



**ENSEMBLE MODEL NAIVE BAYES DAN RANDOM FOREST
UNTUK KLASIFIKASI SENTIMEN DAN PREDIKSI ULASAN
PENGUNA APLIKASI GOOGLE GEMINI**

LAPORAN TUGAS AKHIR

Jesica Walanda Haryanti

41821010021

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA**

2025



**ENSEMBLE MODEL NAIVE BAYES DAN RANDOM FOREST
UNTUK KLASIFIKASI SENTIMEN DAN PREDIKSI ULASAN
PENGGUNA APLIKASI GOOGLE GEMINI**

LAPORAN TUGAS AKHIR

Jesica Walanda Haryanti

41821010021

**UNIVERSITAS
MERCU BUANA**

Diajukan untuk melengkapi salah satu syarat memperoleh gelar sarjana

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA**

2025

HALAMAN PENYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Jesica Walanda Haryanti

NIM : 41821010021

Program Studi : Sistem Informasi

Judul Laporan Skripsi : Ensemble Model Naive Bayes Dan Random Forest Untuk Klasifikasi Sentimen Dan Prediksi Ulasan Pengguna Aplikasi Google Gemini

Menyatakan bahwa Laporan Aplikatif/Tugas Akhir/Jurnal/Media Ilmiah 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 Tugas Akhir saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 28 Juli 2025



Jesica Walanda Haryanti

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

Nama Mahasiswa : Jesica Walanda Haryanti
 NIM : 41821010021
 Judul Tugas Akhir : Ensemble Model Naive Bayes dan Random Forest
 untuk Klasifikasi Sentimen dan Prediksi Ulasan
 Pengguna Aplikasi Google Gemini

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Mercu Buana.

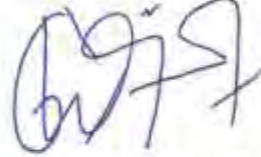
Jakarta, 28-Juli-2025

Menyetujui,

Pembimbing	:	Ifan Prihandi, S.Kom, M.Kom	()
NIDN	:	0313098901	
Ketua Penguji	:	Wawan Gunawan, S.Kom, MT., M.Kom	()
NIDN	:	0424108104	
Penguji 1	:	Abdi Wahab, S.Kom, MT	()
NIDN	:	0305068502	
Penguji 2	:	Yudo Devianto, S.Kom, M.Kom	()
NIDN	:	0315127303	

Mengetahui,


Dr. Bambang Jekonowo, S.Si., M.T.I.
 Dekan Fakultas Ilmu Komputer


Wawan Gunawan, S.Kom, MT., M.Kom.
 Ka.Prodi Sistem Informasi

KATA PENGANTAR

Puji syukur kita panjatkan ke hadirat Tuhan Yang Maha Esa, atas limpahan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir dengan judul "Ensemble Model Naive Bayes dan Random Forest untuk Klasifikasi Sentimen dan Prediksi Ulasan Pengguna Aplikasi Google Gemini".

Tugas akhir ini disusun untuk memenuhi salah satu syarat kelulusan pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Mercu Buana. Dalam proses penyusunan tugas akhir ini, penulis menyadari bahwa pencapaian ini tidak lepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

- 1) Bapak Ifan Prihandi, S.Kom., M.Kom., selaku Dosen Pembimbing Tugas Akhir, atas arahan, masukan, serta bimbingannya yang sangat berarti selama proses penyusunan tugas akhir ini.
- 2) Bapak Bagus Priambodo, ST, M.TI., Ph.D., selaku Dosen Pembimbing Akademik, Terimakasih atas arahan, dan masukan yang sangat berarti selama proses perkuliahan.
- 3) Ibu Dr. Ruci Meiyanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi, yang telah memberikan fasilitas, dukungan, dan motivasi selama penulis menempuh studi.
- 4) Untuk seluruh Dosen Fasilkom Universitas Mercu Buana Terimakasih atas pembelajaran selama diperkuliahan.
- 5) Untuk cinta pertama saya, Papa Haryanto, terima kasih atas perjuangan dan pengorbanannya demi kehidupan penulis. Beliau telah menjadi sosok pendidik, motivator, dan penyemangat yang luar biasa hingga penulis mampu menyelesaikan pendidikan sampai jenjang sarjana.
- 6) Untuk Mama tercinta, Titin Sumarni, terima kasih atas cinta, doa, dukungan, dan semangat yang tak pernah putus, yang menjadi kekuatan besar bagi penulis dalam menyelesaikan pendidikan ini.

- 7) Untuk Tante tercinta, Rita Sagita, terima kasih atas doa, dukungan, dan bantuan finansial selama masa perkuliahan. Tanpa bantuan dan ketulusan hati Tante, penulis tidak akan sampai pada titik ini.
- 8) Untuk kedua adik penulis, Jason Kaymala Ganesha dan Melvin Irawan Putra, terima kasih atas doa dan dukungan kalian yang selalu menguatkan langkah penulis.
- 9) Untuk seseorang yang sangat penulis cintai, Gusti Ngurah Adi Premana, terima kasih atas kehadiranmu yang berarti dalam perjalanan hidup ini. Terima kasih telah memberikan semangat, dukungan, dan perjuangan bersama hingga penulis bisa sampai pada tahap ini.
- 10) Untuk diri saya sendiri, apresiasi sebesar-besarnya atas segala perjuangan yang telah dilalui untuk menyelesaikan apa yang telah dimulai. Terima kasih karena telah bertahan, dan terus berusaha, meskipun sering kali dihadapkan pada keputusan. Tetaplah menjadi manusia yang tidak lelah untuk mencoba dan belajar, apa pun yang terjadi.

Akhir kata, penulis berharap tugas akhir ini dapat memberikan manfaat bagi pembaca, khususnya dalam pengembangan teknologi informasi dan pengolahan data sentimen. Penulis juga terbuka terhadap kritik dan saran demi perbaikan di masa mendatang.

UNIVERSITAS
MERCU BUANA

Jakarta, 17 Juli 2025

Jesica Walanda Haryanti

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

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

Nama : Jesica Walanda Haryanti
NIM : 41821010021
Program Studi : Sistem Informasi
Judul Laporan Skripsi : Ensemble Model Naive Bayes Dan Random Forest Untuk Klasifikasi Sentimen Dan Prediksi Ulasan Pengguna Aplikasi Google Gemini

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, 28 Juli 2025

Yang menyatakan,



(Jesica Walanda Haryanti)

ABSTRAK

Nama Mahasiswa : Jesica Walanda Haryanti
NIM Mahasiswa : 41821010021
Pembimbing TA : Ifan Prihandi, S.Kom, M.Kom
Ensemble Model Naive Bayes Dan Random Forest
Judul Tugas Akhir : Untuk Klasifikasi Sentimen Dan Prediksi Ulasan
Pengguna Aplikasi Google Gemini

Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi Google Gemini di Google Play Store dan mengevaluasi tren persepsi harian menggunakan pendekatan machine learning. Variabel yang diteliti adalah sentimen pengguna, dikategorikan menjadi positif dan negatif. Sebanyak 3.000 ulasan dikumpulkan menggunakan metode purposive sampling melalui teknik web scraping dengan Python di Google Colaboratory. Data diproses melalui tahapan preprocessing mencakup normalisasi teks, tokenisasi, penghapusan stopword, dan stemming. Algoritma Multinomial Naive Bayes digunakan untuk klasifikasi awal, kemudian ditingkatkan melalui teknik SMOTE untuk penyeimbangan data dan model ensemble berbasis soft voting yang menggabungkan Naive Bayes dan Random Forest. Hasil analisis menunjukkan bahwa model Naive Bayes menghasilkan akurasi 80% namun kurang sensitif terhadap kelas negatif. Model ensemble meningkatkan akurasi menjadi 85% dan F1-score makro sebesar 0.80. Tren sentimen harian selama satu bulan terakhir juga berhasil dimodelkan dengan akurasi prediksi 85,3%. Kesimpulan penelitian ini menunjukkan bahwa pendekatan ensemble dan balancing data efektif dalam meningkatkan performa klasifikasi dan relevan untuk pemantauan opini pengguna secara berkala.

Kata Kunci: Analisis Sentimen, Google Gemini, Naive Bayes, Ensemble

ABSTRACT

Nama Mahasiswa : Jesica Walanda Haryanti
NIM Mahasiswa : 41821010021
Pembimbing TA : Ifan Prihandi, S.Kom, M.Kom
Ensemble Model Naive Bayes Dan Random Forest Untuk
Judul Tugas Akhir : Klasifikasi Sentimen Dan Prediksi Ulasan Pengguna
Aplikasi Google Gemini

This study aims to analyze user sentiment toward the Google Gemini application on the Google Play Store and to evaluate daily perception trends using a machine learning approach. The variable examined is user sentiment, categorized into positive and negative classes. A total of 3,000 reviews were collected using purposive sampling through web scraping with Python on Google Colaboratory. The data underwent preprocessing steps including text normalization, tokenization, stopword removal, and stemming. The initial classification used the Multinomial Naive Bayes algorithm, enhanced through data balancing with the SMOTE technique and an ensemble model combining Naive Bayes and Random Forest with soft voting. The Naive Bayes model achieved 80% accuracy but had low recall for negative reviews. The ensemble model improved accuracy to 85% and reached a macro F1-score of 0.80. Furthermore, the daily sentiment trend over the past month was successfully modeled with a prediction accuracy of 85.3%. The study concludes that ensemble and data balancing methods are effective in improving sentiment classification performance and are suitable for real-time user opinion monitoring.

Keywords: Sentiment Analysis, Google Gemini, Naive Bayes, Ensemble

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
LEMBAR PENGESAHAN	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI.....	ix
DAFTAR TABLE	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Maslaah	2
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Google Gemini.....	5
2.2 Analisis Sentimen	5
2.3 Preprocessing Teks	5
2.4 Algoritma Naive Bayes	6
2.5 Random Forest.....	7
2.6 Inverse Document Frequency (IDF)	8
2.7 SMOTE (Synthetic Minority Over-sampling Technique)	9
2.8 Ensemble Learning (Voting Classifier)	10
2.9 Evaluasi Model	11
2.10 Analisis Tren Sentime Berbasis Waktu	12
2.11 Penelitian Terdahulu	13
2.11.1 Analisis Literatur Review	18
2.11.2 Kontribusi Penulis.....	19
BAB III METODOLOGI PENELITIAN	20

3.1 Deskripsi Sumber Data	20
3.2 Teknik Pengumpulan Data.....	20
3.3 Digram Alir Penelitian.....	20
3.3.1 Proses Pengumpulan dan Pra-pemrosesan Data	21
3.3.2 Pembangunan dan Pelatihan Model Klasifikasi.....	22
3.3.3 Penanganan Ketidakseimbangan Data dan Ensemble Learning	22
3.3.4 Analisis Tren Sentimen Harian (1 Bulan Terakhir)	23
3.4 Pemogramaan.....	24
3.4.1 Instalasi dan Persiapan Library	24
3.4.2 Pengambilan dan Penyimpanan Data Ulasan.....	24
3.4.3 Pelabelan Sentimen	24
3.4.4 Preprocessing Teks	25
3.4.5 Visualisasi Awal	25
3.4.6 Pelatihan dan Evaluasi Model Naive Bayes.....	25
3.4.7 Model Ensemble dan Penyeimbangan Data.....	26
3.4.8 Analisis dan Visualisasi Tren Sentimen Harian.....	26
BAB IV HASIL DAN PEMBAHASAN	27
4.1 Tahap Pra-pemrosesan Data.....	27
4.1.1 Pengumpulan Data dan Persiapan Awal untuk Analisis	27
4.1.2 Pemberian Label Sentimen	27
4.1.3 Pembersihan Teks	28
4.2 Word Cloud.....	29
4.3 Pemodelan dan Evaluasi Naive Bayes.....	30
4.4 Model Ensemble (Naive Bayes + Random Forest).....	31
4.5 Analisis Tren Sentimen Harian.....	33
4.6 Pembahasan.....	34
4.6.1 Kontribusi Penelitian.....	34
4.6.2 Kekuatan Penelitian	35
4.6.3 Tantangan Penelitian.....	35
4.6.4 Perbandingan Hasil Penelitian dengan Penelitian Terdahulu	36
BAB V PENUTUP	37
5.1 Kesimpulan	37
5.2 Saran	37
DAFTAR PUSTAKA	38
LAMPIRAN.....	43

DAFTAR TABLE

Table 2.1 Confusion Matrix	11
Table 2.2 Ukuran Evaluasi Model Klasifikasi	12
Table 2.3 Penelitian Terdahulu.....	13
Table 4.1 Distribusi Data Label Sentimen	28
Table 4.2 Hasil Evaluasi Model Naive Bayes	30
Table 4.3 Hasil Evaluasi Ensemble Model	32
Table 4.4 Agregasi Sentimen Harian Aktual dan Prediksi Ensemble	33
Table 4.5 Perbandingan Hasil	36



DAFTAR GAMBAR

Gambar 3.1 Alur Diagram Penelitian.....	21
Gambar 3.2 Arsitektur Model Ensemble	23
Gambar 4.1 Data Awal Pengguna	27
Gambar 4. 2 Data Awal Pengguna Tahap Pembersihan	28
Gambar 4. 3 Hasil Casefolding	28
Gambar 4. 4 Hasil Clening.....	28
Gambar 4. 5 Hasil Token.....	29
Gambar 4. 6 Hasil Stopwords	29
Gambar 4. 7 Hasil Stemming.....	29
Gambar 4. 8 Hasil Final	29
Gambar 4. 9 Word Cloud Ulasan	29
Gambar 4.10 Confusion Matrix Naive Bayes	31
Gambar 4.11 Confusion Matrix Naive Bayes	32
Gambar 4.12 Sentimen Harian Aktual vs Prediksi Ensemble.....	34

DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi Bimbingan	43
Lampiran 2. Curriculum Vitae	44
Lampiran 3. Surat Keterangan Telah Mengikuti BNSP	45
Lampiran 4. Hasil Cek Turnitin	46

