

# Klasifikasi Ulasan Pengguna Menggunakan Metode *Support Vector Machine* Pada Aplikasi Halodoc

Fachran Sandi<sup>1</sup>, Iin Ernawati<sup>2</sup>

Fakultas Ilmu Komputer

Universitas Pembangunan Nasional Veteran Jakarta

Jl. Rs. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12450, Indonesia

fachransandi@upnvj.ac.id<sup>1</sup>, iinernawati@upnvj.ac.id<sup>2</sup>

**Abstrak.** 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:** Halodoc, ISO 9126, *Support Vector Machine*

## 1. Pendahuluan

Dengan adanya pandemic Covid-19, semua layanan publik juga beralih secara online, layana publik kesehatan menjadi layanan yang harus diutamakan di masa ini. Telemedicine adalah praktek kesehatan yang dilakukan secara jarak jauh menggunakan komunikasi audio, visual, dan data, termasuk dengan perawatan, diagnosa, konsultasi dan pengobatan [1].

Perkembangan telemedicine di Indonesia telah diterapkan dengan adanya peraturan Menteri Kesehatan Nomor 20 Tahun 2019 terkait penyelenggara pelayanan telemedicine antar fasilitas pelayanan kesehatan. Salah satu pelayanan *telemedicine* yang ada di Indonesia adalah aplikasi Halodoc. Halodoc adalah apalikasi telemedicine yang ada di Indonesia yang didirikan oleh Jonathan Sudharta pada tahun 2016 di Jakarta.

Minat masyarakat dalam menggunakan layanan kesehatan secara jarak jauh menggunakan *telemedicine* meningkat sejak *pandemic* Covid-19. Berdasarkan survei pada tahun 2020 yang dilakukan oleh McKinsey & Company, 76% konsumen tertarik menggunakan *telemedice* dalam masa mendatang, sedangkan pada tahun sebelumnya tahun 2019 angka minat konsumen hanya sebesar 11% [2]. Fitur layanan pada aplikasi halodoc antara lain layanan konsultasi dengan dokter, pembelian obat dengan memesan obat sesuai resep dokter atau obat bebas, fitur buat janji rumah sakit [3]

Dikarenakan adanya lonjakan pengguna pada aplikasi kesehatan halodoc pada saat Covid-19 kemarin, terdapat peningkatan 600% pengguna sebelum adanya Covid-19, berdasarkan wawancara Jonathan Sudharta Bersama CNBC pada Juni 2020 [4]. Aplikasi HaloDoc diperlukan adanya perbaikan untuk mengurangi kekurangannya dalam menyediakan informasi kepada pengguna, ulasan pengguna akan digunakan untuk analisa sentimen yang dimana ulasan pengguna tersebut dapat digunakan untuk mengukur kualitas perangkat lunak yang ada pada aplikasi Halodoc. Hal ini diharapkan dapat meningkatkan kualitas perangkat lunak untuk menghindari adanya kesalahan pada sistem atau kurangnya informasi kepada pengguna pada versi selanjutnya.

Penelitian yang dilakukan oleh Hendra & Fitriyani pada tahun 2021, dimana metode yang digunakan adalah naïve bayes, tujuan dari penelitian ini untuk mengetahui kinerja algoritma Naïve Bayes pada ulasan Halodoc, memiliki kemiripan yang sama pada penelitian ini yaitu memiliki fokus objek penelitian yang sama (aplikasi HaloDoc) [5]. Penelitian terdahulu yang dilakukan oleh Rohanah, Rianti, & Sari, serta Ilmawan dan Mude, pada tahun 2021 dan 2020. Melakukan penelitian dengan membandingkan Metode Naïve Bayes dan Metode SVM pada analisis sentimen, dimana dari kedua penelitian tersebut menunjukkan bahwa metode SVM memiliki akurasi lebih tinggi, hasil dari penelitian tersebut mempengaruhi dalam menentukan metode yang dilakukan pada penelitian ini [6] [7]. Penelitian yang dilakukan oleh Rohanah, Dermawan, & Purnamasari, mengenai klasifikasi pengguna memiliki kemiripan yang sama yaitu fokus penelitian yang sama menggunakan ISO/IEC 9126 sebagai kelas

kategori dalam klasifikasi pengguna [8].

Tujuan penelitian ini untuk meningkatkan pelayanan pada aplikasi halodoc serta untuk mengurangi kekurangannya dalam menyediakan informasi kepada pengguna, penelitian ini akan melakukan klasifikasi ulasan pengguna pada aplikasi Halodoc menggunakan algoritma *Support Vector Machine (SVM)* dengan data yang diperoleh dari hasil ulasan yang ada pada google playstore. Pada penelitian ini menggunakan ISO/IEC 9126 sebagai metode dalam mengukur kualitas perangkat lunak.

## 2. Metodologi Penelitian

### 2.1 Text Mining

*Text Mining* terdiri dari algoritma inti yang memungkinkan untuk mendapatkan wawasan dan pengetahuan dari data masif [9]. *Text mining* mengacu pada proses *extraction* pengetahuan yang di perlukan untuk membuat keputusan penting dari data tekstual [10].

### 2.2 Analisis Sentimen

Analisis sentimen adalah bidang study yang mempelajari analisis pendapat seseorang, sentimen, evaluasi, penilaian, emosi, dan sikap terhadap suatu event, produk, layanan, isu yang sedang terjadi [11]. Analisis sentimen memiliki fungsi yang penting untuk perusahaan, dengan melihat opini yang diberikan untuk produk / jasa pada suatu perusahaan, memiliki kepuasan pelayanan publik, serta dapat membantu pihak development untuk mengembangkan aplikasi kedepannya [12].

### 2.3 ISO/IEC 9126

Penerapan ISO 9126 merupakan standar pengukuran jaminan kualitas perangkat lunak yang dibuat oleh *International Organization for Standardization (ISO)* dan *International Electrotechnical Commission (IEC)*. Pada dasarnya ISO 9126 digunakan untuk menentukan dan mengevaluasi produk perangkat lunak untuk mengetahui kualitas internal dan kualitas eksternal serta hubungannya dengan atribut kualitas [13]. 6 *sub*-karakteristik dari model kualitas software ISO 9126, sebagai berikut:

1. *Functionality* adalah kemampuan fungsi produk perangkat lunak yang menyediakan kepuasan kebutuhan user.
2. *Reliability* adalah kemampuan perangkat lunak dapat mempertahankan level kinerjanya dalam kondisi dan waktu yang telah ditetapkan.
3. *Usability* adalah kemampuan perangkat lunak yang menunjukkan kemudahannya pengoperasian.
4. *Efficiency* adalah kemampuan perangkat lunak yang menyangkut waktu eksekusi ketika perangkat lunak sdijalankan.
5. *Maintainability* adalah kemampuan perangkat lunak dalam melakukan perubahan perangkat lunak yang mencakup pembaruan, perbaikan dan sebagainya.
6. *Portability* adalah kemampuan perangkat lunak yang menyangkut kemampuan perangkat tersebut digunakan dilingkungan yang berbeda.

### 2.4 Telemedicine

Telemedicine adalah penggunaan teknologi informasi dan komunikasi (TIK) untuk memberikan pelayanan kesehatan di mana ada pemisah fisik antara penyedia layanan atau penerima baik jarak dekat maupun jauh [14]. *Telemedicine* adalah pemanfaatan teknologi dengan memakai komunikasi audio, visual dan data, data termasuk diagnose atau data medis lainnya yang digunakan untuk praktek kesehatan melalui jarak jauh.

### 2.5 Praproses Data

Data yang baru dikumpulkan kemungkinan terdapat banyak noise, oleh karena itu sebelum masuk ke klasifikasi, data harus terlebih dahulu di lakukan *preprocessing*. Ada beberapa tahapan yaitu:

- a) *Case Folding*  
Tahapan ini mengubah semua teks huruf besar yang ada pada kalimat menjadi huruf kecil.
- b) Pembersihan Data

- Tahapan ini menghapus karakter yang tidak dibutuhkan, seperti tanda baca, angka, dan sebagainya.
- c) *Normalization*  
Tahapan ini mengubah kata yang tidak baku, disingkat, atau salah eja menjadi sesuai dengan kamus KBBI (Kamus Besar Bahasa Indonesia)
  - d) *Stemming*  
Pada tahapan ini mengubah kata menjadi bentuk kata dasar sesuai dengan KBBI (Kamus Besar Bahasa Indonesia)
  - e) *Stopword Removal*  
Tahapan ini menghapus kata yang tidak bermakna, menggunakan stopwords sastrawi.
  - f) *Tokenization*  
Pada tahap ini kalimat yang terdapat pada ulasan di pecah menjadi kata per kata dengan spasi sebagai pemisahannya.

## 2.6 Pembobotan Kata

TF-IDF (*Term Frequency-Inverse Document Frequency*) adalah skema pembobotan kata yang populer [10]. Fungsi dari TF-IDF untuk mencari kata yang penting sesuai dengan berapa kali kata itu muncul dalam dokumen dan dengan mengimbangi dengan frekuensi kemunculannya di setiap dokumen.

Rumus perhitungan *term frequency* (TF) pada Persamaan (1).

$$TF_i = N(t_i, d) \quad (1)$$

Rumus perhitungan *inverse document frequency* (IDF) pada Persamaan (2).

$$IDF_i = \log\left(\frac{N}{df_i}\right) \quad (2)$$

Rumus perhitungan TF-IDF pada Persamaan (3).

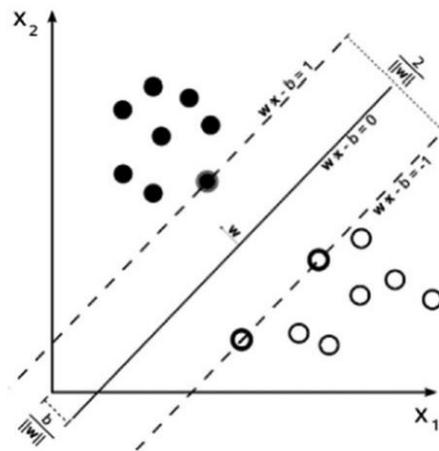
$$TF\_IDF_i = TF_i \cdot IDF_i \quad (3)$$

Keterangan :

- N = Total dokumen.
- $df_i$  = Total dokumen yang memiliki kata (t)

## 2.7 Klasifikasi SVM

*Support Vector Machine* (SVM), adalah metode klasifikasi yang berdasarkan pembatas linier margin maksimum. Tujuan dari SVM ini untuk menemukan *hyperplane* optimal yang memaksimalkan kerenggan antar kelas [9]. Pada penelitian ini akan dibangun 2 model dimana model yang pertama untuk klasifikasi ulasan penggunaan berdasarkan kelas sentimen yaitu kelas negatif dan positif, sedangkan model lainnya akan digunakan untuk klasifikasi ulasan pengguna berdasarkan ISO 9126 menggunakan metode SVM *One – Against – All* dengan persamaan (15) sebagai solusi untuk penyelesaian klasifikasi untuk model kelas kategori.



**Gambar 1.** SVM 2 kelas yang menggambarkan *hyperplane* dan *support vector*

Pada Gambar 2 diasumsikan bahwa 2 kelas objek terpisah sempurna oleh *hyperplane* didefinisikan sebagai berikut:

$$w \cdot x + b = 0 \quad (4)$$

keterangan:

$x$  = input vector

$w$  = berat vektor

$b$  = bias

Pola  $x$  yang termasuk ke dalam kelas positif (+1) didefinisikan sebagai berikut:

$$w \cdot x + b \geq 1 \quad (5)$$

Pola  $x$  yang termasuk ke dalam kelas negatif (-1) didefinisikan sebagai berikut:

$$w \cdot x + b \leq -1 \quad (6)$$

Dalam menentukan nilai  $w$  dapat menggunakan persamaan sebagai berikut:

$$w = \sum_{i=1}^n \alpha_i y_i x_i \quad (7)$$

Dalam menentukan nilai  $w$  kita membutuhkan nilai  $\alpha$ , untuk mencari nilai  $\alpha$  kita dapat menggunakan *lagrange multiplier*.

$$\max_{\alpha} L_d = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i x_j) \quad (8)$$

Dalam menentukan nilai  $b$  dapat menggunakan persamaan sebagai berikut:

$$b = -\frac{1}{2} (x_i^+ \cdot w + x_i^- \cdot w) \quad (9)$$

Dalam penentuan kelas dapat menggunakan persamaan dibawah ini:

$$f(x) = \text{sign}[\sum_{i=1}^m \alpha_i y_i K(x_i x_j) + b] \quad (10)$$

Pemisahan kelas yang tidak bisa dilakukan secara linier, dapat dilakukan oleh fungsi kernel, dengan mentransformasi vektor ruang dimensi. Sebagai contoh untuk transformasi ke fungsi kernel:

Fungsi kernel *Linear*:

$$K(X_i, X_j) = X_i \cdot X_j \quad (11)$$

Fungsi kernel *Polynomial* derajat  $h$ :

$$K(X_i, X_j) = (X_i \cdot X_j + 1)^h \quad (12)$$

Fungsi kernel *Gaussian radial basis* :

$$K(X_i, X_j) = e^{-\|X_i - X_j\|^2 / 2\sigma^2} \quad (13)$$

Fungsi kernel *Sigmoid*:

$$K(X_i, X_j) = \tanh(\kappa X_i \cdot X_j - \delta) \quad (14)$$

Untuk perhitungan SVM *One – Against – All* pada persamaan (15)

$$\begin{cases} \text{Cari } w^i \text{ dan } b^i \text{ untuk meminimalkan } \frac{\|w^i\|^2}{2} + C \sum_{j=1}^m \xi_j^i \\ \text{Subject untuk } y_j (w^i \cdot x_j - b^i) - 1 - \xi_j^i, \xi_j^i \geq 0, \text{ untuk semua } j = 1, 2, \dots, m. \end{cases} \quad (15)$$

## 2.8 Evaluasi

Pada saat melakukan klasifikasi, penting untuk mengetahui seberapa besar kesalahan dalam melakukan klasifikasi. Akurasi klasifikasi merupakan ukuran (statistik) yang menunjukkan seberapa baik model klasifikasi yang kita buat [15].

**Tabel 1** *Confusion Matrix* untuk 2 Kelas

		<i>Predicted Class</i>	
		<i>Positive</i> ( $c_1$ )	<i>Negative</i> ( $c_2$ )
<i>Actual Class</i>	<i>Positive</i> ( $c_1$ )	<i>True Positive</i> (TP)	<i>False Positive</i> (FP)
	<i>Negative</i> ( $c_2$ )	<i>False Negative</i> (FN)	<i>True Negative</i> (TN)

**Tabel 2** *Confusion Matrix* untuk *Multiclass*

		<i>Predicted Class</i>			
		$C_1$	$C_2$	...	$C_N$
<i>Actual Class</i>	$C_1$	$C_{1,1}$	FP	...	$C_{1,N}$
	$C_2$	FN	TP	...	FN
	...	...	...	...	...
	$C_N$	$C_{N,1}$	FP	...	$C_{N,N}$

Keterangan:

- True Positive* (TP) : jumlah benar prediksi oleh pengklasifikasi dengan benar sebagai positif
- False Positive* (FP) : jumlah prediksi oleh pengklasifikasi menjadi positif yang sebenarnya negatif
- False Negative* (FN) : jumlah prediksi oleh pengklasifikasi negatif yang sebenarnya kelas positif
- True Negative* (TN) : jumlah prediksi oleh pengklasifikasi dengan benar sebagai negative

Matrik kinerja *confusion matrix* untuk klasifikasi biner

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (17)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (18)$$

$$\text{F-1 score} = 2 \cdot \frac{\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}} \quad (19)$$

Matrik Kinerja *confusion matrix* untuk *multiclass*

$$\text{Akurasi} = \frac{\sum_{i=1}^N TP(C_i)}{\sum_{i=1}^N \sum_{j=1}^N C_{ij}} \quad (20)$$

$$\text{Precision} = \frac{TP(C_i)}{TP(C_i) + FP(C_i)} \quad (21)$$

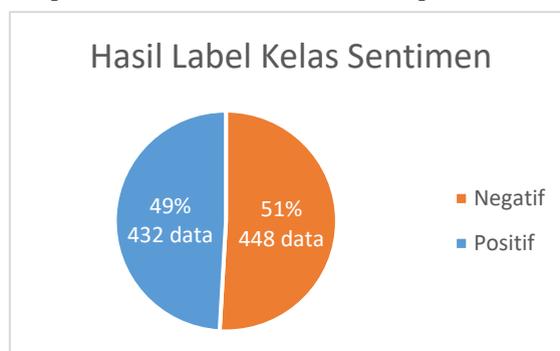
$$\text{Recall} = \frac{TP(C_i)}{TP(C_i) + FN(C_i)} \quad (22)$$

$$\text{F-1 score} = 2 \cdot \frac{\text{Recall}(C_i) \cdot \text{Precision}(C_i)}{\text{Recall}(C_i) + \text{Precision}(C_i)} \quad (23)$$

### 3. Hasil dan Pembahasan

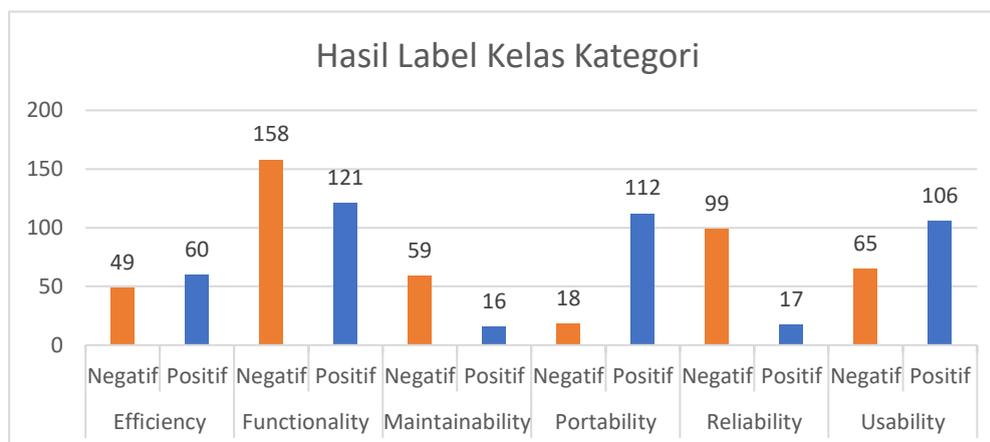
Data yang digunakan dalam penelitian ini diperoleh dari hasil review aplikasi halodoc versi 10, menggunakan bahasa pemrograman *python* dengan menggunakan *library googlescraper*. Proses yang dilakukan diambil secara otomatis dari review yang ada pada *google playstore* dalam periode data mulai dari tanggal 20 Juli 2021 sampai 12 November 2021, data yang didapatkan sebanyak 880 *record*.

Tahap pelabelan dilakukan terhadap 880 data ulasan secara manual oleh penulis sendiri dibantu dengan beberapa referensi yang menunjang dalam melakukan pelabelan, setiap data ulasan akan diberikan 2 label yaitu label sentimen dan label kategori, hasil pelabelan untuk kelas sentimen dapat dilihat di pada Gambar 2



**Gambar 2.** Grafik Hasil Pelabelan kelas Sentimen

Hasil dari pelabelan untuk kelas sentimen dapat dilihat pada Gambar 2, menunjukkan bahwa untuk aplikasi halodoc untuk versi ke-10, pengguna kurang puas akan aplikasi halodoc ini, ditunjukkan oleh banyaknya ulasan yang negatif sebesar 51% atau 448 data ulasan.



**Gambar 3.** Grafik Hasil Pelabelan kelas Sentimen

Pada Gambar 3, kelas *efficiency*, kelas *porability*, dan kelas *usability*, adalah kelas kategori dengan ulasan positif lebih banyak dari ulasan negatif, sedangkan kelas *functionality*, kelas *maintainability*, dan kelas *reliability*, merupakan kelas kategori dengan ulasan negatif lebih banyak dari ulasan positif.

Setelah dilakukan pelabelan secara manual, tahapan selanjutnya adalah praproses data metode- metode yang dilakukan pada tahapan ini antara lain, *CaseFolding*, *Pembersihan Data*, *Nomalization*, *Stopword Removal*, *Stemming*, dan *Tokenization*.

**Tabel 3.** Hasil Praproses Data Sampel

Ulasan	Sebelum Praproses	Setelah Praproses
U1	Aplikasi Gak jelas, sudah buat janji tapi di undur terus	['jelas', 'janji', 'undur']
U2	Aplikasi mudah di mengerti	['mudah', 'erti']
U3	Bagus banget abis bayar ada hitungan waktu menunggu konsultasi dg dokter.	['bagus', 'habis', 'bayar', 'hitung', 'tunggu', 'konsultasi']
U4	Perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan	['modifikasi', 'tampil', 'menu', 'mudah']
U5	Error lebih dari 10 menit pas janji ama dokter.	['error', 'janji']
U6	Barang sdh diterima dgn baik sesuai order, pengiriman cukup cepat, tks Halodoc	['barang', 'sesuai', 'order', 'kirim', 'cepat']
U7	Sdh seminggu ini aplikasinya kok g bs dipakai yaa.. Sdh saya update sampe download ulang tetep bgitu.. kenapa yaa..	['update', 'download', 'ulang']
U8	Membabtu banget, apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes. Terima kasih Halodoc	['bantu', 'pandemic', 'sakit', 'ringan', 'malas', 'badan', 'lemas']

Tabel 3 adalah hasil dari praproses, setelah melewati tahap Praproses data, data akan dibagi terlebih dahulu menjadi data *training* (80%) dengan data berjumlah 704 data dan data *testing*(20%) dengan data berjumlah 176

data, setelah itu dilanjutkan dengan pem bobotan TF-IDF. Nilai pem bobotan TF-IDF ukuran statistik dari perkalian antara nilai TF dengan nilai IDF yang menggambarkan pentingnya istilah tersebut. Perhitungan bobot TF-IDF dapat dilakukan dengan menggunakan persamaan rumus (3).

**Tabel 4.** Perhitungan TF-IDF Data Sampel

Kata	TF			DF	IDF	TF-IDF		
	U1	...	U8			U1	...	U8
badan	0	...	1	1	0.903	0	...	0.903
bagus	0	...	0	1	0.903	0	...	0
bantu	0	...	1	1	0.903	0	...	0.903
barang	0	...	0	1	0.903	0	...	0
bayar	0	...	0	1	0.903	0	...	0
cepat	0	...	0	1	0.903	0	...	0
download	0	...	0	1	0.903	0	...	0
error	0	...	0	1	0.903	0	...	0
erti	0	...	0	1	0.903	0	...	0
habis	0	...	0	1	0.903	0	...	0
hitung	0	...	0	1	0.903	0	...	0
janji	1	...	0	2	0.602	0.602	...	1
jelas	1	...	0	1	0.903	0.903	...	1
kirim	0	...	0	1	0.903	0	...	0
konsultasi	0	...	0	1	0.903	0	...	0
lemas	0	...	1	1	0.903	0	...	0.903
malas	0	...	1	1	0.903	0	...	0.903
menu	0	...	0	1	0.903	0	...	0
modifikasi	0	...	0	1	0.903	0	...	0
mudah	0	...	0	2	0.602	0	...	0
order	0	...	0	1	0.903	0	...	0
pandemic	0	...	1	1	0.903	0	...	0.903
ringan	0	...	1	1	0.903	0	...	0.903
sakit	0	...	1	1	0.903	0	...	0.903
sesuai	0	...	0	1	0.903	0	...	0
tampil	0	...	0	1	0.903	0	...	0
tunggu	0	...	0	1	0.903	0	...	0
ulang	0	...	0	1	0.903	0	...	0
undur	1	...	0	1	0.903	0.903	...	0
update	0	...	0	1	0.903	0	...	0

Setelah dilakukan pem bobotan kata, kita akan melakukan pengklasifikasian menggunakan SVM. Pada Tabel 5 kita akan mencoba pengklasifikasian dengan data di bawah ini, U1 sebagai data ujinya.

**Tabel 5.** Data Sampel untuk Klasifikasi

Ulasan	
U1	Aplikasi Gak jelas, sudah buat janji tapi di undur terus
U2	Aplikasi mudah di mengerti

U3	Bagus banget abis bayar ada hitungan waktu menunggu konsultasi dg dokter.
U4	Perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan
U5	Eror lebih dari 10 menit pas janji ama dokter.
U6	Barang sdh diterima dgn baik sesuai order, pengiriman cukup cepat, tks Halodoc
U7	Sdh seminggu ini aplikasinya kok g bs dipakai yaa.. Sdh saya update sampe download ulang tetep bgitu.. kenapa yaa..
U8	Membabtu banget, apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes. Terima kasih Halodoc

Dari Tabel 5 kita dapat menentukan bidang pembatasnya menggunakan rumus persamaan (5), dan (6). Dapat dilihat dari Tabel 6 bahwa ulasan yang diberi label negatif akan diberi nilai -1, sebaliknya untuk ulasan yang diberi label positif akan diberi nilai 1. Untuk U1 tidak diproses karena U1 sebagai data uji.

**Tabel 6.** Bidang Pembatas

Bidang Pembatas
U2: $0.845w_1 + 0,544w_2 + b \geq 1$
U3: $0.845w_1 + 0.845w_2 + 0.845w_3 + 0.845w_4 + 0.845w_5 + 0.845w_6 + b \geq 1$
U4: $0.845w_1 + 0.845w_2 + 0.845w_3 + 0,544w_4 + b \leq -1$
U5: $0.845w_1 + 0.845w_2 + b \geq -1$
U6: $0.845w_1 + 0.845w_2 + 0.845w_3 + 0.845w_4 + 0.845w_5 + b \geq 1$
U7: $0.845w_1 + 0.845w_2 + 0.845w_3 + b \leq -1$
U8: $0.845w_1 + 0.845w_2 + 0.845w_3 + 0.845w_4 + 0.845w_5 + 0.845w_6 + 0.845w_7 + b \geq 1$
U2: $0.845w_1 + 0,544w_2 + b \geq 1$

Setelah didapatkan bidang pembatas pada Tabel 6, akan dilakukan transformasi data dapat dilihat pada Tabel 7 dengan U1 tidak dilakukan perhitungan TF-IDF.

**Tabel 7.** Transformasi Data

Kata	TF						
	U2	U3	U4	U5	U6	U7	U8
hitung	0	0,845	0	0	0	0	0
konsultasi	0	0,845	0	0	0	0	0
tunggu	0	0,845	0	0	0	0	0
menu	0	0	0,845	0	0	0	0

modifikasi	0	0	0,845	0	0	0	0
tampil	0	0	0,845	0	0	0	0
error	0	0	0	0,845	0	0	0
janji	0	0	0	0,845	0	0	0
barang	0	0	0	0	0,845	0	0
cepat	0	0	0	0	0,845	0	0
kirim	0	0	0	0	0,845	0	0
order	0	0	0	0	0,845	0	0
sesuai	0	0	0	0	0,845	0	0
download	0	0	0	0	0	0,845	0
ulang	0	0	0	0	0	0,845	0
update	0	0	0	0	0	0,845	0
badan	0	0	0	0	0	0	0,845
bantu	0	0	0	0	0	0	0,845
lemas	0	0	0	0	0	0	0,845
malas	0	0	0	0	0	0	0,845
pandemic	0	0	0	0	0	0	0,845
ringan	0	0	0	0	0	0	0,845
sakit	0	0	0	0	0	0	0,845
y (kelas)	1	1	-1	-1	1	-1	1

Setelah format fitur sudah diubah selanjutnya kita akan menggunakan kernel *linear* dengan persamaan rumus(11). Pada Tabel 8 adalah hasil *matriks* dari perhitungan kernel *linear*.

**Tabel 8.** Matrix Kernel Linear

Matriks							
$x_i x_j^T =$	1,010	0	0,296	0	0	0	0
	0	4,285	0	0	0	0	0
	0,296	0	2,438	0	0	0	0
	0	0	0	1,428	0	0	0
	0	0	0	0	3,570	0	0
	0	0	0	0	0	2,142	0
	0	0	0	0	0	0	4,999

Selanjutnya kita akan mencari nilai  $\alpha$  dengan persamaan rumus(8) nilai ini untuk menentukan margin terbesar dengan teknik komputasi *lagrange multiplier*, yang dapat dilihat pada Tabel 9.

**Tabel 9.** Permisalan Nilai  $\alpha$

y (kelas)		1	1	-1	-1	1	-1	1
Kemungkinan alfa	$\alpha_1$	2	1	1	1	1	1	1
	$\alpha_2$	1	2	1	1	1	1	1
	$\alpha_3$	1	1	2	1	1	1	1

Dari Tabel 9 dalam menentukan kemungkinan alfa kita dapat memilih nilainya dari range 1 sampai banyak dokumen kita yaitu 7. Kita akan memilih  $\alpha_1$  karena memiliki nilai terbesar,  $\max Ld \alpha_1 = -2,860$ ,  $\max Ld \alpha_2 = -8,069$ , dan  $\max Ld \alpha_3 = -5,003$ .

Setelah mendapatkan nilai  $\alpha$  selanjutnya kita akan mencari nilai  $w$  untuk setiap fitur yang kita punya, untuk mencarinya kita akan menggunakan persamaan rumus(7), hasil dari perhitungan nilai  $w$  dapat dilihat pada Tabel 10.

Tabel 10. Hasil Nilai W

Fitur /Term	Nilai w
erti	1,690196
mudah	0,544068
bayar	0,845098
bagus	0,845098
habis	0,845098
hitung	0,845098
konsultasi	0,845098
tunggu	0,845098
menu	-0,8451
modifikasi	-0,8451
tampil	-0,8451
error	-0,8451
janji	-0,8451
barang	0,845098
cepat	0,845098
kirim	0,845098
order	0,845098
sesuai	0,845098
download	-0,8451
ulang	-0,8451
update	-0,8451
badan	0,845098
bantu	0,845098
lemas	0,845098
malas	0,845098
pandemic	0,845098
ringan	0,845098
sakit	0,845098

Setelah mendapatkan nilai w kita dapat mencari nilai b, untuk nilai b kita akan menggunakan persamaan rumus (9), untuk hasil perhitungan didapatkan nilai  $b = -4,938$ .

Selanjutnya kita akan menguji model yang sudah kita buat dengan data uji yang ada di atas yaitu U1, dimana U1 ini memiliki kelas *actual* negatif. Dalam menentukan kelas kita akan menggunakan persamaan rumus (10), nilai yang sudah kita dapatkan antara lain:

$$\begin{array}{l} \alpha : \\ y: \\ b: \end{array} \begin{array}{|c|c|c|c|c|c|c|} \hline 2 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & -1 & -1 & 1 & -1 & 1 \\ \hline \end{array} \\ -4,938$$

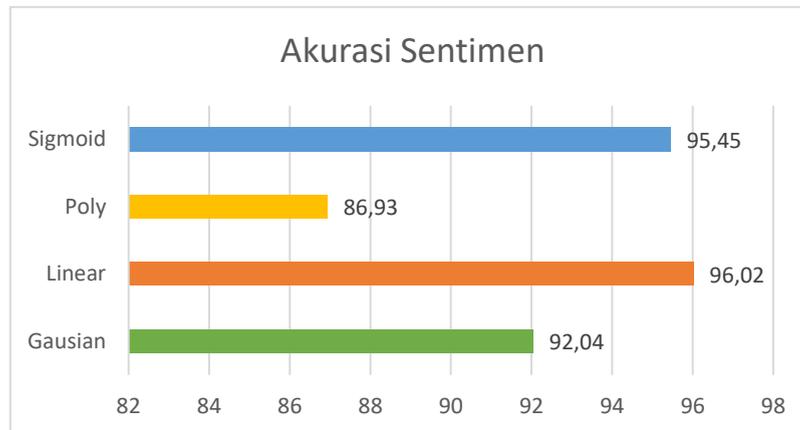
Nilai  $\alpha$ ,  $y$ , dan  $b$  sudah didapatkan nilai yang belum ada adalah perhitungan kernel untuk data uji karena kita menggunakan kernel linear maka dapat kita gunakan persamaan rumus (11), hasil yang didapatkan dari perhitungan tersebut adalah :

$$k(x_i x_j) = \begin{array}{|c|c|c|c|c|c|c|} \hline 0 & 0 & 0 & 0,459 & 0 & 0 & 0 \\ \hline \end{array}$$

Setelah semua nilai yang dibutuhkan sudah ada, kita dapat menggunakan persamaan rumus (10), hasil dari perhitungan tersebut adalah  $-1$ , fungsi  $sign()$  adalah fungsi normalisasi, jika hasil nilai  $\geq 0$ , maka fungsi tersebut bernilai kelas positif, jika hasil nilai  $\leq 0$ , maka fungsi tersebut bernilai kelas negatif. Berdasarkan fungsi keputusan diatas maka data uji masuk ke kelas "negatif", dimana model yang kita buat sesuai dengan *actual class*-nya.

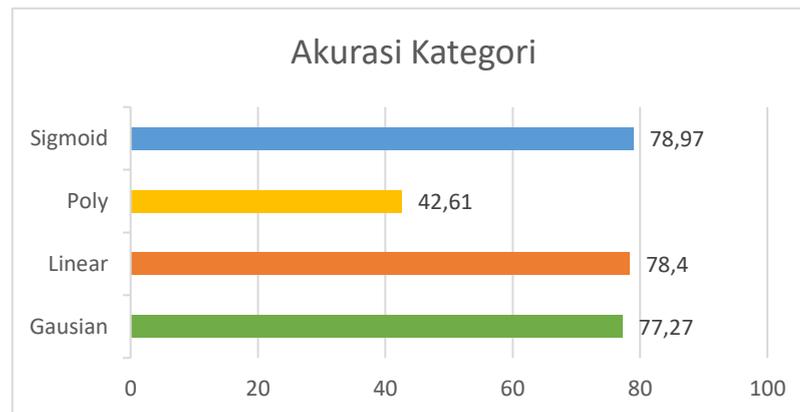
Setelah dilakukan pembobotan kata, di tahap ini akan dibuat 2 model klasifikasi dimana yang pertama untuk klasifikasi sentimen dan klasifikasi kategori. Pada setiap model akan digunakan 4 kernel yaitu Kernel *Linear*

dengan rumus persamaan (11), Kernel *Polynomial* dengan rumus persamaan (12), Kernel RBF dengan rumus persamaan (13), dan Kernel *Sigmoid* dengan rumus persamaan (14).



**Gambar. 4.** Hasil Akurasi Tiap Kernel untuk Klasifikasi Sentimen

Pada Gambar 4 menunjukkan untuk akurasi tertinggi untuk klasifikasi kelas sentimen ada di kernel *linear* dengan akurasi sebesar 96,02%, disusul oleh akurasi tertinggi selanjutnya kernel *sigmoid* sebesar 95,45%, dan kernel *gaussian* sebesar 92,04%, serta kernel *polynomial* dengan akurasi sebesar 86,93%.



**Gambar. 5.** Hasil Akurasi Tiap Kernel untuk Klasifikasi Kategori

Pada Gambar 5 menunjukkan untuk akurasi tertinggi untuk klasifikasi kelas kategori ada di kernel *sigmoid* dengan akurasi sebesar 78,97%, disusul oleh akurasi tertinggi selanjutnya kernel *linear* sebesar 78,40%, dan kernel *gaussian* sebesar 77,27%, serta kernel *polynomial* dengan akurasi sebesar 42,61%.

Proses evaluasi pada penelitian ini akan menggunakan confusion matrix untuk melihat seberapa baik model yang sudah kita buat. Pada confusion matrix akan menghasilkan nilai TP(true positive), TN(true negative), FP(false positive), dan FN(false negative). Tabel 11 adalah *confusion matrix* kelas sentimen dan Tabel 12 adalah *confusion matrix* untuk kelas kategori

**Tabel 11.** *Confusion Matrix* kelas sentimen

True Label	Predict Label	
	positif	negatif
positif	86	4
negatif	3	83

$$Accuracy = \frac{86 + 83}{86 + 83 + 4 + 3} = 0,96$$

$$\begin{aligned}
 \text{Positif} = \quad & \text{Precision} = \frac{86}{86+4} = 0,955 \\
 & \text{Recall} = \frac{86}{86+3} = 0,966 \\
 & \text{F1 - score} = 2 * \frac{0,955 \times 0,966}{0,955 + 0,966} = 0,960 \\
 \text{Negatif} = \quad & \text{Precision} = \frac{83}{83+3} = 0,965 \\
 & \text{Recall} = \frac{83}{83+4} = 0,954 \\
 & \text{F1 - score} = 2 * \frac{0,965 \times 0,954}{0,965 + 0,954} = 0,959
 \end{aligned}$$

Tabel 12. Confusion Matrix kelas Kategori

True Label	Predict Label					
	Effeciency	Functionality	Maintainability	Portability	Reliability	Usability
Effeciency	20	1	0	0	0	1
Functionality	1	46	1	3	2	3
Maintability	1	1	11	0	1	1
Portability	0	3	0	22	1	0
Reliability	0	3	1	1	18	0
Usability	2	6	1	1	2	22

$$\begin{aligned}
 \text{Accuracy} &= \frac{20+46+11+22+18+22}{20+46+11+22+18+22+37} = 0,789 \\
 \text{Effeciency} = \quad & \text{Precision} = \frac{20}{20+4} = 0,833 \\
 & \text{Recall} = \frac{20}{20+2} = 0,909 \\
 & \text{F1 - score} = 2 * \frac{0,833 \times 0,909}{0,833 + 0,909} = 0,869 \\
 \text{Functionality} = \quad & \text{Precision} = \frac{46}{46+14} = 0,766 \\
 & \text{Recall} = \frac{46}{46+10} = 0,821 \\
 & \text{F1 - score} = 2 * \frac{0,766 \times 0,821}{0,766 + 0,821} = 0,792 \\
 \text{Maintainability} = \quad & \text{Precision} = \frac{11}{11+3} = 0,785 \\
 & \text{Recall} = \frac{11}{11+4} = 0,733 \\
 & \text{F1 - score} = 2 * \frac{0,785 \times 0,733}{0,785 + 0,733} = 0,758 \\
 \text{Portability} = \quad & \text{Precision} = \frac{22}{22+5} = 0,814 \\
 & \text{Recall} = \frac{22}{22+4} = 0,846 \\
 & \text{F1 - score} = 2 * \frac{0,814 \times 0,846}{0,814 + 0,846} = 0,829 \\
 \text{Reliability} = \quad & \text{Precision} = \frac{18}{18+6} = 0,75 \\
 & \text{Recall} = \frac{18}{18+5} = 0,782 \\
 & \text{F1 - score} = 2 * \frac{0,75 \times 0,782}{0,75 + 0,782} = 0,765 \\
 \text{Usability} = \quad & \text{Precision} = \frac{22}{22+4} = 0,846 \\
 & \text{Recall} = \frac{22}{22+12} = 0,647 \\
 & \text{F1 - score} = 2 * \frac{0,846 \times 0,647}{0,846 + 0,647} = 0,733
 \end{aligned}$$

Selanjutnya untuk memudahkan dalam menangkap informasi dengan mudah, tahapan visualisasi akan menggunakan *wordcloud*, *wordcloud* menampilkan kata yang sering muncul pada setiap kelas baik sentimen maupun kategori. Semakin besar kata, maka semakin banyak kata tersebut muncul.



Gambar. 6. Wordcloud Sentimen Positif



Gambar. 7. Wordcloud Sentimen Negatif

Dari Gambar 6 menunjukkan bahwa kata “bantu”, “konsultasi”, “obat”, “cepat”, dan “mudah”, adalah kata yang sering muncul untuk kelas sentimen positif, sedangkan dari Gambar 7 menunjukkan bahwa kata “bayar”, “obat”, “vaksin”, “konsultasi”, dan “pesan”, adalah kata yang sering muncul untuk kelas sentimen negatif,



Gambar. 8. Wordcloud Kategori *Functionality*



Gambar. 9. Wordcloud Kategori *Reliability*

Dari Gambar 8 menunjukkan bahwa kata “konsultasi”, “bayar”, “pesan”, “vaksin”, dan “bantu”, adalah kata yang sering muncul untuk kelas kategori *functionality*, sedangkan dari Gambar 9 menunjukkan bahwa kata “bayar”, “obat”, “konsultasi”, “pesan”, dan “lambat”, adalah kata yang sering muncul untuk kelas kategori *reliability*.



Gambar. 10. Wordcloud Kategori *Usability*



Gambar. 11. Wordcloud Kategori *Efficiency*

Dari Gambar 10 menunjukkan bahwa kata “mudah”, “obat”, “bayar”, “konsultasi”, dan “bantu”, adalah kata yang sering muncul untuk kelas kategori *usability*, sedangkan dari Gambar 11 menunjukkan bahwa kata “cepat”, “obat”, “lambat”, “respons”, dan “apotek”, adalah kata yang sering muncul untuk kelas kategori *efficiency*.



Gambar. 12. Wordcloud Kategori *Maintainability*



Gambar. 13. Wordcloud Kategori *Portability*

Dari Gambar 12 menunjukkan bahwa kata “rekondisi”, “koneksi”, “buka”, “masuk”, dan “update”, adalah kata yang sering muncul untuk kelas kategori *maintainability*, sedangkan dari Gambar 13 menunjukkan bahwa kata “bantu”, “rumah”, “obat”, “konsultasi”, dan “butuh”, adalah kata yang sering muncul untuk kelas kategori *portability*.

## 4. Kesimpulan dan Saran

### 4.1 Kesimpulan

Ulasan yang diberikan oleh pengguna untuk aplikasi halodoc versi 10 pada periode Juli sampai November 2021 sebanyak 880 data ulasan, sebanyak 448 ulasan menunjukkan pengguna tidak suka dengan versi ini, dan sebanyak 432 ulasan pengguna suka dengan versi ini. Berdasarkan kategori ISO 9126, sebanyak 279 ulasan menunjukkan kelas *functionality*, 171 ulasan menunjukkan kelas *usability*, 130 ulasan menunjukkan kelas *portability*, 116 ulasan menunjukkan kelas *reliability*, 109 ulasan menunjukkan kelas *efficiency*, dan 75 ulasan menunjukkan *maintainability*. Algoritma SVM dengan data latih sebesar 80% dan data uji sebesar 20%, dapat mengklasifikasi sentimen dengan nilai akurasi sebesar 96,02% dengan kernel *linear*, dan algoritma SVM dapat mengklasifikasi kategori berdasarkan ISO 9126 dengan nilai akurasi sebesar 78,97% dengan kernel *sigmoid*.

### 4.2 Saran

Dari hasil penelitian yang telah dibuat ada beberapa saran untuk pengembangan penelitian, yaitu:

- Data yang digunakan dalam pelabelan di usahakan seimbang untuk setiap labelnya pada saat proses pembangunan model dan pengujian model klasifikasi.
- Penelitian selanjutnya, diharap dapat mengkombinasikan tools atau algoritma lainnya dalam peningkatan kualitas sehingga dapat menunjukkan hasil lebih representative

## Referensi

- [1] S. Gogia, *Fundamentals of Telemedicine and Telehealth*, India: Stacy Masucci, 2020.
- [2] O. Bestsenny, G. Gilbert, A. Harris and J. Rost, "Telehealth: A quarter-trillion-dollar post-COVID-19 reality?," McKinsey & Company, 9 July 2021. [Online]. Available: <https://www.mckinsey.com/industries/healthcare/our-insights/telehealth-a-quarter-trillion-dollar-post-covid-19-reality>. [Accessed 2023 03 6].
- [3] Halodoc, "Frequently Asked Question," Halodoc, [Online]. Available: <https://www.halodoc.com/faq>. [Accessed 6 03 2023].
- [4] L. Hasibuan, "Gegara Corona, Bisnis Halodoc Tumbuh 600% dalam Dua Bulan," CNBC Indonesia, 08 June 2020. [Online]. Available: <https://www.cnbcindonesia.com/tech/20200608153553-37-163900/gegara-corona-bisnis-halodoc-tumbuh-600-dalam-dua-bulan>. [Accessed 6 03 2023].
- [5] A. Hendra and Fitriyani, "Analisis Sentimen Review Halodoc Menggunakan Naïve Bayes Classifier," *JISKa*, pp. 78-89, 2021.
- [6] A. Rohanah, D. L. Rianti and N. B. Sari, "PERBANDINGAN NAÏVE BAYES DAN SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI ULASAN PELANGGAN INDIHOME," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, pp. 23-30, 2021.
- [7] L. B. Ilmawan and M. A. Mude, "Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve

- Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store," *ILKOM Jurnal Ilmiah*, pp. 154-161, 2020.
- [8] A. Rohanah, B. A. Dermawan and I. Purnamasari, "Klasifikasi Ulasan Pengguna Zoom Cloud Meetings Menggunakan Metode Information Gain dan Naïve Bayes Classifier," *Jurnal Informatika Universitas Pamulang*, pp. 348-157, 2021.
- [9] J. M. Zaki and M. W. JR, *Data Mining and Analysis: Fundamental Concepts and Algorithms*, Cambridge University Press, 2020.
- [10] T. Jo, *Text Mining: Concepts, Implementation, and Big Data Challenge*, Springer, 2019.
- [11] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, Cambridge University Press, 2015.
- [12] G. D. W. B. T. W. Amany Mursianto, "Analisis Sentimen Ulasan Pengguna Pada Aplikasi Google Classroom Menggunakan Metode SVM Dan Seleksi Fitur PSO," *Jurnal Informatik*, vol. 18, 2022.
- [13] W. Suryn, *Software Quality Engineering: A Practitioner's Approach*, Montreal: Wiley-IEEE Computer Society Pr, 2014.
- [14] G. S. Sari and W. Wirman, "Telemedicine sebagai Media Konsultasi Kesehatan di Masa Pandemic COVID 19 di Indonesia," *Jurnal Komunikasi*, pp. 43-54, 2021.
- [15] F. Gorunescu, *Data Mining Concepts, Models and Techniques*, Springer, 2011.
- [16] R. Koch, *The 80/20 Principle and 92 Other Powerful Laws of Nature.*, Lodon: Quercus, 2013.