



SKRIPSI

PENERAPAN *FINE-TUNING* INDOBERT UNTUK *ASPECT-BASED SENTIMENT* *ANALYSIS* (ABSA) PADA ULASAN PENGGUNA APLIKASI JAKI

SAFNA FARADILLAH

22082010095

DOSEN PEMBIMBING

Abdul Rezha Efrat Najaf, S.Kom., M.Kom.

Muhammad Reza Pahlawan, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
SURABAYA
2026**

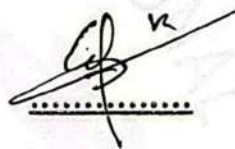
LEMBAR PENGESAHAN

PENERAPAN *FINE-TUNING* INDOBERT UNTUK *ASPECT-BASED SENTIMENT ANALYSIS* (ABSA) PADA ULASAN PENGGUNA APLIKASI JAKI

Oleh :
SAFNA FARADILLAH
NPM. 22082010095

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur pada tanggal 23 bulan Juni tahun 2026.

Abdul Rezha Efrat Najaf, S.Kom., M.Kom.
NIP. 19940929 2022031 008



(Pembimbing I)

Muhammad Reza Pahlawan, S.Kom., M.Kom.
NIP. 19980516 2025061 005



(Pembimbing II)

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



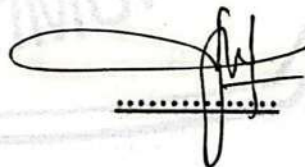
(Ketua Penguji)

Rizka Hadiwiyanti, S.Kom., M.Kom., MBA.
NIP. 19860727 2018032 001



(Anggota Penguji II)

Mohammad Al Hafidz, S. Kom., M.Kom.
NIP. 19910922 2025061 003



(Anggota Pengujii III)

Mengetahui,
Dekan Fakultas Ilmu Komputer

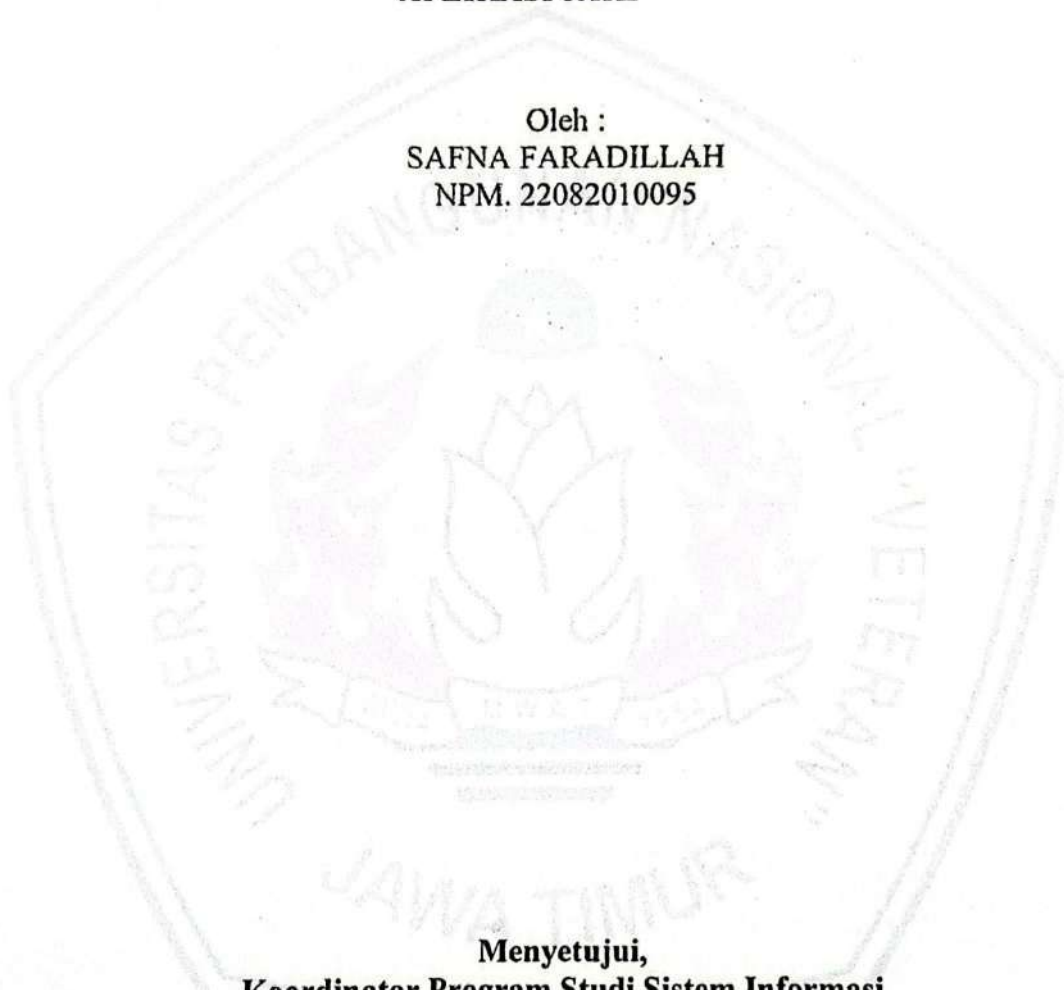


Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

**PENERAPAN *FINE-TUNING* INDOBERT UNTUK *ASPECT-BASED SENTIMENT ANALYSIS* (ABSA) PADA ULASAN PENGGUNA
APLIKASI JAKI**

Oleh :
SAFNA FARADILLAH
NPM. 22082010095



**Menyetujui,
Koordinator Program Studi Sistem Informasi
Fakultas Ilmu Komputer**



**Siti Mukaromah, S.Kom., M.Kom.
NIP. 19810704 2021212 011**

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama : SAFNA FARADILLAH
NPM : 22082010095
Program : Sarjana (S1)
Program Studi : Sistem Informasi
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah 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, 14 Juli 2026
Yang Membuat Pernyataan,



A handwritten signature in black ink, appearing to read 'Safna Faradillah'.

SAFNA FARADILLAH
NPM. 22082010095

ABSTRAK

Nama Mahasiswa / NPM : Safna Faradillah / 22082010095
Judul Skripsi : Penerapan *Fine-Tuning* IndoBERT Untuk *Aspect-Based Sentiment Analysis* (ABSA) Pada Ulasan Pengguna Aplikasi JAKI
Dosen Pembimbing : 1. Abdul Rezha Efrat Najaf, S.Kom., M.Kom.
2. Muhammad Reza Pahlawan, S.Kom., M.Kom.

Seiring berkembangnya transformasi digital, pemanfaatan platform digital dalam pelayanan publik turut terdampak, salah satunya aplikasi JAKI. Aplikasi JAKI (Jakarta Kini) merupakan aplikasi layanan publik digital Pemerintah Provinsi DKI Jakarta yang telah menerima ribuan ulasan yang beragam dari pengguna di Google Play Store. Oleh karena itu, perlu dilakukan analisis sentimen berbasis aspek untuk dapat memahami opini pengguna secara lebih mendalam terhadap suatu aspek. Pada penelitian ini melakukan penerapan model *fine-tuning* IndoBERT dengan pendekatan ABSA pada ulasan aplikasi JAKI di Google Play Store terhadap aspek *predefined*, yaitu *functional*, *service*, dan *usability*, dengan klasifikasi sentimen positif, negatif, dan netral. Proses pemodelan dilakukan dengan membandingkan performa model XGBoost dan *fine-tuning* IndoBERT yang dikombinasi dengan skenario *hyperparameter optimization*, dua rasio pembagian data, yaitu 70:15:15 dan 80:10:10, serta tiga teknik penanganan data tidak seimbang, yaitu Near Miss, SMOTE, dan Class Weighting. Hasil evaluasi menunjukkan model *fine-tuning* IndoBERT dengan parameter *learning rate* 2×10^{-5} , *batch size* 16, dan *epoch* 10, menggunakan teknik Class Weighting pada rasio 70:15:15 sebagai skenario terbaik dengan performa rata-rata nilai *f1-score macro* tertinggi, yaitu pada aspek *service* dan *usability* masing-masing mencapai 76%, sementara aspek *functional* 49% yang disebabkan oleh ketidakseimbangan kelas yang sangat ekstrem. Hasil penelitian ini menunjukkan pemilihan parameter, rasio pembagian data, dan teknik penanganan data tidak seimbang yang tepat mampu memberikan pengaruh yang signifikan terhadap performa model *fine-tuned* IndoBERT dalam melakukan analisis sentimen berbasis aspek pada ulasan aplikasi JAKI.

Kata kunci : Analisis Sentimen, Analisis Sentimen Berbasis Aspek, Data Tidak Seimbang, *Fine-Tuning* IndoBERT, JAKI.

ABSTRACT

Student Name / NPM : Safna Faradillah / 22082010095
Thesis Title : Implementation of IndoBERT Fine-Tuning for Aspect-Based Sentiment Analysis (ABSA) on JAKI Application User Reviews
Advisor : 1. Abdul Rezha Efrat Najaf, S.Kom., M.Kom.
2. Muhammad Reza Pahlawan, S.Kom., M.Kom.

Along with the rapid advancement of digital transformation, the adoption of digital platforms in public services has also increased, one of which is the JAKI application. The JAKI (Jakarta Kini) is a digital public service application developed by the Provincial Government of DKI Jakarta that has received thousands of diverse user reviews on the Google Play Store. Therefore, aspect-based sentiment analysis is necessary to gain a deeper understanding of user opinions regarding specific aspects. This study employs a fine-tuned IndoBERT model, utilizing the ABSA approach, to analyze JAKI app reviews on the Google Play Store, with a focus on predefined aspects: functional, service, and usability, and classifying sentiments into positive, negative, and neutral categories. The modeling process was conducted by comparing the performance of the XGBoost model and fine-tuned IndoBERT model, combined with hyperparameter optimization scenarios, two data split ratios (70:15:15 and 80:10:10), and three data imbalance techniques, namely Near Miss, SMOTE, and Class Weighting. The evaluation results indicate that the fine-tuned IndoBERT model with a learning rate of 2×10^{-5} , a batch size of 16, and 10 epochs, combined with the Class Weighting technique using the 70:15:15 data split ratio, achieved the best performance, yielding the highest average macro F1-score. Specifically, the service and usability aspects each achieved a macro F1-score of 76%, while the functional aspect achieved 49%, due to the extremely severe class imbalance. The results of this study indicate that the selection of appropriate parameters, data splitting ratios, and techniques for handling imbalanced data can have a significant impact on the performance of the *fine-tuned* IndoBERT model in conducting aspect-based sentiment analysis on reviews of the JAKI app.

Keywords: Sentiment Analysis, Aspect-Based Sentiment Analysis, Data Imbalance, Fine-Tuning IndoBERT, JAKI.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, taufik, dan hidayah-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul *"Penerapan Fine-Tuning IndoBERT Untuk Aspect-Based Sentiment Analysis (ABSA) Pada Ulasan Pengguna Aplikasi JAKI"*. Penulisan skripsi ini merupakan tahapan akhir yang harus ditempuh untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur.

Penyusunan skripsi ini memerlukan waktu, tenaga, dan komitmen, serta mendapat dukungan dari berbagai pihak. Untuk itu, penulis menyampaikan terima kasih kepada:

1. Allah SWT atas segala rahmat, karunia, kesehatan, kemudahan, serta kekuatan yang diberikan sehingga penulis dapat menyelesaikan skripsi dan perkuliahan ini.
2. Diri sendiri, Safna Faradillah, terima kasih telah senantiasa berusaha, bertahan, dan tidak menyerah dalam menghadapi berbagai kesulitan selama proses penyusunan skripsi dan perkuliahan di perantauan.
3. Bapak Carto dan Ibu Nilla, orang tua penulis yang senantiasa memberikan do'a, kasih sayang, dukungan moral maupun material, serta menjadi sumber kekuatan terbesar bagi penulis.
4. Kedua kakak penulis, Indriany dan Anissa, yang selalu memberikan dukungan, perhatian, nasihat, dan menjadi teladan dan sumber motivasi bagi penulis, serta adik penulis, Aura, yang selalu menjadi sumber hiburan penulis selama perkuliahan.
5. Ibu Siti Mukaromah, S.Kom., M.Kom., selaku dosen wali yang selalu memberikan motivasi dan arahan kepada penulis selama perkuliahan.
6. Bapak Abdul Rezha Efrat Najaf, S.Kom., M.Kom. dan Bapak Muhammad Reza Pahlawan, S.Kom., M.Kom., selaku dosen pembimbing skripsi, yang telah meluangkan waktu, tenaga, dan pikiran untuk memberikan arahan, masukan, dan motivasi sehingga penulis dapat menyelesaikan skripsi ini.

7. Seluruh Bapak dan Ibu dosen program studi Sistem Informasi yang telah memberikan ilmu, wawasan, pengalaman, serta bimbingan selama masa perkuliahan..
8. Keponakan-keponakan tersayang, Naureen, Alsaki, Kyana, dan Kyara, yang secara tidak langsung telah memberikan hiburan dan semangat bagi penulis selama proses pengerjaan skripsi.
9. Teman seperjuangan, Layla, Nadin, Novanda, dan Nadya, yang telah menemani, mendukung, dan berjuang bersama selama proses penyusunan skripsi maupun perkuliahan.
10. Teman-teman Parcing, Galaksi (Sistem Informasi Angkatan 2022), dan teman-teman lainnya yang tidak bisa disebutkan satu per satu, yang telah menemani perjalanan perkuliahan dan kehidupan perantauan penulis di Surabaya, serta memberikan banyak pengalaman, pembelajaran, dukungan, dan kenangan yang berharga.
11. Grup K-Pop Seventeen yang selalu menghibur penulis, dan temen-temen di server Study Together yang telah menjadi teman belajar dan memberikan motivasi selama proses penyusunan skripsi, serta seluruh pihak yang tidak bisa disebutkan satu per satu yang telah membantu dan mendukung penulis baik secara langsung maupun tidak langsung.

Penulis menyadari skripsi ini masih memiliki kekurangan, sehingga saran dan kritik yang membangun sangat diharapkan. Semoga skripsi ini dapat memberikan manfaat bagi pembaca, peneliti berikutnya, dan pihak-pihak yang berkepentingan.

Surabaya, 14 Juli 2026

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN.....	v
SURAT PERNYATAAN BEBAS PLAGIASI	vii
ABSTRAK	ix
ABSTRACT.....	xi
KATA PENGANTAR	xiii
DAFTAR ISI.....	xv
DAFTAR TABEL	xix
DAFTAR GAMBAR	xxi
DAFTAR LAMPIRAN	xxv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian	7
1.6 Sistematika Penelitian	7
BAB II TINJAUAN PUSTAKA.....	9
2.1 Dasar Teori.....	9
2.1.1 Jakarta Kini (JAKI)	9
2.1.2 Text mining	10
2.1.3 Natural Language Processing (NLP)	11
2.1.4 Analisis sentimen	14

2.1.5 ABSA	15
2.1.6 Transformer	16
2.1.7 BERT	18
2.1.8 IndoBERT	20
2.1.9 Fine-tuning	21
2.1.10 XGBoost.....	22
2.1.11 CRISP-DM	24
2.1.12 Krippendorff's Alpha	25
2.1.13 Confusion matrix.....	25
2.1.14 Metrik kinerja klasifikasi	26
2.1.15 Express.js.....	27
2.1.16 Flask	28
2.1.17 Near Miss	28
2.1.18 SMOTE	28
2.1.19 Class Weighting	28
2.1.20 ISO/IEC 25010:2011.....	29
2.2 Penelitian Terdahulu	29
BAB III METODOLOGI PENELITIAN	33
3.1 Studi Literatur	34
3.2 <i>Business Understanding</i>	35
3.3 <i>Data Collection</i>	37
3.4 <i>Data Filtering</i>	37
3.5 <i>Data Labeling</i>	37
3.6 <i>Additional Filtering</i>	39
3.7 <i>General Preprocessing</i>	39
3.8 <i>Data Splitting</i>	40

3.9 <i>Specific Preprocessing and Feature Extraction</i>	40
3.9.1 Baseline model	40
3.9.2 IndoBERT model	41
3.10 Handling Imbalance	41
3.11 <i>Modeling</i>	42
3.11.1 Model <i>baseline</i>	42
3.11.2 Model transformer	43
3.12 Model Evaluation	44
3.13 <i>Aspect-Based Performance Analysis</i>	44
3.14 Visualisasi	44
BAB IV HASIL DAN PEMBAHASAN	45
4.1 Data Collection.....	45
4.2 Data Filtering	46
4.3 Data Labeling	48
4.4 Additional Filtering.....	50
4.5 General Preprocessing.....	52
4.6 Splitting.....	55
4.7 Specific Preprocessing	56
4.7.1 Model Baseline.....	56
4.7.2 Model IndoBERT	57
4.8 Handling Imbalance	59
4.8.1 Near Miss	59
4.8.2 SMOTE	60
4.8.3 Class Weighting	61
4.9 Modeling	61
4.9.1 Near Miss	65

4.9.2 SMOTE	66
4.9.3 Class Weighting	67
4.9.4 Pemilihan Skenario Terbaik Setiap Teknik.....	68
4.10 Evaluation	69
4.10.1 Perbandingan Performa antar Skenario Terbaik Setiap Teknik	69
4.10.2 Perbandingan Performa Model Terbaik Dengan Tanpa Teknik	78
4.11 ABSA Analysis	79
4.12 Visualisasi	93
4.12.1 Dashboard	94
4.12.2 Analisis Sentimen	96
4.12.3 Tentang JAKI	97
BAB V PENUTUP.....	99
5.1 Kesimpulan	99
5.2 Saran.....	100
DAFTAR PUSTAKA	103
LAMPIRAN.....	115

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	26
Tabel 2.2 Penelitian Terdahulu	30
Tabel 3.1 Skenario Model <i>Baseline</i> XGBoost	42
Tabel 3.2 Skenario Model IndoBERT	43
Tabel 4.1 Contoh Hasil <i>Data Collection</i>	46
Tabel 4.2 Contoh Hasil <i>Data Filtering</i>	48
Tabel 4.3 Contoh Hasil <i>Data Labeling</i> Kelima Anotator	49
Tabel 4.4 Contoh Hasil <i>Data Labeling</i> Akhir.....	50
Tabel 4.5 Contoh Hasil <i>Data Additional Filtering</i>	52
Tabel 4.6 Contoh Hasil <i>Case Folding</i>	52
Tabel 4.7 Contoh Hasil <i>Data Cleaning</i>	53
Tabel 4.8 Contoh Hasil Normalisasi	54
Tabel 4.9 Contoh Hasil <i>Translation</i>	55
Tabel 4.10 Contoh Hasil <i>Splitting</i>	55
Tabel 4.11 Contoh Hasil <i>Specific Preprocessing Model Baseline</i>	57
Tabel 4.12 Konfigurasi TF-IDF	57
Tabel 4.13 Distribusi Data Latih Sebelum dan Sesudah Teknik Near Miss.....	59
Tabel 4.14 Distribusi Data Latih Sebelum dan Sesudah Teknik SMOTE.....	60
Tabel 4.15 Hasil Bobot Kelas Setiap Aspek	61
Tabel 4.16 Hasil Validasi Seluruh Skenario Kombinasi Model	62
Tabel 4.17 Hasil Validasi Skenario Terbaik Setiap Teknik.....	68
Tabel 4.18 Perbandingan Hasil Evaluasi Model Terbaik Setiap Teknik	69
Tabel 4.19 Hasil Evaluasi Model Terbaik Pertama per Kelas	72
Tabel 4.20 Hasil Evaluasi Model Terbaik Kedua per Kelas.....	75

Tabel 4.21 Perbandingan Performa Model Terbaik dengan Tanpa Teknik Penanganan Data.....	79
--	----

DAFTAR GAMBAR

Gambar 2.1 <i>Rating</i> Aplikasi JAKI	9
Gambar 2.2 Teknik <i>Parsing</i>	12
Gambar 2.3 <i>Transformer</i>	16
Gambar 2.4 <i>Multi-Head Attention</i>	17
Gambar 2.5 Struktur <i>Input Token</i> BERT	19
Gambar 2.6. Proses <i>Embedding</i> IndoBERT	21
Gambar 2.7 Alur CRISP-DM Secara Umum	24
Gambar 3.1 Metodologi Penelitian	34
Gambar 4.1 <i>Source Code</i> Proses Pengumpulan Data	45
Gambar 4.2 <i>Source Code Data Filtering</i> (1)	46
Gambar 4.3 <i>Source Code Data Filtering</i> (2)	47
Gambar 4.4 <i>Source Code Data Filtering</i> (3)	47
Gambar 4.5 <i>Source Code</i> Perhitungan <i>Krippendorff's Alpha</i>	50
Gambar 4.6 <i>Source Code</i> Pengecekan <i>Missing Value</i>	51
Gambar 4.7 <i>Source Code</i> Simplifikasi dan Imputasi Data Kosong	51
Gambar 4.8 <i>Source Code</i> Penghapusan Data dan Kolom Tidak Relevan	51
Gambar 4.9 <i>Source Code Case Folding</i>	52
Gambar 4.10 <i>Source Code Data Cleaning</i>	53
Gambar 4.11 <i>Source Code Normalization</i>	54
Gambar 4.12 <i>Source Code Translation</i>	54
Gambar 4.13 <i>Source Code</i> Distribusi Sentimen Pada Data Latih	56
Gambar 4.14 <i>Source Code</i> Konfigurasi Proses Tokenisasi IndoBERT	58
Gambar 4.15 Proses Tokenisasi IndoBERT	59

Gambar 4.16 <i>Source Code</i> Konfigurasi XGBoost Teknik Sampling	65
Gambar 4.17 <i>Source Code</i> Konfigurasi IndoBERT Teknik Sampling.....	65
Gambar 4.18 <i>Source Code</i> Konfigurasi Model XGBoost Class Weighting	67
Gambar 4.19 <i>Source Code</i> Konfigurasi IndoBERT Class Weighting	67
Gambar 4.20 <i>Confusion Matrix</i> Aspek <i>Functional</i> Model Terbaik Pertama.....	73
Gambar 4.21 <i>Confusion Matrix</i> Aspek <i>Service</i> Model Terbaik Pertama.....	74
Gambar 4.22 <i>Confusion Matrix</i> Aspek <i>Usability</i> Model Terbaik Pertama.....	74
Gambar 4.23 <i>Confusion Matrix</i> Aspek <i>Functional</i> Model Terbaik Kedua	76
Gambar 4.24 <i>Confusion Matrix</i> Aspek <i>Service</i> Model Terbaik Kedua	77
Gambar 4.25 <i>Confusion Matrix</i> Aspek <i>Usability</i> Model Terbaik Kedua	77
Gambar 4.26 <i>Pie Chart</i> Distribusi Sentimen Keseluruhan	80
Gambar 4.27 <i>Bar Chart</i> Distribusi Sentimen Tiap Aspek	80
Gambar 4.28 <i>Line Chart</i> Tren Sentimen per Versi Rilis	81
Gambar 4.29 <i>Line Chart</i> Tren Sentimen pada Aspek <i>Functional</i>	82
Gambar 4.30 <i>Line Chart</i> Tren Sentimen pada Aspek <i>Service</i>	83
Gambar 4.31 <i>Line Chart</i> Tren Sentimen pada Aspek <i>Usability</i>	84
Gambar 4.32 Top 20 Kata yang Sering Muncul	85
Gambar 4.33 <i>Wordcloud</i> Seluruh Aspek dan Seluruh Sentimen	86
Gambar 4.34 <i>Wordcloud</i> Aspek <i>Functional</i> Seluruh Sentimen.....	86
Gambar 4.35 <i>Wordcloud</i> Aspek <i>Functional</i> Sentimen Positif.....	87
Gambar 4.36 <i>Wordcloud</i> Aspek <i>Functional</i> Sentimen Negatif	87
Gambar 4.37 <i>Wordcloud</i> Aspek <i>Functional</i> Sentimen Netral	88
Gambar 4.38 <i>Wordcloud</i> Aspek <i>Service</i> Seluruh Sentimen	89
Gambar 4.39 <i>Wordcloud</i> Aspek <i>Service</i> Sentimen Positif	89
Gambar 4.40 <i>Wordcloud</i> Aspek <i>Service</i> Sentimen Negatif	90
Gambar 4.41 <i>Wordcloud</i> Aspek <i>Service</i> Sentimen Netral	90

Gambar 4.42 <i>Wordcloud</i> Aspek <i>Usability</i> Seluruh Sentimen.....	91
Gambar 4.43 <i>Wordcloud</i> Aspek <i>Usability</i> Sentimen Positif.....	91
Gambar 4.44 <i>Wordcloud</i> Aspek <i>Usability</i> Sentimen Negatif	92
Gambar 4.45 <i>Wordcloud</i> Aspek <i>Usability</i> Sentimen Netral	92
Gambar 4.46 Halaman <i>Dashboard</i>	94
Gambar 4.47 Fitur Filter Versi Rilis Aplikasi.....	95
Gambar 4.48 Tampilan Grafik Tren dan Distribusi Sentimen per Versi Rilis.....	95
Gambar 4.49 Tampilan <i>Wordcloud</i>	95
Gambar 4.50 Halaman Analisis Sentimen	96
Gambar 4.51 Contoh Hasil Analisis Sentimen Teks Tunggal	96
Gambar 4.52 Contoh Hasil Analisis Sentimen Unggah Berkas.....	97
Gambar 4.53 Halaman Tentang JAKI.....	97

DAFTAR LAMPIRAN

1.	Dataset.....	115
2.	Bukti Permohonan Izin	115
3.	Panduan Pelabelan Manusia.....	115
4.	Prompt Pelabelan AI	116