



**ANALISIS SENTIMEN NETIZEN TERHADAP ISU “INDONESIA
GELAP” DI APLIKASI X MENGGUNAKAN METODE NAIVE BAYES
DAN SUPPORT VECTOR MACHINE**

LAPORAN TUGAS AKHIR

BAGAS HAIDAR ARHAB
41521010126

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**ANALISIS SENTIMEN NETIZEN TERHADAP ISU “INDONESIA
GELAP” DI APLIKASI X MENGGUNAKAN METODE NAIVE BAYES
DAN SUPPORT VECTOR MACHINE**

LAPORAN TUGAS AKHIR

BAGAS HAIDAR ARHAB

41521010126

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**UNIVERSITAS
MERCU BUANA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Bagas Haidar Arhab
NIM : 41521010126
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Netizen terhadap Isu “Indonesia Gelap” di Aplikasi X Menggunakan Metode Naive Bayes dan Support Vector Machine

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 19 Juli 2025

UNIVERSITAS
MERCU BUANA



Bagas Haidar Arhab

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Bagas Haidar Arhab
NIM : 41521010126
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Netizen terhadap Isu
"Indonesia Gelap" di Aplikasi X
menggunakan Metode Naive Bayes dan
Support Vector Machine

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Fauzi Nur Iman, S.Kom, M.Kom
NIDN : 0318088903

Ketua Penguji : Afiyati, Dr., S.Si, MT
NIDN : 0316106908

Penguji 1 : Dhanny Permatasari Putri, S.Kom, M
NIDN : 0328087903

Penguji 2 : Dwiki Jatikusumo, S.Kom, M.Kom
NIDN : 0301128903

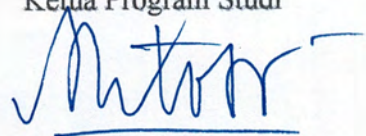
Jakarta, 19 Juli 2025

Mengetahui,

Dekan


Dr. Bambang Jekonowo, S.Si., MTI
NIDN : 0320037002

Ketua Program Studi


Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga saya dapat menyelesaikan proposal penelitian yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Saya menyadari bahwa proposal penelitian ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa saya terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Bapak Fauzi Nur Iman S.Kom., M.Kom. selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini.
5. Ibu Dosen Nur Ani, ST, MMSI selaku Dosen Metodologi Penelitian TI yang telah memberikan saran dan masukan.
6. (Nama dosen penguji) selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
7. Teristimewa kepada orang tua tercinta yang telah memberikan banyak semangat dan dukungan positif baik berupa doa, moral dan terutama materi.
8. Seluruh teman dan sahabat yang telah memberikan motivasi dan dorongan sehingga saya dapat menyelesaikan MPTI ini.
9. Dan semua pihak yang tidak dapat disebutkan satu persatu yang telah memberi bantuan dan dorongan baik moril maupun materil.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 19 Juli 2025



Bagas Haidar Arhab



HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Bagas Haidar Arhab
NIM : 41521010126
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Netizen terhadap Isu
"Indonesia Gelap" di Aplikasi X
Menggunakan Metode Naive Bayes dan
Support Vector Machine

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 19 Juli 2025
Yang menyatakan,
MERCU BUANA



Bagas Haidar Arhab

ABSTRAK

Nama : Bagas Haidar Arhab
NIM : 41521010126
Program Studi : Teknik Informatika
Judul Laporan Skripsi. : Analisis Sentimen Netizen terhadap Isu “Indonesia Gelap” di Aplikasi X menggunakan Metode Naive Bayes dan Support Vector Machine
Dosen Pembimbing : Fauzi Nur Iman, S.Kom., M. Kom

Penelitian ini bertujuan untuk menganalisis sentimen netizen terhadap isu “Indonesia Gelap” di media sosial Twitter serta membandingkan performa dua algoritma klasifikasi teks, yaitu Naive Bayes dan Support Vector Machine (SVM). Data diambil dari tweet yang membahas isu tersebut dan diproses melalui serangkaian tahap pre-processing, termasuk penghapusan karakter khusus, normalisasi, stopword removal, dan tokenisasi. Proses pelabelan sentimen dilakukan secara otomatis menggunakan pendekatan berbasis IndoBERT, yang kemudian diperkuat dengan verifikasi manual untuk meningkatkan akurasi label. Setelah data siap, pembobotan dilakukan dengan menggunakan metode Term Frequency – Inverse Document Frequency (TF-IDF). Menggunakan kedua algoritma Naive Bayes dan SVM, model klasifikasi kemudian dilatih untuk mengukur performa mereka dalam mengklasifikasikan sentimen menjadi tiga kategori: negatif, netral, dan positif. Sementara SVM mengungguli Naive Bayes, kedua algoritma tersebut memiliki performa dan metrik evaluasi yang jauh berbeda satu sama lain. SVM tercatat mendapatkan akurasi 97,9% serta F1-score 72% sedangkan Naive Bayes hanya mendapatkan 94,4% akurasi dan 37% untuk F1-score. Temuan ini menegaskan bahwa algoritma SVM lebih efektif dalam menangani klasifikasi teks berbahasa Indonesia dengan distribusi kelas yang tidak seimbang. Penelitian ini juga membuktikan bahwa Twitter merupakan sumber informasi yang berharga dalam menangkap opini publik terhadap isu-isu sosial dan politik yang berkembang.

Kata kunci: analisis sentimen, Twitter, Indonesia Gelap, Naive Bayes, Support Vector Machine, TF-IDF

ABSTRACT

Name : Bagas Haidar Arhab
Student ID : 41521010126
Study Program : Teknik Informatika
Thesis Tittle : Analisis Sentimen Netizen terhadap Isu “Indonesia Gelap” di Aplikasi X menggunakan Metode Naive Bayes dan Support Vector Machine
Thesis Supervisor : Fauzi Nur Iman, S.Kom., M. Kom

The purpose of this research is to examine the public sentiment on Twitter regarding the “Indonesia Gelap” (“Dark Indonesia”) controversy, as well as to evaluate the performance of two text classification algorithms, Naive Bayes and Support Vector Machine (SVM). The data were collected from tweets discussing the issue and processed through several pre-processing steps, including special character removal, normalization, stopword removal, and tokenization. Sentiment labeling was performed automatically using an IndoBERT-based approach and manually verified to enhance label accuracy. After preprocessing, For feature weighting, TF-IDF (Term Frequency-Inverse Document Frequency) was used. With the dataset thus prepared, classification models based on Naive Bayes and SVM were built with the objective of categorizing tweets into three sentiment levels: negative, neutral, and positive. Based on the evaluation results, SVM performed better than Naive Bayes, achieving 97.9% accuracy and an F1-score of 72%, while Naive Bayes recorded 94.4% accuracy and 37% F1-score. These findings indicate that SVM is more effective in handling sentiment classification for Indonesian-language texts, particularly when the data distribution is imbalanced. Furthermore, this study confirms that Twitter is a valuable source for capturing real-time public opinion on emerging social and political issues.

Key Word: sentiment analysis, Twitter, Indonesia Gelap, Naive Bayes, Support Vector Machine, TF-IDF

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	2
1.5 Batasan Masalah	3
BAB II TINJAUAN PUSTAKA	5
2.1 Penelitian Terdahulu	5
2.2 Teori Pendukung.....	21
2.2.1 Analisis Sentimen	21
2.2.2 Isu “Indonesia Gelap”	22
2.2.3 Text Mining dan Natural Language Processing (NLP).....	22
2.2.4 Algoritma Naive Bayes	22
2.2.5 Algoritma Support Vector Machine (SVM).....	23
2.2.6 TF-IDF (Term Frequency-Inverse Document Frequency).....	24
BAB III METODE PENELITIAN	25
3.1 Jenis Penelitian	25
3.2 Tahapan Penelitian.....	25

BAB IV PEMBAHASAN	28
4.1 Deskripsi Dataset	28
4.2 Pra-pemrosesan Data	29
4.2.1 Cleansing	29
4.2.2 Case folding	30
4.2.3 Stopword Removal	31
4.2.4 Tokenization	32
4.3 Labelling	33
4.4 TF - IDF	35
4.5 Implementasi Naive Bayes dan Support Vector Machine	37
4.5.1 Naive Bayes Classifier	37
4.5.2 Support Vector Machine (SVM)	38
4.6 Evaluasi dan Perbandingan Model.....	39
4.7 Analisis Hasil	40
BAB V KESIMPULAN DAN SARAN.....	42
5.1 Kesimpulan	42
5.2 Saran	44
DAFTAR PUSTAKA.....	46
LAMPIRAN.....	49



 UNIVERSITAS
 MERCU BUANA

DAFTAR TABEL

Table 2.1. 1 Penelitian Terkait	5
Table 4.6. 1 Hasil Evaluasi	40
Table 4.7. 1 Analisi Hasil	41



DAFTAR GAMBAR

Gambar 2.2.4. 1 Rumus NB.....	23
Gambar 2.2.5. 1 Rumus SVM.....	23
Gambar 2.2.6. 1 Rumus TF - IDF	24
Gambar 3.2. 1.....	27
Gambar 4. 1 Proses Crawling	28
Gambar 4.2.1. 1 Proses Cleansing	29
Gambar 4.2.1. 2 Hasil Cleansing	30
Gambar 4.2.2. 1 Case Folding	30
Gambar 4.2.2. 2 Hasil Case Folding	30
Gambar 4.2.3. 1 Install Library.....	31
Gambar 4.2.3. 2 Proses Stopword.....	31
Gambar 4.2.3. 3 Hasil Stopword.....	32
Gambar 4.2.4. 1 Proses Tokenisasi.....	32
Gambar 4.2.4. 2 Hasil Tokenisasi.....	33
Gambar 4.3. 1 Setup Analisis Sentimen	34
Gambar 4.3. 2 Proses Pelabelan.....	34
Gambar 4.3. 3 Hasil Pelabelan.....	35
Gambar 4.4. 1 Persiapan Data.....	35
Gambar 4.4. 2 Proses TF - IDF	36
Gambar 4.4. 3 Bentuk Hasil.....	36
Gambar 4.5.1. 1 Latih Model NB	38
Gambar 4.5.1. 2 Latih Model SVM	39

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi	49
Lampiran 2 Curriculum Vitae	51
Lampiran 3 Sertifikasi BNSP.....	54
Lampiran 4 Hasil Cek Turnitin	57

