

**ANALISIS SENTIMEN TERHADAP FITUR SHOPEE LIVE  
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**



**SKRIPSI**

**Ferdy Willy  
00000043525**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA  
TANGERANG  
2024**

**ANALISIS SENTIMEN TERHADAP FITUR SHOPEE LIVE  
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**



Diajukan sebagai salah satu syarat untuk memperoleh  
Gelar Sarjana Komputer (S.Kom.)

**Ferdy Willy**  
**00000043525**

**UMMN**

**UNIVERSITAS  
MULTIMEDIA  
NUSANTARA**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA**

**TANGERANG**

**2024**

## HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Ferdy Willy  
NIM : 00000043525  
Program Studi : Informatika

Menyatakan dengan sesungguhnya bahwa Tesis/Skripsi/Tugas Akhir/Laporan Magang/MBKM saya yang berjudul:

**Analisis Sentimen Terhadap Fitur Shopee Live Menggunakan Algoritma Support Vector Machine**

merupakan hasil karya saya sendiri, bukan merupakan hasil plagiat, dan tidak pula dituliskan oleh orang lain; Semua sumber, baik yang dikutip maupun dirujuk, telah saya cantumkan dan nyatakan dengan benar pada bagian Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan karya ilmiah, saya bersedia menerima konsekuensi untuk dinyatakan TIDAK LULUS. Saya juga bersedia menanggung segala konsekuensi hukum yang berkaitan dengan tindak plagiarisme ini sebagai kesalahan saya pribadi dan bukan tanggung jawab Universitas Multimedia Nusantara.

Tangerang, 22 Mei 2024



(Ferdy Willy)

**HALAMAN PENGESAHAN**

Skripsi dengan judul

**ANALISIS SENTIMEN TERHADAP FITUR SHOPEE LIVE  
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**

oleh

Nama : Ferdy Willy  
NIM : 00000043525  
Program Studi : Informatika  
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Rabu, 29 Mei 2024

Pukul 13.00 s/s 15.00 dan dinyatakan

**LULUS**

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Eunike Endariahna Surbakti, S.Kom.,  
M.T.I.)

NIDN: 0322099401

Penguji



(Sy Yuliani Yakub, S.Kom., M.T. PhD)

NIDN: 0411037904

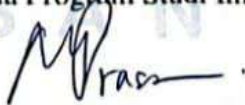
Pembimbing



(Suwito Pomalingo, S.Kom., M.Kom.)

NIDN: 0911098201

Pj Ketua Program Studi Informatika,



(Dr. Eng. Niki Prastomo, S.T., M.Sc.)

NIDN: 0419128203

## HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini:

Nama : Ferdy Willy  
NIM : 00000043525  
Program Studi : Informatika  
Jenjang : S1  
Jenis Karya : Skripsi

Menyatakan dengan sesungguhnya bahwa:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia karena dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)<sup>\*\*</sup>.

Tangerang, 22 Mei 2024

Yang menyatakan



Ferdy Willy

\*\* Jika tidak bisa membuktikan LoA jurnal/HKI selama enam bulan ke depan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

**Halaman Persembahan / Motto**

"A good name is to be more desired than great wealth, Favor is better than silver and gold."

Proverbs 22:1 (NASB)



**UMMN**  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Analisis Sentimen Terhadap Fitur Shopee Live Menggunakan Algoritma Support Vector Machine dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Pjs. Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Suwito Pomalingo, S.Kom., M.Kom., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 22 Mei 2024



Ferdy Willy

# ANALISIS SENTIMEN TERHADAP FITUR SHOPEE LIVE MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE

Ferdy Willy

## ABSTRAK

Dengan perkembangan pengguna internet di Indonesia yang sangat pesat, istilah *ecommerce* sudah tidak asing lagi di masyarakat. Sehingga pertumbuhan *ecommerce* di Indonesia juga sangat tinggi. Shopee merupakan salah satu platform *ecommerce* yang terus mengembangkan aplikasinya. Salah satu fitur yang sedang tren di Indonesia adalah *shopee live*. Fitur ini memungkinkan pembeli untuk berinteraksi langsung dengan penjual sebelum membeli produknya. Fitur ini tentunya mendapatkan berbagai macam ulasan dari penggunanya. Ulasan ini berguna sebagai masukan kepada pengembang dan juga dapat membantu pengguna dalam memilih aplikasi yang tepat. Ulasan yang digunakan pada penelitian ini didapatkan dari ulasan di *playstore* pada aplikasi *shopee*. Penelitian ini bertujuan untuk membuat sebuah model untuk melakukan analisis sentimen terhadap fitur *shopee live*. Analisis sentimen dilakukan menggunakan metode *support vector machine*, dengan mengumpulkan ulasan dari *playstore*. Kemudian akan menggunakan *confusion matrix* untuk mendapatkan nilai *accuracy*, *precision*, *recall*, dan *f1-score* dari model yang sudah dibuat. Hasil dari penelitian, didapatkan algoritma *support vector machine* memiliki nilai *accuracy* 78%, *precision* 70%, *recall* 69%, dan *f1-score* 69%. Selain itu, dari 1011 ulasan yang telah dikumpulkan, didapatkan 58% data memiliki sentimen positif, 31% data memiliki sentimen negatif, dan 11% data memiliki sentimen netral.

**Kata kunci:** *Analisis Sentimen, Confusion Matrix, Review, Shopee Live, Support Vector Machine*

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A



***Sentiment Analysis of Shopee Live Features Using Support Vector Machine  
Algorithm***

Ferdy Willy

***ABSTRACT***

*With the rapid growth of internet users in Indonesia, the term e-commerce has become familiar to society. Consequently, the growth of e-commerce in Indonesia is also very high. Shopee is one of the e-commerce platforms that continues to develop its application. One trending feature in Indonesia is Shopee Live. This feature allows buyers to interact directly with sellers before purchasing their products. This feature receives various reviews from its users. These reviews are useful as feedback to developers and can also help users choose the right application. The reviews used in this study were obtained from the Play Store reviews of the Shopee application. This research aims to create a model for sentiment analysis of the Shopee Live feature. Sentiment analysis is performed using the support vector machine method by collecting reviews from the Play Store. Then, a confusion matrix will be used to obtain the accuracy, precision, recall, and f1-score values of the created model. The results of the research, it was found that the support vector machine algorithm had an accuracy value of 78%, precision of 70%, recall of 69%, and f1-score of 69%. Furthermore, out of 1011 reviews collected, 58% of the data have positive sentiment, 31% have negative sentiment, and 11% have neutral sentiment.*

***Keywords:*** *Confusion Matrix, Review, Sentiment Analysis, Shopee Live, Support Vector Machine*



## DAFTAR ISI

HALAMAN JUDUL . . . . .	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT . . . . .	ii
HALAMAN PENGESAHAN . . . . .	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH . . . . .	iv
HALAMAN PERSEMBAHAN/MOTO . . . . .	v
KATA PENGANTAR . . . . .	vi
ABSTRAK . . . . .	vii
ABSTRACT . . . . .	viii
DAFTAR ISI . . . . .	ix
DAFTAR GAMBAR . . . . .	xi
DAFTAR TABEL . . . . .	xii
DAFTAR KODE . . . . .	xiii
DAFTAR LAMPIRAN . . . . .	xiv
BAB 1 PENDAHULUAN . . . . .	1
1.1 Latar Belakang Masalah . . . . .	1
1.2 Rumusan Masalah . . . . .	3
1.3 Batasan Permasalahan . . . . .	3
1.4 Tujuan Penelitian . . . . .	3
1.5 Manfaat Penelitian . . . . .	4
1.6 Sistematika Penulisan . . . . .	4
BAB 2 LANDASAN TEORI . . . . .	6
2.1 Analisis Sentimen . . . . .	6
2.2 Shopee . . . . .	6
2.3 Text Preprocessing . . . . .	6
2.4 Support Vector Machine . . . . .	7
2.5 TF-IDF . . . . .	9
2.6 <i>Confusion Matrix</i> . . . . .	10
2.7 <i>Synthetic Minority Oversampling Technique (SMOTE)</i> . . . . .	11
BAB 3 METODOLOGI PENELITIAN . . . . .	12
3.1 Alur Penelitian . . . . .	12
3.1.1 Studi Literatur . . . . .	12
3.1.2 Pengumpulan Data . . . . .	13
3.1.3 Perancangan Sistem . . . . .	13
3.1.4 Implementasi . . . . .	13
3.1.5 Pengujian Model dan Evaluasi . . . . .	13
3.1.6 Pembuatan Laporan . . . . .	14
3.2 Perancangan Sistem . . . . .	14
3.2.1 Scraping Data . . . . .	14
3.2.2 Text Preprocessing . . . . .	15
3.2.3 Labelling Data . . . . .	16
3.2.4 Split Data . . . . .	17
3.2.5 Penerapan TF-IDF . . . . .	17
3.2.6 Penerapan SMOTE . . . . .	17
3.2.7 Penerapan Model dengan SVM . . . . .	17
3.2.8 Pengujian dan Evaluasi Model . . . . .	18
BAB 4 HASIL DAN DISKUSI . . . . .	19
4.1 Spesifikasi Sistem . . . . .	19
4.2 Implementasi Sistem . . . . .	19

4.2.1	Pengumpulan Data . . . . .	19
4.2.2	Filtering Data . . . . .	20
4.2.3	Text Preprocessing . . . . .	22
4.2.4	Labelling Data . . . . .	37
4.2.5	Word Cloud . . . . .	40
4.2.6	Split data . . . . .	42
4.2.7	Penerapan TF-IDF . . . . .	43
4.2.8	Penerapan SMOTE . . . . .	43
4.2.9	Penerapan SVM . . . . .	44
4.2.10	Evaluasi Model . . . . .	44
4.3	Uji Coba . . . . .	45
4.3.1	Pengujian dengan Perbandingan Data 60:40 Tanpa SMOTE . . . . .	45
4.3.2	Pengujian dengan Perbandingan Data 60:40 Dengan SMOTE . . . . .	46
4.3.3	Pengujian dengan perbandingan data 70:30 Tanpa SMOTE . . . . .	47
4.3.4	Pengujian dengan perbandingan data 70:30 Dengan SMOTE . . . . .	48
4.3.5	Pengujian dengan perbandingan data 80:20 Tanpa SMOTE . . . . .	49
4.3.6	Pengujian dengan perbandingan data 80:20 Dengan SMOTE . . . . .	50
4.3.7	Evaluasi Hasil Pengujian . . . . .	51
BAB 5	SIMPULAN DAN SARAN . . . . .	53
5.1	Simpulan . . . . .	53
5.2	Saran . . . . .	53
	DAFTAR PUSTAKA . . . . .	54



## DAFTAR GAMBAR

Gambar 2.1	<i>Support Vector Machine</i> . . . . .	8
Gambar 2.2	Tabel <i>Confusion Matrix</i> . . . . .	10
Gambar 3.1	<i>Flowchart</i> alur penelitian . . . . .	12
Gambar 3.2	<i>Flowchart</i> perancangan sistem . . . . .	14
Gambar 3.3	<i>Flowchart Text Preprocessing</i> . . . . .	15
Gambar 4.1	Wordcloud Label Positif . . . . .	41
Gambar 4.2	Wordcloud Label Negatif . . . . .	41
Gambar 4.3	Wordcloud Label Netral . . . . .	42
Gambar 4.4	Hasil <i>Confusion Matrix</i> . . . . .	45
Gambar 4.5	Hasil <i>Confusion Matrix</i> untuk rasio 60:40 Tanpa SMOTE .	46
Gambar 4.6	Hasil <i>Classification Report</i> untuk Rasio 60:40 Tanpa SMOTE . . . . .	46
Gambar 4.7	Hasil <i>Confusion Matrix</i> untuk rasio 60:40 Dengan SMOTE	47
Gambar 4.8	Hasil <i>Classification Report</i> untuk Rasio 60:40 Dengan SMOTE . . . . .	47
Gambar 4.9	Hasil <i>Confusion Matrix</i> untuk rasio 70:30 Tanpa SMOTE .	48
Gambar 4.10	Hasil <i>Classification Report</i> untuk Rasio 70:30 Tanpa SMOTE . . . . .	48
Gambar 4.11	Hasil <i>Confusion Matrix</i> untuk rasio 70:30 Dengan SMOTE	49
Gambar 4.12	Hasil <i>Classification Report</i> untuk Rasio 70:30 Dengan SMOTE . . . . .	49
Gambar 4.13	Hasil <i>Confusion Matrix</i> untuk rasio 80:20 Tanpa SMOTE .	50
Gambar 4.14	Hasil <i>Classification Report</i> untuk Rasio 80:20 Tanpa SMOTE . . . . .	50
Gambar 4.15	Hasil <i>Confusion Matrix</i> untuk rasio 80:20 Dengan SMOTE	51
Gambar 4.16	Hasil <i>Classification Report</i> untuk Rasio 80:20 Dengan SMOTE . . . . .	51
Gambar 4.17	Perbandingan Nilai Akurasi, Presisi, Recall, dan F1-score berdasarkan Setiap Skenario . . . . .	52

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A

## DAFTAR TABEL

Tabel 1.1	Perbandingan Algoritma	5
Tabel 4.1	Data Hasil filtering	21
Tabel 4.1	Data Hasil filtering (lanjutan)	22
Tabel 4.2	Hasil Proses <i>Cleaning</i>	24
Tabel 4.2	Hasil Proses <i>Cleaning</i> (lanjutan)	25
Tabel 4.3	Hasil Proses <i>Case Folding</i>	26
Tabel 4.3	Hasil Proses <i>Case Folding</i> (lanjutan)	27
Tabel 4.4	Hasil Proses <i>Tokenizing</i>	28
Tabel 4.4	Hasil Proses <i>Tokenizing</i> (lanjutan)	29
Tabel 4.5	Hasil Proses Normalisasi	31
Tabel 4.5	Hasil Proses Normalisasi (lanjutan)	32
Tabel 4.6	Hasil Proses <i>Stopword Removal</i>	33
Tabel 4.6	Hasil Proses <i>Stopword Removal</i> (lanjutan)	34
Tabel 4.7	Hasil Proses <i>Stemming</i>	36
Tabel 4.8	Hasil Proses <i>Labelling</i>	39
Tabel 4.8	Hasil Proses <i>Labelling</i> (lanjutan)	40
Tabel 4.9	Perbandingan Nilai Akurasi, Presisi, Recall, dan F1-score berdasarkan Setiap Skenario	52

UMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

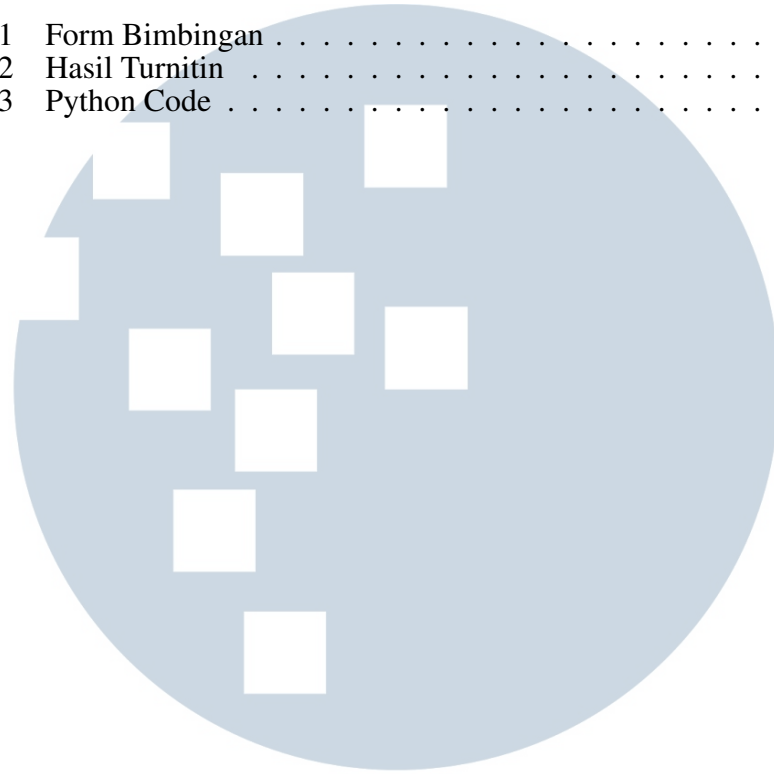
## DAFTAR KODE

4.1	Potongan Kode Scraping Data	20
4.2	<i>Pseudocode Filtering</i>	20
4.3	<i>Pseudocode Cleaning</i>	23
4.4	<i>Pseudocode Case Folding</i>	25
4.5	<i>Pseudocode Tokenizing</i>	27
4.6	<i>Pseudocode Normalisasi</i>	30
4.7	<i>Pseudocode Stopword Removal</i>	32
4.8	<i>Pseudocode Stemming</i>	35
4.9	<i>Pseudocode Translate</i>	37
4.10	<i>Pseudocode VaderSentiment</i>	37
4.11	<i>Pseudocode Labelling</i>	38
4.12	<i>Pseudocode Split Data</i>	42
4.13	<i>Pseudocode Penerapan TFIDF</i>	43
4.14	<i>Pseudocode Penerapan SMOTE</i>	43
4.15	<i>Pseudocode Penerapan SVM</i>	44
4.16	<i>Pseudocode Evaluasi Model</i>	44
5.1	Potongan Kode Scraping Data	68
5.2	Potongan Kode Filtering	68
5.3	Potongan Kode Cleaning Data	68
5.4	Potongan Kode Casefolding	69
5.5	Potongan Kode Tokenizing	69
5.6	Potongan Kode Normalisasi	69
5.7	Potongan Kode Stopword Removal	70
5.8	Potongan Kode Stemming	70
5.9	Potongan Kode Translate	70
5.10	Potongan Kode Vadersentiment	71
5.11	Potongan Kode Labelling	71
5.12	Potongan Kode Wordcloud	71
5.13	Potongan Kode Split Data	72
5.14	Potongan Kode Penerapan TFIDF	72
5.15	Potongan Kode Penerapan SMOTE	72
5.16	Potongan Kode Penerapan SVM	72
5.17	Potongan Kode Confusion Matrix label	73

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A

## DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan . . . . .	57
Lampiran 2	Hasil Turnitin . . . . .	58
Lampiran 3	Python Code . . . . .	68



**UMMN**  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA