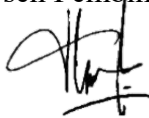


1810511049_Fachran Sandi

by 1 1

Dosen Pembimbing



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

Submission date: 11-Jan-2023 11:17PM (UTC-0800)

Submission ID: 1991649389

File name: turnitin_fachran_sandi.docx (2.73M)

Word count: 13269

Character count: 78123



⁹¹
**KLASIFIKASI ULASAN PENGGUNA MENGGUNAKAN
METODE SUPPORT VECTOR MACHINE PADA APLIKASI
HALODOC**

SKRIPSI

**Fachran Sandi
1810511049**

¹
**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



Lin 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:

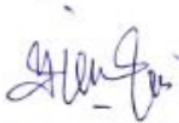
Nama : Fachran Sandi

NIM : 1810511049

Program Studi : SI 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 SI Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Yuni Widiastiwati, S.Kom., M.Si.
Penguji 1



Henki Bayu Seta, S.Kom, MTL.
Penguji 2



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



Dr. Widva 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 ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia di tuntutan dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 4 Januari 2023

Yang Menyatakan,

Fachran Sandi

3

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

5
**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.

1 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 Penelitian.....	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

1		
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

40	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

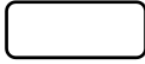


Daftar Tabel

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.....	31
Tabel 6 Hasil <i>Output</i> dari Proses <i>Case Folding</i>	34
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>	Error! Bookmark not defined.

1 Daftar Simbol

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

BAB I PENDAHULUAN

1.1 Latar Belakang

Dengan adanya *pandemic Covid-19*, semua layanan publik juga beralih secara online, layana publik kesehatan menjadi layanan yang harus diutamakan di masa ini. Tercatat bahwa penderita *covid* di Indonesia saat ini lebih dari 4jt jiwa (PHEOC Kemkes RI, 2021). Pemerintah sudah melakukan upaya dalam mencegah penyebaran *covid-19* dengan melakukan protokol kesehatan.

Penggunaan Smartphone di Indonesia berkembang dengan cepat. Pengguna aktif smartphone di Indonesia diperkirakan sekitar 100 juta orang (Lembaga Riset Digital Marketing Emarketer). Dengan meningkatnya Penggunaan *mobile phone*, maka meningkat pula layanan berbasis *electronic*, khususnya di *pandemic* ini, yang mewajibkan kita untuk melakukan protokol kesehatan. *Telemedicine* adalah praktek kesehatan yang dilakukan secara jarak jauh menggunakan komunikasi audio, visual, dan data, termasuk dengan perawatan, diagnosa, konsultasi dan pengobatan (Sari & Wirman, 2021), salah satunya adalah aplikasi HaloDoc.

Aplikasi HaloDoc adalah aplikasi *Telemedicine* yang populer di Indonesia, aplikasi HaloDoc diperlukan adanya perbaikan untuk mengurangi kekurangannya dalam menyediakan informasi kepada pengguna. Kegiatan tersebut didasari dari ulasan yang ada pada *google playstore* terkait aplikasi HaloDoc. Ulasan pengguna akan digunakan untuk analisa sentimen pengguna yang dimana ulasan pengguna tersebut dapat digunakan untuk mengukur kualitas perangkat lunak yang ada pada aplikasi HaloDoc.

Dengan latar belakang tersebut, 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.

1.2 Rumusan Masalah

Berdasarkan latar belakang, maka berikut rumusan masalah pada penelitian ini:

1. Bagaimana gambaran umum dari pengguna aplikasi HaloDoc berdasarkan ulasan di *google playstore*?
2. Bagaimana hasil akurasi dari metode SVM dalam klasifikasi ulasan pengguna pada aplikasi HaloDoc?
3. Bagaimana hasil informasi ulasan pengguna berdasarkan ISO 9126 untuk membantu pengembangan aplikasi?

56 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang ada, tujuan penelitian ini adalah:

1. Mengidentifikasi gambaran umum pengguna aplikasi halodoc berdasarkan ulasan di *google playstore*.
2. Mengukur hasil kinerja dan akurasi dari metode *Support Vector Machine* dalam mengklasifikasi ulasan pengguna pada aplikasi halodoc.
3. Mengetahui informasi ulasan terhadap aplikasi HaloDoc berdasarkan ISO 9126.

47 1.4 Ruang Lingkup

Agar penelitian ini berjalan lancar dan fokus ke tujuan penelitian, maka berikut adalah batasan / ruang lingkup pada penelitian ini:

- 6
1. Algoritma yang digunakan pada algoritma ini adalah *Support Vector Machine* (SVM).
2. Sumber data yang digunakan berasal dari hasil ulasan yang ada pada *google playstore* sebanyak 880 *record* dalam Bahasa Indonesia yang diambil pada tanggal 20 Juli 2021 sampai 12 November 2021.
- 21
3. Klasifikasi kelas kategori berdasarkan *model quality* perangkat lunak ISO/IEC 9126 yaitu *functionality*, *reliability*, *usability*, *efficiency*, *maintainability* dan *portability*.
4. Dataset aplikasi halodoc menggunakan versi 10.

1.5 Luaran yang diharapkan

Luaran yang diperoleh berupa klasifikasi kelas sentimen (positif dan negatif) dan klasifikasi kelas kategori (*functionality*, *reliability*, *usability*, *efficiency*, *maintainability* dan *portability*).

1.6 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini, pembaca dapat mengetahui kinerja dari algoritma *support vector machine* (SVM), apabila ingin membangun model. Sehingga dapat mengklasifikasi setimen dari data tekstual khususnya ulasan aplikasi HaloDoc. Manfaat bagi pihak development aplikasi HaloDoc untuk mempermudah penyampaian dan persepsi proyek pengembangan perangkat lunak dari opini yang ada pada ulasan di google playstore sehingga dapat dijadikan evaluasi untuk kedepannya.

1.7 Sistematika Penulisan

Dalam penulisan skripsi ini, dibagi menjadi beberapa bab, yaitu:

BAB I PENDAHULUAN

Bab ini menjelaskan mengenai latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup, luaran dari penelitian ini dan sistematik dari penelitian ini.

BAB II TINJAUAN PUSTAKA

Bab ini menjelaskan teori-teori yang digunakan pada penelitian ini, Teori yang dibahas dapat berupa definis konsep, metode, model, algoritma, atau prosedur yang terkait dengan topik penelitian.

BAB III METODE PENELITIAN

Pada bab ini memaparkan tahap riset, sumber data, metode pengumpulan data, metode pengolahan data, dan metode menganalisis data yang sesuai dengan tujuan dari penelitian.

BAB IV HASIL DAN PEMBAHASAN

Di bab ini menjelaskan proses mengolah dataset, dimulai dari pengumpulan data, klasifikasi data, pembersihan data, dan pemodelan dataset hingga dapat divisulisasi.

BAB V PENUTUP

Pada bab ini bersisi kesimpulan hasil dari penelitian ini dan saran yang berguna untuk penelitian yang serupa.

⁴⁷ BAB II TINJAUAN PUSTAKA

2.1 Ulasan / Review

Ulasan adalah ulasan yang artinya dalam KBBI memiliki arti komentar, kupasan, tafsiran. Ulasan umumnya berguna bagi suatu perusahaan atau kelompok dalam mengetahui produknya diterima di masyarakat atau tidak. Ulasan sendiri bisa menjadi bahan rekomendasi dalam pengembangan aplikasi. ¹⁷ Menganalisis pendapat, sentimen, evaluasi, sikap, penilai, emosi seseorang dapat membantu kita mengetahui ¹³ apakah berkenaan dengan suatu topik, layanan, organisasi, individu, atau kegiatan tertentu (Liu, 2015).

2.2 Telemedicine

¹⁷ 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 (Gogia, 2020). ⁸⁹ 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.

Perkembangan telemedicine dapat ditelusuri ke penemuannya Alexander Graham Bell, pada tahun 1876. Dia meminta bantuan dari asistennya Mr. Watson setelah menumpahkan asam ke celananya. The Lancet pada tahun 1879 menggambarkan keberhasilan diagnosis seorang anak oleh dokter. Konsep telemedicine berkembang pada tahun 1905 berkat seorang ahli fisiologi Belanda yang memanfaatkan telepon untuk transmisi dan pemantauan suara jantung dan ritme. Penggunaan pertama telemedicine untuk mengirimkan video, gambar, dan data medis yang kompleks terjadi ⁷² pada akhir 1950-an dan awal 1960-an. Pada tahun 1959 Universitas Nebraska menggunakan televisi interaktif (IATV) untuk mengirimkan pemeriksaan neurologis, yang secara luas dianggap sebagai kasus pertama dari konsultasi telemedicine video real-time (Gogia, 2020).

⁴² 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 aplikasi telemedicine yang ada di Indonesia yang didirikan oleh Jonathan Sudharta pada tahun 2016 di Jakarta (<https://id.wikipedia.org/wiki/Halodoc>, diakses pada tanggal 15 November 2021).

2.3 Software Quality

Produk perangkat lunak bagi pengguna harus secara efektif mendukung proses bisnis. Dari pernyataan ini kebutuhan bisnis menjadi alasan bahwa pengembangan produk perangkat lunak yang berkualitas dan kebutuhan bergerak ke posisi pembeli. Dominasi pandangan manufaktur dalam rekayasa perangkat lunak ini dapat ditelusuri kembali di tahun 1960-an, Ketika Departemen Pertahanan AS dan IBM (*International Business Machines*) melahirkan jaminan kualitas perangkat lunak. Hal ini menyebabkan keyakinan bahwa kepatuhan terhadap proses pengembangan, seperti di bidang manufaktur, akan menghasilkan produk yang berkualitas. Selama beberapa tahun terakhir, para peneliti telah mengusulkan pendekatan dan model baru yang mencoba mencakup lebih banyak perspektif kualitas daripada sekadar pandangan manufaktur (Suryan, 2014).

Model Quality menyajikan pendekatan untuk menyatukan atribut kualitas yang berbeda dengan tujuan dasar, sebagai berikut:

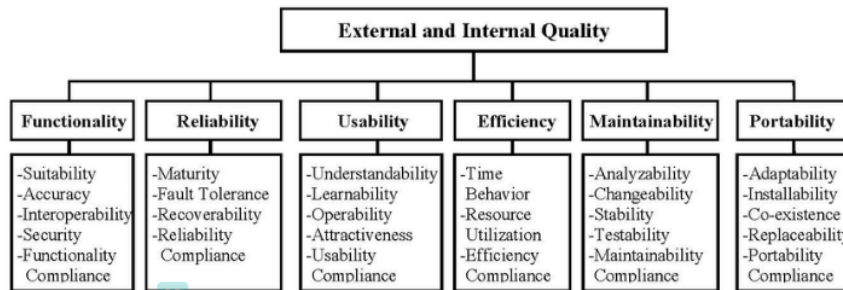
1. Membantu memahami bagaimana beberapa aspek kualitas berkontribusi secara keseluruhan
2. Kualitas perangkat lunak bukan hanya sekedar kesalahan dan kegagalan
3. Membantu mengidentifikasi dan menentukan persyaratan kualitas
4. Bantuan untuk menavigasi peta karakteristik kualitas, subkarakteristik, dan ukuran yang sesuai (rumus pengukuran dan skala)
5. Bantuan untuk menentukan profil evaluasi (apa tepatnya yang harus dievaluasi).

2.4 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 (Suryan, 2014).

Sumber: (Suryan, 2014)



Gambar 1 ISO 9126 quality model untuk external dan internal quality

Model kualitas internal dan eksternal ISO 9126 ada 3 lapis model yang terdiri dari karakteristik kualitas, subkarakteristik kualitas, dan ukuran kualitas dapat dilihat pada Gambar 1.

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.

2.5 Data Mining

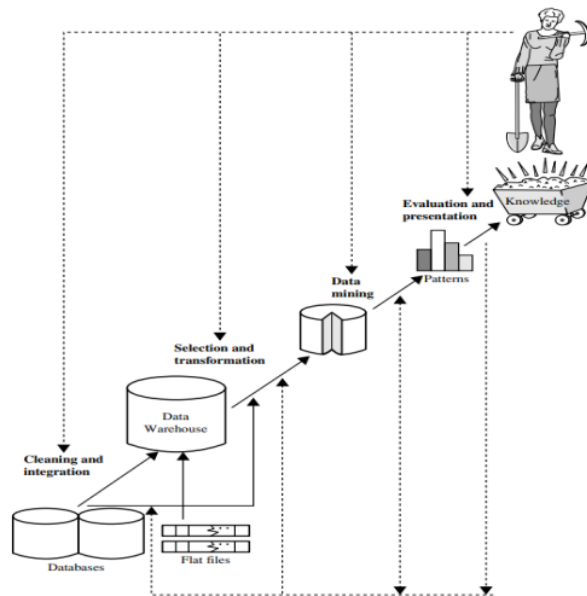
Data mining adalah suatu proses pencarian informasi baru, berharga, dan nontrivial dalam jumlah volume data yang besar (Kantardzic, 2020). Tujuan utama data mining cenderung *prediction* dan *description*, dimana prediksi lebih melibatkan penggunaan beberapa variable untuk memprediksi nilai yang tidak diketahui atau nilai masa depan, sedangkan deskriptif lebih mencari pola yang menggambarkan data yang dapat dianalisis oleh manusia.

Tujuan utama data mining *prediction* adalah untuk menghasilkan suatu model, yang dinyatakan dalam *code* yang dapat di eksekusi, yang digunakan dalam melakukan klasifikasi, prediksi, atau tugas lainnya yang serupa. Sedangkan *description* tujuannya untuk mendapatkan pemahaman tentang sistem yang dianalisis dengan mencari pola dan hubungannya dalam kupulaman data yang besar (Kantardzic, 2020)

Tugas data mining yang utama menurut (Kantardzic, 2020) sebagai berikut:

1. Klasifikasi, mengklasifikasikan item data ke salah satu kelas dari beberapa kelas yang telah ada sebelumnya.
2. Regresi, memetakan item data ke variable nilai nyata.
3. Clustering, data berusaha diidentifikasi ke suatu kategori / cluster yang menggambarkan data.
4. Summarization, untuk menemukan deksripsi atau rincian ringkas untuk subset data.
5. Dependency, mencari model local yang menjelaskan dependensi signifikan antara variable dalam data.
6. Change and deviation detection, mencari perubahan yang paling signifikan dalam kumpulan data.

Terkadang data mining di sinonimkan dengan istilah *Knowledge Discovery in Database* (KDD), namun sebaliknya data mining merupakan salah satu dalam proses KDD, tahapan dalam proses KDD ada pada Gambar 2



Gambar 2 Tahapan KDD

Sumber: (Han, Kamber, & Pei, 2012)

Pada Gambar 2 adalah tahapan proses KDD, menurut (Han, Kamber, & Pei, 2012) dijelaskan sebagai berikut:

1. Data cleaning, tahapan ini menghilangkan data yang tidak konsisten dan menghilangkan *noise*
2. Data integration, tahapan yang menggabungkan data dari berbagai sumber data
3. Data Selection, tahapan ini memilih data yang relevant untuk tugas yang akan di analisis dari database
4. Data transformation, proses mengubah bentuk data yang tujuannya agar dapat di gunakan pada proses datamining
5. Data mining, merupakan proses penting yang menerapkan metode untuk mengekstrak pola data

6. Pattern Evaluation, untuk mentafsirkan pola informasi yang menarik untuk mewakili pengetahuan di dalam data.
7. Knowledge Presentation, dimana Teknik visualisasi dan representasi pengetahuan digunakan untuk menunjukkan pengetahuan yang sudah di dapatkan.

2.6 Analisis Sentimen

Sebelum tahun 2000 penelitian yang berkaitan tentang analisis sentimen hampir tidak ada penelitian baik dari komunitas *linguistic* atau komunitas NLP (*Neuro-Linguistic Programming*). Hal ini terjadi karena hampir tidak ada opini yang tercatat secara digital. Dengan pertumbuhan media sosial dan web secara pesat adalah awal mula dan pertumbuhan yang cepat analisis sentimen. Selama bertahun-tahun, sistem media sosial di web telah menyediakan layanan yang sangat baik, platform untuk memfasilitasi dan memungkinkan pengguna memberikan ulasannya. Dari ulasan blog ke youtube, facebook, dan twitter, orang-orang menggunakan platform ini secara antusias karena memungkinkan pengguna dapat dengan bebas dan nyaman menyuarakan pendapat dan ulasan mereka, mengkomunikasikan pandangan mereka tentang subjek apa pun (Liu, 2015).

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 (Liu, 2015). 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.

2.7 Text Mining

Text Mining terdiri dari algoritma inti yang memungkinkan untuk mendapatkan wawasan dan pengetahuan dari data masif (Zaki & JR, 2020). Text mining mengacu pada proses extraction pengetahuan yang di perlukan untuk membuat keputusan penting dari data tekstual (Jo, 2019).

Tugas text mining menurut Taeho Jo (2019) ada 3, yaitu:

1. Classification, didefinisikan sebagai proses untuk menentukan kategori atau beberapa kategori yang telah ditentukan untuk setiap item.
2. Clustering, didefinisikan sebagai proses segmentasi kelompok dari berbagai item ke dalam subkelompok yang serupa.
3. Association, didefinisikan sebagai proses ekstraksi aturan asosiasi dalam bentuk *if-then*.

2.8 Pra Proses Data

Pra Proses Data dimana text akan difilter dan dimanipulasi, filter pada text diperlukan untuk menghilangkan istilah yang tidak mengandung banyak konten atau informasi yang dianggap tidak diinginkan sebagai fitur vektor dokumen, sedangkan manipulasi teks dilakukan untuk menormalkan istilah (Hoffmann & Chisholm, 2016). Proses Pra Proses Data menurut Feldman dan Sanger pada (Fatmawati & Affandes, 2017) sebagai berikut:

1. CaseFolding, mengubah kata bentuk huruf besar menjadi bentuk huruf kecil
2. Pembersihan Data, menghapus noise seperti karakter yang tidak dibutuhkan, tanda baca, angka, dan sebagainya.
3. Normalization, mengubah kata yang tidak baku, atau tidak sesuai dengan penulisan sebenarnya.
4. Stemming, mengubah kata menjadi kata dasar sesuai dengan KBBI (Kamus Besar Bahasa Indonesia).
5. Stopword Removal, menghilangkan kata yang tidak penting / bermakna.
6. Tokenization, memisahkan kata pada kalimat menjadi kata individu.

2.9 Pembobotan Kata

Pembobotan kata (*term weighting*) adalah istilah yang menjelaskan proses menghitung dan menetapkan bobot pada setiap kata sebagai tingkat kepentingannya. TF-IDF (*Term Frequency-Inverse Document Frequency*) adalah skema pembobotan kata yang populer (Jo, 2019). 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 (Markus dan Andrew, 2016).

Rumus perhitungan TF-IDF (Zong, Xia, & Zhang, 2021) sebagai berikut:

- Term frequency (TF): bobot mewakili frekuensi istilah dalam dokumen saat ini. TF mengasumsikan bahwa semakin sering kata tersebut muncul maka semakin penting kata tersebut.

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

untuk beberapa kata berfrekuensi tinggi, seperti *stopword*, frekuensi absolut akan jauh lebih tinggi dari pada rata-rata, dan ini mempengaruhi representasi teks. Untuk menurunkan dampak tersebut menggunakan frekuensi istilah logaritmik

$$f_i = \log(tf_i + 1) \quad (2)$$

- Inverse Document Frequency (IDF): IDF adalah fitur *global statistical* yang mencerminkan pentingnya istilah di seluruh korpus. Document Frequency (DF) menunjukkan dokumen dari istilah tertentu dari korpus. Semakin tinggi DF maka semakin rendah jumlah informasi efektif yang dikandungnya.

$$IDF_i = \log\left(\frac{N}{df_i}\right) \quad (3)$$

keterangan:

N : jumlah dokumen teks

df_i : jumlah dokumen yang mengandung kata i

- *Term Frequency – Inverted Document Frequency* (TF-IDF): TF-IDF adalah hasil pembobotan kata dari perkalian TF dengan IDF.

$$TF_IDF_i = TF_i \cdot IDF_i \quad (4)$$

2.10 Klasifikasi

Klasifikasi mengacu pada tugas untuk memprediksi sebuah label kelas yang belum di beri kelas. (Mohammed dan Wagner, 2020). Klasifikasi adalah proses menempatkan objek tertentu ke dalam suatu set kategori, berdasarkan properti yang dimiliki (Gorunescu, 2011). Proses klasifikasi mempunyai 4 komponen dasar menurut Gorunescu (2011), yaitu:

1. *Class* (kelas), adalah variabel terikat dari model, yang merupakan variabel kategori yang mewakili “label” di letakan setelah proses klasifikasi.

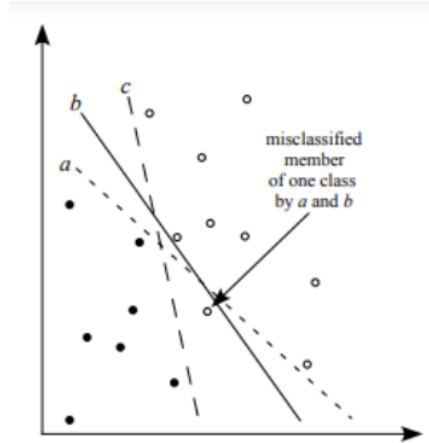
2. *Predictors* (prediktor), adalah variable bebas dari model yang merupakan karakteristik (atribut) data, yang akan diklasifikasikan dan berdasarkan klasifikasi yang dibuat.
3. *Training set* (data latih), adalah kumpulan data yang berisi nilai untuk 2 komponen sebelumnya yaitu (*class* dan *predictors*) yang digunakan untuk “*training*” model untuk mengenali kelas yang sesuai berdasarkan prediktor yang tersedia.
4. *Testing set* (data uji), adalah data baru yang akan di klasifikasi oleh model yang sudah dibuat menggunakan data latih, dan akurasi klasifikasi dapat di evaluasi.

2.11 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 (Zaki & JR, 2020). Pada tahun 1963 dasar SVM telah dikembangkan oleh matematikawan asal Rusia (Soviet) oleh Vladimir Vapnik, Colleagues Bernhard Boser, dan Isabelle Guyon, dengan menerbitkan sebuah paper dengan judul “*I’ve heard that SVMs have attracted a great deal of attention lately. Why?*”. (Han, Kamber, & Pei, 2012) (Zizka, Darena, & Svoboda, 2020). Terdapat permasalahan dalam menentukan hyperplane dalam pengklasifikasi, sebagai berikut:

2.11.1 Menemukan hyperplane pemisah yang optimal

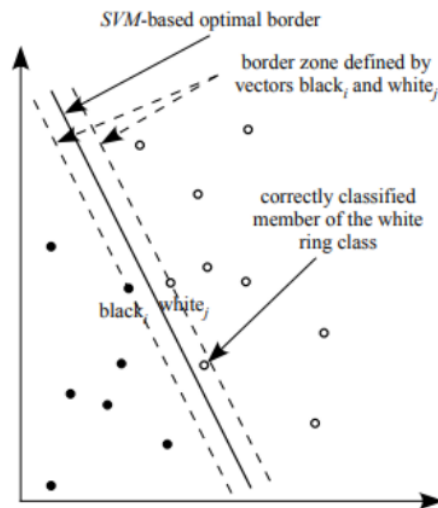
SVM adalah algoritma klasifikasi yang bertujuan untuk memisahkan 2 kelas menggunakan hyperplane – pengklasifikasi linier (Zizka, Darena, & Svoboda, 2020). Dalam menempatkan hyperplane terdapat beberapa cara yang dapat memisahkan elemen dari 2 kelas. Masalahnya adalah jika pilihan tersebut ada maka, seseorang dapat membuat banyak batasan seperti Gambar 3.



Gambar 3 Beberapa hyperlane yang dapat salah membagi kelas

Sumber: (Zizka, Darena, & Svoboda, 2020)

Solusi yang ditawarkan oleh algoritma SVM dapat dilihat pada Gambar 4. SVM akan mencari *gap* terbesar antar 2 kelas, selanjutnya akan dibuat bidang batas hyperplane optimum yang ditempatkan pada pusat zona.



Gambar 4 Optimal hyperplane menggunakan algoritma SVM

Sumber: (Zizka, Darena, & Svoboda, 2020)

Persamaan *hyperplane* sebagai berikut:

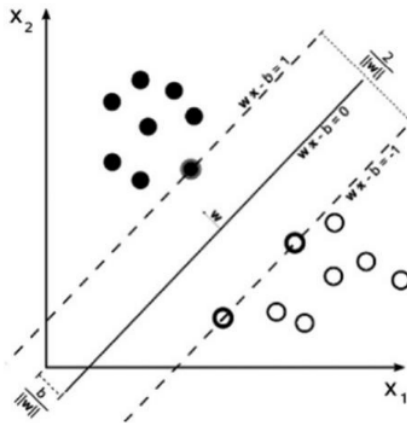
$$f(x) = w^T x + b \quad (5)$$

keterangan:

x = input vector

w = berat vektor

b = bias



Gambar 5 SVM 2 kelas yang menggambarkan hyperplane dan support vector

Sumber: (Sarkar, 2016)

Pada Gambar 5 diasumsikan bahwa 2 kelas objek terpisah sempurna oleh hyperplane didefinisikan sebagai berikut:

$$w \cdot x + b = 0 \quad (6)$$

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

$$w \cdot x + b \geq 1 \quad (7)$$

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

$$w \cdot x + b \leq -1 \quad (8)$$

Dalam menentukan nilai w dapat menggunakan persamaan sebagai berikut:

$$w = \sum_{i=1}^n \alpha_i y_i x_i \quad (9)$$

Dalam menentukan nilai w kita membutuhkan nilai α , untuk mencari nilai α kita dapat menggunakan *lagrange multiplier*.

$$\max_{\alpha} L_d = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i x_j) \quad (10)$$

Dalam menentukan nilai b dapat menggunakan persamaan sebagai berikut:

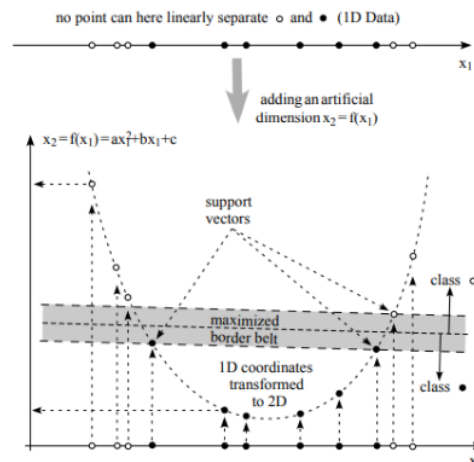
$$b = -\frac{1}{2} (x_i^+ \cdot w + x_i^- \cdot w) \quad (11)$$

Dalam penentuan kelas dapat menggunakan persamaan dibawah ini:

$$f(x) = \text{sign}[\sum_{i=1}^m \alpha_i y_i K(x_i x_j) + b] \quad (12)$$

2.11.2 Klasifikasi non-Linier dan fungsi kernel

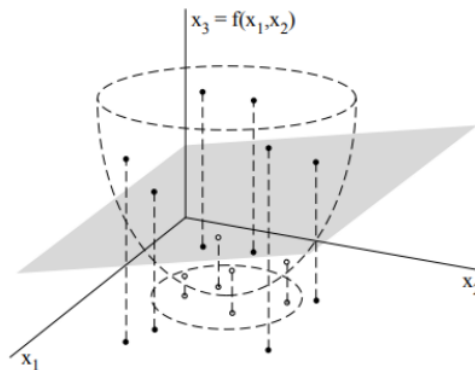
Masalah kedua adalah seringkali batas linier tidak mencakupi tugas secara inheren memerlukan batas nonlinier, hal ini menyebabkan kesulitan dalam menentukan bentuk batas non linier. Masalah yang berkaitan dengan klasifikasi nonlinier ini diselesaikan dengan apa yang disebut transformasi vektor berdasarkan fungsi kernel (Zizka, Darena, & Svoboda, 2020). Proses transformasi diilustrasikan pada Gambar 6 (dari ruang satu dimensi ke dua dimensi) dan Gambar 7 (dari ruang dua dimensi ke tiga dimensi).



Gambar 6 Ilustrasi transformasi dari 1D ke 2D

Sumber: (Zizka, Darena, & Svoboda, 2020)

Dalam ruang 1 dimensi, tidak dapat untuk memisahkan lingkaran kelas hitam dan kelas putih menggunakan hyperplane pada sumbu x_1 . Hanya setelah dilakukan transformasi menjadi ruang 2 dimensi, di mana dibuat koordinat baru x_2 dibuat sebagai fungsi x_1 , $x_2 = f(x_1)$, dan titik-titik di distribusikan pada parabola (fungsi kernel), hyperplane yang memisahkan antar 2 kelas dapat di temukan lihat garis warna abu-abu pada Gambar 7. Demikian juga untuk transformasi dari dua dimensi (satu kelas dibatasi oleh oval) menjadi tiga dimensi dengan dimensi buatan baru $x_3 = f(x_1, x_2)$, di mana lingkaran putih dan hitam ditempatkan pada permukaan paraboloid (fungsi kernel) Gambar 7.



Gambar 7 Ilustrasi transformasi dari 2D ke 3D

Sumber: (Zizka et al., 2020)

Pemisahan kelas yang tidak bisa dilakukan secara linier, dapat dilakukan oleh fungsi kernel, dengan mentransformasi vektor ruang dimensi. intinya tidak perlu mengetahui representasi dari Φ fungsi transformasi, tetapi cukup untuk menghitung skalar dari Φ fungsi yang sudah diterapkan pada vektor transformasi, menggunakan fungsi kernel K :

$$K(X_i, X_j) = \Phi(X_i) \cdot \Phi(X_j) \quad (13)$$

Sebagai contoh untuk transformasi ke fungsi kernel:

84
Fungsi kernel Linear:

$$K(X_i, X_j) = X_i \cdot X_j \quad (14)$$

Fungsi kernel Polynomial derajat h :

$$K(X_i, X_j) = (X_i \cdot X_j + 1)^h \quad (15)$$

Fungsi kernel Gaussian radial basis :

$$K(X_i, X_j) = e^{-\|X_i - X_j\|^2 / 2\sigma^2} \quad (16)$$

Fungsi kernel Sigmoid:

$$K(X_i, X_j) = \tanh(\kappa X_i \cdot X_j - \delta) \quad (17)$$

2.11.3 Multiclass SVM Klasifikasi

Pada penjelasan sebelumnya dijelaskan bahwa SVM diformulasi untuk 2 masalah kelas saja. Untuk mengatasi masalah *multiclass* dapat diselesaikan dengan metode SVM sebagai berikut:

1) One – Against – All

One against all adalah algoritma yang membuat klasifikasi k . untuk setiap kelas i . SVM akan melabelkan menjadi positif dan sisa kelas lainnya menjadi negatif (Stoean & Stoean, 2014). Keputusan untuk setiap kelas i untuk menentukan koefisien optimal w dan b dari hyperlane untuk memisahkan sampel dengan dengan hasil kelas i dengan kelas lainnya, maka (13)

$$\begin{cases} \text{Cari } w^i \text{ dan } b^i \text{ untuk meminimkan } \frac{\|w^i\|^2}{2} + C \sum_{j=1}^m \xi_j^i \\ \text{Subject untuk } y_j (w^i \cdot x_j - b^i) \geq 1 - \xi_j^i, \xi_j^i \geq 0, \text{ untuk semua } j = 1, 2, \dots, m. \end{cases} \quad (18)$$

Setelah semua hyperplan sudah ditentukan, kelas untuk sampel uji x diberikan kategori yang memiliki nilai maksimum untuk fungsi *training* lihat persamaan (4)

$$\text{Class}(x) = \operatorname{argmax}_{i=1, 2, \dots, k} (w^i \cdot \Phi(x)) - b^i \quad (19)$$

2) One – Against – One

One against one adalah Teknik membuat hyperplan svm, setiap kelas pada data latih. setiap 2 kelas, kelas i dan j , dimana kelas i diberi label positive dan kelas j di beri label negative (Stoean & Stoean, 2014).

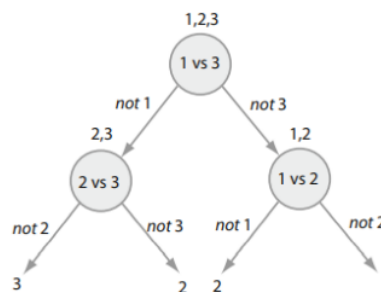
Untuk menentukan koefisien optimal dari hyperplane untuk memisahkan sample i dan j , maka (15)

$$\left\{ \begin{array}{l} \text{Cari } w^{ij} \text{ dan } b^{ij} \text{ untuk meminimalkan } \frac{\|w^{ij}\|^2}{2} + C \sum_{j=1}^m \xi_l^{ij} \\ \text{Subject untuk } y_i (w^{ij} \cdot x_j - b^{ij}) - 1 - \xi_l^{ij}, \xi_l^{ij} \geq 0, \text{ untuk semua } l = 1, 2, \dots, m. \end{array} \right. \quad (20)$$

Ketika hyperplane sudah ditemukan, metode voting digunakan untuk menentukan kelas untuk sampel uji x^* . Untuk setiap SVM, kelas x^* dihitung dengan mengikuti tanda dari hasil fungsi keputusan yang diterapkan pada x^* . Selanjutnya jika tanda menyatakan x^* ada di kelas i , perolehan suara untuk kelas i bertambah satu; sebaliknya, suara akan bertambah ke kelas j . Akhirnya x^* diambil sebagai kelas dengan suara terbesar. Jika 2 kelas memiliki jumlah suara yang sama maka dipilih berdasarkan indeks yang kecil.

3) Decision Directed Acyclic Graph

Klasifikasi dengan Teknik *Decision Directed Acyclic Graph* dilakukan sama dengan klasifikasi *One – Against – One* (Stoean & Stoean, 2014). Setelah mendapatkan hyperplane, sistem grafik Gambar 8 digunakan untuk menentukan kelas sampel uji X .



Gambar 8 Contoh masalah label 3 kelas menggunakan Decision Directed Acyclic Graph

Sumber: (Zizka, Darena, & Svoboda, 2020)

Simpul akar sesuai dengan semua kelas k . Ketika contoh uji x dievaluasi, satu turun dari simpul ke simpul, dengan kata lain, menghilangkan satu kelas dari setiap daftar yang sesuai, sampai daun tercapai.

2.12 Evaluasi Model

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 (Gorunescu, 2011). Pada penelitian ini evaluasi model yang digunakan adalah *confusion matrix*. *Confusion matrix* sering disebut juga klasifikasi matrik, ini digunakan untuk mengukur akurasi prediksi model (Kantardzic, 2020). *Confusion matrix* untuk masalah klasifikasi biner dan *multiclass* dilihat pada Tabel 1 dan Tabel 2.

Tabel 1 Confusion Matrix untuk 2 Kelas

		<i>Predicted Class</i>	
		<i>Positive (c₁)</i>	<i>Negative (c₂)</i>
<i>Actual Class</i>	<i>Positive (c₁)</i>	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
	<i>Negative (c₂)</i>	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

Tabel 2 Confusion Matrix untuk Multiclass

		<i>Predicted Class</i>			
		<i>C₁</i>	<i>C₂</i>	...	<i>C_N</i>
<i>Actual Class</i>	<i>C₁</i>	<i>C_{1,1}</i>	<i>FP</i>	...	<i>C_{1,N}</i>
	<i>C₂</i>	<i>FN</i>	<i>TP</i>	...	<i>FN</i>

	<i>C_N</i>	<i>C_{N,1}</i>	<i>FP</i>	...	<i>C_{N,N}</i>

Keterangan:

True Positive (TP) : jumlah benar prediksi oleh pengklasifikasi dengan benar sebagai positive

False Positive (FP) : jumlah prediksi oleh pengklasifikasi menjadi positif yang sebenarnya negative

False Negative (FN) : jumlah prediksi oleh pengklasifikasi negatif yang sebenarnya kelas positif

True Negative (TN) : jumlah prediksi oleh pengklasifikasi dengan benar sebagai negative

Setelah mendapatkan nilai yang diperlukan dari confusion matrix, kita dapat menghitung 4 performa matrik satu per satu (*Accuracy, Precision, Recall, F-1 Score*):

1. **Accuracy** didefinisikan sebagai akurasi dari seluruh prediksi yang benar pada setiap model (Sarkar, 2016).
2. **Precision** didefinisikan sebagai jumlah prediksi yang dibuat yang benar dari semua kelas positif (Sarkar, 2016).
3. **Recall** didefinisikan sebagai jumlah kasus dari kelas positif yang di prediksi benar (Sarkar, 2016).
4. **F-1 Score** didefinisikan sebagai ukuran akurasi lain yang dihitung dengan mengambil *mean harmonic* dari presisi dan recall (Sarkar, 2016).

Matrik kinerja *confusion matrix* untuk klasifikasi biner

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (21)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (22)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (23)$$

$$\text{F-1 score} = 2 \cdot \frac{\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}} \quad (24)$$

Matrik Kinerja *confusion matrix* untuk *multiclass*

$$\text{Akurasi} = \frac{\sum_{i=1}^N TP(C_i)}{\sum_{i=1}^N \sum_{j=1}^N C_{ij}} \quad (25)$$

$$\text{Precision} = \frac{TP(C_i)}{TP(C_i) + FP(C_i)} \quad (26)$$

$$\text{Recall} = \frac{TP(C_i)}{TP(C_i) + FN(C_i)} \quad (27)$$

$$\text{F-1 score} = 2 \cdot \frac{\text{Recall}(C_i) \cdot \text{Precision}(C_i)}{\text{Recall}(C_i) + \text{Precision}(C_i)} \quad (28)$$

2.13 Penelitian Terdahulu

Penelitian yang dilakukan oleh (Rohanah, Dermawan, & Purnamasari, 2021) dengan judul "Klasifikasi Ulasan Pengguna Zoom Cloud Meetings Menggunakan Metode Information Gain dan Naïve Bayes Classifier". Tujuan dari penelitian ini adalah mengklasifikasi ulasan pengguna berdasarkan faktor kualitas *software* ISO/IEC 9126 untuk mengetahui keluhan pengguna yang ada pada ulasan *google playstore*. Pada penelitian ini, menggunakan metode Naïve Bayes Classifier dan Information Gain, dengan akurasi terbaik sebesar 79% dan pengujian k-fold cross validation 80,51%.

Penelitian yang dilakukan oleh (Ilmawan & Mude, 2020) dengan judul "Perbandingan Metode Klasifikasi *Support Vector Machine* dan *Naïve Bayes* untuk Analisis Sentimen pada Ulasan Tekstual di *Google Play Store*". Pada penelitian ini, bertujuan untuk membandingkan antara algoritma SVM dan Naïve bayes dalam mengklasifikasi analisis sentimen di *google playstore*. Hasil akurasi pada SVM classifier mendapatkan akurasi sebesar 81,46% dan Naïve Bayes classifier sebesar 75,41%.

Penelitian yang dilakukan oleh (Rohanah, Rianti, & Sari, 2021) dengan judul "Perbandingan Naïve Bayes dan Support Vector Machine untuk Klasifikasi Ulasan Pelanggan Indihome". Tujuan dari penelitian ini adalah membandingkan algoritma antara *naïve bayes* dengan *Support Vector Machine* (SVM). Hasil dari perbandingan algoritma dalam penelitian ini adalah SVM memiliki akurasi lebih baik dari pada *Naïve bayes* dengan akurasi secara beurut sebesar 82,11% dan 77,03%.

Penelitian yang dilakukan oleh (Fransiska, Rianto, & Gufroni, 2020) dengan judul "Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method". Tujuan dari penelitian ini adalah untuk mengetahui pandangan pengguna mengenai pelayanan yang

dilakukan oleh provider by.U. Pada penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) dan TF-IDF sebagai *feature extraction*. Dengan akurasi sebesar 84,7%, precision 84,9%, recall 84,7%, dan f-measure sebesar 84,8%.

Penelitian yang dilakukan oleh (Hendra & Fitriyani, 2021) dengan judul “Analisis Sentimen Review Halodoc Menggunakan *Nai’ve Bayes Classifier*”. Tujuan dari penelitian ini adalah untuk mengetahui berapa besar akurasi pada ulasan HaloDoc dan mengetahui kinerja algoritma *naive bayes* pada klasifikasi *sentiment analysis* Halodoc. Hasil akurasi dari penelitian ini sebesar 81,68%.

Penelitian yang dilakukan oleh (Alita, Fernando, & Sulistiani, 2020) dengan judul “Implementasi Algoritma *Multiclass SVM* pada Opini Publik Berbahasa Indonesia di Twitter”. Tujuan dari penelitian ini adalah untuk mengetahui performa pada proses analisis opini Bahasa Indonesia yang memiliki 3 kelas positif, negatif, dan netral dari metode SVM yang telah di optimasi dan diterapkan. Hasil dari penelitian ini adalah klasifikasi yang menggunakan lebih dari 2 kelas dapat dilakukan klasifikasi menggunakan metode SVM dengan metode *SVM One Against One* dan *One Against All*, dimana *One Against All* memiliki nilai akurasi lebih baik dengan perbedaan 0,06.

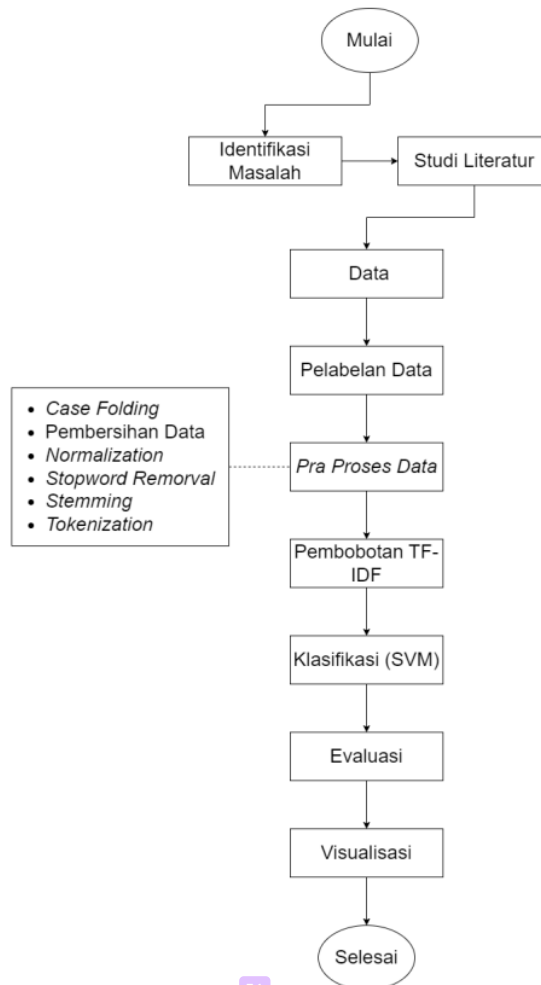
Penelitian yang dilakukan oleh (Rohanah, Dermawan, & Purnamasari, 2021) mengenai klasifikasi pengguna memiliki kemiripan yang sama yaitu fokus penelitian yang sama menggunakan ISO/IEC 9126 sebagai kelas kategori dalam klasifikasi pengguna, selain itu penelitian yang dilakukan oleh (Ilmawan & Mude, 2020) dan (Rohanah, Rianti, & Sari, 2021) berfokus pada pengujian algoritma pada klasifikasi *sentiment analysis*, hasil dari penelitian mempengaruhi dalam menentukan metode yang dilakukan dalam penelitian ini. Penelitian yang dilakukan oleh (Fransiska, Rianto, & Gufroni, 2020) memiliki kemiripan yang sama yaitu menggunakan metode dan *feature extraction* yang sama yaitu SVM dan TF-IDF. Dan penelitian yang dilakukan oleh (Hendra & Fitriyani, 2021) memiliki kemiripan yang sama yaitu memiliki fokus objek penelitian yang sama (aplikasi HaloDoc). Serta penelitian yang dilakukan oleh (Alita, Fernando, & Sulistiani, 2020) menjelaskan bahwa metode SVM dapat digunakan untuk klasifikasi lebih dari 2 kelas dengan menggunakan *one against all* dengan akurasi lebih baik dari *one*

against one. Jadi dalam penelitian ini akan dilakukan klasifikasi ulasan pengguna berdasarkan ISO 9126 dan sentimen positif dan negatif pada aplikasi HaloDoc dengan metode SVM dan TF-IDF sebagai *feature extraction* serta untuk klasifikasi *multiclass* akan menggunakan metode SVM *one against all*.

BAB III METODELOGI PENELITIAN

3.1 Tahapan Penelitian

Demi mencapai tujuan dari penelitian, penelitian ini akan mengikuti tahapan alur kerja yang telah dibuat. Berikut beberapa tahapan proses penelitian yang akan dilakukan pada Gambar 9 sebagai berikut.



Gambar 9 Tahapan Penelitian

3.1.1 Identifikasi Masalah

Identifikasi masalah merupakan tahapan untuk mengidentifikasi masalah yang akan diangkat pada penelitian ini, yakni klasifikasi ulasan pengguna pada aplikasi halodoc yang bertujuan membantu pengembang dalam memperbaiki atau meningkatkan kinerja aplikasi halodoc.

3.1.2 Studi Literatur

Studi literatur digunakan sebagai bahan/ sumber dalam penelitian ini, untuk mendukung dalam memecahkan permasalahan dalam penelitian sebagai sumber pustaka. Sumber yang digunakan berasal dari *e-book dan jurnal yang berkaitan dengan* permasalahan pada *data mining, software quality, analisis sentimen, preprocessing, dan klasifikasi menggunakan support vector machine.*

3.1.3 Data

Proses pengumpulan data menggunakan google scrapper, data yang diambil berasal dari ulasan aplikasi HaloDoc versi 10 yang ada pada *google playstore* dengan data sejumlah 880 record, yang diambil dari tanggal 20 Juli 2021 sampai 12 November 2021 menggunakan Bahasa Indonesia dan disimpan dalam bentuk *.csv*.

3.1.4 Pelabelan Data

Pelabelan data dilakukan sebelum preprocessing yang berguna untuk mempertahankan nilai asli dari data tersebut. Pelabelan dilakukan secara manual terdiri dari 2 kelas sentimen (positif dan negatif) dan 6 kelas kategori berdasarkan ISO 9126 (*functionality, reliability, usability, efficiency, maintainability dan, portability*).

3.1.5 Pra Proses Data

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

1. *Case Folding*, mengubah semua teks huruf besar yang ada pada kalimat menjadi huruf kecil.
2. Pembersihan Data, tahapan ini menghapus karakter yang tidak dibutuhkan, seperti tanda baca, angka, dan sebagainya.
3. *Normalization*, mengubah kata yang tidak baku, disingkat, atau salah eja menjadi sesuai dengan kamus KBBI (Kamus Besar Bahasa Indonesia)
4. *Stemming*, mengubah kata menjadi bentuk kata dasar sesuai dengan KBBI (Kamus Besar Bahasa Indonesia).
5. *Stopword Removal*, menghapus kata yang tidak bermakna, menggunakan *stopword* sastrawi.
6. *Tokenization*, kalimat yang terdapat pada ulasan di pecah menjadi kata per kata dengan spasi sebagai pemisahannya

3.1.6 Pembobotan TF-IDF

Sebelum pembobotan kata menggunakan TF-IDF, data terlebih dahulu dibagi menjadi dua yaitu data training (80%) dan data testing (20%), penggunaan rasio 80:20 karena rasio ini sering digunakan pada berbagai bidang kehidupan (Koch, 2013).

Pada penelitian ini pembobotan kata menggunakan TF-IDF dimana persamaan dari TF-IDF dapat dilihat dengan menggunakan persamaan (4). Nilai pembobotan TF-IDF ukuran statistik dari perkalian antara nilai TF dengan nilai IDF yang menggambarkan pentingnya istilah tersebut.

3.1.7 Klasifikasi

Data training digunakan untuk membuat model klasifikasi menggunakan algoritma SVM, dan data testing digunakan untuk mengimplementasi model yang sudah dibuat menggunakan algoritma SVM dengan melakukan prediksi label.

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 (18) sebagai solusi untuk penyelesaian klasifikasi untuk model kelas kategori.

3.1.8 Evaluasi

Pada tahapan ini dilakukan evaluasi untuk mengetahui kinerja dari model klasifikasi yang sudah kita buat menggunakan *confusion matrix*, pada *confusion matrix* kita bisa mengetahui nilai akurasi, *precision*, *recall* dan *f1 score* dari model yang sudah dibangun dengan persamaan (21) untuk akurasi, (22) untuk *precision*, (23) untuk *recall*, dan (24) untuk *f1 score*. Untuk *confusion matrix multiclass* kita dapat menggunakan persamaan (25) untuk akurasi, (26) untuk *precision*, (27) untuk *recall*, dan (28) untuk *f1 score*.

3.1.9 Visualisasi

Tahapan ini menampilkan informasi yang sudah di dapatkan dan ditampilkan dalam bentuk gambar dengan maksud mempermudah dalam mengerti informasi yang terkandung. Data akan dibentuk secara visualisasi berupa gambar *wordcloud* pada setiap kelas.

3.2 Alat yang digunakan

Penelitian ini, menggunakan alat bantu yang terdiri dari perangkat keras dan lunak, spesifikasi sebagai berikut:

1. Perangkat Keras

Spesifikasi perangkat keras sebagai berikut:

- 1) *Processor* : AMD Ryzen 5 2600
- 2) *Motherboard* : B450M Pro-VDH Max (MS-7A38)
- 3) *Memory* : DDR4 1600MHz 16 GB
- 4) *Graphic* : NVIDIA GeForce GTX 1650 Super
- 5) *Monitor* : 22" 16:9 IPS (1920 x 1080) LED
- 6) *SSD* : 250GB 5400RPM

2. Perangkat Lunak

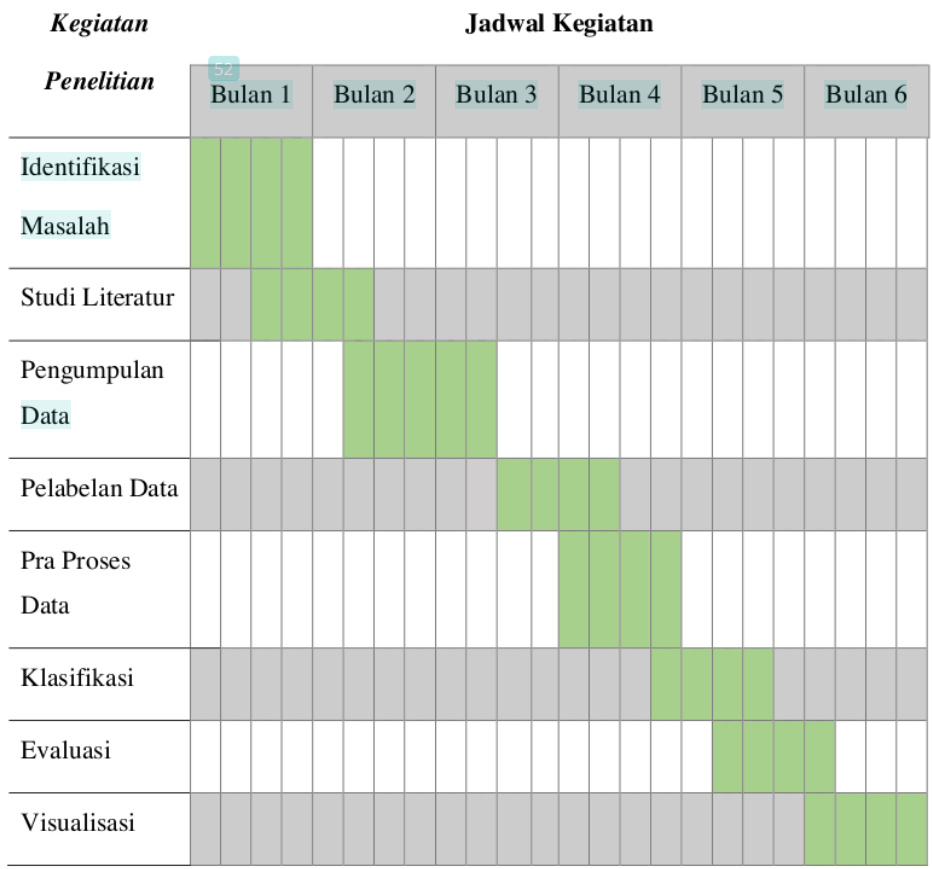
Perangkat lunak yang digunakan sebagai berikut:

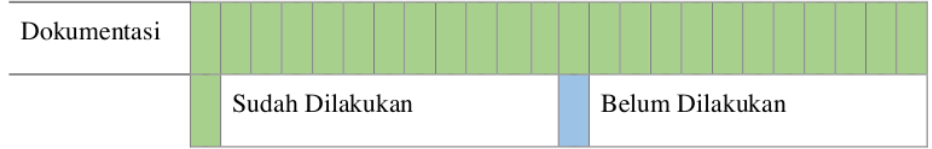
- 1) *Sistem Operasi* : Microsoft® Windows 10 Professional

- 2) Perangkat Lunak
 - a) Microsoft word 2019
 - b) Microsoft Excel 2019
 - c) Google Collab
 - d) Draw io
 - e) Google Sheets

3.3 Jadwal Penelitian

Penelitian ini dimulai dari mengidentifikasi masalah, pemahaman terhadap studi pustaka mengenai penelitian terkait, melakukan pengumpulan data dan pelabelan data, selanjutnya dilakukan preprocessing untuk memaksimalkan pengklasifikasian, lalu melakukan klasifikasi menggunakan SVM yang selanjutnya akan dievaluasi menggunakan *confusion matrix*.





BAB IV HASIL DAN PEMBAHASAN

4.1 Data

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*. Berikut deskripsi dari proses penarikan data dengan *googlescraper* pada Tabel 3.

Tabel 3 Luaran dari Googlescraper

Variable	Deskripsi
<i>ReviewID</i>	Data unik dimana setiap review memiliki id yang berbeda beda
<i>userName</i>	Nama yang mengidentifikasi pengguna akun google
<i>userImage</i>	Gambar yang mengidentifikasi pengguna akun google
<i>Content</i>	Review atau ulasan pengguna
<i>Score</i>	Merupakan bintang atau penilaian pengguna terhadap aplikasi berisi score dari 1- 5
<i>thumbsUp</i>	Jumlah <i>likes</i> yang diberikan oleh pengguna lain terhadap suatu ulasan
<i>reviewCreatedVersion</i>	Versi aplikasi yang diulas
<i>At</i>	Tanggal dan waktu pembuatan ulasan pengguna

Variable	Deskripsi
<i>replyContent</i>	Berisi balasan dari pihak aplikasi terhadap ulasan yang telah diberikan
<i>repliedAt</i>	Tanggal dan waktu <i>reply</i> ulasan dari pihak aplikasi

Variabel yang digunakan pada penelitian ini hanya konten saja, karena variabel konten berisi ulasan para pengguna yang nantinya akan kita gunakan untuk proses klasifikasi sentimen dan kategori berdasarkan ulasan pengguna.

4.2 Pelabelan Data

Tahap pelabelan dilakukan terhadap 880 data ulasan. Berikut dibawah ini pada Tabel 4 dan Tabel 5 adalah contoh pelabelan yang dilakukan pada 880 data ulasan tadi.

Tabel 4 Contoh dataset kelas sentimen

Ulasan	Sentimen
“Aplikasi Gak jelas, sudah buat janji tapi di undur terus”	Negatif
“Aplikasi mudah di mengerti”	Positif

Pada kolom ulasan baris pertama adalah contoh untuk sentimen negatif karena pengguna mengeluh mengenai pesannya yang belum di terima dan susah nya menghubungi *customer service*, sedangkan untuk yang baris kedua contoh untuk sentimen positif dimana pengguna puas saat menggunakan aplikasi halodoc versi 10.

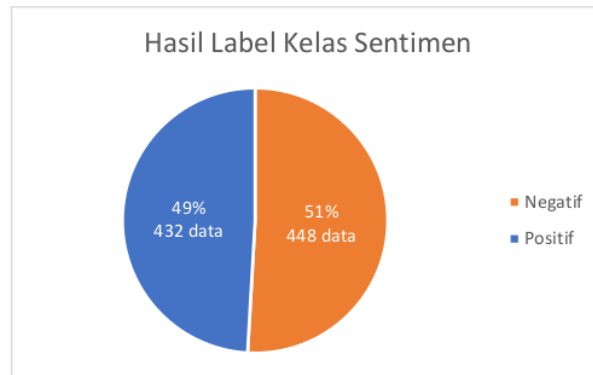
Tabel 5 Contoh dataset kelas kategori

Ulasan	Kategori
“Bagus banget abis bayar ada hitungan waktu menunggu konsultasi dg dokter.”	<i>Functionality</i>
“Perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan”	<i>Usability</i>
“Eror lebih dari 10 menit pas janji ama dokter.”	<i>Reliability</i>
“Barang sdh diterima dgn baik sesuai order, pengiriman cukup cepat, tks Halodoc”	<i>Efficiency</i>
“Sdh seminggu ini aplikasinya kok g bs dipakai yaa.. Sdh saya update sampe download ulang tetep bgitu.. kenapa yaa..”	<i>Maintainability</i>
“Membantu banget, apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes. Terima kasih Halodoc”	<i>Portability</i>

Baris pertama pada kolom ulasan adalah contoh untuk kategori *functionality* karena ulasan tersebut menyinggung fitur yang ada pada aplikasi, untuk baris kedua adalah kategori *usability* karena ulasan tersebut menunjukkan kemudahan pengoperasian aplikasi, selanjutnya baris ketiga adalah kategori *reliability* karena menunjukkan kemampuan perangkat lunak dalam mempertahankan kinerjanya dalam kondisi atau waktu yang sudah ditetapkan.

Baris keempat adalah kelas *efficiency* karena menyangkut waktu eksekusi Ketika perangkat lunak dijalankan dan memberikan respon yang sesuai saat melakukan fungsinya, baris kelima adalah kelas *maintainability* karena mencangkup perbaikan, pembaruan, atau kesalahan setelah perubahan aplikasi. Dan baris keenam adalah kelas *portability* karena menyangkut kemampuan aplikasi pada saat digunakan pada situasi dan kondisi lingkungan yang berbeda.

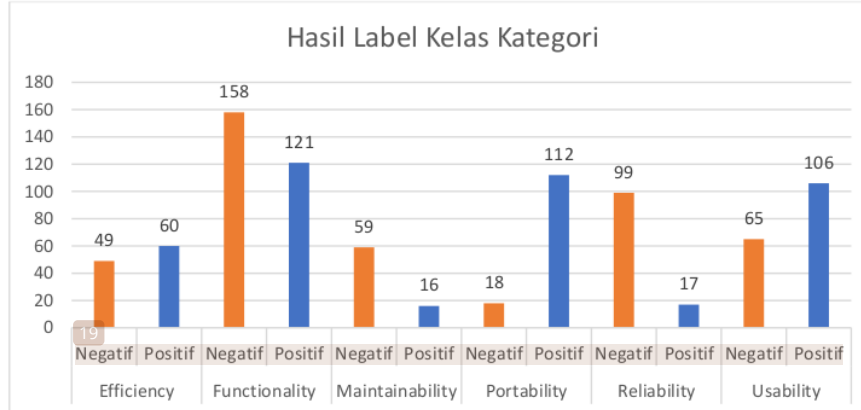
Pada pelabelan data, ulasan yang sudah didapatkan akan dilakukan pelabelan kelas sentimen dan kelas kategori secara manual. Pada pelabelan kelas sentimen akan di berikan nilai negatif dan positif, hasil dari pelabelan dapat dilihat pada Gambar 10.



Gambar 10 Grafik Hasil Pelabelan kelas Sentimen

Hasil dari pelabelan untuk kelas sentimen dapat dilihat pada Gambar 10, 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.

Setelah dilakukan pelabelan kelas sentimen (positif dan negatif), selanjutnya data akan di berikan label kelas kategori dimana nilai dari label kelas kategori ada (functionality, reliability, usability, efficiency, maintainability dan, portability), hasil dari pelabelan dapat dilihat pada Gambar 11.



Gambar 11 Hasil Pelabelan untuk Kelas Kategori

Pada Gambar 11, 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.

Urutan kelas kategori untuk ulasan positif terbanyak yaitu functionality, porability, usability, efficiency, reliability, dan mainability. Sedangkan untuk kelas kategori urutan kelas negatif terbanyak yaitu functionality, reliability, usability, mainability, efficiency, dan porability.

4.3 Pra Proses Data

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

4.3.1 Case Folding

Tahapan ini mengubah ulasan penggunaan menjadi *lowercase*, artinya mengubah dari huruf besar ke huruf kecil. Tabel 6 adalah contoh dari proses *case folding*.

Tabel 6 Hasil Output dari Proses *Case Folding*

Sebelum	Sesudah
Aplikasi Gak jelas, sudah buat janji tapi di undur terus	aplikasi gak jelas, sudah buat janji tapi di undur terus
Aplikasi mudah di mengerti	aplikasi mudah di mengerti
Bagus banget abis bayar ada hitungan waktu menunggu konsultasi dg dokter.	bagus banget abis bayar ada hitungan waktu menunggu konsultasi dg dokter.
Perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan	perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan
Eror lebih dari 10 menit pas janji ama dokter.	eror lebih dari 10 menit pas janji ama dokter.
Barang sdh diterima dgn baik sesuai order, pengiriman cukup cepat, tks Halodok	barang sdh diterima dgn baik sesuai order, pengiriman cukup cepat, tks halodok
Sdh seminggu ini aplikasinya kok g bs dipakai yaa.. Sdh saya update sampe download ulang tetep bgitu.. kenapa yaa..	sdh seminggu ini aplikasinya kok g bs dipakai yaa.. sdh saya update sampe download ulang tetep bgitu.. kenapa yaa..
Membabtu banget, apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes. Terima kasih Halodoc	membabtu banget, apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes. terima kasih halodoc

Pada Tabel 6, kolom sebelum adalah ulasan yang ada pada *google playstore*, ditahapan pertama ini adalah case folding, terlihat pada kolom

sesudah bahwa semua huruf besar pada kolom sebelum, diubah menjadi huruf kecil.

4.3.2 Pembersihan Data

Tahapan ini menghilangkan tanda – tanda atau simbol, tanda baca, dan angka yang terdapat pada ulasan. Tabel 7 adalah contoh dari proses pembersihan data.

Tabel 7 Hasil *Output* dari Proses Pembersihan Data

Sebelum	Sesudah
aplikasi gak jelas, sudah buat janji tapi di undur terus	aplikasi gak jelas sudah buat janji tapi di undur terus
aplikasi mudah di mengerti	aplikasi mudah di mengerti
bagus banget abis bayar ada hitungan waktu menunggu konsultasi dg dokter.	bagus banget abis bayar ada hitungan waktu menunggu konsultasi dg dokter
perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan	perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan
eror lebih dari 10 menit pas janji ama dokter.	eror lebih dari menit pas janji ama dokter
barang sdh diterima dgn baik sesuai order, pengiriman cukup cepat, tks halodok ÃƒfÃƒfÃƒ,ÃƒÃƒfÃƒ,Ãƒ,ÃƒÃƒfÃƒ,Ãƒ	barang sdh diterima dgn baik sesuai order pengiriman cukup cepat tks halodok
sdh seminggu ini aplikasinya kok g bs dipakai yaa.. sdh saya update	sdh seminggu ini aplikasinya kok g bs dipakai yaa sdh saya update

Sebelum	Sesudah
sampe download ulang tetep bgitu.. kenapa yaa..	sampe download ulang tetep bgitu kenapa yaa
membabtu banget, apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes. terima kasih halodoc	membabtu banget apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes terima kasih halodoc

Pada Tabel 7, pada kolom sebelum adalah hasil dari proses casefolding, setelah dilakukan *case folding* adalah pembersihan data terlihat pada kolom sesudah bahwa tanda huruf, angka atau simbol lainnya dihilangkan.

4.3.3 Normalization

Tahapan ini mengubah singkatan atau kata typo yang terdapat pada ulasan aplikasi yang sesuai dengan Kamus Besar Bahasa Indonesia (KBBI). Tabel 8 adalah contoh dari proses *normalization*.

Tabel 8 Hasil *Output* dari Proses *Normalization*

Sebelum	Sesudah
aplikasi gak jelas sudah buat janji tapi di undur terus	aplikasi tidak jelas sudah buat janji tapi di undur terus
aplikasi mudah di mengerti	aplikasi mudah di mengerti
bagus banget abis bayar ada hitungan waktu menunggu konsultasi dg dokter	bagus sangat habis bayar ada hitungan waktu menunggu konsultasi dengan dokter
perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan	perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan

Sebelum	Sesudah
error lebih dari menit pas janji ama dokter	error lebih dari menit pas janji sama dokter
barang sdh diterima dgn baik sesuai order pengiriman cukup cepat tks halodoc	barang sudah diterima dengan baik sesuai order pengiriman cukup cepat terima kasih halodoc
sdh seminggu ini aplikasinya kok g bs dipakai yaa sdh saya update sampe download ulang tetep bgitu kenapa yaa	sudah seminggu ini aplikasinya kok tidak bisa dipakai iya sudah saya update sampai download ulang tetap begitu kenapa iya
membantu banget apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes terima kasih halodoc	membantu sangat apalagi di masa pandemic atau sakit ringan tapi malas buat ke dokter karena badan lemas terima kasih halodoc

Tahapan selanjutnya adalah *normalization*, pada Tabel 8 kolom sebelum adalah hasil dari proses pembersihan data, hasil dari proses *normalization* ada pada kolom sesudah dimana kata yang disingkat, atau typo, akan diubah sesuai dengan KBBI.

4.3.4 Stemming

Tahapan dimana mengubah kata yang ada pada ulasan menjadi bentuk dasar sesuai dengan KBBI menggunakan library sastrawi. Tabel 9 adalah contoh dari proses Stemming.

Tabel 9 Hasil *Output* dari Proses *Stemming*

Sebelum	Sesudah
aplikasi tidak jelas sudah buat janji tapi di undur terus	aplikasi tidak jelas sudah buat janji tapi di undur terus

Sebelum	Sesudah
aplikasi mudah di mengerti	aplikasi mudah di erti
bagus sangat habis bayar ada hitungan waktu menunggu konsultasi dengan dokter	bagus sangat habis bayar ada hitung waktu tunggu konsultasi dengan dokter
perlu modifikasi tampilan menu supaya lebih memudahkan lagi dalam penggunaan	perlu modifikasi tampil menu supaya lebih mudah lagi dalam guna
error lebih dari menit pas janji sama dokter	error lebih dari menit pas janji sama dokter
barang sudah diterima dengan baik sesuai order pengiriman cukup cepat terima kasih halodoc	barang sudah terima dengan baik sesuai order kirim cukup cepat terima kasih halodoc
sudah seminggu ini aplikasinya kok tidak bisa dipakai iya sudah saya update sampai download ulang tetap begitu kenapa iya	sudah minggu ini aplikasi kok tidak bisa pakai iya sudah saya update sampai download ulang tetap begitu kenapa iya
membantu sangat apalagi di masa pandemic atau sakit ringan tapi malas buat ke dokter karena badan lemas terima kasih halodoc	bantu sangat apalagi di masa pandemic atau sakit ringan tapi malas buat ke dokter karena badan lemas terima kasih halodoc

Dapat dilihat dari Tabel 9, pada kolom sebelum adalah hasil dari proses *normalization*, kolom sesudah adalah hasil dari proses *stemming* dimana kata akan diubah menjadi kata dasarnya sesuai dengan KBBI.

4.3.5 Stopword Removal

44

Tahapan ini menghapus atau menghilangkan kata pada ulasan yang tidak memiliki makna seperti kata hubung dan kata yang sering muncul. Tabel 10 adalah contoh dari proses *Stopword Removal*.

Tabel 10 Hasil *Output* dari Proses *Stopword Removal*

Sebelum	Sesudah
aplikasi tidak jelas sudah buat janji tapi di undur terus	jelas janji undur
aplikasi mudah di erti	mudah erti
bagus sangat habis bayar ada hitung waktu tunggu konsultasi dengan dokter	bagus habis bayar hitung tunggu konsultasi
perlu modifikasi tampil menu supaya lebih mudah lagi dalam guna	modifikasi tampil menu mudah
error lebih dari menit pas janji sama dokter	error janji
barang sudah terima dengan baik sesuai order kirim cukup cepat terima kasih halodoc	barang sesuai order kirim cepat
sudah minggu ini aplikasi kok tidak bisa pakai iya sudah saya update sampai download ulang tetap begitu kenapa iya	update download ulang
bantu sangat apalagi di masa pandemic atau sakit ringan tapi	bantu pandemic sakit ringan malas badan lemas

Sebelum	Sesudah
malas buat ke dokter karena badan lemas terima kasih halodoc	

Dari Tabel 10, pada kolom sebelum adalah hasil dari proses *stemming*, kolom sesudah adalah hasil dari proses *stopword removal* dimana kata yang tidak bermakna seperti kata hubung dan kata yang sering muncul akan dihapus.

4.3.6 Tokenization

Tahapan ini adalah memisahkan atau memecah kalimat pada ulasan menjadi kata per kata. Tabel 11 adalah contoh dari proses dari *Tokenization*.

Tabel 11 Hasil *Output* dari Proses *Tokenization*

Sebelum	Sesudah
jelas janji undur	['jelas', 'janji', 'undur']
mudah erti	['mudah', 'erti']
bagus habis bayar hitung tunggu konsultasi	['bagus', 'habis', 'bayar', 'hitung', 'tunggu', 'konsultasi']
modifikasi tampil menu mudah	['modifikasi', 'tampil', 'menu', 'mudah']
error janji	['error', 'janji']
barang sesuai order kirim cepat	['barang', 'sesuai', 'order', 'kirim', 'cepat']
update download ulang	['update', 'download', 'ulang']
bantu pandemic sakit ringan malas badan lemas	['bantu', 'pandemic', 'sakit', 'ringan', 'malas', 'badan', 'lemas']

Pada kolom sebelum pada Tabel 11 adalah hasil dari proses *stopword removal*, kolom sesudah adalah hasil dari proses *tokenization* dimana setiap kalimat akan dipecah menjadi kata perkata.

4.4 Pembobotan TF-IDF

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 pembobotan TF-IDF. Nilai pembobotan 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 (4). Untuk sampel pembobotan TF-IDF ditunjukkan pada table 12.

Tabel 12 Data Sampel untuk Perhitungan TF-IDF Setelah Pra Proses Data

Ulasan	Sebelum	Sesudah
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 Halodok	['barang', 'sesuai', 'order', 'kirim', 'cepat']

Ulasan	Sebelum	Sesudah
U7	Sdh seminggu ini aplikasinya kok g bs dipakai yaa.. Sdh saya update sampe download ulang tetep bgitu.. kenapa yaa..	['update', 'download', 'ulang']
U8	Membantu 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']

Pada Tabel 12 menggunakan 8 dokumen dimana dokumen tersebut sudah dilakukan praproses data sebelumnya, untuk dijadikan sampel untuk perhitungan TF-IDF. Berikut ini adalah perhitungan TF-IDF pada sampel U1.

Perhitungan diawali dengan menghitung TF (Term Frequency) menggunakan persamaan rumus (1) dilakukan dengan menghitung frekuensi kata yang muncul pada satu dokumen sebagai berikut:

$$Tf_{jelas,u1}=1$$

$$Tf_{janji,u1}=1$$

$$Tf_{undur,u1} = 1$$

Selanjutnya perhitungan IDF (*Inverted Document Frequency*) menggunakan persamaan rumus (3) dilakukan dengan menghitung frekuensi kata yang muncul pada seluruh dokumen sebagai berikut:

$$IDF_{jelas,u1} = \log\left(\frac{8}{1}\right) = 0.903$$

$$IDF_{janji,u1} = \log\left(\frac{8}{2}\right) = 0.602$$

$$IDF_{undur,u1} = \log\left(\frac{8}{1}\right) = 0.903$$

Setelah IDF selanjutnya perhitungan TF-IDF sehingga seluruh hasil perhitungan TF-IDF pada kelima dokumen sampel data ditunjukkan pada tabel 13.

Tabel 13 Hasil DF dan IDF

Kata	TF								DF	IDF
	U1	U2	U3	U4	U5	U6	U7	U8		
badan								1	1	0.903
bagus			1						1	0.903
bantu								1	1	0.903
barang						1			1	0.903
bayar			1						1	0.903
cepat						1			1	0.903
download							1		1	0.903
error					1				1	0.903
erti		1							1	0.903
habis			1						1	0.903
hitung			1						1	0.903
janji	1				1				2	0.602
jelas	1								1	0.903
kirim						1			1	0.903
konsultasi			1						1	0.903
lemas								1	1	0.903
malas								1	1	0.903

Kata	TF								DF	IDF
	U1	U2	U3	U4	U5	U6	U7	U8		
menu				1					1	0.903
modifikasi				1					1	0.903
mudah		1		1					2	0.602
order						1			1	0.903
pandemic								1	1	0.903
ringan								1	1	0.903
sakit								1	1	0.903
sesuai						1			1	0.903
tampil				1					1	0.903
tunggu			1						1	0.903
ulang							1		1	0.903
undur	1								1	0.903
update							1		1	0.903

Setelah mendapatkan perhitungan tf dan idf pada Tabel 13, selanjutnya perhitungan TF-IDF menggunakan perasamaan rumus (4) guna mendapatkan nilai bobot setiap kata sebagai berikut:

$$TF-IDF_{jelas,u1} = 1 \cdot 0.903 = 0.903$$

$$TF-IDF_{janji,u1} = 1 \cdot 0.602 = 0.602$$

$$TF-IDF_{undur,u1} = 1 \cdot 0.903 = 0.903$$

Pada Tabel 14 dibawah ini adalah hasil seluruh perhitungan TF-IDF pada setiap kata pada 5 ulasan.

Tabel 14 Hasil Pembobotan TF-IDF

Kata	TF-IDF							
	U1	U2	U3	U4	U5	U6	U7	U8
badan								0.903
bagus			0.903					
bantu								0.903
barang						0.903		
bayar			0.903					
cepat						0.903		
download							0.903	
error					0.903			
erti		0.903						
habis			0.903					
hitung			0.903					
janji	0.602				0.602			
Jelas	0.903							
kirim						0.903		
konsultasi			0.903					
lemas								0.903

Kata	TF-IDF							
	U1	U2	U3	U4	U5	U6	U7	U8
malas								0.903
menu				0.903				
modifikasi				0.903				
mudah		0.602		0.602				
order						0.903		
pandemic								0.903
ringan								0.903
sakit								0.903
sesuai						0.903		
tampil				0.903				
tunggu			0.903					
ulang							0.903	
undur	0.903							
update							0.903	

Pembobotan kata menggunakan tf-idf ini dilakukan hanya pada data training saja. Setiap kata yang sudah dihitung akan digunakan untuk melakukan klasifikasi menggunakan metode SVM (*Support Vector Machine*).

4.5 Klasifikasi (SVM)

Setelah dilakukan pembobotan kata, kita akan melakukan pengklasifikasian menggunakan SVM. Pada Tabel 15 kita akan mencoba pengklasifikasian dengan data di bawah ini, U1 sebagai data ujinya.

Tabel 15 Data Sampel untuk Klasifikasi

Ulasan	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	Error 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 bgiu.. kenapa yaa..
U8	Membantu banget, apalagi di masa pandemi atau sakit ringan tapi mager buat ke dokter karena badan lemes. Terima kasih Halodoc

Dari Tabel 15 kita dapat menentukan bidang pembatasnya menggunakan rumus persamaan (7), dan (8). Dapat dilihat dari Tabel 16 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 16 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 : $0.845w_1 + 0.845w_2 + 0.845w_3 + 0.845w_4 + 0.845w_5 + 0.845w_6 + 0.845w_7 + b \geq 1$

Setelah didapatkan bidang pembatas pada Tabel 16, akan dilakukan transformasi data dapat dilihat pada Tabel 17 dengan U1 tidak dilakukan perhitungan TF-IDF.

Tabel 17 Transformasi Data

Fitur /term	TF-IDF						
	U2 (x1)	U3 (x2)	U4 (x3)	U5 (x4)	U6 (x5)	U7 (x6)	U8 (x7)
erti	0,845	0	0	0	0	0	0
mudah	0,544	0	0,544	0	0	0	0
bayar	0	0,845	0	0	0	0	0
bagus	0	0,845	0	0	0	0	0
habis	0	0,845	0	0	0	0	0
hitung	0	0,845	0	0	0	0	0

Fitur /term	TF-IDF						
	U2 ^{z6} (x1)	U3 (x2)	U4 (x3)	U5 (x4)	U6 (x5)	U7 (x6)	U8 (x7)
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

Fitur /term	TF-IDF						
	U2 (x1)	U3 (x2)	U4 (x3)	U5 (x4)	U6 (x5)	U7 (x6)	U8 (x7)
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(14). Pada Tabel 18 adalah hasil *matriks* dari perhitungan kernel *linear*.

Tabel 18 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

Matriks							
	0	0	0	0	0	0	4,999

Selanjutnya kita akan mencari nilai α dengan persamaan rumus(10) nilai ini untuk menentukan margin terbesar dengan teknik komputasi *lagrange multiplier*, yang dapat dilihat pada Tabel 19.

Tabel 19 Permisalan Nilai α

y (kelas)		1	1	-1	-1	1	-1	1
Kemungkinan alfa	α_1	2	1	1	1	1	1	1
	α_2	1	2	1	1	1	1	1
	α_3	1	1	2	1	1	1	1

Dari Tabel 19 dalam menentukan kemungkinan alfa kita dapat memilih nilainya dari range 1 sampai banyak dokumen kita yaitu 7. Kita akan memilih α_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 α selanjutnya kita akan mencari nilai w untuk setiap fitur yang kita punya, untuk mencarinya kita akan menggunakan persamaan rumus(9), hasil dari perhitungan nilai w dapat dilihat pada Tabel 20.

Tabel 20 Hasil Nilai w

Fitur /Term	Nilai w
erti	1,690196
mudah	0,544068
bayar	0,845098
bagus	0,845098

Fitur /Term	Nilai w
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

Fitur /Term	Nilai w
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 (11), 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 (12), nilai yang sudah kita dapatkan antara lain:

α :	2	1	1	1	1	1	1
y:	1	1	-1	-1	1	-1	1
b:	-4,938						

Nilai α , 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(14), hasil yang didapatkan dari perhitungan tersebut adalah :

$$k(x_i x_j) = \begin{bmatrix} 0 & 0 & 0 & 0,459 & 0 & 0 & 0 \end{bmatrix}$$

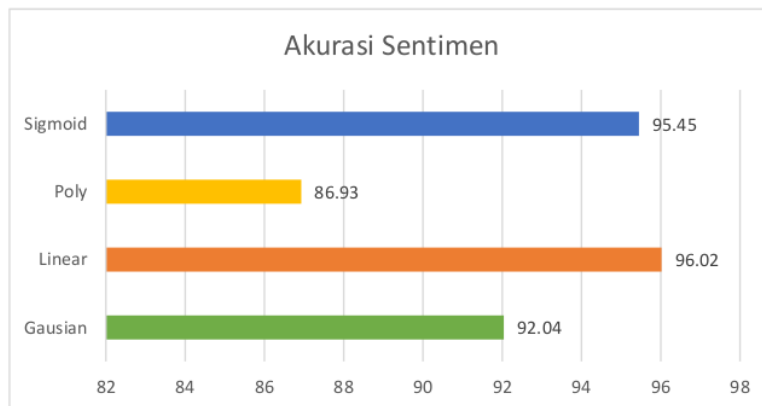
Setelah semua nilai yang dibutuhkan sudah ada, kita dapat menggunakan persamaan rumus (12), hasil dari perhitungan tersebut adalah -1, fungsi *sign()* adalah fungsi normalisasi, jika hasil nilai ≥ 0 , maka fungsi tersebut bernilai kelas positif, jika hasil nilai ≤ 0 , maka fungsi tersebut bernilai kelas 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 (14), Kernel Polynomial dengan rumus persamaan (15), Kernel RBF dengan rumus persamaan (16), dan Kernel Sigmoid dengan rumus persamaan (17).

4.5.1 Klasifikasi Sentimen

Data yang sudah dipisah menjadi data training dan data testing selanjutnya akan diklasifikasi menggunakan library python yang sudah tersedia. Model klasifikasi yang sudah terbentuk selanjutnya akan digunakan untuk pengujian data *testing* untuk mendapatkan nilai akurasinya.



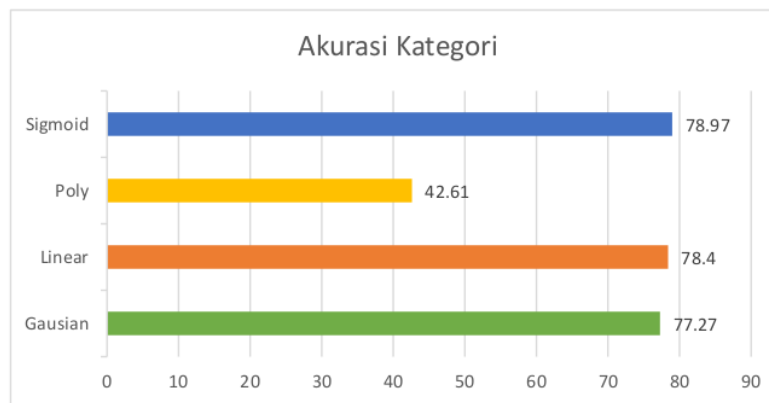
Gambar 12 Hasil Akurasi Tiap Kernel untuk Klasifikasi Sentimen

Pada Gambar 12 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%.

4.5.2 Klasifikasi Kategori

Pada klasifikasi kategori akan dilakukan pengklasifikasian multiclass, dengan metode *One vs Rest* mengacu pada penelitian terdahulu yang mempunyai hasil bahwa nilai akurasi *One vs Rest* lebih tinggi dari pada *One vs One* (Alita, Fernando, & Sulistiani, 2020).

Maka penelitian ini akan menggunakan *One vs Rest* dengan menggunakan library python yang sudah tersedia. Model klasifikasi yang sudah terbentuk selanjutnya akan digunakan untuk pengujian data testing untuk mendapatkan nilai akurasinya.



Gambar 13 Hasil Akurasi Tiap Kernel untuk Klasifikasi Kategori

Pada Gambar 13 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 72,27%, serta kernel *polynomial* dengan akurasi sebesar 42,61%.

4.6 Evaluasi

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).

4.6.1 Evaluasi Model Sentimen

Pada Tabel 21 adalah hasil confusion matrix untuk model kelas sentimen, mengacu pada Tabel 1 yang ada pada bab 2, tabel untuk confusion matrix 2 dimensi, menunjukkan bahwa untuk nilai TP terdapat 86 ulasan yang terklasifikasi benar, dan nilai TN terdapat 83 ulasan yang terklasifikasi benar. Nilai FP dan NP terdapat 4 ulasan dan 3 ulasan yang menunjukkan bahwa nilai yang terklasifikasi salah.

Tabel 21 Confusion Matrix Kelas Sentimen

True Label	Perdict Label	
	Positif	Negatif
Positif	86	4
Negatif	3	83

Pada Tabel 21, setelah mendapatkan nilai yang diperlukan dari confusion matrix, kita dapat menghitung 4 performa matrik satu per satu dengan rumus persamaan untuk *akurasi*(21), untuk *precision* (22), untuk *recall* (23), dan untuk *f1-score* (24) seperti berikut:

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

$$Positif = Precision = \frac{86}{86+4} = 0,955$$

$$Recall = \frac{86}{86+3} = 0,966$$

$$F1 - score = 2 * \frac{0,955 \times 0,966}{0,955 + 0,966} = 0,960$$

$$Negatif = Precision = \frac{83}{83+3} = 0,965$$

$$Recall = \frac{83}{83+4} = 0,954$$

$$F1 - score = 2 * \frac{0,965 \times 0,954}{0,965 + 0,954} = 0,959$$

4.6.2 Evaluasi Model Kategori

Pada Tabel 16 adalah confusion matrix untuk model kelas kategori, confusion matrix yang digunakan mengacu pada tabel yang ada di bab2 yaitu Tabel 2 confusion matrix untuk multi class, pada confusion matrix untuk multiclass tidak ada nilai untuk TN (true negative).

Tabel 22 Confusion Matrix Kelas Kategori

True Label	Predict Label					
	Effecie ncy	Function ality	Maintaina bility	Portabi lity	Reliabi lity	Usabil ity
Effeciency	20	1	0	0	0	1
Functionali ty	1	46	1	3	2	3
Maintaina bility	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

Pada Tabel 22, setelah mendapatkan nilai yang diperlukan pada confusion matrix, kita dapat menghitung 4 performa matrik satu per satu dengan rumus persamaan untuk *akurasi*(25), untuk *precision* (26), untuk *recall* (27), dan untuk *f1-score* (28) seperti berikut:

$$Accuracy = \frac{20+46+11+22+18+22}{20+46+11+22+18+22+37} = 0,789$$

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

$$\text{Recall} = \frac{22}{22+12} = 0,647$$

$$\text{F1-score} = 2 * \frac{0,846 \times 0,647}{0,846 + 0,647} = 0,733$$

4.7 Visualisasi

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

4.7.1 Sentimen Positif

Pada Gambar 14 adalah *worldcloud* untuk kelas sentimen positif, yang di dapat dari ulasan pengguna aplikasi halodoc versi 10.



Gambar 14 *Worldcloud* Sentimen Positif

Dari Gambar 14 menunjukkan bahwa kata “bantu”, “konsultasi”, “obat”, “cepat”, dan “mudah”, adalah kata yang sering muncul untuk kelas sentimen positif, menunjukkan bahwa aplikasi halodoc secara fungsional sesuai dengan yang dibutuhkan oleh user, konsultasi dan pengiriman obat, mudah digunakan, serta aplikasi halodoc sebagai aplikasi utama dalam pengobatan pertama.

4.7.2 Sentimen Negatif

Pada Gambar 15 adalah worlcoud untuk kelas sentimen negatif, yang di dapat dari ulasan pengguna aplikasi halodoc versi 10.



Gambar 15 Wordcloud Sentimen Negatif

Dari Gambar 15 menunjukkan bahwa kata “bayar”, “obat”, “vaksin”, “konsultasi”, dan “pesan”, adalah kata yang sering muncul untuk kelas sentimen negatif, menunjukkan bahwa pengguna aplikasi halodoc protes dengan aplikasi halodoc yang berbayar, pengiriman obat yang terkadang tidak lengkap, pengiriman lambat, pembatalan pesanan, dan aplikasi halodoc sebagai penyelenggara untuk vaksinasi dinilai gagal.

4.7.3 Kategori Functionality

Pada Gambar 16 adalah worlcoud untuk kelas kategori *functionality*, yang di dapat dari ulasan pengguna aplikasi halodoc versi 10.



Gambar 16 Wordcloud Kategori *Functionality*

Dari Gambar 16 menunjukkan bahwa kata “konsultasi”, “bayar”, “pesan”, “vaksin”, dan “bantu”, adalah kata yang sering muncul untuk kelas kategori *functionality*, menunjukkan bahwa pengguna aplikasi halodoc cukup puas dengan fungsi yang ada pada halodoc dari pengiriman dan konsultasi, tetapi disuatu sisi pengguna juga mengeluh dengan adanya tarif biaya penggunaan aplikasi, kebocoran data pengguna, dan program vaksinasi yang tidak puas dari pengguna aplikasi halodoc.

4.7.4 Kategori *Reliability*

Pada Gambar 17 adalah *wordcloud* untuk kelas kategori *reliability*, yang di dapat dari ulasan pengguna aplikasi halodoc versi 10.



Gambar 17 Wordcloud Kategori *Reliability*

Dari Gambar 17 menunjukkan bahwa kata “bayar”, “obat”, “konsultasi”, “pesan”, dan “lambat”, adalah kata yang sering muncul untuk kelas kategori *reliability*, menunjukkan bahwa pengguna aplikasi halodoc rata rata mengeluh akan transaksi yang sering batal padahal pengguna sudah melakukan pembayaran, terutama dalam transaksi pembelian obat, dan dana refund setelah terjadinya kesalahan aplikasi lambat untuk di kembalikan. Dan terkadang sering terjadi kendala pada saat sesi konsultasi dimana aplikasi terkadang keluar sendiri dan gagal dalam membangun kembali aplikasi yang sempat error.

4.7.5 Kategori Usability

Pada Gambar 18 adalah wordcloud untuk kelas kategori *usability*, yang di dapat dari ulasan pengguna aplikasi halodoc versi 10.



Gambar 18 Wordcloud Kategori Usability

Dari Gambar 18 menunjukkan bahwa kata “mudah”, “obat”, “bayar”, “konsultasi”, dan “bantu”, adalah kata yang sering muncul untuk kelas kategori *usability*, menunjukkan bahwa pengguna aplikasi halodoc mudah dalam pengoperasiannya, mudah di pelajari dan dipahami dalam penggunaan fungsi yang ada pada aplikasi halodoc.

4.7.6 Kategori Efficiency

Pada Gambar 19 adalah wordcloud untuk kelas kategori *efficiency*, yang di dapat dari ulasan pengguna aplikasi halodoc versi 10.



Gambar 19 Wordcloud Kategori Efficiency

Dari Gambar 19 menunjukkan bahwa kata “cepat”, “obat”, “lambat”, “respons”, dan “apotek”, adalah kata yang sering muncul untuk kelas kategori *efficiency*, menunjukkan bahwa aplikasi halodoc cepat dalam memberikan respons yang sesuai fungsi, dan mampu dengan baik menggunakan sumber daya yang dimilikinya ketika fungsi tertentu dijalankan, akan tetapi pengguna terkadang protes akan pengiriman obat yang dinilai lambat untuk beberapa daerah di Indonesia.

4.7.7 Kategori Maintainability

Pada Gambar 20 adalah wordcloud untuk kelas kategori *maintainability*, yang di dapat dari ulasan pengguna aplikasi halodoc versi 10.



Gambar 20 Wordcloud Kategori *Maintainability*

Dari Gambar 20 menunjukkan bahwa kata “rekondisi”, “koneksi”, “buka”, “masuk”, dan “update”, adalah kata yang sering muncul untuk kelas kategori *maintainability*, menunjukkan bahwa aplikasi halodoc dalam meminimalkan efek tak terduga dari modifikasi perangkat lunak masih gagal, terbukti dari banyaknya keluhan dari pengguna dari aplikasi yang tidak bisa dibuka, pengguna yang mengeluh setelah dilakukannya update, dan seringkali terjadi kesalahan yang diakibatkan koneksi dari aplikasinya.

4.7.8 Kategori Portability

Pada Gambar 21 adalah wordcloud untuk kelas kategori *portability*, yang di dapat dari ulasan pengguna aplikasi halodoc versi 10.



Gambar 21 Wordcloud Kategori Portability

Dari Gambar 21 menunjukkan bahwa kata “bantu”, “rumah”, “obat”, “konsultasi”, dan “butuh”, adalah kata yang sering muncul untuk kelas kategori *portability*, menunjukkan bahwa pengguna aplikasi halodoc terbantu dengan kemampuan perangkat lunak untuk digunakan sebagai pengganti perangkat lain, dan kemampuan perangkat lunak yang bisa digunakan dalam kondisi dan lingkungan tertentu.

5.1 Kesimpulan

Bedasarkan hasil dan analisis yang dilakukan pada bab 4 dapat disimpulkan sebagai berikut:

1. 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*.
2. 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*.
3. Berdasarkan hasil klasifikasi, untuk kelas *functionality* pengguna mengeluh akan aplikasi halodoc yang menjadi berbayar, kebocoran data pengguna, dan program penyaluran vaksinasi. Untuk kelas *reliability* pengguna mengeluh dengan aplikasi yang terkadang error ketika sedang digunakan dan pengembalian uang refund yang lama setelah terjadinya error pada aplikasi, untuk kelas *usability* pengguna tidak mengalami kesulitan dalam pengoperasian aplikasi, untuk kelas *efficiency* pengguna cukup puas dengan hasil ketika suatu fungsi dijalankan tetapi pengguna juga mengeluh akan pengiriman obat yang dinilai lama. Untuk kelas *maintaibility* aplikasi terkadang tidak bisa dibuka setelah dilakukan update dan sering terjadinya kelasahan diakibatkan koneksi terputus. Dan untuk kelas *porability* menunjukkan bahwa aplikasi halodoc dapat digunakan sebagai perangkat lunak lain.

5.2 Saran

Bedasarkan penelitian ini, saran pada penelitian selanjutnya adalah:

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

Daftar Pustaka

- Alita, D., Fernando, Y., & Sulistiani, H. (2020). IMPLEMENTASI ALGORITMA MULTICLASS SVM PADA OPINI PUBLIK BERBAHASA INDONESIA DI TWITTER. *Jurnal TEKNOKOMPAK*, 86-91.
- 58 Fatmawati, & Affandes, M. (2017). Klasifikasi Keluhan Menggunakan Metode Support Vector Machine (SVM). *Jurnal CoreIT*, 24-30.
- 1 Fransiska, S., Rianto, & Gufroni, A. I. (2020). Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method. *Scientific Journal of Informatics*, 203-212.
- Gogia, S. (2020). *FUNDAMENTALS OF TELEMEDICINE AND TELEHEALTH*. India: Stacy Masucci.
- 27 Gorunescu, F. (2011). *Data Mining Concepts, Models and Techniques*. Springer.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques*. Waltham: Morgan Kaufmann.
- 62 Hendra, A., & Fitriyani. (2021). Analisis Sentimen Review Halodoc Menggunakan Naïve Bayes Classifier. *JISKA*, 78-89.
- 11 Hoffmann, M., & Chisholm, A. (2016). *Text Mining and Visualization: Case Studies Using Open-Source Tools*. CRC Press.
- 24 Imawan, L. B., & Mude, M. A. (2020). Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store. *ILKOM Jurnal Ilmiah*, 154-161.
- 12 Jo, T. (2019). *Text Mining: Concepts, Implementation, and Big Data Challenge*. Springer.
- Kantardzic, M. (2020). *DATA MINING Concepts, Models, Methods, and Algorithm*. Wiley - IEEE Press.
- 29 Koch, R. (2013). *The 80/20 Principle And 92 Other Powerful Laws Of Nature*. Lodon: Quercus.
- 1 Liu, B. (2015). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.
- 9 Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A., & Doulamis, N. (2021). Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. *MDPI*, 1-22.

- ¹ Rohanah, A., Dermawan, B. A., & Purnamasari, I. (2021). Klasifikasi Ulasan Pengguna Zoom Cloud Meetings Menggunakan Metode Information Gain dan Naïve Bayes Classifier. *Jurnal Informatika Universitas Pamulang*, 348-357.
- ⁵ Rohanah, A., Rianti, D. L., & Sari, N. B. (2021). PERBANDINGAN NAÏVE BAYES DAN SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI ULASAN PELANGGAN INDIHOME . *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 23-30.
- ³⁶ Sari, G. S., & Wirman, W. (2021). Telemedicine sebagai Media Konsultasi Kesehatan di Masa Pandemic COVID 19 di Indonesia. *Jurnal Komunikasi*, 43-54.
- ³⁵ Sarkar, D. (2016). *Text Analytics with Python: A Pratical Real-World Approach to Gaining Actionable Insights from your Data*. Apress.
- ⁶⁵ Stoean, C., & Stoean, R. (2014). *Support Vector Machines and Evolutionary Algorithms for Classification*. London: Springer.
- ⁸³ Suryn, W. (2014). *Software Quality Engineering: A Practitioner's Approach*. Montreal: Wiley-IEEE Computer Society Pr.
- ⁷⁰ Zaki, J. M., & JR, M. W. (2020). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press.
- ⁴⁹ Zizka, J., Darena, F., & Svoboda, A. (2020). *Text Mining with Machine Learning Priciples and Techniques*. London: CRC Press.

Riwayat Hidup

Nama : Fachran Sandi
Tempat/ Tanggal Lahir : Jakarta, 20 Juli 2000
Jenis Kelamin : Pria
Agama : Islam
Kewarganegaraan : Indonesia
Alamat : Jalan Basuki RT/RW 05/06 NO.56, Cilangkap
Cipayung Jakarta Timur 13870
No. Telp : +62 81398688964
Email : fachransandi@gmail.com

Pendidikan Formal

SDN Kedoya Selatan 04 Pagi (2006 – 2012)
SMPN 180 Jakarta (2012 – 2015)
SMA Bina Dharma Jakarta (2015 – 2018)

Pengalaman Organisasi

Ketua Divis Akademik KSM Multimedia (2019 – 2020)
Anggota KSM Robotika (2019 – 2020)
Anggota KSM Multimedia (2020 – 2021)
Anggota KSM Robotika (2020 – 2021)

Lampiran

Lampiran 1

Daftar Kata *Normalization*

Sebelum	Sesudah
aaamiin	amin
aamiin	amin
abal	jelek
abes	habis
abis	habis
abisss	habis
abrek	hancur
abs	habis
absen	hadir
acc	terima
acoount	account
acu	aku
ad	ada
ade	ada
adek	adik
adh	waduh
adkaan	ada
adm	admin
adminya	admin
ads	ada
adu	lawan
aduin	dilaporkan
ae	saja
aer	air
afdol	manjur
again	lagi
agak	sedikit

Sebelum	Sesudah
apitik	apotek
apk	aplikasi
apkikasi	aplikasi
apknya	aplikasi
apl	aplikasi
aplg	apalagi
apliaksi	aplikasi
aplikadi	aplikasi
aplikasih	aplikasi
aplikasihnya	aplikasi
aplikasiii	aplikasi
aplikasiny	aplikasi
apliksi	aplikasi
aplilashinya	aplikasinya
aplilasinya	aplikasi
aplsi	aplikasi
aplx	aplikasi
apotik	apotek
app	aplikasi
aplikasi	aplikasi
aplikasinya	aplikasi
applikasi	aplikasi
appointment	bertemu
apps	aplikasi
appsnya	aplikasi
aps	aplikasi
arahin	arah

ahlak	perilaku
aing	saya
aj	saja
aja	saja
ajaa	saja
ajja	saja
ajr	sialan
ake	pake
akhir	akhir
akrifannya	aktif
aks	alias
aktfkn	aktifkan
aku	saya
akutuh	saya
akuw	saya
alami	alam
alesan	alasan
alesannya	alasan
alhamdllh	syukur
alhamdulillah	syukur
alhamdulillah	syukur
alhamdulillah	syukur
alikasi	aplikasi
all	semua
alhamdulillah	syukur
almost	hampir
alpa	bolos
also	juga
am	sama
ama	sama

as	sebagai
asa	ada
asbun	asal bunyi
asik	asik
asis	asik
aslik	asik
astagfirullah	astaga
asyu	anjing
atend	hadir
attention	perhatian
ato	atau
atsan	atasan
auah	tidak tahu
auto	otomatis
ayah	orang tua
back	kembali
badmod	kesal
badmood	kesal
bagusss	bagus
baiikk	baik
baikdan	baik dan
bakal	dibuat
bales	balas
balesan	balasan
balik	kembali
balikin	membaliki
balim	kembali
bandel	nakal
bang	abang
banget	sangat

ammiinn	amin
amp	sampai
ampas	jelek
ampe	sampai
ampun	kapok
ampunnn	ampun
ampyun	ampun
ana	saya
anabul	hewan
anakya	anak
and	dan
andelin	mengandalkan
angep	anggap
angota	anggota
anh	aneh
anj	anjing
anjir	kesal
anjlok	turun
anjrit	kesal
ank	anak
anoying	kesal
anteng	damai
anter	antar
anyep	hangus
anything	segala sesuatu
aoa	apa
ap	apa
apadaya	apa daya
apalikasi	aplikasi
apdet	update

bangett	banget
BANGETTTT	banget
bangettttt	sangat
bangke	bangsat
bangking	bank
bangrt	sangat
bangt	sangat
bantuanya	bantuan
banyakin	banyak
banyak	banyak
bapa	bapak
bapuk	jelek
bareng	bersama
barusn	barusan
base	dasar
bat	banget
batalin	batal
batre	daya
bayangin	bayang
bayarny	bayarnya
bayik	bayi
bberapa	beberapa
bbrapa	beberapa
bbrp	beberapa
bbrpa	beberapa
bcs	karena
be	menjadi
becus	bisa
bego	bodoh
beh	waduh

ape	apa
-----	-----

belakanh	belakang
----------	----------

Lampiran 2
Daftar Stopword

ada	banget	bye	dipunyai	hendaknya
a	bapak	cara	diri	hey
aamiin	baru	caranya	dirinya	hihii
aamin	bawah	cilacap	disampaikan	hingga
abang	beberapa	coy	disebut	hm
achmad	begini	cukup	disebutkan	hmm
adakan	beginian	cukupkah	disebutkannya	hmmmm
adalah	beginikah	cukuplah	disini	huh
adanya	beginilah	cuma	disinilah	i
adapun	begitu	cuss	ditambahkan	ia
adik	begitukah	cuus	ditandaskan	ialah
admin	begitulah	cuuzzz	ditanya	ibarat
agak	begitupun	dah	ditanyai	ibaratkan
agaknya	bekerja	dahulu	ditanyakan	ibaratnya
agar	belakang	dalam	ditegaskan	ibu
agustus	belakangan	dan	ditujukan	ikut
ah	beliau	dapat	ditunjuk	in
ahh	belum	dari	ditunjuki	indomaret
aju	belumlah	daripada	ditunjukkan	indonesia
akan	benar	dasar	ditunjukkannya	ingat
akankah	benarkah	datang	ditunjuknya	ingat-ingat
akhir	benarlah	db	dituturkan	ingin
akhiri	ber	dear	dituturkannya	inginkah
akhirnya	berada	deh	diucapkan	inginkan
aku	berakhir	dekat	diucapkannya	inhealth
akulah	berakhirilah	demi	diungkapkan	ini
al	berakhirnya	demikian	djp	inikah
ala	berapa	demikianlah	doang	inilah
alaikum	berapakah	dengan	doc	insyaallah

alfamart	berapalah	depan	dokter	istri
alhasil	berapapun	depok	dong	itu
allah	berarti	des	dua	itukah
allianz	berawal	desember	duh	itulah
amat	berbagai	detik	duluan	iya
amatlah	berdatangan	di	dwi	iyak
amen	beri	dia	effendy	jadi
amiin	berikan	diakhiri	eh	jadilah
amin	berikut	diakhirinya	ehh	jadinya
an	berikutnya	dialah	eka	jalan
anak	berjumlah	diandra	empat	jam
anandita	berkali-kali	diantara	end	janet
anandya	berkata	diantaranya	endap	jangan
anda	berkehendak	diberi	enggak	jangan
andalah	berkeinginan	diberikan	enggaknya	janganlah
andrey	berkenaan	diberikannya	entah	january
anggap	berlainan	dibuat	entahlah	jateng
anjim	berlalu	dibuatnya	entar	jati
antara	berlangsung	didapat	era	jatim
antaranya	berlebihan	didatangkan	esa	jawab
anthony	bermacam	digunakan	firmansyah	jawaban
antonius	bermacam- macam	diibaratkan	fitri	jawabnya
apa	bermaksud	diibaratkannya	forte	jawi
apaan	bermula	diingat	gara	jaya
apabila	bersama	diingatkan	gatau	jebret
apakah	bersama-sama	diinginkan	gataukan	jelaskan
apalagi	bersiap	dijawab	gebuh	jelaslah
apannya	bersiap-siap	dijelaskan	gimana	jelasnya
apatah	bertanya	dijelaskannya	gitu	jika
aplq	bertanya- tanya	dikarenakan	gresik	jikalau

aplikasi	berturut	dikatakan	guna	jl
artinya	berturut-turut	dikatakannya	gunakan	jm
arya	bertutur	dikerjakan	gunawan	jos
asa	berujar	diketahui	h	juga
asal	berupa	diketuainya	had	juli
asalanualaikum	besar	dikira	hadap	july
asalkan	betul	dilakukan	hadeeeh	jumlah
assalamu	betulkah	dilalui	haduuuuuuu	jumlahnya
astafirullah	biar	dilihat	hal	juni
astrazenece	biasa	dimaksud	haldoc	justru
atas	biasanya	dimaksudkan	hallo	k
atau	bikin	dimaksudkannya	hallo doc	kab
ataukah	bila	dimaksudnya	halo	kadang
ataupun	bilakah	diminta	halodoc	kagak
awal	bilang	dimintai	haloo	kakak
awalnya	bintang	dimisalkan	hampir	kala
b	bisa	dimulai	han	kalau
bagai	bisakah	dimulailah	hanya	kalaulah
bagaikan	bismillah	dimulainya	hanyalah	kalaupun
bagaimana	boleh	dimungkinkan	harap	kali
bagaimanakah	bolehkah	dini	hari	kalian
bagaimanapun	bolehlah	dipastikan	harus	kalo
bagi	bos	diperbuat	haruslah	kami
bagian	buat	diperbuatnya	harusnya	kamilah
bahkan	bukan	dipergunakan	hasan	kamu
bahwa	bukankah	diperkirakan	hehehe	kamulah
bahwasanya	bukanlah	diperlihatkan	hei	kan
baik	bukannya	diperlukan	heidy	kapan
bakal	bulan	diperlukannya	hello	kapankah
bakalan	bung	dipersoalkan	hendak	kapanpun
balik	but	dipertanyakan	hendaklah	karena

ORIGINALITY REPORT

19%

SIMILARITY INDEX

17%

INTERNET SOURCES

7%

PUBLICATIONS

9%

STUDENT PAPERS

PRIMARY SOURCES

1	repository.upnvj.ac.id Internet Source	3%
2	dspace.uii.ac.id Internet Source	1%
3	Submitted to Bellevue Public School Student Paper	1%
4	Submitted to Sriwijaya University Student Paper	1%
5	repository.uin-suska.ac.id Internet Source	<1%
6	Submitted to Universitas Brawijaya Student Paper	<1%
7	repository.its.ac.id Internet Source	<1%
8	Submitted to Universitas Muria Kudus Student Paper	<1%
9	theses.uin-malang.ac.id Internet Source	<1%

10	de.scribd.com Internet Source	<1 %
11	repository.ub.ac.id Internet Source	<1 %
12	www.researchgate.net Internet Source	<1 %
13	docplayer.info Internet Source	<1 %
14	123dok.com Internet Source	<1 %
15	dl.icdst.org Internet Source	<1 %
16	digilib.uinsby.ac.id Internet Source	<1 %
17	repository.usd.ac.id Internet Source	<1 %
18	1library.org Internet Source	<1 %
19	adoc.pub Internet Source	<1 %
20	ena.lp.edu.ua Internet Source	<1 %
21	eprints.uny.ac.id Internet Source	<1 %

22	powcoder.com Internet Source	<1 %
23	web.archive.org Internet Source	<1 %
24	journal.ipm2kpe.or.id Internet Source	<1 %
25	ejournal.uin-suka.ac.id Internet Source	<1 %
26	Submitted to UIN Sultan Syarif Kasim Riau Student Paper	<1 %
27	www.scribd.com Internet Source	<1 %
28	Submitted to Udayana University Student Paper	<1 %
29	conference.upnvj.ac.id Internet Source	<1 %
30	jim.teknokrat.ac.id Internet Source	<1 %
31	www.mdpi.com Internet Source	<1 %
32	www.vegagerdin.is Internet Source	<1 %
33	ejournal.unesa.ac.id Internet Source	<1 %

34	media.neliti.com Internet Source	<1 %
35	Submitted to Universidad Rey Juan Carlos Student Paper	<1 %
36	wisuda.unissula.ac.id Internet Source	<1 %
37	yunipedia.blogspot.com Internet Source	<1 %
38	pdfs.semanticscholar.org Internet Source	<1 %
39	pedagogika.tc-h.ru Internet Source	<1 %
40	repository.bsi.ac.id Internet Source	<1 %
41	text-id.123dok.com Internet Source	<1 %
42	Submitted to Universitas Airlangga Student Paper	<1 %
43	downloads.hindawi.com Internet Source	<1 %
44	jurnal.kemendag.go.id Internet Source	<1 %
45	Submitted to Universitas Pendidikan Ganesha Student Paper	<1 %

46	es.scribd.com Internet Source	<1 %
47	johannessimatupang.wordpress.com Internet Source	<1 %
48	repub.eur.nl Internet Source	<1 %
49	sutir.sut.ac.th:8080 Internet Source	<1 %
50	Ernianti Hasibuan, Elmo Allistair Heriyanto. "ANALISIS SENTIMEN PADA ULASAN APLIKASI AMAZON SHOPPING DI GOOGLE PLAY STORE MENGUNAKAN NAIVE BAYES CLASSIFIER", Jurnal Teknik dan Science, 2022 Publication	<1 %
51	Submitted to Universitas Atma Jaya Yogyakarta Student Paper	<1 %
52	digilib.esaunggul.ac.id Internet Source	<1 %
53	openjournal.unpam.ac.id Internet Source	<1 %
54	repository.unja.ac.id Internet Source	<1 %
55	Submitted to Universitas Dian Nuswantoro Student Paper	<1 %

56	ecampus.pelitabangsa.ac.id Internet Source	<1 %
57	jurnal.fikom.umi.ac.id Internet Source	<1 %
58	jurnal.umj.ac.id Internet Source	<1 %
59	repository.teknokrat.ac.id Internet Source	<1 %
60	Submitted to Ajou University Graduate School Student Paper	<1 %
61	Submitted to Fakultas Ekonomi Universitas Indonesia Student Paper	<1 %
62	ejournal.kresnamediapublisher.com Internet Source	<1 %
63	etd.umy.ac.id Internet Source	<1 %
64	Submitted to Universitas Negeri Jakarta Student Paper	<1 %
65	Submitted to University Politehnica of Bucharest Student Paper	<1 %
66	lib.ibs.ac.id Internet Source	<1 %

67	lib.uin-malang.ac.id Internet Source	<1 %
68	openlibrarypublications.telkomuniversity.ac.id Internet Source	<1 %
69	www.ncbi.nlm.nih.gov Internet Source	<1 %
70	dokumen.pub Internet Source	<1 %
71	fr.scribd.com Internet Source	<1 %
72	darusmansurya.wordpress.com Internet Source	<1 %
73	elibrary.unikom.ac.id Internet Source	<1 %
74	ijeais.org Internet Source	<1 %
75	jurnal.polibatam.ac.id Internet Source	<1 %
76	medium.com Internet Source	<1 %
77	repositori.usu.ac.id Internet Source	<1 %
78	simdos.unud.ac.id Internet Source	<1 %

79	digilib.unila.ac.id Internet Source	<1 %
80	doku.pub Internet Source	<1 %
81	publikasi.mercubuana.ac.id Internet Source	<1 %
82	www.helmirfansah.com Internet Source	<1 %
83	www.springerprofessional.de Internet Source	<1 %
84	Carlos Enrique Paucar Farfán. "Classificação dos estados cognitivos orientados pelo sujeito baseada na variabilidade da frequência cardíaca", Universidade de Sao Paulo, Agencia USP de Gestao da Informacao Academica (AGUIA), 2021 Publication	<1 %
85	Diah Angraina, Aprillia Putri. "Analisis Sentimen Pengguna Aplikasi Google Meet Menggunakan Algoritma Support Vector Machine", Jurnal CoSciTech (Computer Science and Information Technology), 2022 Publication	<1 %
86	Eko Budi Susanto, Paminto Agung Christianto, Mohammad Reza Maulana, Satriedi Wahyu Binabar. "Analisis Kinerja Algoritma Naïve	<1 %

Bayes Pada Dataset Sentimen Masyarakat Aplikasi NEWSAKPOLE Samsat Jawa Tengah",
Jurnal CoSciTech (Computer Science and Information Technology), 2022

Publication

87

Hendry Cipta Husada, Adi Suryaputra Paramita. "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)", Teknika, 2021

Publication

<1 %

88

citeseerx.ist.psu.edu

Internet Source

<1 %

89

journal.trunojoyo.ac.id

Internet Source

<1 %

90

journal.uc.ac.id

Internet Source

<1 %

91

journal2.uad.ac.id

Internet Source

<1 %

92

library.palcomtech.com

Internet Source

<1 %

93

qdoc.tips

Internet Source

<1 %

94

repository.radenfatah.ac.id

Internet Source

<1 %

repository.univ-tridinanti.ac.id

95

Internet Source

<1 %

96

Mukhtar Mukhtar, Munawir Munawir.
"Aplikasi Decision Support System (DSS)
dengan Metode Fuzzy Multiple Attribute
Decision Making (FMADM) Studi Kasus :
AMIK Indonesia Dan STMIK Indonesia", Jurnal
JTIC (Jurnal Teknologi Informasi dan
Komunikasi), 2018

Publication

<1 %

97

Surya Rahayuda. "Evaluasi Penggunaan
Framework Laravel Pada E-government
Menggunakan ISO/IEC 25010:2011", JURNAL
IPTEKKOM : Jurnal Ilmu Pengetahuan &
Teknologi Informasi, 2017

Publication

<1 %

98

Septiyati Purwandari, Siwi Mukti Wati,
Uswatun Khasanah, Risye Insiamaufida C, Afie
Masithoh, Rinda Asmara. "Pendampingan
Masyarakat Mewujudkan Desa Tangguh
dalam Upaya Preventif Penyebaran Covid-19",
Community Empowerment, 2021

Publication

<1 %

99

kelasprogrammer.com

Internet Source

<1 %

Exclude quotes Off

Exclude bibliography Off

Exclude matches Off