

**ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES,
SUPPORT VECTOR MACHINE, DAN RANDOM FOREST
UNTUK SENTIMEN ULASAN APLIKASI 'DANA' DI PLAY
STORE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

YANU MARIO

21.11.3853

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES,
SUPPORT VECTOR MACHINE, DAN RANDOM FOREST
UNTUK SENTIMEN ULASAN APLIKASI 'DANA' DI PLAY
STORE**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

YANU MARIO

21.11.3853

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

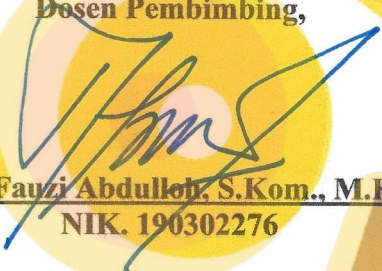
ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES, SUPPORT VECTOR MACHINE, DAN RANDOM FOREST

yang disusun dan diajukan oleh

Yanu Mario
21.11.3853

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 08 Juli 2025

Dosen Pembimbing,


Ferian Fauzi Abdulloh, S.Kom., M.Kom.
NIK. 190302276

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES, SUPPORT VECTOR MACHINE, DAN RANDOM FOREST UNTUK SENTIMEN ULASAN APLIKASI 'DANA' DI PLAY STORE

yang disusun dan diajukan oleh

Yanu Mario
21.11.3853

Telah dipertahankan di depan Dewan Penguji
pada tanggal 28 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Ali Mustopa, S.Kom., M.Kom.
NIK. 190302192

Dwi Nurani, S.Kom., M.Kom.
NIK. 190302236

Ferian Fauzi Abdulloh, S.Kom., M.Kom.
NIK. 190302276

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 28 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Yanu Mario
NIM : 21.11.3853

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES, SUPPORT VECTOR MACHINE, DAN RANDOM FOREST UNTUK SENTIMEN ULASAN APLIKASI 'DANA' DI PLAY STORE

Dosen Pembimbing : Ferian Fauzi Abdulloh, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 28 Juli 2025

Yang Menyatakan,



Yanu Mario

HALAMAN PERSEMBAHAN

Dengan segala kerendahan hati, skripsi ini aku persembahkan kepada:

1. Orang tua dan keluarga tercinta yang selalu memberikan dukungan dan doa tiada henti.
2. Dosen pembimbing bapak Ferian Fauzi Abdulloh yang dengan sabar membimbing dan memberikan ilmu berharga.
3. Teman-teman terutama Fathan yang sering membantu saya dalam bertukar pikiran dan memperbaiki kode, Lian yang sering mengingatkan *progress* skripsi dan semua teman saya yang selalu memberi semangat dan membantu selama masa penyusunan skripsi ini.
4. Semua pihak yang telah membantu dan memberikan inspirasi dalam perjalanan akademik ini.

Semoga karya sederhana ini bermanfaat dan menjadi langkah awal untuk mencapai tujuan yang lebih besar.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat, karunia, dan kesempatan sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Sentimen Ulasan Pengguna Aplikasi DANA Menggunakan Algoritma Naive Bayes, Support Vector Machine, dan Random Forest". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi.

Penulis menyadari bahwa tanpa bantuan dan dukungan dari berbagai pihak, penulisan skripsi ini tidak akan dapat terselesaikan dengan baik. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak dosen pembimbing Ferian Fauzi Abdulloh, S.Kom., M.Kom., yang telah memberikan bimbingan, arahan, dan motivasi selama proses penyusunan skripsi ini. Dosen penguji bapak Ali Mustopa, S.Kom., M.Kom., dan ibu Dwi Nurani, S.Kom., M.Kom., yang telah menguji penelitian saya
2. Keluarga tercinta yang selalu memberikan dukungan moral dan doa.
3. Teman-teman serta semua pihak yang telah membantu dengan berbagai cara baik secara langsung maupun tidak langsung.

Penulis berharap skripsi ini dapat memberikan kontribusi positif bagi pengembangan ilmu pengetahuan terutama dalam bidang analisis sentimen dan aplikasi machine learning di Indonesia. Penulis juga menyadari bahwa skripsi ini masih jauh dari sempurna, oleh karena itu kritik dan saran yang membangun sangat penulis harapkan demi perbaikan ke depan.

Akhir kata, semoga skripsi ini bermanfaat bagi pembaca dan semua pihak yang terkait.

Yogyakarta, 28 Juli 2025



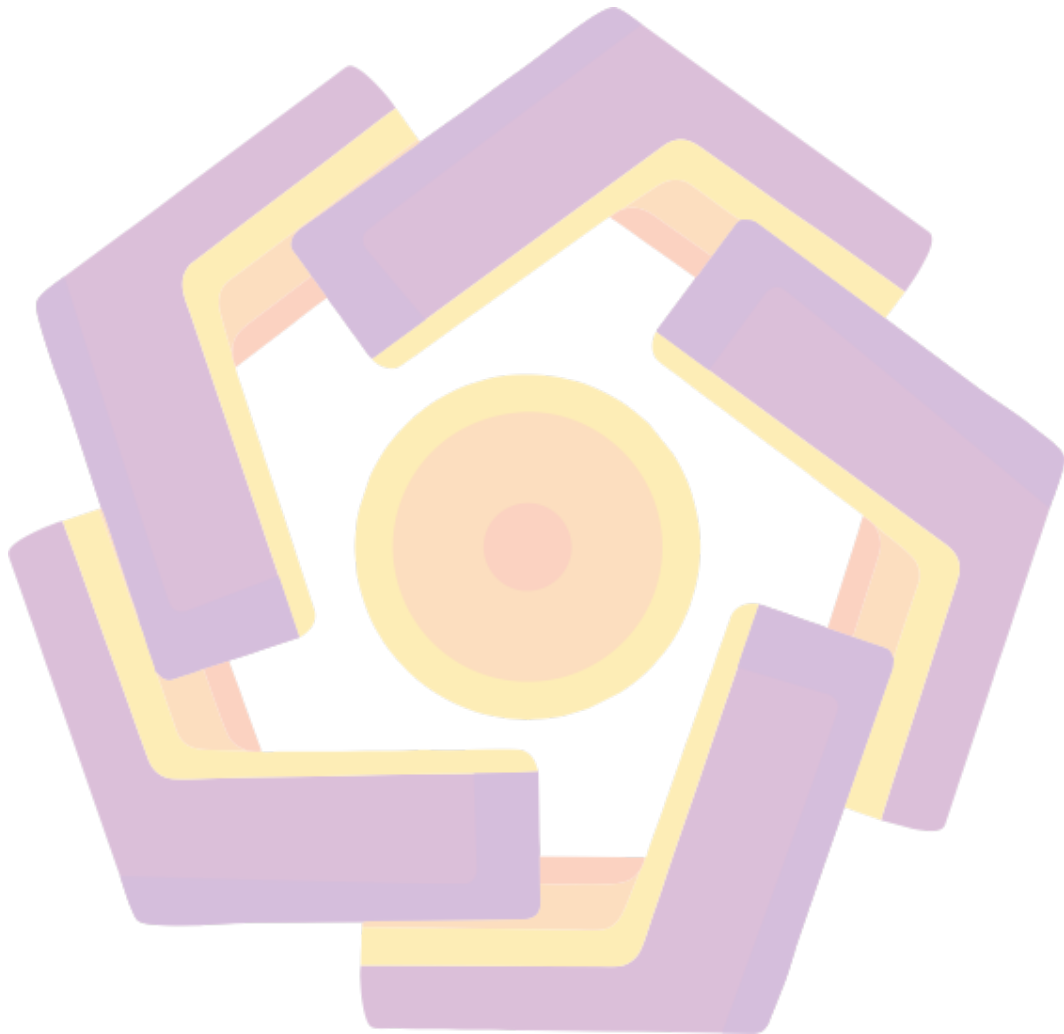
Yanu Mario

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
INTISARI	xiii
<i>ABSTRACT</i>	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	13
2.2.1 Analisis Sentimen	13
2.2.2 <i>Machine Learning</i>	13

2.2.3 Naïve Bayes	13
2.2.4 Support Vector Machine	14
2.2.5 Random Forest	14
2.2.6 <i>Preprocessing</i>	14
2.2.7 TF-IDF	16
2.2.8 SMOTE	16
2.2.9 <i>Confusion Matrix</i>	16
BAB III METODE PENELITIAN	18
3.1 Objek Penelitian	18
3.2 Alur Penelitian.....	18
3.3 Alat dan Bahan	21
BAB IV HASIL DAN PEMBAHASAN	23
4.1 Hasil Penelitian	23
4.1.1 Pengumpulan Data	23
4.1.2 <i>Preprocessing</i>	23
4.1.3 <i>Data Labeling</i>	26
4.1.4 Visualisasi TF IDF	26
4.1.5 Pembagian Data	27
4.1.6 SMOTE	27
4.1.7 Implementasi Model	29
4.1.8 Evaluasi Kinerja Algoritma	29
4.1.9 Perbandingan Algoritma	37
BAB V PENUTUP	38
5.1 Kesimpulan.....	38
5.2 Saran.....	39

REFERENSI	40
LAMPIRAN.....	44

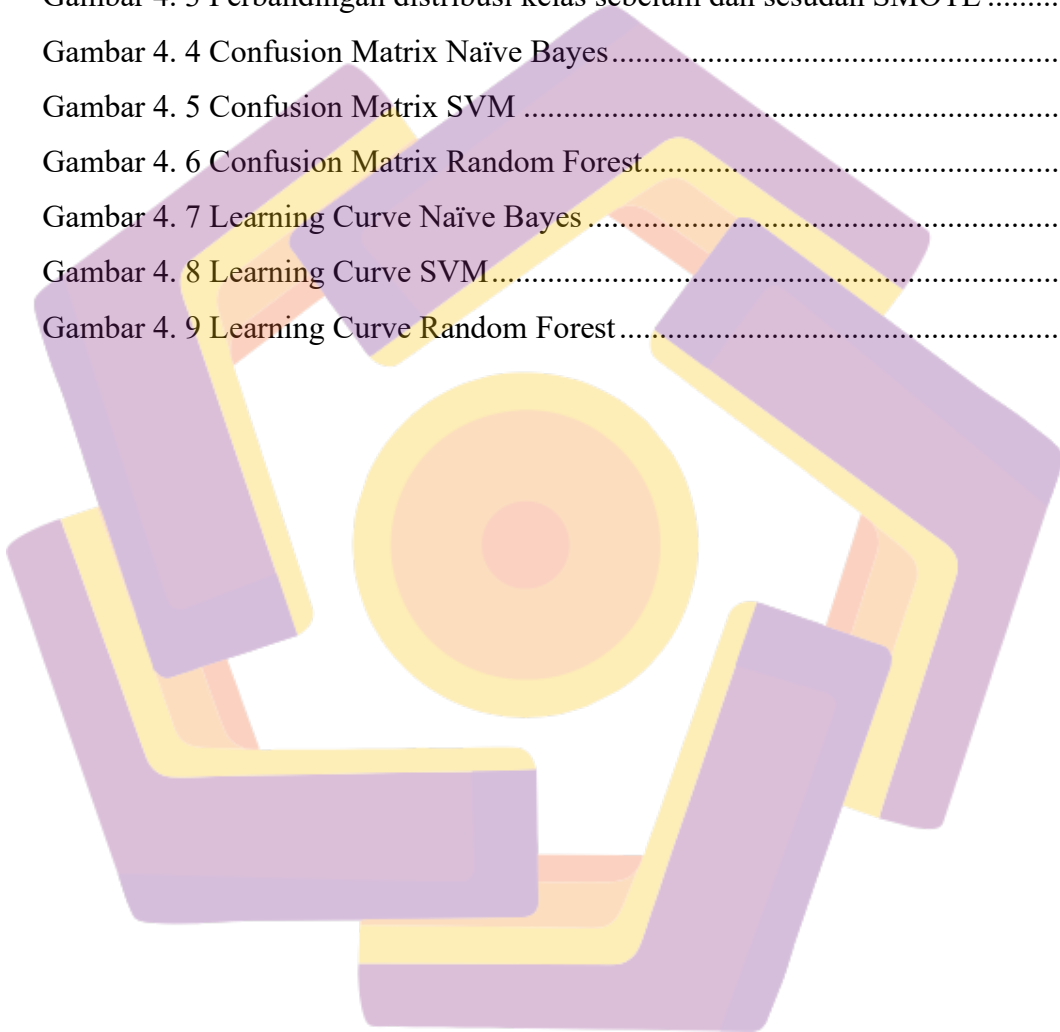


DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	10
Tabel 4. 1 Data Cleansing dan casefolding.....	23
Tabel 4. 2 Stopword Removal	24
Tabel 4. 3 Tokenize.....	25
Tabel 4. 4 Stemming	25
Tabel 4. 5 Data Labeling.....	26
Tabel 4. 6 Performa sebelum SMOTE.....	28
Tabel 4. 7 Performa setelah SMOTE	28
Tabel 4. 8 Classification Report Naive Bayes	33
Tabel 4. 9 Classification Report SVM.....	34
Tabel 4. 10 Classification Report Random Forest	34
Tabel 4. 11 Perbandingan Algoritma	37

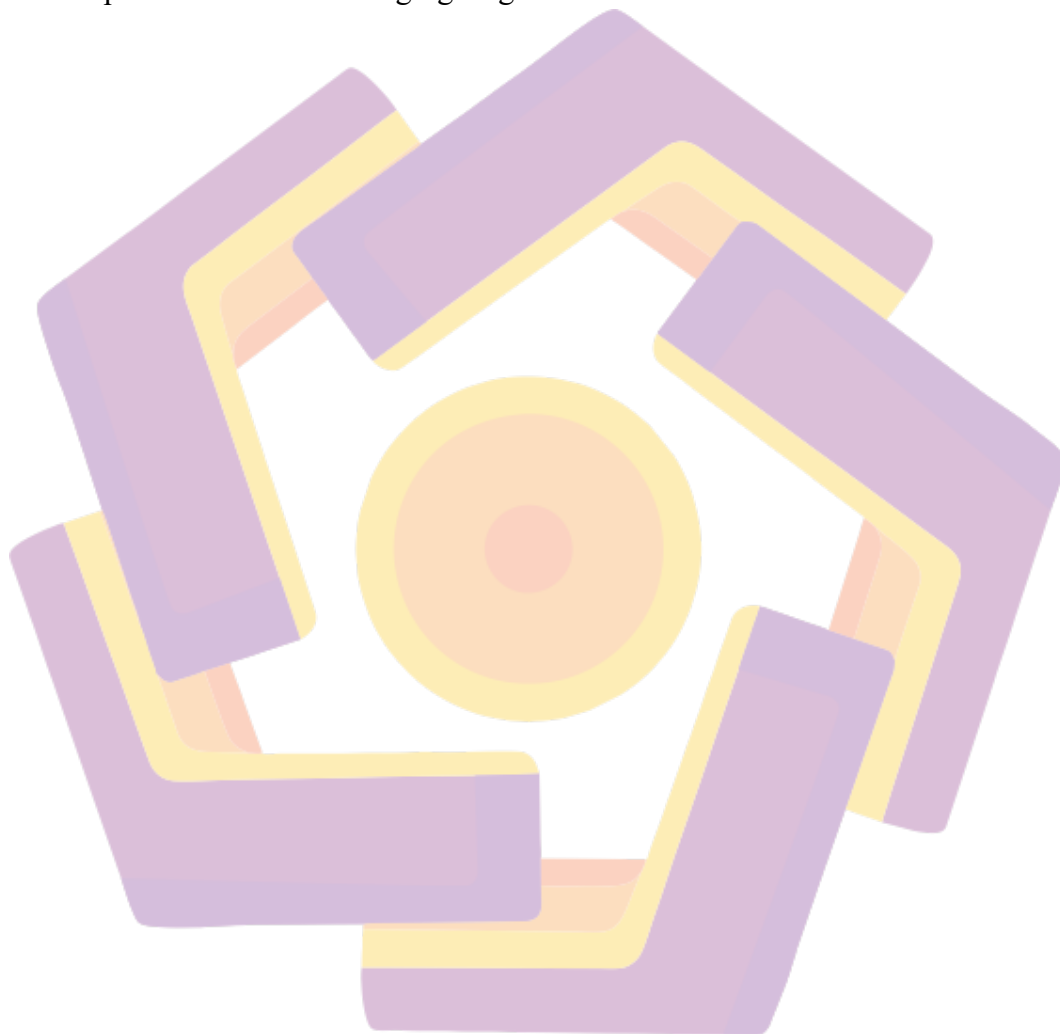
DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	19
Gambar 4. 1 Pengumpulan Data	23
Gambar 4. 2 Visualisasi TF IDF Ulasan DANA	27
Gambar 4. 3 Perbandingan distribusi kelas sebelum dan sesudah SMOTE	27
Gambar 4. 4 Confusion Matrix Naïve Bayes	30
Gambar 4. 5 Confusion Matrix SVM	31
Gambar 4. 6 Confusion Matrix Random Forest	32
Gambar 4. 7 Learning Curve Naïve Bayes	35
Gambar 4. 8 Learning Curve SVM	35
Gambar 4. 9 Learning Curve Random Forest	36



DAFTAR LAMPIRAN

Lampiran 1. Kode Pengumpulan Data ulasan DANA	44
Lampiran 2. Kode Modeling tiga algoritma.....	48



INTISARI

Pesatnya perkembangan teknologi digital telah membawa perubahan signifikan dalam kehidupan masyarakat, salah satunya melalui kehadiran aplikasi dompet digital seperti DANA. Sebagai salah satu aplikasi finansial yang populer di Indonesia, DANA digunakan untuk berbagai transaksi seperti pembayaran, transfer, dan pembelian online. Ulasan pengguna di platform Google Play Store menjadi sumber informasi yang penting untuk menilai kepuasan atau ketidakpuasan pengguna terhadap layanan yang diberikan. Namun, banyaknya data dalam bentuk teks yang bersifat tidak terstruktur menyulitkan proses analisis manual oleh pengembang. Oleh karena itu, diperlukan metode analisis sentimen berbasis algoritma *Machine Learning* yang mampu mengklasifikasikan ulasan pengguna secara otomatis dan akurat.

Penelitian ini bertujuan untuk membandingkan kinerja tiga algoritma *Machine Learning*, yaitu Naive Bayes, Support Vector Machine (SVM), dan Random Forest dalam melakukan klasifikasi sentimen terhadap ulasan aplikasi DANA. Data yang digunakan dalam penelitian ini diperoleh melalui proses web scraping terhadap 2.134 ulasan pengguna dari Google Play Store. Tahapan yang dilakukan meliputi *preprocessing* data (*casefolding*, *cleansing*, *tokenizing*, *Stopword Removal*, *stemming*), pelabelan sentimen (positif dan negatif), ekstraksi fitur menggunakan metode TF-IDF, serta penyeimbangan data menggunakan teknik SMOTE. Dataset kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20 sebelum dilakukan pelatihan model. Kinerja model dievaluasi menggunakan *Confusion Matrix* dan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-Score*.

Hasil dari penelitian menunjukkan bahwa algoritma Naive Bayes memberikan performa terbaik dengan nilai akurasi sebesar 88%. Sementara itu, Random Forest menempati posisi kedua dengan akurasi 86%, dan SVM di posisi terakhir dengan akurasi 84%. Penelitian ini memberikan kontribusi praktis bagi pengembang aplikasi dalam memahami lebih dalam pengalaman dan persepsi pengguna, sehingga dapat digunakan sebagai dasar dalam pengembangan fitur dan peningkatan layanan. Selain itu, hasil penelitian ini juga memberikan referensi bagi peneliti dan akademisi dalam menentukan algoritma yang paling efektif untuk tugas analisis sentimen, khususnya dalam konteks ulasan aplikasi berbasis bahasa Indonesia.

Kata kunci: analisis sentimen, Naive Bayes, Support Vector Machine, Random Forest, aplikasi DANA

ABSTRACT

The rapid advancement of digital technology has significantly transformed society, particularly through the rise of digital wallet applications such as DANA. As one of Indonesia's leading financial apps, DANA facilitates various transactions, including payments, transfers, and online purchases. User reviews on platforms like Google Play Store offer valuable insights into customer satisfaction or dissatisfaction. However, the unstructured and voluminous nature of these textual data makes manual analysis inefficient and impractical. Therefore, implementing sentiment analysis using Machine Learning algorithms is essential for automatically and accurately classifying user reviews.

This study aims to compare the performance of three Machine Learning algorithms—Naive Bayes, Support Vector Machine (SVM), and Random Forest—in classifying the sentiment of user reviews for the DANA application. The dataset consists of 2,134 user reviews collected through web scraping from Google Play Store. The analysis process includes text preprocessing (casefolding, cleansing, tokenizing, Stopword Removal, stemming), sentiment labeling (positive and negative), feature extraction using TF-IDF, and data balancing using the SMOTE technique. The dataset is split into 80% training and 20% testing sets. Model performance is evaluated using a Confusion Matrix and metrics such as accuracy, precision, recall, and F1-Score.

The results indicate that the Naïve Bayes algorithm delivers the best performance, achieving 88% across all evaluation metrics. Random Forest ranks second with an accuracy of 86%, followed by SVM with 84% accuracy. This research provides practical insights for app developers to better understand user experience and sentiment, allowing for more informed decisions in feature development and service improvements. Furthermore, the findings serve as a reference for researchers and academics in selecting effective algorithms for sentiment analysis tasks, especially in the context of Indonesian-language user reviews.

Keyword: *sentiment analysis, Naive Bayes, Support Vector Machine, Random Forest, DANA application.*