



SKRIPSI

KOMPARASI ALGORITMA *TOPIC MODELING* DAN *MACHINE LEARNING* UNTUK KLASIFIKASI SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI PLN MOBILE

YULIANI PURWITASARI

NPM 22082010168

DOSEN PEMBIMBING

Eka Dyar Wahyuni, S.Kom.,M.Kom.

Tri Luhur Indayanti Sugata, S.ST., M.IM.

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

LEMBAR PENGESAHAN

KOMPARASI ALGORITMA *TOPIC MODELING* DAN *MACHINE LEARNING* UNTUK KLASIFIKASI SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI PLN MOBILE

Oleh :

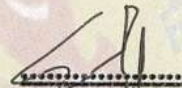
YULIANI PURWITASARI

NPM. 22082010168

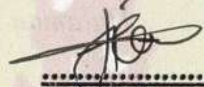
Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 27 Februari 2026.

Menyetujui

Eka Dyar Wahyuni, S.Kom., M.Kom.
NIP. 198412012021212005


..... (Pembimbing I)

Tri Luhur Indayanti Sugata, S.ST., M.IM.
NIP. 199206162024062001


..... (Pembimbing II)

Rizka Hadiwiyanti, S.Kom., M.Kom., MBA
NIP. 198607272018032001


..... (Ketua Penguji)

Prasasti Karunia F. A., S.Kom., M.Kom., M.IM.
NIP. 199707042024062001


..... (Penguji II)

Nambi Sembilu, S.Kom., M. Kom.
NIP. 199005162024061003


..... (Penguji III)

Mengetahui,

Dekan Fakultas Ilmu Komputer


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

LEMBAR PERSETUJUAN

**KOMPARASI ALGORITMA *TOPIC MODELING* DAN *MACHINE LEARNING* UNTUK KLASIFIKASI SENTIMEN BERBASIS ASPEK
PADA ULASAN APLIKASI PLN MOBILE**

Oleh:

YULIANI PURWITASARI

NPM. 22082010168

Menyetujui,

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



Agung Brastama Putra, S.Kom., M.Kom.
NIP. 19851124 202121 1 003

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertandatangan di bawah ini:

Nama Mahasiswa : Yuliani Purwitasari
NPM : 22082010168
Program : Sarjana (S1)
Program Studi : Sistem Informasi
Fakultas : 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, 27 Februari 2026

Yang Membuat Pernyataan,



Yuliani Purwitasari

NPM. 22082010168

ABSTRAK

Nama Mahasiswa / NPM : Yuliani Purwitasari / 22082010168
Judul Skripsi : Komparasi Algoritma *Topic Modeling* dan *Machine Learning* untuk Klasifikasi Sentimen Berbasis Aspek pada Ulasan Aplikasi PLN Mobile
Dosen Pembimbing : 1. Eka Dyar Wahyuni, S.Kom.,M.Kom.
2. Tri Luhur Indayanti Sugata, S.ST., M.IM.

Perkembangan teknologi informasi dan komunikasi mendorong digitalisasi layanan publik, termasuk transformasi layanan PT PLN (Persero) melalui aplikasi PLN Mobile. Meskipun aplikasi ini memperoleh *rating* tinggi di Google Play Store dan App Store, masih terdapat ulasan negatif yang menunjukkan kendala pengguna. Volume ulasan yang besar dan tidak terstruktur memerlukan pendekatan otomatis melalui *Aspect-Based Sentiment Analysis* (ABSA) untuk mengidentifikasi aspek utama dan sentimen secara lebih spesifik. Penelitian ini bertujuan mengevaluasi metode *topic modeling* (LDA dan BERTopic), mengomparasikan algoritma klasifikasi SVM dan Random Forest, serta menentukan kombinasi terbaik dalam analisis sentimen berbasis aspek pada ulasan PLN Mobile. Data ulasan berbahasa Indonesia dikumpulkan melalui *scraping*, kemudian melalui tahap *pre-processing*, pemodelan topik, dan klasifikasi multilabel dengan skenario pembagian data serta penerapan SMOTE-Tomek Links untuk menangani ketidakseimbangan kelas. Hasil pemodelan menunjukkan bahwa LDA menghasilkan dua aspek utama, yaitu *user experience* dan *service quality*, sedangkan BERTopic menghasilkan tiga aspek utama, yaitu *user experience*, *service quality*, dan *system reliability*. Secara evaluasi, BERTopic memperoleh *coherence score* (c_v) tertinggi sebesar 0,6707 pada empat topik, melampaui LDA yang mencapai 0,5543 pada dua topik. Namun, pada implementasi akhir, BERTopic menghasilkan tiga topik utama berdasarkan proses *clustering* otomatis. Kombinasi model klasifikasi terbaik diperoleh pada model LDA + Random Forest + SMOTE-Tomek Links (80:20) dengan *Macro F1-Score* 0,769 dan AUC 0,85. Model validasi memperoleh *Macro F1-Score* 0,739, *Macro AUC* 0,85, dan *Hamming Loss* 0,058. Model diimplementasikan ke dalam sistem berbasis website yang mampu memprediksi aspek dan sentimen melalui input teks maupun *file* CSV serta menampilkan visualisasi distribusi, proporsi, dan tren sentimen.

Kata Kunci: *Aspect-Based Sentiment Analysis*, *Topic Modeling*, Algoritma Klasifikasi, Klasifikasi Multilabel, PLN Mobile.

ABSTRACT

Student Name / NPM : Yuliani Purwitasari / 22082010168
Thesis Title : Comparison of Topic Modeling and Machine Learning Algorithms for Aspect-Based Sentiment Classification in PLN Mobile Application Reviews
Dosen Pembimbing : 1. Eka Dyar Wahyuni, S.Kom.,M.Kom.
2. Tri Luhur Indayanti Sugata, S.ST., M.IM.

Advances in information and communication technology have accelerated the digitalization of public services, including the transformation of PT PLN (Persero) services through the PLN Mobile application. Although the application has received high ratings on the Google Play Store and App Store, negative reviews still indicate user-reported issues. The large volume of unstructured reviews requires an automated approach using Aspect-Based Sentiment Analysis (ABSA) to more specifically identify key aspects and sentiments. This study aims to evaluate topic modeling methods (LDA and BERTopic), compare the performance of SVM and Random Forest classification algorithms, and determine the best combination for aspect-based sentiment analysis of PLN Mobile reviews. Indonesian language reviews were collected through scraping, followed by preprocessing, topic modeling, and multilabel classification with various data splitting scenarios and the application of SMOTE-Tomek Links to address class imbalance. The modeling results show that LDA yields two key aspects: user experience and service quality, while BERTopic yields three key aspects: user experience, service quality, and system reliability. In terms of evaluation, BERTopic achieved the highest coherence score (c_v) of 0.6707 with four topics, outperforming LDA which achieved 0.5543 with two topics. However, in the final implementation, BERTopic generated three main topics based on its automatic clustering process. The best classification performance was obtained using LDA + Random Forest + SMOTE-Tomek Links (80:20 split), achieving a Macro F1-Score of 0.769 and an AUC of 0.85. The validation model achieved a Macro F1-Score of 0.739, a Macro AUC of 0.85, and a Hamming Loss of 0.058. The model is implemented into a website-based system that is able to predict aspects and sentiments through text input or CSV files and displays visualizations of distribution, proportion, and sentiment trends.

Keywords: Aspect-Based Sentiment Analysis, Topic Modeling, Classification Algorithm, Multi-label Classification, PLN Mobile.

KATA PENGANTAR

Puji dan syukur penulis ucapkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat, kasih, serta penyertaan-Nya sehingga penulis dapat menuntaskan skripsi yang berjudul “Komparasi Algoritma Topic Modeling dan Machine Learning untuk Klasifikasi Sentimen Berbasis Aspek pada Ulasan Aplikasi PLN Mobile” dengan baik. Penyusunan skripsi ini dilakukan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana (Strata 1) pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Proses penyusunan skripsi ini menjadi perjalanan panjang yang sarat dengan pembelajaran, tantangan, dan refleksi diri, yang tidak mungkin penulis lalui seorang diri tanpa sokongan dari berbagai pihak. Dengan penuh kerendahan hati, rasa hormat dan rasa syukur yang mendalam, dalam kesempatan ini ingin penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Koordinator Prodi Sistem Informasi, Bapak Agung Brastama Putra, S.Kom., M.Kom., yang telah memberikan arahan, dukungan, serta kebijakan sehingga proses akademik penulis dapat berjalan dengan lancar.
2. Ibu Eka Dyar Wahyuni, S.Kom., M.Kom., selaku Dosen Pembimbing I, yang telah membimbing dengan penuh kesabaran, memberikan motivasi, arahan, serta masukan yang sangat berarti bagi penulis sejak tahap awal hingga selesainya penyusunan penelitian skripsi ini.
3. Ibu Tri Luhur Indayanti Sugata, S.ST., M.IM., selaku Dosen Pembimbing II, yang telah memberikan ilmu, arahan, serta evaluasi yang konstruktif sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
4. Seluruh dosen dan tenaga kependidikan Program Studi Sistem Informasi, yang telah memberikan ilmu, pengalaman, serta pelayanan akademik dengan penuh dedikasi selama penulis menempuh masa perkuliahan.
5. Keluarga tercinta, terutama Ibu, alm. Kakung, Om, Tante, dan kedua Sepupu penulis, Aishwa dan Alisha yang selalu menjadi tempat pulang ternyaman, mendukung setiap cita-cita penulis, serta memberikan dukungan sejak penulis lahir hingga berhasil menjadi sarjana pertama di keluarga. Pengorbanan, ketulusan, dan kepercayaan yang diberikan menjadi kekuatan terbesar bagi

penulis untuk terus melangkah dan tidak menyerah dalam menghadapi setiap proses kehidupan.

6. Kedua orang tua dan adik terkasih, yang senantiasa memberikan doa, dukungan, serta harapan bagi penulis. Dukungan dan doa yang diberikan menjadi motivasi terbesar bagi penulis untuk menyelesaikan studi ini.
7. Sahabat penulis, Sundari Cintya Meilasari dan Fransiscus Patric Diaz Wibisono, yang telah menjadi bagian berharga dalam perjalanan hidup penulis sejak masa putih abu-abu. Terima kasih atas segala dukungan, kehadiran, serta kepercayaan yang diberikan kepada penulis sehingga menjadi penguat langkah penulis sampai hari ini.
8. Teduh Daffa Maulana, yang senantiasa membersamai dan memberikan dukungan emosional, kepercayaan, serta semangat tanpa henti kepada penulis. Kehadiran yang diberikan menjadi penyeimbang di tengah tekanan selama proses penyusunan skripsi.
9. Teman seperkuliah penulis, Putri Intan Octavia Br Sipayung, Hellawati Ayu Rizmadita, Kanaya Deva Cahyarani, Keysha Naila Chadijah, Syilvana Efendi, dan Navy Nurlyn Ajrina yang telah menemani perjalanan akademik penulis, berbagi cerita, tawa, serta perjuangan selama masa perkuliahan.
10. Teman-teman Advokesma HIMASIFO periode 2023/2024 (Advokeos) dan periode 2024/2025 (AdvOMG), yang telah memberikan ruang bagi penulis untuk belajar, bertumbuh, serta mengembangkan diri melalui kebersamaan dan pengalaman organisasi.
11. Rekan-rekan satu angkatan penulis, Galaksi 2022, yang telah menjadi bagian dari perjalanan panjang perkuliahan, saling menguatkan dalam proses belajar, dan berbagi semangat hingga tahap akhir studi.
12. BTS (*Bangtan Sonyeondan*), yang melalui karya-karyanya telah menjadi sumber motivasi dan penguat semangat bagi penulis, khususnya di masa-masa sulit sehingga penulis dapat terus bertahan dan percaya diri untuk menuntaskan skripsi ini.
13. Rekan seperjuangan dalam bimbingan skripsi, Rifda Nasywatul Affah dan Nadiyah Syaidatus Shofa Abdul Hayat, yang telah menjadi rekan diskusi, berbagi informasi, serta saling menguatkan dalam setiap proses bimbingan hingga pelaksanaan yudisium. Kebersamaan dalam menghadapi revisi,

menunggu jadwal bimbingan, hingga saling menyemangati di tengah tekanan penyusunan skripsi menjadi bagian berharga dalam perjalanan ini.

14. Semua pihak yang tidak dapat disebutkan satu per satu, yang baik secara langsung maupun tidak langsung telah memberikan bantuan, kontribusi, dukungan, dan doa dalam proses penyusunan skripsi ini. Setiap bentuk kebaikan, sekecil apa pun, memiliki arti yang besar bagi perjalanan penulis hingga mencapai tahap ini.
15. Dan yang terakhir, diri penulis sendiri, atas keteguhan, keberanian, dan semangat yang tidak pernah benar-benar padam. Terima kasih telah bertahan, berproses, dan tidak menyerah meskipun jalan yang dilalui tidak selalu mudah. Hari ini bukan merupakan akhir, namun awal dari perjalanan panjang dalam hidup untuk bertumbuh menjadi pribadi yang lebih baik. Semoga setiap langkah ke depan senantiasa diiringi dengan kerendahan hati, rasa syukur, dan tekad untuk memberikan manfaat bagi orang lain serta membanggakan mereka yang telah percaya dan mendukung sejak awal.

Penulis menyadari bahwa skripsi ini masih memiliki sejumlah keterbatasan dan belum mencapai kesempurnaan sepenuhnya. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang bersifat konstruktif untuk perbaikan di masa yang akan datang. Semoga penelitian skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan serta menjadi kontribusi yang berarti bagi pihak-pihak yang berkepentingan.

Surabaya, 27 Februari 2026

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN BEBAS PLAGIASI	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI.....	xvii
DAFTAR GAMBAR	xxi
DAFTAR TABEL	xxv
DAFTAR SINGKATAN, ISTILAH, DAN SIMBOL	xxix
DAFTAR LAMPIRAN	xxxii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Batasan Masalah.....	7
1.4 Tujuan Penelitian	7
1.5 Manfaat Penelitian	8
1.6 Sistematikan Penulisan.....	8
BAB II TINJAUAN PUSTAKA	11
2.1 Dasar Teori.....	11
2.1.1 PLN Mobile.....	11
2.1.2 Google Play Store	12
2.1.3 Apple App Store.....	12
2.1.4 Python	13
2.1.5 <i>Text Mining</i>	14
2.1.6 <i>Aspect-Based Sentiment Analysis (ABSA)</i>	14
2.1.7 <i>Topic Modeling</i>	15
2.1.8 Latent Dirichlet Allocation (LDA).....	16
2.1.9 BERTopic.....	17
2.1.10 Support Vector Machine (SVM).....	18
2.1.11 Random Forest	20

2.1.12 <i>Web Scraping</i>	21
2.1.13 <i>Data Cleaning</i>	21
2.1.14 <i>Text Pre-processing</i>	21
2.1.15 Krippendorff's Alpha	23
2.1.16 Holdout.....	24
2.1.17 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	25
2.1.18 Teknik Penanganan Imbalanced Data	26
2.1.19 <i>Confusion Matrix</i>	28
2.1.20 <i>Hamming Loss</i>	31
2.1.21 <i>Area Under the Curve (AUC)</i>	32
2.1.22 Flask	32
2.2 Penelitian Terdahulu	33
BAB III METODOLOGI PENELITIAN.....	37
3.1 Identifikasi Masalah	37
3.2 Studi Literatur	38
3.3 Analisis Kebutuhan	38
3.3.1 Kebutuhan Perangkat Keras	38
3.3.2 Kebutuhan Perangkat Lunak	39
3.3.3 Kebutuhan Data.....	39
3.4 Pengumpulan Data	40
3.5 Eksplorasi Data	41
3.6 Pembersihan Data.....	42
3.7 <i>Text Pre-Processing</i>	42
3.7.1 <i>Text Cleaning</i>	43
3.7.2 <i>Case Folding</i>	43
3.7.3 <i>Tokenization</i>	44
3.7.4 <i>Normalization</i>	44
3.7.5 <i>Stopword Removal</i>	45
3.7.6 <i>Stemming</i>	45
3.8 <i>Topic Modeling</i>	46
3.8.1 Latent Dirichlet Allocation (LDA).....	47
3.8.2 BERTopic.....	47
3.9 Evaluasi dan Interpretasi Topik.....	48

3.10 Pelabelan Data.....	49
3.10.1 Pelabelan Aspek.....	49
3.10.2 Pelabelan Sentimen.....	49
3.11 Pembagian Data	51
3.12 Pembangunan Model Klasifikasi	51
3.12.1 Pembagian Data Model Klasifikasi.....	52
3.12.2 Pembobotan Kata	55
3.12.3 Penanganan Ketidakseimbangan Data	55
3.12.4 Pembentukan Model Klasifikasi	56
3.12.5 Hasil Pemodelan Klasifikasi	56
3.13 Evaluasi Performa Model.....	56
3.14 Analisis Model Terbaik.....	57
3.15 Penyimpanan Model.....	57
3.16 Implementasi Sistem	58
3.17 Pengujian Validasi Sistem.....	60
3.18 Penarikan Kesimpulan	61
3.19 Penyusunan Laporan.....	61
3.20 Jadwal Pelaksanaan Penelitian.....	62
BAB IV HASIL DAN PEMBAHASAN	63
4.1 Hasil	63
4.1.1 Identifikasi Masalah.....	63
4.1.2 Studi Literatur	64
4.1.3 Analisis Kebutuhan	64
4.1.4 Pengumpulan Data	66
4.1.5 Eksplorasi Data	72
4.1.6 Pembersihan Data.....	77
4.1.7 <i>Text Pre-Processing</i>	79
4.1.8 <i>Topic Modeling</i>	92
4.1.9 Evaluasi dan Interpretasi Topik	104
4.1.10 Pelabelan Data.....	113
4.1.11 Pembagian Data	122
4.1.12 Pembangunan Model Klasifikasi	124
4.1.13 Evaluasi Performa Model.....	139

4.1.14 Analisis Model Terbaik.....	141
4.1.15 Penyimpanan Model.....	144
4.1.16 Implementasi Sistem	146
4.1.17 Pengujian Validasi Sistem.....	149
4.2 Pembahasan.....	155
BAB V PENUTUP	163
5.1 Kesimpulan.....	163
5.2 Saran.....	164
DAFTAR PUSTAKA.....	166
LAMPIRAN	177

DAFTAR GAMBAR

Gambar 1.1 Digital Indonesia 2025 oleh We Are Social	1
Gambar 1.2 <i>Rating</i> Aplikasi PLN Mobile pada Google Play Store dan Apple App Store	2
Gambar 1.3 Ulasan Pengguna Aplikasi PLN Mobile pada Google Play Store dan Apple App Store	3
Gambar 2.1 Tampilan Aplikasi PLN Mobile	11
Gambar 2.2 Representasi Model Grafis dari LDA [37]	16
Gambar 2.3 BERTopic Flowchart [41]	17
Gambar 2.4 Ilustrasi Hyperplane Pemisah pada Ruang Fitur [45]	19
Gambar 2.5 Visualisasi <i>Margin</i> dan <i>Support Vector</i> [45]	19
Gambar 2.6 Representasi Algoritma Random Forest [16]	20
Gambar 3.1 Alur Penelitian.....	37
Gambar 3.2 Tahapan Pengumpulan Data	40
Gambar 3.3 Alur Proses <i>Text Pre-processing</i>	43
Gambar 3.4 Tahapan Proses <i>Text Cleaning</i>	43
Gambar 3.5 Tahapan Proses <i>Case Folding</i>	43
Gambar 3.6 Tahapan Proses <i>Tokenization</i>	44
Gambar 3.7 Tahapan Proses <i>Normalization</i>	44
Gambar 3.8 Tahapan Proses <i>Stopword Removal</i>	45
Gambar 3.9 Tahapan Proses <i>Stemming</i>	45
Gambar 3.10 Tahapan <i>Topic Modeling</i>	46
Gambar 3.11 Alur Pembangunan Model Klasifikasi	52
Gambar 3.12 <i>Wireframe</i> Analisis Teks.....	59
Gambar 3.13 <i>Wireframe</i> Analisis File CSV	60
Gambar 3.14 Jadwal Pelaksanaan Penelitian	62
Gambar 4.1 Kode <i>Scraping</i> Ulasan PLN Mobile pada Google Play Store.....	66
Gambar 4.2 Kode <i>Scraping</i> Ulasan PLN Mobile pada Apple App Store	69
Gambar 4.3 Kode Deteksi Bahasa Ulasan PLN Mobile pada Apple App Store	69
Gambar 4.4 Kode Penggabungan Dataset Ulasan PLN Mobile	71
Gambar 4.5 Kode Analisis Deskriptif Dataset	73
Gambar 4.6 Distribusi Ulasan Berdasarkan Sumber Data	75

Gambar 4.7 Visualisasi Distrbusi <i>Rating</i> Ulasan Pengguna.....	76
Gambar 4.8 <i>Word Cloud</i> Ulasan Pengguna PLN Mobile	76
Gambar 4.9 Kode Eliminasi Data Duplikat	77
Gambar 4.10 Kode Eliminasi <i>Missing Values</i>	78
Gambar 4.11 Kode Eliminasi Ulasan yang Terlalu Pendek	79
Gambar 4.12 Kode Implementasi <i>Text Cleaning</i>	81
Gambar 4.13 Kode Implementasi <i>Case Folding</i>	82
Gambar 4.14 Kode Implementasi <i>Tokenization</i>	84
Gambar 4.15 Kode Implementasi <i>Normalization</i>	85
Gambar 4.16 Kode Implementasi <i>Stopword Removal</i> Tahap Pertama.....	87
Gambar 4.17 Kode Analisis Frekuensi Kata	87
Gambar 4.18 Kode Implementasi <i>Stopword Removal</i> Tahap Kedua	88
Gambar 4.19 Kode Implementasi <i>Stemming</i>	90
Gambar 4.20 <i>Word Cloud</i> Hasil <i>Text Pre-Processing</i>	91
Gambar 4.21 Kode Implementasi Pelatihan Model LDA	93
Gambar 4.22 Kode Evaluasi <i>Coherence Score</i> LDA	97
Gambar 4.23 Kode Implementasi Pelatihan Model BERTopic.....	99
Gambar 4.24 Kode Evaluasi <i>Coherence Score</i> BERTopic	103
Gambar 4.25 Visualisasi Perbandingan <i>Coherence Score</i> LDA dan BERTopic .	105
Gambar 4.26 <i>Intertopic Distance Map</i> LDA.....	106
Gambar 4.27 <i>Intertopic Distance Map</i> BERTopic	106
Gambar 4.28 Visualisasi Probabilitas Kata Topik 0	109
Gambar 4.29 Visualisasi Probabilitas Kata Topik 1	109
Gambar 4.30 Visualisasi Top 10 Skor c-TF-IDF Topik 0	111
Gambar 4.31 Visualisasi Top 10 Skor c-TF-IDF Topik 1	111
Gambar 4.32 Visualisasi Top 10 Skor c-TF-IDF Topik 2	112
Gambar 4.33 Kode Implementasi Krippendorff's Alpha	118
Gambar 4.34 Visualisasi Distribusi Sentimen per Aspek pada Model LDA.....	121
Gambar 4.35 Visualisasi Distribusi Sentimen per Aspek pada Model BERTopic	122
Gambar 4.36 Kode Implementasi Pembagian Data <i>Modeling</i> dan <i>Validation</i>	123
Gambar 4.37 Perbandingan <i>Confusion Matrix</i> pada Kelas Minoritas	143
Gambar 4.38 Kode Penyimpanan Model Terbaik	145
Gambar 4.39 Kode Inisialisasi Flask.....	146

Gambar 4.40 Kode Load Model dan Komponen	147
Gambar 4.41 Tampilan Utama Sistem	147
Gambar 4.42 Halaman Analisis Teks	148
Gambar 4.43 Tampilan Analisis File CSV	149
Gambar 4.44 Kode Pengujian Model Terbaik Menggunakan Data Validasi	150
Gambar 4.45 Tampilan Hasil Prediksi Ulasan Teks	152
Gambar 4.46 Tampilan Hasil Prediksi Ulasan File CSV	153
Gambar 4.47 Tampilan Visualisasi Hasil Analisis (<i>Bar Chart</i> dan <i>Pie Chart</i>)....	154
Gambar 4.48 Tampilan Tabel Detail <i>Bar Chart</i> dan <i>Pie Chart</i>	154
Gambar 4.49 Tampilan <i>Line Chart</i> Tren Sentimen Tiap Aspek.....	155
Gambar 4.50 Perbandingan <i>Confusion Matrix</i> Kelas Minoritas.....	158

DAFTAR TABEL

Tabel 2.1 Komponen <i>Confusion Matrix</i> [55]	28
Tabel 2.2 Kriteria Nilai AUC [79].....	32
Tabel 2.3 Penelitian Terdahulu.....	33
Tabel 3.1 Interpretasi Nilai Krippendorff's Alpha [59].....	50
Tabel 3.2 Skenario Pengujian.....	53
Tabel 4.1 Spesifikasi Perangkat Keras	65
Tabel 4.2 Contoh Hasil Ulasan PLN Mobile pada Google Play Store.....	67
Tabel 4.3 Contoh Hasil Ulasan PLN Mobile pada Apple App Store	70
Tabel 4.4 Contoh Hasil Ulasan PLN Mobile pada Kedua Platform	71
Tabel 4.5 Hasil Analisis Deskriptif Dataset	73
Tabel 4.6 Hasil Statistik Kolom Numerik dari Dataset.....	74
Tabel 4.7 Jumlah Ulasan Berdasarkan <i>Rating</i>	75
Tabel 4.8 Hasil Eliminasi Duplikasi Data	78
Tabel 4.9 Hasil Eliminasi <i>Missing Values</i>	78
Tabel 4.10 Hasil Eliminasi Ulasan yang Terlalu Pendek	79
Tabel 4.11 Contoh Hasil Implementasi <i>Text Cleaning</i>	81
Tabel 4.12 Contoh Hasil Implementasi <i>Case Folding</i>	82
Tabel 4.13 Contoh Hasil Implementasi <i>Tokenization</i>	84
Tabel 4.14 Contoh Hasil Implementasi <i>Normalization</i>	86
Tabel 4.15 Analisis Kata Paling Sering Muncul dari Stopword Pertama	88
Tabel 4.16 Contoh Hasil Implementasi <i>Stopword Removal</i>	89
Tabel 4.17 Contoh Hasil Implementasi <i>Stemming</i>	90
Tabel 4.18 Hasil Eliminasi <i>Missing Values</i>	91
Tabel 4.19 Daftar Topik untuk Setiap Jumlah Topik LDA	93
Tabel 4.20 Hasil Evaluasi <i>Coherence Score</i> untuk Setiap Jumlah Topik LDA	97
Tabel 4.21 Daftar Topik untuk Setiap Jumlah Topik BERTopic	100
Tabel 4.22 Hasil Evaluasi <i>Coherence Score</i> untuk Setiap Jumlah Topik BERTopic.....	103
Tabel 4.23 Hasil Perbandingan <i>Coherence Score</i> LDA dan BERTopic.....	104
Tabel 4.24 Hasil Topik Optimal LDA dan BERTopic.....	108
Tabel 4.25 Identifikasi Fokus Utama dan Aspek setiap Topik LDA.....	110
Tabel 4.26 Identifikasi Fokus Utama dan Aspek setiap Topik BERTopic	113

Tabel 4.27 Contoh Pengisian Pelabelan Aspek LDA	114
Tabel 4.28 Contoh Pengisian Pelabelan Aspek BERTopic.....	114
Tabel 4.29 Hasil Perhitungan Krippendorff's Alpha pada Pelabelan Aspek LDA dan BERTopic	115
Tabel 4.30 Hasil Majority Voting pada Pelabelan Aspek LDA.....	116
Tabel 4.31 Hasil Majority Voting pada Pelabelan Aspek BERTopic	116
Tabel 4.32 Contoh Pengisian Pelabelan Sentimen LDA	117
Tabel 4.33 Contoh Pengisian Pelabelan Sentimen BERTopic.....	118
Tabel 4.34 Hasil Perhitungan Krippendorff's Alpha pada Pelabelan Sentimen...	119
Tabel 4.35 Hasil <i>Majority Voting</i> pada Pelabelan Sentimen LDA	119
Tabel 4.36 Hasil <i>Majority Voting</i> pada Pelabelan Sentimen BERTopic.....	120
Tabel 4.37 Hasil Pembagian Data	124
Tabel 4.38 Evaluasi Skenario 1	125
Tabel 4.39 Evaluasi Skenario 2	126
Tabel 4.40 Evaluasi Skenario 3	127
Tabel 4.41 Evaluasi Skenario 4	128
Tabel 4.42 Evaluasi Skenario 5	129
Tabel 4.43 Evaluasi Skenario 6	129
Tabel 4.44 Evaluasi Skenario 7	130
Tabel 4.45 Evaluasi Skenario 8	131
Tabel 4.46 Evaluasi Skenario 9	132
Tabel 4.47 Evaluasi Skenario 10	133
Tabel 4.48 Evaluasi Skenario 11	134
Tabel 4.49 Evaluasi Skenario 12	135
Tabel 4.50 Evaluasi Skenario 13	136
Tabel 4.51 Evaluasi Skenario 14	137
Tabel 4.52 Evaluasi Skenario 15	138
Tabel 4.53 Evaluasi Skenario 16	139
Tabel 4.54 Rekapitulasi Performa 16 Skenario Pengujian	140
Tabel 4.55 Perbandingan Performa Beberapa Skenario Terbaik.....	142
Tabel 4.56 Perbandingan Performa Kelas Minoritas (<i>Service Quality Negative</i>)	142
Tabel 4.57 Evaluasi Performa Model Menggunakan Data Validasi.....	150
Tabel 4.58 Perbandingan Performa Model pada Data Test dan Validasi.....	151

Tabel 4.59 Perbandingan Performa Seluruh Skenario Model.....	156
Tabel 4.60 Perbandingan Performa Kelas Minoritas (<i>Service Quality Negative</i>)	157
Tabel 4.61 Contoh Hasil Prediksi Model Terbaik pada Data Validasi	159

DAFTAR SINGKATAN, ISTILAH, DAN SIMBOL

SINGKATAN/ ISTILAH/ SIMBOL	KETERANGAN
α	: Koefisien tingkat kesepakatan Krippendorff's Alpha
Do	: Tingkat ketidaksepakatan yang diamati (<i>observed disagreement</i>)
De	: Tingkat ketidaksepakatan yang diharapkan terjadi secara kebetulan (<i>expected disagreement</i>)
t	: Kata yang frekuensinya dihitung.
d	: Dokumen yang sedang dianalisis.
$f_{t,d}$	: Banyaknya kemunculan kata t dalam dokumen d .
$\sum_k f_{k,d}$	: Total kata yang terdapat dalam dokumen d .
D	: Sekumpulan dokumen atau <i>corpus</i> .
N	: Jumlah keseluruhan dokumen dalam <i>corpus</i> .
$d_f(t, D)$	: Banyak dokumen dalam <i>corpus</i> D yang memuat kata t .
TP	: Jumlah data positif yang benar-benar diprediksi sebagai positif.
TN	: Jumlah data negatif yang benar-benar diprediksi sebagai negatif.
FP	: Jumlah data negatif yang keliru diprediksi sebagai positif.
FN	: Jumlah data positif yang keliru diprediksi sebagai negatif.
L	: Jumlah label.
y_{ij}	: Label sebenarnya untuk data ke- i dan label ke- j .
$\hat{y}_{ij}$	: Label hasil prediksi untuk data ke- i dan label ke- j

DAFTAR LAMPIRAN

Lampiran 1. Surat Pengantar Penelitian Skripsi	177
Lampiran 2. Izin Pengambilan Data Ulasan	178
Lampiran 3. Kode Implementasi Sistem ABSA PLN Mobile.....	179