



**KLASIFIKASI ULASAN PENGGUNA MENGGUNAKAN
METODE SUPPORT VECTOR MACHINE PADA APLIKASI
HALODOC**

SKRIPSI

Fachran Sandi

1810511049

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN
JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
2023**



**KLASIFIKASI ULASAN PENGGUNA MENGGUNAKAN
METODE SUPPORT VECTOR MACHINE PADA APLIKASI
HALODOC**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
Sarjana Komputer**

Fachran Sandi

1810511049

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN
JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
2023**

LEMBAR PERSETUJUAN

Dengan ini menyatakan bahwa skripsi/tugas akhir berikut:

Nama : Fachran Sandi
NIM : 1810511049
Program Studi : Informatika
Judul : Klasifikasi Ulasan Pengguna Menggunakan Metode Support Vector Machine Pada Aplikasi Halodoc

Sebagai bagian persyaratan yang diperlukan untuk mengikuti ujian Sidang Tugas Akhir/Skripsi pada Program Studi Informatika Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Menyetujui,
Dosen Pembimbing



Iin Ernawati, S.Kom., M.Si.

Mengetahui,
a.n Ketua Program Studi



Dr. Widya Cholil, S.Kom., M.I.T.

Ditetapkan : Jakarta,

Tanggal Persetujuan : 22 November 2021

Lembar Pengesahan

Dengan ini dinyatakan bahwa Skripsi berikut:

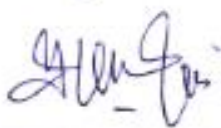
Nama : Fachran Sandi

NIM : 1810511049

Program Studi : S1 Informatika

Judul Skripsi : Klasifikas Ulasan Pengguna Menggunakan Metode Support Vector Machine Pada Aplikasi Halodoc

Telah berhasil dipertahankan dihadapan tim penguji dan diterima sebagai persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Yuni Widlastiwi, S.Kom., M.Si.

Penguji 1



Henki Bayu Seta, S.Kom., M.TI.

Penguji 2



Iin Ernawati, S.Kom., M.Si.

Pembimbing



Dr. Ermatita, M.Kom.

Dekan



Dr. Widya Cholil, S.Kom., M.I.T.

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Pengesahan : 4 Januari 2023



Pernyataan Orisinalitas

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

Nama : Fachran Sandi

NIM : 1810511049

Tanggal : 4 Januari 2023

Bila Mana di kemudian hari ditemukan ketidak sesuaian dengan pernyataan saya ini, maka saya bersedia di tuntutan dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 4 Januari 2023

Yang Menyatakan,



Fachran Sandi

Pernyataan Persetujuan Publikasi Skripsi Untuk Kepentingan Akademis

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan dibawah ini :

Nama : Fachran Sandi

NIM : 1810511049

Fakultas : Ilmu Komputer

Program Studi : Informatika

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:

Klasifikas Ulasan Pengguna Menggunakan Metode Support Vector Machine Pada Aplikasi Halodoc

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 4 Januari 2023

Yang menyatakan,



Fachran Sandi

KLASIFIKASI ULASAN PENGGUNA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE PADA APLIKASI HALODOC

Fachran Sandi

Abstraksi

Dengan adanya wabah covid-19, semua layanan beralih menjadi online. Aplikasi HaloDoc adalah aplikasi *telemedicine* yang populer di Indonesia yang menyediakan layanan Kesehatan secara online, aplikasi HaloDoc diperlukan adanya perbaikan untuk mengurangi kekurangannya dalam menyediakan informasi kepada pengguna. Analisis sentimen dapat melakukan klasifikasi pengguna, Pada penelitian ini akan menggunakan 2 kelas yaitu kelas sentimen yaitu positif dan negatif, serta kelas kategori, dimana kelas kategori diambil dari atribut ISO 9126, ISO 9126 adalah standar perangkat lunak yang dibuat oleh ISO dan IEC sebagai standar pengukuran jaminan kualitas perangkat lunak. Penelitian ini akan dibuat 2 model, model dengan dua kelas untuk kelas sentimen dan model *multiclass* untuk kelas kategori, masing – masing model akan menggunakan algoritma *Support Vector Machine* dan dengan algoritma TF – IDF. Hasil dari klasifikasi terhadap aplikasi halodoc, untuk kelas sentimen didapatkan hasil sebesar 96,02% dengan kernel *linear*, dan hasil untuk kelas kategori menggunakan metode *one vs rest* dan kernel *sigmoid* didapatkan hasil sebesar 78,97%.

Kata Kunci: klasifikasi, *Support Vector Machine*, Halodoc, analisis sentimen

CLASSIFICATION OF USER REVIEWS USING THE SUPPORT VECTOR MACHINE METHOD IN THE HALODOC APPLICATION

Fachran Sandi

Abstraction

Due to the COVID-19 outbreak, all services have switched to online. The Halodoc application is a popular telemedicine application in Indonesia that provides online health services. The Halodoc application needs improvement to reduce its shortcomings in providing information to users. Sentiment analysis can carry out user classification. In this study, 2 classes will be used, namely the sentiment class, namely positive and negative, as well as the category class, where the category class is taken from the attributes of ISO 9126, which is a software standard made by ISO and IEC as a measurement standard for software quality assurance. This research will create two models: a model with two classes for the sentiment class and a multiclass model for the category class; each model will use the support vector machine algorithm and the TF-IDF algorithm. The results of the classification of the Halodoc application for the sentiment class obtained a result of 96.02% with a linear kernel, and the results for the class category used the one vs. rest method, and the results for the sigmoid kernel were obtained at 78.97%.

Keywords: classification, support vector machine, Halodoc, sentiment analysis.

Daftar Isi

Lembar Persetujuan.....	iii
Lembar Pengesahan	iv
Pernyataan Orisinalitas.....	v
Pernyataan Persetujuan Publikasi Skripsi Untuk Kepentingan Akademis	vi
Abstraksi	vii
Abstraction	viii
Daftar Isi.....	ix
Daftar Gambar.....	xiii
Daftar Tabel	xiv
Daftar Lampiran	xv
Daftar Simbol.....	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan Penelitian.....	2
1.4 Ruang Lingkup	2
1.5 Luaran yang diharapkan	2
1.6 Manfaat Peneilitan.....	3
1.7 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Ulasan / Review.....	4
2.2 Telemedicine	4
2.3 Software Quality.....	5
2.4 ISO/IEC 9126	5

2.5 Data Mining.....	7
2.6 Analisis Sentimen.....	9
2.7 Text Mining.....	9
2.8 Pra Proses Data.....	10
2.9 Pembobotan Kata	10
2.10 Klasifikasi.....	11
2.11 SVM	12
2.11.1 Menemukan hyperplane pemisah yang optimal	12
2.11.2 Klasifikasi non-Linier dan fungsi kernel	15
2.11.3 Multiclass SVM Klasifikasi.....	17
2.12 Evaluasi Model.....	19
2.13 Penelitian Terdahulu.....	21
BAB III METODELOGI PENELITIAN	24
3.1 Tahapan Penelitian	24
3.1.1 Identifikasi Masalah.....	25
3.1.2 Studi Literatur	25
3.1.3 Data.....	25
3.1.4 Pelabelan Data	25
3.1.5 Pra Proses Data	25
3.1.6 Pembobotan TF-IDF	26
3.1.7 Klasifikasi	26
3.1.8 Evaluasi.....	27
3.1.9 Visualisasi	27
3.2 Alat yang digunakan.....	27
3.3 Jadwal Penelitian	28
BAB IV HASIL DAN PEMBAHASAN	30

4.1 Data	30
4.2 Pelabelan Data	31
4.3 Pra Proses Data	34
4.3.1 Case Folding	34
4.3.2 Pembersihan Data	36
4.3.3 <i>Normalization</i>	37
4.3.4 <i>Stemming</i>	38
4.3.5 <i>StopwordRemoval</i>	40
4.3.6 <i>Tokenization</i>	41
4.4 Pembobotan TF-IDF	42
4.5 Klasifikasi (SVM)	48
4.5.1 Klasifikasi Sentimen	55
4.5.2 Klasifikasi Kategori	56
4.6 Evaluasi	56
4.6.1 Evaluasi Model Sentimen	57
4.6.2 Evaluasi Model Kategori	58
4.7 Visualisasi	60
4.7.1 Sentimen Positif	60
4.7.2 Sentimen Negatif	61
4.7.3 Kategori Functionality	61
4.7.4 Kategori Reliability	62
4.7.5 Kategori Usability	63
4.7.6 Kategori Efficiency	63
4.7.7 Kategori Maintainability	64
4.7.8 Kategori Portability	65
Bab V PENUTUP	66

5.1 Kesimpulan.....	66
5.2 Saran.....	67
Daftar Pustaka.....	68
Riwayat Hidup.....	69
Lampiran.....	70

Daftar Gambar

Gambar 1 ISO 9126 quality model untuk external dan internal quality.....	6
Gambar 2 Tahapan KDD.....	8
Gambar 3 Beberapa hyperlane yang dapat salah membagi kelas.....	13
Gambar 4 Optimal hyperplane menggunakan algoritma SVM.....	13
Gambar 5 SVM 2 kelas yang menggambarkan hyperplane dan support vector .	14
Gambar 6 Ilustrasi transformasi dari 1D ke 2D.....	15
Gambar 7 Ilustrasi transformasi dari 2D ke 3D.....	16
Gambar 8 Contoh masalah label 3 kelas menggunakan Decision Directed Acyclic Graph.....	18
Gambar 9 Tahapan Penelitian.....	24
Gambar 10 Grafik Hasil Pelabelan kelas Sentimen.....	33
Gambar 11 Hasil Pelabelan untuk Kelas Kategori	34
Gambar 12 Hasil Akurasi Tiap Kernel untuk Klasifikasi Sentimen	55
Gambar 13 Hasil Akurasi Tiap Kernel untuk Klasifikasi Kategori.....	56
Gambar 14 <i>Worldcloud</i> Sentimen Positif.....	60
Gambar 15 <i>Wordcloud</i> Sentimen Negatif	61
Gambar 16 <i>Wordcloud</i> Kategori <i>Functionality</i>	62
Gambar 17 <i>Wordcloud</i> Kategori <i>Reliability</i>	62
Gambar 18 <i>Wordcloud</i> Kategori <i>Usability</i>	63
Gambar 19 <i>Wordcloud</i> Kategori <i>Efficiency</i>	63
Gambar 20 <i>Wordcloud</i> Kategori <i>Maintainability</i>	64
Gambar 21 <i>Wordcloud</i> Kategori <i>Portability</i>	65




Daftar Tabel

Tabel 1 Confusion Matrix untuk 2 Kelas	19
Tabel 2 Confusion Matrix untuk Multiclass.....	19
Tabel 3 Luaran dari Googlescraper	30
Tabel 4 Contoh dataset kelas sentimen	31
Tabel 5 Contoh dataset kelas kategori.....	32
Tabel 6 Hasil <i>Output</i> dari Proses <i>Case Folding</i>	35
Tabel 7 Hasil <i>Output</i> dari Proses Pembersihan Data.....	36
Tabel 8 Hasil <i>Output</i> dari Proses <i>Normalization</i>	37
Tabel 9 Hasil <i>Output</i> dari Proses <i>Stemming</i>	38
Tabel 10 Hasil <i>Output</i> dari Proses <i>Stopword Removal</i>	40
Tabel 11 Hasil <i>Output</i> dari Proses <i>Tokenization</i>	41
Tabel 12 Data Sampel untuk Perhitungan TF-IDF Setelah Pra Proses Data	42
Tabel 13 Hasil DF dan IDF	44
Tabel 14 Hasil Pembobotan TF-IDF	46
Tabel 15 Data Sampel untuk Klasifikasi	48
Tabel 16 Bidang Pembatas	48
Tabel 17 Transformasi Data	49
Tabel 18 Matrix <i>Kernel Linear</i>	51
Tabel 19 Permisalan Nilai α	52
Tabel 20 Hasil Nilai W	52
Tabel 21 Confusion Matrix Kelas Sentimen	57
Tabel 22 Confusion Matrix Kelas Kategori	58

Daftar Lampiran

Lampiran 1 Daftar Kata <i>Normalization</i>	71
Lampiran 2 Daftar <i>Stopword</i>	75
Lampiran 3 <i>Similarity Index</i>	78

Daftar Simbol

<i>Simbol Flowchart</i>			
<i>No</i>	Simbol	Nama Simbol	Keterangan
<i>1</i>		<i>Terminator (strat, end)</i>	Simbol yang menjelaskan dimulainya atau berakhirnya suatu kegiatan.
<i>2</i>		<i>Flow Line</i>	Simbol yang menggambarkan hubungan antar simbol
<i>3</i>		<i>Process</i>	Simbol yang menjelaskan proses yang sedang di lakukan.