

**ANALISIS SENTIMEN ULASAN PENGGUNA KAI ACCESS
PADA GOOGLE PLAY STORE MENGGUNAKAN
ALGORITMA *SUPPORT VECTOR MACHINE* (SVM)**

SKRIPSI



Oleh :

EKA RIZQI MAR'ATUS SHOLIHAH

NPM. 19081010022

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR**

2024

**ANALISIS SENTIMEN ULASAN PENGGUNA KAI ACCESS
PADA GOOGLE PLAY STORE MENGGUNAKAN
ALGORITMA *SUPPORT VECTOR MACHINE* (SVM)**

SKRIPSI

Diajukan Untuk Memenuhi Sebagai Persyaratan Dalam Menempuh Gelar Sarjana
Komputer Program Studi Informatika



Oleh :

EKA RIZQI MAR'ATUS SHOLIHAH

NPM. 19081010022

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR**

2024

LEMBAR PENGESAHAN SKRIPSI

**Judul : ANALISIS SENTIMEN ULASAN PENGGUNA KAI ACCESS
PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA
SUPPORT VECTOR MACHINE (SVM)**

Oleh : EKA RIZQI MAR'ATUS SHOLIHAH

NPM : 19081010022

Telah Diseminarkan Dalam Ujian Skripsi Pada

Hari Jumat, Tanggal 5 Januari 2024

Mengetahui

Dosen Pembimbing

1

Dr. Ir. I Gede Susrama Mas Divasa, S.T., M.T., IPU
NIP. 19700619 2021211 009

Dosen Penguji

1

Fauwaz Ali Akbar, S.Kom., M.Kom
NIP. 19920317 2018031 002

2

Eva Yulia Puspaningrum, S.Kom., M. Kom.
NIP. 19890705 2021212 002

2

Achmad Junaidi, S.Kom., M.Kom
NPT : 3 7811 04 0199 1

Menyetujui

**Dekan
Fakultas Ilmu Komputer**

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

**Koordinator Program Studi
Informasi**

Fetty Tri Angraeni, S.Kom., M.Kom
NIP : 19820211 2021212 005

SURAT PERNYATAAN ANTI PLAGIAT

Saya, mahasiswa program studi Informatika UPN “Veteran” Jawa Timur,
yang bertanda tangan dibawah ini:

Nama : Eka Rizqi Mar’atus Sholihah

NPM : 19081010022

Menyatakan bahwa judul skripsi/tugas akhir yang saya ajukan dan kerjakan
dengan judul:

“ANALISIS SENTIMEN ULASAN PENGGUNA KAI ACCESS PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)”

Bukan merupakan plagiat dari skripsi/ tugas akhir/ penelitian orang lain
juga bukan merupakan produk dan atau perangkat lunak yang beli dari pihak lain.
Saya juga menyatakan bahwa skripsi/ tugas akhir ini adalah pekerjaan saya sendiri,
kecuali yang dinyatakan dalam daftar pustaka dan tidak pernah diajukan untuk
memperoleh gelar di UPN “Veteran” Jawa Timur maupun di institusi pendidikan
lain.

Jika ternyata di kemudian hari ini terbukti tidak benar, maka saya
siap menerima segala konsekuensinya.

Surabaya, 30 Desember 2023

Hormat Saya,



EKA RIZQI MAR'ATUS SHOLIHAH

NPM.19081010022

ANALISIS SENTIMEN ULASAN PENGGUNA KAI ACCESS PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)

Nama Mahasiswa : Eka Rizqi Mar'atus Sholihah

NPM : 19081010022

Program Studi : Informatika

Dosen Pembimbing : Dr. Ir. I Gede Susrama Mas Diyasa, S.T., M.T., IPU

Eva Yulia Puspaningrum, S. Kom., M.Kom.

ABSTRAK

KAI Access merupakan aplikasi *mobile online ticketing* milik PT Kereta Api Indonesia (Persero). Aplikasi ini dapat diunduh melalui platform penyedia aplikasi seperti Google Play Store. Pada Google Play Store, pengguna aplikasi dapat memberikan ulasan serta *rating* terkait aplikasi KAI Access. Ulasan serta *rating* ini akan menjadi pertimbangan untuk pelanggan lain yang akan mengunduh aplikasi karena hal tersebut akan membentuk persepsi pengguna terkait aplikasi tersebut. Untuk memahami ulasan atau mendeteksi sentimen yang diberikan pengguna terhadap aplikasi, penting dilakukan teknik analisis sentimen. Pada penelitian ini akan diimplementasikan analisis sentimen menggunakan algoritma *support vector machine* (SVM) pada ulasan aplikasi KAI Access di Google Play Store. Penelitian ini bertujuan untuk mengukur tingkat performansi algoritma SVM kernel linier dan kernel rbf dalam melakukan klasifikasi sentimen ke dalam dua kelas, yaitu kelas sentimen positif dan kelas sentimen negatif. Penelitian ini menggunakan 10000 data ulasan aplikasi KAI Acces di Google Play Store. Dari beberapa pengujian yang dilakukan, SVM Kernel linear mampu menghasilkan model analisis sentimen dengan akurasi tertinggi yaitu 83,1 %. Sedangkan untuk SVM Kernel RBF, dari beberapa pengujian yang telah dilakukan mampu menghasilkan mode analisis sentimen dengan akurasi tertinggi yaitu 86,1 %.

Kata kunci: *Support Vector Machine, Analisis Sentimen, KAI Access, Kernel Linear, Kernel RBF.*

KATA PENGANTAR

Puji dan syukur penulis panjatkan atas kehadiran Allah SWT yang telah memberikan rahmat dan petunjuk-Nya sehingga penulis dapat menyusun Laporan Akhir Mata Kuliah Skripsi dengan judul "Analisis Sentimen Ulasan Pengguna KAI Access Pada Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)". Skripsi ini merupakan bagian dari persyaratan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur.

Penulis menyadari bahwa skripsi ini belum mencapai kata sempurna dan masih memiliki kekurangan, baik dalam aspek penulisan maupun pembahasan. Oleh karena itu, penulis sangat mengharapkan adanya masukan, kritik, dan saran yang bersifat membangun dan dapat menyempurnakan skripsi ini.

Akhir kata, penulis mengucapkan banyak terima kasih yang sebesar-besarnya kepada semua pihak yang turut membantu dalam penyusunan laporan skripsi ini, sehingga dapat diselesaikan dengan baik. Penulis berharap agar skripsi ini dapat memberikan manfaat dan menambah informasi bagi semua pihak yang membaca. Penulis juga berharap agar skripsi ini dapat menjadi pelajaran bagi penulis maupun pembaca di kemudian hari.

Surabaya, 30 Desember 2023

Penulis

UCAPAN TERIMA KASIH

Puji dan syukur atas kehadiran Allah SWT atas berkah dan karunia-Nya yang telah memberikan kerahmatan dan kesehatan sehingga penulis dapat menyelesaikan penelitian dan penyusunan tugas akhir/skripsi ini. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Ir. Akhmad Fauzi, MMT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom., M. Kom. selaku Koordinator Program Studi Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Bapak Dr. Ir. I Gede Susrama Mas Diyasa, S.T., M.T., IPU. selaku dosen pembimbing pertama dan Ibu Eva Yulia Puspaningrum, S.Kom., M.Kom. selaku dosen pembimbing kedua yang telah meluangkan waktunya memberikan arahan kepada penulis selama penyusunan tugas akhir/skripsi.
5. Staff Dosen dan Tenaga Kependidikan program studi Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah memberikan ilmu dan pengalaman selama masa perkuliahan.
6. Orang Tua penulis, yang dengan penuh kasih sayang, motivasi dan dukungan dalam penyusunan tugas akhir/skripsi. Terima kasih atas dedikasi tanpa pamrih, kesabaran, dan cinta yang luar biasa yang telah diberikan kepada penulis.
7. Teman-teman tercinta penulis terutama Aditya Rizqi Ardhana, Alfina Damayanti, Almira Aurora Prameswaty, Anwar Sanusi, Fatimatus Zahroh, Maurisa Arimbi Putri, Kristian Rolando Limantara, Veronika Paskalia Sinambela dan Wariyanti Nugroho Putri yang telah memberikan semangat, wawasan, motivasi, perhatian dan pencerahan saat mengalami kendala, hiburan saat jenuh dan bosan sehingga penulis

dapat menyelesaikan penyusunan tugas akhir/skripsi ini serta berbagi cerita suka dan duka selama masa perkuliahan.

8. Seluruh teman-teman dari program studi Informatika Angkatan 2019 lainnya yang tentunya tidak dapat penulis sebutkan satu persatu yang telah membantu dan menjadi bagian dari kisah dan cerita perkuliahan.

Penulis berharap semoga Allah SWT membalas seluruh kebaikan yang telah diberikan selama proses penyusunan tugas akhir/skripsi. Semoga tugas akhir/skripsi ini dapat memberikan manfaat bagi pembaca ataupun penulis sendiri.

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	i
SURAT PERNYATAAN ANTI PLAGIAT.....	ii
ABSTRAK.....	iii
KATA PENGANTAR	iv
UCAPAN TERIMA KASIH	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
2.1. Penelitian Terdahulu	5
2.2. KAI <i>Access</i>	6
2.3. Google Play Store	6
2.4. <i>Scraping</i>	7
2.5. Analisis Sentimen	7
2.6. Sastrawi.....	7
2.7. <i>Preprocessing</i>	7
2.7.1. <i>Case Folding</i>	8
2.7.2. <i>Cleaning</i>	8

2.7.3. <i>Tokenizing</i>	8
2.7.4. <i>Normalization</i>	9
2.7.5. <i>Stop Word Removal</i>	9
2.7.6. <i>Stemming</i>	9
2.8. <i>Term Frequency Inverse Document Frequency (TF-IDF)</i>	10
2.9. <i>Support Vector Machine (SVM)</i>	10
2.10. <i>Confusion Matrix</i>	11
BAB III METODOLOGI PENELITIAN	13
3.1. <i>Metode Penelitian</i>	13
3.2. <i>Pengumpulan Data dan Labeling Data</i>	14
3.3. <i>Preprocessing Data</i>	14
3.3.1. <i>Case Folding</i>	14
3.3.2. <i>Cleaning</i>	15
3.3.3. <i>Tokenizing</i>	15
3.3.4. <i>Normalization</i>	15
3.3.5. <i>Stop Word Removal</i>	16
3.3.6. <i>Stemming</i>	17
3.4. <i>Transformation</i>	17
3.5. <i>Implementasi Algoritma</i>	21
3.5.1. <i>Kernel Linear</i>	22
3.5.2. <i>Kernel RBF</i>	26
3.6. <i>Skenario Pengujian</i>	30
BAB IV HASIL DAN PEMBAHASAN	32
4.1. <i>Kebutuhan Perangkat</i>	32
4.1.1. <i>Perangkat Keras</i>	32
4.1.2. <i>Perangkat Lunak</i>	32

4.2.	Pengumpulan Data.....	32
4.3.	Labeling Data.....	33
4.4.	<i>Preprocessing Data</i>	35
4.4.1.	<i>Case Folding</i>	36
4.4.2.	<i>Cleaning</i>	36
4.4.3.	<i>Tokenizing</i>	37
4.4.4.	<i>Normalization</i>	38
4.4.5.	<i>Stop Word Removal</i>	39
4.4.6.	<i>Stemming</i>	40
4.5.	Word Cloud.....	42
4.6.	<i>Transformation Data (TF-IDF)</i>	43
4.7.	Pembagian Data	43
4.8.	Proses Implementasi SVM.....	44
4.8.1.	Kernel Linear.....	44
4.8.2.	Kernel RBF	46
4.9.	Evaluasi Model	49
4.9.1.	Pengujian SVM Kernel Linear.....	50
1.	Evaluasi Model Pengujian Skenario 1	50
2.	Evaluasi Model Pengujian Skenario 2.....	52
3.	Evaluasi Model Pengujian Skenario 3	53
4.	Evaluasi Model Pengujian Skenario 4.....	54
5.	Evaluasi Model Pengujian Skenario 5	56
6.	Evaluasi Model Pengujian Skenario 6.....	57
4.9.2.	Pengujian SVM Kernel RBF.....	58
1.	Evaluasi Model Pengujian Skenario 1	58
2.	Evaluasi Model Pengujian Skenario 2.....	60

3.	Evaluasi Model Pengujian Skenario 3	61
4.	Evaluasi Model Pengujian Skenario 4	62
5.	Evaluasi Model Pengujian Skenario 5	64
6.	Evaluasi Model Pengujian Skenario 6	65
7.	Evaluasi Model Pengujian Skenario 7	66
8.	Evaluasi Model Pengujian Skenario 8	68
9.	Evaluasi Model Pengujian Skenario 9	69
10.	Evaluasi Model Pengujian Skenario 10	70
11.	Evaluasi Model Pengujian Skenario 11	72
12.	Evaluasi Model Pengujian Skenario 12	73
4.9.3.	Perbandingan Hasil Skenario Pengujian	74
BAB V	PENUTUP	79
5.1.	Kesimpulan	79
5.2.	Saran	80
DAFTAR PUSTAKA		81

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	13
Gambar 3.2 Diagram Alur TF-IDF.....	18
Gambar 3.3 Diagram Alur Algoritma SVM	21
Gambar 4.1 Output Hasil Scraping.....	33
Gambar 4.2 Rating Aplikasi di Google Play Store.....	33
Gambar 4.3 Data Hasil Labeling	34
Gambar 4.4 Jumlah masing-masing sentiment	35
Gambar 4.5 Atribut yang akan di proses	35
Gambar 4.6 Hasil Case Folding.....	36
Gambar 4.7 Hasil Proses Data Cleaning.....	37
Gambar 4.8 Hasil Proses Tokenisasi	38
Gambar 4.9 Hasil Proses Normalization.....	39
Gambar 4.10 Hasil Proses Stop Word Removal.....	40
Gambar 4.11 Hasil Proses Stemming	41
Gambar 4.12 Word Cloud Sentimen Positif.....	42
Gambar 4.13 Word Cloud Sentimen Negatif.....	42
Gambar 4.14 Output Prediksi 10 Data Teratas Dari Data Uji	45
Gambar 4.15 Output Prediksi 6 Data Teratas dari Data Uji	48
Gambar 4.16 Confusion Matrix SVM Kernel Linear Skenario 1	51
Gambar 4.17 Classification Report SVM Kernel Linear Skenario 1.....	51
Gambar 4.18 Confusion Matrix SVM Kernel Linear Skenario 2.....	52
Gambar 4.19 Classification Report SVM Kernel Linear Skenario 2.....	52
Gambar 4.20 Confusion Matrix SVM Kernel Linear Skenario 3.....	53
Gambar 4.21 Classification Report SVM Kernel Linear Skenario 3.....	54
Gambar 4.22 Confusion Matrix SVM Kernel Linear Skenario 4.....	55
Gambar 4.23 Classification Report SVM Kernel Linear Skenario 4.....	55
Gambar 4.24 Confusion Matrix SVM Kernel Linear Skenario 5.....	56
Gambar 4.25 Classification Report SVM Kernel Linear Skenario 5.....	56
Gambar 4.26 Confusion Matrix SVM Kernel Linear Skenario 6.....	57
Gambar 4.27 Classification Report SVM Kernel Linear Skenario 6.....	58
Gambar 4.28 Confusion Matrix SVM Kernel RBF Skenario 1	59

Gambar 4.29 Classification Report SVM Kernel RBF Skenario 1	59
Gambar 4.30 Confusion Matrix SVM Kernel RBF Skenario 2.....	60
Gambar 4.31 Classification Report SVM Kernel RBF Skenario 2	60
Gambar 4.32 Confusion Matrix SVM Kernel RBF Skenario 3.....	61
Gambar 4.33 Classification Report SVM Kernel RBF Skenario 3	62
Gambar 4.34 Confusion Matrix SVM Kernel RBF Skenario 4.....	63
Gambar 4.35 Classification Report SVM Kernel RBF Skenario 4	63
Gambar 4.36 Confusion Matrix SVM Kernel RBF Skenario 5.....	64
Gambar 4.37 Classification Report SVM Kernel RBF Skenario 5	64
Gambar 4.38 Confusion Matrix SVM Kernel RBF Skenario 6.....	65
Gambar 4.39 Classification Report SVM Kernel RBF Skenario 6	66
Gambar 4.40 Confusion Matrix SVM Kernel RBF Skenario 7.....	66
Gambar 4.41 Classification Report SVM Kernel RBF Skenario 7	67
Gambar 4.42 Confusion Matrix SVM Kernel RBF Skenario 8.....	68
Gambar 4.43 Classification Report SVM Kernel RBF Skenario 8	68
Gambar 4.44 Confusion Matrix SVM Kernel RBF Skenario 9.....	69
Gambar 4.45 Classification Report SVM Kernel RBF Skenario 9	70
Gambar 4.46 Confusion Matrix SVM Kernel RBF Skenario 10.....	71
Gambar 4.47 Classification Report SVM Kernel RBF Skenario 10	71
Gambar 4.48 Confusion Matrix SVM Kernel RBF Skenario 11.....	72
Gambar 4.49 Classification Report SVM Kernel RBF Skenario 11	72
Gambar 4.50 Confusion Matrix SVM Kernel RBF Skenario 12.....	73
Gambar 4.51 Classification Report SVM Kernel RBF Skenario 12	74
Gambar 4.52 Perbandingan Hasil Evaluasi SVM Kernel Linear & RBF	78

DAFTAR TABEL

Tabel 3.1 Contoh Tahapan <i>Case Folding</i>	14
Tabel 3.2 Contoh Tahapan Data <i>Cleaning</i>	15
Tabel 3.3 Contoh Tahapan <i>Tokenizing</i>	15
Tabel 3.4 Contoh Proses <i>Normalization</i>	16
Tabel 3.5 Contoh Tahapan <i>Stop Word Removal</i>	16
Tabel 3.6 Contoh Tahapan <i>Stemming</i>	17
Tabel 3.7 Contoh Kalimat Data Latih.....	18
Tabel 3.8 Contoh Kalimat Data Uji	19
Tabel 3.9 Frekuensi Kata dan Document Frequency (DF)	19
Tabel 3.10 TF dan IDF	20
Tabel 3.11 Hasil TF-IDF	21
Tabel 3.12 Skenario Pengujian	30
Tabel 4.1 Spesifikasi Perangkat Keras	32
Tabel 4.2 Hasil Prediksi	46
Tabel 4.3 Hasil Prediksi	49
Tabel 4.4 Hasil Pengujian Model	75