



SKRIPSI

**ANALISIS SENTIMEN MENGGUNAKAN MODEL
SUPPORT VECTOR MACHINE MENGENAI
KERUSAKAN RANGKA ESAF DENGAN
PENERAPAN STREAMLIT**

KINANTHI PUTRI ARIYANI
NPM 21083010047

DOSEN PEMBIMBING
Aviolla Terza Damaliana, S.Si., M.Stat.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

**ANALISIS SENTIMEN MENGGUNAKAN MODEL
SUPPORT VECTOR MACHINE MENGENAI
KERUSAKAN RANGKA ESAF DENGAN
PENERAPAN STREAMLIT**

KINANTHI PUTRI ARIYANI
NPM 21083010047

DOSEN PEMBIMBING
Aviollla Terza Damaliana, S.Si., M.Stat.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

**ANALISIS SENTIMEN MENGGUNAKAN MODEL
SUPPORT VECTOR MACHINE MENGENAI
KERUSAKAN RANGKA ESAF DENGAN
PENERAPAN STREAMLIT**

KINANTHI PUTRI ARIYANI
NPM 21083010047

DOSEN PEMBIMBING
Aviollla Terza Damaliana, S.Si., M.Stat.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**

LEMBAR PENGESAHAN

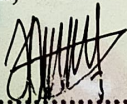
ANALISIS SENTIMEN MENGGUNAKAN MODEL *SUPPORT VECTOR MACHINE* MENGENAI KERUSAKAN RANGKA ESAF DENGAN PENERAPAN STREAMLIT

Oleh:
KINANTHI PUTRI ARIYANI
NPM. 21083010047

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Sidang Skripsi Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 7 Maret 2025:

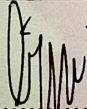
Menyetujui,

Aviolla Terza Damaliana, S.Si., M.Stat.
NIP. 19940802 202203 2 015


.....


(Pembimbing I)

Trimono, S.Si., M.Si.
NIP. 19950908 202203 1 003


.....


(Pembimbing II)

Kartika Maulida Hindrayani, S.Kom.,
M.Kom.
NIP. 19920909 202203 2 009


.....

(Ketua Penguji)

Shindi Shella May Wara, M.Stat.
NIP. 19960518 202406 2 003


.....

(Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer




Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

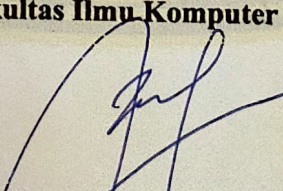
ANALISIS SENTIMEN MENGGUNAKAN MODEL *SUPPORT VECTOR MACHINE* MENGENAI KERUSAKAN RANGKA ESAF DENGAN PENERAPAN STREAMLIT

Oleh:
KINANTHI PUTRI ARIYANI
NPM. 21083010047

Telah disetujui untuk mengikuti Ujian Skripsi

Menyetujui,

**Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer**


Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.
NIP. 19801205 200501 1 002

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan dibawah ini:

Nama : Kinanthi Putri Ariyani
NPM : 21083010047
Program : Sarjana (S1)
Program Studi : Sains Data
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 7 Maret 2025

Yang Membuat Pernyataan



Kinanthi Putri Ariyani
NPM. 21083010047

ABSTRAK

Nama Mahasiswa / NPM : Kinanthi Putri Ariyani / 21083010047
Judul Skripsi : Analisis Sentimen Menggunakan Model *Support Vector Machine* Mengenai Kerusakan Rangka eSAF dengan Penerapan Streamlit
Dosen Pembimbing : 1. Aviolla Terza Damaliana, S.Si., M.Stat.
2. Trimono, S.Si., M.Si.

Kerusakan rangka *Enhanced Smart Architecture Frame* (eSAF) pada sepeda motor Honda telah memicu kekhawatiran konsumen dan menjadi sorotan publik. Penelitian ini mengkaji sentimen masyarakat terhadap masalah tersebut menggunakan *Support Vector Machine* (SVM) serta dampaknya terhadap penjualan di salah satu dealer di Surabaya. Data yang digunakan berupa komentar dari media sosial Twitter dengan membagi data ke dalam dua kelompok klasifikasi utama yakni sentimen positif dan sentimen negatif. Berdasarkan hasil analisis, mayoritas 589 sentimen masyarakat (59,7%) cenderung negatif terhadap masalah kerusakan rangka eSAF, sedangkan 397 sentimen masyarakat (40,3%) menunjukkan sentimen positif. Hasil penjualan menunjukkan sentimen negatif di media sosial tidak memengaruhi angka penjualan. Model SVM dengan *kernel Linear* dan *Polynomial* memberikan hasil terbaik dengan akurasi 85%, *precision* 85%, *recall* 85%, dan *f1-score* 85%. SVM dipilih karena unggul dalam klasifikasi teks dibandingkan dengan algoritma seperti *K-Nearest Neighbors* (KNN), C4.5, dan *Naïve Bayes*, serta telah diterapkan dalam berbagai bidang seperti deteksi wajah, bioinformatika, dan pemrosesan teks. Penelitian ini memberikan wawasan bagi produsen untuk memperbaiki kualitas produk, meningkatkan layanan konsumen, dan memulihkan kepercayaan publik. Selain itu, penggunaan algoritma SVM dalam analisis sentimen ini dapat menjadi referensi bagi penelitian serupa di bidang lain.

Kata kunci: Analisis Sentimen, Penjualan, Rangka eSAF, *Support Vector Machine*

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Kinanthi Putri Ariyani / 21083010047
Thesis Title : *Sentiment Analysis Using Support Vector Machine Model Regarding eSAF Frame Damage with Streamlit Implementation*
Advisor : 1. Aviolla Terza Damaliana, S.Si., M.Stat.
2. Trimono, S.Si., M.Si.

The damage to the Enhanced Smart Architecture Frame (eSAF) on Honda motorcycles has raised consumer concerns and become a public spotlight. This study examines public sentiment towards the problem using Support Vector Machine (SVM) and its impact on sales at a dealer in Surabaya. The data used are comments from Twitter social media by dividing the data into two main classification groups, namely positive sentiment and negative sentiment. Based on the analysis results, the majority of 589 public sentiments (59.7%) tend to be negative towards the eSAF frame damage problem, while 397 public sentiments (40.3%) show positive sentiment. The sales results show that negative sentiment on social media does not affect sales figures. The SVM model with Linear and Polynomial kernels gave the best results with 85% accuracy, 85% precision, 85% recall, and 85% f1-score. SVM was chosen because it excels in text classification compared to algorithms such as K-Nearest Neighbors (KNN), C4.5, and Naïve Bayes, and has been applied in various fields such as face detection, bioinformatics, and text processing. This research provides insights for manufacturers to improve product quality, enhance consumer service, and restore public trust. In addition, the use of the SVM algorithm in this sentiment analysis can be a reference for similar research in other fields.

Keywords: *eSAF Frame, Sales, Sentiment Analysis, Support Vector Machine*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Analisis Sentimen Menggunakan Model *Support Vector Machine* Mengenai Kerusakan Rangka eSAF dengan Penerapan Streamlit”** dapat terselesaikan dengan baik.

Penulis mengucapkan terima kasih kepada Ibu Aviolla Terza Damaliana, S.Si., M.Stat. selaku Dosen Pembimbing utama yang telah meluangkan waktunya untuk memberikan bimbingan, nasehat serta motivasi kepada penulis. Dan penulis juga banyak menerima bantuan dari berbagai pihak, baik itu berupa moril, spiritual maupun materiil. Untuk itu penulis mengucapkan terima kasih kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Bapak Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., Asean. Eng selaku Ketua Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Bapak/Ibu Dosen Program Studi Sains Data yang telah memberikan dukungan, ilmu pengetahuan, inspirasi, dan wawasan yang sangat berharga selama masa perkuliahan.
4. Orang Tua yang telah memberikan kesempatan saya dalam menempuh pendidikan, memberikan cinta, dukungan, dan doa yang tulus.
5. Teman-teman yang senantiasa selalu mendukung, memberi semangat, dan saling membantu dalam melewati masa-masa sulit.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, Maret 2025

Penulis

Halaman ini sengaja dikosongkan

DAFTAR ISI

LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	iv
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
DAFTAR NOTASI.....	xxi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA.....	7
2.1. Penelitian Terdahulu	7
2.2. Landasan Teori.....	10
2.1.1. Sentimen.....	10
2.1.2. Penjualan	11
2.1.3. Rangka eSAF	11
2.1.4. Pembobotan Kata TF-IDF.....	12
2.1.5. Seleksi Fitur <i>Mutual Information</i>	13
2.1.6. <i>Grid Search</i>	15
2.1.7. <i>Support Vector Machine</i>	15
2.1.8. <i>Confusion Matrix</i>	19
2.1.9. <i>Word Cloud</i>	20
2.1.10. <i>Graphical User Interface (GUI) Streamlit</i>	20

2.1.11.	Uji <i>Mann-Whitney</i>	21
BAB III DESAIN DAN IMPLEMENTASI SISTEM		23
3.1.	Metode dan Langkah Penelitian.....	23
3.1.1.	Variabel Penelitian dan Sumber Data	23
3.1.2.	Langkah Analisis.....	23
3.2.	Desain Sistem.....	32
BAB IV PENGUJIAN DAN ANALISA		35
4.1.	Pengumpulan Data	35
4.2.	<i>Preprocessing Data</i>	36
4.2.1.	<i>Data Cleansing</i>	36
4.2.2.	<i>Case Folding</i>	37
4.2.3.	<i>Tokenizing</i>	38
4.2.4.	Normalisasi	39
4.2.5.	<i>Filtering</i>	41
4.2.6.	<i>Stemming</i>	42
4.2.7.	Penggabungan Kata.....	44
4.3.	Pelabelan Data.....	45
4.3.1.	<i>Handling Missing Value and Remove Duplicate</i>	45
4.3.2.	Pelabelan Data.....	45
4.4.	Pembobotan Kata TF-IDF.....	48
4.5.	Seleksi Fitur <i>Mutual Information</i>	51
4.6.	Pembagian Data	54
4.7.	Klasifikasi Model <i>Support Vector Machine</i>	55
4.7.1.	<i>Grid Search</i>	55
4.7.2.	Mengambil Nilai w , b , dan α	57
4.7.3.	<i>Support Vector Machine</i> dengan <i>Kernel Linear</i>	59
4.7.4.	<i>Support Vector Machine</i> dengan <i>Kernel Polynomial</i>	60
4.7.5.	<i>Support Vector Machine</i> dengan Kernel RBF	62
4.8.	Evaluasi Model.....	63
4.8.1.	<i>Kernel Linear</i>	64
4.8.2.	<i>Kernel Polynomial</i>	65
4.8.3.	Kernel RBF	67

4.8.4.	Rangkuman Pengujian <i>Support Vector Machine</i>	68
4.9.	Visualisasi <i>WordCloud</i>	69
4.10.	Penerapan GUI Streamlit	70
4.10.1.	<i>Sidebar</i>	71
4.10.2.	<i>Navigation</i>	71
4.11.	Analisis Grafik Data Penjualan	77
BAB V	PENUTUP	81
5.1.	Kesimpulan	81
5.2.	Saran Pengembangan	82
DAFTAR PUSTAKA	83
LAMPIRAN	89

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2. 1 Hyperplane Terbaik dan Margin Maksimum [30].....	16
Gambar 3. 1 Diagram Alir Penelitian.....	24
Gambar 3. 2 Diagram Alir <i>Preprocessing</i>	25
Gambar 3. 3 Diagram Alir <i>Support Vector Machine</i>	29
Gambar 4. 1 Penghapusan Data Kosong	45
Gambar 4. 2 Pelabelan Textblob	45
Gambar 4. 3 Pelabelan VADER	46
Gambar 4. 4 Visualisasi Pelabelan Manual	47
Gambar 4. 5 Classification Report Kernel Linear	64
Gambar 4. 6 Confusion Matrix Kernel Linear	65
Gambar 4. 7 Classification Report Kernel Polynomial	65
Gambar 4. 8 Confusion Matrix Kernel Polynomial	66
Gambar 4. 9 Clasification Report Kernel RBF.....	67
Gambar 4. 10 Confusion Matrix Kernel RBF	67
Gambar 4. 11 WordCloud Sentimen Positif.....	69
Gambar 4. 12 WordCloud Sentimen Negatif	70
Gambar 4. 13 Sidebar Dashboard Streamlit.....	71
Gambar 4. 14 Tampilan Home	72
Gambar 4. 15 Tampilan Dataset.....	72
Gambar 4. 16 Tampilan Preprocessing (1).....	73
Gambar 4. 17 Tampilan Preprocessing (2).....	73
Gambar 4. 18 Tampilan Modeling (1).....	74
Gambar 4. 19 Tampilan Modeling (2).....	74
Gambar 4. 20 Tampilan Modeling (3).....	75
Gambar 4. 21 Tampilan Visualization (1).....	75
Gambar 4. 22 Tampilan Visualization (2).....	76
Gambar 4. 23 Tampilan Prediction Sentimen Positive.....	76
Gambar 4. 24 Tampilan Prediction Sentimen Negative	76
Gambar 4. 25 Grafik Penjualan	77

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	7
Tabel 2. 2. Kontingensi Mutual Information.....	14
Tabel 2. 3 Confusion Matrix	19
Tabel 3. 1 Parameter Kernel.....	30
Tabel 4. 1 Data Komentar	35
Tabel 4. 2 Hasil Data Cleansing	37
Tabel 4. 3 Hasil Case Folding	38
Tabel 4. 4 Hasil Tokenizing	39
Tabel 4. 5 Hasil Normalisasi	40
Tabel 4. 6 Hasil Filtering.....	42
Tabel 4. 7 Hasil Stemming	44
Tabel 4. 8 Hasil Penggabungan Kata	44
Tabel 4. 9 Hasil Pelabelan Menggunakan Library	46
Tabel 4. 10 Hasil Pelabelan Manual.....	47
Tabel 4. 11 Contoh Data Pembobotan TF-IDF.....	49
Tabel 4. 12 Hasil Pembobotan TF-IDF	50
Tabel 4. 13 Contoh Data Seleksi Fitur Mutual Information.....	52
Tabel 4. 14 Kontingensi Jumlah Kemunculan Term Honda.....	52
Tabel 4. 15 Hasil Seleksi Fitur Mutual Information	53
Tabel 4. 16 Hasil Pembobotan Pada Seleksi Fitur	54
Tabel 4. 17 Pembagian Data.....	55
Tabel 4. 18 Hasil Grid Search	57
Tabel 4. 19 Hasil Prediksi Data Uji Kernel Linear	60
Tabel 4. 20 Hasil Prediksi Data Uji Kernel Polynomial	61
Tabel 4. 21 Hasil Prediksi Data Uji Kernel RBF	63
Tabel 4. 22 Rangkuman Pengujian Support Vector Machine	68

Halaman ini sengaja dikosongkan

DAFTAR LAMPIRAN

Lampiran 1. Code Analisis dan Dashboard.....	89
Lampiran 2. Data Penelitian, File Pelabelan, Normalisasi Kata	90

Halaman ini sengaja dikosongkan

DAFTAR NOTASI

$f(w, d)$:	Frekuensi kemunculan kata w dalam dokumen d
w	:	Parameter hyperplane yang sedang dicari
x	:	Data masukan SVM dimana x_1 adalah indeks kata dan x_2 adalah bobot kata
b	:	Nilai bias atau parameter hyperplane yang juga dicari
$K(x_i, x)$:	Fungsi kernel
$f(x)$:	Fungsi keputusan
α	:	<i>alpha</i>
$P(U = et, C = ec)$:	Probabilitas gabungan dari U yang bernilai et dan C yang bernilai ec
TP	:	<i>True Positive</i>
TN	:	<i>True Negative</i>
FP	:	<i>False Positive</i>
FN	:	<i>False Negative</i>
c	:	Nilai Konstanta
d	:	Parameter <i>degree</i>
γ	:	Parameter <i>gamma</i>

Halaman ini sengaja dikosongkan