

**ANALISIS SENTIMEN PEMANFAATAN CHATGPT DI EKOSISTEM
AKADEMIK MENGGUNAKAN INDOBERT BERBASIS DATA
TWITTER/X**

SKRIPSI

Oleh:

IRSYAD MAULI NASUTION

NIM.221113468



**PROGRAM STUDI S-1 TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2026**

SENTIMEN ANALYSIS OF CHATGPT ADAPTION IN THE ACADEMIC ECOSYSTEM USING INDOBERT AND TWITTER/X DATA

UNDERGRADUATE THESIS

By:

IRSYAD MAULI NASUTION

NIM.221113468



**UNDERGRADUATE PROGRAM OF COMPUTER SCIENCE
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN**

2026

LEMBARAN PENGESAHAN

**ANALISIS SENTIMEN PEMANFAATAN CHATGPT DI EKOSISTEM
AKADEMIK MENGGUNAKAN INDOBERT BERBASIS DATA
TWITTER/X**

SKRIPSI

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana
Program Studi S-1 Teknik Informatika

Oleh:

IRSYAD MAULI NASUTION
NIM. 221113468

Disetujui Oleh:

Dosen Pembimbing I,

Syanti Irviantina, S.Kom., M.Kom.

Dosen Pembimbing II,

Heru Kurniawan, S.Kom., M.Kom.

Medan, 28 Juli 2026
Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-1 Teknik Informatika,

Carles Juliandy, S.Kom., M.Kom.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Irsyad Mauli Nasution

NIM : 221113468

Saya telah melaksanakan penelitian dan penulisan Tugas Akhir dengan judul dan tempat penelitian sebagai berikut:

Judul Tugas Akhir : Analisis Sentimen Pemanfaatan ChatGPT Di Ekosistem Akademik Menggunakan IndoBERT Berbasis Data Twitter/X

Tempat Penelitian : Universitas Mikroskil

Alamat Tempat Penelitian : Jl. M.H Thamrin No.140, Pusat Ps., Kec. Medan Kota, Kota Medan, Sumatra Utara 20212

No. Telp. Tempat Penelitian : 0812-6243-2853

Sehubungan dengan Tugas Akhir tersebut, dengan ini saya menyatakan dengan sebenarnya bahwa penelitian dan penulisan Tugas Akhir tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tugas Akhir saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tugas Akhir saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 30 Juni 2026

Saya yang membuat pernyataan,



Irsyad Mauli Nasution

ANALISIS SENTIMEN PEMANFAATAN CHATGPT DI EKOSISTEM AKADEMIK MENGGUNAKAN INDOBERT BERBASIS DATA TWITTER/X

Abstrak

Perkembangan penggunaan ChatGPT dalam bidang akademik memunculkan beragam tanggapan di media sosial, sehingga diperlukan analisis sentimen untuk memahami persepsi masyarakat terhadap pemanfaatannya. Penelitian ini bertujuan menganalisis sentimen pengguna ChatGPT pada konteks akademik berdasarkan data Twitter/X berbahasa Indonesia serta membandingkan kinerja model Support Vector Machine (SVM) berbasis TF-IDF dengan model IndoBERT. Dataset yang digunakan berjumlah 6.377 data hasil Preprocessing, pelabelan otomatis menggunakan mdhugol/indonesia-bert-sentiment-classification, verifikasi manual terhadap 1.200 data sampel, serta penghapusan data berlabel Remove. Data kemudian dibagi menggunakan rasio 90:05:05 untuk proses pelatihan, validasi, dan pengujian. Hasil penelitian menunjukkan bahwa sentimen positif mendominasi dengan 2.660 data, diikuti sentimen netral sebanyak 2.227 data dan sentimen negatif sebanyak 1.490 data. Model SVM memperoleh nilai accuracy sebesar 0,6771, sedangkan IndoBERT mencapai accuracy sebesar 0,7931 dengan selisih performa sebesar 11,60%. Hasil tersebut menunjukkan bahwa IndoBERT lebih efektif dibandingkan SVM berbasis TF-IDF dalam mengklasifikasikan sentimen berbahasa Indonesia pada konteks pemanfaatan ChatGPT di bidang akademik.

Kata kunci: Analisis Sentimen, ChatGPT, Twitter/X, SVM, IndoBERT

Abstract

The growing use of ChatGPT in academic settings has sparked a variety of reactions on social media, making sentiment analysis necessary to understand public perceptions of its use. This study aims to analyze the sentiment of ChatGPT users in an academic context based on Indonesian-language Twitter/X data and to compare the performance of a TF-IDF-based Support Vector Machine (SVM) model with the IndoBERT model. The dataset consisted of 6,377 preprocessed data points, automatically labeled using mdhugol/indonesia-bert-sentiment-classification, with manual verification of 1,200 sample data points and the removal of data labeled "Remove." The data was then split in an 90:05:05 ratio for training, validation, and testing. The results showed that positive sentiment dominated with 2,660 data points, followed by neutral sentiment with 2,227 data points and negative sentiment with 1,490 data points. The SVM model achieved an accuracy of 0.6771, while IndoBERT reached an accuracy of 0.7931—a performance difference of 11.60%. These results indicate that IndoBERT is more effective than TF-IDF-based SVM in classifying Indonesian-language sentiment in the context of using ChatGPT in academic settings.

Keywords: Sentiment Analysis, ChatGPT, Twitter/X, SVM, IndoBERT

KATA PENGANTAR

Ucapan syukur Ucapan syukur yang tulus saya panjatkan kepada Tuhan yang Maha Esa atas segala limpahan Rahmat, petunjuk, dan karunia-Nya yang telah menyertai perjalanan saya dalam menyelesaikan skripsi ini. Dengan penuh rasa Syukur, saya mengakui bahwa tanpa pertolongan-Nya, saya tidak akan mampu mengatasi berbagai tantangan dan hambatan yang saya hadapi selama proses penyusunan skripsi ini.

Penyusunan skripsi ini merupakan bagian integral dari tahapan akhir dalam menyelesaikan pendidikan saya di Program Studi Teknik Informatika, Fakultas Informatika Universitas Mikroskil Medan.

Saya ingin menyampaikan ucapan terima kasih yang sebesar besarnya kepada semua pihak yang telah memberikan dukungan, bimbingan, dan motivasi dalam proses penyusunan skripsi ini.

1. Syanti Irviantina, S.Kom., M.Kom., selaku Dosen Pembimbing I.
2. Heru Kurniawan, S.Kom., M.Kom., selaku Dosen Pembimbing II.
3. Bapak Hardy, S.kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Sunaryo Winardi, S.kom., M.T.,selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Bapak Carles Juliandy, S.kom., M.kom., selaku Ketua Program Studi S-1 Fakultas Teknik Informatika Universitas Mikroskil Medan.
6. Orang tua dan keluarga

Penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada kedua orang tua dan seluruh keluarga atas doa, kasih sayang, dukungan, kesabaran, serta semangat yang selalu diberikan kepada penulis selama menempuh pendidikan hingga penyusunan skripsi ini dapat diselesaikan.

Medan, 30 Juni 2026

Penulis

DAFTAR ISI

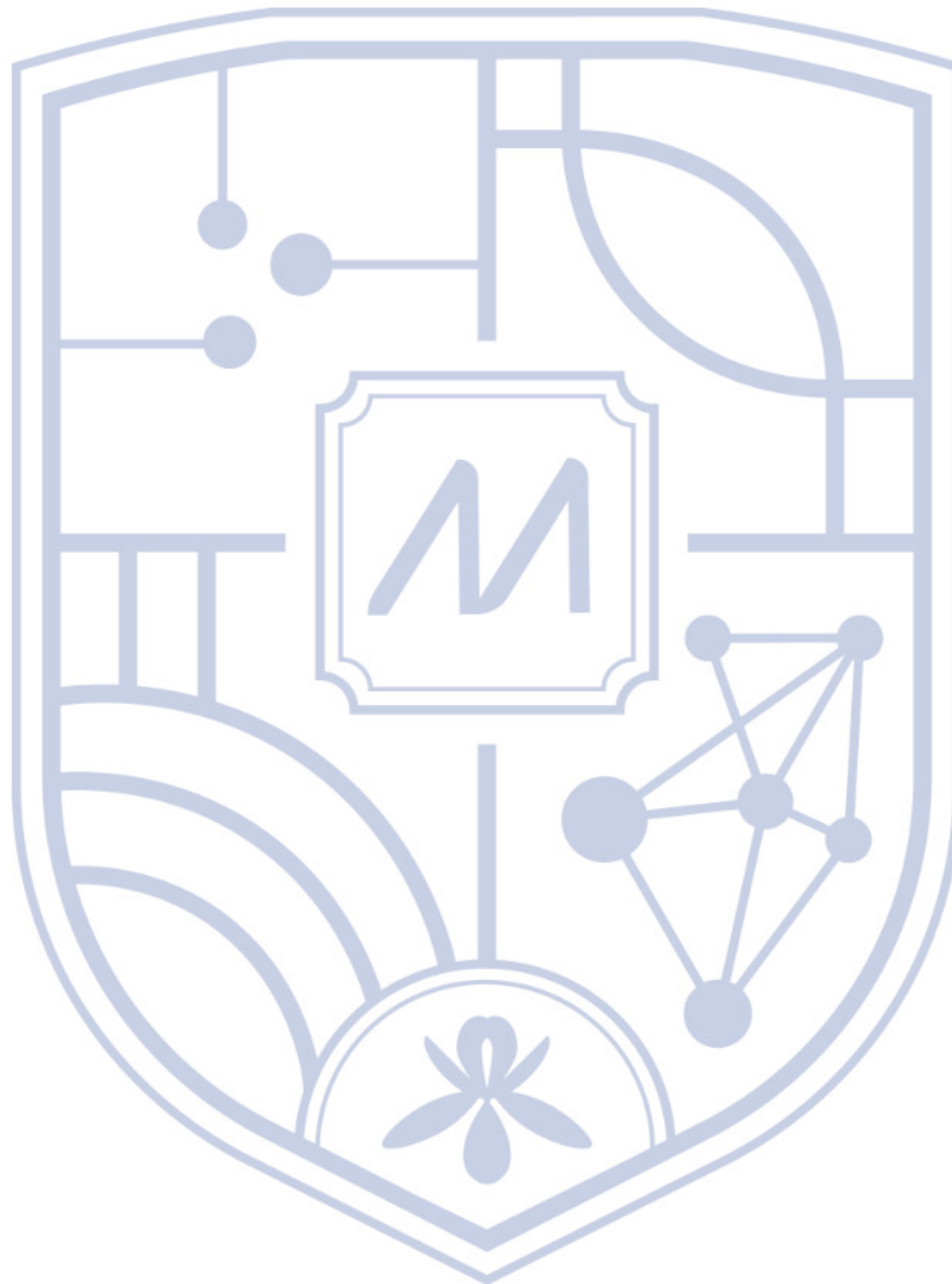
Abstrak	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
DAFTAR GAMBAR	v
DAFTAR TABEL	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
BAB II KAJIAN LITERATUR	5
2.1 Kecerdasan Buatan (Artificial Intelligence/AI)	5
2.1.1 Machine Learning	5
2.1.2 Deep Learning	6
2.1.3 Natural Language Processing (NLP)	6
2.2 Analisis Sentimen	7
2.3 Preprocessing dan Feature Extraction	8
2.4 Support Vector Machine (SVM)	11
2.5 Transformer	12
2.5.1 BERT	13
2.5.2 IndoBERT	14
2.6 Metrik Evaluasi	16
2.7 Penelitian Terdahulu	18
BAB III TAHAPAN PELAKSANAAN	20
3.1 Kerangka Tahapan Pelaksanaan	20
3.2 Metode Pengembangan Perangkat Lunak	22
3.3 Analisis Proses	23
3.3.1 Input Dataset	24
3.3.2 Preprocessing Data	26
3.3.3 Pelabelan Data	28
3.3.4 Pembagian Data	30

3.4 Pembangunan Model Klasifikasi	31
3.4.1 Representasi Fitur	34
3.4.2 Model IndoBERT	36
3.4.3 Model Support Vector Machine (SVM)	36
3.5 Evaluasi Model.....	37
3.6 Perancangan Prototype.....	40
BAB IV HASIL DAN PEMBAHASAN	53
4.1 Hasil Penelitian	53
4.1.1 Pengumpulan dan Distribusi Dataset	53
4.1.2 Pengambilan Data	54
4.1.3 Preprocessing Data.....	57
4.1.4 Pelabelan Data	60
4.1.5 Pembagian Dataset.....	63
4.1.6 Implementasi dan Evaluasi IndoBERT	64
4.1.7 Implementasi dan Evaluasi TF-IDF + SVM.....	68
4.1.8 Evaluasi Perbandingan Model	72
4.1.9 Analisis Sentimen Pemanfaatan ChatGPT.....	73
4.1.10 Implementasi Prototype Sistem	76
4.2 Pembahasan.....	86
4.2.1 Analisis Sentimen Pemanfaatan ChatGPT.....	87
4.2.2 Perbandingan Kinerja Model SVM dan IndoBERT	87
BAB V PENUTUP	89
5.1 Kesimpulan	89
5.2 Saran.....	89
DAFTAR PUSTAKA	91
DAFTAR RIWAYAT HIDUP	99

DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian.....	23
Gambar 3.2 Proses Training dan Evaluasi Model SVM	32
Gambar 3.3 Proses Fine-Tuning dan Evaluasi Model IndoBERT	33
Gambar 3.4 Rancangan Tampilan Beranda.....	41
Gambar 3.5 Rancangan Tampilan Import Dataset	42
Gambar 3.6 Rancangan Tampilan Preprocessing Data	43
Gambar 3.7 Rancangan Tampilan Split Dataset.....	45
Gambar 3.8 Rancangan Halaman Upload	46
Gambar 3.9 Rancangan Tampilan SVM (TF-IDF)	48
Gambar 3.10 Rancangan Tampilan IndoBERT	49
Gambar 3.11 Rancangan Tampilan Hasil dan Perbandingan Model.....	51
Gambar 4.1 Distribusi jumlah data.....	63
Gambar 4.2 Confusin Matrix IndoBERT	66
Gambar 4.3 Grafik Loss Fuction	68
Gambar 4.4 Confusion Matrix SVM.....	71
Gambar 4.5 Visulisasi Perbandingan Kedua Model.....	73
Gambar 4.6 Hasil Sentimen Sebelum label manual (labeling).....	74
Gambar 4.7 Hasil Sentimen Sesudah Labeling manual	74
Gambar 4.8 Tampilan Halaman Beranda	76
Gambar 4.9 Tampilan Halaman Import Dataset.....	77
Gambar 4.10 Tampilan Halaman Preprocessing	78
Gambar 4.11 Tampilan Halaman Pelabelan Data	79
Gambar 4.12 Halaman Upload	79
Gambar 4.13 Hasil Visualisasi Pada Dataset Final	80
Gambar 4.14 Halaman Split Dataset	81
Gambar 4.15 Halaman Klasifikasi SVM (TF-IDF).....	81
Gambar 4.16 Confusion Matrix SVM Streamlit	82
Gambar 4.17 Classification Report Streamlit.....	83
Gambar 4.18 Halaman Klasifikasi IndoBERT	83
Gambar 4.19 Grafik Loss function Indobert	84
Gambar 4.20 Confusion Matrix IndoBERT Streamlit	84
Gambar 4.21 Classification Report Streamlit.....	85

Gambar 4.22 Halaman Perbandingan Model	85
Gambar 4.23 Visualisasi Hasil Perbadningan	86



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	18
Tabel 3.1 Contoh Cuitan.....	24
Tabel 3.2 Contoh hasil Case Folding	27
Tabel 3.3 Contoh hasil Data Cleaning.....	27
Tabel 3.4 Contoh hasil Normalisasi Teks.....	27
Tabel 3.5 Contoh hasil Stopword Removal.....	28
Tabel 3.6 Contoh hasil Tokenization.....	28
Tabel 3.7 Contoh Data Sesudah Dilabel.....	29
Tabel 3.8 Contoh Probabilitas Softmax.....	30
Tabel 3.9 Representasi TF-IDF pada SVM.....	34
Tabel 3.10 Tokenizationi Teks	35
Tabel 3.11 Konversi ke Input Model.....	36
Tabel 3.12 Confusion Matrix (3×3).....	37
Tabel 3.13 Ringkasan nilai TP, FP, FN, TN per kelas	38
Tabel 3.14 Ringkasan Hasil Perhitungan	39
Tabel 4.1 Contoh hasil pengambilan data	54
Tabel 4.2 Contoh hasil Cleaning data.....	57
Tabel 4.3 Contoh hasil Case folding	58
Tabel 4.4 Contoh hasil normalisasi kata.....	58
Tabel 4.5 Contoh hasil Tokenization.....	59
Tabel 4.6 Contoh hasil Stopword removal	59
Tabel 4.7 Contoh Hasil Verifikasi Pelabelan Manual	61
Tabel 4.8 Perbandingan Sebelum dan Sesudah Verifikasi	62
Tabel 4.9 Skenario pembagian data.....	64
Tabel 4.10 Hasil eksperimen Hyperparameter Training	65
Tabel 4.11 Hasil Evaluasi IndoBERT	65
Table 4.12 Hasil Evaluasi Loss Model IndoBERT	68
Tabel 4.13 Inisialisasi Parameter Model SVM.....	69
Table 4.14 Hasil Perhitungan Manual TF-IDF.....	69
Tabel 4.15 Hasil Evaluasi SVM	70
Tabel 4.16 Hasil Perbandingan Performa Model	72