



**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI  
GOJEK MENGGUNAKAN ALGORITMA NAIVE BAYES DAN  
K-NEAREST NEIGHBOR**

**LAPORAN TUGAS AKHIR**

**HALAMAN JUDUL**

**RAFFI ZAIDAN ATHALLA**

**41522010268**

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

**JAKARTA**

**2026**



**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI  
GOJEK MENGGUNAKAN ALGORITMA NAIVE BAYES  
DAN K-NEAREST NEIGHBOR**

**LAPORAN TUGAS AKHIR**

**RAFFI ZAIDAN ATHALLA  
41522010268**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

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

**2026**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Raffi Zaidan Athalla  
NIM : 41522010268  
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:

"Analisis Sentimen Ulasan Pengguna Aplikasi Gojek Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbor" adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

UNIVERSITAS  
MERCU BUANA

Jakarta, 20 Januari 2026

  
METERAI  
TEMPEL  
PAS 45 ANX 24 00502  
Raffi Zaidan Athalla

**PERNYATAAN SIMILARITY CHECK**  
*(SIMILARITY CHECK STATEMENT)*

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh  
*The undersigned, hereby declare that the scientific work written by*

Nama /Name : Raffi Zaidan Athalla  
NIM /Student ID Number : 41522010268  
Program Studi /Study Program : Teknik Informatika

Dengan Judul Tugas Akhir

*(The title)*

**“ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI GOJEK  
MENGUNAKAN ALGORITMA NAIVE BAYES DAN K-NEAREST NEIGHBOR”**

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

*Similarity checks have been carried out with the Turnitin system on the date:*

31 Desember 2025

dengan nilai persentase sebesar :

*/with a percentage value of:*

30%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer**  
Universitas Mercu Buana. *declared to meet standards in accordance with applicable*  
*regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

File hasil cek similarity turnitin:

*(Turnitin similarity report file)*

[https://drive.google.com/file/d/1AmssK7B2Hldbfy3JrfTRMux5iRdSyE/view?usp=drive\\_mk](https://drive.google.com/file/d/1AmssK7B2Hldbfy3JrfTRMux5iRdSyE/view?usp=drive_mk)



Jakarta, 31 Desember 2025  
Admin Turnitin Fasilkom UMB



**Agung Prawoto, S.Kom., B.Sc**

NIK : 322970503

## HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : Raffi Zaidan Athalla  
NIM : 41522010268  
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika  
Judul Tugas Akhir : Analisis Sentimen Ulasan Pengguna Aplikasi Gojek Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbor

Telah berhasil dipertahankan pada sidang tanggal 20 Januari 2026 dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing



Sabar Rudiarto, S.Kom., M.Kom.

NIDN/NUPTK: 0309036902

UNIVERSITAS  
MERCU BUANA

Jakarta, 20 Januari 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer



Dr. Bambang Jokonowo, S.Si., MTI  
NIDN/NUPTK: 0320037002

Ketua Program Studi  
Teknik Informatika



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

## KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini. Penulisan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Teknik Informatika pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Tugas Akhir ini, sangatlah sulit bagi saya untuk menyelesaikan Tugas Akhir ini. Oleh karena itu, 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 Mercu Buana.
4. Bapak/Ibu Dosen Pembimbing 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 mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercu Buana.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini membawa manfaat bagi pengembangan ilmu.



Jakarta, 20 Desember 2025

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB

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

Nama : Raffi Zaidan Athalla  
NIM : 41522010268  
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika  
Judul Laporan Skripsi : Analisis Sentimen Ulasan Pengguna Aplikasi Gojek Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbor

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 Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 20 Januari 2026

Yang menandatangani,

UNIVERSITAS  
MERCU BUANA



Raffi Zaidan Athalla

# ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI GOJEK MENGUNAKAN ALGORITMA NAIVE BAYES DAN K-NEAREST NEIGHBOR

RAFFI ZAIDAN ATHALLA

## ABSTRAK

Penelitian ini bertujuan untuk membandingkan kinerja algoritma *Naive Bayes* dan *K-Nearest Neighbor (KNN)* dalam menganalisis sentimen ulasan pengguna aplikasi Gojek yang diambil dari Google Play Store sebanyak 20.000 data ulasan. Setiap ulasan diproses melalui tahapan *text preprocessing* yang meliputi *case folding*, *tokenizing*, *stopword removal*, *stemming*, serta pembobotan menggunakan metode *Term Frequency–Inverse Document Frequency (TF-IDF)*. Kedua algoritma diuji dengan pendekatan *supervised learning* menggunakan rasio pembagian data 80% untuk pelatihan dan 20% untuk pengujian. Evaluasi performa dilakukan menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa algoritma *Naive Bayes* menghasilkan performa yang lebih baik dengan akurasi mencapai **89%**, sedangkan *K-Nearest Neighbor (KNN)* hanya mencapai **69%**. *Naive Bayes* juga lebih stabil dalam mengenali sentimen positif dan negatif serta memiliki efisiensi komputasi yang lebih tinggi. Sebaliknya, *KNN* lebih sensitif terhadap distribusi data yang tidak seimbang dan membutuhkan waktu komputasi yang lebih lama. Berdasarkan hasil tersebut, dapat disimpulkan bahwa algoritma *Naive Bayes* lebih unggul dan sesuai untuk digunakan dalam klasifikasi sentimen ulasan pengguna Gojek berskala besar, sementara *KNN* dapat menjadi alternatif untuk dataset yang lebih kecil dan seimbang.

**Kata kunci:** *Naive Bayes*, *K-Nearest Neighbor*, Analisis Sentimen, Ulasan Gojek, *Machine Learning*, TF-IDF.

# ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI GOJEK MENGUNAKAN ALGORITMA NAIVE BAYES DAN K-NEAREST NEIGHBOR

RAFFI ZAIDAN ATHALLA

## ABSTRACT

*This study aims to compare the performance of the Naive Bayes and K-Nearest Neighbor (KNN) algorithms in sentiment analysis of Gojek user reviews obtained from Google Play Store, consisting of 20,000 data entries. Each review underwent several text preprocessing stages including case folding, tokenizing, stopword removal, stemming, and weighting using the Term Frequency–Inverse Document Frequency (TF-IDF) method. Both algorithms were tested using a supervised learning approach with an 80:20 ratio for training and testing data. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. The experimental results indicate that the Naive Bayes algorithm achieved superior performance with an accuracy of 93%, while K-Nearest Neighbor (KNN) reached 69%. Naive Bayes also demonstrated higher stability in identifying positive and negative sentiments and provided better computational efficiency. Conversely, KNN was more sensitive to unbalanced data distribution and required longer computation time. Based on these findings, it can be concluded that Naive Bayes is more suitable for large-scale sentiment classification of Gojek user reviews, whereas KNN serves as an alternative for smaller and more balanced datasets.*

**Keywords:** *Naive Bayes, K-Nearest Neighbor, Sentiment Analysis, Gojek Reviews, Machine Learning, TF-IDF.*

## DAFTAR ISI

HALAMAN JUDUL .....	0
LAPORAN TUGAS AKHIR .....	i
HALAMAN PERNYATAAN KARYA SENDIRI .....	ii
HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN .....	iii
HALAMAN PENGESAHAN .....	iv
KATA PENGANTAR .....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASITUGAS AKHIR DI REPOSITORI UMB .....	vi
ABSTRAK .....	vii
<i>ABSTRACT</i> .....	viii
DAFTAR ISI .....	ix
DAFTAR TABEL .....	xi
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN .....	xiii
BAB I .....	1
PENDAHULUAN .....	1
1.1 Latar Belakang Penelitian .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	5
1.4.1 Manfaat Teoretis .....	5
1.4.2. Manfaat Praktis .....	5
1.5 Batasan Masalah .....	6
BAB II .....	8
TINJAUAN PUSTAKA .....	8
2.1 Teori Utama .....	8
2.1.1 Algoritma Naive Bayes .....	8
2.1.2 Proses Klasifikasi dengan Naive Bayes .....	9
2.1.3 Keunggulan dan Kelemahan Algoritma Naive Bayes .....	11
2.2.1 Pengertian Algoritma K-Nearest Neighbor (KNN) .....	13
2.2.2 Kelebihan dan Kelemahan Algoritma K-Nearest Neighbor (KNN) .....	14
2.2.3 Proses Klasifikasi dengan Algoritma K-Nearest Neighbor (KNN) .....	17
2.2 Teori Pendukung .....	19
2.3 Penelitian Terdahulu .....	20
2.4 Gap Penelitian .....	25

BAB III .....	26
METODE PENELITIAN .....	26
3.1 Jenis Penelitian .....	26
3.2 Desain Penelitian .....	28
3.2.1 Pengumpulan Data (Data Collection) .....	28
3.2.2 Pemahaman dan Eksplorasi Data (Data Understanding & Exploration) .....	29
3.2.3 Pra-Pemrosesan Data (Data Preprocessing) .....	29
3.2.4 Pelabelan Sentimen (Sentiment Labeling) .....	30
3.2.5 Ekstraksi Fitur (Feature Extraction) .....	30
3.2.6 Pembagian Data (Training dan Testing Data) .....	31
3.2.7 Pelatihan Model Klasifikasi .....	31
3.2.8 Evaluasi Model .....	32
3.2.9 Pengujian Model pada Data Baru .....	32
BAB IV .....	33
PEMBAHASAN .....	33
4.1 Gambaran Umum Hasil Penelitian .....	33
4.2 Distribusi Data dan Karakteristik Sentimen .....	33
4.3 Hasil Klasifikasi Menggunakan Algoritma Naive Bayes .....	34
4.4 Hasil Klasifikasi Menggunakan Algoritma K-Nearest Neighbor (KNN) .....	35
4.5 Perbandingan Kinerja Naive Bayes dan KNN .....	36
4.6 Implikasi Hasil Penelitian .....	37
BAB V .....	38
KESIMPULAN DAN SARAN .....	38
5.1 Kesimpulan .....	38
5.2 Saran .....	39
DAFTAR PUSTAKA .....	40
LAMPIRAN .....	43

## DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu .....	20
--------------------------------------	----



## DAFTAR GAMBAR

Gambar 2.1 Logo Google Colab .....	19
Gambar 3.1 Alur Desain Penelitian .....	28
Gambar 4.1 Hasil Distribusi Sentimen .....	33
Gambar 4.2 Hasil WordCloud .....	34
Gambar 4.3 Hasil Klasifikasi Naive Bayes .....	34
Gambar 4.4 Hasil Confusion Matrix NB .....	35
Gambar 4.5 Hasil Klasifikasi KNN .....	35
Gambar 4.6 Hasil Confusion Matrix KNN .....	36
Gambar 4.7 Hasil Perbandingan NB dan KNN .....	36



## DAFTAR LAMPIRAN

Lampiran 1	kartu asitensi .....	43
Lampiran 2	Cek Turnitin .....	44
Lampiran 3	CV .....	45
Lampiran 4	Surat Pernyataan HAKI .....	46
Lampiran 5	BNSP .....	48
Lampiran 6	Form Revisi & Penguji .....	49

