



**ANALISIS SENTIMEN MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE*
BERDASARKAN ULASAN PENGGUNA TERKAIT FAKTOR-FAKTOR
KELEBIHAN DAN KEKURANGAN FITUR APLIKASI SEGARI**

SKRIPSI

**FARHAN HABIB
2010512124**

**S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
PROGRAM STUDI SISTEM INFORMASI PROGRAM SARJANA
2024**



**ANALISIS SENTIMEN MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE*
BERDASARKAN ULASAN PENGGUNA TERKAIT FAKTOR-FAKTOR
KELEBIHAN DAN KEKURANGAN FITUR APLIKASI SEGARI**

SKRIPSI

Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Komputer

FARHAN HABIB

2010512124

**S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
PROGRAM STUDI SISTEM INFORMASI PROGRAM SARJANA
2024**

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Farhan Habib

NIM : 2010512124

Tanggal : 22 Juli 2024

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses dengan ketentuan berlaku

Jakarta, 22 Juli 2024

Yang Menyatakan,


Farhan Habib

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional "Veteran" Jakarta, saya yang bertanda tangan di bawah ini.

Nama : Farhan Habib

NIM : 2010512124

Fakultas : Ilmu Komputer

Program Studi : S1 – Sistem Informasi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional "Veteran" Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Analisis Sentimen Menggunakan Metode *Support Vector Machine* Berdasarkan Ulasan Pengguna Terkait Faktor-Faktor Kelebihan dan Kekurangan Fitur Aplikasi Aplikasi Segari

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional "Veteran" Jakarta berhak menyimpan, mengalih data/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenar-benarnya.

Dibuat di : Jakarta

Pada tanggal : 22 Juli 2024

Yang Menyatakan,



Farhan Habib

LEMBAR PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Farhan Habib

NIM : 2010512124

Program Studi : S1 – Sistem Informasi

Judul Tugas Akhir : Analisis Sentimen Menggunakan Metode *Support Vector Machine* Berdasarkan Ulasan Pengguna Terkait Faktor-Faktor Kelebihan dan Kekurangan Fitur Aplikasi Aplikasi Segari

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.



Tri Rahayu, S.Kom., M.M.

Pembimbing 1



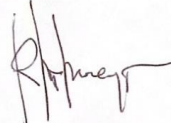
Muhammad Panji Muslim, S.Pd., M.Kom.

Pembimbing 2



Ika Nurlaiti Isnainivah, S.Kom., M.Sc.

Penguji 1



Rudhy Ho Purabaya, S.E., M.M.S.I.

Penguji 2



Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM

Dekan Fakultas Ilmu Komputer



Anita Muliawati, S.Kom., M.T.I.

Ketua Program Studi S1 Sistem Informasi

Ditetapkan di : Jakarta

Tanggal Ujian : 5 Juli 2024

ANALISIS SENTIMEN MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE* BERDASARKAN ULASAN PENGGUNA TERKAIT FAKTOR-FAKTOR KELEBIHAN DAN KEKURANGAN FITUR APLIKASI SEGARI

Farhan Habib

ABSTRAK

Latar belakang penelitian ini adalah banyak perusahaan *e-groceries* mengalami kesulitan bertahan di pasar pasca pandemi. Contoh nyata adalah HappyFresh yang menghentikan layanannya karena kesulitan finansial dan TaniHub yang menutup layanan karena kesulitan mempertahankan profitabilitas. Namun, Segari berhasil bertahan dan mengalami pertumbuhan bisnis positif. Penelitian ini bertujuan untuk memahami faktor-faktor yang membuat aplikasi Segari mampu bertahan hingga saat ini, mengidentifikasi keunggulan dan kekurangannya, serta mengukur seberapa besar sentimen pengguna terhadap penggunaan aplikasi tersebut melalui analisis sentimen. Manfaatnya adalah Mengetahui faktor-faktor yang disukai dan yang perlu ditingkatkan pada aplikasi Segari. Metode penelitian yang digunakan adalah metode klasifikasi *support vector machine* dan kernel *linear* dengan 5437 data ulasan Segari dari Google Play Store dan App Store. Hasil penelitian menunjukkan komentar kelas positif dominan pada ulasan aplikasi Segari, sebanyak 5107 dari total data dan sisanya kelas negatif sebanyak 330 data ulasan. Kesimpulan dari penelitian ini adalah bahwa topik yang paling sering dibahas untuk sentimen positif mencakup kesegaran produk, kecepatan proses pemesanan dan pengiriman, promo menarik, dan kualitas tinggi, sedangkan sentimen negatif mencakup masalah teknis dan layanan, pengiriman dan kurir, serta pengalaman pengguna. Disarankan untuk Segari bisa meningkatkan layanannya dari sentimen negatif yang diperoleh. Performa model yang dibuat dengan menggunakan metode *support vector machine* dibantu dengan kernel *linear* mendapatkan hasil akurasi sebesar 96 persen.

Kata Kunci: Analisis sentimen, *Support Vector Machine*, *Kernel Linear*, *Segari*, *E-groceries*

**SENTIMENT ANALYSIS UTILIZING SUPPORT VECTOR MACHINE METHOD
BASED ON USER REVIEWS RELATED TO FACTORS OF ADVANTAGES AND
DISADVANTAGES OF SEGARI APPLICATION FEATURES**

Farhan Habib

ABSTRACT

The background of this research is that many e-groceries companies have difficulty surviving in the post-pandemic market. Obvious examples are HappyFresh which discontinued its services due to financial difficulties and TaniHub which closed its services due to difficulties maintaining profitability. However, Segari managed to survive and experience positive business growth. This research aims to understand the factors that make the Segari application able to survive until now, identify its advantages and disadvantages, and measure how much user sentiment towards using the application through sentiment analysis. The benefit is knowing the factors that are favored and which need to be improved on the Segari application. The research method used is the support vector machine and linear kernel classification method with 5437 Segari review data from the Google Play Store and App Store. The results showed that positive class comments were dominant in Segari application reviews, as much as 5107 of the total data and the remaining negative class as much as 330 review data. The conclusion of this study is that the most frequently discussed topics for positive sentiments include product freshness, speed of the ordering and delivery process, attractive promos, and high quality, while negative sentiments include technical and service issues, shipping and couriers, and user experience. It is suggested that Segari can improve its services from the negative sentiments obtained. The performance of the model created using the support vector machine method assisted by a linear kernel obtained an accuracy result of 96 percent.

Keywords: *Sentimen Analysis, Support Vector Machine, Kernel Linear, Segari, E-groceries*

KATA PENGANTAR

Puji dan syukur dipanjatkan kepada Tuhan Yang Maha Esa atas rahmat dan karunianya sehingga Proposal Tugas Akhir ini dapat diselesaikan dengan baik. Proposal Tugas Akhir ini dibuat guna memenuhi persyaratan untuk menyusun Tugas Akhir/Skripsi.

Dalam penyusunan proposal ini hingga selesai, terdapat banyak bantuan dari banyak pihak. Demikian, segala rasa hormat dan terima kasih disampaikan setinggi-tingginya kepada.

1. Allah SWT. atas segala Rahmat, Ridho, dan Karunia-Nya.
2. Kedua orang tua yang selalu memberikan dukungan baik moril maupun materil.
3. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM, selaku Dekan Fakultas Ilmu Komputer UPN “Veteran” Jakarta.
4. Ibu Anita Muliawati, S.Kom. M.T.I., selaku Ketua Program Studi S1 Sistem Informasi.
5. Ibu Tri Rahayu, S.Kom, M.M, selaku Dosen Pembimbing 1.
6. Teman-teman jurusan di Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta angkatan 2020.
7. Keluarga, kerabat, teman dan semua pihak yang tidak dapat saya sebutkan satu persatu yang telah mendukung dalam penyelesaian Proposal Skripsi ini.

Masih banyak kekurangan yang terdapat di Proposal Tugas Akhir ini, baik dari segi isi maupun penulisan, mengingat kurangnya pengetahuan dan pengalaman serta hakikat sebagai manusia yang selalu salah. Oleh karena itu, saran dan kritik yang membangun akan sangat berarti.

Jakarta, 15 Maret 2024

Farhan Habib

DAFTAR ISI

PERNYATAAN ORISINALITAS	i
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xi
BAB I.....	1
PENDAHULUAN	1
4.1 Latar belakang	1
4.2 Rumusan Masalah	4
4.3 Ruang Lingkup.....	4
4.4 Tujuan Penelitian.....	4
4.5 Manfaat Penelitian.....	4
4.6 Luaran yang Diharapkan	5
4.7 Sistematika Penulisan.....	5
BAB II.....	6
TINJAUAN PUSTAKA	6
2.1 Analisis Sentimen.....	6
2.2 <i>Text Mining</i>	6
2.3 Segari.....	7
2.4 <i>Text Preprocessing</i>	7
2.4.1 <i>Case Folding</i>	8
2.4.2 <i>Data Cleaning</i>	8
2.4.3 <i>Stemming</i>	8
2.4.4 <i>Normalization</i>	9
2.4.5 <i>Stopword</i>	9

2.5	<i>Indonesian sentiment Lexicon</i>	10
2.6	Pembobotan Kata	10
2.6.1.	TF (<i>Term Frequency</i>).....	10
2.6.2.	IDF (<i>Inverse Document Frequency</i>)	10
2.6.3.	TF-IDF	11
2.7	<i>Support Vector Machine</i>	11
2.8	Python.....	13
2.9	Penelitian Terdahulu	14
BAB III		17
METODE PENELITIAN.....		17
3.1	Alur Penelitian.....	17
3.1.1	Identifikasi Masalah.....	18
3.1.2	Studi Literatur	19
3.1.3	<i>Data Scraping</i>	19
3.1.4	<i>Text Preprocessing</i>	19
3.1.5	<i>Labelling</i>	20
3.1.6	Pembobotan Kata.....	20
3.1.7	Model SVM	21
3.1.8	Hasil Evaluasi Model.....	21
3.1.9	Analisis Hasil Klasifikasi.....	21
3.2	Alat Penelitian	22
3.2.1	Perangkat Keras	22
3.2.2	Perangkat Lunak	22
BAB IV		23
HASIL DAN PEMBAHASAN.....		23
4.1	Pengumpulan Data	23
4.2	Pre-processing Data.....	24
4.2.1	<i>Case Folding</i>	24
4.2.2	<i>Data Cleaning</i>	25
4.2.3	<i>Stemming</i>	26
4.2.4	<i>Normalization</i>	27
4.2.5	<i>Stopword</i>	28
4.3	<i>Labelling</i>	29

4.4	Pembobotan Kata	30
4.5	Split Data.....	32
4.6	Klasifikasi dengan <i>Support Vector Machine</i> (SVM)	33
4.7	Hasil Evaluasi Model	34
4.8	<i>Labelling</i> dengan SVM	36
4.9	Hasil dan Visualisasi	37
4.10	Sentimen Positif	37
4.11	Sentimen Negatif.....	39
BAB V		41
PENUTUP		41
DAFTAR PUSTAKA		43
RIWAYAT HIDUP		47
LAMPIRAN		49

DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi Metode Support Vector Machine	12
Gambar 4. 1 Sintaks Implementasi Proses Labelling	29
Gambar 4. 2 Sintaks Implementasi Proses Pembobotan Kata	31
Gambar 4. 3 Sintaks Implementasi Split Data	32
Gambar 4. 4 Sintaks Klasifikasi SVM.....	33
Gambar 4. 5 Hasil Confusion Matrix.....	35
Gambar 4. 6 Sintaks Implementasi Proses Labelling Dengan Model yang Telah Dibuat.....	36
Gambar 4. 7 Persentase Dokumen Positif dan Negatif.....	37
Gambar 4. 8 Top 20 Kata Dalam Dokumen Positif.....	38
Gambar 4. 10 Top 20 Kata Dalam Dokumen Negatif	39

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	12
Tabel 2. 2 Penelitian Terdahulu	14
Tabel 4. 1 Kolom Data yang di pakai	23
Tabel 4. 2 Hasil Scrapping	23
Tabel 4. 3 Hasil Case Folding.....	24
Tabel 4. 4 Hasil Cleaning Data	25
Tabel 4. 5 Hasil Stemming.....	26
Tabel 4. 6 Hasil Normalization.....	27
Tabel 4. 7 Hasil Stopword	28
Tabel 4. 8 Hasil Pelabelan Dengan InSet Lexicon	29
Tabel 4. 9 Hasil Persebaran Data Tiap Kelas/Label	30
Tabel 4. 10 Hasil dari Pembobotan Kata	31
Tabel 4. 11 Banyaknya Data Uji dan Data Latih	32
Tabel 4. 12 Top 10 Koefisien Kata Positif	33
Tabel 4. 13 Top 10 Koefisien Kata Negatif.....	34
Tabel 4. 14 Hasil Evaluasi Model.....	36
Tabel 4. 15 Hasil Implementasi Labelling Dengan Model yang Telah Dibuat	36