

**ANALISIS SENTIMEN ULASAN PENGGUNA
APLIKASI FACEBOOK DI GOOGLE PLAY STORE
MENGUNAKAN ALGORITMA NAÏVE BAYES DAN
K-NEAREST NEIGHBOR**

SKRIPSI SARJANA SISTEM INFORMASI

Oleh

Tulus Ramadha Triputra

217006516024



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA**

UNIVERSITAS NASIONAL

2025

**ANALISIS SENTIMEN ULASAN PENGGUNA
APLIKASI FACEBOOK DI GOOGLE PLAY STORE
MENGUNAKAN ALGORITMA NAÏVE BAYES DAN
K-NEAREST NEIGHBOR**

SKRIPSI SARJANA SISTEM INFORMASI

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Sistem
Informasi dari Fakultas Teknologi Komunikasi dan Informatika

Oleh

Tulus Ramadha Triputra

217006516024



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA**

UNIVERSITAS NASIONAL

2025

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI FACEBOOK DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN K-NEAREST NEIGHBOR



Tulus Ramadha Triputra

217006516024

Dosen Pembimbing

Albaar Rubhasy, S.Si, MTI

NIDN. 0317108104

PERNYATAAN KEASLIAN TUGAS AKHIR

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI FACEBOOK DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN K-NEAREST NEIGHBOR

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 04 Maret 2025


METERAN
TEMPER
8AMX141695673
Ratus Kanaqha Iriputra

217006516024

HALAMAN PERNYATAAN ORISINALITAS

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwa karya tulis ini menyalahi peraturan yang ada berkaitan etika dan kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku

Yang menyatakan,

Nama : Tulus Ramadha Triputra

NIM : 217006516024

Tanda Tangan :



(Tulus Ramadha Triputra)

Tanggal : 4 Maret 2025

Mengetahui

Pembimbing : Albaar Rubhasy, S.Si, MTI



LEMBAR PERSETUJUAN REVIEW AKHIR

LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI FACEBOOK DI
GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAÏVE
BAYES DAN K-NEAREST NEIGHBOR**

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2024-2025 pada tanggal 26 Februari Tahun 2025.

Dosen Pembimbing 1



Albaar Rubhasy, S.Si, MTI
NIDN. 0317108104

Ketua Program Studi



Dr. Andrianingsih, S.Kom., MMSI
NIDN. 0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Tulus Ramadha Triputra
 NPM : 217006516024
 Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
 Program Studi : Sistem Informasi
 Tanggal Sidang : 26 Februari 2025

JUDUL DALAM BAHASA INDONESIA :

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI FACEBOOK DI
 GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN
 K-NEAREST NEIGHBOR

JUDUL DALAM BAHASA INGGRIS :

SENTIMENT ANALYSIS OF FACEBOOK APP USER REVIEWS ON GOOGLE
 PLAY STORE USING NAIVE BAYES AND K-NEAREST NEIGHBOR
 ALGORITHMS

TANDA TANGAN DAN TANGGAL

Pembimbing I	Ka. Prodi	Mahasiswa
TGL : 04 Maret 2025	TGL : 04 Maret 2025	TGL : 04 Maret 2025
		

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia sehingga penulis dapat menyelesaikan skripsi dengan judul **“Analisis Sentimen Ulasan Pengguna Aplikasi Facebook di Google Play Store Menggunakan Algoritma Naive Bayes Dan K-Nearest Neighbor”** sebagai salah satu syarat kelulusan Program Studi Sarjana Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika.

Dalam proses penyusunan skripsi ini, penulis mendapatkan banyak bantuan dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya terutama kepada dosen pembimbing Tugas Akhir, Bapak Albaar Rubhasy, S.Si, MTI atas waktu, tenaga, dan pikiran dalam membimbing dan memberikan saran kepada penulis selama penelitian dan penulisan skripsi. Penulis turut menyampaikan rasa terima kasih kepada.

1. Allah SWT yang telah memberikan kesehatan, kemudahan dan kelancaran dalam menyelesaikan penelitian serta penulisan skripsi ini dengan sebaik mungkin.
2. Kedua orang tua penulis, Bapak Madani dan Ibu Nurhayati yang telah memberikan dukungan dalam berbagai bentuk.
3. Bapak Albaar Rubhasy, S.Si, MTI, selaku Dosen Pembimbing Skripsi.
4. Seluruh dosen pengajar di Program Studi Sistem Informasi FTKI maupun dosen di Program Studi lain yang telah memberikan banyak ilmu.
5. Teman – teman seangkatan yang sudah memberikan bantuan serta dukungan.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun. Semoga skripsi ini dapat memberikan manfaat di bidang IT.

Jakarta, 04 Maret 2025



Tulus Ramadha Triputra

ABSTRAK

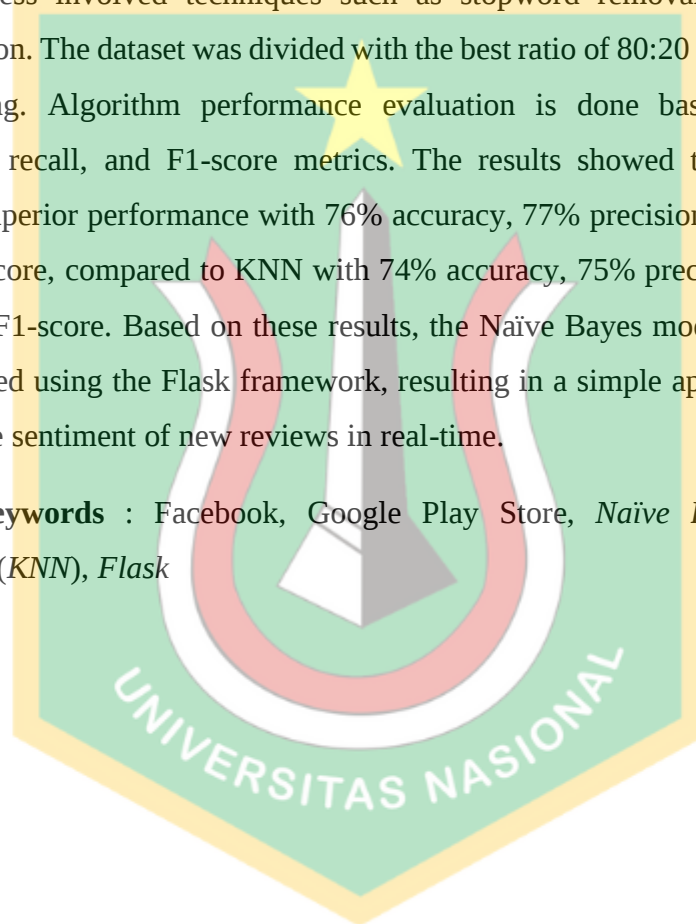
Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Facebook di Google Play Store dengan menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbor (KNN)*. Data ulasan sebanyak 5000 dikumpulkan melalui proses *web scraping* menggunakan *Python*, kemudian dilakukan *preprocessing* untuk membersihkan dan mempersiapkan data. Proses ini melibatkan teknik seperti *stopword removal*, *stemming*, dan tokenisasi. Dataset dibagi dengan rasio terbaik 80:20 untuk pelatihan dan pengujian model. Evaluasi kinerja algoritma dilakukan berdasarkan metrik *akurasi*, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa, *Naïve Bayes* menunjukkan kinerja yang lebih unggul yaitu *akurasi* 76%, *precision* 77%, *recall* 76%, dan *F1-score* 76%, dibandingkan *KNN* yaitu *akurasi* 74%, *precision* 75%, *recall* 74%, dan *F1-score* 74%. Berdasarkan hasil ini, model *Naïve Bayes* dipilih untuk di-deploy menggunakan *framework Flask*, menghasilkan aplikasi sederhana yang dapat memprediksi sentimen ulasan baru secara real-time.

Kata kunci : Facebook, Google Play Store, *Naïve Bayes*, *K-Nearest Neighbor (KNN)*, *Flask*

ABSTRACT

This research aims to analyze the sentiment of Facebook application user reviews on the Google Play Store using the Naïve Bayes and K-Nearest Neighbor (KNN) algorithms. The 5000 review data was collected through a web scraping process using Python, then preprocessing was done to clean and prepare the data. This process involved techniques such as stopword removal, stemming, and tokenization. The dataset was divided with the best ratio of 80:20 for model training and testing. Algorithm performance evaluation is done based on accuracy, precision, recall, and F1-score metrics. The results showed that Naïve Bayes showed superior performance with 76% accuracy, 77% precision, 76% recall, and 76% F1-score, compared to KNN with 74% accuracy, 75% precision, 74% recall, and 74% F1-score. Based on these results, the Naïve Bayes model was chosen to be deployed using the Flask framework, resulting in a simple application that can predict the sentiment of new reviews in real-time.

Keywords : Facebook, Google Play Store, Naïve Bayes, K-Nearest Neighbor (KNN), Flask



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN TUGAS AKHIR.....	ii
PERNYATAAN KEASLIAN TUGAS AKHIR.....	iii
HALAMAN PERNYATAAN ORISINALITAS.....	iv
LEMBAR PERSETUJUAN REVIEW AKHIR.....	v
LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI	vi
KATA PENGANTAR.....	vii
ABSTRAK.....	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian.....	3
1.4. Batasan Masalah.....	4
1.5. Kontribusi	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Data Mining	6
2.2. Text Mining.....	7
2.3. Natural Language Processing (NLP).....	7
2.4. Analisis Sentimen.....	8
2.5. Kepuasan Pengguna.....	10
2.6. Machine Learning.....	11
2.7. Web Scraping.....	12
2.8. Confusion Matrix.....	12
2.9. Python	13
2.10. Algoritma Naïve Bayes.....	14
2.11. Algoritma K-Nearest Neighbor (KNN)	15
2.12. Flask	15

2.13.	Facebook	16
2.14.	Google Play Store	17
2.15.	Google Colab	17
2.16.	Visual Studio Code	17
2.17.	Studi Literatur	18
BAB III METODOLOGI PENELITIAN		25
3.1.	Lokasi Penelitian	25
3.2.	Subjek Penelitian	25
3.3.	Desain Penelitian	25
3.4.	Pengumpulan Data	26
3.5.	<i>Pre-processing</i>	27
3.6.	Pengolahan Data	28
3.7.	Pelatihan Model (<i>Naïve Bayes</i> dan <i>KNN</i>)	29
3.8.	<i>Evaluation</i>	29
3.9.	<i>Deployment</i>	30
BAB IV HASIL DAN PEMBAHASAN		31
5.1.	Perhitungan Manual Algoritma <i>Naïve Bayes</i>	31
5.2.	Perhitungan Manual Algoritma <i>K-Nearest Neighbor</i>	36
5.3.	Implementasi Analisis Sentimen Menggunakan Machine Learning	39
5.4.	Pengumpulan Data	41
5.5.	<i>Pre-Processing</i>	42
4.5.1.	<i>Cleaning Data</i>	42
4.5.2.	<i>Case Folding</i>	43
4.5.3.	Normalisasi	44
4.5.4.	<i>Stopword Removal</i>	44
4.5.5.	<i>Tokenizing</i>	45
4.5.6.	<i>Stemming</i>	46
5.6.	Pengolahan Data	46
4.6.1.	<i>Labeling</i>	47
4.6.2.	<i>Exploratory Data Analysis (EDA)</i>	47
4.6.3.	<i>Splitting Data</i>	53
5.7.	Pemodelan dengan Algoritma <i>Naïve Bayes</i> dan <i>K-Nearest Neighbor</i>	54
5.8.	Visualisasi <i>Confusion Matrix Naïve Bayes</i> dan <i>K-Nearest Neighbor</i>	56

5.9.	Analisis Hasil Evaluasi.....	58
5.9.1.	Perbandingan 3 Skenario Pembagian Data <i>Naive Bayes</i> dan <i>KNN</i>	58
5.9.2.	Perbandingan Hasil Algoritma <i>Naive Bayes</i> dan <i>KNN</i>	61
5.10.	Deployment	62
BAB V KESIMPULAN DAN SARAN		66
5.1.	Kesimpulan	66
5.2.	Saran	66
DAFTAR PUSTAKA		68
LAMPIRAN		71



DAFTAR GAMBAR

Gambar 3.1 Desain Penelitian	26
Gambar 4.1 Bar chart 10 kata teratas sebelum pre-processing	42
Gambar 4.2 WordCloud pada data ulasan.....	48
Gambar 4.3 WordCloud pada data ulasan berlabel positif	48
Gambar 4.4 WordCloud pada data ulasan berlabel negatif.....	49
Gambar 4.5 Bar chart 10 kata teratas setelah pre-processing.....	50
Gambar 4.6 Bar chart 10 kata teratas pada data berlabel positif	50
Gambar 4.7 Bar chart 10 kata teratas pada data berlabel negatif	51
Gambar 4.8 Bar chart jumlah rating untuk setiap rating.....	52
Gambar 4.9 Bar chart jumlah distribusi sentimen	52
Gambar 4.10 Jumlah Data Latih & Data Uji.....	53
Gambar 4.11 Hasil Classification Report Naive Bayes	54
Gambar 4.12 Hasil Classification Report KNN	55
Gambar 4.13 Confusion Matrix Naive Bayes	56
Gambar 4.14 Confusion Matrix K-Nearest Neighbor	57
Gambar 4.15 Confusion Matrix Naive Bayes 60:40, 70:30, 80:20	58
Gambar 4.16 Confusion Matrix K-Nearest Neighbor 60:40, 70:30, 80:20	59
Gambar 4.17 Classification Report Naive Bayes	61
Gambar 4.18 Classification Report K-Nearest Neighbor	61
Gambar 4.19 Menu awal Cek Sentimen	63
Gambar 4.20 Hasil dari pengecekan data baru sentimen positif	63
Gambar 4.21 Hasil dari pengecekan data baru sentimen negatif	64
Gambar 4.22 Hasil salah mengklasifikasikan data baru	65
Gambar 4.23 Menu Data Ulasan	65

DAFTAR TABEL

Tabel 2.1 Tabel Studi Literatur	19
Tabel 4.1 Sampel Data Ulasan	31
Tabel 4.2 Ulasan Positif	32
Tabel 4.3 Ulasan Negatif.....	32
Tabel 4.4 Frekuensi kata dalam ulasan positif	33
Tabel 4.5 Frekuensi kata dalam ulasan negatif	33
Tabel 4.6 Sampel Data Ulasan	36
Tabel 4.7 Representasi Numerik pada Data Positif	37
Tabel 4.8 Representasi Numerik pada Data Negatif.....	37
Tabel 4.9 Hasil Cleaning Data.....	43
Tabel 4.10 Hasil Case Folding	43
Tabel 4.11 Hasil Normalisasi	44
Tabel 4.12 Hasil Stopword Removal	45
Tabel 4.13 Hasil Tokenizing	45
Tabel 4.14 Hasil Stemming.....	46
Tabel 4.15 Menghapus baris yang memiliki nilai yang hilang	47
Tabel 4.16 Hasil Labeling	47
Tabel 4.17 Perbandingan Classification report Naive Bayes	58
Tabel 4.18 Perbandingan Classification report K-Nearest Neighbor	59