

**ANALISIS SENTIMEN ULASAN PRODUK AMAZON
MENGUNAKAN ALGORITMA SUPPORT VECTOR
MACHINE (SVM)**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

BAYU AHMAD NUGROHO

21.12.2229

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN ULASAN PRODUK AMAZON
MENGUNAKAN ALGORITMA SUPPORT VECTOR
MACHINE (SVM)**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

BAYU AHMAD NUGROHO

21.12.2229

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS SENTIMEN ULASAN PRODUK AMAZON

MENGGUNAKAN ALGORITMA SUPPORT VECTOR

MACHINE (SVM)

yang disusun dan diajukan oleh

Bayu Ahmad Nugroho

21.12.2229

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Juli 2025

Dosen Pembimbing,


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

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN ULASAN PRODUK AMAZON
MENGUNAKAN ALGORITMA SUPPORT VECTOR
MACHINE (SVM)**

yang disusun dan diajukan oleh

Bayu Ahmad Nugroho

21.12.2229

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302282

Ikma, S.Kom., M.Kom.
NIK. 190302282

Agung Nugroho, S.Kom., M.Kom.
NIK. 190302242

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusri, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Bayu Ahmad Nugroho
NIM : 21.12.2229

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN ULASAN PRODUK AMAZON
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE
(SVM)**

Dosen Pembimbing : Hanif Al Fatta, S.Kom., M.Kom., Ph.D

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, 21 Juli 2025

Yang Menyatakan,

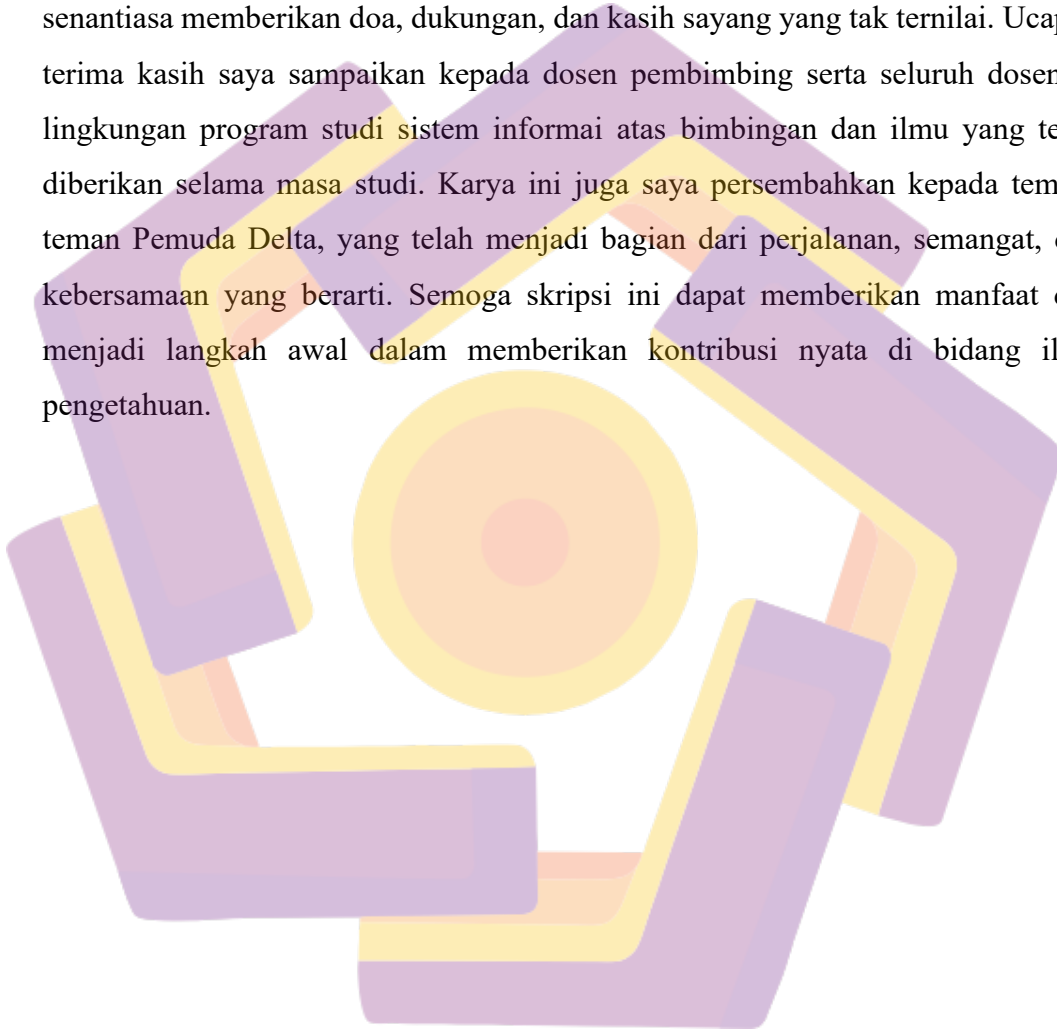


The image shows a red rectangular meter stamp with the text 'METERAN TEMPEL' and a serial number '27AMX424442557'. A handwritten signature in black ink is written over the stamp.

Bayu Ahmad Nugroho

HALAMAN PERSEMBAHAN

Dengan rasa syukur kepada Allah SWT atas segala limpahan rahmat dan karunia-Nya, skripsi ini saya persembahkan kepada kedua orang tua tercinta yang senantiasa memberikan doa, dukungan, dan kasih sayang yang tak ternilai. Ucapan terima kasih saya sampaikan kepada dosen pembimbing serta seluruh dosen di lingkungan program studi sistem informasi atas bimbingan dan ilmu yang telah diberikan selama masa studi. Karya ini juga saya persembahkan kepada teman-teman Pemuda Delta, yang telah menjadi bagian dari perjalanan, semangat, dan kebersamaan yang berarti. Semoga skripsi ini dapat memberikan manfaat dan menjadi langkah awal dalam memberikan kontribusi nyata di bidang ilmu pengetahuan.



KATA PENGANTAR

Puji syukur saya panjatkan ke hadirat Allah SWT atas limpahan rahmat, taufik, dan hidayah-Nya, sehingga saya dapat menyelesaikan skripsi yang berjudul Analisis Sentimen Ulasan Produk Amazon Menggunakan Algoritma Support Vector Mechine (SVM) sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Amikom Yogyakarta. Dalam proses penyusunan skripsi ini, saya mendapat banyak dukungan dan bimbingan dari berbagai pihak. Oleh karena itu, saya menyampaikan terima kasih kepada Bapak Hanif Al Fatta, S.Kom., M.Kom., Ph.D, selaku dosen pembimbing, serta Tim Dosen Penguji atas masukan dan arahan yang diberikan. Saya juga mengucapkan terima kasih kepada kedua orang tua tercinta atas doa dan dukungannya, serta kepada teman-teman Pemuda Delta dan rekan-rekan yang telah memberi semangat selama proses ini. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan menjadi kontribusi bagi pengembangan ilmu pengetahuan.

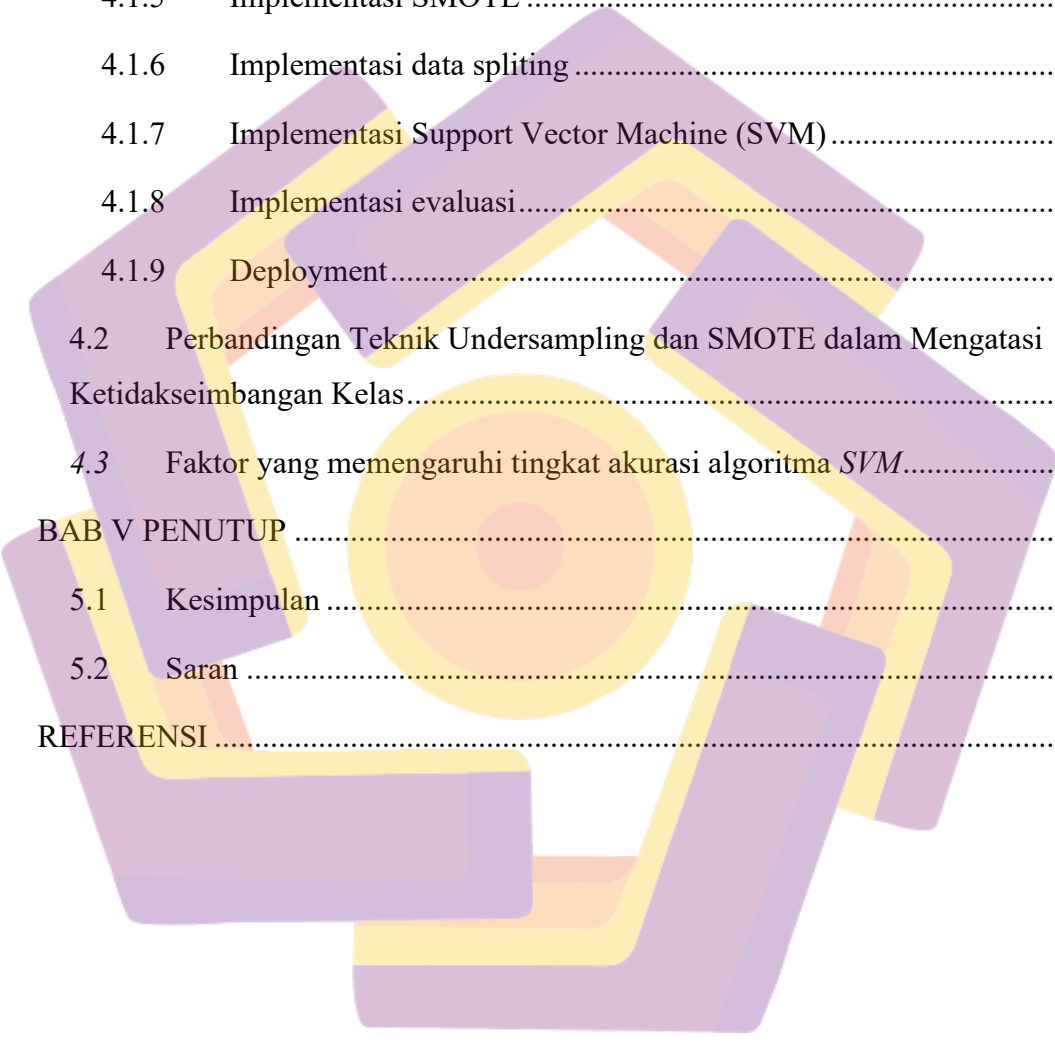
Yogyakarta, 21 Juli 2025

Penulis

DAFTAR ISI

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.....	xi
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
<i>ABSTRACT</i>	xvi
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
1.5.1 Manfaat Bagi Objek Penelitian	4
1.5.2 Manfaat Bagi Peneliti Selanjutnya.....	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6

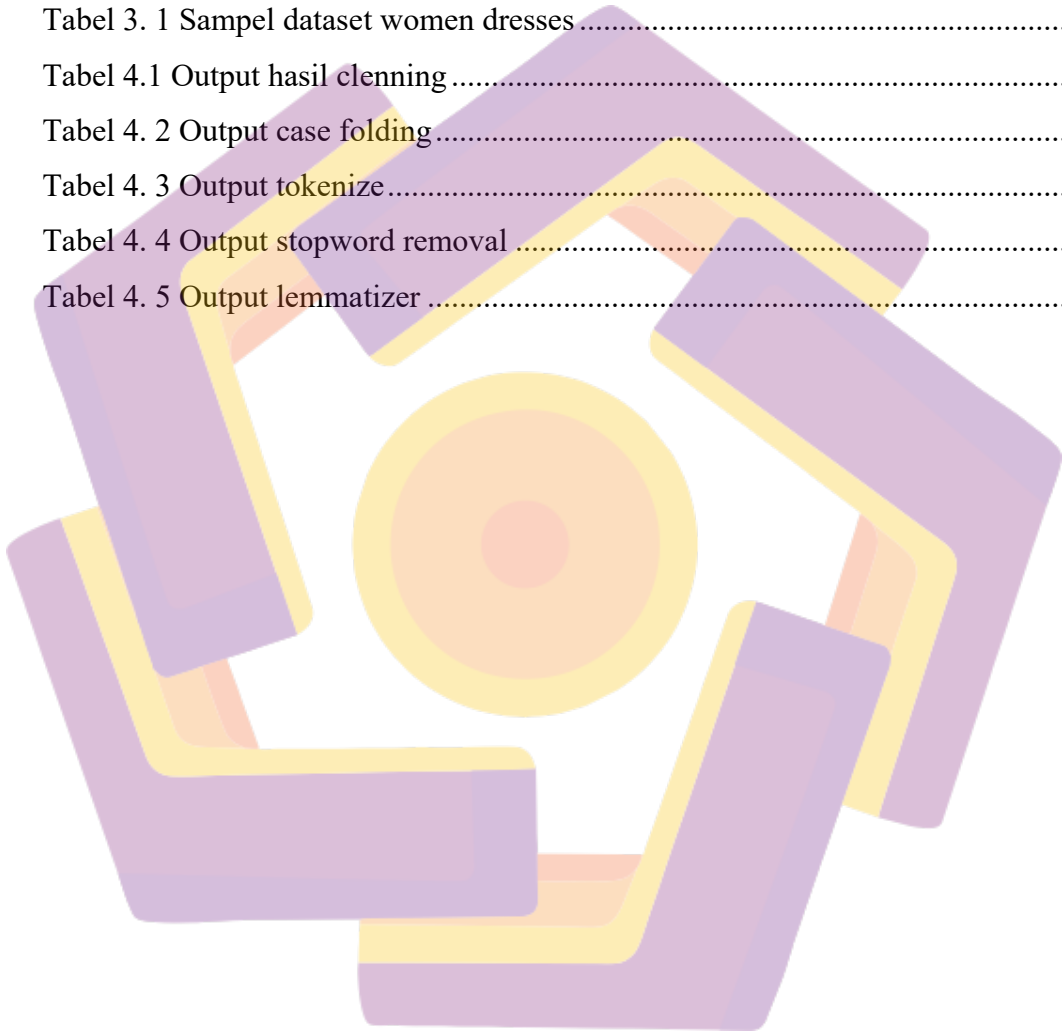
2.1	Studi Literatur	6
2.2	Dasar Teori.....	15
2.2.1	Analisis Sentimen	15
2.2.2	E-commerce	16
2.2.3	Amazon	16
2.2.4	Support Vector Machine (SVM).....	16
2.2.5	Pemrosesan Data Teks (Text Preprocessing).....	18
2.2.6	TF-IDF (Term Frequency-Inverse Document Frequency)	19
2.2.7	Evaluasi Model	20
BAB III METODE PENELITIAN		23
3.1	Objek Penelitian.....	23
3.2	Alur Penelitian	23
3.2.1	Pengumpulan Data	25
3.2.2	Preprocessing Data.....	25
3.3.3	Labeling Data.....	26
3.3.4	TF-IDF	26
3.3.5	Smote	26
3.3.6	Data Splitting	27
3.3.7	Pemodelan SVM	27
3.3.8	Evaluasi.....	27
3.3.9	`Deployment	28
3.3	Alat dan Bahan.....	28
3.3.1	Data Penelitian	28
BAB IV HASIL DAN PEMBAHASAN		32
4.1	Implementasi algoritma Support Vector Machine (SVM).....	32



4.1.1	Implementasi pengumpulan data	32
4.1.2	Implementasi preprocessing Data	33
4.1.3	Implementasi labeling Data	44
4.1.4	Implementasi TF-IDF	48
4.1.5	Implementasi SMOTE	51
4.1.6	Implementasi data splitting	52
4.1.7	Implementasi Support Vector Machine (SVM)	55
4.1.8	Implementasi evaluasi.....	56
4.1.9	Deployment.....	62
4.2	Perbandingan Teknik Undersampling dan SMOTE dalam Mengatasi Ketidakseimbangan Kelas.....	68
4.3	Faktor yang memengaruhi tingkat akurasi algoritma SVM.....	70
BAB V PENUTUP		71
5.1	Kesimpulan	71
5.2	Saran	71
REFERENSI		73

DAFTAR TABEL

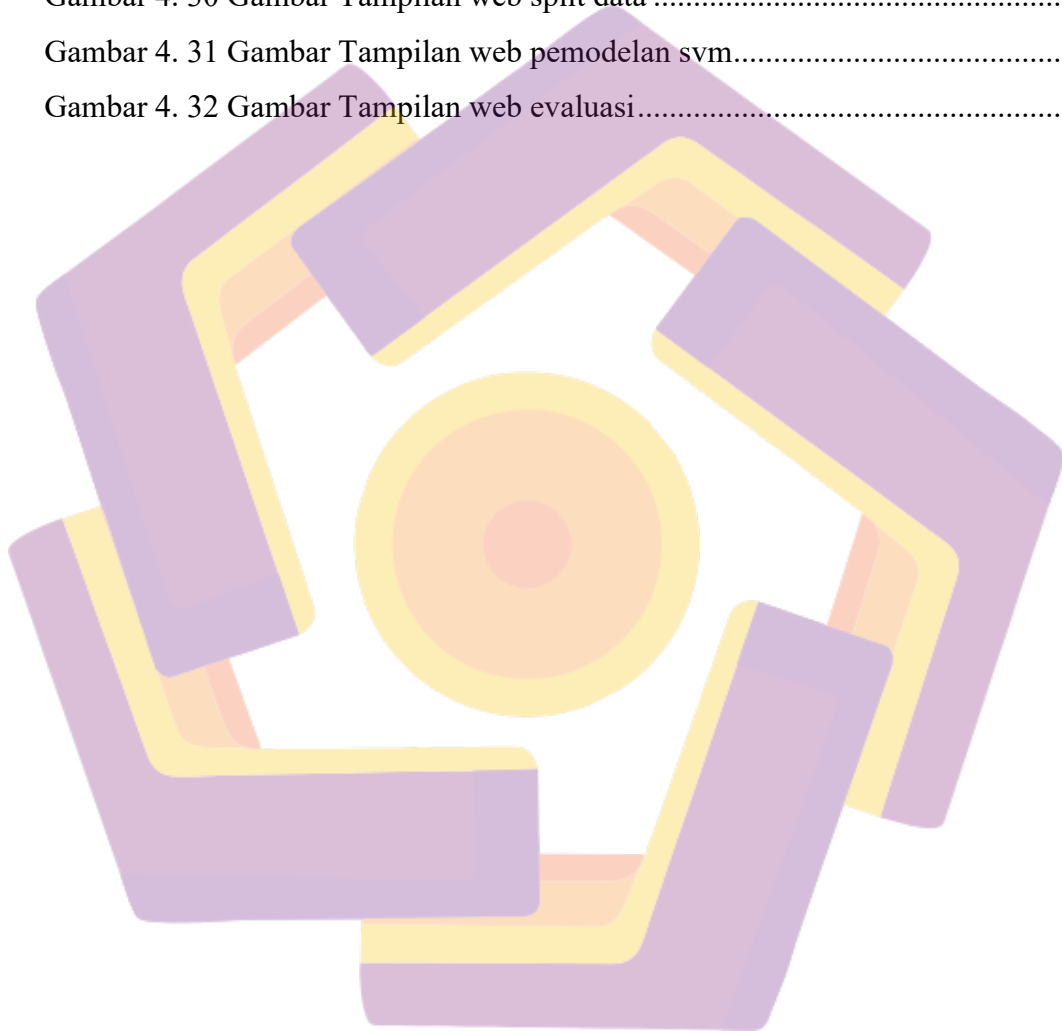
Tabel 2. 1 Keaslian Penelitian	10
Tabel 2. 2 Confusion Matrix untuk dua kelas	20
Tabel 3. 1 Sampel dataset women dresses	28
Tabel 4.1 Output hasil clenning	35
Tabel 4. 2 Output case folding	37
Tabel 4. 3 Output tokenize	39
Tabel 4. 4 Output stopword removal	41
Tabel 4. 5 Output lemmatizer	43



DAFTAR GAMBAR

Gambar 2. 1 analisis sentimen	15
Gambar 2. 2 Model SVM.....	17
Gambar 3. 1 Alur Penelitian	24
Gambar 4. 1 Source code upluod file csv	32
Gambar 4. 2 Source code filter dataset	33
Gambar 4. 3Source code clenning	34
Gambar 4. 4 Source code case folding.....	36
Gambar 4. 5 Source code case tokenize.....	38
Gambar 4. 6 Source code case mengimpor stopwords	40
Gambar 4. 7 Source code case stopwords.....	41
Gambar 4. 8 Source code case mengimpor WordNetLemmatizer.....	42
Gambar 4. 9 Source code case Lemmatizer	43
Gambar 4. 10 Source code case labeling	45
Gambar 4. 11 Source code case visualisasi labeling.....	46
Gambar 4. 12 Output visualisasi labeling	47
Gambar 4. 13 Source code vektorisasi tf-idf.....	49
Gambar 4. 14 Bentuk matriks tf-idf.....	49
Gambar 4. 15 Source code smote.....	51
Gambar 4. 16 Hasil imbalnceing smote	52
Gambar 4. 17 Source code data spliting	52
Gambar 4. 18 Hasil data spliting.....	53
Gambar 4. 19 Source code visualisasi data spliting.....	54
Gambar 4. 20 Hasil visualisasi data spliting.....	55
Gambar 4. 21 Source code pemodelan svm.....	55
Gambar 4. 22 Source code evaluasi	56
Gambar 4. 23 Hasil heatmap confusion matrix	57
Gambar 4. 24 Hasil classification report.....	58

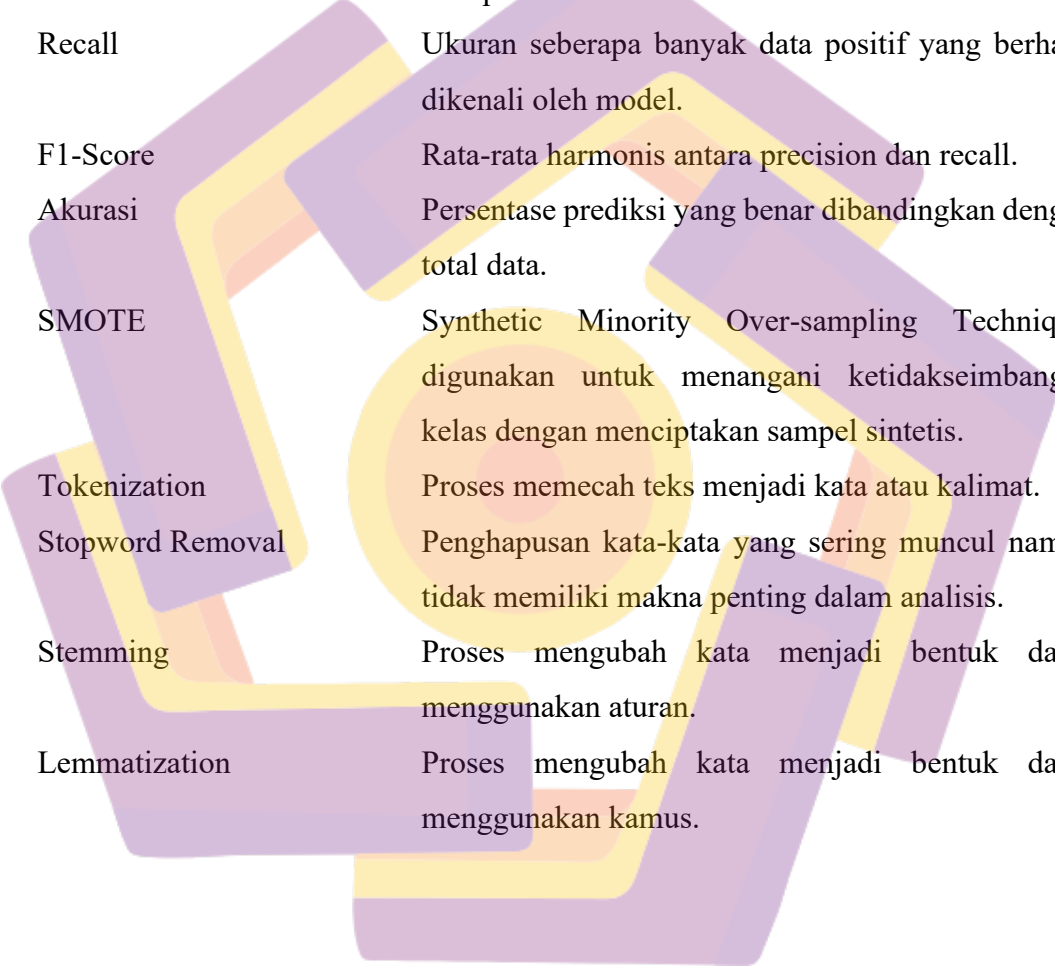
Gambar 4. 25 Hasil akurasi.....	61
Gambar 4. 26 Gambar Tampilan web analisis sentiment teks.....	63
Gambar 4. 27 Gambar Tampilan web analisis dataset.....	63
Gambar 4. 28 Gambar Tampilan web preview data	63
Gambar 4. 29 Gambar Tampilan web tf-idf.....	64
Gambar 4. 30 Gambar Tampilan web split data	65
Gambar 4. 31 Gambar Tampilan web pemodelan svm.....	66
Gambar 4. 32 Gambar Tampilan web evaluasi.....	67



DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machine
TF-IDF	Term Frequency-Inverse Document Frequency
TP	True Positive
FP	False Positive
TN	True Negative
FN	False Negative
SMOTE	Synthetic Minority Over-sampling Technique
POS	Part of Speech

DAFTAR ISTILAH



Confusion Matrix	Matriks yang menunjukkan jumlah prediksi yang benar dan salah untuk setiap kelas dalam klasifikasi.
Precision	Ukuran seberapa tepat model mengklasifikasikan data positif.
Recall	Ukuran seberapa banyak data positif yang berhasil dikenali oleh model.
F1-Score	Rata-rata harmonis antara precision dan recall.
Akurasi	Persentase prediksi yang benar dibandingkan dengan total data.
SMOTE	Synthetic Minority Over-sampling Technique, digunakan untuk menangani ketidakseimbangan kelas dengan menciptakan sampel sintetis.
Tokenization	Proses memecah teks menjadi kata atau kalimat.
Stopword Removal	Penghapusan kata-kata yang sering muncul namun tidak memiliki makna penting dalam analisis.
Stemming	Proses mengubah kata menjadi bentuk dasar menggunakan aturan.
Lemmatization	Proses mengubah kata menjadi bentuk dasar menggunakan kamus.

INTISARI

Penelitian ini bertujuan untuk menganalisis sentimen ulasan produk *women dresses* di Amazon menggunakan algoritma *Support Vector Machine (SVM)*. Mengingat banyaknya ulasan produk, analisis manual menjadi tidak efisien, sehingga penelitian ini menggunakan SVM untuk mengklasifikasikan ulasan menjadi sentimen positif, negatif, atau netral. Data ulasan diperoleh dari dataset *Kaggle* yang berisi 23.486 ulasan produk. Proses *preprocessing* dilakukan dengan membersihkan teks, menghapus kata-kata tidak relevan (*stopwords*), serta melakukan *tokenisasi* dan *lemmatization*. Fitur utama diekstraksi menggunakan *TF-IDF* untuk mengubah teks menjadi representasi numerik. Hasil evaluasi model *SVM* menunjukkan akurasi 83% dalam mengklasifikasikan sentimen. Penelitian ini juga menemukan bahwa penggunaan *SMOTE* untuk menangani ketidakseimbangan kelas meningkatkan kinerja model. Temuan ini memberikan kontribusi bagi sistem analisis sentimen di platform *e-commerce* dan meningkatkan pengalaman belanja konsumen.

Kata kunci: Analisis Sentimen, Support Vector Machine (SVM), TF-IDF, SMOTE, E-commerce.

ABSTRACT

This study aims to analyze the sentiment of product reviews for women's dresses on Amazon using the Support Vector Machine (SVM) algorithm. Given the large number of product reviews, manual analysis becomes inefficient, so this research utilizes SVM to classify reviews into positive, negative, or neutral sentiments. The review data is obtained from a Kaggle dataset containing 23,486 product reviews. The preprocessing process involves cleaning the text, removing irrelevant words (stopwords), and performing tokenization and lemmatization. Key features are extracted using TF-IDF to convert the text into numerical representations. The evaluation results of the SVM model show an accuracy of 83% in classifying sentiment. The study also found that the use of SMOTE to address class imbalance improved the model's performance. These findings contribute to sentiment analysis systems on e-commerce platforms and enhance consumer shopping experiences.

Keyword: Sentiment Analysis, Support Vector Machine (SVM), TF-IDF, SMOTE, E-commerce.