

**ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP ULASAN  
PENGGUNA DI TRAVELOKA MENGGUNAKAN  
METODE SUPPORT VECTOR MACHINE**

**SKRIPSI**

Oleh:  
**MAULIDYA RAHMAH**  
**NIM. 210605110007**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2025**

**ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP ULASAN  
PENGGUNA DI TRAVELOKA MENGGUNAKAN  
METODE SUPPORT VECTOR MACHINE**

**SKRIPSI**

Diajukan kepada:  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
Untuk memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh:  
**MAULIDYA RAHMAH**  
**NIM. 210605110007**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2025**

**HALAMAN PERSETUJUAN**

**ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP ULASAN  
PENGGUNA DI TRAVELOKA MENGGUNAKAN METODE SUPPORT  
VECTOR MACHINE**

**SKRIPSI**

Oleh:  
**MAULIDYA RAHMAH**  
**NIM. 210605110007**

Telah Diperiksa dan Disetujui untuk Diuji:  
Tanggal: 15 April 2025

Pembimbing I,

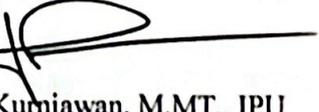
  
**Okta Comaruddin Aziz, M.Kom**  
**NIP. 19911019 201903 1 013**

Pembimbing II,

  
**Dr. Irwan Budi Santoso, M.Kom**  
**NIP. 19770103 201101 1 004**

Mengetahui,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



  
**Dr. Ir. Achrul Kurniawan, M.MT., IPU**  
**NIP. 19771020 200912 1 001**

## HALAMAN PENGESAHAN

### ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP ULASAN PENGUNA DI TRAVELOKA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

#### SKRIPSI

Oleh:

MAULIDYA RAHMAH

NIM. 210605110007

Telah Dipertahankan di Depan Dewan Penguji Skripsi  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
Untuk Memperoleh Gelar Sarjana Komputer ( S.Kom )  
Tanggal: 05 Mei 2025

#### Susunan Dewan Penguji

Ketua Penguji : Dr. Ririen Kusumawati, S.Si., M.Kom  
NIP. 19720309 200501 2 002

Anggota Penguji I : Dr. M. Ainul Yaqin, M.Kom  
NIP. 19761013 200604 1 004

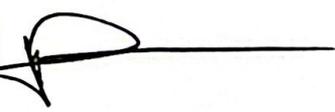
Anggota Penguji II : Okta Qomaruddin Aziz, M.Kom  
NIP. 19911019 201903 1 013

Anggota Penguji III : Dr. Irwan Budi Santoso, M.Kom  
NIP. 19770103 201101 1 004

(  )  
(  )  
(  )  
(  )

Mengetahui dan Mengesahkan,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



  
Dr. Ir. Fachrul Kurniawan, M.MT., IPU  
NIP. 19771020 200912 1 001

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Maulidya Rahmah  
NIM : 210605110007  
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika  
Judul Skripsi : Analisis Sentimen Berbasis Aspek Terhadap Ulasan Pengguna Di Traveloka Menggunakan Metode Support Vector Machine.

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 10 April 2025  
Yang membuat pernyataan,



Maulidya Rahmah  
NIM.210605110007

## **MOTTO**

*"Berpikirlah positif, tidak peduli seberapa keras kehidupanmu."*

*"Kesabaran itu ada dua macam: sabar atas sesuatu yang tidak kau ingin dan sabar menahan diri dari sesuatu yang kau ingini"*

*(Ali Bin Abi Thalib)*

## **HALAMAN PERSEMBAHAN**

Dengan rasa syukur yang mendalam kepada Allah SWT, yang telah memberikan rahmat, kesehatan, dan kekuatan sehingga skripsi ini dapat terselesaikan. Penulis mempersembahkan karya ini kepada orang paling berharga dalam hidup, ibu tercinta dan tersayang, Almh. Nurul Hidayah. Meskipun engkau telah tiada, kasih sayang dan nasihatmu akan selalu hidup dalam setiap langkah dan doa-doaku, mengiringi setiap perjuangan dan keberhasilanku. Terima kasih telah menjadi orang tua yang sempurna, menjadi teladan dalam hidupku.

Tak lupa, kepada adik saya, Maulida Rosdiana, yang selalu menjadi sumber semangat saya untuk terus berusaha menjadi kakak yang lebih baik setiap harinya. Terima kasih atas doa, dukungan, dan kebersamaan yang tak ternilai.

## KATA PENGANTAR

*Bismillahirrahmaanirrahiim, Assalamu 'alaikum wr. wb.*

Segala puji dan syukur yang tak terhingga penulis panjatkan kepada Allah SWT atas rahmat dan petunjuk-Nya. Berkat karunia-Nya, penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Analisis Sentimen Berbasis Aspek terhadap Ulasan Pengguna di Traveloka Menggunakan Metode Support Vector Machine”. Oleh karena itu, dengan penuh rendah hati, penulis mengucapkan terima kasih kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Hj. Sri Hariani, M.Si., selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Ir. Fachrul Kurniawan, M.MT, selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang, atas motivasi dan arahan yang diberikan selama proses perkuliahan.
4. Pak Okta Qomaruddin Aziz, M.Kom, selaku dosen pembimbing I yang dengan penuh kesabaran memberikan arahan, bimbingan, dan solusi dalam penulisan skripsi ini.
5. Dr. Irwan Budi Santoso, M.Kom, selaku dosen pembimbing II, yang dengan penuh kesabaran membimbing penulis hingga selesai.
6. Dr. Ririen Kusumawati, M.Kom, dan Dr. M. Ainul Yaqin, M.Kom., selaku dosen penguji saya, yang telah memberikan kritik dan saran yang berharga selama proses ujian skripsi.

7. Seluruh staf dan dosen Program Studi Teknik Informatika, atas ilmu, dukungan, dan fasilitas yang telah diberikan kepada penulis selama masa studi.
8. Kedua orang tua yang penulis hormati dan sayangi, Almh. Nurul Hidayah selaku ibu penulis yang telah meninggal pada Januari 2021. Dan untuk Alm. Achmad Suhairi selaku ayah penulis yang telah meninggal pada Agustus 2021, terima kasih telah menjadi orang tua yang luar biasa di sisa umur terakhirnya, terima kasih untuk semua kasih sayang, cinta, pengorbanan, perjuangan dan pelajaran hidup bagi penulis.
9. Teruntuk adik penulis, Maulida Rosdiana, yang telah mendoakan, menjadi tempat bercerita, dan memberi semangat yang menjadikan penulis terus berusaha menjadi kakak yang lebih baik.
10. Untuk Prasetyo Putra Pratama, terima kasih telah menemani dan menjadi pendukung yang baik. Kesediaan untuk selalu mendengarkan setiap keluh kesah, dan selalu memberikan semangat untuk penulis.
11. Sahabat semasa perkuliahan, Nabila Mahdiya Putri, yang ikut menemani proses pembelajaran dan selalu memberikan semangat kepada penulis. Semoga persahabatan ini tetap erat dan terjaga selamanya.
12. Kepada rekan-rekan Badan Pengurus Harian “BPH Lope Sudah Demis”, yaitu Reyhan, Gianda, Aisha, Renata, dan Atsila, terima kasih atas kebersamaan serta kerja sama yang luar biasa selama menjalankan tugas di HMPS Teknik Informatika “Encoder”. Semoga kebersamaan dan

pengalaman yang telah kita bangun tetap menjadi kenangan berharga dan komunikasi yang telah terjalin dapat terus berlanjut tanpa batas waktu.

13. Kepada rekan-rekan kerja Ruangguru, khususnya Tim English Academy Center Malang, terima kasih atas dukungan dan pengertiannya yang telah memberikan izin kepada penulis untuk tetap bisa mengikuti bimbingan, ujian komprehensif, dan sidang skripsi di sela waktu jam kerja.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan, sehingga masukan, kritik, dan saran yang membangun dari pembaca sangat diharapkan untuk perbaikan di masa mendatang. Semoga karya ini dapat bermanfaat bagi para pembaca sekalian serta memberikan kontribusi positif dalam pengembangan ilmu pengetahuan.

Malang, 10 April 2025

Penulis

## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	1
<b>HALAMAN PENGAJUAN</b> .....	ii
<b>HALAMAN PERSETUJUAN</b> .....	iii
<b>HALAMAN PENGESAHAN</b> .....	iv
<b>PERNYATAAN KEASLIAN TULISAN</b> .....	v
<b>MOTTO</b> .....	vi
<b>HALAMAN PERSEMBAHAN</b> .....	vii
<b>KATA PENGANTAR</b> .....	viii
<b>DAFTAR ISI</b> .....	xi
<b>DAFTAR GAMBAR</b> .....	xiii
<b>DAFTAR TABEL</b> .....	xiv
<b>ABSTRAK</b> .....	xvi
<b>ABSTRACT</b> .....	xvii
البحث مستخلص .....	xviii
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	7
1.3 Tujuan Penelitian .....	7
1.4 Batasan Penelitian .....	7
1.5 Manfaat Penelitian .....	8
<b>BAB II STUDI PUSTAKA</b> .....	9
2.1 Online Travel Agencies .....	9
2.2 Analisis Sentimen Berbasis Aspek .....	13
2.3 <i>Preprocessing Data</i> .....	17
2.4 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> .....	17
2.5 <i>Support Vector Machine</i> .....	18
<b>BAB III DESAIN DAN IMPLEMENTASI</b> .....	23
3.1 Desain Sistem .....	23
3.2 Pengumpulan Data .....	24
3.3 <i>Preprocessing Data</i> .....	26
3.4 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> .....	28
3.5 <i>Support Vector Machine</i> .....	31
3.6 Evaluasi Pengujian .....	38
3.7 Skenario Pengujian .....	41
<b>BAB IV UJI COBA DAN PEMBAHASAN</b> .....	44
4.1 Hasil Hasil Uji Coba .....	44
4.1.1 Hasil Skenario Uji 1 .....	44
4.1.2 Hasil Skenario Uji 2 .....	46
4.1.3 Hasil Skenario Uji 3 .....	47
4.1.4 Hasil Skenario Uji 4 .....	52
4.2 Pembahasan .....	57
4.2.1 Pembahasan Skenario Uji 1 .....	58

4.2.2 Pembahasan Skenario Uji 2.....	60
4.2.3 Pembahasan Skenario Uji 3.....	63
4.2.4 Pembahasan Skenario Uji 4.....	67
4.3 Integrasi Islam.....	81
<b>BAB V KESIMPULAN DAN SARAN</b> .....	<b>85</b>
5.1 Kesimpulan .....	85
5.2 Saran .....	86
<b>DAFTAR PUSTAKA</b>	
<b>LAMPIRAN</b>	

## DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	23
Gambar 3.2 Kode Preprocessing Data .....	28
Gambar 3.3 Kode Term Frequency-Inverse Document Frequency .....	31
Gambar 3.4 Desain Sistem Penelitian.....	33
Gambar 3.5 Flowchart Sequential Minimal Optimization .....	34
Gambar 3.6 Kode Support Vector Machine Optimasi Sequential Minimal Optimization .....	38
Gambar 3.7 Kode Confusion Matrix.....	41
Gambar 4. 1 Hasil Visualisasi Skenario Uji 1.....	45
Gambar 4. 2 Visual Performa Skenario Uji 1 Kernel Linear Single Stage Classification .....	46
Gambar 4. 3 Hasil Visualisasi Skenario Uji 2.....	46
Gambar 4. 4 Visual Performa Skenario Uji 2 Kernel Polinomial Single Stage Classification .....	47
Gambar 4. 5 Hasil Visualisasi Klasifikasi Aspek Skenario Uji 3 .....	48
Gambar 4. 6 Hasil Visualisasi Data dari Aspek Harga Skenario Uji 3.....	49
Gambar 4. 7 Hasil Visualisasi Data dari Aspek Keamanan Skenario Uji 3 .....	50
Gambar 4. 8 Hasil Visualisasi Data dari Aspek Layanan Skenario Uji 3.....	51
Gambar 4. 9 Visual Performa Skenario Uji 3 Kernel Linear Multistage Classification .....	52
Gambar 4. 10 Hasil Visualisasi Klasifikasi Aspek Skenario Uji 4.....	53
Gambar 4. 11 Hasil Visualisasi Data dari Aspek Harga Skenario Uji 4.....	54
Gambar 4. 12 Hasil Visualisasi Data dari Aspek Keamanan Skenario Uji 4 .....	55
Gambar 4. 13 Hasil Visualisasi Data dari Aspek Layanan Skenario Uji 4.....	56
Gambar 4. 14 Visual Performa Skenario Uji 4 Kernel Polinomial Multistage Classification .....	57
Gambar 4. 15 Visual Performa Parameter Skenario Uji 1 .....	58
Gambar 4. 16 Visual Performa Parameter Skenario Uji 2 .....	61
Gambar 4. 17 Visual Performa Parameter Skenario Uji 3 .....	66
Gambar 4. 18 Visual Performa Parameter Skenario Uji 4 .....	78
Gambar 4. 19 Visual Performa Klasifikasi Single Stage Classification dan Multistage Classification .....	79

## DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu Tentang Online Travel Agencies .....	11
Tabel 2.2 Penelitian Terdahulu Analisis Sentimen Berbasis Aspek .....	14
Tabel 2.3 Penelitian Terdahulu Tentang Support Vector Machine.....	20
Tabel 3.1 Atribut Data.....	24
Tabel 3.2 Deskripsi Tabel Kategori .....	25
Tabel 3.3 Contoh Data .....	25
Tabel 3.4 Proses Cleansing .....	26
Tabel 3.5 Proses Case Folding.....	27
Tabel 3.6 Proses Tokenizing.....	27
Tabel 3.7 Proses Stopword Removal .....	27
Tabel 3.8 Proses Stemming.....	28
Tabel 3.9 Contoh Perhitungan TF-IDF .....	30
Tabel 3.10 Contoh Confusion Matrix 3 kelas .....	39
Tabel 3.11 Contoh Confusion Matrix 2 kelas .....	40
Tabel 3.17 Skenario Pengujian .....	42
Tabel 4. 1 Score Precision, Recall, F1-Score pada 6 Kelas Skenario Uji 1 .....	45
Tabel 4. 2 Score Precision, Recall, F1-Score pada 6 Kelas Skenario Uji 2.....	47
Tabel 4. 3 Score Precision, Recall, F1-Score pada 3 Kelas Skenario Uji 3.....	48
Tabel 4. 4 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Harga Skenario Uji 3 .....	49
Tabel 4. 5 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Keamanan Skenario Uji 3.....	50
Tabel 4. 6 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Layanan Skenario Uji 3.....	51
Tabel 4. 7 Performa Hasil Uji dengan Kernel Linear dengan Metode Multistage classification .....	51
Tabel 4. 8 Score Precision, Recall, F1-Score pada 3 Kelas Skenario Uji 4.....	53
Tabel 4. 9 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Harga Skenario Uji 4 .....	54
Tabel 4. 10 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Keamanan Skenario Uji 4.....	55
Tabel 4. 11 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Layanan Skenario Uji 4.....	56
Tabel 4. 12 Performa Hasil Uji dengan Kernel Polinomial dengan Metode Multistage classification .....	56
Tabel 4. 13 Performa Uji Hyperparameter pada Skenario Uji 1 .....	58
Tabel 4. 14 Performa Hasil Uji Aspek Harga, Keamanan, Layanan pada Skenario Uji Pertama .....	59
Tabel 4. 15 Performa Uji Hyperparameter pada Skenario Uji 2.....	60
Tabel 4. 16 Performa Hasil Uji Aspek Harga, Keamanan, Layanan pada Skenario Uji 2 .....	62
Tabel 4. 17 Performa Hasil Uji Klasifikasi Aspek pada Skenario Uji 3.....	63

Tabel 4. 18 Performa Hasil Uji Klasifikasi Sentimen Aspek Harga pada Skenario Uji 3.....	64
Tabel 4. 19 Performa Hasil Uji Klasifikasi Sentimen Aspek Keamanan pada Skenario Uji 3.....	65
Tabel 4. 20 Performa Hasil Uji Klasifikasi Sentimen Aspek Layanan pada Skenario Uji 3.....	65
Tabel 4. 21 Performa Hasil Uji Aspek Harga, Keamanan, Layanan pada Skenario Uji 3.....	67
Tabel 4. 22 Performa Hasil Uji Klasifikasi Aspek pada Skenario Uji 4.....	67
Tabel 4. 23 Performa Hasil Uji Klasifikasi Sentimen Aspek Harga pada Skenario Uji 4.....	72
Tabel 4. 24 Performa Hasil Uji Klasifikasi Sentimen Aspek Keamanan pada Skenario Uji 4.....	74
Tabel 4. 25 Performa Hasil Uji Klasifikasi Sentimen Aspek Layanan pada Skenario Uji 4.....	76
Tabel 4. 26 Performa Hasil Uji Aspek Harga, Keamanan, Layanan pada Skenario Uji 4.....	79

## ABSTRAK

Rahmah, Maulidya. 2025. **Analisis Sentimen Berbasis Aspek terhadap Ulasan Pengguna di Traveloka Menggunakan Metode Support Vector Machine.** Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Okta Qomaruddin Aziz, M.Kom (II) Dr. Irwan Budi Santoso, M.Kom.

**Kata kunci:** Analisis Sentimen Berbasis Aspek, Support Vector Machine, Multi-Stage Classification, Single Stage Classification, Sequential Minimal Optimization

Analisis sentimen berbasis aspek terhadap ulasan pengguna di Traveloka digunakan untuk mengidentifikasi sentimen positif dan negatif pada aspek layanan, harga, dan keamanan. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi sentimen berbasis aspek menggunakan metode Support Vector Machine (SVM) dengan optimasi Sequential Minimal Optimization (SMO) serta teknik ekstraksi fitur TF-IDF. Data diperoleh melalui scraping ulasan pengguna di Google Play Store dan melalui tahapan preprocessing serta klasifikasi aspek. Hasil penelitian menunjukkan bahwa skenario percobaan terbaik diperoleh pada skenario 3 dengan pendekatan multi-stage classification. Pada tahap klasifikasi aspek, parameter  $C = 0.01$  dan nilai toleransi  $= 0.01$  menghasilkan akurasi sebesar 86%. Selanjutnya, pada tahap klasifikasi sentimen, aspek harga mencapai akurasi 87% dengan parameter  $C = 1$  dan nilai toleransi 0,01, aspek keamanan mencapai akurasi 88% dengan parameter  $C = 0.1$  dengan nilai toleransi  $= 0.01$ , dan aspek layanan mencapai akurasi tertinggi sebesar 91% dengan parameter  $C = 1$  dengan nilai toleransi 0.01. Implementasi metode ini memberikan hasil analisis yang lebih akurat, sehingga dapat digunakan untuk memahami preferensi pengguna serta membantu perusahaan dalam meningkatkan kualitas layanan.

## ABSTRACT

Rahmah, Maulidya. 2025. **Analisis Sentimen Berbasis Aspek terhadap Ulasan Pengguna di Traveloka Menggunakan Metode Support Vector Machine.** Thesis. Informatics Engineering Study Program, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University Malang. Supervisor: (I) Okta Qomaruddin Aziz, M.Kom (II) Dr. Irwan Budi Santoso, M.Kom.

**Key words:** Aspect-Based Sentiment Analysis, Support Vector Machine, Multi-Stage Classification, Single-Stage Classification, Sequential Minimal Optimization.

Aspect-based sentiment analysis of user reviews on Traveloka is used to identify positive and negative sentiments regarding the aspects of service, price, and security. This study aims to develop an aspect-based sentiment classification system using the Support Vector Machine (SVM) method optimized with Sequential Minimal Optimization (SMO) and the TF-IDF feature extraction technique. The data was obtained through scraping user reviews on the Google Play Store and processed through preprocessing and aspect classification stages. The results show that the best experimental scenario was achieved in Scenario 3 using a multi-stage classification approach. In the aspect classification stage, the parameters  $C = 0.01$  and tolerance = 0.01 produced an accuracy of 86%. Furthermore, in the sentiment classification stage, the price aspect achieved an accuracy of 87% with  $C = 1$  and tolerance = 0.01, the security aspect reached 88% accuracy with  $C = 0.1$  and tolerance = 0.01, and the service aspect achieved the highest accuracy of 91% with  $C = 1$  and tolerance = 0.01. The implementation of this method provides more accurate sentiment analysis results, which can be used to understand user preferences and assist companies in improving service quality.

## البحث مستخلص

.رحمة، مولدية. 2025. تحليل المشاعر المستندة إلى الجوانب لمراجعات المستخدمين على ترافلوكا باستخدام طريقة آلة دعم المتجهات الأطروحة. قسم هندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم مالانج الإسلامية الحكومية المشرف: (الأول) أوكتا قمر الدين عزيز، م. كوم (الثاني) الدكتور إروان بودي سانتوسو، م. كوم در

**الكلمات المفتاحية:** تحليل المشاعر المستند إلى الجوانب، وآلة دعم المتجهات، والتصنيف متعدد المراحل والتصنيف أحادي المرحلة والتحسين الأدي المتسلسل

يُستخدم لتحديد المشاعر Traveloka في تحليل المشاعر المستند إلى الجوانب في تقييمات المستخدمين على منصة الإيجابية والسلبية المتعلقة بجوانب الخدمة، والسعر، والأمان. يهدف هذا البحث إلى تطوير نظام تصنيف للمشاعر المستندة إلى الجوانب وتقنية استخراج الميزات (SMO) مع تحسين عبر خوارزمية التحسين الأدي المتسلسلي (SVM) باستخدام خوارزمية آلة الدعم الناقل وتمت معالجتها عبر Google Play، تم الحصول على البيانات من خلال تجميع تقييمات المستخدمين من متجر TF-IDF. مراحل التحضير المسبق وتصنيف الجوانب. أظهرت نتائج البحث أن أفضل سيناريو تجريبي تم تحقيقه في السيناريو الثالث باستخدام قيمة التسامح = 0.01 إلى تحقيق  $C = 0.01$  نُجج التصنيف متعدد المراحل. في مرحلة تصنيف الجوانب، أدى استخدام المعامل، وتسامح =  $C = 1$  0.01 = 0.86%. بعد ذلك، في مرحلة تصنيف المشاعر، حقق جانب السعر دقة بلغت 87% باستخدام % وتسامح = 0.01، بينما حقق جانب الخدمة أعلى دقة بلغت  $C = 0.1$  91 = 0.88% باستخدام وتسامح = 0.01. إن تطبيق هذه الطريقة يوفر نتائج أكثر دقة في تحليل المشاعر، مما يساعد في فهم تفضيلات  $C = 1$  باستخدام المستخدمين ودعم الشركات في تحسين جودة خدماتها.

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Sebelum adanya teknologi digital, pemesanan hotel, tiket kendaraan, dan layanan perjalanan lainnya dilakukan secara manual. Calon wisatawan harus mengunjungi agen perjalanan atau langsung ke tempat layanan terkait untuk melakukan pemesanan. Proses ini seringkali memakan waktu dan tidak praktis, terutama bagi mereka yang tinggal jauh dari pusat layanan (Miftahul & Slamet, L., 2020). Proses pemesanan manual memiliki beberapa kelebihan dan kekurangan. Kelebihannya termasuk interaksi langsung dengan agen yang dapat memberikan saran dan rekomendasi. Namun, kekurangannya mencakup keterbatasan informasi, waktu yang dibutuhkan untuk proses pemesanan, dan risiko kesalahan manusia. Selain itu, efisiensi pemesanan manual sangat rendah dibandingkan dengan metode modern (Wibisono, 2022).

Saat ini, pemesanan hotel, tiket kendaraan, dan layanan perjalanan lainnya telah bertransformasi dengan penggunaan teknologi digital. Melalui platform *social media*, pengguna dapat dengan mudah memesan layanan perjalanan secara *online* (Indria, T., & Sari, I., 2021). Hanya dengan beberapa klik, pengguna bisa membandingkan harga, melihat ulasan, dan melakukan pemesanan kapan saja dan di mana saja. Kelebihan pemesanan di era digital meliputi kemudahan akses, kecepatan, dan ketersediaan informasi yang lebih lengkap. Namun, kekurangannya termasuk risiko keamanan data dan kemungkinan adanya kesalahan teknis. Meski

demikian, efisiensi pemesanan secara digital jauh lebih tinggi dibandingkan dengan metode tradisional, membuat proses perjalanan lebih nyaman dan terorganisir.

Kemajuan teknologi digital telah memberikan dampak signifikan pada berbagai industri, termasuk pariwisata. *Online Travel Agencies* adalah platform digital yang menyediakan berbagai layanan perjalanan seperti pemesanan tiket pesawat, hotel, dan paket wisata (Iswahyuniarto, 2023). *Online Travel Agencies* memungkinkan pengguna untuk merencanakan dan memesan perjalanan secara online dengan mudah. Jenis-jenis pelayanan *Online Travel Agencies* sangat beragam, seperti fokus pada pemesanan hotel, tiket pesawat, atau kombinasi dari keduanya. Beberapa *Online Travel Agencies* juga menawarkan paket wisata lengkap yang mencakup transportasi, akomodasi, dan aktivitas selama perjalanan. Setiap jenis *Online Travel Agencies* memiliki spesialisasi dan keunggulan tersendiri dalam memenuhi kebutuhan pengguna. Contoh *Online Travel Agencies* yang terkenal di Asia Tenggara adalah Traveloka (Putro, 2022). Selain Traveloka, ada juga *Online Travel Agencies* lainnya seperti Booking.com dan Expedia yang populer secara global. Platform-platform ini memanfaatkan teknologi digital untuk memberikan kemudahan dan kenyamanan bagi pengguna dalam merencanakan perjalanan mereka.

Pemahaman tentang kepuasan pelanggan di antara manajer penyedia layanan juga menjadi bagian penting dalam mendukung transformasi layanan berbasis teknologi. Meskipun teknologi digital telah memudahkan proses pemesanan dan memberikan kemudahan akses informasi bagi pelanggan, banyak manajer hotel yang belum sepenuhnya memahami faktor-faktor utama yang memengaruhi tingkat

kepuasan tersebut. Faktor seperti kualitas layanan, kebersihan, fasilitas, dan perilaku staf sering kali tidak mendapatkan prioritas yang tepat dalam pengambilan Keputusan (Mohanty, 2023).

*Online Travel Agencies* seperti Traveloka telah memanfaatkan teknologi untuk memberikan kemudahan kepada pengguna dalam merencanakan dan memesan perjalanan. Salah satu keunggulan utama platform ini adalah kemampuannya mengumpulkan data dari ulasan pengguna, yang dapat menjadi sumber informasi berharga bagi manajer penyedia layanan. Dengan semakin populernya platform *Online Travel Agencies*, jumlah ulasan pengguna yang tersedia juga meningkat secara signifikan. Ulasan ini tidak hanya sekadar memberikan penilaian umum, tetapi juga menawarkan wawasan yang lebih rinci mengenai pengalaman pengguna pada berbagai aspek layanan yang disediakan. Analisis ulasan pengguna menjadi penting untuk memahami persepsi mereka, sehingga dapat digunakan untuk meningkatkan kualitas layanan (Novita & Helena, 2021).

Namun, untuk memanfaatkan potensi besar dari data ulasan pengguna, diperlukan pendekatan analisis yang efisien. Seiring dengan bertambahnya jumlah ulasan yang tersedia, melakukan analisis secara manual menjadi tidak efisien dan memakan waktu. Menurut laporan dari Statista (2023), jumlah ulasan pengguna di platform *e-commerce* dan aplikasi perjalanan terus meningkat pesat, dengan Traveloka mencatatkan lebih dari 10 juta ulasan di *Google Play Store* pada tahun 2023. Dengan volume data sebesar ini, analisis manual akan membutuhkan waktu sehari-hari bahkan minggu untuk memproses setiap ulasan secara individu. Hal ini

tidak hanya membebani sumber daya manusia tetapi juga mengurangi efektivitas dalam mengambil keputusan yang cepat dan akurat. Selain itu, analisis manual juga rentan terhadap bias subjektif dan inkonsistensi. Dengan demikian, diperlukan pendekatan yang lebih canggih untuk mengotomatisasi proses analisis ulasan ini. Dalam konteks ini, metode *machine learning* seperti *Support Vector Machine* menjadi pilihan yang tepat karena kemampuannya dalam menangani data teks yang kompleks dan menghasilkan hasil klasifikasi yang akurat (Jordan Sihombing & Irvandi Sitorus, 2022).

*Support Vector Machine* dikenal karena kemampuannya dalam menangani masalah klasifikasi, terutama dalam konteks data yang memiliki dimensi tinggi seperti teks. *Support Vector Machine* bekerja dengan mencari hyperplane yang optimal untuk memisahkan data ke dalam kelas yang berbeda. Dalam analisis sentimen, *Support Vector Machine* dapat digunakan untuk mengklasifikasikan ulasan pengguna menjadi sentimen positif, negatif, atau netral (Handayani, 2021). Keunggulan *Support Vector Machine* dalam generalisasi terhadap data yang tidak terlihat membuatnya menjadi pilihan yang baik untuk analisis ulasan pengguna, di mana variasi dalam bahasa dan konteks sangat umum ditemui (Jordan Sihombing & Irvandi Sitorus, 2022).

Untuk meningkatkan efektivitas *Support Vector Machine* dalam analisis teks, digunakanlah teknik ekstraksi fitur TF-IDF (*Term Frequency-Inverse Document Frequency*). TF-IDF membantu mengidentifikasi kata-kata yang paling relevan dalam sebuah dokumen dengan memberikan bobot yang lebih tinggi pada kata-kata yang lebih sering muncul dalam satu dokumen, tetapi jarang muncul dalam

dokumen lain (Paulina, 2020). Dengan demikian, TF-IDF membantu mengurangi pengaruh kata-kata umum yang kurang informatif, sehingga meningkatkan akurasi analisis sentiment.

Implementasi metode *Support Vector Machine* dengan TF-IDF sebagai teknik ekstraksi fitur telah terbukti memberikan hasil yang konsisten dan akurat dalam berbagai studi kasus (Novita & Helena, 2021). Metode ini memungkinkan untuk mengotomatisasi analisis ulasan pengguna secara efisien dan efektif, tanpa mengorbankan akurasi. Dalam konteks Traveloka, analisis ulasan pengguna dengan metode ini dapat memberikan wawasan berharga mengenai aspek-aspek layanan yang perlu ditingkatkan.

Penggunaan *Support Vector Machine* dan TF-IDF juga sejalan dengan prinsip-prinsip yang diajarkan dalam Islam, terutama dalam hal menjaga keseimbangan dan ketepatan dalam segala hal. Islam mengajarkan pentingnya berbicara dengan kata-kata yang baik dan benar, serta menjaga keseimbangan dalam interaksi sosial. Prinsip ini dapat diterapkan dalam konteks analisis ulasan pengguna, di mana setiap ulasan diperlakukan dengan ketelitian dan perhatian yang sama, memastikan bahwa analisis dilakukan dengan adil dan akurat.

Pada perspektif Islam, umat Muslim diajarkan untuk berbicara dengan kata-kata yang baik dan menghindari perkataan kasar atau menyakitkan. Ini karena kata-kata yang baik dapat membawa kebaikan dan mempererat hubungan antar sesama manusia, sementara kata-kata yang kasar dapat menyebabkan permusuhan dan keretakan dalam hubungan sosial. Islam menekankan pentingnya menjaga lisan dan

menggunakan kata-kata yang baik dalam berinteraksi dengan orang lain. Allah SWT berfirman dalam Surah Al-Isra (17:53):

وَقُلْ لِعِبَادِي يَقُولُوا الَّتِي هِيَ أَحْسَنُ إِنَّ الشَّيْطَانَ يَنْزِعُ بَيْنَهُمْ إِنَّ الشَّيْطَانَ كَانَ  
لِلْإِنْسَانِ عَدُوًّا مُّبِينًا

*"Dan katakanlah kepada hamba-hamba-Ku, 'Hendaklah mereka mengucapkan perkataan yang lebih baik (benar).' Sesungguhnya syaitan itu menimbulkan perselisihan di antara mereka. Sesungguhnya syaitan adalah musuh yang nyata bagi manusia." (Q.S. Al-Isra: 53).*

Ayat ini memerintahkan umat Muslim untuk selalu berbicara dengan perkataan yang baik dan benar, menghindari perselisihan yang dihasut oleh syaitan. Pentingnya analisis ulasan pengguna tidak dapat diabaikan dalam industri pariwisata yang sangat kompetitif. Pengalaman pengguna merupakan faktor kunci yang menentukan kesuksesan sebuah platform *Online Travel Agencies*. Dengan memahami persepsi dan preferensi pengguna, perusahaan dapat membuat keputusan yang lebih baik dalam meningkatkan layanan. Oleh karena itu, penelitian ini menekankan perlunya penerapan metode yang tepat, seperti metode *Support Vector Machine* dan TF-IDF, untuk membantu perusahaan dalam mengidentifikasi area-area yang perlu diperbaiki, sehingga mereka dapat terus bersaing dalam pasar yang semakin ketat.

Dengan mempertimbangkan semua faktor ini, dapat disimpulkan bahwa penggunaan *Support Vector Machine* dalam analisis ulasan pengguna merupakan langkah yang tepat dan relevan. Pendekatan ini tidak hanya memberikan hasil yang akurat tetapi juga sejalan dengan prinsip-prinsip Islam dalam menjaga keseimbangan dan ketepatan. Oleh karena itu, penelitian ini memberikan kontribusi yang signifikan dalam memahami pentingnya teknologi dalam analisis sentimen,

serta bagaimana teknologi ini dapat digunakan untuk meningkatkan kualitas layanan dalam industri pariwisata.

## **1.2 Rumusan Masalah**

Berikut merupakan rumusan masalah pada penelitian ini:

1. Bagaimana hasil analisis ulasan pengguna di Traveloka untuk mengidentifikasi sentiment positif dan negatif menggunakan metode *Support Vector Machine*?
2. Bagaimana cara meningkatkan efisiensi dalam analisis sentimen terhadap komentar pengguna pada platform Traveloka?

## **1.3 Tujuan Penelitian**

1. Menganalisis ulasan pengguna pada *platform* Traveloka dan mengidentifikasi sentimen positif dan negatif.
2. Meningkatkan efisiensi pengguna terhadap aspek layanan, harga, dan keamanan pada *platform* Traveloka.

## **1.4 Batasan Penelitian**

Berikut batasan masalah yang digunakan dalam penelitian ini:

1. Analisis terbatas pada ulasan pengguna terkait dengan aspek layanan, harga, dan keamanan.
2. Data yang digunakan adalah ulasan dalam bahasa Indonesia.
3. Data yang diambil merupakan komentar dari tahun 2020-2024.
4. Data yang digunakan sebanyak 1.000 pada setiap aspek.

## 1.5 Manfaat Penelitian

Beberapa manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Bagi pengguna, penelitian ini memberikan wawasan yang mendalam tentang bagaimana platform *Online Travel Agencies* seperti Traveloka dapat lebih memahami dan memenuhi kebutuhan.
2. Bagi perusahaan *Online Travel Agencies*, penelitian ini menyediakan metode analisis yang efisien dan akurat untuk meningkatkan layanan berdasarkan umpan balik pengguna

## **BAB II**

### **STUDI PUSTAKA**

#### **2.1 *Online Travel Agencies***

*Online Travel Agencies* merupakan platform digital yang secara khusus dirancang untuk memudahkan pengguna dalam melakukan pemesanan berbagai kebutuhan perjalanan, termasuk tiket pesawat, hotel, penyewaan mobil, serta paket wisata (Iswahyuniarto, 2023). Platform ini memungkinkan pengguna untuk mengakses berbagai layanan perjalanan secara *online* tanpa perlu mengunjungi agen perjalanan fisik, menjadikannya solusi yang lebih efisien dan praktis dibandingkan dengan metode pemesanan tradisional. Kemajuan teknologi digital telah memainkan peran penting dalam transformasi *Online Travel Agencies*, mengubahnya dari sekadar alternatif untuk pemesanan manual menjadi komponen esensial dalam industri pariwisata modern.

Sebelum adanya *Online Travel Agencies*, pemesanan perjalanan dilakukan secara manual dengan mengunjungi agen perjalanan atau tempat layanan terkait. Proses ini sering kali memakan waktu dan bergantung pada interaksi langsung dengan agen. Sementara interaksi ini memberikan keuntungan berupa saran dan rekomendasi yang dipersonalisasi, metode manual memiliki berbagai kekurangan, termasuk keterbatasan informasi yang tersedia dan waktu yang dibutuhkan untuk menyelesaikan pemesanan (Putro G., 2020). Dengan hadirnya *Online Travel Agencies*, banyak dari tantangan ini dapat diatasi. Pengguna kini dapat membandingkan berbagai pilihan layanan perjalanan, seperti harga tiket pesawat,

jenis akomodasi, dan fasilitas tambahan lainnya, hanya dengan beberapa klik dari perangkat mereka (Rosyidin & Prihatini, 2020).

Platform *Online Travel Agencies* seperti Traveloka, Booking.com, Expedia, Agoda, dan Airbnb telah menjadi pemain utama dalam industri pariwisata, berkat kemampuan mereka untuk menawarkan akses yang cepat dan mudah ke berbagai layanan perjalanan. Pengguna dapat mencari, membandingkan, dan memesan layanan perjalanan kapan saja dan di mana saja, tanpa perlu terikat oleh waktu operasional agen perjalanan fisik. *Online Travel Agencies* juga menawarkan kelebihan berupa kemudahan akses ke berbagai informasi, termasuk ulasan pengguna, foto, serta detail spesifik tentang layanan yang tersedia. Dengan semua informasi ini berada di ujung jari pengguna, mereka dapat membuat keputusan yang lebih informasi, yang pada akhirnya meningkatkan pengalaman perjalanan mereka.

*Online Travel Agencies* memainkan peran penting dalam pengumpulan dan analisis data ulasan pengguna, yang tidak hanya memberikan wawasan langsung kepada calon pelanggan, tetapi juga menjadi sumber data berharga bagi perusahaan *Online Travel Agencies*. Melalui analisis ulasan ini, perusahaan dapat memahami tren, preferensi, dan tingkat kepuasan pelanggan terhadap berbagai layanan, memungkinkan mereka untuk mengidentifikasi area yang perlu diperbaiki dan mengembangkan strategi peningkatan layanan (Putro G., 2020). Pengaruh *Online Travel Agencies* terhadap industri pariwisata global sangat signifikan, mendorong perusahaan-perusahaan pariwisata tradisional untuk beradaptasi dengan tren digital agar tetap relevan. Dengan memberikan akses yang lebih luas dan beragam pilihan kepada wisatawan, *Online Travel Agencies* telah membantu meningkatkan jumlah

wisatawan secara global, yang pada akhirnya memberikan dampak positif bagi perekonomian destinasi wisata.

Tabel 2.1 Penelitian Terdahulu Tentang *Online Travel Agencies*

No.	Referensi Penelitian	Input	Metode yang Digunakan	Hasil
1.	Pratama, Y. T., Bachtiar, F. A., & Setiawan, N. Y. (2021)	Ulasan pengguna aplikasi booking hotel	<i>Support Vector Machine</i>	Menganalisis sentimen terhadap aplikasi booking hotel dengan akurasi klasifikasi rata-rata 85% dan visualisasi data yang mendukung evaluasi dan pengawasan destinasi wisata.
2.	Naquitasia, R., Fudholi, D. H., & Iswari, L. (2022)	Ulasan pengguna tentang aspek wisata halal	<i>Convolutional Neural Networks (CNN)</i>	Menilai sentimen terhadap aspek wisata halal dengan akurasi klasifikasi aspek sebesar 98.299% dan sentimen sebesar 93.96%.
3.	Radiena, G., & Nugroho, A. (2023)	Ulasan pengguna aplikasi KAI Access	<i>Support Vector Machine</i>	Evaluasi sentimen aplikasi KAI Access menunjukkan sentimen positif pada aspek kepuasan dan negatif pada aspek learnability, efisiensi, dan error.
4.	Wulandari, R., & Mulyani, Y. (2023)	Ulasan pengguna terhadap layanan <i>Online Travel Agencies</i>	<i>Random Forest</i>	Menganalisis sentimen ulasan pengguna terhadap layanan <i>Online Travel Agencies</i> , dengan hasil akurasi klasifikasi aspek sebesar 90% dan pengidentifikasian area perbaikan layanan.
5.	Prasetio, A., Nursandi, W. (2022)	Data survei faktor UTAUT 2 dari pengguna Tiket.com	<i>Structural Equation Modelling (SEM)</i>	Enam faktor UTAUT 2 memengaruhi minat penggunaan website Tiket.com, dengan Behavioral Intention dan Habit berdampak positif pada Use Behavior. Gender tidak memoderasi, sedangkan usia hanya memoderasi Price Value terhadap Behavioral Intention.

Dari penjelasan tabel 2.1 mengenai penelitian terdahulu terkait analisis terhadap layanan *Online Travel Agencies* menunjukkan beragam pendekatan metode yang digunakan serta hasil yang bervariasi. Pratama (2021) menggunakan

*Support Vector Machine* untuk menganalisis sentimen terhadap aplikasi booking hotel, dengan hasil akurasi klasifikasi rata-rata mencapai 85%. Penelitian ini juga menekankan pentingnya visualisasi data dalam mendukung evaluasi dan pengawasan destinasi wisata. Naquitasia (2022) menerapkan *Convolutional Neural Networks* (CNN) dalam analisis sentimen terhadap aspek wisata halal. Mereka mencapai akurasi klasifikasi aspek sebesar 98.299% dan sentimen sebesar 93.96%, yang menunjukkan efektivitas CNN dalam menangani sentimen spesifik pada aspek tertentu.

Penelitian Radiena dan Nugroho (2023) juga menggunakan *Support Vector Machine* untuk mengevaluasi aplikasi KAI Access. Hasilnya menunjukkan sentimen positif terhadap kepuasan pengguna, tetapi negatif pada aspek learnability, efisiensi, dan *error*. Wulandari dan Mulyani (2023), dalam penelitian mereka menggunakan *Random Forest*, berhasil mengidentifikasi sentimen pengguna layanan *Online Travel Agencies* dengan akurasi klasifikasi aspek sebesar 90%, serta mengidentifikasi area layanan yang memerlukan perbaikan. Prasetio dan Nursandi (2022) menggunakan SEM untuk menganalisis faktor UTAUT 2 pada minat penggunaan website Tiket.com. Enam faktor signifikan, dengan Behavioral Intention dan Habit berdampak positif pada *Use Behavior*. Gender tidak memoderasi, sedangkan usia hanya memoderasi *Price Value* terhadap Behavioral Intention. Secara keseluruhan, penelitian-penelitian ini memperlihatkan berbagai metode yang efektif untuk menganalisis pengguna *Online Travel Agencies*.

Kebaruan penelitian ini adalah objek dan data yang akan diuji, yaitu penggunaan data ulasan pengguna dari platform Traveloka dengan fokus pada

aspek layanan, harga, dan keamanan. Penelitian ini menggunakan kombinasi metode *Support Vector Machine* dan ekstraksi fitur TF-IDF untuk menganalisis sentimen berbasis aspek.

## 2.2 Analisis Sentimen Berbasis Aspek

Analisis Sentimen Berbasis Aspek (ASBA) adalah metode yang digunakan untuk mengevaluasi sentimen terhadap berbagai aspek spesifik dari entitas seperti produk, layanan, atau fitur. Pendekatan ini memungkinkan analisis sentimen secara lebih mendetail dibandingkan dengan analisis sentimen umum yang hanya mengklasifikasikan sentimen keseluruhan sebagai positif, negatif, atau netral. Metode ini memecah teks menjadi komponen-komponen kecil, seperti fitur produk atau layanan, dan menilai sentimen terhadap masing-masing komponen secara individu (Ismet, 2022). Pada penelitian ini, analisis sentimen hanya mencakup sentimen positif dan negatif, tanpa memasukkan sentimen netral. Hal ini dikarenakan sentimen netral cenderung tidak memberikan performa yang signifikan dalam analisis, sehingga fokus hanya pada sentimen yang lebih jelas dan relevan, yaitu positif dan negatif (Galih Pradana, 2020).

Dalam penerapan Analisis Sentimen Berbasis Aspek, pemilihan aspek yang akan dianalisis menjadi langkah penting, karena aspek-aspek tersebut harus relevan dengan entitas yang dikaji, seperti *platform Online Travel Agencies*. Namun, tidak ada teori khusus yang secara eksplisit menyatakan bahwa pemilihan aspek yang dianalisis dalam analisis sentimen untuk *platform Online Travel Agencies* harus selalu melibatkan aspek harga, layanan, dan keamanan. Pemilihan aspek-aspek ini lebih didorong oleh penelitian terdahulu dalam penelitian analisis sentimen, yang

didasarkan pada faktor-faktor yang mempengaruhi keputusan konsumen dan persepsi terhadap produk atau layanan. Aspek harga sering kali menjadi fokus utama karena harga merupakan faktor krusial dalam keputusan pembelian di *Online Travel Agencies* (Paulina, 2020). Begitu pula dengan aspek layanan, yang mencakup kualitas pelayanan, kenyamanan, dan kemudahan penggunaan platform, yang secara langsung mempengaruhi pengalaman pengguna. Sementara itu, keamanan menjadi perhatian penting karena berkaitan dengan perlindungan data pribadi dan transaksi *online*, yang merupakan aspek utama dalam menjaga kepercayaan pengguna terhadap platform *Online Travel Agencies*. Meskipun tidak ada teori yang mengharuskan pemilihan aspek-aspek ini, terdapat penelitian terdahulu seperti pada table 2.2 yang memilih aspek harga, layanan, dan keamanan karena relevansinya yang tinggi terhadap kepuasan pengguna dan keputusan pembelian di platform *Online Travel Agencies*.

Seiring kemajuan teknologi, analisis sentimen berbasis aspek semakin banyak diterapkan di berbagai industri seperti perbankan, perhotelan, teknologi, dan layanan publik. Tabel 2.2 menjelaskan penelitian sebelumnya terkait analisis sentimen berbasis aspek.

Tabel 2.2 Penelitian Terdahulu Analisis Sentimen Berbasis Aspek

No.	Referensi Penelitian	Input	Metode yang Digunakan	Hasil
1.	(Jessica Widyadhana Iskandar & Yessica Nataliani, 2021)	Ulasan tentang Samsung Galaxy Z Flip 3 (aspek desain, harga, spesifikasi, citra merk)	CRISP-DM, Naïve Bayes (NB), Support Vector Machine, k-Nearest Neighbor (k-NN)	<i>Support Vector Machine</i> menunjukkan performa terbaik dengan rata-rata akurasi 96.43%, khususnya pada aspek desain, harga, spesifikasi, dan citra merk dalam analisis sentimen terhadap Samsung Galaxy Z Flip 3.

No.	Referensi Penelitian	Input	Metode yang Digunakan	Hasil
2.	(Yoga Tika Pratama, Fitra Abdurrachman Bachtiar, dan Nanang Yudi Setiawan, 2021)	Data ulasan pengguna untuk klasifikasi sentimen	Support Vector Machine	Menunjukkan hasil akurasi klasifikasi rata-rata 85% dengan precision, recall, dan F1-score yang mendukung visualisasi hasil dalam dashboard.
3.	(Risca Naquitasia, Dhomas Hatta Fudholi, dan Lizda Iswari, 2022)	Ulasan wisata halal	Deep Learning (CNN)	CNN memberikan akurasi terbaik dalam analisis sentimen wisata halal, dengan akurasi klasifikasi aspek sebesar 98.299% dan sentimen sebesar 93.96%.
4.	(Rifqy Mikoriza Turjaman & Indra Budi, 2022)	Data ulasan pengguna LinkAja	String Matching, Support Vector Machine dengan Undersampling	Sentimen negatif pada aspek produk (98%) dan tempat (100%), netral pada aspek harga (89%), dan positif pada aspek promosi (98%) dalam analisis sentimen LinkAja.
5.	(Gracia Radiena & Adi Nugroho, 2023)	Data ulasan aplikasi KAI Access	CRISP-DM, Support Vector Machine	Hasil klasifikasi sentimen aplikasi KAI Access menunjukkan opini positif pada aspek satisfaction dan opini negatif pada aspek learnability, efficiency, dan errors.

Dari penjelasan tabel 2.2 mengenai penelitian terdahulu mengenai analisis sentimen pada berbagai aspek layanan teknologi, termasuk *Online Travel Agencies*, menunjukkan hasil yang signifikan dengan berbagai metode. Iskandar dan Nataliani (2021) mengaplikasikan CRISP-DM, *Naïve Bayes*, *Support Vector Machine*, dan *k-Nearest Neighbor* untuk menganalisis sentimen terhadap Samsung Galaxy Z Flip 3, dengan *Support Vector Machine* memberikan performa terbaik dan akurasi mencapai 96.43%, terutama pada aspek desain, harga, spesifikasi, dan citra merek. Pratama (2021) juga menggunakan *Support Vector Machine* dalam analisis

sentimen, menghasilkan akurasi rata-rata sebesar 85% dengan metrik *precision*, *recall*, dan F1-score yang terintegrasi ke dalam visualisasi dashboard.

Naquitasia (2022) menggunakan *Deep Learning* (CNN) untuk menganalisis sentimen terhadap wisata halal. Hasil penelitian menunjukkan bahwa CNN memberikan akurasi klasifikasi aspek sebesar 98.299% dan sentimen sebesar 93.96%, menjadikannya metode yang efektif untuk konteks ini. Radiena dan Nugroho (2023) kembali menerapkan metode CRISP-DM dan *Support Vector Machine* dalam analisis aplikasi KAI Access, yang mengidentifikasi sentimen positif pada aspek kepuasan pengguna, namun sentimen negatif pada aspek *learnability*, efisiensi, dan *error*. Penelitian-penelitian ini menunjukkan bahwa *Support Vector Machine* dan CNN menjadi metode yang dominan dalam menganalisis sentimen pengguna terhadap layanan teknologi dan *Online Travel Agencies*, dengan fokus pada berbagai aspek seperti kepuasan, promosi, dan harga.

Dengan berbagai aplikasi dan temuan dari penelitian-penelitian ini, analisis sentimen berbasis aspek terbukti sebagai metode analisis yang efektif dalam mengevaluasi sentimen terkait aspek-aspek spesifik dari produk, layanan, dan fitur. Metode ini lebih relevan dan memberikan hasil yang lebih baik dibandingkan dengan analisis sentimen umum, karena dalam analisis sentimen berbasis aspek, sentimen dianalisis sesuai dengan aspek yang lebih spesifik terlebih dahulu. Hal ini penting karena tidak semua kata memiliki arti yang sama dalam konteks aspek yang berbeda. Dengan demikian, pendekatan berbasis aspek memungkinkan pemahaman yang lebih mendalam dan kontekstual mengenai sentimen pengguna, yang dapat digunakan untuk perbaikan dan pengembangan produk atau layanan. Kebaruan

penelitian ini adalah penggunaan data ulasan pengguna yang diambil dari platform Traveloka melalui *review* yang terdapat di *Google Play Store*. Sementara, pada table 2.2 para peneliti menggunakan data ulasan dari penyebaran formulir, terdapat juga data ulasan pengguna yang diambil dari *review* di *Goole Play Store*.

### **2.3 *Preprocessing Data***

*Preprocessing Data* adalah salah satu langkah dalam *text mining* yang berfungsi untuk mengubah data mentah yang tidak terstruktur menjadi data terstruktur yang siap dianalisis. Data mentah tidak dapat langsung diproses, sehingga perlu dilakukan tahapan *text preprocessing* agar data tersebut siap untuk dievaluasi (Deviyanto, 2018).

### **2.4 *Term Frequency-Inverse Document Frequency (TF-IDF)***

Pembobotan kata atau *weighting word*, adalah teknik yang digunakan untuk memberikan nilai pada kemunculan kata dalam dokumen. Metode ini berguna untuk menilai tingkat kesamaan antar kata serta memastikan data yang dianalisis siap untuk diproses dalam pengembangan model yang lebih lanjut (Yulita, 2021). Salah satu teknik yang paling umum digunakan dalam pembobotan kata adalah *Term Frequency-Inverse Document Frequency (TF-IDF)*, yang menghitung pentingnya sebuah kata dalam dokumen tertentu dibandingkan dengan dokumen lain dalam satu koleksi. TF-IDF adalah kombinasi dari dua metode, yaitu *Term Frequency (TF)* dan *Inverse Document Frequency (IDF)*, yang digabungkan untuk memberikan bobot yang lebih akurat pada kata-kata penting dalam sebuah dokumen.

*Term Frequency* (TF) mengukur frekuensi kemunculan kata dalam suatu dokumen. Semakin sering sebuah kata muncul, semakin besar bobot yang akan diberikan padanya. Namun, TF memiliki kekurangan yaitu jika kata tersebut muncul di banyak dokumen, kata tersebut menjadi common term yang dianggap kurang penting (Rofiqi, 2019). Ini karena kata yang muncul di banyak dokumen biasanya memiliki makna yang lebih umum, sehingga tidak banyak membantu dalam menentukan relevansi dokumen terhadap suatu topik.

Untuk mengatasi kelemahan ini, digunakan *Inverse Document Frequency* (IDF). IDF menghitung seberapa jarang sebuah kata muncul di seluruh koleksi dokumen, dan semakin jarang kemunculannya, semakin tinggi bobot yang diberikan. Dengan demikian, IDF mengurangi dominasi dari common term, memberikan nilai yang lebih besar pada kata-kata yang lebih spesifik dan jarang muncul (Noviyanti, 2019). Proses perhitungan TF-IDF melibatkan perhitungan nilai TF untuk setiap kata dalam dokumen, diikuti dengan perhitungan *Document Frequency* (DF) atau jumlah dokumen yang mengandung kata tersebut. Nilai TF dan IDF dikalikan untuk mendapatkan bobot TF-IDF yang memberikan nilai akurat pada pentingnya suatu kata dalam dokumen.

## **2.5 *Support Vector Machine***

*Support Vector Machine* adalah algoritma pembelajaran mesin yang efektif untuk klasifikasi dan regresi, bertujuan menemukan hyperplane optimal yang memisahkan data ke dalam dua kelas berbeda (Tommy Rustandi, 2023). *Hyperplane* ini membagi ruang fitur menjadi dua bagian, dan *Support Vector Machine* berfokus pada memaksimalkan margin—jarak antara hyperplane dan titik

data terdekat dari masing-masing kelas, yang dikenal sebagai *support vectors* (Hanna, 2020). Titik-titik ini sangat penting karena mereka menentukan posisi hyperplane dan mempengaruhi hasil klasifikasi. Untuk menangani data yang tidak dapat dipisahkan secara linear, *Support Vector Machine* menggunakan teknik kernel trick yang memetakan data ke ruang fitur yang lebih tinggi, memungkinkan pemisahan linear dengan kernel seperti linear, polynomial, atau radial basis function (RBF). Parameter C pada *Support Vector Machine* mengatur trade-off antara margin yang lebih besar dan kesalahan klasifikasi, membantu menyeimbangkan fitting data pelatihan dan generalisasi model.

Meskipun *Support Vector Machine* sangat efektif, pelatihan algoritma ini pada dataset besar dapat menghadapi tantangan komputasi yang signifikan. Untuk mengatasi masalah ini, *Sequential Minimal Optimization* (SMO) dirancang untuk meningkatkan efisiensi pelatihan *Support Vector Machine* dengan mengurangi kompleksitas komputasi (Iwan Nurhidayat & Fatrianto, 2021). *Sequential Minimal Optimization* menyederhanakan proses optimisasi yang kompleks dengan membagi masalah menjadi sub-masalah yang lebih kecil, menyelesaikan masalah optimisasi untuk dua variabel (atau *multiplicator Lagrange*) secara bersamaan, dan memperbarui nilai-nilai tersebut secara iteratif. Dengan pendekatan ini, *Sequential Minimal Optimization* mengubah masalah optimisasi besar menjadi masalah yang lebih kecil dan lebih mudah dipecahkan, memungkinkan pelatihan *Support Vector Machine* dilakukan dengan lebih efisien (Kundu, 2023).

Kombinasi antara *Support Vector Machine* dan *Sequential Minimal Optimization* memberikan solusi yang kuat untuk menangani dataset besar dan

kompleks (Elizabeth, 2022). *Support Vector Machine* dengan kernel *trick* dapat menangani data non-linear dan memberikan kinerja yang baik dalam berbagai aplikasi seperti pengklasifikasian teks, pengenalan pola, dan analisis sentimen. *Sequential Minimal Optimization* memungkinkan proses pelatihan yang lebih cepat dan efisien tanpa memerlukan sumber daya komputasi yang berlebihan, menjadikannya sangat berguna dalam skenario dengan data besar (Irfani, 2020). Dengan metode ini, pelatihan *Support Vector Machine* menjadi lebih praktis dan efektif, mempertahankan kemampuan *Support Vector Machine* untuk memberikan performa tinggi dan akurat dalam analisis data yang kompleks.

Tabel 2.3 Penelitian Terdahulu Tentang Support Vector Machine

No.	Referensi Penelitian	Input	Metode yang Digunakan	Hasil
1.	(Hanna, 2020)	Dataset dengan pola non-linear	<i>Support Vector Machine</i> dengan kernel RBF	<i>Support Vector Machine</i> berhasil mengklasifikasikan data dengan akurasi 92.5% pada dataset dengan pola non-linear. Kernel RBF membantu dalam menangani kompleksitas data yang tidak terpisahkan secara linear.
2.	(Iwan Nurhidayat & Fatrianto, 2021)	Pola gambar	<i>Support Vector Machine</i> dengan Kernel Trick	Penggunaan kernel polynomial pada <i>Support Vector Machine</i> menghasilkan akurasi klasifikasi sebesar 88.7% dalam pengenalan pola gambar. Kernel trick memungkinkan pemetaan ke ruang fitur yang lebih tinggi, meningkatkan kemampuan klasifikasi.
3.	(Elizabeth, 2022)	Data teks	<i>Support Vector Machine</i> dengan <i>Sequential Minimal Optimization</i>	<i>Sequential Minimal Optimization</i> meningkatkan efisiensi pelatihan <i>Support Vector Machine</i> hingga 30% pada dataset besar tanpa mengurangi akurasi model, menjadikan <i>Support Vector Machine</i> lebih praktis untuk diterapkan pada data skala besar.
4.	(Tommy Rustandi, 2023)	Data teks	<i>Support Vector Machine</i> dengan	Kombinasi kernel trick dan <i>Sequential Minimal Optimization</i> menghasilkan

No.	Referensi Penelitian	Input	Metode yang Digunakan	Hasil
			Kernel Trick dan SMO	model <i>Support Vector Machine</i> dengan akurasi 90.3% dalam klasifikasi teks. Teknik ini juga mempercepat proses pelatihan pada data berukuran besar.
5.	(Kundu, 2023)	Data teks dan pola	<i>Support Vector Machine</i> dengan <i>Sequential Minimal Optimization</i>	<i>Sequential Minimal Optimization</i> memungkinkan pelatihan <i>Support Vector Machine</i> pada dataset besar dengan efisiensi tinggi, menghasilkan akurasi 91.8% dalam analisis sentimen. Model ini menunjukkan kinerja yang optimal dalam berbagai aplikasi, termasuk pengklasifikasian teks dan pengenalan pola.

Tabel 2.3 menunjukkan beberapa penelitian terdahulu tentang metode *Support Vector Machine* dengan berbagai pendekatan. Hanna et al. (2020), menggunakan kernel RBF dan mencapai akurasi 92.5% pada dataset non-linear. Selanjutnya, penelitian yang dilakukan oleh Iwan Nurhidayat & Fatrianto (2021) memanfaatkan kernel polynomial, menghasilkan akurasi 88.7% dalam pengenalan pola gambar. Elizabeth (2022), menerapkan *Sequential Minimal Optimization*, yang meningkatkan efisiensi pelatihan hingga 30% pada dataset besar. Kemudian, Tommy Rustandi (2023) menggabungkan kernel trick dan *Sequential Minimal Optimization* untuk mencapai akurasi 90.3% dalam klasifikasi teks. Terakhir, Kundu (2023) menunjukkan bahwa *Sequential Minimal Optimization* dapat menangani dataset besar dengan efisiensi tinggi, mencapai akurasi 91.8% dalam analisis sentimen.

Kebaruan penelitian ini terletak pada tipe klasifikasi yang digunakan. Penelitian ini menerapkan dua tipe klasifikasi, yaitu *single stage classification*

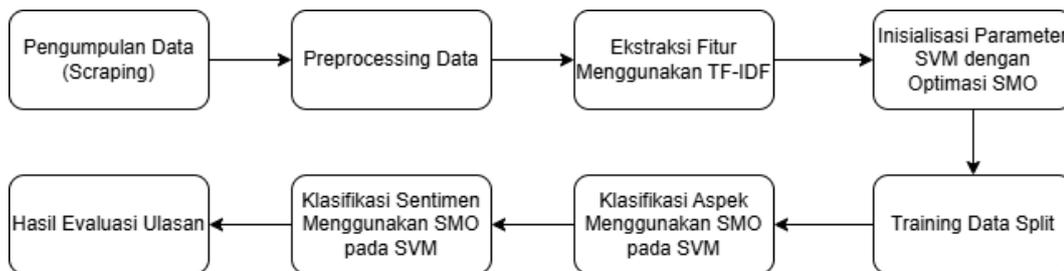
untuk menentukan sentimen secara langsung dan *multistage classification* untuk menganalisis sentimen pada aspek tertentu secara bertahap. Pendekatan ini memberikan kontribusi baru dalam mengoptimalkan analisis sentimen berbasis aspek dengan memanfaatkan kombinasi strategi klasifikasi yang lebih terstruktur.

## BAB III

### DESAIN DAN IMPLEMENTASI

#### 3.1 Desain Sistem

Desain sistem dalam penelitian ini merupakan gambaran alur kerja yang menjelaskan langkah-langkah yang dilakukan oleh sistem, dimulai dari tahap awal pengumpulan dataset hingga tahap akhir evaluasi. Secara umum, alur penelitian dijelaskan pada Gambar 3.1.



Gambar 3.1 Alur Penelitian

Berdasarkan Gambar 3.1, diagram blok ini menggambarkan tahapan-tahapan proses desain sistem yang diterapkan dalam penelitian ini. Proses ini dirancang agar setiap tahap dapat dilaksanakan secara terstruktur, dimulai dari pengumpulan data ulasan dari platform *Google Play Store* melalui proses scraping. Setelah data terkumpul, tahapan selanjutnya adalah *preprocessing data* yang mencakup pembersihan dan persiapan data untuk dianalisis lebih lanjut. Kemudian, data yang telah diproses akan melalui tahap ekstraksi fitur menggunakan teknik TF-IDF untuk mengidentifikasi istilah-istilah penting yang terdapat dalam ulasan pengguna. Selanjutnya, data yang telah diekstrak akan diproses melalui dua tahap klasifikasi, yakni klasifikasi sentimen dan klasifikasi aspek. Kedua tahap klasifikasi ini

menggunakan algoritma *Support Vector Machine* yang dioptimalkan dengan *Sequential Minimal Optimization* untuk memaksimalkan akurasi prediksi. Proses klasifikasi sentimen bertujuan untuk menentukan apakah ulasan bersifat positif atau negatif, sedangkan klasifikasi aspek berfungsi untuk mengidentifikasi aspek-aspek tertentu yang dibahas dalam ulasan, seperti harga atau aksesibilitas. Setelah seluruh proses klasifikasi selesai, sistem akan mengevaluasi hasil klasifikasi tersebut untuk mengukur kinerja model yang dibangun, dengan menggunakan metrik evaluasi yang sesuai.

### 3.2 Pengumpulan Data

Penelitian ini menggunakan ulasan pengguna aplikasi Traveloka yang diambil dari platform *Google Play Store*. Objek yang akan digunakan berupa ulasan pengguna berbahasa Indonesia. Pengambilan data dilakukan dengan teknik scraping menggunakan kata kunci pada tiap kategori. Pelabelan pada tiap aspek akan dilakukan oleh sistem berdasarkan kata kunci yang digunakan. Data yang digunakan berjumlah 3000 ulasan pengguna, dengan pembagian kategori harga, layanan, dan keamanan sebanyak 1000 data per-aspek. Atribut pada dataset meliputi:

Tabel 3.1 Atribut Data

<b>Kolom</b>	<b>Deskripsi</b>	<b>Tipe Data</b>
Username	Nama atau pengenal unik pengguna yang memberikan ulasan.	<i>String</i>
Tanggal dan Waktu	Tanggal dan waktu saat ulasan diberikan oleh pengguna.	<i>Date</i>
<i>Content</i>	Teks ulasan pengguna yang berisi opini terkait aspek harga, layanan, atau keamanan.	<i>String</i>

Pada aspek harga digunakan kata kunci 'harga', 'mahal', 'murah', dan 'biaya'. Pada aspek layanan digunakan kata kunci 'layanan', 'servis', 'pelayanan', 'support', dan 'customer service'. Pada aspek keamanan digunakan kata kunci 'keamanan', 'aman', 'penipuan', 'transaksi', dan 'verifikasi'.

Kemudian, data ulasan akan diklasifikasikan ke dalam kategori layanan, keamanan, atau harga. Tabel 3.1 menjelaskan label kategori yang digunakan dalam penelitian ini, yang merujuk pada penelitian yang dilakukan oleh Paulina (2020) mengenai aspek harga, layanan, dan keamanan.

Tabel 3.2 Deskripsi Tabel Kategori

Harga	Ulasan pengguna mengenai harga mencakup kepuasan terhadap nilai yang diberikan, daya saing harga, dan biaya tambahan yang mungkin dikenakan selama transaksi.
Layanan	Ulasan pengguna mengenai layanan mencakup kualitas pelayanan pelanggan, responsivitas terhadap masalah, dan kehandalan aplikasi dalam menangani keluhan dan umpan balik.
Keamanan	Ulasan pengguna mengenai keamanan mencakup tingkat perlindungan data pribadi, keandalan sistem dalam mencegah pelanggaran, dan efektivitas fitur keamanan aplikasi.

Setelah klasifikasi aspek selesai, langkah berikutnya adalah melakukan klasifikasi sentimen untuk setiap aspek yang diidentifikasi.

Tabel 3.3 Contoh Data

Ulasan	Aspek
Aplikasinya mudah banget dipakai, sangat membantu untuk mencari destinasi liburan yang sudah direncanakan gak perlu ribet harus telfon2 ke tempat/hotel yang kita mau.	Layanan
Harga aplikasi ini sangat kompetitif dibandingkan dengan pesaing. Harganya terjangkau dan memberikan banyak fitur	Harga
Kecewa berat Tiket udah dibeli mana mahal kan lagi ramadhan, eh ada kendala sistem keamanan ga bisa login ke akun yg udah ada tiketnya. Gimana saya mau mudik ini aduhhhh!!!!	Keamanan

### 3.3 *Preprocessing Data*

*Preprocessing Data* adalah salah satu langkah dalam *text mining* yang berfungsi untuk mengubah data mentah yang tidak terstruktur menjadi data terstruktur yang siap dianalisis. Data mentah tidak dapat langsung diproses, sehingga perlu dilakukan tahapan *text preprocessing* agar data tersebut siap untuk dievaluasi (Deviyanto, 2018). Langkah-langkah *text preprocessing* yang digunakan dalam penelitian ini adalah sebagai berikut:

#### a. *Cleansing*

*Cleansing* adalah tahap dimana elemen-elemen yang tidak relevan, karakter yang tidak standar, atau komponen yang tidak berhubungan dengan konten dokumen dihapus. Ini termasuk emotikon, angka, tanda baca, simbol, spasi berlebih, enter, tautan, dan tagar.

Tabel 3.4 Proses Cleansing

Sebelum <i>Cleansing</i>	Traveloka bagus Cuma..harganya kurang byk diskonnya 😊😄 klo diskon dikit banget..boleh dong promo diskonnya digedein dikit ya min.. Sukses terus traveloka.. Trimakasih sdh bantu byk buat saya..
Sesudah <i>Cleansing</i>	Traveloka bagus, cuma harganya kurang banyak diskonnya. Jika diskon sedikit sekali, boleh dong promo diskonnya diperbesar sedikit ya admin. Sukses terus Traveloka. Terima kasih sudah membantu banyak buat saya.

#### b. *Case Folding*

*Case folding* berfungsi sebagai teknik untuk mengubah huruf kapital menjadi huruf kecil, sehingga memastikan bahwa semua kata yang digunakan berada dalam bentuk huruf kecil. Konversi ini diterapkan secara konsisten, dimulai dari awal kalimat, nama pribadi, nama kota, dan contoh lainnya. Contoh implementasi *case folding* diberikan di bawah ini.

Tabel 3.5 Proses Case Folding

Sebelum <i>Case Folding</i>	Traveloka bagus, cuma harganya kurang banyak diskonnya. Jika diskon sedikit sekali, boleh dong promo diskonnya diperbesar sedikit ya admin. Sukses terus Traveloka. Terima kasih sudah membantu banyak buat saya.
Sesudah <i>Case Folding</i>	traveloka bagus, cuma harganya kurang banyak diskonnya. jika diskon sedikit sekali, boleh dong promo diskonnya diperbesar sedikit ya admin. sukses terus traveloka. terima kasih sudah membantu banyak buat saya.

c. *Tokenizing*

Merupakan proses pemecahan kalimat menjadi kata-kata individual.

Tahap ini memudahkan dalam menghitung frekuensi kemunculan kata dalam dokumen.

Tabel 3.6 Proses Tokenizing

Sebelum <i>Tokenizing</i>	traveloka bagus, cuma harganya kurang banyak diskonnya. jika diskon sedikit sekali, boleh dong promo diskonnya diperbesar sedikit ya admin. sukses terus traveloka. terima kasih sudah membantu banyak buat saya.
Sesudah <i>Tokenizing</i>	[traveloka, bagus, cuma, harganya, kurang, banyak, diskonnya, jika, diskon, sedikit, sekali, boleh, dong, promo, diskonnya, diperbesar, sedikit, ya, admin, sukses, terus, traveloka, terima, kasih, sudah, membantu, banyak, buat, saya]

d. *Stopword Removal*

*Stopword Removal* adalah tahap untuk menghilangkan kata-kata umum yang sering muncul dan tidak memiliki makna signifikan. *Stopword removal* bertujuan untuk menghapus kata-kata tersebut dari daftar token.

Tabel 3.7 Proses Stopword Removal

Sebelum <i>Stopword Removal</i>	[traveloka, bagus, cuma, harganya, kurang, banyak, diskonnya, jika, diskon, sedikit, sekali, boleh, dong, promo, diskonnya, diperbesar, sedikit, ya, admin, sukses, terus, traveloka, terima, kasih, sudah, membantu, banyak, buat, saya]
Sesudah <i>Stopword Removal</i>	[traveloka, bagus, harganya, kurang, diskonnya, diskon, sedikit, promo, diskonnya, diperbesar, admin, sukses, traveloka, terima, kasih, membantu, buat]

e. *Stemming*

*Stemming* merupakan langkah terakhir dalam *text preprocessing* dimana setiap kata dalam dokumen diubah menjadi bentuk dasarnya. Kata-kata dengan imbuhan seperti awalan (*prefix*), akhiran (*suffix*), sisipan (*infix*), dan kombinasi antara awalan dan akhiran (*confix*) akan dihilangkan.

Tabel 3.8 Proses Stemming

Sebelum <i>Stemming</i>	[traveloka, bagus, harganya, kurang, diskonnya, diskon, sedikit, promo, diskonnya, diperbesar, admin, sukses, traveloka, terima, kasih, membantu, buat]
Sesudah <i>Stemming</i>	[traveloka, bagus, harga, kurang, diskon, diskon, sedikit, promo, diskon, besar, admin, sukses, traveloka, terima, kasih, bantu, buat]

```

all_data['content'] = all_data['content'].str.replace('\d+', '', regex=True)
all_data['content'] = all_data['content'].str.replace('[^\w\s]', '', regex=True)
all_data['content'] = all_data['content'].str.lower()
all_data['content'] = all_data['content'].apply(word_tokenize)
all_data['content'] = all_data['content'].apply(lambda x: [word for word in x if word not in stop_words])
all_data['content'] = all_data['content'].apply(lambda x: [stemmer.stem(word) for word in x])

```

Gambar 3.2 Kode *Preprocessing Data*

Pada gambar 3.2 merupakan implementasi dari kode *preprocessing data* yang digunakan untuk membersihkan teks pada kolom *content*. Proses ini mencakup penghapusan angka dan karakter non-alfabet, mengubah teks menjadi huruf kecil, tokenisasi, penghapusan *stop words*, dan stemming. Tahapan ini bertujuan untuk mengubah data mentah menjadi lebih terstruktur dan siap digunakan dalam analisis sentimen.

### 3.4 *Term Frequency-Inverse Document Frequency (TF-IDF)*

Pembobotan kata atau *weighting word* adalah teknik yang memberikan nilai pada kemunculan kata dalam dokumen untuk menilai tingkat kesamaan antar kata dan mempersiapkan data untuk analisis lebih lanjut. Salah satu metode yang umum digunakan adalah TF-IDF (*Term Frequency-Inverse Document Frequency*), yang

menggabungkan frekuensi kemunculan kata dalam dokumen tertentu (TF) dengan keunikannya di seluruh koleksi dokumen (IDF) untuk menghasilkan bobot yang akurat pada kata-kata penting.

TF mengukur seberapa sering sebuah kata muncul dalam dokumen, sedangkan IDF memberikan bobot lebih tinggi pada kata-kata yang jarang muncul di banyak dokumen, mengurangi dominasi kata-kata umum. Hasil dari TF-IDF menunjukkan pentingnya suatu kata dalam konteks spesifik, dengan bobot yang lebih besar untuk kata-kata yang signifikan dan jarang digunakan.

Setelah data melalui tahap *preprocessing*, maka selanjutnya data akan di ekstraksi menggunakan TF terlebih dahulu seperti persamaan 3.1:

$$TF(t, d) = \frac{f_{t,d}}{N_d} \quad (3.1)$$

Keterangan:

- $f_{t,d}$  = Jumlah kemunculan term/kata dalam dokumen  $d$ .
- $N_d$  = Jumlah total term/kata dalam dokumen  $d$ .

Kemudian setelah dilakukan perhitungan TF, selanjutnya akan dilakukan perhitungan IDF seperti persamaan 3.2:

$$IDF(t, D) = \log \frac{N}{n_t} \quad (3.2)$$

Keterangan:

- $N$  = Jumlah total dokumen dalam koleksi  $D$ .
- $n_t$  = Jumlah dokumen yang mengandung term/kata.

Setelah menghitung nilai TF dan IDF, langkah terakhir adalah menghitung nilai bobot TF-IDF untuk menentukan sejauh mana pentingnya suatu kata dalam dokumen, seperti persamaan 3.3:

$$ITF - IDF(t, d, D) = TF(t, d) * IDF(t, D) \quad (3.3)$$

Keterangan:

- $t$  = Term/kata tertentu yang dianalisis.
- $d$  = Dokumen tertentu dalam koleksi dokumen  $D$ .
- $D$  = Koleksi atau kumpulan seluruh dokumen yang dianalisis.
- $TF(t, d)$  = Frekuensi kemunculan term  $t$  dalam dokumen  $d$ .
- $IDF(t, d)$  = Ukuran seberapa umum atau jarang term  $t$  muncul di seluruh dokumen dalam koleksi  $D$ .

Setelah membobotkan setiap kata dalam dokumen, setiap kata di seluruh dokumen kini memiliki bobot yang sesuai. Contoh perhitungan TF-IDF dapat dilihat pada Tabel 3.9 untuk penjelasan lebih lanjut:

Tabel 3.9 Contoh Perhitungan TF-IDF

Kata	Term TF			DF	IDF	TF-IDF		
	D1	D2	D3			D1	D2	D3
Cocok	1	1	0	2	0.176	0.176	0.176	0
Langganan	0	1	0	1	0.477	0	0.477	0
Tiket	1	1	1	3	0	0	0	0
Mudah	1	0	1	2	0.176	0.176	0	0.176
Murah	1	0	1	2	0.176	0.176	0	0.176
Pemesanan	0	0	1	1	0.477	0	0	0.477

Pada Tabel 3.9, ditampilkan hasil perhitungan manual untuk ekstraksi fitur TF-IDF dari enam kata kunci dalam tiga dokumen (D1, D2, dan D3). Proses ini dimulai dengan menghitung *Term Frequency* (TF) dari setiap kata dalam masing-masing dokumen, yang menunjukkan frekuensi kemunculan kata. Misalnya, kata "cocok" muncul sekali di D1 dan D2, sementara tidak ada di D3, menghasilkan TF 1 untuk D1 dan D2. Selanjutnya, dihitung *Document Frequency* (DF) untuk setiap kata, yang menunjukkan jumlah dokumen yang mengandung kata tersebut. Kata "cocok" muncul di 2 dari 3 dokumen, sehingga DF-nya adalah 2.

*Inverse Document Frequency* (IDF) kemudian dihitung menggunakan persamaan  $IDF = \log \frac{N}{n_t}$ , nilai  $N$  adalah total jumlah dokumen. Untuk kata "cocok",

IDF yang dihitung adalah 0.176, menunjukkan kata ini tidak muncul di seluruh dokumen secara merata. Terakhir, nilai TF-IDF dihitung dengan mengalikan TF dan IDF. Misalnya, nilai TF-IDF untuk kata "cocok" di D1 adalah 0.176, yang menunjukkan pentingnya kata tersebut dalam konteks dokumen. Tabel ini menggambarkan seberapa relevan setiap kata dalam dokumen berdasarkan frekuensi kemunculan dan kepentingannya di seluruh koleksi dokumen.

```
def compute_tf_idf(corpus):
    tf = []
    idf = Counter()
    num_docs = len(corpus)

    print("Menghitung TF-IDF...")
    for document in corpus:
        word_count = Counter(document.split())
        total_words = len(document.split())
        tf.append({word: count / total_words for word, count in word_count.items()})
        idf.update(set(word_count.keys()))

    idf = {word: np.log(num_docs / freq) for word, freq in idf.items()}

    tf_idf = []
    for document_tf in tf:
        tf_idf.append({word: tf_val * idf[word] for word, tf_val in document_tf.items()})

    return tf_idf

# Menghitung TF-IDF untuk data
tf_idf = compute_tf_idf(all_data['content'])
```

Gambar 3.3 Kode *Term Frequency-Inverse Document Frequency*

Pada gambar 3.3 menunjukkan implementasi perhitungan *Term Frequency-Inverse Document Frequency* (TF-IDF). Kode ini menghitung frekuensi kata dalam setiap dokumen (TF), lalu menghitung IDF berdasarkan jumlah dokumen yang mengandung kata tersebut. Hasil perkalian antara TF dan IDF menghasilkan bobot TF-IDF yang digunakan sebagai fitur dalam proses klasifikasi sentimen.

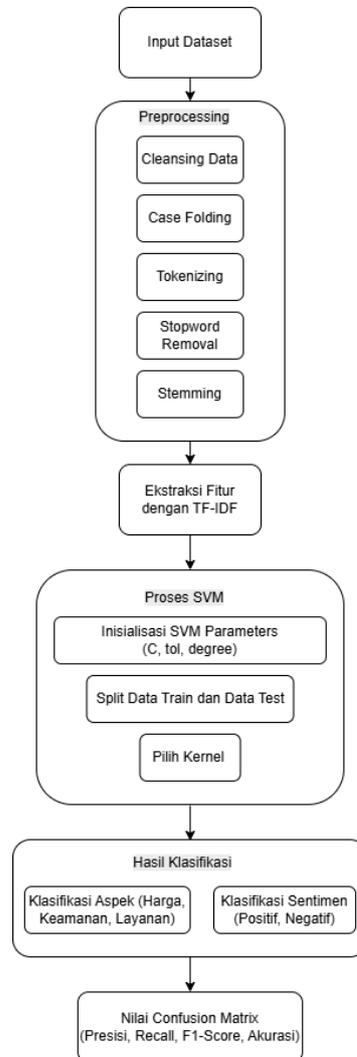
### 3.5 *Support Vector Machine*

*Support Vector Machine* adalah algoritma pembelajaran mesin yang digunakan terutama untuk klasifikasi biner. *Support Vector Machine* mencari *hyperplane* yang memisahkan dua kelas dengan margin maksimum. Titik-titik data

yang paling dekat dengan hyperplane, disebut *support vectors*, menentukan posisi *hyperplane* (Ma Y., Guo G., 2014). *Support Vector Machine* juga dapat menangani klasifikasi non-linear menggunakan kernel *trick*, yang memetakan data ke ruang dimensi lebih tinggi untuk pemisahan linear yang lebih mudah (Mahendro I., & Abimanto D., 2023). Untuk mengoptimalkan fungsi biaya dalam *Support Vector Machine* yang kompleks, terutama pada dataset besar, optimisasi *Sequential Minimal Optimization* digunakan sebagai metode yang efisien.

*Sequential Minimal Optimization* adalah algoritma optimasi yang dirancang khusus untuk menyelesaikan masalah optimasi kuadratik dalam *Support Vector Machine*. Pada dasarnya, pelatihan *Support Vector Machine* melibatkan optimasi fungsi biaya yang kompleks dengan batasan-batasan tertentu, yang dapat menjadi sangat sulit dihitung secara efisien, terutama pada dataset yang besar. *Sequential Minimal Optimization* memecahkan masalah ini dengan memecah optimasi besar menjadi serangkaian submasalah kecil yang lebih mudah diselesaikan. Setiap submasalah hanya melibatkan dua parameter *Lagrange* ( $\alpha$ ), yang dapat dioptimalkan secara analitis dan iteratif, sehingga mengurangi kebutuhan akan perhitungan komputasi yang mahal. Dengan cara ini, *Sequential Minimal Optimization* membuat pelatihan *Support Vector Machine* menjadi lebih cepat dan lebih efisien, terutama untuk dataset yang besar dan kompleks. Hubungan antara *Support Vector Machine* dan *Sequential Minimal Optimization* adalah bahwa *Sequential Minimal Optimization* merupakan metode yang memungkinkan *Support Vector Machine* untuk dioptimalkan dengan cara yang lebih praktis dan terukur.

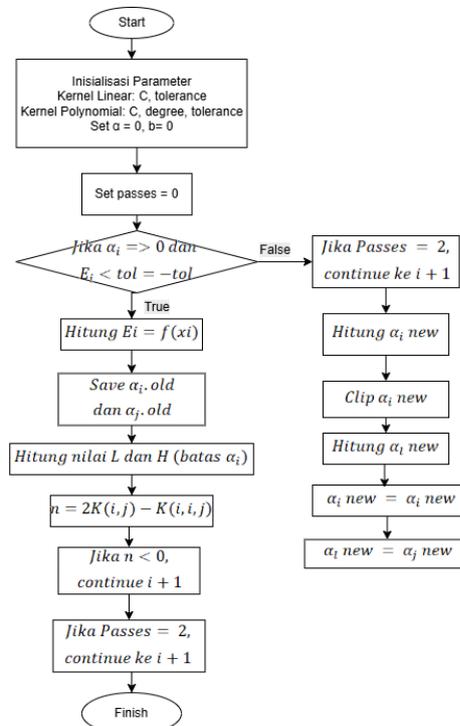
Berikut gambar 3.4 merupakan diagram alur pengujian pada algoritma *Support Vector Machine* yang di optimasi menggunakan *Sequential Minimal Optimization*.



Gambar 3.4 Desain Sistem Penelitian

Proses pelatihan dan pengujian model *Support Vector Machine* terdiri dari beberapa langkah utama yang saling terhubung. Pertama, dimulai dengan data latihan Hasil TF-IDF, di mana data teks diubah menjadi vektor fitur menggunakan metode TF-IDF untuk mempersiapkan data latihan. Selanjutnya adalah Inisialisasi Parameter

*Support Vector Machine*, di mana parameter seperti nilai *degree*, toleransi, pemilihan kernel dan parameter  $C$  diatur.



Gambar 3.5 Flowchart Sequential Minimal Optimization

Pada gambar 3.5 merupakan flowchart algoritma *Support Vector Machine* yang dioptimasi menggunakan *Sequential Minimal Optimization*. Dimulai dengan proses inisialisasi beberapa parameter dasar sesuai dengan jenis kernel yang digunakan, pada penelitian ini menggunakan kernel linear atau polynomial. Setelah itu, semua nilai penting seperti variabel  $\alpha$  dan bias di-set ke nol. Dilanjutkan proses pelatihan mulai berjalan secara berulang-ulang dengan memeriksa setiap data. Sistem mengecek apakah nilai  $\alpha$  untuk data sudah sesuai atau belum, dan juga menghitung kesalahan prediksi. Jika nilai ini tidak terpenuhi, maka proses dilanjutkan ke tahap berikutnya. Jika nilai terpenuhi, maka algoritma akan menghitung batas untuk nilai  $\alpha$  yang akan di-update.

Jika syaratnya terpenuhi, algoritma akan memperbarui nilai  $\alpha$  dan memastikan tetap dalam batas yang ditentukan. Jika perubahan cukup signifikan, nilai bias juga diperbarui agar model lebih akurat. Jika tidak, proses lanjut ke data berikutnya. Proses ini diulang terus tidak ada perubahan signifikan atau sudah mencapai batas iterasi. Dengan metode ini, *Sequential Minimal Optimization* secara efektif menyempurnakan nilai  $\alpha$  sehingga *Support Vector Machine* dapat menentukan batas pemisah optimal dengan lebih cepat dan akurat. Persamaan dari hyperplane dalam bentuk linear untuk model *Support Vector Machine* seperti pada persamaan 3.4:

$$w = x + b = 0 \quad (3.4)$$

Keterangan:

- $w$  adalah vektor bobot,
- $x$  adalah vektor fitur (dalam kasus ini, nilai TF-IDF dari dokumen),
- $b$  adalah bias.

Setelah itu, iterasi pembaruan *Support Vector Machine* dengan Inisialisasi

Awal  $w$  dan  $b$  menggunakan persamaan 3.5:

$$y_1 (w * x_1 + b) \quad (3.5)$$

Keterangan:

- $y_1$  : Label kelas data (misalnya +1 atau -1)
- $w$  : Vektor bobot
- $x$  : Vektor fitur (Nilai TF-IDF dari dokumen)
- $b$  : Bias.

Kemudian dilakukan perhitungan pembaruan dengan persamaan 3.6:

$$w = w + \eta \times y_1 \times x_1 \quad (3.6)$$

$$b = b + \eta * y_1 \quad (3.7)$$

Setelah iterasi, hasil pembaruan bobot dan bias adalah dengan persamaan

3.8:

$$w = \sum \alpha_i y_i x_i \quad (3.8)$$

Keterangan:

- $w$  : Vektor bobot akhir setelah semua alpha diperbarui.
- $\alpha_i$  : Pengali Lagrange yang dioptimalkan untuk data latih ke-i.
- $y_i$  : Label kelas untuk data latih ke-i.
- $x_i$  : Vektor fitur untuk data latih ke-i.

Untuk mengoptimalkan model *Support Vector Machine* menggunakan *Sequential Minimal Optimization*, Langkah pertama yaitu menghitung *error* (E) untuk setiap data dengan persamaan 3.9:

$$E_i = f(x_i) - y_i \quad (3.9)$$

$$f(x_i) = w * x_i + b \quad (3.10)$$

Keterangan:

- $w$  : Vektor bobot akhir setelah semua alpha diperbarui.
- $y_i$  : Label kelas untuk data latih ke-i.
- $x_i$  : Vektor fitur untuk data latih ke-i.
- $E_i$  : Nilai error untuk data  $x_i$

Lalu menghitung nilai eta ( $\eta$ ) dengan persamaan 3.11:

$$\eta = 2x_i * x_j - x_i * x_i - x_j * x_j \quad (3.11)$$

Keterangan:

- $\eta$  : Sebuah nilai yang digunakan untuk menentukan seberapa besar perubahan yang harus diterapkan pada  $\alpha_j$ .
- $x_j$  dan  $x_i$  : Vektor fitur untuk data latih ke-i dan j.

Setelah menghitung nilai  $\eta$ , langkah selanjutnya adalah menghitung nilai  $\alpha_j$  dengan persamaan 3.12:

$$\alpha_j = \alpha_j - \frac{y_j(E_i - E_j)}{\eta} \quad (3.12)$$

Keterangan:

- $\alpha_j$  : Nilai alpha j yang baru setelah pembaruan.
- $E_i$  dan  $E_j$  : Kesalahan prediksi untuk data latih ke-i dan j.

Selanjutnya adalah menghitung nilai  $\alpha_i$  dengan persamaan 3.13:

$$\alpha_i = \alpha_i + y_i y_j (\alpha_j - \alpha_j^{old}) \quad (3.13)$$

Langkah selanjutnya adalah menghitung bias dengan persamaan 3.14:

$$b_1 = b - E_i - y_i (\alpha_i - \alpha_i^{old}) x_i * x_i - y_j (\alpha_j - \alpha_j^{old}) x_i * x_j \quad (3.14)$$

$$b_2 = b - E_j - y_i (\alpha_i - \alpha_i^{old}) x_i * x_i - y_j (\alpha_j - \alpha_j^{old}) x_i * x_j \quad (3.15)$$

Keterangan:

- $b_1$  : Bias yang dihitung dari data latih ke-i.
- $b_2$  : Bias yang dihitung dari data latih ke-j.
- $E_j$  dan  $E_i$  : Kesalahan prediksi untuk data latih ke-I dan ke-j.
- $\alpha_i$  dan  $\alpha_j$  : Nilai alpha setelah pembaruan.

Langkah selanjutnya adalah menghitung rata-rata nilai  $b$  dengan persamaan

3.16:

$$b = \frac{b_1 + b_2}{2} \quad (3.16)$$

Setelah menghitung rata-rata bias, dilanjutkan menghitung nilai  $w$  dengan persamaan 3.17:

$$w = \sum \alpha_i y_i x_i \quad (3.17)$$

Keterangan:

- $w$  : Vektor bobot akhir setelah semua alpha diperbarui.
- $\alpha_i$  : Pengali Lagrange yang dioptimalkan untuk data latih ke-i.
- $y_i$  : Label kelas untuk data latih ke-i.
- $x_i$  : Vektor fitur untuk data latih ke-i.

Dengan  $w$  dan  $b$  yang baru, dapat menghitung prediksi. Rumus persamaan untuk prediksi adalah 3.18:

$$\text{prediksi} = \text{sign}(w * x + b) \quad (3.18)$$

Keterangan:

- $w * x$  : Dot product antara vektor bobot yang telah dioptimalkan dan vektor fitur dari data yang akan diprediksi
- $b$  : Bias yang telah diperbarui.
- $\text{sign}()$  : Fungsi sign yang menghasilkan +1 jika nilainya positif (kelas 1), dan negative (-1).

```
# Iterasi untuk kombinasi parameter C dan max_iter
for C in C_values:
    for max_iter in max_iter_values:
        # Inisialisasi dan pelatihan model
        svm_smo = LinearSVC(C=C, max_iter=max_iter)
        svm_smo.fit(X_train, y_train)

        # Prediksi
        y_pred = svm_smo.predict(X_test)

        # Evaluasi
        precision = precision_score(y_test, y_pred, average='weighted')
        recall = recall_score(y_test, y_pred, average='weighted')
        f1 = f1_score(y_test, y_pred, average='weighted')
        accuracy = accuracy_score(y_test, y_pred)

        # Simpan hasil
        results.append({
            'C': C,
            'max_iter': max_iter,
            'precision': precision,
            'recall': recall,
            'f1_score': f1,
            'accuracy': accuracy
        })
```

Gambar 3.6 Kode *Support Vector Machine* Optimasi *Sequential Minimal Optimization*

Pada gambar 3.6 menunjukkan proses pelatihan dan evaluasi model *Support Vector Machine* dengan optimasi *Sequential Minimal Optimization* menggunakan kombinasi parameter  $C$  dan  $\text{max\_iter}$ , kemudian hasil evaluasinya disimpan berdasarkan matrix akurasi, presisi, recall, dan F1-score.

### 3.6 Evaluasi Pengujian

Langkah berikutnya adalah melakukan evaluasi pengujian untuk mengukur kinerja sistem. Salah satu cara untuk mengukur kinerja ini adalah dengan menggunakan *confusion matrix*. Melalui *confusion matrix*, performa sistem dievaluasi berdasarkan empat parameter: akurasi, presisi, *recall*, dan *f1-score*.

Akurasi menggambarkan sejauh mana sistem mampu mengklasifikasikan ulasan dengan benar (Pratiwi, 2020).

*Confusion matrix* adalah sebuah matriks dua dimensi yang menunjukkan perbandingan antara kelas sebenarnya dan kelas yang diprediksi. Dalam konteks penelitian ini, terdapat dua jenis evaluasi pengujian, yaitu evaluasi berdasarkan aspek dan evaluasi berdasarkan sentimen. Tabel 3.9 di bawah ini menunjukkan *confusion matrix* untuk evaluasi berdasarkan aspek yang terdiri dari 3 aspek atau kelas.

Tabel 3.10 Contoh Confusion Matrix 3 kelas

<i>Actual Aspect</i>	<i>Predicted Aspect</i>		
	<b>Layanan</b>	<b>Harga</b>	<b>Keamanan</b>
Layanan	3	1	7
Harga	4	2	5
Keamanan	2	5	3

Akurasi adalah metrik yang digunakan untuk mengukur seberapa sering model membuat prediksi yang benar secara keseluruhan. Rumus persamaan adalah:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.19)$$

Presisi adalah metrik yang mengukur seberapa akurat model dalam memprediksi kelas positif. Persamaannya adalah:

$$Presisi = \frac{TP}{TP + FP} \quad (3.20)$$

*Recall* digunakan untuk mengukur seberapa baik model dalam menemukan semua contoh positif yang sebenarnya. Rumus persamaannya adalah:

$$Recall = \frac{TP}{TP + FN} \quad (3.21)$$

*F1-score* adalah metrik yang menggabungkan presisi dan recall menjadi satu nilai, memberikan gambaran seimbang antara keduanya. Persamaan *F1-score* adalah:

$$F1 - score = \frac{2x \text{ Presisi } x \text{ Recall}}{\text{Presisi} + \text{Recall}} \quad (3.22)$$

*F1-score* sangat berguna dalam situasi di mana ada ketidakseimbangan antara kelas positif dan negatif, atau ketika perlu memperhitungkan keseimbangan antara presisi dan recall.

Untuk mengevaluasi confusion matrix dengan 3 kelas berdasarkan aspek, nilai TP, TN, FP, dan FN dihitung untuk setiap kelas secara terpisah. Setelah melakukan evaluasi berdasarkan aspek, evaluasi selanjutnya dilakukan berdasarkan sentimen. Tabel 3.11 memperlihatkan kombinasi nilai parameter confusion matrix untuk kelas sentimen.

Tabel 3.11 Contoh Confusion Matrix 2 kelas

Actual Class	Predicted Class	
	Positif	Negatif
Positif	TP	FP
Negatif	FN	TN

Tabel 3.11 menampilkan parameter yang digunakan dalam pengujian menggunakan *confusion matrix*. Nilai *True Positive* (TP) adalah jumlah ulasan dengan nilai positif yang diprediksi secara benar sebagai positif. Nilai *True Negative* (TN) adalah jumlah ulasan dengan nilai negatif yang diprediksi secara benar sebagai negatif. Nilai *False Positive* (FP) adalah jumlah ulasan dengan nilai

negatif yang diprediksi sebagai positif. Nilai *False Negative* (FN) adalah jumlah ulasan dengan nilai positif yang diprediksi sebagai negatif. Untuk memahami penggunaannya dalam evaluasi model, dapat dilihat pada metrik kinerja seperti akurasi, presisi, *recall*, dan *F1-score* yang menjelaskan kinerja model.

```
def plot_confusion_matrix(y_true, y_pred, title, labels):
    # Menghitung confusion matrix
    conf_matrix = confusion_matrix(y_true, y_pred)

    # Membuat plot confusion matrix
    plt.figure(figsize=(10, 8))
    sns.heatmap(conf_matrix, annot=True, fmt="d", cmap="Blues", xticklabels=labels, yticklabels=labels)
    plt.title(title)
    plt.xlabel("Predicted Label")
    plt.ylabel("True Label")
    plt.show()
```

Gambar 3.7 Kode *Confusion Matrix*

Pada gambar 3.7 menunjukkan implementasi fungsi untuk menghitung dan memvisualisasikan *confusion matrix* dari hasil prediksi model, dengan tampilan visual menggunakan heatmap untuk memudahkan interpretasi label aktual dan label prediksi.

### 3.7 Skenario Pengujian

Tahapan selanjutnya adalah skenario pengujian yang bertujuan untuk mengevaluasi kinerja sistem dalam melakukan klasifikasi. Dalam penelitian ini, pengujian dilakukan untuk menemukan parameter terbaik pada algoritma *Support Vector Machine*, termasuk pemilihan kernel, pengaturan parameter C, nilai toleransi, dan nilai *degree* pada kernel polinomial.

Pada penelitian ini, pembagian data melibatkan rasio seperti 80:20. Pengujian dengan rasio *split data* 80:20 dipilih karena memberikan keseimbangan optimal antara pelatihan dan evaluasi model. Dengan 80% data untuk pelatihan, model dapat mempelajari pola yang cukup tanpa kehilangan informasi penting, sementara

20% data pengujian menyediakan sampel yang memadai untuk mengevaluasi generalisasi model. Rasio ini juga mengurangi resiko *overfitting* dan telah menjadi standar industri karena terbukti efektif dalam berbagai jenis dataset, mempermudah perbandingan hasil penelitian (Muraina, 2022). Selain itu, dilakukan perbandingan antara dua tipe klasifikasi, yaitu *Single Stage Classification* dan *Multistage Classification*.

Untuk memastikan bahwa model yang dikembangkan telah bekerja sesuai kebutuhan, pengujian fungsional dilakukan dengan metrik evaluasi seperti akurasi, *presisi*, *recall*, dan *F1-score*. Jika hasil evaluasi menunjukkan nilai metrik yang tinggi, maka sistem dapat dianggap memiliki performa yang baik (Jordan Sihombing & Irvandi Sitorus, 2022). Sebaliknya, nilai metrik yang rendah menunjukkan bahwa diperlukan perbaikan pada parameter atau algoritma yang diterapkan. Pengujian ini dirancang untuk memastikan bahwa algoritma *Support Vector Machine* mampu mengklasifikasikan sentimen berbasis aspek secara efektif dan akurat.

Tabel 3.12 Skenario Pengujian

Uji Coba	Type of Stage Classification	Data Train	Data Test	Kernel
1	Single Stage Classification	80%	20%	Linear
2				Polinomial
3	Multistage Classification			Linear
4				Polinomial

Tabel 3.12 menunjukkan 4 skenario pengujian yang dilakukan untuk mengevaluasi kinerja algoritma *Support Vector Machine* dalam klasifikasi sentimen berbasis aspek. Pengujian ini mencakup dua tipe klasifikasi, yaitu *Single Stage Classification* dan *Multistage Classification*. Pada tipe *Single Stage*

*Classification*, klasifikasi dilakukan dalam satu tahap, sementara pada *Multistage Classification*, klasifikasi dilakukan dalam beberapa tahap yaitu pembagian berdasarkan aspek terlebih dahulu, kemudian dilanjutkan tahap klasifikasi sentimennya.

## BAB IV

### HASIL DAN PEMBAHASAN

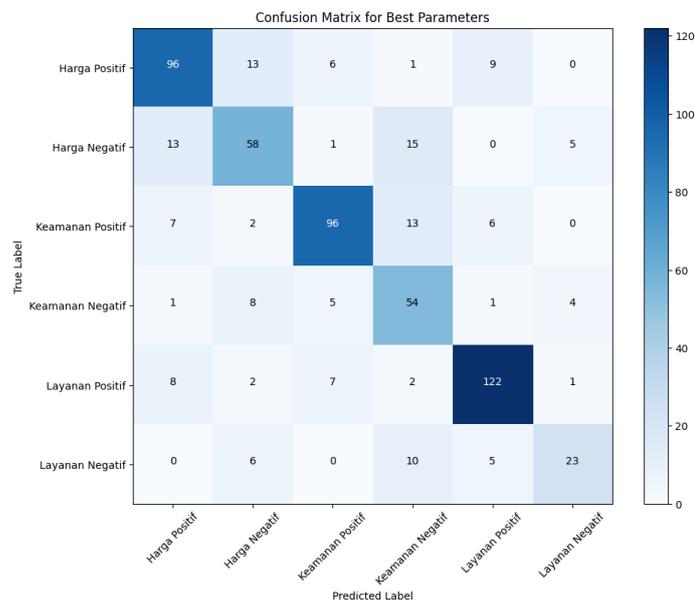
#### 4.1 Hasil Hasil Uji Coba

Pada tabel 3.17 menunjukkan berbagai uji coba klasifikasi menggunakan dua jenis tahap, yaitu *Single Stage Classification* dan *Multistage Classification*, dengan dua jenis kernel, yaitu Linear dan Polinomial.

Pada *Single Stage Classification*, hanya melalui satu tahap klasifikasi, yaitu klasifikasi aspek dan sentiment (positif dan negatif) secara langsung. Pada *Multistage Classification*, tahap pertama mengklasifikasikan aspek (harga, keamanan, atau layanan), dan tahap kedua mengklasifikasikan sentiment (positif dan negatif). Semua uji coba menggunakan pembagian data 80% untuk pelatihan dan 20% untuk pengujian, untuk memastikan model dapat menggeneralisasi dengan baik.

##### 4.1.1 Hasil Skenario Uji 1

Hasil skenario uji 1 menggunakan metode *single stage classification* dengan kernel linear menggunakan hyperparameter  $C=0.1$  dan nilai toleransi 0.01. Data dibagi dengan rasio 80:20, terdiri dari 2.400 data train dan 600 data test.

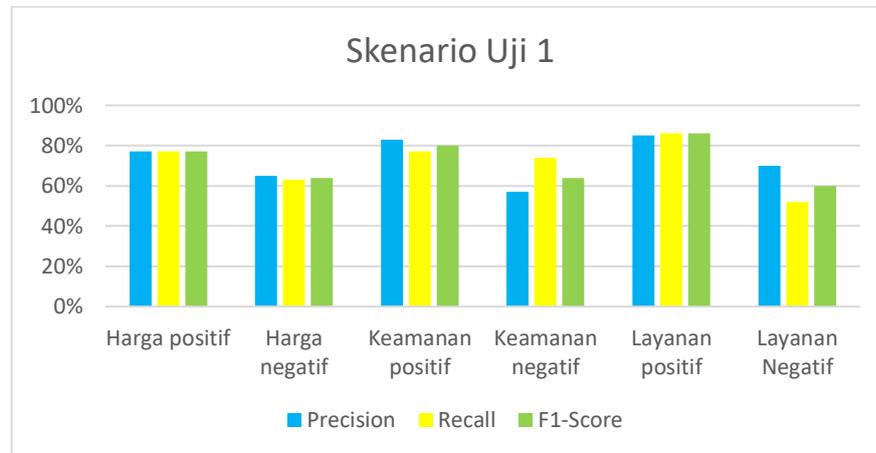


Gambar 4. 1 Hasil Visualisasi Skenario Uji 1

Pada Gambar 4.1 menunjukkan hasil *confusion matrix* pada data testing yang berhasil memprediksi kelas dengan benar sebanyak 449 data dan gagal memprediksi kelas dengan benar sebanyak 151 data, dari total 600 data testing. Hasil tersebut menunjukkan performa model dalam mengklasifikasikan data testing ke dalam 6 kelas. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confusion matrix* dengan akurasi model sebesar 75%, dan hasil disajikan pada Tabel 4.1.

Tabel 4. 1 Score Precision, Recall, F1-Score pada 6 Kelas Skenario Uji 1

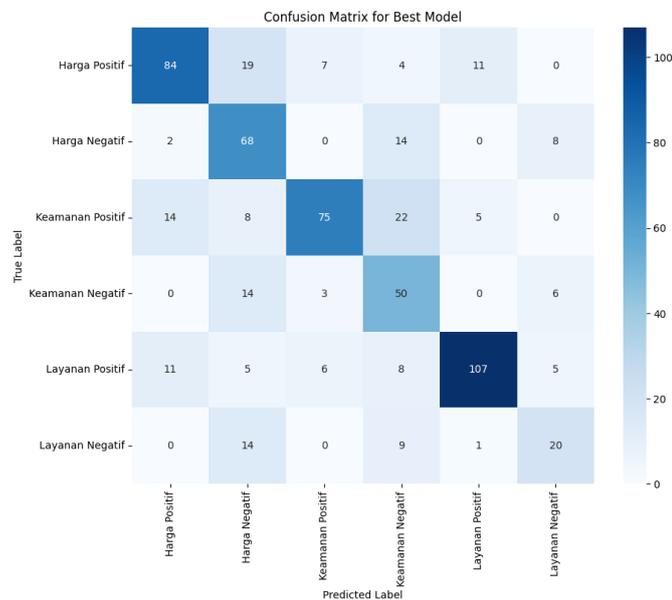
Kelas Sentimen	Confusion Matrix		
	Precision	Recall	F1-Score
Harga Positif	77%	77%	77%
Harga Negatif	65%	63%	64%
Layanan Positif	83%	77%	80%
Layanan Negatif	57%	74%	64%
Keamanan Positif	85%	86%	86%
Keamanan Negatif	70%	52%	60%



Gambar 4. 2 Visual Performa Skenario Uji 1 Kernel Linear Single Stage Classification

#### 4.1.2 Hasil Skenario Uji 2

Hasil skenario uji 2 menggunakan metode *single stage classification* dengan kernel polynomial dengan hyperparameter  $C=100$ ,  $degree=2$ , dan nilai toleransi=0.01. Data dibagi dengan rasio 80:20, terdiri dari 2.400 data train dan 600 data test.

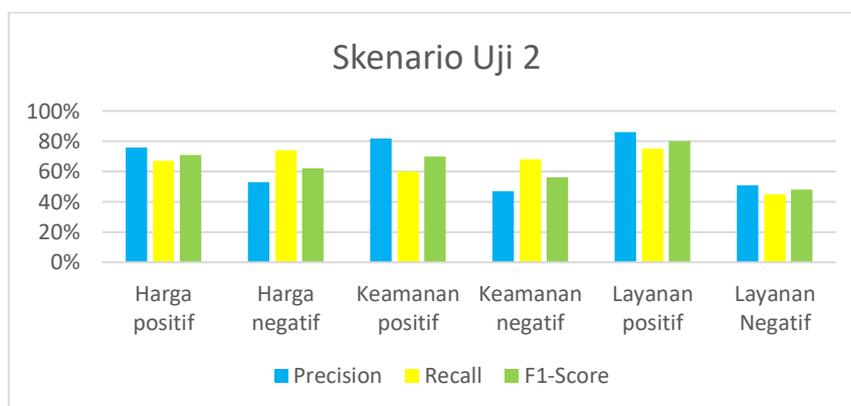


Gambar 4. 3 Hasil Visualisasi Skenario Uji 2

Pada Gambar 4.3 menunjukkan hasil *confusion matrix* pada data testing yang berhasil memprediksi kelas dengan benar sebanyak 404 data dan gagal memprediksi kelas dengan benar sebanyak 196 data, dari total 600 data testing. Hasil tersebut menunjukkan performa model dalam mengklasifikasikan data testing ke dalam 6 kelas. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confusion matrix* dengan akurasi model sebesar 67%, dan hasil lengkap disajikan pada Tabel 4.2.

Tabel 4. 2 Score Precision, Recall, F1-Score pada 6 Kelas Skenario Uji 2

Kelas Sentimen	Confusion Matrix		
	Precision	Recall	F1-Score
Harga Positif	76%	67%	71%
Harga Negatif	53%	74%	62%
Layanan Positif	82%	60%	70%
Layanan Negatif	47%	68%	56%
Keamanan Positif	86%	75%	80%
Keamanan Negatif	51%	45%	48%

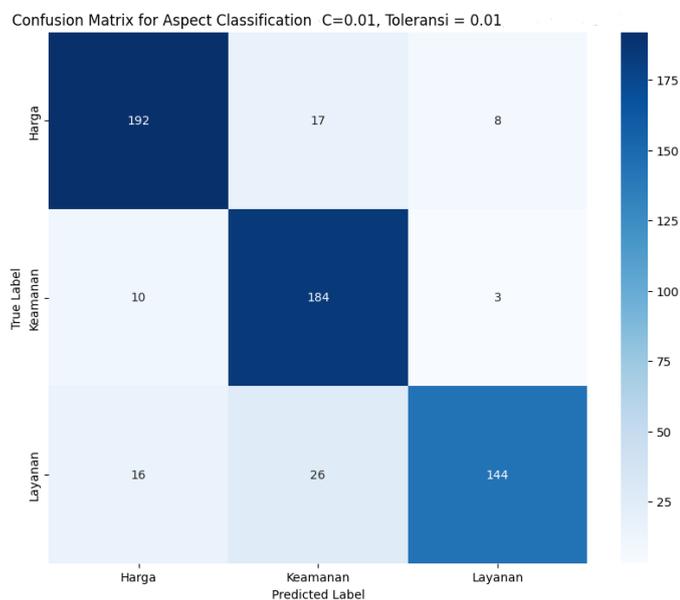


Gambar 4. 4 Visual Performa Skenario Uji 2 Kernel Polinomial Single Stage Classification

### 4.1.3 Hasil Skenario Uji 3

Hasil skenario uji 3 menggunakan metode *multistage classification* dengan kernel linear dengan hyperparameter  $C=0.01$  dan nilai toleransi=0.01. Data dibagi dengan rasio 80:20, terdiri dari 2.400 data train dan 600 data test pada tahap

klasifikasi pertama, dan 800 data train dan 200 data test pada tahap klasifikasi kedua.



Gambar 4. 5 Hasil Visualisasi Klasifikasi Aspek Skenario Uji 3

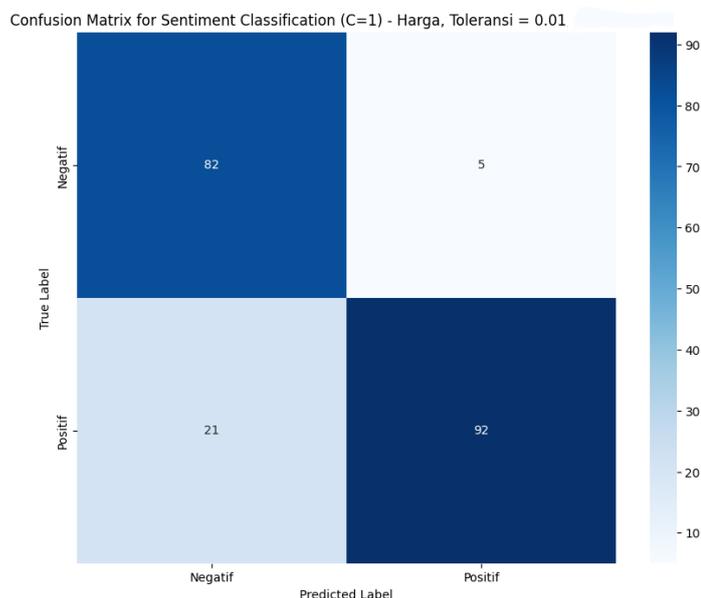
Pada Gambar 4.5 menunjukkan hasil *confusion matrix* pada data testing yang berhasil memprediksi kelas dengan benar sebanyak 520 data dan gagal memprediksi kelas dengan benar sebanyak 80 data, dari total 600 data testing. Hasil tersebut menunjukkan performa model dalam mengklasifikasikan data testing ke dalam 3 kelas aspek. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confusion matrix*, hasil lengkap disajikan pada Tabel 4.3.

Tabel 4. 3 Score Precision, Recall, F1-Score pada 3 Kelas Skenario Uji 3

Kelas Aspek	Confusion Matrix		
	Precision	Recall	F1-Score
Harga	88%	88%	88%
Layanan	81%	93%	86%
Keamanan	92%	77%	84%

Dari *confusion matrix* tersebut, didapatkan nilai akurasi pada skenario uji 3 adalah sebesar 86%. Selanjutnya, dilakukan pelatihan untuk tahap kedua pada

aspek harga. Hasil dari klasifikasi *multistage classification* dengan nilai  $C=1$  dan nilai toleransi=0.01 dan kernel linear adalah sebagai berikut.



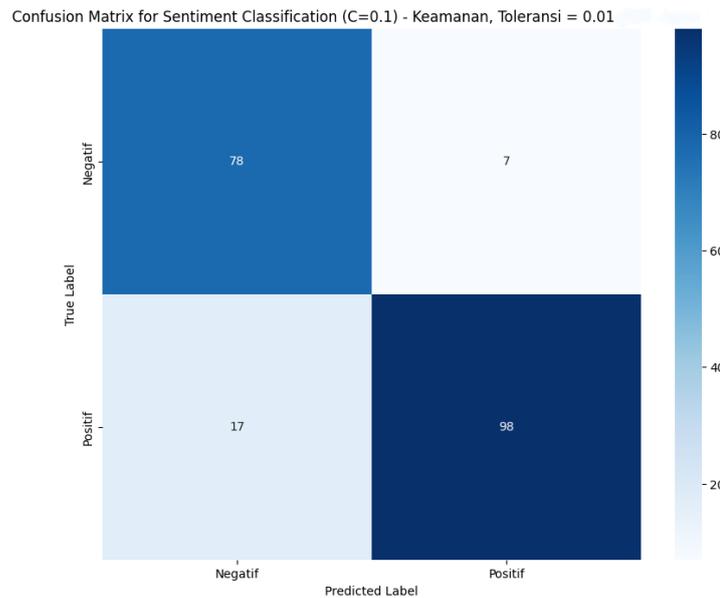
Gambar 4. 6 Hasil Visualisasi Data dari Aspek Harga Skenario Uji 3

Pada Gambar 4.6 menunjukkan hasil *confusion matrix* pada aspek harga dengan hasil akurasi 87%. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confusion matrix* untuk 2 kategori kelas, hasil disajikan pada Tabel 4.4.

Tabel 4. 4 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Harga Skenario Uji 3

Kelas Sentimen	Confusion Matrix		
	Precision	Recall	F1-Score
Negatif	79%	94%	86%
Positif	94%	81%	87%

Selanjutnya, dilakukan pelatihan untuk tahap kedua pada aspek keamanan. Hasil dari klasifikasi *multistage classification* dengan  $C=0.1$ , nilai toleransi=0.01 dan kernel linear adalah sebagai berikut.



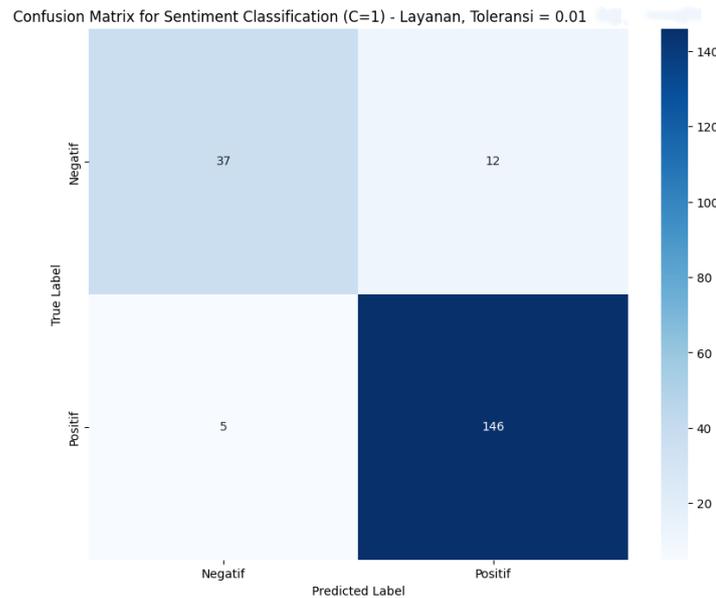
Gambar 4. 7 Hasil Visualisasi Data dari Aspek Keamanan Skenario Uji 3

Pada Gambar 4.7 menunjukkan hasil *confussion matrix* pada aspek keamanan dengan hasil akurasi 88%. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confussion matrix* untuk 2 kategori kelas, hasil disajikan pada Tabel 4.5.

Tabel 4. 5 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Keamanan Skenario Uji 3

Kelas Sentimen	Confusion Matrix		
	Precision	Recall	F1-Score
Negatif	82%	91%	86%
Positif	93%	85%	89%

Selanjutnya, dilakukan pelatihan untuk tahap kedua pada aspek layanan. Hasil dari klasifikasi *multistage classification* dengan C=1, nilai toleransi=0.01 dan kernel linear adalah sebagai berikut.



Gambar 4. 8 Hasil Visualisasi Data dari Aspek Layanan Skenario Uji 3

Pada Gambar 4.8 menunjukkan hasil *confusion matrix* pada aspek layanan dengan hasil akurasi 91%. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari confusion matrix untuk 2 kategori kelas, hasil disajikan pada Tabel 4.6.

Tabel 4. 6 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Layanan Skenario Uji 3

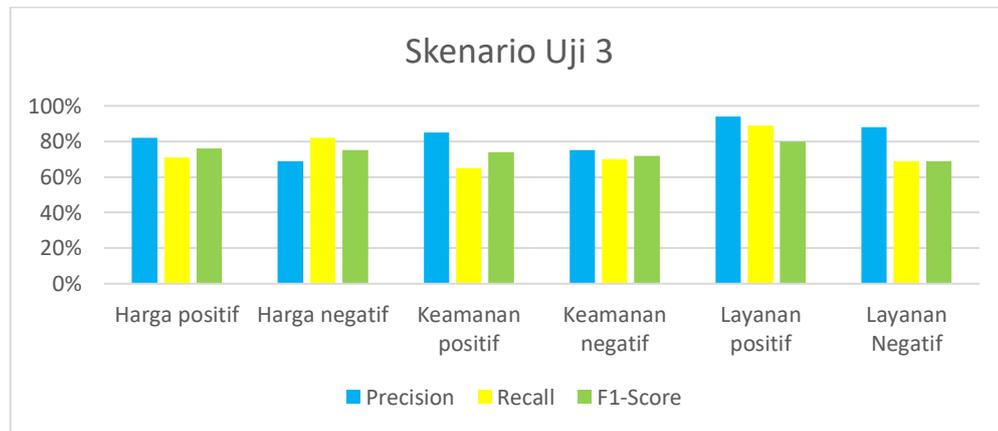
Kelas Sentimen	Confusion Matrix		
	Precision	Recall	F1-Score
Negatif	88%	75%	81%
Positif	92%	96%	94%

Selanjutnya, setelah diketahui semua nilai dari setiap tahap, maka selanjutnya adalah menggabungkan performa yang didapatkan dan memberikan nilai rata-rata dengan cara mengalikan antara tahap ke-1 dan tahap ke-2, dan hasilnya disajikan pada Tabel 4.7.

Tabel 4. 7 Performa Hasil Uji dengan Kernel Linear dengan Metode Multistage classification

Multistage Performa	Confusion Matrix		
	Precision	Recall	F1-Score
Harga Positif	82%	71%	76%
Harga Negatif	69%	82%	75%
Keamanan Positif	85%	65%	74%

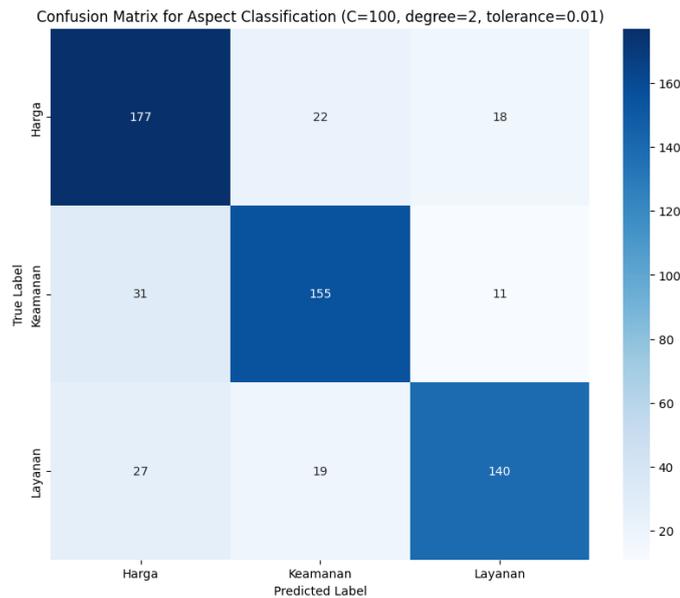
Keamanan Negatif	75%	70%	72%
Layanan Positif	94%	89%	80%
Layanan Negatif	88%	69%	69%



Gambar 4. 9 Visual Performa Skenario Uji 3 Kernel Linear Multistage Classification

#### 4.1.4 Hasil Skenario Uji 4

Hasil skenario uji 4 menggunakan metode *multistage classification* dengan kernel polynomial dengan hyperparameter  $C=100$ ,  $degree=2$ , dan nilai toleransi=0.01. Data dibagi dengan rasio 80:20, terdiri dari 2.400 data train dan 600 data test pada tahap klasifikasi pertama, dan 800 data train dan 200 data test pada tahap klasifikasi kedua.



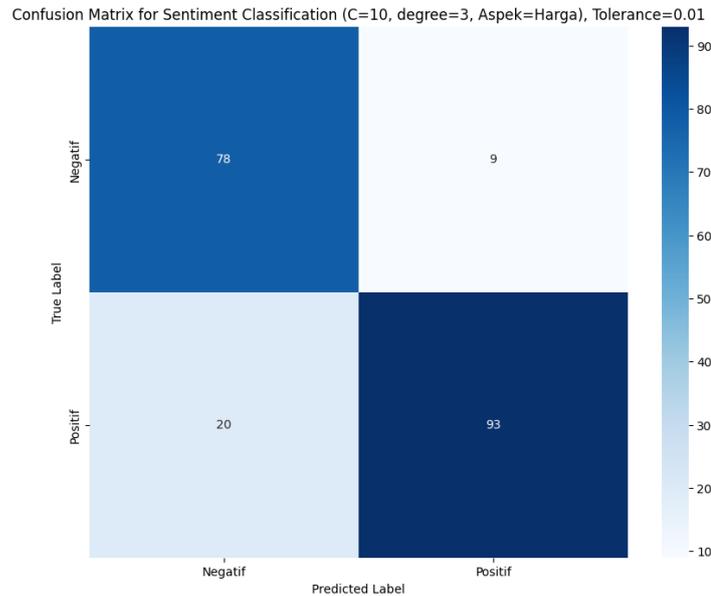
Gambar 4. 10 Hasil Visualisasi Klasifikasi Aspek Skenario Uji 4

Pada Gambar 4.10 menunjukkan hasil *confusion matrix* pada data testing yang berhasil memprediksi kelas dengan benar sebanyak 472 data dan gagal memprediksi kelas dengan benar sebanyak 128 data, dari total 600 data testing. Hasil tersebut menunjukkan performa model dalam mengklasifikasikan data testing ke dalam 3 kelas aspek. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confusion matrix*, hasil lengkap disajikan pada Tabel 4.8.

Tabel 4. 8 Score Precision, Recall, F1-Score pada 3 Kelas Skenario Uji 4

Kelas Aspek	<i>Confusion Matrix</i>		
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Harga	75%	81%	78%
Layanan	79%	78%	78%
Keamanan	82%	75%	78%

Dari *confusion matrix* tersebut, didapatkan nilai akurasi pada skenario uji 4 adalah sebesar 78%. Selanjutnya, dilakukan pelatihan untuk tahap kedua pada aspek harga. Hasil dari klasifikasi *multistage classification* dengan C=10, degree=3, nilai toleransi=0.01, dan kernel polynomial adalah sebagai berikut.



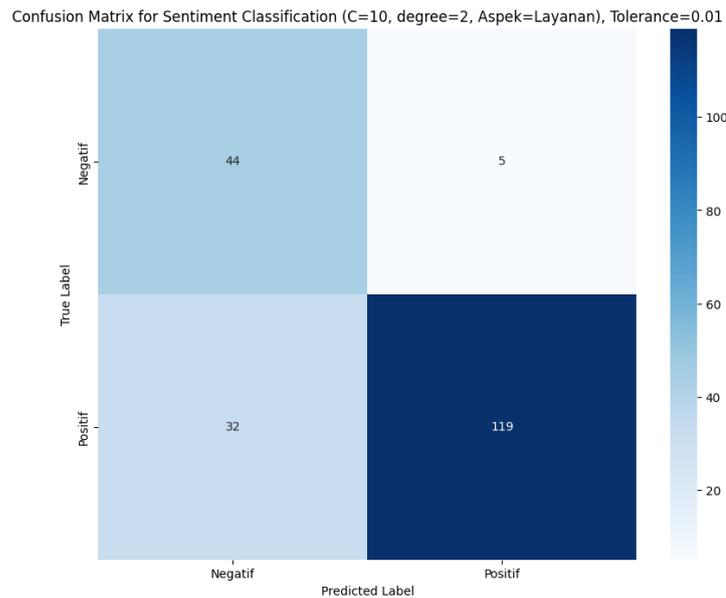
Gambar 4. 11 Hasil Visualisasi Data dari Aspek Harga Skenario Uji 4

Pada Gambar 4.11 menunjukkan hasil confusion matrix pada aspek harga dengan hasil akurasi 85%. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confusion matrix* untuk 2 kategori kelas, hasil disajikan pada Tabel 4.9.

Tabel 4. 9 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Harga Skenario Uji 4

Kelas Sentimen	Confusion Matrix		
	Precision	Recall	F1-Score
Negatif	79%	90%	84%
Positif	91%	82%	86%

Selanjutnya, dilakukan pelatihan untuk tahap kedua pada aspek keamanan. Hasil dari klasifikasi *multistage classification* dengan C=100, degree=2, nilai toleransi=0.01, dan kernel polynomial adalah sebagai berikut.



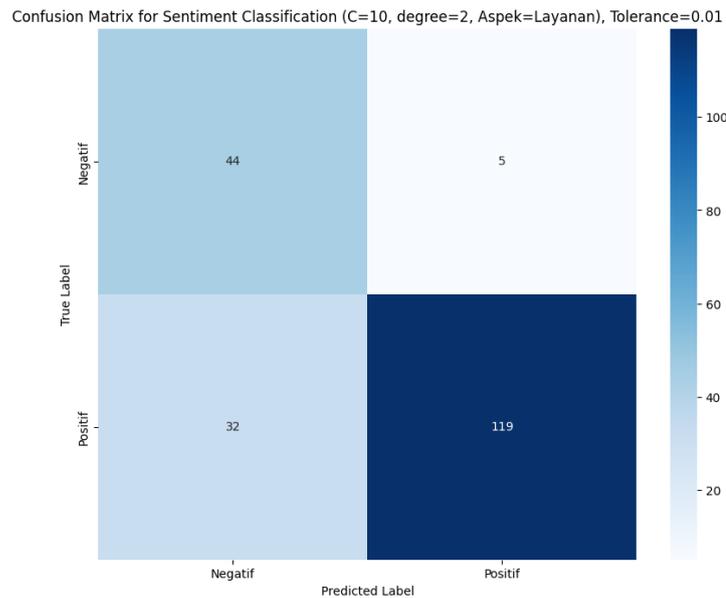
Gambar 4. 12 Hasil Visualisasi Data dari Aspek Keamanan Skenario Uji 4

Pada Gambar 4.12 menunjukkan hasil *confussion matrix* pada aspek keamanan dengan hasil akurasi 79%. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confussion matrix* untuk 2 kategori kelas, hasil disajikan pada Tabel 4.10.

Tabel 4. 10 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Keamanan Skenario Uji 4

Kelas Sentimen	Confusion Matrix		
	Precision	Recall	F1-Score
Negatif	67%	96%	79%
Positif	96%	66%	78%

Selanjutnya, dilakukan pelatihan untuk tahap kedua pada aspek layanan. Hasil dari klasifikasi *multistage classification* dengan C=100, degree=2, nilai toleransi=0.01, dan kernel polynomial adalah sebagai berikut.



Gambar 4. 13 Hasil Visualisasi Data dari Aspek Layanan Skenario Uji 4

Pada Gambar 4.13 menunjukkan hasil *confusion matrix* pada aspek layanan dengan hasil akurasi 81%. Berikut adalah nilai-nilai yang dihitung untuk *precision*, *recall*, dan *F1-Score* dari *confusion matrix* untuk 2 kategori kelas, hasil disajikan pada Tabel 4.11.

Tabel 4. 11 Score Precision, Recall, F1-Score pada 2 Kelas Aspek Layanan Skenario Uji 4

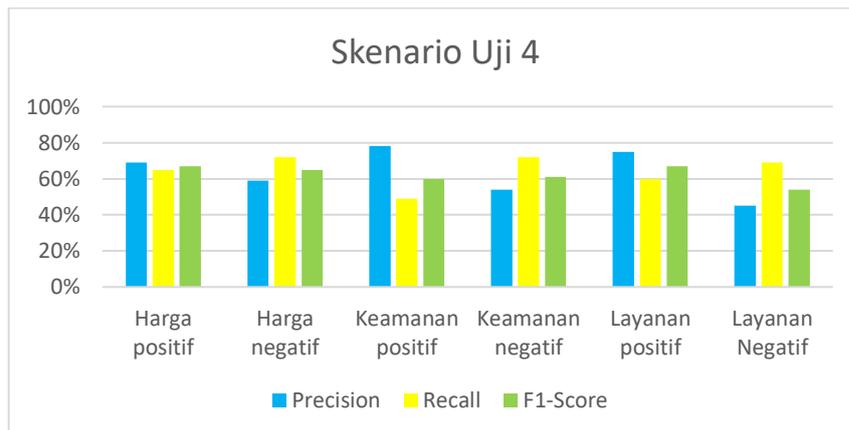
Kelas Sentimen	<i>Confusion Matrix</i>		
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Negatif	57%	89%	70%
Positif	95%	78%	86%

Selanjutnya, setelah diketahui semua nilai dari setiap tahap, maka selanjutnya adalah menggabungkan performa yang didapatkan dan memberikan nilai rata-rata dengan cara mengalikan antara tahap ke-1 dan tahap ke-2, dan hasilnya disajikan pada Tabel 4.12.

Tabel 4. 12 Performa Hasil Uji dengan Kernel Polinomial dengan Metode Multistage classification

<i>Multistage Performa</i>	<i>Confusion Matrix</i>		
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Harga Positif	69%	65%	67%
Harga Negatif	59%	72%	65%

<i>Multistage Performa</i>	<i>Confusion Matrix</i>		
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Keamanan Positif	78%	49%	60%
Keamanan Negatif	54%	72%	61%
Layanan Positif	75%	60%	67%
Layanan Negatif	45%	69%	54%



Gambar 4. 14 Visual Performa Skenario Uji 4 Kernel Polinomial Multistage Classification

## 4.2 Pembahasan

Berdasarkan hasil uji coba yang dilakukan, terdapat empat skenario pengujian dengan pembagian data menggunakan rasio 80:20, di mana 80% data digunakan untuk pelatihan dan 20% untuk pengujian. Pengujian dilakukan menggunakan dua metode klasifikasi, yaitu *single stage classification* dan *multistage classification*, dengan penerapan dua jenis kernel yang berbeda, yaitu kernel linear dan polinomial. Pembagian data dilakukan secara manual sesuai dengan skenario yang telah ditentukan untuk memastikan proporsi data latih dan data uji sesuai. Setelah data dipisahkan, model dilatih menggunakan data latih, dan performanya diuji menggunakan data uji. Kombinasi metode klasifikasi dan kernel yang digunakan dalam setiap skenario bertujuan untuk mengevaluasi efektivitas model secara menyeluruh.

#### 4.2.1 Pembahasan Skenario Uji 1

Setelah melakukan pengujian pada skenario uji 1 dengan tipe klasifikasi *single stage classification* dengan kernel linear menggunakan hyperparameter kombinasi nilai  $C = 0.01, 0.1, 1, 10, 100$  dan nilai tolerance = 0.1, 0.01, 0.001 didapatkan hasil presisi, *recall*, *F1-Score*, dan akurasi seperti table 3.14.

Tabel 4. 13 Performa Uji Hyperparameter pada Skenario Uji 1

C	Tolerance	Presisi	Recall	F1-Score	Akurasi
0.01	0.1	65%	70%	67%	70%
	0.01	65%	70%	67%	70%
	0.001	65%	70%	67%	70%
0.1	0.1	75%	74%	74%	74%
	0.01	75%	74%	74%	74%
	0.001	75%	74%	74%	74%
1	0.1	72%	71%	72%	71%
	0.01	72%	71%	72%	71%
	0.001	72%	71%	72%	71%
10	0.1	69%	68%	68%	68%
	0.01	69%	68%	68%	68%
	0.001	69%	68%	68%	68%
100	0.1	67%	67%	67%	67%
	0.01	67%	67%	67%	67%
	0.001	68%	67%	67%	67%



Gambar 4. 15 Visual Performa Parameter Skenario Uji 1

Berdasarkan gambar 4.15, penggunaan hyperparameter  $C = 0.1$  dan nilai toleransi=0.01 menghasilkan akurasi tertinggi sebesar 74%, sehingga dipilih sebagai parameter terbaik dalam proses pengujian. Berdasarkan gambar 4.15 sebaran data membentuk persamaan kuadrat logaritmis, yaitu di dapatkan melalui perhitungan di Desmos, dengan persamaan  $y_{1,2,3}=a \log(x_1) + b$ . Nilai parameter regresi  $a = -1.2$ ,  $b = 70$ , dengan nilai statistic  $R^2 = 0.48$ .

Tabel 4. 14 Performa Hasil Uji Aspek Harga, Keamanan, Layanan pada Skenario Uji 1

Aspek	Sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Harga	Positif	77%	77%	77%
	Negatif	65%	63%	64%
Keamanan	Positif	83%	77%	80%
	Negatif	57%	74%	64%
Layanan	Positif	85%	86%	86%
	Negatif	70%	52%	60%

Tabel 4.14 menunjukkan hasil pengujian dari percobaan uji hyperparameter terbaik dalam proses pengujian. Pengujian dilakukan dengan mencoba berbagai kombinasi nilai parameter  $C$  (0.01, 0.1, 1, 10, 100) dan toleransi (0.01, 0.001, 0.0001) untuk menemukan kombinasi yang menghasilkan kinerja terbaik. Berdasarkan percobaan, hyperparameter terbaik ditemukan pada  $C = 0.1$  dan  $tol=0.01$ , yang memberikan performa klasifikasi dengan rincian pada aspek Harga, sentimen positif memiliki nilai *precision*, *recall*, dan *F1-Score* masing-masing sebesar 77%, sedangkan sentimen negatif mencapai nilai *precision* 65%, *recall* 63%, dan *F1-Score* 64%. Pada aspek Keamanan, sentimen positif menghasilkan nilai *precision* 83%, *recall* 77%, dan *F1-Score* 80%, sementara sentimen negatif memiliki nilai *precision* 57%, *recall* 74%, dan *F1-Score* 64%. Pada aspek Layanan, sentimen positif mencapai nilai *precision* 85%, *recall* 86%, dan *F1-Score* sebesar

86%, sedangkan sentimen negatif mencapai nilai *precision* 70%, *recall* 52%, dan *F1-Score* 60%.

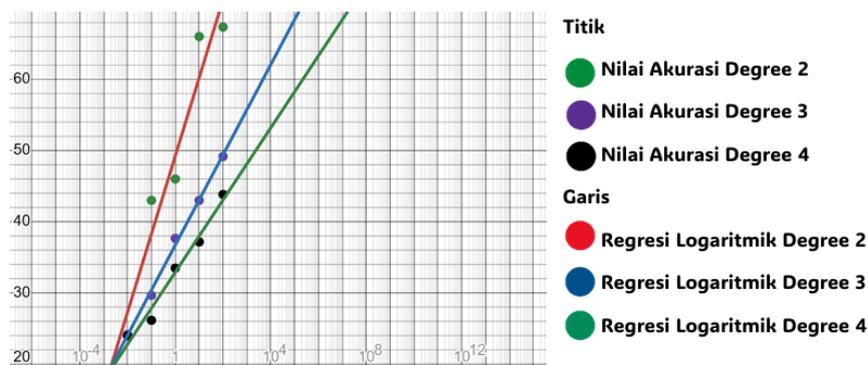
#### 4.2.2 Pembahasan Skenario Uji 2

Setelah melakukan pengujian pada skenario uji 2 dengan tipe klasifikasi *single stage classification* dengan kernel polynomial menggunakan hyperparameter kombinasi nilai  $C = 0.01, 0.1, 1, 10, 100$ , *degree* = 2, 3, 4 dan nilai *tolerance* = 0.1, 0.01, 0.001 didapatkan hasil presisi, *recall*, *F1-Score*, dan akurasi seperti table 3.15.

Tabel 4. 15 Performa Uji Hyperparameter pada Skenario Uji 2

C	Degrre	Tolerance	Presisi	Recall	F1-Score	Akurasi
0.01	2	0.01	26%	24%	10%	24%
		0.001	26%	24%	10%	24%
		0.001	26%	24%	10%	24%
	3	0.01	26%	24%	10%	24%
		0.001	26%	24%	10%	24%
		0.001	26%	24%	10%	24%
	4	0.01	26%	24%	10%	24%
		0.001	26%	24%	10%	24%
		0.001	26%	24%	10%	24%
0.1	2	0.01	45%	43%	36%	43%
		0.001	45%	43%	36%	43%
		0.001	45%	43%	36%	43%
	3	0.01	37%	29%	19%	29%
		0.001	37%	29%	19%	29%
		0.001	37%	29%	19%	29%
	4	0.01	38%	26%	14%	26%
		0.001	38%	26%	14%	26%
		0.001	38%	26%	14%	26%
1	2	0.01	62%	46%	41%	46%
		0.001	62%	46%	41%	46%
		0.001	62%	46%	41%	46%
	3	0.01	60%	37%	32%	37%
		0.001	60%	37%	32%	37%
		0.001	60%	37%	32%	37%
	4	0.01	59%	33%	27%	33%
		0.001	59%	33%	27%	33%

C	Degree	Tolerance	Pesisi	Recall	F1-Score	Akurasi
10	2	0.001	59%	33%	27%	33%
		0.01	70%	66%	66%	66%
		0.001	70%	66%	66%	66%
	3	0.01	55%	43%	38%	43%
		0.001	55%	43%	38%	43%
		0.001	55%	43%	38%	43%
	4	0.01	55%	37%	32%	37%
		0.001	55%	37%	32%	37%
		0.001	55%	37%	32%	37%
	100	2	0.01	70%	67%	68%
0.001			70%	67%	68%	67%
0.001			70%	67%	68%	67%
3		0.01	63%	49%	50%	49%
		0.001	63%	49%	50%	49%
		0.001	63%	49%	50%	49%
4		0.01	60%	43%	44%	43%
		0.001	60%	43%	44%	43%
		0.001	60%	43%	44%	43%



Gambar 4. 16 Visual Performa Parameter Skenario Uji 2

Berdasarkan Gambar 4.16, penggunaan parameter  $C = 100$ , nilai toleransi=0.01, dan degree = 2 dipilih karena menghasilkan akurasi tertinggi sebesar 67%. Pada gambar 4.16, terlihat bahwa nilai hyperparameter  $C = 100$ , nilai toleransi=0.01, dan degree = 2 memiliki performa yang lebih tinggi daripada hyperparameter dan degree yang lain. Berdasarkan gambar 4.15 sebaran data

membentuk persamaan kuadrat logaritmis, yaitu di dapatkan melalui perhitungan di Desmos, dengan persamaan  $y = a \log(x_1) + b$ . Nilai parameter regresi pada nilai  $y_1$  adalah  $a = 10.966$ ,  $b = 49.266$ , dengan nilai statistic  $R^2 = 0.9289$ , untuk nilai  $y_2$  nilai parameter regresinya adalah  $a = 6.333$ ,  $b = 36.736$ , dengan nilai statistic  $R^2 = 0.9963$ , kemudian untuk nilai  $y_3$  nilai parameter regresinya adalah  $a = -5.066$ ,  $b = 32.934$ , dengan nilai statistic  $R^2 = 0.9775$ .

Tabel 4. 16 Performa Hasil Uji Aspek Harga, Keamanan, Layanan pada Skenario Uji 2

Aspek	Sentimen	Precision	Recall	F1-Score
Harga	Positif	76%	67%	71%
	Negatif	53%	74%	62%
Keamanan	Positif	82%	60%	70%
	Negatif	47%	68%	56%
Layanan	Positif	86%	75%	80%
	Negatif	51%	45%	48%

Tabel 4.16 menunjukkan hasil uji percobaan pengujian dengan hyperparameter terbaik, yaitu nilai  $C = 100$  dan  $degree = 2$ . Dalam eksperimen ini, parameter lainnya, seperti nilai  $C$  (0.01, 0.1, 1, 10, 100) dan  $degree$  (2, 3, 4), telah diuji untuk menemukan kombinasi terbaik dalam meningkatkan performa model. Berdasarkan hasil percobaan,  $C = 100$  dan  $degree = 2$  menunjukkan hasil terbaik dengan *precision* 76%, *recall* 67%, dan *F1-score* 71% untuk sentimen positif pada aspek Harga, serta *precision* 53%, *recall* 74%, dan *F1-score* 62% untuk sentimen negatif pada aspek yang sama.

Selain itu, pada aspek Keamanan, *precision* untuk sentimen positif adalah 82%, dengan *recall* 60% dan *F1-score* 70%, sedangkan untuk sentimen negatif, *precision* 47%, *recall* 68%, dan *F1-score* 56%. Pada aspek Layanan, *precision* untuk sentimen positif adalah 86%, *recall* 75%, dan *F1-score* 80%, namun untuk sentimen negatif, *precision* turun menjadi 51%, *recall* 45%, dan *F1-score* 48%.

### 4.2.3 Pembahasan Skenario Uji 3

Setelah melakukan pengujian pada skenario uji 2 dengan tipe klasifikasi *multistage classification* dengan kernel linear menggunakan hyperparameter kombinasi nilai  $C = 0.01, 0.1, 1, 10, 100$  dan nilai *tolerance* = 0.1, 0.01, 0.001 didapatkan hasil presisi, *recall*, *F1-Score*, dan akurasi seperti table 3.17.

Tabel 4. 17 Performa Hasil Uji Klasifikasi Aspek pada Skenario Uji 3

C	Tolerance	Kelas Aspek	Presisi	Recall	F1-Score	Akurasi
0.01	0.01	Harga	88%	88%	88%	86%
		Keamanan	81%	93%	86%	
		Layanan	92%	77%	84%	
	0.001	Harga	88%	88%	88%	86%
		Keamanan	81%	93%	86%	
		Layanan	92%	77%	84%	
	0.0001	Harga	88%	88%	88%	86%
		Keamanan	81%	93%	86%	
		Layanan	92%	77%	84%	
0.1	0.01	Harga	87%	86%	87%	86%
		Keamanan	82%	91%	86%	
		Layanan	90%	81%	85%	
	0.001	Harga	87%	86%	87%	86%
		Keamanan	82%	91%	86%	
		Layanan	90%	81%	85%	
	0.0001	Harga	87%	86%	87%	86%
		Keamanan	82%	91%	86%	
		Layanan	90%	81%	85%	
1	0.01	Harga	84%	81%	82%	83%
		Keamanan	82%	85%	83%	
		Layanan	85%	85%	85%	
	0.001	Harga	84%	81%	82%	84%
		Keamanan	82%	85%	84%	
		Layanan	85%	85%	85%	
	0.0001	Harga	84%	81%	82%	84%
		Keamanan	82%	85%	84%	
		Layanan	85%	85%	85%	
10	0.01	Harga	84%	79%	81%	83%
		Keamanan	82%	82%	82%	
		Layanan	81%	87%	84%	

<b>C</b>	<b>Tolerance</b>	<b>Kelas Aspek</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>
	0.001	Harga	84%	79%	81%	82%
		Keamanan	82%	81%	82%	
		Layanan	80%	87%	83%	
	0.0001	Harga	84%	79%	81%	82%
		Keamanan	82%	81%	82%	
		Layanan	80%	87%	83%	
100	0.01	Harga	84%	78%	81%	82%
		Keamanan	81%	82%	82%	
		Layanan	81%	87%	84%	
	0.001	Harga	84%	78%	81%	82%
		Keamanan	82%	83%	82%	
		Layanan	81%	87%	84%	
	0.0001	Harga	84%	79%	81%	82%
		Keamanan	82%	81%	82%	
		Layanan	80%	87%	83%	

Tabel 4. 18 Performa Hasil Uji Klasifikasi Sentimen Aspek Harga pada Skenario Uji 3

<b>C</b>	<b>Tolerance</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>
0.01	0.01	Positif	56%	99%	72%	56%
		Negatif	50%	1%	2%	
	0.001	Positif	56%	99%	72%	
		Negatif	50%	1%	2%	
	0.0001	Positif	56%	99%	72%	
0.1	0.01	Positif	88%	84%	85%	85%
		Negatif	81%	85%	83%	
	0.001	Positif	88%	84%	85%	
		Negatif	81%	85%	83%	
	0.0001	Positif	88%	84%	85%	
1	0.01	Positif	94%	81%	87%	87%
		Negatif	79%	94%	86%	
	0.001	Positif	94%	81%	87%	
		Negatif	79%	94%	86%	
	0.0001	Positif	94%	81%	87%	
10	0.01	Positif	93%	82%	87%	86%
		Negatif	80%	91%	85%	
	0.001	Positif	93%	82%	87%	
		Negatif	80%	91%	85%	
	0.0001	Positif	93%	82%	87%	
100	0.01	Positif	93%	82%	87%	86%
		Negatif	80%	91%	85%	
	0.001	Positif	93%	82%	87%	
		Negatif	80%	91%	85%	
	0.0001	Positif	93%	82%	87%	

<b>C</b>	<b>Tolerance</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>
		Negatif	80%	91%	85%	

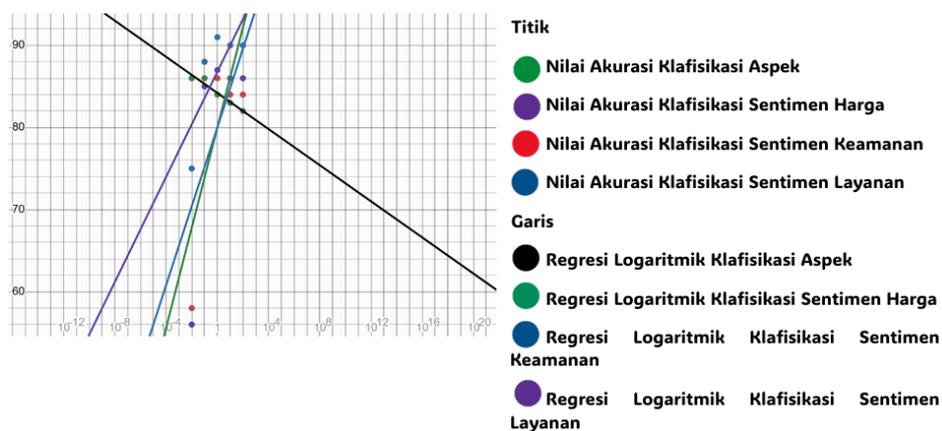
Tabel 4. 19 Performa Hasil Uji Klasifikasi Sentimen Aspek Keamanan pada Skenario Uji 3

<b>C</b>	<b>Tolerance</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>
0.01	0.01	Positif	57%	100%	73%	58%
		Negatif	100%	1%	2%	
	0.001	Positif	57%	100%	73%	
		Negatif	100%	1%	2%	
	0.0001	Positif	57%	100%	73%	
		Negatif	100%	1%	2%	
0.1	0.01	Positif	93%	85%	89%	88%
		Negatif	82%	91%	86%	
	0.001	Positif	93%	85%	89%	
		Negatif	82%	91%	86%	
	0.0001	Positif	93%	85%	89%	
		Negatif	82%	91%	86%	
1	0.01	Positif	93%	82%	87%	86%
		Negatif	79%	91%	85%	
	0.001	Positif	93%	82%	87%	
		Negatif	79%	91%	85%	
	0.0001	Positif	93%	82%	87%	
		Negatif	79%	91%	85%	
10	0.01	Positif	90%	81%	85%	84%
		Negatif	78%	88%	82%	
	0.001	Positif	90%	81%	85%	
		Negatif	78%	88%	82%	
	0.0001	Positif	90%	81%	85%	
		Negatif	78%	88%	82%	
100	0.01	Positif	90%	80%	85%	84%
		Negatif	77%	88%	82%	
	0.001	Positif	90%	80%	85%	
		Negatif	77%	88%	82%	
	0.0001	Positif	90%	80%	85%	
		Negatif	77%	88%	82%	

Tabel 4. 20 Performa Hasil Uji Klasifikasi Sentimen Aspek Layanan pada Skenario Uji 3

<b>C</b>	<b>Tolerance</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>
0.01	0.01	Positif	0%	0%	0%	75%
		Negatif	75%	100%	86%	
	0.001	Positif	0%	0%	0%	
		Negatif	75%	100%	86%	
	0.0001	Positif	0%	0%	0%	
		Negatif	75%	100%	86%	
0.1	0.01	Positif	88%	98%	92%	88%
		Negatif	90%	59%	71%	
	0.001	Positif	88%	98%	92%	
		Negatif	90%	59%	71%	
	0.0001	Positif	88%	98%	92%	
		Negatif	90%	59%	71%	
1	0.01	Positif	92%	96%	94%	91%
		Negatif	88%	75%	81%	

C	Tolerance	Sentimen	Presisi	Recall	F1-Score	Akurasi
	0.001	Positif	92%	96%	94%	
		Negatif	88%	75%	81%	
	0.0001	Positif	92%	96%	94%	
		Negatif	88%	75%	81%	
10	0.01	Positif	92%	95%	93%	90%
		Negatif	84%	75%	79%	
	0.001	Positif	92%	95%	93%	
		Negatif	84%	75%	79%	
	0.0001	Positif	92%	95%	93%	
		Negatif	84%	75%	79%	
100	0.01	Positif	92%	94%	93%	90%
		Negatif	82%	75%	78%	
	0.001	Positif	92%	94%	93%	
		Negatif	82%	75%	78%	
	0.0001	Positif	92%	94%	93%	
		Negatif	82%	75%	78%	



Gambar 4. 17 Visual Performa Parameter Skenario Uji 3

Berdasarkan gambar 4.17, penggunaan parameter  $C = 0.01$  dan nilai toleransi=0.01 menghasilkan akurasi aspek tertinggi sebesar 86%, sehingga dipilih sebagai parameter terbaik dalam proses pengujian klasifikasi aspek. Pada aspek harga, hyperparameter  $C$  yg terbaik adalah  $C=1$  dan nilai toleransi=0.01 yang menghasilkan akurasi 87%, pada aspek Harga, nilai hyperparameter terbaik adalah  $C=0.1$  dan nilai toleransi=0.01 yang menghasilkan akurasi 88%, dan pada aspek layanan, hyperparameter terbaik dalah  $C=1$  dan nilai toleransi=0.01 yang menghasilkan akurasi 91%. Berdasarkan gambar 4.17 sebaran data membentuk

persamaan kuadrat logaritmis, yaitu di dapatkan melalui perhitungan di Desmos, dengan persamaan  $y = a \log(x_1) + b$ . Nilai parameter regresi pada nilai  $y_1$  adalah  $a = -1.1$ ,  $b = 84.2$ , dengan nilai statistic  $R^2 = 0.9453$ , untuk nilai  $y_2$  nilai parameter regresinya adalah  $a = 6.1$ ,  $b = 80$ , dengan nilai statistic  $R^2 = 0.5154$ , untuk nilai  $y_3$  nilai parameter regresinya adalah  $a = 4.8$ ,  $b = 80$ , dengan nilai statistic  $R^2 = 0.374$ , kemudian untuk nilai  $y_4$  nilai parameter regresinya adalah  $a = 3.2$ ,  $b = 86.8$ , dengan nilai statistic  $R^2 = 0.5727$ .

Tabel 4. 21 Performa Hasil Uji Aspek Harga, Keamanan, Layanan pada Skenario Uji 3

<i>Multistage Performa</i>	<i>Confusion Matrix</i>		
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Harga Positif	82%	71%	76%
Harga Negatif	69%	82%	75%
Keamanan Positif	85%	65%	74%
Keamanan Negatif	75%	70%	72%
Layanan Positif	94%	89%	80%
Layanan Negatif	88%	69%	69%

Tabel 4.21 menunjukkan hasil uji performa model yang diperoleh dari pengujian nilai hyperparameter C. Pengujian dilakukan dengan berbagai nilai parameter C, yaitu 0.01, 0.1, 1, 10, dan 100, dan nilai toleransi 0.01, 0.001, dan 0.0001 untuk menentukan nilai yang menghasilkan performa optimal dalam nilai *precision*, *recall*, dan *F1-score*.

#### 4.2.4 Pembahasan Skenario Uji 4

Setelah melakukan pengujian pada skenario uji 2 dengan tipe klasifikasi *multistage classification* dengan kernel linear menggunakan hyperparameter

kombinasi nilai  $C = 0.01, 0.1, 1, 10, 100$ ,  $degree = 2, 3, 4$ , dan nilai  $tolerance = 0.1, 0.01, 0.001$  didapatkan hasil presisi, *recall*, *F1-Score*, dan akurasi seperti table 3.22.

Tabel 4. 22 Performa Hasil Uji Klasifikasi Aspek pada Skenario Uji 4

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Aspek</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>	
0.01	0.01	2	Harga	0%	0%	0%	31%	
			Keamanan	100%	2%	3%		
			Layanan	31%	100%	47%		
		3	Harga	0%	0%	0%	36%	
			Keamanan	33%	100%	50%		
			Layanan	100%	10%	18%		
		4	Harga	66%	0%	0%	35%	
			Keamanan	33%	100%	50%		
			Layanan	100%	0%	1%		
	0.001	0.001	2	Harga	0%	0%	0%	31%
				Keamanan	100%	2%	3%	
				Layanan	31%	100%	47%	
			3	Harga	0%	0%	0%	36%
				Keamanan	33%	100%	50%	
				Layanan	100%	10%	18%	
			4	Harga	66%	0%	0%	35%
				Keamanan	33%	100%	50%	
				Layanan	100%	0%	1%	
	0.0001	0.0001	2	Harga	0%	0%	0%	31%
				Keamanan	100%	2%	3%	
				Layanan	31%	100%	47%	
			3	Harga	0%	0%	0%	36%
				Keamanan	33%	100%	50%	
				Layanan	100%	10%	18%	
4			Harga	66%	0%	0%	35%	
			Keamanan	33%	100%	50%		
			Layanan	100%	0%	1%		
0.1	0.01	2	Harga	42%	100%	59%	51%	

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Aspek</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>	
			Keamanan	100%	19%	33%		
			Layanan	100%	27%	43%		
		3	Harga	40%	100%	57%	46%	
			Keamanan	100%	11%	20%		
			Layanan	100%	19%	32%		
		4	Harga	39%	100%	56%	43%	
			Keamanan	100%	0%	11%		
			Layanan	100%	0%	30%		
		0.001	2	Harga	42%	100%	59%	51%
				Keamanan	100%	19%	33%	
				Layanan	100%	27%	43%	
			3	Harga	40%	100%	57%	45%
	Keamanan			100%	11%	20%		
	Layanan			100%	19%	32%		
	4		Harga	39%	100%	56%	43%	
			Keamanan	100%	0%	11%		
			Layanan	100%	0%	30%		
	0.0001		2	Harga	42%	100%	59%	51%
				Keamanan	100%	19%	33%	
				Layanan	100%	27%	43%	
		3	Harga	40%	100%	57%	45%	
			Keamanan	100%	11%	20%		
			Layanan	100%	19%	32%		
		4	Harga	39%	100%	56%	43%	
Keamanan			100%	0%	11%			
Layanan			100%	0%	30%			
1		0.01	2	Harga	50%	97%	66%	63%
				Keamanan	90%	40%	55%	
				Layanan	97%	47%	64%	
	3		Harga	42%	98%	59%	51%	
			Keamanan	94%	17%	29%		
			Layanan	96%	33%	50%		

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Aspek</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>
	0.001	4	Harga	40%	98%	57%	47%
			Keamanan	91%	10%	19%	
			Layanan	95%	25%	40%	
		2	Harga	50%	97%	66%	63%
			Keamanan	90%	40%	55%	
			Layanan	97%	47%	64%	
		3	Harga	42%	98%	59%	51%
			Keamanan	94%	17%	29%	
			Layanan	96%	33%	50%	
	4	Harga	40%	98%	57%	47%	
		Keamanan	91%	10%	19%		
		Layanan	95%	25%	40%		
	0.0001	2	Harga	50%	97%	66%	63%
			Keamanan	90%	40%	55%	
			Layanan	97%	47%	64%	
		3	Harga	42%	98%	59%	51%
			Keamanan	94%	17%	29%	
			Layanan	96%	33%	50%	
4		Harga	40%	98%	57%	47%	
		Keamanan	91%	10%	19%		
		Layanan	95%	25%	40%		
10	0.01	2	Harga	69%	84%	76%	76%
			Keamanan	78%	77%	77%	
			Layanan	84%	65%	73%	
		3	Harga	47%	94%	62%	58%
			Keamanan	81%	29%	43%	
			Layanan	91%	46%	61%	
		4	Harga	43%	95%	59%	52%
			Keamanan	82%	19%	31%	
			Layanan	91%	36%	52%	
0.001	2	Harga	69%	84%	76%	76%	
		Keamanan	78%	77%	77%		

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Aspek</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>	
		3	Layanan	84%	65%	73%	58%	
			Harga	47%	94%	62%		
			Keamanan	81%	29%	43%		
		4	Layanan	91%	46%	61%	52%	
			Harga	43%	95%	59%		
			Keamanan	82%	19%	31%		
		0.0001	2	Layanan	91%	36%	52%	76%
				Harga	69%	84%	76%	
				Keamanan	78%	77%	77%	
	3		Harga	47%	94%	62%	58%	
			Keamanan	81%	29%	43%		
			Layanan	91%	46%	61%		
	4	Harga	43%	95%	59%	52%		
		Keamanan	82%	19%	31%			
		Layanan	91%	36%	52%			
100	0.01	2	Harga	75%	81%	78%	78%	
			Keamanan	79%	78%	78%		
			Layanan	82%	75%	78%		
		3	Harga	51%	86%	64%	62%	
			Keamanan	79%	44%	57%		
			Layanan	77%	53%	63%		
		4	Harga	45%	89%	60%	55%	
			Keamanan	74%	26%	38%		
			Layanan	84%	45%	58%		
	0.001	2	Harga	75%	81%	78%	78%	
			Keamanan	79%	78%	78%		
			Layanan	82%	75%	78%		
		3	Harga	51%	86%	64%	62%	
			Keamanan	79%	44%	57%		
			Layanan	77%	53%	63%		
4	Harga	45%	89%	60%	55%			

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Aspek</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>		
			Keamanan	74%	26%	38%			
			Layanan	84%	45%	58%			
	0.0001	2	2	Harga	75%	81%	78%	78%	
				Keamanan	79%	78%	78%		
				Layanan	82%	75%	78%		
			3	Harga	51%	86%	64%		62%
				Keamanan	79%	44%	57%		
				Layanan	77%	53%	63%		
		4	Harga	45%	89%	60%	55%		
			Keamanan	74%	26%	38%			
			Layanan	84%	45%	58%			

Tabel 4. 23 Performa Hasil Uji Klasifikasi Sentimen Aspek Harga pada Skenario Uji 4

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>
0.01	0.01	2	Positif	56%	100%	72%	56%
			Negatif	0%	0%	0%	
		3	Positif	56%	100%	72%	
			Negatif	0%	0%	0%	
		4	Positif	56%	100%	72%	
			Negatif	0%	0%	0%	
	0.001	2	Positif	56%	100%	72%	56%
			Negatif	0%	0%	0%	
		3	Positif	56%	100%	72%	
			Negatif	0%	0%	0%	
		4	Positif	56%	100%	72%	
			Negatif	0%	0%	0%	
0.0001	2	Positif	56%	100%	72%	56%	
		Negatif	0%	0%	0%		
	3	Positif	56%	100%	72%		
		Negatif	0%	0%	0%		
	4	Positif	56%	100%	72%		
		Negatif	0%	0%	0%		
0.1	0.01	2	Positif	56%	100%	72%	56%
			Negatif	0%	0%	0%	
		3	Positif	56%	100%	72%	
			Negatif	0%	0%	0%	
		4	Positif	56%	100%	72%	
			Negatif	0%	0%	0%	
	0.001	2	Positif	56%	100%	72%	56%
			Negatif	0%	0%	0%	
		3	Positif	56%	100%	72%	
			Negatif	0%	0%	0%	
		4	Positif	56%	100%	72%	
			Negatif	0%	0%	0%	

C	Tolerance	Degree	Sentimen	Presisi	Recall	F1-Score	Akurasi	
1	0.0001	2	Positif	56%	100%	72%	56%	
			Negatif	0%	0%	0%		
		3	Positif	56%	100%	72%		
			Negatif	0%	0%	0%		
		4	Positif	56%	100%	72%		
			Negatif	0%	0%	0%		
	1	0.01	2	Positif	56%	100%	72%	57%
				Negatif	100%	0%	0%	
			3	Positif	56%	100%	72%	
				Negatif	0%	0%	0%	
			4	Positif	56%	100%	72%	
				Negatif	0%	0%	0%	
1		0.001	2	Positif	56%	100%	72%	57%
				Negatif	100%	0%	0%	
			3	Positif	56%	100%	72%	
				Negatif	0%	0%	0%	
			4	Positif	56%	100%	72%	
				Negatif	0%	0%	0%	
	1	0.0001	2	Positif	56%	100%	72%	57%
				Negatif	100%	0%	0%	
			3	Positif	56%	100%	72%	
				Negatif	0%	0%	0%	
			4	Positif	56%	100%	72%	
				Negatif	0%	0%	0%	
10		0.01	2	Positif	98%	57%	72%	75%
				Negatif	64%	98%	77%	
			3	Positif	91%	82%	86%	85%
				Negatif	79%	90%	84%	
			4	Positif	56%	99%	72%	57%
				Negatif	66%	2%	4%	
	10	0.001	2	Positif	98%	57%	72%	75%
				Negatif	64%	98%	77%	
			3	Positif	91%	82%	86%	85%
				Negatif	79%	90%	84%	
			4	Positif	56%	99%	72%	57%
				Negatif	66%	2%	4%	
10		0.0001	2	Positif	98%	57%	72%	75%
				Negatif	64%	98%	77%	
			3	Positif	91%	82%	86%	85%
				Negatif	79%	90%	84%	
			4	Positif	56%	99%	72%	57%
				Negatif	66%	2%	4%	
	100	0.01	2	Positif	98%	59%	74%	76%
				Negatif	65%	98%	78%	
			3	Positif	100%	23%	38%	57%
				Negatif	50%	100%	66%	
			4	Positif	100%	13%	23%	51%
				Negatif	47%	100%	63%	
100		0.001	2	Positif	98%	59%	74%	76%
				Negatif	65%	98%	78%	
			3	Positif	100%	23%	38%	57%
		Negatif		50%	100%	66%		
		Negatif		50%	100%	66%		

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>	
		4	Positif	100%	13%	23%	51%	
			Negatif	47%	100%	63%		
	0.0001	2	2	Positif	98%	59%	74%	76%
				Negatif	65%	98%	78%	
		3	3	Positif	100%	23%	38%	57%
				Negatif	50%	100%	66%	
		4	4	Positif	100%	13%	23%	51%
				Negatif	47%	100%	63%	

Tabel 4. 24 Performa Hasil Uji Klasifikasi Sentimen Aspek Keamanan pada Skenario Uji 4

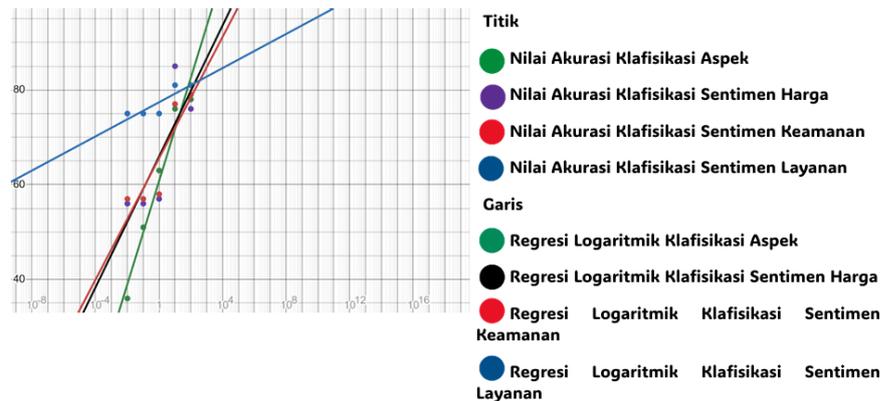
<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>		
0.01	0.01	2	Positif	57%	100%	73%	57%		
			Negatif	0%	0%	0%			
		3	3	Positif	57%	100%		73%	
				Negatif	0%	0%		0%	
		4	4	Positif	57%	100%		73%	
				Negatif	0%	0%		0%	
		0.0001	2	2	Positif	57%		100%	73%
					Negatif	0%		0%	0%
			3	3	Positif	57%		100%	73%
					Negatif	0%		0%	0%
			4	4	Positif	57%		100%	73%
					Negatif	0%		0%	0%
	0.1	0.01	2	Positif	57%	100%	73%	57%	
				Negatif	0%	0%	0%		
			3	3	Positif	57%	100%		73%
					Negatif	0%	0%		0%
			4	4	Positif	57%	100%		73%
					Negatif	0%	0%		0%
		0.0001	2	2	Positif	57%	100%		73%
					Negatif	0%	0%		0%
			3	3	Positif	57%	100%		73%
					Negatif	0%	0%		0%
			4	4	Positif	57%	100%		73%
					Negatif	0%	0%		0%
1	0.01	2	Positif	57%	100%	73%	58%		
			Negatif	100%	0%	0%			
		3	3	Positif	57%	100%	73%	58%	
				Negatif	100%	1%	2%		

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>	
		4	Positif	57%	100%	73%	57%	
			Negatif	0%	0%	0%		
	0.001	2	2	Positif	57%	100%	73%	58%
				Negatif	100%	0%	0%	
		3	3	Positif	57%	100%	73%	58%
				Negatif	100%	1%	2%	
	4	4	Positif	57%	100%	73%	57%	
			Negatif	100%	1%	2%		
	0.0001	2	2	Positif	57%	100%	73%	58%
				Negatif	100%	0%	0%	
		3	3	Positif	57%	100%	73%	58%
				Negatif	100%	0%	0%	
4	4	Positif	57%	100%	73%	57%		
		Negatif	0%	0%	0%			
10	0.01	2	Positif	96%	63%	76%	77%	
			Negatif	66%	96%	78%		
		3	3	Positif	96%	49%	65%	70%
				Negatif	58%	97%	73%	
	4	4	Positif	57%	100%	73%	58%	
			Negatif	100%	1%	2%		
	0.001	2	2	Positif	96%	63%	76%	77%
				Negatif	57%	89%	70%	
		3	3	Positif	96%	49%	65%	70%
				Negatif	58%	97%	73%	
	4	4	Positif	57%	100%	73%	58%	
			Negatif	100%	1%	2%		
0.0001	2	2	Positif	96%	63%	76%	77%	
			Negatif	57%	89%	70%		
	3	3	Positif	96%	49%	65%	70%	
			Negatif	58%	97%	73%		
4	4	Positif	57%	100%	73%	58%		
		Negatif	100%	1%	2%			
100	0.01	2	Positif	96%	66%	78%	79%	
			Negatif	67%	96%	79%		
		3	3	Positif	100%	24%	39%	56%
				Negatif	49%	100%	66%	
	4	4	Positif	100%	16%	28%	52%	
			Negatif	46%	100%	63%		
	0.001	2	2	Positif	96%	66%	78%	79%
				Negatif	67%	96%	79%	
		3	3	Positif	100%	24%	39%	56%
				Negatif	49%	100%	66%	
	4	4	Positif	100%	16%	28%	52%	
			Negatif	46%	100%	63%		
0.0001	2	2	Positif	96%	66%	78%	79%	
			Negatif	67%	96%	79%		
	3	3	Positif	100%	24%	39%	56%	
			Negatif	49%	100%	66%		
4	4	Positif	100%	16%	28%	52%		
		Negatif	46%	100%	63%			

Tabel 4. 25 Performa Hasil Uji Klasifikasi Sentimen Aspek Layanan pada Skenario Uji 4

<b>C</b>	<b>Tolerance</b>	<b>Degree</b>	<b>Sentimen</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Akurasi</b>
0.01	0.01	2	Positif	75%	100%	86%	75%
			Negatif	0%	0%	0%	
		3	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
		4	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
	0.001	2	Positif	75%	100%	86%	75%
			Negatif	0%	0%	0%	
		3	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
		4	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
0.0001	2	Positif	75%	100%	86%	75%	
		Negatif	0%	0%	0%		
	3	Positif	75%	100%	86%		
		Negatif	0%	0%	0%		
	4	Positif	75%	100%	86%		
		Negatif	0%	0%	0%		
0.1	0.01	2	Positif	75%	100%	86%	75%
			Negatif	0%	0%	0%	
		3	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
		4	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
	0.001	2	Positif	75%	100%	86%	75%
			Negatif	0%	0%	0%	
		3	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
		4	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
0.0001	2	Positif	75%	100%	86%	75%	
		Negatif	0%	0%	0%		
	3	Positif	75%	100%	86%		
		Negatif	0%	0%	0%		
	4	Positif	75%	100%	86%		
		Negatif	0%	0%	0%		
1	0.01	2	Positif	75%	100%	86%	75%
			Negatif	0%	0%	0%	
		3	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
		4	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
	0.001	2	Positif	75%	100%	86%	75%
			Negatif	0%	0%	0%	
		3	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
		4	Positif	75%	100%	86%	
			Negatif	0%	0%	0%	
0.0001	2	Positif	75%	100%	86%	75%	

<i>C</i>	<i>Tolerance</i>	<i>Degree</i>	<i>Sentimen</i>	<i>Presisi</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Akurasi</i>	
10	0.01	3	Negatif	0%	0%	0%	75%	
			Positif	75%	100%	86%		
		4	Negatif	0%	0%	0%		
			Positif	75%	100%	86%		
		0.001	2	Positif	95%	78%		86%
				Negatif	57%	89%		70%
	3		Positif	75%	100%	86%		
			Negatif	0%	0%	0%		
	4		Positif	75%	100%	86%		
			Negatif	0%	0%	0%		
	0.0001	2	Positif	95%	78%	86%		
			Negatif	57%	89%	70%		
3		Positif	75%	100%	86%			
		Negatif	0%	0%	0%			
4		Positif	75%	100%	86%			
		Negatif	0%	0%	0%			
100	0.01	2	Positif	95%	78%	86%	81%	
			Negatif	57%	89%	70%		
		3	Positif	96%	50%	66%	61%	
			Negatif	38%	93%	54%		
		4	Positif	95%	43%	60%	56%	
			Negatif	35%	93%	51%		
	0.001	2	Positif	95%	78%	86%	81%	
			Negatif	57%	89%	70%		
		3	Positif	96%	50%	66%	61%	
			Negatif	38%	93%	54%		
		4	Positif	95%	43%	60%	56%	
			Negatif	35%	93%	51%		
	0.0001	2	Positif	95%	78%	86%	81%	
			Negatif	57%	89%	70%		
		3	Positif	96%	50%	66%	61%	
			Negatif	38%	93%	54%		
		4	Positif	95%	43%	60%	56%	
			Negatif	35%	93%	51%		



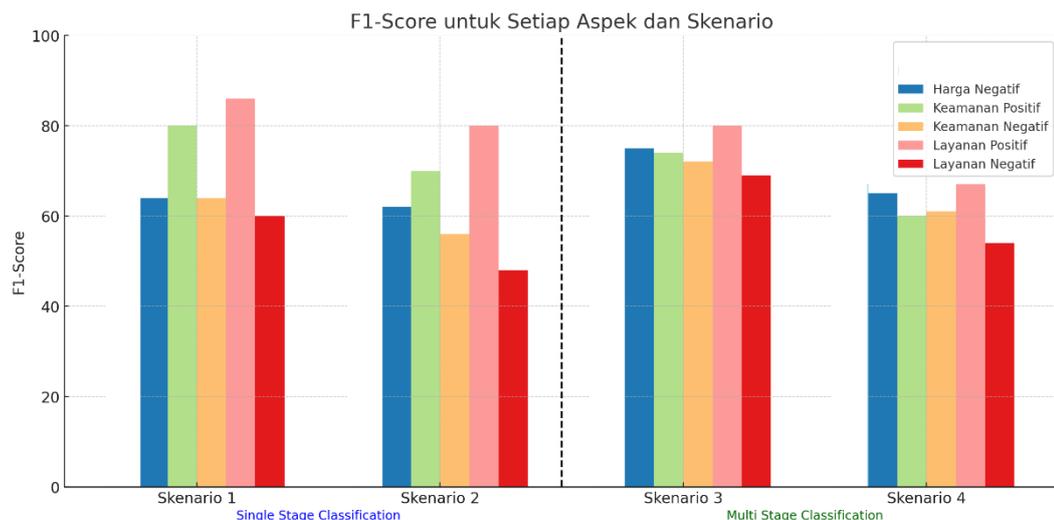
Gambar 4. 18 Visual Performa Parameter Skenario Uji 4

Berdasarkan gambar 4.18, penggunaan parameter  $C = 100$ ,  $degree = 2$ , dan nilai toleransi=0.01 menghasilkan akurasi aspek tertinggi sebesar 78%, sehingga dipilih sebagai parameter terbaik dalam proses pengujian klasifikasi aspek. Pada aspek harga, hyperparameter  $C$  yg terbaik adalah  $C=100$ ,  $degree=2$  dan nilai toleransi=0.01, yang menghasilkan akurasi 85%, pada aspek keamanan, nilai hyperparameter terbaik adalah  $C=100$ ,  $degree=2$ , dan nilai toleransi=0.01 yang menghasilkan akurasi 79%, dan pada aspek layanan, hyperparameter terbaik dalah  $C=10$ ,  $degree=2$ , dan nilai toleransi=0.01 yang menghasilkan akurasi 81%. Berdasarkan gambar 4.18 sebaran data membentuk persamaan kuadrat logaritmis, yaitu di dapatkan melalui perhitungan di Desmos, dengan persamaan  $y=a \log(x_1) + b$ . Nilai parameter regresi pada nilai  $y_1$  adalah  $a = 10.9$ ,  $b = 60.8$ , dengan nilai statistic  $R^2 = 0.956$ , untuk nilai  $y_2$  nilai parameter regresinya adalah  $a = 6.9$ ,  $b = 66$ , dengan nilai statistic  $R^2 = 0.6416$ , untuk nilai  $y_3$  nilai parameter regresinya adalah  $a = 6.4$ ,  $b = 65.6$ , dengan nilai statistic  $R^2 = 0.795$ , kemudian untuk nilai  $y_4$  nilai parameter regresinya adalah  $a = 1.8$ ,  $b = 77.4$ , dengan nilai statistic  $R^2 = 0.75$ .

Tabel 4. 26 Performa Hasil Uji Aspek Harga, Keamanan, Layanan pada Skenario Uji 4

<i>Multistage Performa</i>	<i>Confusion Matrix</i>		
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Harga Positif	69%	65%	67%
Harga Negatif	59%	72%	65%
Keamanan Positif	78%	49%	60%
Keamanan Negatif	54%	72%	61%
Layanan Positif	75%	60%	67%
Layanan Negatif	45%	69%	54%

Tabel 4.19 menunjukkan hasil uji performa model yang diperoleh dari pengujian nilai hyperparameter C, nilai toleransi, dan nilai *degree*. Pengujian dilakukan dengan berbagai nilai parameter C, yaitu 0.01, 0.1, 1, 10, dan 100, dan nilai *degree* 2, 3, dan 4 untuk menentukan nilai yang menghasilkan performa optimal dalam nilai *precision*, *recall*, dan *F1-score*.



Gambar 4. 19 Visual Performa Klasifikasi *Single Stage Classification* dan *Multistage Classification*

Pada gambar 4.19 merupakan grafik yang menunjukkan hasil performa analisis sentimen untuk empat skenario berbeda, terdiri dari skenario 1 dan skenario 2 yang menggunakan *single stage classification*, skenario 3 dan skenario 4 menggunakan *multi-stage classification*. Grafik ini didasarkan pada nilai F1-Score

untuk setiap aspek, yaitu Harga, Keamanan, dan Layanan, dengan sentimen positif dan negatif.

Hasil analisis menunjukkan bahwa *multistage classification* memberikan performa yang lebih baik dibandingkan dengan *single stage classification*. Skenario 1 dan skenario 3 memiliki persamaan dalam penggunaan kernel linear, namun berbeda pada tipe klasifikasinya. Skenario 1 menggunakan pendekatan *single stage classification*, sedangkan skenario 3 menerapkan *multistage classification*. Hasil perbandingan menunjukkan bahwa skenario 3 menghasilkan performa yang lebih baik dibandingkan skenario 1. Berdasarkan hasil pengujian, terdapat perbedaan hasil klasifikasi antara kedua skenario tersebut. Pada kelas harga, terdapat selisih 1% untuk sentimen positif dan 11% untuk sentimen negatif. Pada kelas keamanan, selisihnya mencapai 14% untuk sentimen positif dan 8% untuk sentimen negatif. Sementara itu, pada kelas layanan, selisih yang diperoleh adalah 6% untuk sentimen positif dan 9% untuk sentimen negatif.

Selanjutnya, persamaan antara skenario 2 dan skenario 4 menggunakan kernel polinomial, namun berbeda dalam pendekatan klasifikasinya. Skenario 2 menggunakan *single stage classification*, sedangkan skenario 4 menerapkan *multistage classification*. Hasil perbandingan menunjukkan bahwa skenario 4 menghasilkan performa yang lebih baik dibandingkan skenario 2. Berdasarkan hasil pengujian, terdapat perbedaan hasil klasifikasi antara kedua skenario tersebut. Pada kelas harga, terdapat selisih 6% untuk sentimen positif dan 3% untuk sentimen negatif. Pada kelas keamanan, selisihnya mencapai 10% untuk sentimen positif dan 5% untuk sentimen negatif. Sementara itu, pada kelas layanan, selisih yang

diperoleh adalah 13% untuk sentimen positif dan 6% untuk sentimen negatif. Selisih-selisih hasil yang diperoleh dari pengujian skenario menunjukkan bahwa pendekatan *multistage classification* lebih efektif dalam menangani klasifikasi sentimen berbasis aspek.

Dengan membagi proses klasifikasi menjadi beberapa tahap, model dapat menganalisis data dalam konteks yang lebih spesifik, sehingga mampu menangani sentimen positif dan negatif secara lebih terarah. Secara umum, *multistage classification* mampu menghasilkan akurasi yang lebih tinggi dibandingkan *single stage classification*, serta memberikan fleksibilitas yang lebih besar dalam mengurangi ambiguitas antar kelas, terutama saat diterapkan pada data yang kompleks atau memiliki banyak kategori (G. Logewari, 2025).

### 4.3 Integrasi Islam

Dalam Islam, hubungan manusia dengan Allah SWT (*hablum minallah*) menjadi dasar utama dalam setiap aktivitas, termasuk jual beli. Jual beli harus dilakukan dengan kejujuran dan keadilan sebagai bentuk kepatuhan kepada Allah SWT. Dalam Al-Qur'an, Allah SWT telah menghalalkan praktek jual beli yang sesuai dengan ketentuan dan syari'atNya. Hal ini sesuai dengan firman Allah SWT dalam Q.S. al-Baqarah: 275:

...وَأَحَلَّ اللَّهُ الْبَيْعَ وَحَرَّمَ الرِّبَا...

“...Dan Allah telah menghalalkan jual beli dan mengharamkan riba” (Q.S. Al-Baqarah: 275)

Manfaat penelitian ini juga mengajarkan pentingnya prinsip amanah dalam memberikan pelayanan kepada pelanggan. Amanah dalam konteks jual beli

tercermin dalam kejujuran dan keadilan, seperti menyempurnakan takaran atau timbangan. Hal ini sesuai dengan firman Allah SWT dalam Q.S. Asy-Syu'ara: 181-183:

أَوْفُوا الْكَيْلَ وَلَا تَكُونُوا مِنَ الْمُخْسِرِينَ ﴿١٨١﴾ وَزِنُوا بِالْقِسْطِ الْمُسْتَقِيمِ ﴿١٨٢﴾ وَلَا تَبْخَسُوا النَّاسَ أَشْيَاءَهُمْ ﴿١٨٣﴾

“Sempurnakanlah takaran dan janganlah kamu termasuk orang-orang yang merugi, dan timbanglah dengan timbangan yang lurus. Janganlah kamu merajalela di muka bumi dengan membuat kerusakan” (Q.S. Asy-Syu'ara: 181-183).

Dalam tafsir Ibnu Katsir, ayat ini menegaskan bahwa dalam setiap transaksi, baik memberi maupun menerima, kita harus berlaku adil dan memastikan pembayaran dilakukan dengan jumlah yang tepat tanpa mengurangi hak orang lain. Hal ini juga mencakup prinsip jual beli yang jujur dan adil, yang tidak hanya sebagai bentuk ibadah untuk mendekatkan diri kepada Allah SWT, tetapi juga menjadi dasar bagi hubungan yang harmonis dalam bermuamalah. Dalam konteks penelitian ini, wawasan yang dihasilkan dapat memberikan manfaat bagi perusahaan layanan *Online Travel Agencies*, sehingga pemiliknya dapat memahami preferensi pengguna secara mendalam, memberikan pelayanan yang lebih baik, serta menjalankan prinsip keadilan dan kejujuran sebagai wujud amanah dalam pelayanan.

Sedangkan, hubungan antar manusia (*hablum minannas*) dalam jual beli juga memerlukan kejelasan dan persetujuan bersama untuk menghindari konflik. Islam mengajarkan pentingnya ijab kabul sebagai bentuk kesepakatan dalam transaksi. Hal ini ditegaskan dalam Surat An-Nisa ayat 29:

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا تَأْكُلُوا أَمْوَالِكُمْ بَيْنَكُمْ بِالْبَاطِلِ إِلَّا أَنْ تَكُونَ تِجَارَةً عَنْ تَرَاضٍ مِّنْكُمْ ۗ وَلَا تَقْتُلُوا أَنْفُسَكُمْ ۗ إِنَّ اللَّهَ كَانَ بِكُمْ رَحِيمًا ۝

*"Wahai orang-orang yang beriman! Janganlah kamu saling memakan harta sesamamu dengan jalan yang batil, kecuali dengan jalan perniagaan yang berlaku dengan suka sama suka di antara kamu" (Q.S. An-Nisa: 29).*

Dalam tafsir Ibnu Katsir, ayat ini menegaskan bahwa umat Islam dilarang saling memakan harta sesama mereka dengan cara yang batil, kecuali melalui perdagangan yang dilakukan dengan suka sama suka dan sesuai dengan prinsip keadilan. Transaksi harus dilakukan dengan kerelaan kedua belah pihak dan syarat sah jual beli adalah adanya ijab kabul, untuk memastikan bahwa hubungan antar manusia tetap terjaga dengan harmonis melalui keadilan dan kejujuran. Selain itu, Allah SWT juga melarang tindakan yang dapat merugikan diri sendiri, seperti membunuh diri, karena Allah SWT Maha Penyayang terhadap hamba-Nya.

Dalam konteks ini, penggunaan teknologi seperti aplikasi *Online Travel Agencies* dalam jual beli juga diperbolehkan selama tidak melanggar prinsip syariah. Hukum Islam mengizinkan transaksi berbasis teknologi asalkan memenuhi kriteria kejujuran, transparansi, dan tidak mengandung unsur haram seperti riba atau gharar (ketidakjelasan). Oleh karena itu, aplikasi *Online Travel Agencies* dapat menjadi sarana yang memudahkan transaksi selama memenuhi syarat syariat, seperti kejelasan informasi harga, layanan, dan adanya kesepakatan yang sah antara pihak-pihak yang terlibat. Pada hadits perdagangan (tjjarah), Nabi Muhammad SAW bersabda,

التَّاجِرُ الصَّدُوقُ الْأَمِينُ مَعَ النَّبِيِّ، وَالصِّدِّيقِينَ، وَالشُّهَدَاءِ

*"Seorang pedagang yang jujur dan amanah akan bersama para nabi dan orang-orang shiddiqin dan para syuhada." (HR. Tirmidzi 3/515 no. 1209).*

Hadis ini menunjukkan betapa tingginya kedudukan pedagang yang jujur dalam Islam, yang juga relevan dengan kejujuran dalam transaksi melalui aplikasi modern.

Selain itu, umat Muslim diajarkan untuk berbicara dengan kata-kata yang baik dan menghindari perkataan kasar atau menyakitkan. Ini karena kata-kata yang baik dapat membawa kebaikan dan mempererat hubungan antar sesama manusia, sementara kata-kata yang kasar dapat menyebabkan permusuhan dan keretakan dalam hubungan sosial. Allah SWT berfirman dalam Surah Al-Isra (17:53):

وَقُلْ لِعِبَادِي يَقُولُوا الَّتِي هِيَ أَحْسَنُ إِنَّ الشَّيْطَانَ يَنْزِعُ بَيْنَهُمْ إِنَّ الشَّيْطَانَ كَانَ لِلْإِنْسَانِ عَدُوًّا مُّبِينًا

*"Dan katakanlah kepada hamba-hamba-Ku, 'Hendaklah mereka mengucapkan perkataan yang lebih baik (benar).' Sesungguhnya syaitan itu menimbulkan perselisihan di antara mereka. Sesungguhnya syaitan adalah musuh yang nyata bagi manusia". (Q.S. Al-Isra: 53).*

Dalam tafsir Ibnu Katsir, Ayat ini menjelaskan bahwa Allah SWT memerintahkan kepada hamba dan Rasul-Nya yaitu Nabi Muhammad SAW agar memerintahkan kepada hamba-hamba Allah SWT yang beriman, hendaklah mereka dalam khotbah dan pembicaraannya mengucapkan kata-kata yang terbaik dan kalimat yang menyenangkan. Karena sesungguhnya jika mereka tidak melakukan hal ini, tentulah setan akan menimbulkan permusuhan di antara mereka dengan membakar emosi mereka, sehingga terjadilah pertengkaran dan peperangan serta keburukan. Hal ini selaras dengan dengan dengan tujuan penelitian yaitu untuk menganalisis ulasan pengguna dalam mengidentifikasi sentiment positif dan negatif.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil penelitian mengenai Analisis Sentimen Berbasis Aspek terhadap Ulasan Pengguna di Traveloka dengan metode *Support Vector Machine* dan optimasi *Sequential Minimal Optimization*, diperoleh beberapa kesimpulan sebagai berikut:

1. Metode *Support Vector Machine* yang dioptimasi dengan *Sequential Minimal Optimization* terbukti mampu mengklasifikasikan sentimen berbasis aspek dengan performa yang baik, mencapai akurasi hingga 92%. Keberhasilan ini turut didukung oleh penerapan teknik *preprocessing* dan ekstraksi fitur menggunakan TF-IDF, yang efektif dalam mengurangi bias terhadap kata-kata umum sehingga model dapat lebih fokus pada kata-kata bermakna dalam analisis sentimen. Berdasarkan hasil pengujian, pendekatan *multistage classification* menunjukkan peningkatan akurasi, presisi, recall, dan F1-score dibandingkan dengan *single stage classification*. Hal ini terlihat dari selisih performa pada berbagai skenario pengujian. Pada Skenario 1 dan Skenario 3 yang sama-sama menggunakan kernel linear, peningkatan performa tercatat antara 1% hingga 14%. Sementara itu, pada Skenario 2 dan Skenario 4 yang menggunakan kernel polinomial, selisih performa berkisar antara 3% hingga 13%. Perbedaan ini menunjukkan bahwa pendekatan *multistage classification* lebih efektif dalam menghasilkan keputusan klasifikasi yang

akurat, terutama dalam menangani variasi sentimen pada masing-masing aspek. Dari keempat skenario yang diuji, Skenario 3 memberikan hasil terbaik dengan nilai akurasi, presisi, recall, dan F1-score tertinggi dibandingkan skenario lainnya.

2. Penelitian ini memberikan wawasan bagi perusahaan *Online Travel Agencies* seperti Traveloka dalam memahami tingkat kepuasan pelanggan berdasarkan aspek-aspek spesifik, yaitu harga, layanan, dan keamanan. Dengan menerapkan pendekatan *multistage classification*, hasil analisis sentimen menunjukkan performa yang lebih baik dibandingkan dengan *single stage classification*, sehingga dapat meningkatkan efisiensi dalam mengidentifikasi sentimen pengguna secara lebih spesifik dan fokus.

## 5.2 Saran

Berdasarkan hasil penelitian ini, terdapat beberapa saran untuk penelitian selanjutnya:

1. Menggunakan dataset yang lebih besar dan lebih beragam untuk meningkatkan generalisasi model. Dalam penelitian ini, dataset yang digunakan berasal dari ulasan pengguna di *Google Play Store*. Untuk penelitian selanjutnya, dataset dapat diperluas dengan menambahkan ulasan dari platform lain seperti *media social* dari Traveloka.
2. Selain itu, penelitian juga dapat mencoba menganalisis berbagai aspek lain selain harga, layanan, dan keamanan, seperti kenyamanan, kecepatan pemrosesan transaksi, kebijakan *refund*, serta pengalaman pengguna dalam

navigasi aplikasi, yang dapat memberikan wawasan lebih mendalam mengenai faktor-faktor yang memengaruhi kepuasan pelanggan.

3. Menguji metode klasifikasi lain seperti *Deep Learning* untuk membandingkan performa dengan *Support Vector Machine*. Meskipun *Support Vector Machine* memberikan hasil yang baik, metode berbasis Deep Learning seperti LSTM dapat dieksplorasi untuk menangani kompleksitas bahasa dalam ulasan pengguna secara lebih mendalam. Selain itu, penelitian selanjutnya dapat mencoba metode lain seperti *Random Forest*, *Naïve Bayes*, atau *Decision Tree* untuk membandingkan efektivitas berbagai pendekatan klasifikasi.

## DAFTAR PUSTAKA

- Anita M., Ambarish. G. Mohapatra, S. K. Mohanty. (2023). Exploring Factors Influencing Customer Satisfaction in the Hotel Industry and Facilitating Decision-Making through an Analytical Hierarchy Process (AHP) Based Model. *1st International Conference on Circuits, Power and Intelligent Systems (CCPIS), Bhubaneswar, India*. Doi: 10.1109/CCPIS59145.2023.10291347.
- Deviyanto A, Didik Wahyudi M. (2018). Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor. *JISKA (Jurnal Informatika Sunan Kalijaga)*. Vol. 3, No. 1, MEI, 2018, Pp. 1–13 ISSN: 2527-5836 (print) | 2528-0074 (online).
- Elizabeth T. (2022). Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode *Support Vector Machine*. *Jurnal Teknik Informatika dan Sistem Informasi*. Vol. 9, No. 4, Hal. 3436-3442 ISSN 2407-4322 E- ISSN 2503-2933.
- G. Logewari, J. Deepika Roselind, K. TAMILASARI, V. NIVETHITHA. (2025). A Comprehensive Approach to Intrusion Detection in IoT Environments Using Hybrid Feature Selection and Multi-Stage Classification Techniques. *IEEE Access*. Volume 13, 2025.
- Handayani R. (2021). Optimasi Algoritma Support Vector Machine Untuk Analisis Sentimen Pada Ulasan Produk Tokopedia Menggunakan PSO. *Media Informatika*. Vol.20 No.2.
- Indria, Tio. (2021). Aplikasi E-Business Travel Berbasis Android. *Jurnal SANTI (Sistem Informasi dan Teknologi Informasi)*. Vol.1 No.1 Tahun 2021.
- Irfani F. (2020). Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine. *JBMI (Jurnal Bisnis, Manajemen, dan Informatika)*. JBMI Vol. 16 No. 3 Februari 2020.
- Irwan B. S., Y. Adrianto, A. D. Sensusiaty, D. P. Wulandari, I. K. E. Purnama. (2022), Ensemble Convolutional Neural Networks with Support Vector Machine for Epilepsy Classification Based on Multi-Sequence of Magnetic Resonance Images. *IEEE Access*, vol. 10, pp. 32034-32048, 2022, doi: 10.1109/ACCESS.2022.3159923.

- Iskandar J, Nataliani Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*. Vol. 5 No. 6 (2021) 1120 – 1126 ISSN Media Elektronik: 2580-0760.
- Ismet H. Mustaqim T. Purwitasari D. (2020). Aspect Based Sentiment Analysis of Product Review Using Memory Network. *Scientific Journal of Informatics*. p-ISSN 2407-7658 e-ISSN 2460-0040.
- Iswahyuniarto. (2023). Pengaruh Online Travel Agent (OTA) Traveloka Terhadap Tingkat Hunian Kamar Pada Hotel Calista Beach Baubau. *Jurnal Mahasiswa Pariwisata Dan Bisnis*. EISSN: 2828-3325. Vol 01 No 07, 2022: 1683 -1700.
- Iwan A, Fatrianto D. (2021). Prediksi Kinerja Akademik Mahasiswa Menggunakan Machine Learning dengan Sequential Minimal Optimization untuk Pengelola Program Studi. *Journal of Information Engineering and Educational Technology*. Volume 05 Nomor 02, 2021. (Journal Information Engineering and Educational Technology). ISSN: 2549-869X.
- Jordan Sihombing, Jeremia. (2022). Implementasi Algoritma Support Vector Machine Untuk Analisis Sentimen Aplikasi OLX di Playstore. *Journal of Informatics and Data Science (J-IDS)*. Vol. 1, No. 2, Desember 2022 ISSN 2964-0415.
- Hanna A., Bambang S., Lilik L. Suryasatriya T., Yohanes S. (2020). Pembelajaran Vektor Untuk Klasifikasi Data Pada Bidang. *SJME (Supremum Journal of Mathematics Education)*. Vol.4, No.2, July 2020, pp. 147-158 ISSN: 2548-8163 (online)| ISSN: 2549-3639 (print).
- Kundu A, Karim R, Ave A. (2023). Escalating SVM-Efficiency by Sequential QP LP and Slack Variable Analysis. *Journal of Computer Science*. Vol 19 (10): 1253.1262 DOI: 10.3844/jcssp.2023.1253.1262.
- Ma Y., Guo G., (2014). Support Vector Machines Applications. *Springer Science Business Media*
- Mahendro I., Abimanto D. (2023). Algoritma Support Vector Machine Untuk Menganalisa Kepuasan Mahasiswa Terhadap Pembelajaran Learning. *CV Pustaka Stimart AMNI Semarang*

- Musu W, Ibrahim A. (2021). Pengaruh Komposisi Data *Training* dan *Testing* terhadap Akurasi Algoritma C4.5. Prosiding Seminar Ilmiah Sistem Informasi Dan Teknologi Informasi. Vol. X, No. 1, Pebruari 2021: 186 – 195
- Muraina I. (2022). *Ideal Dataset Splitting Ratios in Machine Learning Algorithms: General Concerns for Data Scientists and Data Analysts. 7th International Mardin Artuklu Scientific Researches Conference*. ORCID ID: 0000-0002-9633-6080
- Naquitasia R, Hatta Fudholi D, Iswari L. (2020). Analisis Sentimen Berbasis Aspek Pada Wisata Halal Dengan Metode *Deep Learning*. Jurnal Teknoinfo. Volume 16, Nomor 2, Juli 2022, Page 156-164 ISSN: 1693-0010(Print), ISSN: 2615-224X(Online).
- Novita, Dien. (2021). Analisis Kepuasan Pengguna Aplikasi Traveloka Menggunakan Metode *Technology Acceptance Model* (TAM) Dan *End-User Computing Satisfaction* (EUCS). JTISI, Vol. 2, No. 1, April 2021: 22-37.
- Noviyanti I. (2019). Pengaruh Kualitas Produk Dan Kualitas Pelayanan Terhadap Kepuasan Pelanggan Pada CV Usaha Mandiri Jakarta Selatan. JIMF (Jurnal Ilmiah Manajemen Forkamma). ISSN (print): 2598-9545 & ISSN (online) : 2599-171X.
- Paulina, Wirdhayanti. (2020). Analisis Sentimen Berbasis Aspek Ulasan Pelanggan Terhadap Kertanegara Premium Guest House Menggunakan *Support Vector Machine*. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. Vol. 4, No. 4, April 2020, hlm. 1141-1149.
- Putro, Galuh Angger Adji. (2022). Pengaruh Digital Marketing Terhadap Keputusan Pembelian Dan Kepuasan Pelanggan Pada Jasa *Online Travel Agent*. Jurnal Mahasiswa Pariwisata dan Bisnis. Volume 4, No. 3 Maret 2023. p-ISSN 2722-7782 | e-ISSN 2722-5356.
- Radiena G, Nugroho A. (2023). Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi KAI Access Menggunakan Metode *Support Vector Machine*. Jurnal Pendidikan Teknologi Informasi (JUKANTI). Volume (6) No (1) April 2023– eISSN : 2621-1467.
- Rifqy O, Turjaman M, Budi I. (2022). Analisis Sentimen Berbasis Aspek *Marketing Mix* Terhadap Ulasan Aplikasi Dompot Digital (Studi Kasus: Aplikasi Linkaja Pada Twitter). Jurnal Darma Agung, [S.l.], v. 30, n. 2, p. 266-275, aug. 2022. ISSN 2654-3915.

- Rofiqi M, Fauzan A, Agustin A Saputra A. (2019). Implementasi *Term-Frequency Inverse Document Frequency* (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*. Vol. 1, No. 2, Desember 2019, Halaman 58-64.
- Rosyidin D, Prihatini A. (2020). Pengaruh *E-Service Quality* Dan *Brand Trust* Terhadap Keputusan Pembelian Pada Traveloka.Com (Studi Pada Konsumen Traveloka.Com Kota Semarang). *JIAB (Jurnal Ilmu Administrasi Bisnis)*. *Jurnal Ilmu Administrasi Bisnis*, vol. 9, no. 2, pp. 41-49.
- Tommy Rustandi S, Suhaedi D, Pemasari Y. (2023). Pemetaan *Hyperplane* Pada *Support Vector Machine*. *Bandung Conference Series: Mathematics*. <https://doi.org/10.29313/bcsm.v3i2.8187>.
- Wibisono, Indra Perdana. (2022). Efektifitas *Digital Concierge* Dalam Mendukung Operasional Hotel. *Wahana: Tridarma Perguruan Tinggi Efektifitas Digital Concierge Dalam Mendukung Operasional Hotel*. Volume 74 No 1 (2022).
- Yulita W, Dwi Nugroho E, Habib Algifari M. (2021). Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma *Naïve Bayes Classifier*. *JDMSI*, Vol. 2, No. 2, 2021, 1-9, ISSN: 2745-8458.

# LAMPIRAN

## Lampiran 1: Hasil Prediksi Sentimen Setiap Aspek

### 1. Aspek Harga

No.	Aktual	Prediksi	Ulasan
1.	Positif	Positif	aman dan murah
2.	Positif	Positif	sejauh ini ini yang paling jauh lebih baik terlebih paylater nya sangat membantu refund proses cepat saran saja harga nya jangan terlalu mahal nanti bakal tersaingi oleh kompetitor
3.	Positif	Positif	bolak balik samarinda surabaya atau surabaya samarinda slalu beli tiket di travelokabayar di idm slalu aman lebih murah bisa hemat banyak
4.	Negatif	Negatif	harga tiketnya gk ngotak nyari untung gk wajar
5.	Positif	Positif	terbaik dan termurah
6.	Positif	Positif	traveloka sangat memudahkan pemesanan cicilan paylaternya pun murah terima kasih traveloka
7.	Negatif	Negatif	sekarang kok traveloka untuk pemesanan hotel ada pajak dan biayanya besar bgd
8.	Negatif	Negatif	hati hati ya pengalaman pribadi juli pesan hotel dismd melalui aplikasi ini dari rmh jam malm bpp dpt harga booking sdh mau dekat hotel telp hotel utk konfirmasi ternyata kamar yg sdh dibayar full menurut resepsionis panik minta refaund lewat aplikasi tdk ada jawaban makin mlm makin panik uang segitu lumayan akhirnya datang ke hotelx minta refaund tdk bisa nunggu hari takut uang hilang minta kehotelnya minta solusi ada kmr yg kosong sharga rb terpaksa tdk sesuai pesanan
9.	Positif	Positif	alhamdulillah selama pake aplikasi ini selalu mulus dan lancar jalannya walaupun harga sdikit tinggi dibanding aa tp cukup terpercaya cancel penerbangan dari maskapai jg bisa refund trima kasih traveloka
10.	Positif	Positif	mudah murah n cepat
11.	Negatif	Negatif	voucher gagal terbit di hubungi cs nya lama sekali solusinya hampir jam sudah nunggu jam solusinya di upgrade kamar dengan extra biaya kesalahan dari sistem nya tapi konsumen yg dirugikan tidak ada kompensasi
12.	Positif	Positif	lbih murah dari aplikasi lain
13.	Positif	Positif	alhamdulillah memudahkan untuk bepergian harga terjangkau waktu berangkat dan tiba tepat waktumantap lah makasih
14.	Negatif	Negatif	aplikasi tolol area byk hotel rekomendasi dikit mahal doang
15.	Negatif	Negatif	kok harga hotel di traveloka sekarang pada mahal banget gak masuk akal harganya
16.	Positif	Positif	murah banget
17.	Negatif	Negatif	katanya epic ngibul aja kerjanya dibanding tiketcom ga ada embel epic harganya sama aja dikira kami bego apa ya
18.	Negatif	Negatif	gak jelas lah saya udah minta tolong dan tlvon sama orang traveloka nya buat bantu ganti nomor tlvon sama email sayamalah paylater saya di matiin kalo gak niat bantuin costumerngomong dong jangan bilang nya bisa saya udah senang minta tolong ganti nomor telepon baru sama email nya malah paylater saya di matiin ribet gak jelaass
19.	Negatif	Negatif	harga hotel di aplikasi ini paling mahal sejagad di banding aplikasi sebelah
20.	Negatif	Negatif	mending gak usah ada pilihan request khusus late check out kalau ujungujungnya pihak hotel masih minta tambahan biaya charge

### 2. Aspek Keamanan

No.	Aktual	Prediksi	Ulasan
1.	Negatif	Positif	jual macam apa di treveloka harganya selangit semua mending di aplikasi lain murah
2.	Negatif	Negatif	sebenarnya traveloka sangat membantu untuk pesan tiket pesawat dan hotel dengan mudah cukup modal hp saja tapi cukup menyebalkan rasanya ketika sedang memproses order tiket

			pesawat lalu isi biodata penumpang dll dan pas banget waktu tiket mau dicheckout mendadak ada pemberitahuan kalau tiket harganya sudah naik beda harganya bisa sampai ratusan ribu juga ini agak payah sih bisa lah diperbaiki sistemnya
3.	Positif	Positif	pembayaran nya sangat mudah dan harga juga cocok
4.	Positif	Positif	trimah kasih traveloka sellu memberikan harga tiket yg murahperjalanan jd menyenangkan
5.	Positif	Positif	memudahkan untk liburan dan bisnis cuma ada beberapa tiket hotel lebih mahal dari pada kita datang langsung
6.	Negatif	Negatif	saya suka heran sama algoritma harga di traveloka banyak sekali menampilkan harga murah ending nya saat mau di check out harga berlipat mahalnya belum lagi kalau mau bayar dengan paylater nambah ga jelas hitungan nya sebagai pengguna traveloka dari dulu saya sangat kecewa
7.	Negatif	Negatif	kenapa pas diliat di tanggal harganya beda sm yg pas mau di check out mana hampir semua tanggal berubah harganya jd lebih naik harusnya klo naik berubah jg di tanggal awal yg saya lihat
8.	Positif	Positif	semuanya sangat mudah dan nyamanlengkap dan juga murah banyak promo traveloka aplikasi terbaik
9.	Positif	Positif	alhamdulillah sangat membantu dalam perjalanan liburan semua ada di satu apkikasi biar tiket masuk ancol tersedia dan pastinya harga lebih murah
10.	Positif	Positif	seru ga akan kecewa harga ramah di kantong staf juga ramah
11.	Negatif	Positif	ditungguin minggu harga tiket pesawatnya naik terus pas udh beli dpt hari lgsg turun itu mksdnya apa emng lbih murah dri aplikasi lainnya tpi untuk grafik harganya hancur berasa nyesel beli di traveloka gk lagi dah mending yg lain
12.	Negatif	Positif	hati dgn iklan easy refund di aplikasi inikatanya bisa refund dgn alasan apapunbahkan ketika memilih easy refund dikenakan biaya jaminan bisa refund hampir jt tapi kenyataan nya nol besar boro bisa refund yang ada di refund biaya jaminan refund itupun gak full masa refund jt di kembalikan cuma jt yang bener aja traveloka
13.	Positif	Positif	biaya layanan terlalu besar skrng banyak aplikator yg menawarkan pembelian tiket murah klo ada pungutan tanpa ada persetujuan dr konsumen so pasti pelanggan akan berpindah menggunakan aplikasi lain semoga bs menjadi pertimbangan terima kasih
14.	Negatif	Negatif	meskipun demikian kami sangat menghargai minat anda terhadap paylater dan karena itu kami ingin mengundang anda untuk kembali mendaftar paylater setelah bulan pada pendaftaran berikutnya mohon ikuti petunjuk di bawah ini kami mohon maaf atas segala ketidaknyamanan yang ditimbulkan dan terima kasih atas minat anda terhadap paylater
15.	Negatif	Positif	ribet tampilan aplikasi jelek membingungkan harga semua layanan kali lipat dari harga penyedia hotel tiket aplikasi mark up biaya
16.	Positif	Positif	setiap saya menggunakan jasa trapeloka untuk beli tiket pesawat pas selalu mahal
17.	Positif	Positif	aku selalu beli tiket pesawathotelbusdll di traveloka karna murah bangetklau ad kendala lgsung aku tanyain dari ig langsung gercep bgt ditanggapinsuka bgt de jdi klau lgi stres bukany langsung traveloka buat healing
18.	Positif	Positif	salah satu kelebihan traveloka menurut saya adalah harga yg ditampilkan sudah termasuk semuanya aplikasi online booking lain kadang menampilkan harga yg belum termasuk pajak service charge atau surcharge pada saat pencarianjadi bagus buat kalian yang mau jalan jalan bisa download aplikasi traveloka bisa
19.	Positif	Positif	aplikasi best sih sekarang mau pergi liburan tinggal pake traveloka aplikasi recommended yang mau healing pake traveloka ya guys di jamin murah
20.	Positif	Positif	sangat menarikuntuk pembelian tiket wahana sangat murah harga jadi lebih terjangkauharga tiket pesawat juga banyak promo nyaterima kasih traveloka

### 3. Aspek Layanan

No.	Aktual	Prediksi	Ulasan
1.	Negatif	Positif	perusahaan gila makan darah kalau mau reschedule tiket lama di anggap nol nilainya di arahkan harga tiket baru dengan harga lebih mahal dari harga normalnya pada tiket lama diambil di jual lagi dengan harga ekonomi aturan gila di negara dan perusahaan konoha lebih baik pakai jasa aplikasi perusahaan asing dan transportasi asing daripada lokal biar makin ancur ni lokalnya
2.	Positif	Positif	selama ini membantu sih aplikasi traveloka ini tapi saat pembaharuan ini saya agak kecewa kenapa beli tiket pesawat diwajibkan tambahan bayar asuransi lagi kan jadi harganya naik dan tidak sesuai pada saat pengecekan awal
3.	Negatif	Negatif	fungsi tida tidak termasuk sarapan oke wifi oke tidak boleh merokok dengan harga yang beda gk fungsi percuma ambil yg lebih mahal tapi resepsionis gk tau
4.	Negatif	Negatif	aplikasi cukup interaktif layanan pembayaran jg lumayan mudah namun ada kendala utk pembayaran di tempat sangat ribet hrs menampilkan kartu kredit dll kdg gak bisa terupload gagal trus harga hotel terlalu mahal terutama pada saat liburan jauh dak msuk akal dg fasilitas hotel yg kdg ala kadarnya kadang bed g sesuai dg kondisi di etalase size king ternyata sampai di hotel size dan kurang nya promo voucher diskon pdahal saya seringkali pesan d trvlk bertahun gt aja
5.	Negatif	Negatif	admin atas nama arum sangat pandai dalam berbelit dan membuat nama traveloka jadi jelek semoga kedepannya traveloka lebih pandai lagi dalam mencari sdm costumer service yang attitudenya bagus dan jujur terimakasih
6.	Positif	Positif	mudah murah n cepat
7.	Positif	Positif	wajib dan cukup traveloka ada di smartphone kalian gaessss simplemurahnyaman dan yg paling penting nih cs nya pun gercep tidak seperti aplikasi lain tengkyu
8.	Negatif	Negatif	sangat buruk utk di ungkapkan begitu mau di pesan seketika trjadi perubahan harga yg signifikan sangat kecewa
9.	Positif	Positif	oke terimakasih traveloka atas responnya atas kendala saya full senyum dah sekarang dengan aplikasi inikeren dah atas responnya sangat menghargai pengguna dan pelanggan setianya
10.	Positif	Positif	jatuhnya malah mahal katanya dapat kupon ternyata tidak bisa dipakai karena kuota kupon habis
11.	Positif	Positif	waw karena adanya traveloka sekarang kalo mau liburan tinggal cek in hotel yang murah udah murah aesthetic dan enak banget untuk berlibur bareng keluarga
12.	Negatif	Neagatif	gak ada kebijakan buat pengguna baru tentang salah masukan nama pada pembelian tiket pesawat sangat merugikan konsumen kalau di refund dikenakan biaya
13.		Positif	untuk pemesan hotel memang sedikit lebih mahal dari aplikasi sebelah namun disini tidak perlu menghitung berapa kalkulasi akhir untuk booking hotel kecuali tambahan administrasi saja sebab tidak ada biaya tambahan pajak dll pernah beli tiket bus dan tidak jadi terpakai karena kesalahan informasi pada titik loket penjemputan juga uang di refund dengan mudah bahkan pernah tidak salah buat kode akhir pembayaran bisa selesai dibantu admin dengan keterangan sebab saya sudah jadi pelanggan
14.	Negatif	Positif	hitaveloka menggap saya sudah cancel dan menerima refund full paylater dari akomodasi tetapi tetap tertagih bunga cicilan dari pihak hotel saja saya sama sekali tidak ada biaya pembatalan sebelum nya saya cancel booking tidak pernah ada biaya bunga cicilan seperti ini
15.	Positif	Positif	lengkap dan murah rekomended bngt buat yg gk suka ribet
16.	Positif	Positif	traveloka pesan hotel lebih murah dan mudah
17.	Negatif	Positif	mau beli tiket pesawat mau checkout tiba harga telah naik astagaaa apaan si nyebel bgt
18.	Negatif	Positif	kenapa pas mau pembayaran harga pesawat tiba berubah naik padahal di cek harga masih belum naik uda refresh berkali jg harga msh promo tp tiap mau pembayaran harga berubah naik gimana nih
19.	Positif	Positif	apk ini bener membantu banget buat kita kaum mager cuma duduk santai d rumah sudah bisa booking tiket lewat apk ini bukan

			hanya tiket tp kita juga bisa booking hotel lewat sini dan tentunya harga nya lebih terjangkau sejauh ini mantap sih
20.	Positif	Positif	sangat membantu mencari tiket yg murah

## Lampiran 2: Perhitungan Manual Proses Implementasi *Metode Support Vector Machine*

Dari data table 3.9, diperlukan melabeli dokumen-dokumen berdasarkan aspek

harga:

- Dokumen D1: Berkaitan dengan harga → Label: +1
- Dokumen D2: Tidak berkaitan dengan harga → Label: -1
- Dokumen D3: Tidak berkaitan dengan harga → Label: -1

Maka label untuk setiap dokumen adalah:

- D1: +1 (positif untuk aspek harga)
- D2: -1 (negatif, karena tidak terkait aspek harga)
- D3: -1 (negatif, karena tidak terkait aspek harga)

Vektor fitur untuk setiap dokumen berasal dari nilai TF-IDF untuk kata-kata:

- D1: [0.176,0,0.176]
- D2: [0.176,0.477,0]
- D3: [0,0,0.176]

Dalam kasus ini, kita akan melatih *Support Vector Machine* untuk aspek harga.

Persamaan dari hyperplane dalam bentuk linear untuk model *Support Vector Machine* adalah:

$$w \cdot x + b = 0 \quad (3.4)$$

Keterangan:

- $w$  adalah vektor bobot,
- $x$  adalah vektor fitur (dalam kasus ini, nilai TF-IDF dari dokumen),
- $b$  adalah bias.

*Support Vector Machine* mencari  $w$  dan  $b$  yang memaksimalkan margin antara dokumen yang diberi label +1 dan -1. Setelah itu, iterasi pembaruan *Support Vector Machine* dengan Inisialisasi Awal  $w$  dan  $b$  menggunakan persamaan 3.5:

$$y_1 (w \cdot x_1 + b) \quad (3.5)$$

Keterangan:

- $y_1$  : Label kelas data (misalnya +1 atau -1)
- $w$  : Vektor bobot
- $x$  : Vektor fitur (Nilai TF-IDF dari dokumen)
- $b$  : Bias.

Berikut adalah perhitungan iterasi pembaruan *Support Vector Machine* dengan persamaan 3.5:

- Menghitung Dokumen D1 (Label: +1)

Vektor fitur D1: [0.176,0,0.176], Label: +1.

$$\begin{aligned} y_1 (w \cdot x_1 + b) &= 1 \times ([0.5, 0.5, 0.5] \cdot [0.176, 0, 0.176] + 0.5) \\ &= 1 * ((0.5 \times 0.176 + 0.5 \times 0 + 0.5 \times 0.176) + 0.5) \\ &= 1 * (0.088 + 0.088 + 0.5) = 1 \times 0.676 = 0.676 \end{aligned}$$

Karena 0.676 lebih kecil dari 1, diperlukan memperbarui  $w$  dan  $b$ . Kemudian dilakukan perhitungan pembaruan 3.6:

$$w = w + \eta \times y_1 \times x_1 \quad (3.6)$$

Berikut adalah perhitungan pembaruan nilai  $w$  dan  $b$  dengan inisialisasi nilai  $b=0.5$  dan  $\eta=1$ .

$$\begin{aligned} w &= w + \eta \times y_1 \times x_1 = [0.5, 0.5, 0.5] + 1 \times [0.176, 0, 0.176] \\ &= [0.676, 0.5, 0.676] \\ b &= b + \eta * y_1 \quad (3.7) \\ b &= b + \eta * y_1 = 0.5 + 1 = 0.5 \end{aligned}$$

- Menghitung Dokumen D2 (Label: -1)

Vektor fitur D2: [0.176,0.477,0], Label: -1.

$$\begin{aligned} y_2(w * x_2 + b) &= -1 \times ([0.676, 0.5, 0.676] [0.176, 0.477, 0] + 1.5) \\ &= -1 \times ((0.676 \times 0.176 + 0.5 \times 0.477 + 0.676 \times 0) + 1.5) \\ &= -1 \times (0.119 + 0.239 + 0 + 1.5) = -1 \times 1.858 = -1.858 \end{aligned}$$

Karena  $-1.858 \leq 1$ , maka tidak perlu memperbarui  $w$  dan  $b$ .

- Menghitung Dokumen D3 (Label: -1)

Vektor fitur D3: [0,0,0.176], Label: -1.

$$\begin{aligned} y_3(w * x_3 + b) &= -1 \times ([0.676, 0.5, 0.676] [0,0,0.176] + 1.5) \\ &= -1 \times (0.676 \times 0.176 + 0.5 \times 0 + 0.676 \times 0.176 + 1.5) \\ &= -1 \times (0.119 + 0 + 0.119 + 1.5) = -1 \times 1.738 = -1.738 \end{aligned}$$

Karena  $-1.738 \leq 1$ , maka tidak perlu memperbarui  $w$  dan  $b$ .

Setelah iterasi, hasil pembaruan bobot dan bias adalah dengan persamaan 3.8:

$$w = \sum \alpha_i y_i x_i \quad (3.8)$$

Keterangan:

- $w$  : Vektor bobot akhir setelah semua alpha diperbarui.
- $\alpha_i$  : Pengali Lagrange yang dioptimalkan untuk data latih ke-i.
- $y_i$  : Label kelas untuk data latih ke-i.
- $x_i$  : Vektor fitur untuk data latih ke-i.

Berikut perhitungan dari pembaruan bobot dan bias:

$$w = [0.676, 0.5, 0.676]$$

$$b = 1.5$$

Untuk mengoptimalkan model *Support Vector Machine* menggunakan *Sequential Minimal Optimization*, bisa menggunakan nilai yang sudah diketahui dan berikut ini merupakan langkah-langkah perhitungan manual.

Diketahui nilai:

- $x_1 = [0.176, 0, 0.176]$  (D1, label +1)
- $x_2 = [0.176, 0.477, 0]$  (D2, label -1)
- $x_3 = [0, 0, 0.176]$  (D3, label -1)

Dan inialisasi awal  $\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0, b = 1.5, w = [0.676, 0.5, 0.676]$ . Langkah pertama yaitu menghitung error (E) untuk setiap data dengan persamaan 3.9:

$$E_i = f(x_i) - y_i \quad (3.9)$$

$$f(x_i) = w * x_i + b \quad (3.10)$$

Keterangan:

- $w$  : Vektor bobot akhir setelah semua alpha diperbarui.
- $y_i$  : Label kelas untuk data latih ke-i.
- $x_i$  : Vektor fitur untuk data latih ke-i.
- $E_i$  : Nilai error untuk data  $x_i$

➤ Untuk  $x_1$ :

$$f(x_1) = 0.676 \times 0.176 + 0.5 \times 0 + 0.676 \times 0.176 + 1.5 = 1.737152$$

$$E_1 = 1.737152 - 1 = 0.737152$$

➤ Untuk  $x_2$ :

$$f(x_2) = 0.676 \times 0.176 + 0.5 \times 0.477 + 0.676 \times 0 + 1.5 = 1.866152$$

$$E_2 = 1.866152 + 1 = 0.866152$$

➤ Untuk  $x_3$ :

$$f(x_3) = 0.676 \times 0 + 0.5 \times 0 + 0.676 \times 0.176 + 1.5 = 1.596576$$

$$E_3 = 1.596576 + 1 = -0.403424$$

Langkah selanjutnya adalah memilih nilai  $i$  dan nilai  $j$ . Nilainya di aumsikan  $i =$

1 (D1) dan  $j = 2$  (D2). Lalu menghitung nilai eta ( $\eta$ ) dengan persamaan 3.11:

$$\eta = 2x_i * x_j - x_i * x_i - x_j * x_j \quad (3.11)$$

Keterangan:

- $\eta$  : Sebuah nilai yang digunakan untuk menentukan seberapa besar perubahan yang harus diterapkan pada  $\alpha_j$ .
- $x_j$  dan  $x_i$  : Vektor fitur untuk data latih ke-i dan j.

Namun perlu menghitung nilai  $x_i$  dan nilai  $x_2$ :

$$x_1 * x_2 = 0.176 \times 0.176 + 0 \times 0.477 + 0.176 \times 0 = 0.031776$$

$$x_1 * x_1 = 0.176 \times 0.176 + 0 + 0.176 \times 0.176 = 0.031776$$

$$x_2 * x_2 = 0.1762 + 0.4772 + 0 = 0.031776 + 0.227529 = 0.259305$$

Setelah menghitung nilai  $x_i$  dan nilai  $x_2$ , selanjutnya adalah menghitung nilai  $\eta$ :

$$\begin{aligned}\eta &= 2 \cdot 0.031776 - 0.031776 - 0.259305 \\ &= 0.063552 - 0.031776 - 0.259305 \\ &= -0.227529\end{aligned}$$

Setelah menghitung nilai  $\eta$ , langkah selanjutnya adalah menghitung nilai  $\alpha_j$

dengan persamaan 3.12:

$$\alpha_j = \alpha_j - \frac{y_j(E_i - E_j)}{\eta} \quad (3.12)$$

Keterangan:

- $\alpha_j$  : Nilai alpha j yang baru setelah pembaruan.
- $E_i$  dan  $E_j$  : Kesalahan prediksi untuk data latih ke-i dan j.

Berikutnya adalah perhitungan manualnya:

$$\begin{aligned}&= 0 - \frac{-1 * (0.737152 - 0.866152)}{-0.227529} \\ &= 0 - \frac{-1 * (-0.129)}{-0.227529} = 0 - 0.5663 \\ &= 0.5663\end{aligned}$$

Selanjutnya adalah menghitung nilai  $\alpha_i$  dengan persamaan 3.13:

$$\alpha_i = \alpha_i + y_i y_j (\alpha_j - \alpha_j^{old}) \quad (3.13)$$

Berikut adalah nilai  $\alpha_i$  yang dihitung menggunakan persamaan 3.13:

$$\begin{aligned}&= 0 + 1 * (-1) * (0.5663 - 0) \\ &= 0 - 0.5663\end{aligned}$$

$$= -0.5663$$

Langkah selanjutnya adalah menghitung bias dengan persamaan 3.14:

$$b_1 = b - E_i - y_i(\alpha_i - \alpha_i^{old})x_i * x_i - y_j(\alpha_j - \alpha_i^{old})x_i * x_j \quad (3.14)$$

$$b_2 = b - E_j - y_i(\alpha_i - \alpha_i^{old})x_i * x_i - y_j(\alpha_j - \alpha_i^{old})x_i * x_j \quad (3.15)$$

Keterangan:

- $b_1$  : Bias yang dihitung dari data latih ke-i.
- $b_2$  : Bias yang dihitung dari data latih ke-j.
- $E_j$  dan  $E_i$  : Kesalahan prediksi untuk data latih ke-I dan ke-j.
- $\alpha_i$  dan  $\alpha_j$  : Nilai alpha setelah pembaruan.

Kemudian menghitung nilai  $b_1$  dengan persamaan 3.14:

$$\begin{aligned} b_1 &= 1.5 - 0.737152 - 1 * (-0.5663 - 0) * 0.031776 - (-1) (0.5663 - 0).0.031776 \\ &= 1.5 - 0.737152 + 0.018058 + 0.018058 \approx 1.5 - 0.737152 + 0.036116 \approx 0.799964 \end{aligned}$$

Setelah menghitung nilai  $b_1$ , dilanjutkan menghitung nilai  $b_2$  dengan persamaan

3.15:

$$\begin{aligned} b_2 &= 1.5 - 0.866152 - 1 * (-0.5663 - 0) * 0.477 - (-1) * (0.5663 - 0).0.259305 \\ &= 1.5 - 0.866152 + 0.270735 + 0.14745 \approx 1.5 - 0.866152 + 0.418185 \approx 1.052033 \end{aligned}$$

Langkah selanjutnya adalah menghitung rata-rata nilai  $b$  dengan persamaan 3.16:

$$\begin{aligned} b &= \frac{b_1 + b_2}{2} \quad (3.16) \\ b &= \frac{b_1 + b_2}{2} = \frac{0.799964 + 1.052033}{2} = 0.926999 \end{aligned}$$

Setelah menghitung rata-rata bias, dilanjutkan menghitung nilai  $w$  dengan persamaan 3.17:

$$w = \sum \alpha_i y_i x_i \quad (3.17)$$

Keterangan:

- $w$  : Vektor bobot akhir setelah semua alpha diperbarui.
- $\alpha_i$  : Pengali Lagrange yang dioptimalkan untuk data latih ke-i.
- $y_i$  : Label kelas untuk data latih ke-i.
- $x_i$  : Vektor fitur untuk data latih ke-i.

$$\begin{aligned}
w &= (-0.5663)(1) ([0.176, 0, 0.176]) + (0.5663)(-1)([0.176, 0.477, 0]) \\
&= [-0.5663 \times 0.176, 0, -0.5663 \times 0.176] + [-0.5663 \times 0.176, -0.5663 \times 0.477, 0] \\
&= [-0.099835, 0, -0.099835] + [-0.099835, -0.27057, 0] \\
&= [-0.19967, -0.27057, -0.099835]
\end{aligned}$$

Dengan  $w$  dan  $b$  yang baru, dapat menghitung prediksi. Rumus persamaan untuk prediksi adalah 3.18:

$$prediksi = sign(w * x + b) \quad (3.18)$$

Keterangan:

- $w * x$  : Dot product antara vektor bobot yang telah dioptimalkan dan vektor fitur dari data yang akan diprediksi
- $b$  : Bias yang telah diperbarui.
- $sign()$  : Fungsi sign yang menghasilkan +1 jika nilainya positif

(kelas 1), dan negative (-1).

Diketahui nilai  $w = [-0.19967, -0.27057, -0.099835]$ ,  $b \approx 0.926999$ , Vektor fitur  $x = [0.176, 0, 0.176]$ .

$$w \cdot x = (-0.19967 \times 0.176) + (-0.27057 \times 0) + (-0.099835 \times 0.176)$$

$$w \cdot x = -0.03514352 + 0 - 0.01759976 = -0.05274328$$

$$w \cdot x + b = -0.05274328 + 0.926999 \approx 0.87425572$$

Karena menggunakan fungsi sign:

$$prediksi = sign(0.87425572)$$

Karena nilai  $0.87425572 > 0$ , maka: prediksi = +1. Prediksi untuk vektor fitur  $x = [0.176, 0, 0.176]$  adalah +1, yang berarti sentimen untuk aspek yang diuji adalah positif.

### Lampiran 3: Perhitungan Manual Proses *Confusion Matrix* 3 Kelas

Pada table 3.11, untuk mengevaluasi confusion matrix dengan 3 kelas berdasarkan aspek, nilai TP, TN, FP, dan FN dihitung untuk setiap kelas secara terpisah. Berikut adalah cara perhitungan parameter confusion matrix untuk kelas ‘Keamanan’.

$$Akurasi = \frac{TP}{Jumlah\ Data} = \frac{8}{32}$$

$$Presisi_{layanan} = \frac{TP}{TP + FP} = \frac{3}{3 + 8} = \frac{3}{11} = 0,273$$

$$Presisi_{harga} = \frac{TP}{TP + FP} = \frac{2}{2 + 9} = \frac{2}{11} = 0,182$$

$$Presisi_{keamanan} = \frac{TP}{TP + FP} = \frac{3}{3 + 7} = \frac{3}{10} = 0,3$$

$$Rata-rata\ Presisi = \frac{Presisi\ 1 + Presisi\ 2 + Presisi\ 3}{Jumlah\ Kelas} = \frac{0,273 + 0,182 + 0,3}{3} = 0,252$$

$$Recall_{layanan} = \frac{TP}{TP + FN} = \frac{3}{3 + 4} = \frac{3}{7} = 0,429$$

$$Recall_{harga} = \frac{TP}{TP + FN} = \frac{2}{2 + 2} = \frac{2}{4} = 0,5$$

$$Recall_{keamanan} = \frac{TP}{TP + FN} = \frac{3}{3 + 5} = \frac{3}{8} = 0,375$$

$$Rata-rata\ Recall = \frac{Recall\ 1 + Recall\ 2 + Recall\ 3}{Jumlah\ Kelas} = \frac{0,429 + 0,5 + 0,375}{3} = 0,435$$

$$F1-score = 2x \frac{Presisi * Recall}{Presisi + Recall} = 2x \frac{0,252 * 0,435}{0,252 + 0,435} = \frac{0,21924}{0,687} = 0,319\ \text{atau}\ 31,9\%$$