



SKRIPSI

**ANALISIS SENTIMEN BERBASIS ASPEK
TERHADAP PERSEPSI PENGGUNA APLIKASI
DOMPET DIGITAL OVO MENGGUNAKAN
*SUPPORT VECTOR MACHINE (SVM)***

EKA FAHIRA APRILIA

NPM 21082010218

DOSEN PEMBIMBING

Amalia Anjani Arifiyanti, S.Kom., M.Kom

Nambi Sembilu, S.Kom., M.Kom

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
SURABAYA
2025**



SKRIPSI

ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP PERSEPSI PENGGUNA APLIKASI DOMPET DIGITAL OVO MENGGUNAKAN *SUPPORT VECTOR MACHINE (SVM)*

EKA FAHIRA APRILIA

NPM 21082010218

DOSEN PEMBIMBING

Amalia Anjani Arifiyanti, S.Kom., M.Kom

Nambi Sembilu, S.Kom., M.Kom

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
SURABAYA
2025**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP PERSEPSI PENGGUNA APLIKASI DOMPET DIGITAL OVO MENGUNAKAN *SUPPORT VECTOR MACHINE* (SVM)

Oleh:
EKA FAHIRA APRILIA
NPM. 21082010218

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 9 Mei 2025.

Amalia Anjani Arifiyanti, S.Kom., M.Kom
NIP. 19920812 2018032 001

..... (Pembimbing I)

Nambi Sembilu, S.Kom., M.Kom
NIP. 19900516 2024061 003

..... (Pembimbing II)

Eka Dyar Wahyuni, S.Kom., M.Kom
NIP. 19841201 2021212 005

..... (Ketua Penguji)

Dhian Satria Yudha Kartika, S.Kom., M.Kom
NPT. 20119860522249

..... (Anggota Penguji II)

Seftin Fitri Ana Wati, S.Kom., M.Kom
NPT. 212199 10 320267

..... (Anggota Penguji III)

Mengetahui,
Dekan Fakultas Ilmu Komputer

Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP
PERSEPSI PENGGUNA APLIKASI DOMPET DIGITAL OVO
MENGUNAKAN *SUPPORT VECTOR MACHINE* (SVM)**

Oleh:

EKA FAHIRA APRILIA

NPM. 21082010218

Menyetujui,

Koordinator Program Studi Sistem Informasi

Fakultas Ilmu Komputer

Agung Brastama Putra, S.Kom., M.Kom

NIP. 19851124 2021211 003

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertandatangan di bawah ini:

Nama Mahasiswa : Eka Fahira Aprilia
NPM : 21082010218
Program : Sarjana (S1)
Program Studi : Sistem Informasi
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 4 Juni 2025
Yang Membuat Pernyataan,



Eka Fahira Aprilia
NPM. 21082010218

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Eka Fahira Aprilia / 21082010218
Judul Skripsi : Analisis Sentimen Berbasis Aspek Terhadap Persepsi Pengguna Aplikasi Dompet Digital OVO Menggunakan *Support Vector Machine* (SVM)
Dosen Pembimbing : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom
2. Nambi Sembilu, S.Kom., M.Kom

Kemajuan teknologi finansial di Indonesia mendorong penggunaan dompet digital sebagai alternatif pembayaran yang praktis. Salah satu aplikasi dompet digital yang populer di Indonesia adalah OVO, yang meskipun memiliki jumlah pengguna aktif yang tinggi, tetap menerima berbagai persepsi negatif dari penggunaannya. Penelitian ini bertujuan untuk mengelompokkan topik ulasan pengguna OVO menggunakan metode *Latent Dirichlet Allocation* (LDA) untuk menemukan aspek utama, serta membangun model analisis sentimen berbasis aspek menggunakan algoritma *Support Vector Machine* (SVM) yang diimplementasikan dalam sistem berbasis web. Data ulasan dikumpulkan dari aplikasi OVO versi 3.115 hingga 3.119 yang tersedia di Google Play Store dan Apple Store. Hasil pemodelan LDA mengidentifikasi empat aspek utama, yaitu *Transaction Efficiency*, *User Experience*, *Account Access and Registration*, serta *Balance and Charges*. Namun, tingkat kesesuaian pelabelan aspek otomatis LDA terhadap pelabelan manual hanya mencapai 11.46% dan meningkat menjadi 40.60% setelah perbaikan *keyword*, yang menunjukkan keterbatasan LDA dalam memahami konteks ulasan. Model klasifikasi terbaik menggunakan SVM dengan kernel RBF tanpa *resampling*, menghasilkan *macro average F1-Score* sebesar 0.715 dan *hamming loss* sebesar 0.099 pada data pemodelan. Model tersebut diuji kembali menggunakan data validasi sistem dan menghasilkan *F1-Score* sebesar 0.794 serta *hamming loss* sebesar 0.102. Penerapan *oversampling* ML-SMOTE tidak memberikan peningkatan signifikan. Sistem berbasis web yang dikembangkan mampu memprediksi aspek dan sentimen dari ulasan pengguna secara interaktif melalui input teks maupun *file* CSV, serta menyajikan hasil prediksi dalam bentuk visual yang informatif.

Kata Kunci: Analisis Sentimen Berbasis Aspek, *Latent Dirichlet Allocation*, *Support Vector Machine*, OVO, Ulasan Pengguna.

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Eka Fahira Aprilia / 21082010218
Thesis Title : Aspect-Based Sentiment Analysis on User Perceptions of the OVO Digital Wallet Application Using Support Vector Machine (SVM)
Dosen Pembimbing : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom
2. Nambi Sembilu, S.Kom., M.Kom

ABSTRACT

The advancement of financial technology in Indonesia has driven the widespread adoption of digital wallets as a practical payment alternative. One of the most popular digital wallet applications in Indonesia is OVO, which, despite having a high number of active users, continues to receive various negative perceptions from its users. This study aims to categorize topics in OVO user reviews using the Latent Dirichlet Allocation (LDA) method to identify key aspects, and develop an aspect-based sentiment analysis model using the Support Vector Machine (SVM) algorithm, implemented within a web-based system. Review data were collected from OVO application versions 3.115 to 3.119 available on the Google Play Store and Apple Store. The LDA modeling results identified four main aspects: Transaction Efficiency, User Experience, Account Access and Registration, and Balance and Charges. However, the accuracy of automatic aspect labeling by LDA compared to manual labeling was only 11.46%, increasing to 40.60% after keyword refinement, indicating LDA's limitations in understanding review context. The best classification model used an SVM with an RBF kernel without resampling, achieving a macro average F1-Score of 0.715 and a Hamming Loss of 0.099 on the modeling dataset. The model was further evaluated using system validation data, resulting in an F1-Score of 0.794 and a Hamming Loss of 0.102. The application of ML-SMOTE oversampling did not yield significant improvements. The developed web-based system can predict aspects and sentiments of user reviews interactively via text input or CSV file, presenting the prediction results through informative visualizations.

Keywords: Aspect-Based Sentiment Analysis, Latent Dirichlet Allocation, Support Vector Machine, OVO, User Reviews.

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, karunia, dan petunjuk-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Analisis Sentimen Berbasis Aspek Terhadap Persepsi Pengguna Aplikasi Dompot Digital OVO Menggunakan *Support Vector Machine* (SVM)”** dengan sebaik-baiknya. Skripsi ini disusun sebagai bagian dari pemenuhan salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Penyusunan skripsi ini tidak lepas dari bantuan, dukungan, doa, serta motivasi dari berbagai pihak yang telah memberikan semangat dan inspirasi di setiap prosesnya. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua tercinta, atas segala cinta, doa, dan semangat tiada henti yang selalu mengiringi setiap langkah penulis. Terima kasih telah menjadi sumber kekuatan dan motivasi terbesar dalam menyelesaikan studi ini.
2. Keluarga besar penulis, khususnya adik-adik tercinta, yang senantiasa memberikan semangat, keceriaan, dan menjadi tempat pulang yang penuh kehangatan di tengah proses yang penuh tantangan.
3. Ibu Amalia Anjani Arifiyanti, S.Kom., M.Kom, selaku dosen pembimbing I, yang telah dengan penuh kesabaran membimbing, memberikan masukan, serta arahan yang sangat berharga sejak awal penyusunan proposal hingga akhir laporan skripsi ini.
4. Bapak Nambi Sembilu, S.Kom., M.Kom, selaku dosen pembimbing II, atas segala bimbingan, dorongan, serta ilmu yang telah membantu penulis dalam menyelesaikan tugas akhir skripsi ini dengan baik.
5. Bapak Agung Brastama Putra, S.Kom., M.Kom, selaku Koordinator Program Studi Sistem Informasi, atas segala perhatian, dukungan, serta arahnya selama studi berlangsung.

6. Seluruh dosen Program Studi Sistem Informasi UPN “Veteran” Jawa Timur, atas ilmu, wawasan, serta inspirasi yang telah diberikan kepada penulis, yang menjadi bekal berharga dalam proses akademik.
7. Farrel Farhan, yang senantiasa kebersamai penulis, menjadi sumber semangat dan dukungan kepada penulis. Terima kasih atas kesabaran dalam menemani penulis melewati masa perkuliahan serta atas kehadirannya yang selalu ada dalam setiap proses yang dijalani.
8. Aidah Maryam Barmin, Efriza Cahya Narendra, Imamah Nur Fadlilah, dan Dianita Puspitasari, yang telah menjadi teman diskusi dan berbagi pemikiran selama proses penyusunan skripsi ini. Terima kasih atas kebersamaan dan dukungan yang sangat berarti bagi penulis dalam menyelesaikan skripsi.
9. Teman-teman grup “Penghuni Bayangan Kos Parlin”, atas kebersamaan, canda tawa, dan dukungan yang berarti selama masa studi ini. Terima kasih atas semangat, berbagi informasi, serta pengalaman suka dan duka yang telah mewarnai perjalanan akademik penulis.
10. Sahabat-sahabat SMA penulis, yang tetap hadir dan setia menjadi bagian dari perjalanan hidup penulis hingga saat ini. Terima kasih selalu menjadi penghibur penulis saat suntuk dengan perkuliahan.
11. Seluruh rekan-rekan Sistem Informasi angkatan 2021, atas kebersamaan, dukungan, dan proses belajar bersama yang menjadi bagian tak terlupakan dari perjalanan studi ini.
12. Diri sendiri, atas perjuangan yang tidak mudah, atas setiap kerja keras yang telah terbayar dengan tercapainya salah satu pencapaian penting dalam hidup ini. Terima kasih telah bertahan.
13. Serta semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan bantuan, doa, dan dukungan, baik secara langsung maupun tidak langsung, dalam menyelesaikan skripsi ini.

Penulis menyadari sepenuhnya bahwa laporan skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang membangun demi penyempurnaan laporan skripsi ini di masa mendatang. Semoga laporan skripsi ini dapat memberikan manfaat, wawasan, serta menjadi kontribusi

kecil dalam pengembangan ilmu pengetahuan, khususnya di bidang analisis data dan teknologi informasi.

Surabaya, 26 April 2025

Penulis

Halaman ini sengaja dikosongkan

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN BEBAS PLAGIASI.....	vii
ABSTRAK	ix
KATA PENGANTAR.....	xiii
DAFTAR ISI.....	xvii
DAFTAR GAMBAR.....	xxiii
DAFTAR TABEL	xxvii
DAFTAR NOTASI.....	xxix
DAFTAR LAMPIRAN	xxxii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	7
1.3 Batasan Masalah	7
1.4 Tujuan Penelitian	8
1.5 Manfaat Penelitian	8
1.6 Sistematika Penulisan	8
BAB II TINJAUAN PUSTAKA.....	11
2.1 Penelitian Terdahulu	11
2.2 Landasan Teori.....	17
2.2.1 Dompot Digital.....	17
2.2.2 OVO	18
2.2.3 Persepsi	19
2.2.4 Google Play Store	20
2.2.5 Apple Store	21
2.2.6 Python	21
2.2.7 Google Colaboratory	22
2.2.8 <i>Aspect-Based Sentiment Analysis</i>	23
2.2.9 <i>Latent Dirichlet Allocation</i>	23
2.2.10 <i>Support Vector Machine</i>	25

2.2.11	<i>Web Scraping</i>	27
2.2.12	<i>Text Mining</i>	29
2.2.13	<i>Data Cleaning</i>	29
2.2.14	<i>Text Pre-processing</i>	30
2.2.14.1	<i>Text Cleaning</i>	30
2.2.14.2	<i>Case Folding</i>	31
2.2.14.3	<i>Tokenization</i>	31
2.2.14.4	<i>Normalization</i>	32
2.2.14.5	<i>Stopword Removal</i>	33
2.2.14.6	<i>Stemming</i>	34
2.2.15	<i>Fleiss' Kappa</i>	35
2.2.16	<i>Holdout</i>	36
2.2.17	<i>Term Frequency-Inverse Document Frequency</i>	37
2.2.18	<i>Multi-Label Synthetic Minority Oversampling Technique</i>	39
2.2.19	<i>Confusion Matrix</i>	40
2.2.19.1	<i>Accuracy</i>	41
2.2.19.2	<i>Precision</i>	41
2.2.19.3	<i>Recall</i>	42
2.2.19.4	<i>F1-Score</i>	42
2.2.20	<i>Hamming Loss</i>	43
2.2.21	<i>Flask</i>	43
BAB III METODOLOGI PENELITIAN		45
3.1	<i>Identifikasi Masalah</i>	45
3.2	<i>Studi Literatur</i>	46
3.3	<i>Analisis Kebutuhan</i>	46
3.3.1	<i>Kebutuhan Perangkat Keras</i>	46
3.3.2	<i>Kebutuhan Perangkat Lunak</i>	47
3.3.3	<i>Kebutuhan Data</i>	47
3.4	<i>Pengumpulan Data</i>	47
3.5	<i>Eksplorasi Data</i>	49
3.6	<i>Pembersihan Data</i>	49
3.7	<i>Text Pre-processing</i>	50

3.7.1	<i>Text Cleaning</i>	51
3.7.2	<i>Case Folding</i>	52
3.7.3	<i>Tokenization</i>	52
3.7.4	<i>Normalization</i>	53
3.7.5	<i>Stopword Removal</i>	53
3.7.6	<i>Stemming</i>	54
3.8	Pemodelan Topik	54
3.9	Interpretasi Topik	56
3.10	Pelabelan Data	56
3.10.1	Pelabelan Aspek	56
3.10.2	Pelabelan Sentimen	57
3.11	Pembagian Data	57
3.12	Pembangunan Model Klasifikasi	58
3.12.1	Pembagian Data Pemodelan Klasifikasi	59
3.12.2	Pembobotan Kata	60
3.12.3	Penyeimbangan Data	61
3.12.4	Pembentukan Model Klasifikasi	61
3.12.5	Hasil Pemodelan Klasifikasi	62
3.13	Evaluasi Performa Model	62
3.14	Analisis Model Terbaik	62
3.15	Penyimpanan Model	63
3.16	Implementasi Sistem	63
3.17	Pengujian Validasi Sistem	63
3.18	Penarikan Kesimpulan	64
3.19	Penyusunan Laporan	64
BAB IV HASIL DAN PEMBAHASAN		65
4.1	Hasil	65
4.1.1	Identifikasi Masalah	65
4.1.2	Studi Literatur	65
4.1.3	Analisis Kebutuhan	66
4.1.3.1	Kebutuhan Perangkat Keras	66
4.1.3.2	Kebutuhan Perangkat Lunak	66

4.1.3.3	Kebutuhan Data	67
4.1.4	Pengumpulan Data.....	67
4.1.5	Eksplorasi Data.....	72
4.1.6	Pembersihan Data	76
4.1.7	<i>Text Pre-processing</i>	79
4.1.7.1	<i>Text Cleaning</i>	79
4.1.7.2	<i>Case Folding</i>	80
4.1.7.3	<i>Tokenization</i>	82
4.1.7.4	<i>Normalization</i>	83
4.1.7.5	<i>Stopword Removal</i>	85
4.1.7.6	<i>Stemming</i>	88
4.1.8	Pemodelan Topik.....	92
4.1.9	Interpretasi Topik	99
4.1.10	Pelabelan Data	106
4.1.10.1	Pelabelan Aspek	107
4.1.10.2	Pelabelan Sentimen	111
4.1.11	Pembagian Data.....	118
4.1.12	Pembangunan Model Klasifikasi.....	120
4.1.12.1	Model Pengujian 1	120
4.1.12.2	Model Pengujian 2.....	121
4.1.12.3	Model Pengujian 3.....	122
4.1.12.4	Model Pengujian 4.....	123
4.1.12.5	Model Pengujian 5.....	123
4.1.12.6	Model Pengujian 6.....	124
4.1.12.7	Model Pengujian 7.....	125
4.1.12.8	Model Pengujian 8.....	126
4.1.12.9	Model Pengujian 9.....	127
4.1.12.10	Model Pengujian 10.....	128
4.1.12.11	Model Pengujian 11	129
4.1.12.12	Model Pengujian 12.....	130
4.1.13	Evaluasi Performa Model	131
4.1.14	Analisis Model Terbaik	133

4.1.15	Penyimpanan Model	136
4.1.16	Implementasi Sistem	137
4.1.17	Pengujian Validasi Sistem	139
4.2	Pembahasan.....	144
BAB V PENUTUP		149
5.1	Kesimpulan	149
5.2	Saran.....	150
DAFTAR PUSTAKA		153
LAMPIRAN.....		161

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 1.1 Proporsi Pengguna dan Layanan Pembayaran Digital	2
Gambar 1.2 <i>Rating</i> Aplikasi OVO pada Google Play Store	3
Gambar 1.3 Ulasan Pengguna Aplikasi OVO pada Google Play Store	4
Gambar 2.1 Transformasi dan Kontribusi OVO untuk Indonesia	18
Gambar 2.2 Representasi Model Grafis dari LDA.....	24
Gambar 2.3 Ilustrasi <i>Hyperplane</i> Pemisah dalam Ruang Fitur.....	26
Gambar 2.4 Visualisasi Margin Maksimum pada Model Klasifikasi Linear.....	27
Gambar 2.5 Kode Program Proses Dasar <i>Web Scraping</i>	28
Gambar 3.1 Alur Sistematika Penelitian.....	45
Gambar 3.2 Tahapan Pengumpulan Data Ulasan	48
Gambar 3.3 Alur Proses Eliminasi Baris Data Duplikasi	50
Gambar 3.4 Alur Proses Eliminasi Ulasan Satu Kata	50
Gambar 3.5 Alur Proses <i>Text Pre-processing</i>	51
Gambar 3.6 Alur Proses <i>Text Cleaning</i>	52
Gambar 3.7 Alur Proses <i>Case Folding</i>	52
Gambar 3.8 Alur Proses <i>Tokenization</i>	53
Gambar 3.9 Alur Proses <i>Normalization</i>	53
Gambar 3.10 Alur Proses <i>Stopword Removal</i>	54
Gambar 3.11 Alur Proses <i>Stemming</i>	54
Gambar 3.12 Alur Proses Pemodelan Topik.....	55
Gambar 3.13 Alur Pembangunan Model Klasifikasi	58
Gambar 4.1 Kode Program Proses <i>Scraping Data</i> dari Google Play Store	68
Gambar 4.2 Kode Program Proses <i>Scraping Data</i> dari Apple Store	69
Gambar 4.3 Kode Program Menghitung Analisis Deskriptif dari Dataset	72
Gambar 4.4 Visualisasi Jumlah Ulasan Berdasarkan Sumber	74
Gambar 4.5 Visualisasi Jumlah Ulasan Berdasarkan <i>Rating</i>	75
Gambar 4.6 Visualisasi <i>WordCloud</i> dari Ulasan Pengguna.....	76
Gambar 4.7 Kode Program Eliminasi Duplikasi Data Ulasan.....	77
Gambar 4.8 Kode Program Eliminasi Ulasan Terlalu Singkat	78
Gambar 4.9 Kode Program Implementasi Proses <i>Text Cleaning</i>	79

Gambar 4.10 Kode Program Implementasi Proses <i>Case Folding</i>	81
Gambar 4.11 Kode Program Implementasi Proses <i>Tokenization</i>	82
Gambar 4.12 Kode Program Implementasi Proses <i>Normalization</i>	84
Gambar 4.13 Kode Program Implementasi Proses <i>Stopword Removal</i> Pertama ..	86
Gambar 4.14 Kode Program Analisis 150 <i>Top Terms</i> Setelah <i>Stopword Removal</i> Pertama	86
Gambar 4.15 Kode Program Implementasi Proses <i>Stopword Removal</i> Kedua.....	87
Gambar 4.16 Kode Program Implementasi Proses <i>Stemming</i>	89
Gambar 4.17 Kode Program Persiapan Data dan Penerapan LDA	93
Gambar 4.18 Kode Program Evaluasi <i>Coherence Score</i>	98
Gambar 4.19 Visualisasi <i>Coherence Score</i> Setiap Topik.....	99
Gambar 4.20 Kode Program Implementasi Model LDA dengan 4 Topik	100
Gambar 4.21 Visualisasi <i>WordCloud</i> untuk Setiap Topik	103
Gambar 4.22 Visualisasi <i>Bar Chart</i> Probabilitas Kata Topik 0	103
Gambar 4.23 Visualisasi <i>Bar Chart</i> Probabilitas Kata Topik 1	104
Gambar 4.24 Visualisasi <i>Bar Chart</i> Probabilitas Kata Topik 2	104
Gambar 4.25 Visualisasi <i>Bar Chart</i> Probabilitas Kata Topik 3	105
Gambar 4.26 Kode Program Pelabelan Aspek Kata Kunci LDA.....	107
Gambar 4.27 Kode Program Pelabelan Aspek Perbaikan Kata Kunci LDA.....	109
Gambar 4.28 Kode Program Pelabelan Aspek Manual dengan Anotator	109
Gambar 4.29 Antarmuka Pelabelan Aspek Anotator	110
Gambar 4.30 Kode Program Implementasi Pelabelan Sentimen Anotator	112
Gambar 4.31 Antarmuka Pelabelan Sentimen Anotator	113
Gambar 4.32 Kode Program Menghitung Fleiss' Kappa antar Anotator.....	113
Gambar 4.33 Visualisasi Distribusi Label <i>Transaction Efficiency</i>	116
Gambar 4.34 Visualisasi Distribusi Label <i>User Experience</i>	116
Gambar 4.35 Visualisasi Distribusi Label <i>Account Access and Registration</i>	117
Gambar 4.36 Visualisasi Distribusi Label <i>Balance and Charges</i>	118
Gambar 4.37 Kode Program Pembagian Data.....	119
Gambar 4.38 Kode Program Menyimpan Model Terbaik.....	136
Gambar 4.39 Implementasi Flask dan Pemanggilan Model.....	137
Gambar 4.40 Halaman Utama Sistem	138

Gambar 4.41 Halaman Prediksi Sistem	139
Gambar 4.42 Kode Program Pengujian Data Validasi Sistem.....	139
Gambar 4.43 Contoh Hasil Prediksi Ulasan Berdasarkan Teks.....	141
Gambar 4.44 Tampilan Setelah Mengunggah <i>File</i> CSV	141
Gambar 4.45 Hasil Prediksi Ulasan Berdasarkan Unggah <i>File</i> CSV	142
Gambar 4.46 Tampilan <i>Bar Chart</i> Hasil Prediksi Ulasan dari <i>File</i> CSV	142
Gambar 4.47 Tampilan <i>Pie Chart</i> Hasil Prediksi Ulasan dari <i>File</i> CSV	143

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	11
Tabel 2.2 Contoh Proses <i>Text Cleaning</i>	30
Tabel 2.3 Contoh Proses <i>Case Folding</i>	31
Tabel 2.4 Contoh Proses <i>Tokenization</i>	32
Tabel 2.5 Contoh Proses <i>Normalization</i>	33
Tabel 2.6 Contoh Proses <i>Stopword Removal</i>	34
Tabel 2.7 Contoh Proses <i>Stemming</i>	35
Tabel 2.8 Interpretasi Nilai Fleiss' Kappa	36
Tabel 2.9 <i>Confusion Matrix</i>	40
Tabel 3.1 Skenario Pengujian Pembangunan Model Klasifikasi	59
Tabel 4.1 Contoh Hasil Ulasan dari Google Play Store	69
Tabel 4.2 Contoh Hasil Ulasan dari Apple Store	70
Tabel 4.3 Contoh Hasil Ulasan dari Kedua Platform	71
Tabel 4.4 Hasil Analisis Deskriptif dari Dataset	73
Tabel 4.5 Hasil Statistik Kolom Numerik dari Dataset	73
Tabel 4.6 Hasil Statistik Kolom Kategorikal dari Dataset	74
Tabel 4.7 Hasil Eliminasi Duplikasi Data	77
Tabel 4.8 Hasil Eliminasi Ulasan yang Terlalu Singkat	78
Tabel 4.9 Jumlah Total Data Setelah Proses Pembersihan Data	79
Tabel 4.10 Contoh Hasil Implementasi Proses <i>Text Cleaning</i>	80
Tabel 4.11 Contoh Hasil Implementasi Proses <i>Case Folding</i>	81
Tabel 4.12 Contoh Hasil Implementasi Proses <i>Tokenization</i>	82
Tabel 4.13 Contoh Kata dalam Kamus <i>Normalizaton</i>	83
Tabel 4.14 Contoh Hasil Implementasi Proses <i>Normalization</i>	84
Tabel 4.15 Analisis <i>Top Terms</i> Setelah <i>Stopword Removal</i> Pertama	87
Tabel 4.16 Contoh Hasil Implementasi Proses <i>Stopword Removal</i>	88
Tabel 4.17 Contoh Hasil Implementasi Proses <i>Stemming</i>	89
Tabel 4.18 Hasil Pengecekan Informasi Data Setelah <i>Pre-processing</i>	90
Tabel 4.19 Hasil Pengecekan <i>Missing Values</i> Setelah <i>Pre-processing</i>	91
Tabel 4.20 Hasil Pengecekan Informasi Data Final Setelah <i>Pre-processing</i>	92

Tabel 4.21 Hasil Perhitungan <i>Coherence Score</i> untuk Jumlah Topik 2–10	94
Tabel 4.22 Hasil Evaluasi <i>Coherence Score</i> untuk Setiap Topik.....	98
Tabel 4.23 Hasil Probabilitas Kata dalam Setiap Topik.....	100
Tabel 4.24 Identifikasi Topik Bahasan dan Aspek.....	105
Tabel 4.25 Contoh Hasil Pelabelan Aspek Kata Kunci LDA	108
Tabel 4.26 Hasil Perbandingan Pelabelan Aspek Kata Kunci LDA dengan Pelabelan Manual	111
Tabel 4.27 Hasil Perbandingan Pelabelan Aspek Perbaikan Kata Kunci LDA dengan Pelabelan Manual.....	111
Tabel 4.28 Hasil Perhitungan Fleiss' Kappa	114
Tabel 4.29 Contoh Hasil Label Akhir oleh Tiga Anotator	114
Tabel 4.30 Evaluasi Performa Model Pengujian 1	120
Tabel 4.31 Evaluasi Performa Model Pengujian 2	121
Tabel 4.32 Evaluasi Performa Model Pengujian 3	122
Tabel 4.33 Evaluasi Performa Model Pengujian 4.....	123
Tabel 4.34 Evaluasi Performa Model Pengujian 5	124
Tabel 4.35 Evaluasi Performa Model Pengujian 6.....	125
Tabel 4.36 Evaluasi Performa Model Pengujian 7	126
Tabel 4.37 Evaluasi Performa Model Pengujian 8.....	127
Tabel 4.38 Evaluasi Performa Model Pengujian 9	128
Tabel 4.39 Evaluasi Performa Model Pengujian 10.....	129
Tabel 4.40 Evaluasi Performa Model Pengujian 11	129
Tabel 4.41 Evaluasi Performa Model Pengujian 12.....	130
Tabel 4.42 Rekapitulasi Evaluasi Performa 12 Model Pengujian.....	132
Tabel 4.43 5 Model Terbaik dari 12 Model Pengujian.....	134
Tabel 4.44 Evaluasi Performa Model Data Validasi Sistem	140
Tabel 4.45 Perbandingan Performa Model Terbaik dan Data Validasi.....	144
Tabel 4.46 Contoh Ulasan Prediksi Aspek dan Sentimen yang Tepat	145
Tabel 4.47 Contoh Ulasan Prediksi Aspek yang Salah	145
Tabel 4.48 Contoh Ulasan Prediksi Sentimen yang Salah	147
Tabel 4.49 Contoh Ulasan Tidak Dapat Diprediksi Aspek dan Sentimen	148

DAFTAR NOTASI

K	: Koefisien tingkat kesepakatan antar penilai
P_o	: Proporsi kesepakatan aktual
P_e	: Proporsi kesepakatan yang diharapkan secara acak
t	: Kata yang dianalisis dalam dokumen
d	: Dokumen yang sedang dianalisis
$f(t, d)$	: Frekuensi kemunculan kata dalam dokumen
N_d	: Total jumlah kata dalam dokumen
D	: Kumpulan dokumen (<i>corpus</i>)
N	: Jumlah total dokumen dalam <i>corpus</i>
$df(t, D)$	: Jumlah dokumen yang mengandung kata
TF	: Frekuensi kemunculan kata dalam dokumen
IDF	: Ukuran kekhasan kata dalam <i>corpus</i>
$TF-IDF$	: Bobot gabungan berdasarkan TF dan IDF
TP	: <i>True Positive</i> , prediksi benar untuk kelas positif
TN	: <i>True Negative</i> , prediksi benar untuk kelas negatif
FP	: <i>False Positive</i> , prediksi salah untuk kelas positif
FN	: <i>False Negative</i> , prediksi salah untuk kelas negatif

Halaman ini sengaja dikosongkan

DAFTAR LAMPIRAN

Lampiran 1: Permohonan Perizinan Pengambilan Data Ulasan	161
--	-----

Halaman ini sengaja dikosongkan