

**ANALISIS SENTIMEN PADA REVIEW PENGGUNA
APLIKASI MOBILE JKN DI GOOGLE PLAY STORE
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

SKRIPSI

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



disusun oleh

PRADIPTA ALIF PARANTOPO

20.11.3703

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN PADA REVIEW PENGGUNA
APLIKASI MOBILE JKN DI GOOGLE PLAY STORE
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

PRADIPTA ALIF PARANTOPO

20.11.3703

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PADA REVIEW PENGGUNA
APLIKASI MOBILE JKN DI GOOGLE PLAY STORE
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Pradipta Alif Parantopo

20.11.3703

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 24 Januari 2025

Dosen Pembimbing,



Windha Mega Pradnya Dhuhiya, S.Kom., M.Kom.
NIK. 190302185

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS SENTIMEN PADA REVIEW PENGGUNA
APLIKASI MOBILE JKN DI GOOGLE PLAY
MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

yang disusun dan diajukan oleh

Pradipta Alif Parantopo

20.11.3703

Telah dipertahankan di depan Dewan Penguji
pada tanggal 24 Januari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Haryoko, S.Kom., M.Cs.
NIK. 190302286

Tri Susanto, M.Kom.
NIK. 190302115

Windha Mega Pradnya Dhubita, S.Kom., M.Kom.
NIK. 190302185



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 24 Januari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Pradipta Alif Parantopo
NIM : 20.11.3703

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Pada Review Pengguna Aplikasi Mobile JKN Di Google Play Store Menggunakan Metode Support Vector Machine

Dosen Pembimbing : Windha Mega Pradnya Dhuhiita, 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, 24 Januari 2025

Yang Menyatakan,

Pradipta Alif Parantopo

HALAMAN PERSEMBAHAN

Pertama-tama, saya panjatkan puji syukur kepada Allah SWT yang telah memberikan rahmat, taufik, dan hidayah-Nya sehingga saya dapat menyelesaikan skripsi ini . Dengan segenap ketulusan dan kerendahan hati, penulis dapat sampai pada tahap ini yang merupakan sebuah pencapaian yang tentunya tidak diraih dengan mudah. Berbagai tantangan telah dilalui, namun berkat niat yang tulus, bimbingan, dan doa-doa yang teriring, serta dukungan dari orang-orang tercinta, skripsi ini dapat terselesaikan dengan baik. Skripsi ini penulis persembahkan dengan penuh cinta dan rasa hormat kepada:

1. Ibu Saya yang telah berkorban banyak tenaga, pikiran, dan materi tanpa lelah dan mengeluh sedikitpun. Terima kasih atas doa yang tak pernah putus, kasih sayang yang tulus, dan dukungan tanpa batas. Orang tua adalah sumber kekuatan penulis untuk terus maju dan menyelesaikan setiap tantangan.
2. Keluarga tercinta yang selalu mendoakan dan memberi semangat. Kehadiran dan doa dari kalian memberi penulis kekuatan dalam menyelesaikan perjalanan ini. Segala pencapaian ini tak akan terwujud tanpa kehadiran dan cinta kalian.
3. Dosen pembimbing yang telah dengan sabar, bijaksana, dan penuh dedikasi membimbing saya selama proses penyusunan skripsi ini. Terima kasih atas setiap waktu, tenaga, dan perhatian yang Ibu berikan.
4. Teman – teman seperjuangan yang telah berbagi pengetahuan, pengalaman satu sama lain selama perjalanan akademik kami di perguruan tinggi.
5. Almamater Tercinta, semoga karya sederhana ini menjadi sumbangsih kecil yang bermanfaat, bagi perkembangan ilmu pengetahuan dan kemajuan almamater tercinta, dan dapat memberi manfaat yang lebih luas.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, petunjuk, dan karunia-Nya yang telah membantu penulis menyelesaikan skripsi dengan judul "Analisis Sentimen pada Review Pengguna Aplikasi Mobile JKN di Google Play Store Menggunakan Metode Support Vector Machine". Skripsi ini disusun untuk memenuhi persyaratan kelulusan Program Studi SI- Informatika pada Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta. Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari doa, bimbingan, dukungan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Keluarga tersayang, yang senantiasa memberikan doa terbaik, dukungan moral, materi, serta semangat dalam perjalanan penyusunan skripsi ini.
2. Ibu Windha Mega Pradnya Dhuhiya, M.Kom., selaku Ketua Program Studi Informatika Universitas Amikom Yogyakarta serta sebagai Dosen Pembimbing Skripsi yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penulis dalam penyusunan skripsi ini. Terima kasih atas ilmu dan pengetahuan yang telah diberikan.
3. Segenap Dosen Program Studi Informatika yang telah mendidik serta memberikan ilmu selama kuliah dan seluruh staff yang selalu sabar melayani segala administrasi
4. Sahabat dan teman, yang telah berbagi ilmu, pengalaman, dan inspirasi sepanjang perjalanan akademik di Universitas Amikom Yogyakarta.
5. Semua individu yang turut menyumbangkan gagasan, pandangan, dan sokongan dalam berbagai bentuk, yang telah membantu kelancaran perjalanan penyelesaian skripsi ini.

Yogyakarta, 24 Januari 2025

Penulis

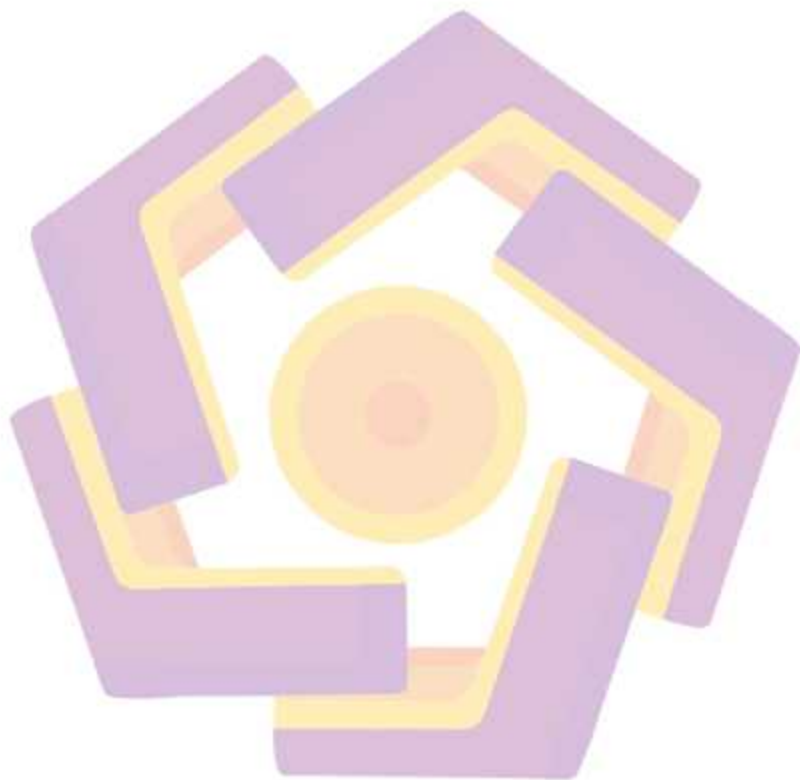
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMBANG DAN SINGKATAN	xi
DAFTAR ISTILAH.....	xii
INTISARI	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
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.....	12

2.2.1	Analisis Sentimen	12
BAB III METODE PENELITIAN		19
3.1	Alur Penelitian	19
3.3	Alat dan Bahan	24
BAB IV HASIL DAN PEMBAHASAN		26
4.1	Scrapping	26
4.2	Preprocessing Data	26
4.2.1	Case Folding	27
4.2.2	Stopword	27
4.2.3	Tokenizing	28
4.2.4	Stemming	29
4.3	Pelabelan Data	30
4.4	Visualisasi	30
4.5	Splitting	33
4.6	TF-IDF	34
4.7	Resampling Data	34
4.8	Modeling dan Evaluasi	35
BAB V PENUTUP		39
5.1	Kesimpulan	39
5.2	Saran	40
REFERENSI		41

DAFTAR TABEL

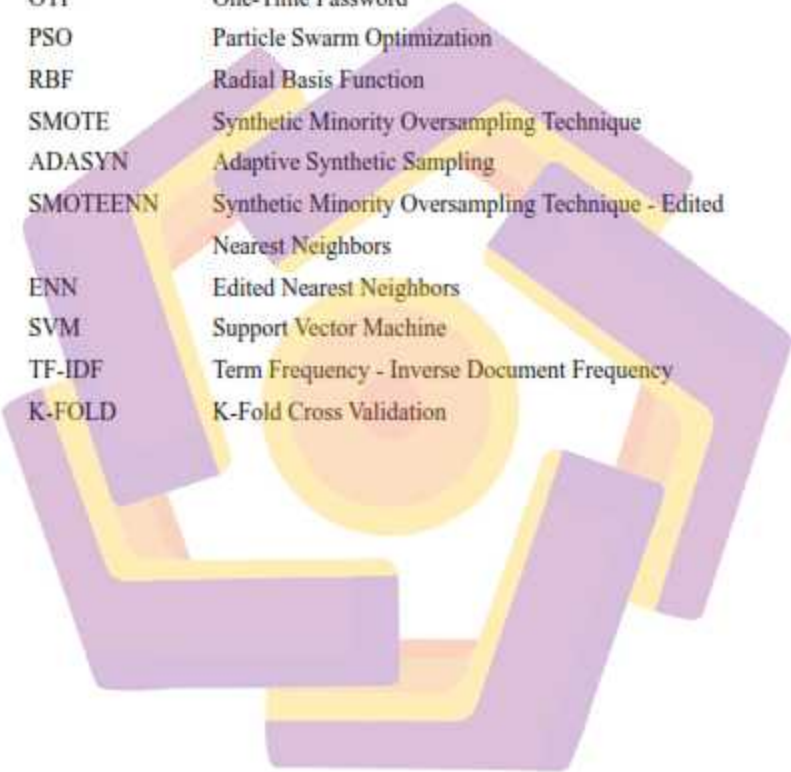
Tabel 2. 1 Keaslian Penelitian	9
Tabel 4. 1 Hasil Proses Evaluasi 10-Fold Cross-Validation.....	36



DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	19
Gambar 4. 1 Hasil Scrapping	26
Gambar 4. 2 Hasil Case Folding	27
Gambar 4. 3 Hasil Stopword	28
Gambar 4. 4 Hasil Tokenizing	28
Gambar 4. 5 Hasil Normalisasi	29
Gambar 4. 6 Hasil Stemming	29
Gambar 4. 7 Hasil Pelabelan Data	30
Gambar 4. 8 WordCloud Sentimen Negatif	31
Gambar 4. 9 WordCloud Sentimen Positif	31
Gambar 4. 10 Frekuensi Kata Negatif	32
Gambar 4. 11 Frekuensi Kata Positif	33
Gambar 4. 12 Proses TF-IDF	34
Gambar 4. 13 Hasil Resampling Data	35
Gambar 4. 14 Sebelum Resampling	37
Gambar 4. 15 SMOTE	37
Gambar 4. 16 ADASYN	37
Gambar 4. 17 Random Undersampling	38
Gambar 4. 18 SMOTEENN	38

DAFTAR LAMBANG DAN SINGKATAN



BPJS	Badan Penyelenggara Jaminan Sosial
JKN	Jaminan Kesehatan Nasional
NLP	Natural Language Processing
OTP	One-Time Password
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
SMOTE	Synthetic Minority Oversampling Technique
ADASYN	Adaptive Synthetic Sampling
SMOTEENN	Synthetic Minority Oversampling Technique - Edited
	Nearest Neighbors
ENN	Edited Nearest Neighbors
SVM	Support Vector Machine
TF-IDF	Term Frequency - Inverse Document Frequency
K-FOLD	K-Fold Cross Validation

DAFTAR ISTILAH

Analisis Sentimen	Identifikasi emosi atau opini dalam teks
Algoritma SVM	Metode klasifikasi untuk memisahkan data dengan hyperplane optimal.
Crawling	Pengumpulan data otomatis dari web
Dataset	Kumpulan data untuk analisis atau pelatihan.
Preprocessing	Proses persiapan data sebelum dianalisis, termasuk cleaning, tokenizing, dan stopwords removal.
Resampling Data	Penyesuaian data untuk mengatasi ketidakseimbangan kelas.
Case Folding	Proses mengubah seluruh huruf dalam teks menjadi huruf kecil dan menghilangkan karakter yang tidak relevan.
Stopword Removal	Penghapusan kata-kata umum yang tidak memiliki makna signifikan dalam analisis data.
Stemming	Proses mengubah kata-kata dalam teks menjadi bentuk dasar atau akar katanya.
Splitting	Proses membagi dataset menjadi data pelatihan dan data pengujian untuk memastikan model tidak overfitting.
Cross Validation	Teknik statistik untuk mengevaluasi kinerja model dengan membagi dataset menjadi data pelatihan dan pengujian secara terstruktur.
Hybrid Sampling	Kombinasi teknik oversampling dan undersampling untuk menangani ketidakseimbangan kelas pada dataset.
Oversampling	Teknik resampling untuk menambah jumlah data pada kelas minoritas.
Undersampling	Teknik resampling yang mengurangi jumlah data pada kelas mayoritas.
Tokenizing	Proses memecah teks menjadi unit-unit kecil, seperti kata atau kalimat.
Visualisasi	Penyajian data dalam bentuk grafis.

Wordcloud	Representasi visual kata berdasarkan frekuensi.
TF-IDF	Metode untuk memberikan bobot pada kata berdasarkan frekuensi kemunculan dalam dokumen dan seberapa jarang kata tersebut muncul dalam korpus.
Sastrawi	Library Python yang digunakan untuk proses stemming dan analisis teks dalam bahasa Indonesia
Pandas	Library Python yang digunakan untuk manipulasi dan analisis data.
Akurasi	Ukuran yang menggambarkan seberapa banyak prediksi yang benar dari keseluruhan prediksi yang dilakukan oleh model.
Precision	Ukuran yang menunjukkan seberapa banyak prediksi positif yang benar dari semua prediksi positif yang dibuat oleh model.
Recall	Ukuran yang menunjukkan seberapa banyak data positif yang berhasil diidentifikasi oleh model dari semua data positif yang ada.
F1-Score	Ukuran performa model yang merupakan rata-rata harmonis dari precision dan recall.
Labeling	Proses pemberian label atau kategori pada data untuk mengklasifikasikan sentimen atau jenis data lainnya.
Overfitting	Kondisi di mana model pembelajaran mesin terlalu menyesuaikan diri dengan data pelatihan, sehingga memiliki performa yang buruk pada data pengujian.

INTISARI

Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap aplikasi Mobile JKN yang diunduh dari Google Play Store menggunakan algoritma Support Vector Machine (SVM). Salah satu tantangan utama adalah ketidakseimbangan kelas pada dataset ulasan, di mana jumlah ulasan positif dan negatif tidak seimbang. Penelitian ini membandingkan empat teknik resampling, yaitu SMOTE, ADASYN, Random Undersampling, dan SMOTEENN, untuk mengatasi masalah ketidakseimbangan kelas.

Tahapan penelitian meliputi pengumpulan data ulasan melalui web scraping, preprocessing untuk membersihkan teks, pemberian label sentimen berdasarkan skor ulasan, dan visualisasi data. Dataset diubah menjadi representasi numerik menggunakan TF-IDF. Keempat teknik resampling diterapkan untuk menyeimbangkan kelas sebelum dilakukan klasifikasi sentimen menggunakan SVM. Evaluasi dilakukan menggunakan K-Fold Cross Validation untuk mengukur kinerja model berdasarkan akurasi, precision, recall, dan F1-score.

Hasil penelitian menunjukkan mayoritas ulasan pengguna mencerminkan sentimen positif terhadap aplikasi, terutama terkait kemudahan penggunaan dan manfaatnya. Namun, beberapa ulasan negatif mengindikasikan masalah pada proses verifikasi dan pendaftaran. Serta menunjukan bahwa teknik resampling mampu meningkatkan kinerja model SVM secara signifikan. Teknik SMOTEENN memberikan akurasi tertinggi sebesar 98,75%, diikuti oleh ADASYN, SMOTE, dan Random Undersampling.

Kata kunci: analisis sentimen, Support Vector Machine, teknik resampling, SMOTEENN, Mobile JKN.

ABSTRACT

This study aims to analyze user sentiment toward the Mobile JKN application, downloaded from Google Play Store, using the Support Vector Machine (SVM) algorithm. One of the main challenges is class imbalance in the dataset, where the number of positive and negative reviews is unequal. This study compares four resampling techniques: SMOTE, ADASYN, Random Undersampling, and SMOTEENN to address the class imbalance issue.

The research stages include collecting user reviews through web scraping, preprocessing the text for cleaning, labeling sentiments based on review scores, and visualizing the data. The dataset was converted into a numerical representation using TF-IDF. The four resampling techniques were applied to balance the classes before performing sentiment classification using SVM. Evaluation was conducted using K-Fold Cross Validation to measure the model's performance in terms of accuracy, precision, recall, and F1-score.

The research results indicate that the majority of user reviews reflect positive sentiment toward the application, particularly regarding ease of use and its benefits. However, some negative reviews highlight issues with the verification and registration processes. Additionally, the findings show that resampling techniques significantly enhance the performance of the SVM model. The SMOTEENN technique achieved the highest accuracy of 98.75%, followed by ADASYN, SMOTE, and Random Undersampling.

Keyword: *sentiment analysis, Support Vector Machine, resampling techniques, SMOTEENN, Mobile JKN*