

**ANALISIS SENTIMEN PENGGUNA APLIKASI SIGNAL-
SAMSAT DIGITAL NASIONAL DI GOOGLE PLAY STORE
MENGUNAKAN MODEL MACHINE LEARNING**

SKRIPSI

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



disusun oleh

ADHE PUTRA KUSUMA DWI WARDHANA

20.11.3440

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**ANALISIS SENTIMEN PENGGUNA APLIKASI SIGNAL-
SAMSAT DIGITAL NASIONAL DI GOOGLE PLAY STORE
MENGUNAKAN MODEL MACHINE LEARNING**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ADHE PUTRA KUSUMA DWI WARDHANA
20.11.3440

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2024

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PENGGUNA APLIKASI SIGNAL-
SAMSAT DIGITAL NASIONAL DI GOOGLE PLAY STORE
MENGUNAKAN MODEL MACHINE LEARNING**

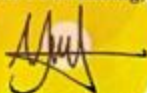
yang disusun dan diajukan oleh

Adhe Putra Kusuma Dwi Wardhana

20.11.3440

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 28 Juni 2024

Dosen Pembimbing,



Yoga Pristvanto, S.Kom, M.Eng

NIK. 190302412

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN PENGGUNA APLIKASI SIGNAL-
SAMSAT DIGITAL NASIONAL DI GOOGLE PLAY STORE
MENGUNAKAN MODEL MACHINE LEARNING**

yang disusun dan diajukan oleh

Adhe Putra Kusuma Dwi Wardhana

20.11.3440

Telah dipertahankan di depan Dewan Penguji
pada tanggal 28 Juni 2024

Susunan Dewan Penguji

Nama Penguji

Anggit Ferdita Nugraha, S.T., M.Eng.
NIK. 190302480

Acihmah Sidauruk, M.Kom
NIK. 190302238

Yoga Pristyanto, S.Kom, M.Eng
NIK. 190302412

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 28 Juni 2024

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 : Adhe Putra Kusuma Dwi Wardhana
NIM : 20.11.3440

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN PENGGUNA APLIKASI SIGNAL-SAMSAT
DIGITAL NASIONAL DI GOOGLE PLAY STORE MENGGUNAKAN
MODEL MACHINE LEARNING**

Dosen Pembimbing : Yoga Pristyanto, S.Kom, M.Eng

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 Juni 2024

Yang Menyatakan,



Adhe Putra Kusuma Dwi Wardhana

HALAMAN PERSEMBAHAN

Alhamdulillahirobbil'alamin rasa syukur yang mendalam, dengan telah diselesaikannya skripsi ini. Penulis mempersembahkan skripsi ini kepada semua pihak yang memberi motivasi, inspirasi dan dukungannya baik secara langsung maupun tidak langsung dalam proses pembuatan.

1. Allah SWT yang dengan segala rahmat dan karunia-Nya lah seluruh rangkaian proses pengerjaan skripsi ini berjalan dengan lancar, dan dapat selesai tepat waktu serta memperoleh hasil yang maksimal.
2. Kedua orang tua saya, mamah dan bapak yang selalu memberikan dukungan, doa, dan kasih sayang yang tak henti – hentinya kalian berikan.
3. Kakak Putri, Mas Oki, Mas Agus, Mas Aji, Tante Dewi, Adik Arya, dan semua keluarga besar yang selalu memberikan semangat dan hiburan dikala sedang menghadapi kesulitan.
4. Bapak Yoga Pristyanto, S.Kom, M.Eng selaku dosen pembimbing yang selalu membantu, membimbing, dan mengarahkan selama proses penyusunan skripsi ini.
5. Bapak dan Ibu Dosen FIK yang telah memberikan banyak ilmu dan pengalaman selama perkuliahan.
6. Kepada sahabat-sahabat saya yang senantiasa membantu dan memberikan dukungan setiap saat terkhusus Sendi, Rolis, Carlos, dan masih banyak lagi yang tidak bisa saya sebutkan.
7. Teman-teman kelas 20-IF-03 yang telah berbagi ilmu serta pengalaman selama masa perkuliahan dikelas.
8. Seluruh pihak yang tidak bisa saya sebutkan satu per satu, terimakasih atas bantuan dan doa – doanya sehingga terselesaikan skripsi ini

KATA PENGANTAR

Alhamdulillah segala puji dan syukur bagi Allah SWT dan segala Rahmat serta hidayahnya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Pengguna Aplikasi SIGNAL–Samsat Digital Nasional Di Google Play Store Menggunakan Model Machine Learning”. Skripsi ini diajukan sebagai salah satu persyaratan untuk mendapatkan gelar sarjana di Program Studi Informatika Universitas AMIKOM Yogyakarta.

Dalam penyusunan dan penulisan skripsi ini tidak lepas dari bimbingan, bantuan, motivasi dan pengarahan dari berbagai pihak, maka pada kesempatan ini saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom., Ph.D. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Windha Mega Pradnya Duhita, M.Kom. selaku Kepala Program Studi Informatika Universitas Amikom Yogyakarta.
4. Bapak Yoga Pristyanto, S.Kom, M.Eng selaku dosen pembimbing yang selalu membantu, membimbing, dan mengarahkan selama proses penyusunan skripsi ini.
5. Bapak dan Ibu Dosen Universitas Amikom Yogyakarta yang telah banyak memberikan ilmunya selama kuliah.
6. Semua pihak yang telah membantu baik secara langsung maupun tidak langsung.

Dalam melakukan penulisan skripsi ini masih banyak kekurangan yang harus diperbaiki, jadi penulis mengharapkan kritik dan saran yang membangun untuk membantu skripsi menjadi lebih baik.

Yogyakarta, 28 Juni 2024

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	x
DAFTAR GAMBAR	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
INTISARI	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	4
2.1 Studi Literatur	4
2.2 Dasar Teori	9
2.2.1. Analisis Sentimen	9

2.2.2. <i>Machine Learning</i>	9
2.2.3. <i>Naïve Bayes</i>	10
2.2.4. <i>Support Vector Machine</i>	12
2.2.5. <i>Random Forest</i>	13
2.2.6. <i>Decision Tree</i>	14
2.2.7. Evaluasi Model	15
2.2.8. <i>Google Colaboratory</i>	17
BAB III METODE PENELITIAN	18
3.1 Objek Penelitian	18
3.2 Alur Penelitian	18
3.2.1. Pengumpulan Data	19
3.2.2. <i>Text preprocessing</i>	19
3.2.3. Pembobotan TF-IDF	23
3.2.4. Evaluasi	24
3.3 Alat dan Bahan	24
BAB IV HASIL DAN PEMBAHASAN	26
4.1 Pengumpulan Data	26
4.2 <i>Text Preprocessing</i>	30
4.3 Pembobotan TF-IDF	39
4.4 Evaluasi	40
4.4.1 Evaluasi <i>Naïve Bayes</i>	41
4.4.2 Evaluasi <i>Support Vector Machine</i>	44
4.4.3 Evaluasi <i>Random Forest</i>	46
4.4.4 Evaluasi <i>Decision Tree</i>	49
4.4.5 Perbandingan Hasil Evaluasi Model <i>Machine Learning</i>	51

4.5 Visualisasi	52
BAB V PENUTUP	58
5.1 Kesimpulan	58
5.2 Saran	59
REFERENSI	60

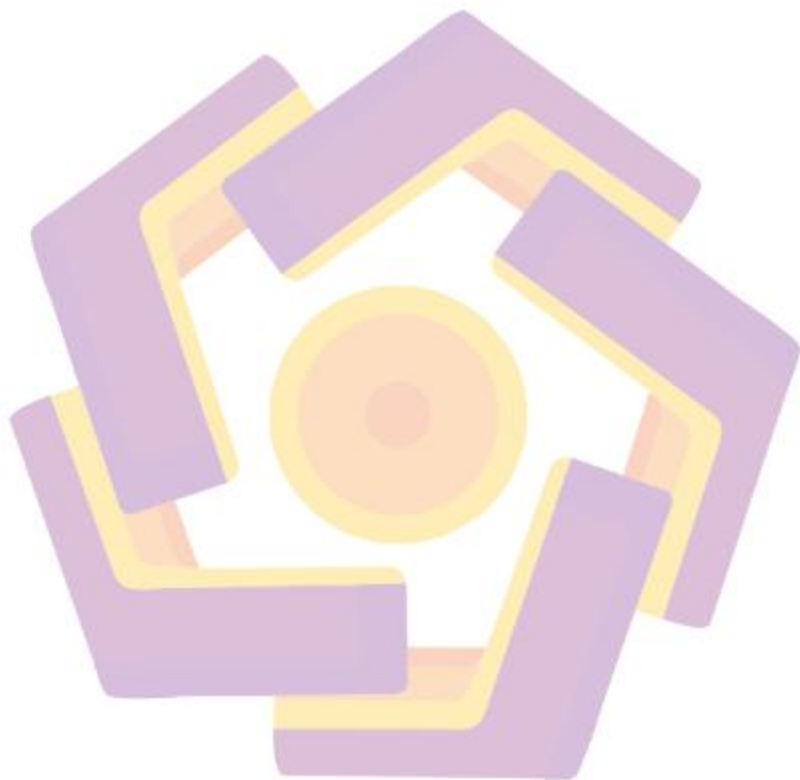


DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian.....	6
Tabel 2.2 <i>Confusion Matrix</i>	15
Tabel 3.1 Pengumpulan Data	19
Tabel 3.2 <i>Labeling</i>	20
Tabel 3.3 <i>Data Cleaning</i>	21
Tabel 3.4 <i>Case Folding</i>	21
Tabel 3.5 <i>Tokenizing</i>	22
Tabel 3.6 <i>Stopword Removal</i>	22
Tabel 3.7 <i>Stemming</i>	23
Tabel 3.8 Pembobotan TF-IDF	24
Tabel 3.9 Spesifikasi Komputer.....	25
Tabel 4.1 Hasil Pengumpulan Data.....	27
Tabel 4.2 30 Kata Dominan Sebelum <i>Preprocessing</i>	29
Tabel 4.3 Hasil <i>Labelling</i>	31
Tabel 4.4 Hasil <i>Data Cleaning</i>	32
Tabel 4.5 Hasil <i>Case Folding</i>	33
Tabel 4.6 Hasil <i>Tokenizing</i>	35
Tabel 4.7 Hasil <i>Stopwords Removal</i>	36
Tabel 4.8 Hasil <i>Stemming</i>	37
Tabel 4.9 30 Kata Dominan Setelah <i>Preprocessing</i>	38
Tabel 4.10 Hasil Pembobotan TF-IDF	39
Tabel 4.11 Hasil Akurasi Model Berdasarkan Data Latih dan Data Uji.....	40
Tabel 4.12 Hasil <i>Confusion Matrix Naïve Bayes</i>	42
Tabel 4.13 Hasil Evaluasi <i>Naïve Bayes</i>	43
Tabel 4.14 Hasil <i>Confusion Matrix Support Vector Machine</i>	44
Tabel 4.15 Hasil Evaluasi <i>Support Vector Machine</i>	45
Tabel 4.16 Hasil <i>Confusion Matrix Random Forest</i>	47
Tabel 4.17 Hasil Evaluasi <i>Random Forest</i>	48
Tabel 4.18 Hasil <i>Confusion Matrix Decision Tree</i>	49

Tabel 4.19 Hasil Evaluasi *Decision Tree*50

Tabel 4.20 Perbandingan Hasil Evaluasi Model *Machine Learning*.....51

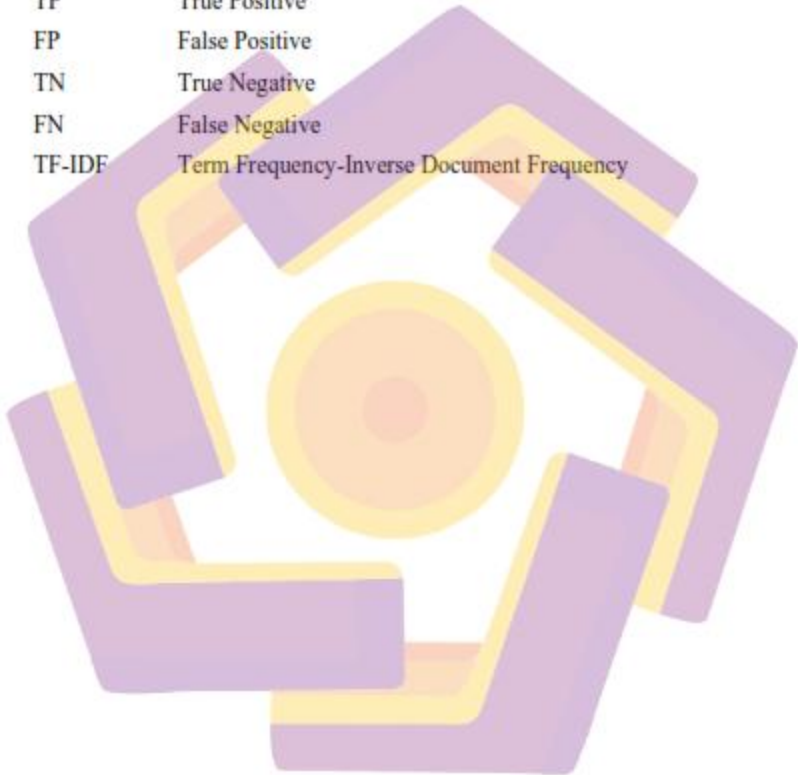


DAFTAR GAMBAR

Gambar 2.1 <i>Machine Learning</i>	9
Gambar 2.2 Model <i>Naïve Bayes</i>	10
Gambar 2.3 Model <i>Support Vector Machine</i>	12
Gambar 2.4 Model <i>Random Forest</i>	13
Gambar 2.5 Model <i>Decision Tree</i>	14
Gambar 3.1 Alur Penelitian.....	18
Gambar 4.1 Jumlah Ulasan Aplikasi SIGNAL	28
Gambar 4.2 <i>Bar Chart</i> sentimen positif.....	53
Gambar 4.3 <i>Word Cloud</i> sentimen positif.....	54
Gambar 4.4 <i>Bar Chart</i> sentimen negatif	56
Gambar 4.5 <i>Word Cloud</i> sentimen negatif.....	56

DAFTAR LAMBANG DAN SINGKATAN

NB	Naïve Bayes
SVM	Support Vector Machine
RF	Random Forest
TP	True Positive
FP	False Positive
TN	True Negative
FN	False Negative
TF-IDF	Term Frequency-Inverse Document Frequency



INTISARI

Aplikasi SIGNAL merupakan inovasi dari POLRI untuk memudahkan masyarakat membayar pajak kendaraan bermotor secara aman dan mudah. Kehadiran aplikasi SIGNAL menimbulkan tanggapan pro dan kontra dari jutaan pengguna aplikasi tersebut. Para pengguna memiliki berbagai komentar positif dan negatif terhadap aplikasi SIGNAL yang dirilis pada tahun 2021.

Pada penelitian ini, penulis menggunakan model *Naïve Bayes*, *Support Vector Machine*, *Random Forest*, dan *Decision Tree* untuk mengklasifikasikan sentimen ulasan pengguna aplikasi SIGNAL di Google Play Store menjadi dua kelas yaitu positif dan negatif. Jumlah ulasan yang digunakan adalah 10.000 data. Tahapan-tahapan yang dilakukan dimulai dari pengumpulan data, *labeling*, *data cleaning*, *case folding*, *tokenizing*, *stopword removal*, *stemming*, pembobotan TF-IDF, dan evaluasi model.

Hasil penelitian diperoleh nilai akurasi masing-masing model *machine learning* yaitu untuk NB sebesar 89%, SVM sebesar 92%, *Random Forest* sebesar 91%, dan *Decision Tree* sebesar 89%. Dalam penelitian ini dapat diketahui bahwa tingkat akurasi yang didapatkan algoritma *Support Vector Machine* (SVM) lebih unggul dibandingkan ke tiga model *machine learning* lainnya. Sehingga model SVM lebih akurat pada analisis sentimen aplikasi SIGNAL.

Kata kunci: Analisis Sentimen, *Machine Learning*, SIGNAL, *Naïve Bayes*, *Support Vector Machine*, *Random Forest*, *Decision Tree*.

ABSTRACT

The SIGNAL application is an innovation from POLRI to make it easier for people to pay motor vehicle taxes safely and easily. The presence of the SIGNAL application has generated pro and con responses from millions of users of the application. Users have various positive and negative comments about the SIGNAL application which was released in 2021.

In this research, the authors used the Naïve Bayes, Support Vector Machine, Random Forest, and Decision Tree models to classify the sentiment of SIGNAL application user reviews on the Google Play Store into two classes, namely positive and negative. The number of reviews used is 10,000 data. The stages carried out start from data collection, labeling, data cleaning, case folding, tokenizing, stopword removal, stemming, TF-IDF weighting, and model evaluation.

The results of the research obtained the accuracy value of each machine learning model, namely for NB of 89%. SVM is 92%, Random Forest is 91%, and Decision Tree is 89%. In this research, the level of accuracy obtained by the Support Vector Machine (SVM) algorithm is superior to the other three machine learning models. So that the SVM model is more accurate in analyzing the sentiment of the SIGNAL application.

Keyword: Sentiment Analysis, Machine Learning, SIGNAL, Naïve Bayes, Support Vector Machine, Random Forest, Decision Tree.