



**PERBANDINGAN DUA EKSTRAKSI FITUR UNTUK ANALISIS
SENTIMEN MENGGUNAKAN NAIVE BAYES PADA MEDIA
SOSIAL STUDI KASUS PEJABAT PUBLIK LULUS S3
KURANG DARI DUA TAHUN**

LAPORAN TUGAS AKHIR

HALAMAN JUDUL

**FERDIANSYAH SAPUTRA
41521010207**

**UNIVERSITAS
MERCU BUANA**

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



**PERBANDINGAN DUA EKSTRAKSI FITUR UNTUK ANALISIS
SENTIMEN MENGGUNAKAN NAIVE BAYES PADA MEDIA
SOSIAL STUDI KASUS PEJABAT PUBLIK LULUS S3
KURANG DARI DUA TAHUN**

LAPORAN TUGAS AKHIR

FERDIANSYAH SAPUTRA

41521010207

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

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

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Ferdiansyah Saputra
NIM : 41521010207
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Dua Ekstraksi Fitur Untuk Analisis Sentimen Menggunakan Naïve Bayes Pada Media Sosial Studi Kasus Pejabat Publik Lulus S3 Kurang Dari Dua Tahun.

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, 24 Juli 2025

MACEANX017050300
Ferdiansyah Saputra.

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Ferdiansyah Saputra
NIM : 41521010207
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Dua Ekstraksi Fitur Untuk Analisis Sentimen Menggunakan Naïve Bayes Pada Media Sosial Studi Kasus Pejabat Publik Lulus S3 Kurang Dari Dua Tahun.

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

Disahkan oleh:

Pembimbing : Dr.Bagus Priambodo, S.T., M. TI
NIDN : 0313057905
Ketua Penguji : Inna Sabily Karima, S.Kom, M.Kom
NIDN : 0324018902
Penguji 1 : Misni, S.Kom, M.Kom
NIDN : 0413046802
Penguji 2 : Lukman Hakim, S.T., M.Kom.
NIDN : 0327107701



Jakarta, 24 Juli 2025

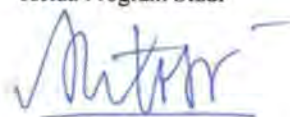
Mengetahui,

Dekan

Ketua Program Studi



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



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 penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis 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 Mercubwana.
4. Bapak Dr.Bagus Priambodo, S.T., M. TI. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubwana..
6. Kepada Dhea Putri Aulia, terima kasih atas segala kasih sayang, kesabaran, serta dukungan yang tidak pernah putus. Kehadiranmu telah menjadi penyemangat dalam setiap proses yang penulis jalani, baik dalam suka maupun duka.

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, 23 Juni 2025

Ferdiansyah Saputra

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR UNTUK KEPENTINGAN AKADEMIS**

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

Nama : Ferdiansyah Saputra
NIM : 41521010207
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Dua Ekstraksi Fitur Untuk Analisis Sentimen Menggunakan Naïve Bayes Pada Media Sosial Studi Kasus Pejabat Publik Lulus S3 Kurang Dari Dua Tahun.

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, 24 Juli 2025

Yang menyatakan,



Ferdiansyah Saputra

ABSTRAK

Nama : Ferdiansyah Saputra
NIM : 41521010207
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Dua Ekstraksi Fitur Untuk Analisis Sentimen Menggunakan Naïve Bayes Pada Media Sosial Studi Kasus Pejabat Publik Lulus S3 Kurang Dari Dua Tahun.
Dosen Pembimbing : Dr.Bagus Priambodo, S.T., M. TI

Media sosial telah menjadi saluran utama masyarakat dalam menyampaikan opini terhadap isu-isu publik, termasuk pencapaian akademik pejabat negara. Salah satu isu yang ramai diperbincangkan adalah disertasi Menteri Energi dan Sumber Daya Mineral yang diselesaikan dalam waktu kurang dari dua tahun. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap disertasi Menteri Energi dan Sumber Daya Mineral, yang diselesaikan dalam waktu 1 tahun 8 bulan. Data yang digunakan dalam penelitian ini diambil dari platform media sosial Twitter, dengan total 2100 komentar yang dikumpulkan menggunakan teknik crawling berdasarkan kata kunci "Disertasi Menteri Energi Dan Sumber Daya Mineral". Metode yang diterapkan dalam penelitian ini adalah analisis sentimen menggunakan algoritma Naive Bayes, dengan perbandingan dua metode ekstraksi fitur, yaitu Term Frequency-Inverse Document Frequency (TF-IDF) dan Bag of Words (BoW). Proses penelitian meliputi tahapan pengumpulan data, preprocessing, pelabelan sentimen, dan evaluasi model. Hasil evaluasi menunjukkan bahwa model Naive Bayes dengan metode BoW memberikan akurasi yang lebih tinggi (83.01%) dibandingkan dengan TF-IDF (81.76%), serta menunjukkan performa yang lebih stabil dalam klasifikasi sentimen positif dan negatif. Penelitian ini diharapkan dapat memberikan wawasan yang lebih dalam mengenai persepsi masyarakat terhadap prestasi akademis pejabat publik dan implikasinya terhadap kepercayaan masyarakat. Selain itu, hasil penelitian ini dapat menjadi referensi bagi penelitian selanjutnya dalam bidang analisis sentimen di media sosial.

Kata kunci: Analisis Sentimen, Naïve Bayes, TF-IDF, BoW, BERT.

ABSTRACT

Nama : Ferdiansyah Saputra
NIM : 41521010207
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Comparison of Two Feature Extractions for Sentiment Analysis Using Naïve Bayes on Social Media Case Study of Public Officials Who Graduated with Doctorate in Less Than Two Years.
Dosen Pembimbing : Dr.Bagus Priambodo, S.T., M. TI

Social media has become the main channel for people to express their opinions on public issues, including the academic achievements of state officials. One of the issues that is widely discussed is the dissertation of the Minister of Energy and Mineral Resources which was completed in less than two years. This study aims to analyze public sentiment towards the dissertation of the Minister of Energy and Mineral Resources, which was completed in 1 year and 8 months. The data used in this study were taken from the social media platform Twitter, with a total of 2100 comments collected using the crawling technique based on the keyword "Disertasi Menteri Energi Dan Sumber Daya Mineral". The method applied in this study is sentiment analysis using the Naïve Bayes algorithm, with a comparison of two feature extraction methods, namely Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words (BoW). The research process includes the stages of data collection, preprocessing, sentiment labeling, and model evaluation. The evaluation results show that the Naïve Bayes model with the BoW method provides higher accuracy (83.01%) compared to TF-IDF (81.76%), and shows more stable performance in the classification of positive and negative sentiments. This study is expected to provide deeper insight into public perceptions of public officials' academic achievements and their implications for public trust. In addition, the results of this study can be a reference for further research in the field of sentiment analysis on social media.

Kata kunci: Sentiment Analysis, Naïve Bayes, TF-IDF, BoW, BERT.

DAFTAR ISI

HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
1.6 Dampak Penulisan Penelitian	5
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu	6
2.2 Teori Pendukung	25
BAB III METODE PENELITIAN	29
3.1 Jenis Penelitian.....	29
3.2 Tahapan Penelitian.....	29
BAB IV PEMBAHASAN	34
4.1 Dataset.....	34
4.2 Pre-processing Data	35
4.3 Labelling Data.....	39
4.4 Pembagian Data	43
4.5 Penerapan Ekstraksi Fitur	45
4.6 Penerapan Algoritma Naïve Bayes	47
4.7 Evaluasi Hasil	49

BAB V KESIMPULAN DAN SARAN	56
5.1 Kesimpulan	56
5.2 Saran	56
DAFTAR PUSTAKA.....	58
LAMPIRAN.....	63



DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait.....	6
Tabel 4. 1 Tabel Data Cleansing.....	35
Tabel 4. 2 Tabel Case Folding	36
Tabel 4. 3 Tabel Tokenization	37
Tabel 4. 4 Tabel Stopword.....	38
Tabel 4. 5 Stemming	39
Tabel 4. 6 Tabel Kelas Labelling	39
Tabel 4. 7 Kalimat Yang Mengandung Sarkasme	40
Tabel 4. 8 Label Yang Sudah Di Perbaiki	42
Tabel 4. 9 Label Positif	43
Tabel 4. 10 Tabel Perbandingan Ekstraksi Fitur.....	54



DAFTAR GAMBAR

Gambar 3. 1 Diagram Alir	30
Gambar 4. 1 Source Code Crawling Data.....	34
Gambar 4. 2 Hasil Crawling Data.....	34
Gambar 4. 3 Source Code Data Cleansing.....	35
Gambar 4. 4 Source Code Case Folding.....	36
Gambar 4. 5 Source Code Tokenization	37
Gambar 4. 6 Source Code Stopword.....	37
Gambar 4. 7 Source Code Stemming.....	38
Gambar 4. 8 Labelling Menggunakan BERT	40
Gambar 4. 9 Pie Chart Sentiment	43
Gambar 4. 10 Source Code Split Data	44
Gambar 4. 11 Source Code TF-IDF.....	45
Gambar 4. 12 Source Code Bag of Word	46
Gambar 4. 13 Hasil Ekstraksi Fitur.....	46
Gambar 4. 14 Source Code NB dengan TF-IDF.....	48
Gambar 4. 15 Source Code NB dengan BoW.....	49
Gambar 4. 16 Confusion Matrix NB dengan TF-IDF.....	50
Gambar 4. 17 Confusion Matrix dengan BoW	50
Gambar 4. 18 Hasil NB dengan TF-IDF.....	51
Gambar 4. 19 Hasil NB dengan Bag of Word (BoW)	53
Gambar 4. 20 Diagram Perbandingan Label Positif	54
Gambar 4. 21 Diagram Perbandingan Label Negatif.....	55



DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi.....	63
Lampiran 2 Curriculum Vitae	64
Lampiran 3 Surat Pernyataan HAKI	65
Lampiran 4 BNSP	67
Lampiran 5 Hasil Cek Turnitin.....	69
Lampiran 6 Form Revisi Dosen Penguji.....	70
Lampiran 7 Similarity Check	72
Lampiran 8 Lembar Persetujuan Sidang	73

