>Term Frequency (TF) Function</i>	6
Rumus 2.2	<i>Inverse Document Frequency (IDF) Function</i>	7
Rumus 2.3	<i>TF-IDF Function</i>	7
Rumus 2.4	<i>Gradient Boosting Prediciton Function</i>	11
Rumus 2.5	<i>XGBoost Prediction Function</i>	13
Rumus 2.6	<i>XGBoost Objective Function</i>	13
Rumus 2.7	<i>XGBoost Regularization Term</i>	13
Rumus 2.8	<i>Second-Order Approximation of XGBoost Loss</i>	13
Rumus 2.9	<i>XGBoost Gain Function</i>	14
Rumus 2.10	<i>Accuracy</i>	16
Rumus 2.11	<i>Precision</i>	16
Rumus 2.12	<i>Recall (Sensitivity)</i>	16
Rumus 2.13	<i>F1-Score</i>	16



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	62
Lampiran 2	Formulir Bimbingan	72



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA