

**ANALISIS SENTIMEN PADA DATA REVIEW PENGGUNA APLIKASI
NOVEL ONLINE (WATTPAD) di GOOGLE PLAY STORE MENGGUNAKAN
METODE SUPPORT VECTOR MACHINE**

SKRIPSI

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



disusun oleh

SINTIYA APRILI YANTI

18.11.2440

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA**

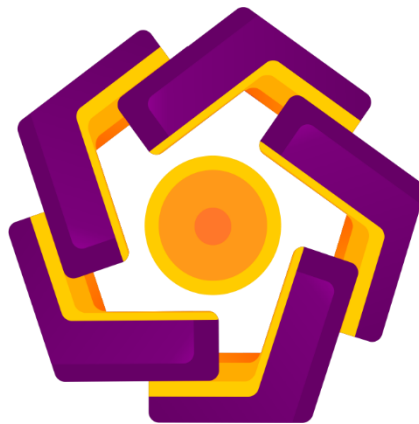
YOGYAKARTA

2023

**ANALISIS SENTIMEN PADA DATA REVIEW PENGGUNA APLIKASI
NOVEL ONLINE (WATTPAD) DI GOOGLE PLAY STORE MENGGUNAKAN
METODE SUPPORT VECTOR MACHINE**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

SINTIYA APRILI YANTI

18.11.2440

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PADA DATA REVIEW PENGGUNA APLIKASI
NOVEL ONLINE (WATTPAD) di GOOGLE PLAY STORE
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Sintiya Aprili Yanti

18.11.2440

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 10 Agustus 2023

Dosen Pembimbing,



Lilis Dwi Farida, S.Kom, M.Eng

NIK. 190302288

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN PADA DATA REVIEW PENGGUNA APLIKASI
NOVEL ONLINE (WATTPAD) di GOOGLE PLAY STORE
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Sintiya Aprili Yanti

18.11.2440

Telah dipertahankan di depan Dewan Penguji
pada tanggal 10 Agustus 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Anna Baita, M.Kom
NIK. 190302290

Donni Prabowo, M.Kom
NIK. 190302253

Lilis Dwi Farida, S.Kom, M.Eng
NIK. 190302288



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 10 Agustus 2023

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Sintiya Aprili Yanti

NIM : 18.11.2440

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Pada Data Review Pengguna Aplikasi Novel Online (Wattpad) di Google Play Store Menggunakan Support Vector Machine

Dosen Pembimbing : Lilis Dwi Farida, 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, 10 Agustus 2023

Yang Menyatakan,



Sintiya Aprili Yanti

HALAMAN PERSEMBAHAN

Dengan rasa syukur yang tak terhingga, penulis ingin mengucapkan penghargaan dan rasa terima kasih kepada semua pihak yang telah turut serta mendukung dan membantu dalam penyelesaian skripsi ini. Segala bantuan, dorongan, dan inspirasi yang diberikan memiliki peran yang sangat penting dalam perjalanan penulisan skripsi ini. Berikut ini adalah ucapan terima kasih penulis:

1. Kepada keluarga yang selalu memberi semangat dan doa dalam setiap langkah perjalanan. Terima kasih kepada orang tua, saudara, dan keluarga yang telah memberikan cinta, dukungan moral, dan doa selama penulis menjalani proses penelitian dan penulisan skripsi ini.
2. Kepada dosen pembimbing Ibu Lilis Dwi Farida, S.Kom., M.Eng., kepada ibu Lilis Dwi Farida terima kasih atas bimbingan, arahan, dan masukan yang berharga sepanjang perjalanan skripsi ini. Dengan kesabaran dan pengetahuan ibu, penulis mendapatkan wawasan yang berharga untuk mengembangkan penelitian ini.
3. Kepada Wahid Rizka Fathurrohman, terima kasih wahid atas diskusi, ide, dan pandangan yang telah membantu penulis dalam merumuskan dan mengembangkan gagasan dalam skripsi ini.
4. Kepada Teman-teman dekat, terima kasih kepada teman-teman KUNCUP, teman-teman kost-an TLUKI 18 (Mitha, Aul, Egi dan Jupita), Teman-teman 18-S1IF-09 atas semangat, dukungan, dan momen berharga yang telah kita bagi bersama selama masa perkuliahan dan penulisan skripsi ini.
5. Kepada Suga dan BTS, terima kasih atas dukungan moral dan mental dalam memberikan semangat tambahan dalam perjalanan skripsi ini.
6. Kepada Andreas Arif Budiman, terima kasih karna selalu mendukung penulis dalam mengerjakan skripsi, memberikan semangat serta membantu penulis lebih disiplin dalam waktu mengerjakan skripsi.
7. Kepada seluruh pihak yang tidak dapat disebutkan satu-persatu. Terima kasih sudah mendukung penulis dengan tulus

KATA PENGANTAR

Dengan rasa rendah hati dan penuh syukur, penulis ingin menyampaikan kata pengantar sebagai ungkapan terima kasih dan penghargaan atas perjalanan panjang dalam menyelesaikan skripsi ini yang berjudul " Analisis Sentimen Pada Data Review Pengguna Aplikasi Novel Online (Wattpad) di Google Play Store Menggunakan Metode Support Vector Machine ". Skripsi ini merupakan bagian integral dari perjalanan akademis penulis sebagai mahasiswa Informatika di Universitas Amikom Yogyakarta. Namun penelitian ini tidak akan selesai tanpa orang-orang yang mendukung penulis secara tulus. Terima kasih penulis sampaikan kepada:

1. Kepada Tuhan Yesus Kristus, atas berkat dan karunia-Nya telah diberikan kepada penulis sehingga penulis dapat menyelesaikan penelitian ini hingga selesai.
2. Bapak Prof. Dr. M. Suyanto, MM. Selaku Rektor Universitas Amikom Yogyakarta.
3. Dosen pembimbing Ibu Lilis Dwi Farida, S.Kom, M.Eng , yang telah memberi bimbingan, arahan, dan masukan yang berharga sepanjang perjalanan skripsi ini.
4. Keluarga dan kedua orang tua penulis, yang telah mendukung dan mendoakan penulis tanpa henti.
5. Para Dosen dan Staf Universitas Amikom Yogyakarta yang telah berperan dalam memberikan ilmu pengetahuan dan pengalaman selama masa perkuliahan.
6. Serta seluruh pihak yang telah membantu penulis, mohon maaf tidak dapat penulis sebutkan satu persatu.

Yogyakarta, <10 Agustus 2023>

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.....	3
1.3 Maksud dan Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	6
2.1 Tinjauan Pustaka.....	6
2.2 Dasar Teori	11
2.1.1 Data Mining	11
2.1.2 Pembelajaran Mesin (Machine Learning).....	11
2.1.3 Analisis Sentimen (Sentimen Analysis).....	13
2.1.4 Wattpad	14
2.1.5 Flowchart	15
2.1.6 Term Frequency Inverse Document Frequency (TF-IDF).....	17

2.1.7	Synthetic Minority Oversampling Technique (SMOTE)	18
2.1.8	Support Vector Machine (SVM)	19
2.1.9	Confusion Matrix	21
BAB III METODE PENELITIAN		24
3.1	Metode yang Digunakan	24
3.2	Alat Penelitian	24
3.1.1	Perangkat Keras (Hardware)	24
3.1.2	Perangkat Lunak (Software)	24
3.3	Alur Penelitian	24
3.4	Pengumpulan Data	25
3.5	Seleksi Data (Data Selection)	26
3.6	Pra-pemrosesan Data (Data Preprocessing)	27
3.6.1.	Pembersihan Teks	29
3.6.2.	Case Folding	29
3.6.3.	Tokenisasi	29
3.6.4.	Formalisasi	30
3.6.5.	Penghapusan Stop Words	30
3.6.6.	Stemming atau Lemmatisasi	30
3.6.7.	Rekonstruksi Kalimat (Detokenization)	30
3.7	Transformasi Data	31
3.7.1.	Vektorisasi Data	32
3.7.2.	Oversampling Data	39
3.8	Klasifikasi Support Vector Machine	39
3.8.1	Pencarian Hyperplane	39
3.9	Rancangan Antarmuka	43
3.9.1	Halaman Analisis Sentimen Ulasan	43
3.9.2	Halaman untuk Melatih Model SVM	44

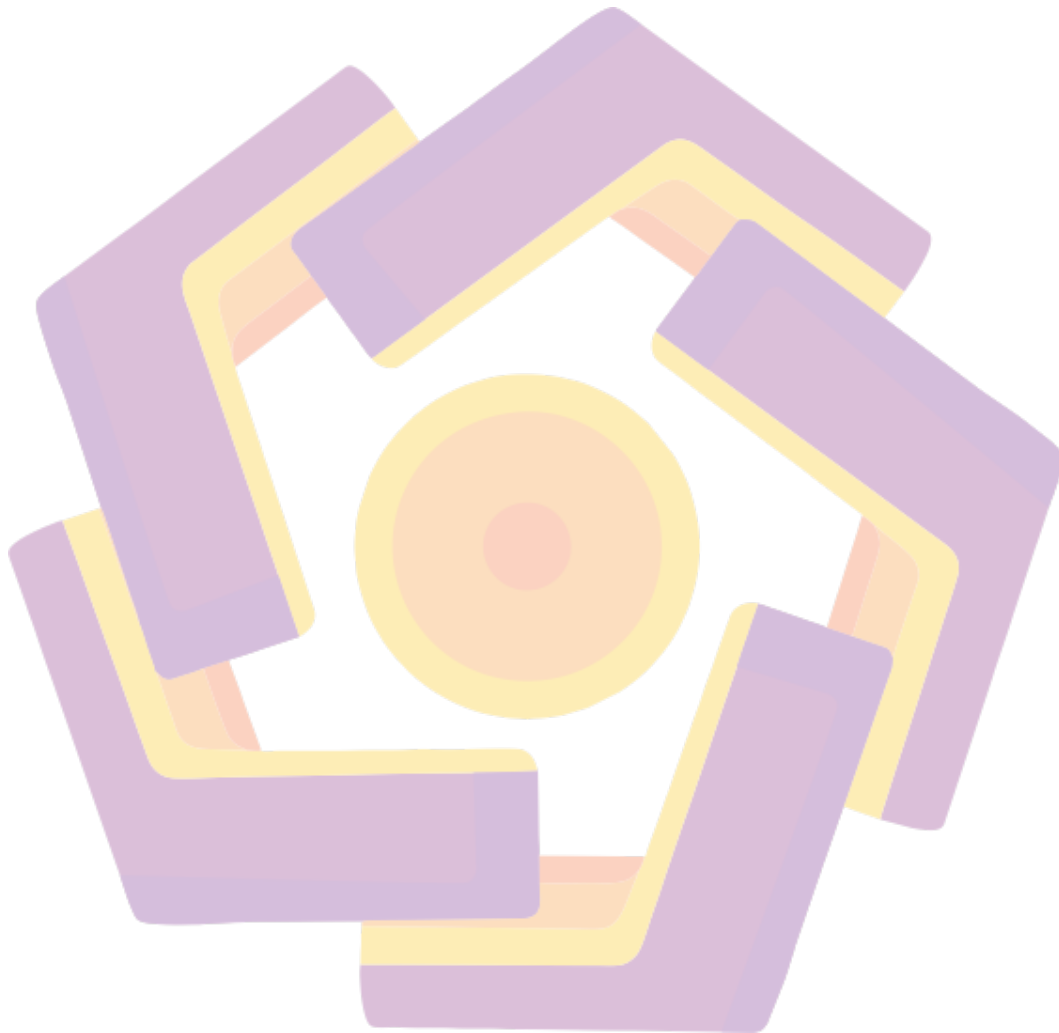
BAB IV HASIL DAN PEMBAHASAN	45
4.1 Pengumpulan Data	45
4.2 Seleksi Data	47
4.3 Pelabelan Data	47
4.4 Pra-pemrosesan Data	48
4.3.1 Pembersihan Teks	48
4.3.2 Case Folding	48
4.3.3 Tokenisasi	49
4.3.4 Formalisasi	50
4.3.5 Penghapusan Stop Words	51
4.3.6 Lemmatisasi atau Stemming	52
4.3.7 Rekonstruksi Kalimat (Detokenisasi)	53
4.5 Transformasi Data	54
4.5.1 Vektorisasi Data	54
4.5.2 Penyeimbangan Data dengan Oversampling	55
4.6 Pembagian Data	56
4.7 Pembelajaran Model Klasifikasi	56
4.8 Evaluasi	57
BAB V KESIMPULAN DAN SARAN	59
5.1 Kesimpulan	59
5.2 Saran	59
Daftar Pustaka	60

DAFTAR TABEL

Tabel 2. 1 Perbandingan Penelitian	9
Tabel 2. 2 Simbol-Simbol Pada Flowchart	16
Tabel 2. 3 Tabel Klasifikasi	22
Table 3. 1 Informasi Atribut Data Ulasan Wattpad	26
Table 3. 2 Sampel Data Teks Ulasan	32
Table 3. 3 Teks Ulasan Setelah Tokenisasi oleh TF-IDF	33
Table 3. 4 Frekuensi Kemunculan Kata Dalam Setiap Dokumen Positif	33
Table 3. 5 Frekuensi Kemunculan Kata Dalam Setiap Dokumen Negatif	34
Table 3. 6 Perhitungan Smooth IDF	35
Table 3. 7 Perhitungan TF-IDF (Dokumen Positif) Sebelum Dilakukan Normalisasi	36
Table 3. 8 Perhitungan TF-IDF (Dokumen Negatif) Sebelum Dilakukan Normalisasi	36
Table 3. 9 TF-IDF (Dokumen Positif) Setelah Dilakukan Normalisasi L2	37
Table 3. 10 TF-IDF (Dokumen Negatif) Setelah Dilakukan Normalisasi L2	38
Table 3. 11 Hasil Akhir TF-IDF	38
Table 3. 12 Klasifikasi SVM	41
Table 3. 13 Confusion Matrix	41
Table 3. 14 Accuracy, Precision, Recall Dan F-Score	42
Table 4. 1 Lima Baris Pertama Ulasan Wattpad	47
Table 4. 2 Data Dengan Label	47
Table 4. 3 Teks Ulasan Setelah Proses Pembersihan dan Case Folding	48
Table 4. 4 Teks Ulasan Setelah Proses Tokenisasi	49
Table 4. 5 Teks Ulasan Setelah Proses Formalisasi	51
Table 4. 6 Teks Ulasan Setelah Penghapusan Stop Words	51
Table 4. 7 Teks Ulasan Setelah Proses Lemmatisasi	52
Table 4. 8 Teks Ulasan Setelah Proses Rekonstruksi Kalimat	53
Table 4. 9 Perbandingan Data Sebelum dan Sesudah Proses Oversampling	55
Table 4. 10 Hasil evaluasi dengan Confusion Matrix	57

Table 4. 11 Hasil perhitungan Akurasi, Presisi, Sensitivitas (Recall), dan F-Score.....57

Table 4. 12 Hasil perhitungan skor pelatihan menggunakan 5 fold validasi silang57



DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi Cara Kerja Teknik Oversampling.....	18
Gambar 2. 2 Pencarian Hyperlane pada SVM.....	20
Gambar 2. 3 Pencarian Hyperlane Terbaik pada SVM.....	20
Gambar 3. 1 Alur Penelitian	25
Gambar 3. 2 Alur Pra-pemrosesan Data	28
Gambar 3. 3 Alur Transformasi Data.....	31
Gambar 3. 4 Ilustrasi Oversampling Data Menggunakan SMOTE	39
Gambar 3. 5 Rancangan Halaman Klasifikasi Ulasan	43
Gambar 3. 6 Rancangan tampilan halaman untuk melatih model SVM	44
Gambar 4. 1 Baris Kode Pengunduhan Data Ulasan Wattpad	45
Gambar 4. 2 Baris kode pengunduhan data ulasan Wattpad 1	45
Gambar 4. 3 Baris Kode Untuk Melakukan Tokenisasi	49
Gambar 4. 4 Contoh Kamus Kata Baku.....	50
Gambar 4. 5 Baris kode Untuk Melakukan Formalisasi	50
Gambar 4. 6 Baris Kode Untuk Penghapusan Stop Words.....	51
Gambar 4. 7 Baris Kode Untuk Lemmatisasi	52
Gambar 4. 8 Baris Kode Untuk Rekonstruksi Kalimat	53
Gambar 4. 9 Baris Kode Untuk Vektorisasi Data.....	54
Gambar 4. 10 Baris Kode Untuk Penyeimbangan Data	55
Gambar 4. 11 Baris Kode Untuk Penyeimbangan Data	56
Gambar 4. 12 Baris Kode untuk Pelatihan Model SVM.....	56
Gambar 4. 13 Hasil perhitungan skor pelatihan menggunakan 5 fold validasi silang	58

DAFTAR LAMBANG DAN SINGKATAN

CSV	:	Comma Separated Values
DF	:	Document Frequency
FN	:	False Negative
FP	:	False Positive
SVM	:	Support Vector Machines
TF	:	Term Frequency
TN	:	True Negative
TP	:	True Positive
TF-IDF	:	Term Frequency-Inverse Document Frequency
IDF	:	Inverse Document frequency
SMOTE	:	Synthetic Minority Over-sampling Technique

INTISARI

Penelitian ini bertujuan untuk menganalisis sentimen dari data review pengguna Aplikasi Novel Online (Wattpad) yang diunggah di Google Play Store. Metode Support Vector Machine (SVM) digunakan sebagai pendekatan dalam menganalisis sentimen pada dataset yang terdiri dari ulasan dan penilaian yang diberikan oleh pengguna aplikasi Wattpad. Langkah-langkah penelitian meliputi pengumpulan dataset review pengguna dari Google Play Store, pra-pemrosesan data untuk membersihkan dan mengubah teks ulasan ke dalam bentuk representasi numerik, serta pembentukan model klasifikasi SVM untuk mengklasifikasikan sentimen dari ulasan tersebut menjadi positif, negatif, atau netral. Penelitian ini menggunakan 1500 data dengan berfokus ke Bahasa Indonesia dan menghasilkan tingkat evaluasi akurasi 85,95%, presisi 84,11%, Sensitivitas 87,80%, dan F-score 85,92%.

Kata kunci: Analisis Sentimen, Data Review Pengguna, Aplikasi Novel Online, Wattpad, Google Play Store, Metode Support Vector Machine.

ABSTRACT

This research aims to analyze the sentiment of user reviews on the Online Novel Application (Wattpad) uploaded on Google Play Store. The Support Vector Machine (SVM) method is employed as an approach to analyze sentiment in a dataset consisting of reviews and ratings given by Wattpad application users. The research steps include collecting user review datasets from Google Play Store, preprocessing the data to clean and convert the text reviews into a numerical representation, and forming an SVM classification model to categorize the sentiment of these reviews into positive, negative, or neutral. This study uses 1500 data with a focus on Indonesian and produces an evaluation level of accuracy of 85.95%, precision of 84.11%, sensitivity of 87.80%, and F-score of 85.92%.

Keyword: *Sentiment Analysis, User Review Data, Online Novel Application, Wattpad, Google Play Store, Support Vector Machine Method.*