

**ANALISIS KLASIFIKASI DATA ULASAN PENGGUNA
MENGUNAKAN RANDOM FOREST DAN
SUPPORT VECTOR MACHINE**

(Studi Kasus : Aplikasi JAKI – Jakarta Kini)

SKRIPSI SARJANA SISTEM INFORMASI

Oleh

Sekar Kesya Meidy Santoso

217006516050



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

2024

**ANALISIS KLASIFIKASI DATA ULASAN PENGGUNA
MENGUNAKAN RANDOM FOREST DAN
SUPPORT VECTOR MACHINE**

(Studi Kasus : Aplikasi JAKI – Jakarta Kini)

SKRIPSI SARJANA SISTEM INFORMASI

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

Oleh

Sekar Kesya Meidy Santoso

217006516050



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

2024

HALAMAN PENGESAHAN

TUGAS AKHIR

**ANALISIS KLASIFIKASI DATA ULASAN PENGGUNA MENGGUNAKAN
RANDOM FOREST DAN SUPPORT VECTOR MACHINE**

(Studi Kasus : Aplikasi JAKI – Jakarta Kini)



Dosen Pembimbing 1

(Gatot Soepriyono, S.Si., M.S.M.)

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

ANALISIS KLASIFIKASI DATA ULASAN PENGGUNA MENGGUNAKAN RANDOM FOREST DAN SUPPORT VECTOR MACHINE

(Studi Kasus : Aplikasi JAKI – Jakarta Kini)

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, 28 Februari 2025



Sekar Kesya Meidy Santoso

NPM. 217006516050

LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

ANALISIS KLASIFIKASI DATA ULASAN PENGGUNA MENGUNAKAN RANDOM FOREST DAN SUPPORT VECTOR MACHINE

(Studi Kasus : Aplikasi JAKI – Jakarta Kini)

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



Gatot Soepriono, S.Si., M.S.M.

NIDN. 0317047803

Ketua Program Studi



Dr. Andrianingsih, S.Kom., MMSI.

NIDN. 0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Sekar Kesya Meidy Santoso
NPM : 217006516050
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Sistem Informasi
Tanggal Sidang : 26 Februari 2025

JUDUL DALAM BAHASA INDONESIA :

**ANALISIS KLASIFIKASI DATA ULASAN PENGGUNA
MENGUNAKAN RANDOM FOREST DAN
SUPPORT VECTOR MACHINE**

(Studi Kasus : Aplikasi JAKI – Jakarta Kini)

JUDUL DALAM BAHASA INGGRIS :

**ANALYSIS OF USER REVIEW CLASSIFICATION
USING RANDOM FOREST AND
SUPPORT VECTOR MACHINE**

(Case Study: JAKI – Jakarta Kini Application)

TANDA TANGAN DAN TANGGAL

Pembimbing 1	Ka. Prodi	Mahasiswa
TGL : 4/3/2025	TGL : 28 Februari 2025	TGL : 1 Maret 2025
	 Dr. Andrianingsih, MM, SI	

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 : Sekar Kesya Meidy Santoso

NIM : 217006516050

Tanda Tangan :



Tanggal : 4 Maret 2025

Mengetahui

Pembimbing 1 : Gatot Soepriyono, S.Si., M.S.M. ()

HALAMAN PENGESAHAN
TUGAS SARJANA

ANALISIS KLASIFIKASI DATA ULASAN PENGGUNA
MENGGUNAKAN RANDOM FOREST DAN

SUPPORT VECTOR MACHINE

(Studi Kasus : Aplikasi JAKI – Jakarta Kini)

Oleh

Sekar Kesya Meidy Santoso

217006516050

Program Studi Sistem Informasi

Fakultas Teknologi Komunikasi dan Informatika

Disetujui pada Tanggal: 4 Maret 2025

Pembimbing I



(Gatot Soepriyono, S.Si., M.S.M.)

KATA PENGANTAR

Puji syukur penulis sampaikan kepada Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya, yang telah memungkinkan penulis menyelesaikan skripsi berjudul "**Analisis Klasifikasi Data Ulasan Pengguna Menggunakan Random Forest Dan Support Vector Machine** (Studi Kasus : Aplikasi JAKI – Jakarta Kini)" sebagai persyaratan kelulusan Program Studi Sarjana Sistem Informasi di Fakultas Teknologi Komunikasi dan Informatika.

Penyelesaian penelitian dan penyusunan skripsi ini tidak lepas dari dukungan berbagai pihak, oleh karena itu, penulis ingin menyampaikan penghargaan kepada dosen pembimbing Tugas Akhir atas dedikasi, bimbingan, motivasi, dan pengertian yang diberikan selama proses penelitian dan penulisan skripsi. Penulis juga mengucapkan terima kasih kepada:

1. Teruntuk kedua orangtua penulis tersayang dan teristimewa, support system terbaik serta panutan bagi penulis. Ayahanda Cinta pertama penulis dan Ibunda Pintu surga bagi penulis. Ucapan terima kasih yang sebesar-besarnya atas segala pengorbanan dan tulus kasih. Senantiasa memberikan yang terbaik, tidak kenal lelah untuk selalu mendoakan dan memberi perhatian serta dukungan hingga penulis mampu menyelesaikan studi sampai sarjana. Penulis berharap, semoga ayah dan bunda diberikan kesehatan, panjang umur dan bahagia selalu.
2. Kepada saudara penulis yang tidak kalah penting kehadirannya, Adik tercinta Wishnu Ghivari Santoso. Terima kasih telah menjadi bagian dari perjalanan hidup penulis. Yang telah mendukung, menghibur dan mendengarkan keluh kesah penulis dengan penuh kesabaran.
3. Bapak Gatot Soepriyono, S.Si., M.S.M selaku Dosen Pembimbing 1 yang telah menyediakan bimbingan, memberikan masukan berharga, motivasi, dan arahan kepada penulis.
4. Bapak Dr. Agung Triayudi, S.Kom., M.Kom. selaku Dekan Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional.

5. Ibu Dr. Andrianingsih, S.Kom., MMSI selaku Ketua Program Studi Sistem Informasi.
6. Seluruh Dosen di Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, atas ilmu dan bimbingan yang telah diberikan selama masa perkuliahan.
7. Semua Dosen di Program Studi Sistem Informasi Universitas Nasional yang telah memberikan banyak ilmu bermanfaat dan memberikan kesempatan untuk terus berkembang selama masa studi.
8. Untuk teman di masa sekolah menengah atas penulis, Dewi, Dina, Fitri dan Elsa yang masih bertahan hingga detik ini. Penulis ucapkan terima kasih karena telah memberikan semangat, dukungan serta doa terbaiknya.
9. Ucapan terima kasih juga penulis sampaikan kepada teman-teman seperjuangan, Abija, Alawi, Aqsal, Arga, Camel, Dhila, Gandhi, Ronald dan Tegar atas dukungan, diskusi yang bermanfaat, serta kerja sama yang telah diberikan selama proses penelitian berlangsung.
10. Kepada sosok yang belum diketahui namanya namun sudah tertulis jelas di *lauhul mahfuz* penulis. Terima kasih sudah menjadi sumber motivasi bagi penulis untuk menyelesaikan skripsi ini sebagai upaya untuk memantaskan diri. Semoga kelak berjumpa pada versi terbaik diri masing-masing.

Penulis mengakhiri kata pengantar ini dengan harapan agar Tuhan Yang Maha Esa membalas kebaikan dan bantuan yang telah diberikan dengan keberkahan. Penulis juga terbuka untuk menerima kritik dan saran yang membangun, sembari berharap skripsi ini dapat memberikan kontribusi positif di bidang Sistem Informasi.

Jakarta, 4 Maret 2025



Sekar Kesya Meidy Santoso

HALAMAN PERNYATAAN PERSETUJUAN

Sebagai sivitas akademik Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika, saya yang bertanda tangan di bawah ini:

Nama : Sekar Kesya Meidy Santoso

NIM : 217006516050

demikian demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Fakultas Teknologi Komunikasi dan Informatika, Hak Bebas Royalti Non Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

“Analisis Klasifikasi Data Ulasan Pengguna Menggunakan Random Forest Dan Support Vector Machine (Studi Kasus : Aplikasi JAKI – Jakarta Kini)”

Beserta perangkat yang ada, dengan hak ini Fakultas Teknologi Komunikasi dan Informatika berhak menyimpan, mengalih media/formatkan, 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.

Dibuat di : Pasar Minggu – Jakarta Selatan

Pada tanggal : 4 Maret 2025

Yang menyatakan



(Sekar Kesya Meidy Santoso)

ABSTRAK

Dengan meningkatnya penggunaan aplikasi digital, jumlah ulasan pengguna juga semakin bertambah. Aplikasi JAKI (Jakarta Kini) berfungsi sebagai platform yang menghubungkan warga dengan berbagai layanan kota untuk meningkatkan kenyamanan dan efisiensi. Analisis sentimen ulasan pengguna menjadi penting untuk meningkatkan kualitas fitur dan pelayanan aplikasi.

Penelitian ini mengevaluasi kinerja algoritma Random Forest dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen ulasan pengguna aplikasi JAKI menjadi kategori positif dan negatif. Data yang digunakan diperoleh melalui web scraping dari Google Play Store, yang mencakup 4506 ulasan dengan variabel seperti ID Ulasan, Nama Pengguna, Tanggal Ulasan, Rating, Isi Ulasan, dan Polarity.

Hasil evaluasi menunjukkan bahwa metode Random Forest memiliki akurasi 91.7%, presisi 95.4%, recall 91.4%, dan F1-score 93.4%. Sementara itu, Support Vector Machine (SVM) mencapai akurasi 92.6%, presisi 95.1%, recall 93.2%, dan F1-score 94.1%. Dengan hasil tersebut, SVM menunjukkan performa yang lebih unggul dalam klasifikasi sentimen ulasan pengguna aplikasi JAKI.

Kata Kunci: Analisis Sentimen, Klasifikasi, Machine Learning, Random Forest, Support Vector Machine, JAKI.

ABSTRACT

With the increasing use of digital applications, the number of user reviews is also growing. The JAKI (Jakarta Kini) application serves as a platform that connects residents with various city services to enhance convenience and efficiency. Analyzing user sentiment is essential for improving the app's features and services.

This study evaluates the performance of the Random Forest and Support Vector Machine (SVM) algorithms in classifying user reviews of the JAKI app into positive and negative sentiment categories. The data used was obtained through web scraping from the Google Play Store, consisting of 4506 reviews with variables such as Review ID, Username, Review Date, Rating, Review Content, and Polarity.

Evaluation results show that the Random Forest method achieved an accuracy of 91.7%, precision of 95.4%, recall of 91.4%, and an F1-score of 93.4%. Meanwhile, the Support Vector Machine (SVM) method achieved the highest performance with an accuracy of 92.6%, precision of 95.1%, recall of 93.2%, and an F1-score of 94.1%. These results indicate that SVM outperforms Random Forest in classifying user sentiment in the JAKI application.

Keyword: Sentiment Analysis, Classification, Machine Learning, Random Forest, Support Vector Machine, JAKI.

DAFTAR ISI

HALAMAN PENGESAHAN TUGAS AKHIR	iii
PERNYATAAN KEASLIAN TUGAS AKHIR	iv
LEMBAR PERSETUJUAN REVIEW AKHIR.....	v
LEMBAR PERSETUJUAN JUDUL.....	vi
HALAMAN PERNYATAAN ORISINALITAS	vii
HALAMAN PENGESAHAN.....	viii
KATA PENGANTAR.....	ix
HALAMAN PERNYATAAN PERSETUJUAN.....	xi
ABSTRAK	xii
ABSTRACT	xiii
DAFTAR ISI.....	xiv
DAFTAR TABEL	xvi
DAFTAR GAMBAR.....	xvii
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Identifikasi Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.6 Batasan Masalah	5
1.7 Kontribusi.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Landasan Teori.....	7
2.1.1 Analisis Sentimen	7
2.1.2 Machine Learning	7
2.1.3 Python.....	8
2.1.4 Text Preprocessing	8
2.1.5 TF-IDF.....	9
2.1.6 Klasifikasi.....	9
2.1.7 Algoritma Random Forest.....	10

2.1.8	Algoritma Support Vector Machine	12
2.1.9	Confusion Matrix	12
2.1.10	Streamlit	14
2.2	Studi Literatur	15
BAB III METODE PENELITIAN		21
3.1	Waktu Penelitian	21
3.2	Objek Penelitian	21
3.3	Sumber Data	21
3.4	Tahapan Penelitian	22
BAB IV HASIL DAN PEMBAHASAN		29
4.1	Hasil Data Scrapping	29
4.2	Hasil Preprocessing Data	29
4.2.1	Hasil Text Preprocessing	29
4.2.2	Hasil Pelabelan Data	33
4.2.3	Hasil TF-IDF	34
4.2.4	Hasil Feature Selection	35
4.2.5	Hasil Pembagian Data	35
4.3	Hasil Modeling Algoritma	36
4.3.1	Perhitungan Manual Algoritma	36
4.3.2	Hasil Implementasi Modeling Algoritma	47
4.4	Implementasi Algoritma ke dalam Sistem	52
BAB V KESIMPULAN DAN SARAN		62
5.1	Kesimpulan	62
5.2	Saran	62
LAMPIRAN		66

DAFTAR TABEL

Table 2. 2 Studi Literatur	15
Table 3. 1 Waktu Penelitian	21
Table 4. 1 Pembagian Data Latih dan Data Uji	36
Table 4. 2 Sample Data Latih Perhitungan Manual	36
Table 4. 3 Sample Data Uji Perhitungan Manual	36
Table 4. 4 Hasil Prediksi Peritugan Random Forest	42
Table 4. 5 Data Latih Perhitungan Manual SVM	43
Table 4. 6 Data Uji Perhitungan Manual SVM.....	47



DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi Confusion Matrix	13
Gambar 3. 1 Flowchart Tahapan Penelitian.....	22
Gambar 3. 2 Flowchart Sistem.....	24
Gambar 4. 1 Dataset Ulasan Pengguna JAKI	29
Gambar 4. 2 Hasil Casefolding	30
Gambar 4. 3 Hasil Slangwords Standardization	30
Gambar 4. 4 Hasil Lemmatisasi	31
Gambar 4. 5 Hasil Stemming	32
Gambar 4. 6 Hasil Stopwords Removal.....	32
Gambar 4. 7 Hasil Unwantedword Removal	33
Gambar 4. 8 Hasil Pelabelan Data	33
Gambar 4. 9 Hasil TF-IDF	34
Gambar 4. 10 Hasil TF-IDF Filter	34
Gambar 4. 11 Hasil Feature Selection.....	35
Gambar 4. 12 Decission tree 1	38
Gambar 4. 13 Decission Tree 2.....	40
Gambar 4. 14 Decission Tree 3.....	41
Gambar 4. 15 Confusion Matrix Random Forest.....	48
Gambar 4. 16 Classification Report Random Forest.....	49
Gambar 4. 17 Hasil Prediksi Random Forest.....	49
Gambar 4. 18 Confusion Matrix Support Vector Machine.....	50
Gambar 4. 19 Classification Report SVM	50
Gambar 4. 20 Hasil Prediksi SVM.....	51
Gambar 4. 21 Perbandingan Evaluasi Model.....	51
Gambar 4. 22 Tampilan Halaman Dashbord.....	53
Gambar 4. 23 Tampilan Halaman Dataset	54
Gambar 4. 24 Tampilan Halaman Klasifikasi Random Forest	55
Gambar 4. 25 Tampilan Halaman Klasifikasi SVM	55

Gambar 4. 26 Tampilan Halaman Model Klasifikasi Random Forest.....	56
Gambar 4. 27 Tampilan Halaman Model Klasifikasi SVM.....	57
Gambar 4. 28 Tampilan Halaman Implementasi Algotirma – Confusion Matrix	58
Gambar 4. 29 Tampilan Halaman Implementasi Algotirma – Random Forest	59
Gambar 4. 30 Tampilan Halaman Implementasi Algotirma – SVM	60
Gambar 4. 31 Tampilan Halaman Implementasi Algotirma – Perbandingan Algoritma	61

