



SKRIPSI

**KOMPARASI KINERJA MODEL DEEP
LEARNING DAN WORD EMBEDDING DALAM
MULTILABEL ASPECT-BASED SENTIMENT
ANALYSIS PADA ULASAN HOTEL**

NAVY NURLYN AJRINA

NPM 22082010158

DOSEN PEMBIMBING

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

Reisa Permatasari, S.T., 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

KOMPARASI KINERJA MODEL DEEP LEARNING DAN WORD EMBEDDING DALAM MULTILABEL ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN HOTEL

Oleh :

NAVY NURLYN AJRINA

NPM. 22082010158

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 13 Mei 2026


Menyetujui

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


.....

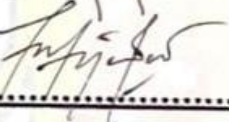
(Pembimbing I)

Reisa Permatasari, S.T, M.Kom.
NIP. 199205142022032007


.....

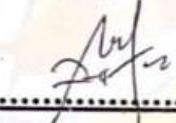
(Pembimbing II)

Seftin Fitri Ana Wati, S.Kom, M.Kom
NPT. 21219910320267


.....

(Ketua Penguji)

Anindo Saka Fitri, S.Kom, M.Kom
NIP. 199303252024062001


.....

(Penguji II)

Abdul Rezha Efrat Najaf, S.Kom, M.Kom
NIP. 199409292022031008


.....

(Penguji III)

Mengetahui,
Dekan Fakultas Ilmu Komputer


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

LEMBAR PERSETUJUAN

KOMPARASI KINERJA MODEL DEEP LEARNING DAN WORD EMBEDDING DALAM MULTILABEL ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN HOTEL

Oleh :

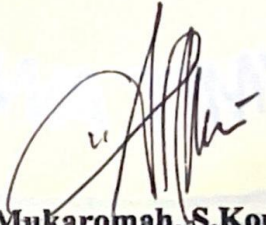
NAVY NURLYN AJRINA

NPM. 22082010158

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 bertandatangan di bawah ini:

Nama Mahasiswa : Navy Nurlyn Ajrina
NPM : 22082010158
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, 13 Mei 2026

Yang Membuat Pernyataan,



NAVY NURLYN AJRINA

NPM. 22082010158

ABSTRAK

Nama Mahasiswa / NPM : Navy Nurlyn Ajrina / 22082010158
Judul Skripsi : Komparasi Kinerja Model Deep Learning dan Word Embedding dalam Multilabel Aspect-Based Sentiment Analysis pada Ulasan Hotel
Dosen Pembimbing : 1. Eka Dyar Wahyuni, S.Kom., M.Kom.
2. Reisa Permatasari, S.T., M.Kom.

Ulasan pengunjung pada platform digital seperti Tripadvisor merupakan sumber informasi yang penting bagi industri perhotelan. Namun, jumlah ulasan yang besar dan tidak terstruktur menyebabkan analisis manual menjadi sulit dilakukan. Penelitian ini bertujuan untuk menganalisis sentimen pada aspek *Room*, *Hotel*, *Location*, dan *Service* menggunakan pendekatan *Multilabel Aspect-Based Sentiment Analysis* (ABSA), serta mengomparasikan kinerja arsitektur *deep learning*, *word embedding*, dan *preprocessing* pada ulasan hotel di Surabaya, Sukabumi, dan Bali. Dataset yang digunakan terdiri dari 7.821 ulasan berbahasa Indonesia dan Inggris yang diproses melalui tahap translasi ke Bahasa Indonesia. Model dikembangkan menggunakan *5-Fold Cross Validation*, mencakup satu model klasifikasi aspek multilabel dan empat model klasifikasi sentimen per aspek. Penelitian ini membandingkan kombinasi arsitektur *deep learning* (CNN dan BiLSTM), *word embedding* (Word2Vec dan FastText), serta teknik *preprocessing* (dengan atau tanpa *stemming*). Hasil penelitian menunjukkan bahwa tidak terdapat satu kombinasi model yang secara konsisten unggul pada seluruh tugas klasifikasi (Task). BiLSTM dengan FastText dan *stemming* menghasilkan performa terbaik pada task klasifikasi aspek multilabel dengan *F1 Macro* 0.9725. Pada task klasifikasi sentimen, CNN cenderung lebih unggul, dengan kombinasi terbaik berupa CNN + FastText + *stemming* untuk aspek *Room* (*F1 Macro* 0.9882), BiLSTM + Word2Vec tanpa *stemming* untuk aspek *Hotel* (*F1 Macro* 0.9516), CNN + Word2Vec tanpa *stemming* untuk aspek *Location* (*F1 Macro* 0.9676), serta CNN + Word2Vec + *stemming* untuk aspek *Service* (*F1 Macro* 0.9725). Hasil penelitian ini menunjukkan bahwa BiLSTM lebih efektif untuk task klasifikasi aspek multilabel, sedangkan CNN lebih efektif untuk task klasifikasi sentimen aspek. Word2Vec cenderung memberikan performa yang lebih baik pada klasifikasi sentimen aspek *Hotel*, *Location*, dan *Service*, sementara pengaruh *stemming* bersifat tidak konsisten dan bergantung pada karakteristik masing-masing tugas. Kelima model terbaik selanjutnya diimplementasikan ke dalam sistem berbasis web untuk melakukan analisis ulasan hotel secara otomatis dan menyediakan visualisasi interaktif sebagai bukti konsep penerapan praktis untuk mendukung pengambilan keputusan.

Kata Kunci: Analisis Sentimen Berbasis Aspek, CNN, BiLSTM, Word2Vec, FastText, *Stemming*

ABSTRACT

Student Name / NPM : Navy Nurlyn Ajrina / 22082010158
Thesis Title : Performance Comparison of Deep Learning Models and Word Embeddings in Multilabel Aspect-Based Sentiment Analysis on Hotel Reviews
Advisors : 1. Eka Dyar Wahyuni, S.Kom., M.Kom.
2. Reisa Permatasari, S.T., M.Kom.

Visitor reviews on digital platforms such as Tripadvisor constitute an important source of information for the hospitality industry. However, the large volume and unstructured nature of these reviews make manual analysis difficult. This study aims to analyze sentiment toward the Room, Hotel, Location, and Service aspects using a Multilabel Aspect-Based Sentiment Analysis (ABSA) approach, as well as to compare the performance of deep learning architectures, word embeddings, and preprocessing techniques on hotel reviews from Surabaya, Sukabumi, and Bali. The dataset consisted of 7,821 hotel reviews written in Indonesian and English, where English-language reviews were translated into Indonesian to maintain linguistic consistency during analysis. The models were developed using 5-Fold Cross Validation, comprising one multilabel aspect classification model and four aspect-based sentiment classification models. This study compared combinations of deep learning architectures (CNN and BiLSTM), word embeddings (Word2Vec and FastText), and preprocessing techniques (with and without stemming). The results indicate that no single model combination consistently outperformed others across all classification tasks. BiLSTM with FastText and stemming achieved the best performance for multilabel aspect classification with an F1 Macro score of 0.9725. For sentiment classification, CNN tended to perform better, with the best-performing combinations being CNN + FastText + stemming for the Room aspect (F1 Macro 0.9882), BiLSTM + Word2Vec without stemming for Hotel (F1 Macro 0.9516), CNN + Word2Vec without stemming for Location (F1 Macro 0.9676), and CNN + Word2Vec + stemming for Service (F1 Macro 0.9725). These findings indicate that BiLSTM is more effective for multilabel aspect classification, whereas CNN performs better for sentiment classification. Word2Vec tended to provide better performance for sentiment classification on the Hotel, Location, and Service aspects, while the effect of stemming was inconsistent and dependent on the characteristics of each task. The five best-performing models were subsequently implemented in a web-based system to automatically analyze hotel reviews and provide interactive visualizations as a proof of concept for practical implementation.

Keyword: *Aspect-Based Sentiment Analysis*, CNN, BiLSTM, Word2Vec, FastText, *Stemming*

KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas segala rahmat, kasih sayang, dan karunia-Nya yang senantiasa menyertai setiap langkah penulis, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Komparasi Kinerja Model Deep Learning dan Word embedding dalam Multilabel Aspect-Based Sentiment Analysis pada Ulasan Hotel”. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa penelitian ini tidak terlepas dari bantuan, dukungan, serta doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada:

1. Kedua orang tua tercinta, Terima kasih atas doa yang tidak pernah terputus, dukungan yang selalu diberikan, serta kasih sayang yang menjadi kekuatan terbesar bagi penulis. Terima kasih karena selalu menjadi tempat pulang yang paling nyaman, dan selalu percaya bahwa penulis mampu melewati semua ini dengan baik. Segala pencapaian ini tidak terlepas dari doa dan dukungan dari Ibu dan Ayah.
2. Ibu Eka Dyar Wahyuni, S.Kom., M.Kom. dan Ibu Reisa Permatasari, S.Kom., M.Kom. selaku dosen pembimbing yang telah meluangkan waktu, membimbing dengan penuh kesabaran, serta memberikan arahan dan masukan yang sangat berarti selama proses penyusunan skripsi ini.
3. Seluruh jajaran dosen Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, yang telah memberikan ilmu, pengalaman, serta wawasan yang sangat berharga selama masa perkuliahan, yang menjadi bekal penting bagi penulis hingga saat ini.
4. Kedua adik penulis, Oryza dan Arawinda. Terima kasih telah menjadi bagian dari kekuatan penulis dalam menyelesaikan skripsi ini. Terima kasih atas kebersamaan, dukungan, serta semangat yang diberikan kepada penulis selama proses penyusunan skripsi ini.

5. Keluarga besar penulis, termasuk yangti, serta seluruh saudara yang selalu memberikan doa, dukungan, dan perhatian yang tulus kepada penulis.
6. Sahabat terdekat penulis, Fathimatus Zahrotun Nisa', Ganes Dwi Febrianti, Yuliani Puwitasari, Putri Intan Octavia Br. Sipayung, yang telah menjadi tempat berbagi cerita, memberikan dukungan tanpa henti, serta menemani penulis dalam setiap fase selama masa perkuliahan. Terima kasih karena membuat perjalanan ini terasa lebih ringan dan bermakna.
7. Teman seperjuangan skripsi, Jihan Hasna Iftinan, yang telah saling mendukung, membantu, dan menguatkan satu sama lain dalam proses penyusunan skripsi ini.
8. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan bantuan baik secara langsung maupun tidak langsung dalam proses penyusunan skripsi ini.
9. Terakhir, penulis ingin berterima kasih kepada diri sendiri yang telah bertahan, terus berusaha, dan tidak menyerah dalam melewati setiap proses hingga sampai pada titik ini. Perjalanan ini tidak selalu mudah, namun setiap langkah yang telah dilalui menjadi bagian berharga dari proses belajar dan bertumbuh. Setiap proses dan pembelajaran yang dilalui menjadi pengingat bahwa usaha yang dilakukan dengan bersungguh-sungguh tidak pernah sia-sia. Semoga pencapaian ini menjadi awal dari perjalanan dan kesempatan yang lebih baik di masa depan.

Penulis menyadari bahwa skripsi ini masih memiliki berbagai keterbatasan. Oleh karena itu, penulis sangat terbuka terhadap kritik dan saran yang membangun sebagai bahan perbaikan di masa mendatang. Akhir kata, penulis berharap skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya di bidang Sistem Informasi, serta memberikan kontribusi bagi kemajuan Universitas Pembangunan Nasional "Veteran" Jawa Timur.

Surabaya, 13 Mei 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 GAMBAR	xix
DAFTAR TABEL	xxiii
DAFTAR LAMPIRAN	xxv
DAFTAR SINGKATAN, ISTILAH DAN SIMBOL	xxvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	8
1.3 Batasan Masalah.....	8
1.4 Tujuan Penelitian	9
1.5 Manfaat Penelitian	10
1.6 Sistematikan Penulisan.....	10
BAB II TINJAUAN PUSTAKA	13
2.1 Dasar Teori.....	13
2.1.1 Industri Perhotelan.....	13
2.1.2 Tripadvisor.....	14
2.1.3 <i>Web Scraping</i>	15
2.1.4 <i>Text Mining</i>	16
2.1.5 <i>Aspect-Based Sentiment Analysis</i>	16
2.1.6 <i>Text Preprocessing</i>	17
2.1.7 Krippendorff's Alpha.....	19
2.1.8 <i>Word embedding</i>	21
2.1.9 Word2Vec	22

2.1.10 FastText.....	23
2.1.11 <i>Deep learning</i>	23
2.1.12 <i>Convolutional Neural Network (CNN)</i>	24
2.1.13 <i>Long-Short Term Memory (LSTM)</i>	27
2.1.14 <i>Bidirectional Long-Short Term Memory (BiLSTM)</i>	29
2.1.15 K-fold Cross Validation.....	31
2.1.16 Confusion Matrix.....	31
2.1.17 AUC-PR.....	33
2.1.18 Flask.....	34
2.2 Penelitian Terdahulu.....	35
BAB III METODOLOGI PENELITIAN.....	41
3.1 Identifikasi Masalah.....	41
3.2 Studi Literatur.....	42
3.3 Analisis Kebutuhan.....	42
3.3.1 Kebutuhan Perangkat Keras (<i>Hardware</i>).....	42
3.3.2 Kebutuhan Perangkat Lunak (<i>Software</i>).....	43
3.3.3 Kebutuhan Data.....	43
3.4 Pengumpulan Data.....	43
3.5 Eksplorasi Data (<i>Exploratory Data Analysis</i>).....	45
3.6 Pembersihan Data (<i>Data Cleaning</i>).....	46
3.7 Penerjemahan Bahasa (<i>Translation</i>).....	46
3.8 Pelabelan Data.....	47
3.9 <i>Text Pre-processing</i>	50
3.9.1 <i>Text Cleaning</i>	50
3.9.2 <i>Case Folding</i>	50
3.9.3 <i>Normalization</i>	51
3.9.4 <i>Stopword Removal</i>	51
3.9.5 <i>Stemming</i>	52
3.10 Pembagian Data.....	52
3.10.1 Subset Data Klasifikasi Aspek Multilabel.....	54
3.10.2 Subset Data Klasifikasi Sentimen Per Aspek.....	54
3.11 <i>Word embedding</i>	55

3.12 Pengembangan Model ABSA	55
3.12.1 Skenario Pelatihan	56
3.12.2 Tokenisasi dan Representasi Teks	57
3.12.3 Arsitektur Model <i>Deep learning</i>	57
3.12.4 Training & Validation Model	59
3.13 <i>Retraining</i> dan Penyimpanan Model.....	60
3.14 Pengujian Model	60
3.15 <i>Deployment</i> dan Implementasi Sistem.....	61
3.16 Penarikan Kesimpulan	62
BAB IV HASIL DAN PEMBAHASAN.....	63
4.1 Hasil	63
4.1.1 Identifikasi Masalah.....	63
4.1.2 Studi Literatur	63
4.1.3 Analisis Kebutuhan.....	64
4.1.4 Pengumpulan Data	65
4.1.5 Eksplorasi Data	67
4.1.6 Pembersihan Data	72
4.1.7 Penerjemahan Data	74
4.1.8 Pelabelan Data	76
4.1.9 <i>Text Pre-Processing</i>	85
4.1.10 Pembagian Data	94
4.1.11 <i>Word embedding</i>	98
4.1.12 Pengembangan Model ABSA.....	102
4.1.13 <i>Training</i> dan <i>Validation</i>	109
4.1.14 <i>Retraining</i> dan Penyimpanan Model	128
4.1.15 Pengujian Model	130
4.1.16 Deployment.....	146
4.2 Pembahasan.....	150
BAB V PENUTUP	167
5.1 Kesimpulan	167
5.2 Saran.....	170
DAFTAR PUSTAKA.....	171

LAMPIRAN	181
-----------------------	------------

DAFTAR GAMBAR

Gambar 1. 1 Grafik Perkembangan TPK Hotel di Indonesia [9].....	2
Gambar 1. 2 Rating Ulasan Hotel Oval di Platform Tripadvisor [21].....	3
Gambar 1. 3 Ulasan Pengunjung pada Platform Tripadvisor [21].....	5
Gambar 2. 1 Arsitektur Model CNN untuk Klasifikasi Teks [61].....	25
Gambar 2. 2 Arsitektur Model LSTM [63].....	27
Gambar 2. 3 Arsitektur Model Bi-LSTM [67].....	29
Gambar 2. 4 Confusion Matrix [69].....	31
Gambar 3. 1 Alur Pelaksanaan Penelitian.....	41
Gambar 3. 2 Alur Proses Penerjemahan Ulasan	46
Gambar 3. 3 Alur Proses Pelabelan Aspek dan Sentimen	47
Gambar 3. 4 Alur Tahap Text <i>Preprocessing</i>	50
Gambar 3. 5 Alur Pembentukan Subset Data	53
Gambar 3. 6 Alur Pengembangan Model.....	56
Gambar 3. 7 Wireframe Analisis Single Teks	61
Gambar 3. 8 Wireframe Analisis Batch Ulasan	62
Gambar 3. 9 Wireframe Visualisasi Data Hasil Analisis File Excel/CSV.....	62
Gambar 4. 1 Tampilan Platform Apify untuk <i>Scraping</i> Ulasan Hotel.....	66
Gambar 4. 2 Kode Perubahan Kolom Dataset.....	67
Gambar 4. 3 Dataset Ulasan Hotel	67
Gambar 4. 4 Kode Program Analisis Deskriptif Dataset	68
Gambar 4. 5 Visualisasi Jumlah Ulasan per Region	69
Gambar 4. 6 Visualisasi Jumlah Ulasan Berdasarkan Rating	70
Gambar 4. 7 Visualisasi Distribusi Bahasa dalam Ulasan	70
Gambar 4. 8 Visualisasi Distribusi Panjang Teks Ulasan.....	71
Gambar 4. 9 Visualisasi WorCloud Ulasan Hotel	72
Gambar 4. 10 Kode Cek Ulasan Tanpa Makna.....	72
Gambar 4. 11 kode Eliminasi Ulasan Pendek	73
Gambar 4. 12 kode Eliminasi Ulasan Duplikat.....	73
Gambar 4. 13 Kode Penerjemahan Ulasan	74

Gambar 4. 14 Pelabelan oleh Anotator Menggunakan Excel.....	78
Gambar 4. 15 Hasil Gabungan Pelabelan Ketiga Anotator	79
Gambar 4. 16 kode Perhitungan Krippendorff's Apha.....	79
Gambar 4. 17 Kode Implementasi <i>Majority voting</i>	80
Gambar 4. 18 Dataset Pelabelan Final.....	81
Gambar 4. 19 Grafik Distribusi Label Final Seluruh Aspek	81
Gambar 4. 20 Wordcloud Aspek <i>Room</i>	83
Gambar 4. 21 Wordcloud Aspek Hotel	83
Gambar 4. 22 Wordcloud Aspek <i>Location</i>	84
Gambar 4. 23 Wordcloud Aspek <i>Service</i>	84
Gambar 4. 24 Kode Implementasi Proses Text Cleaning.....	85
Gambar 4. 25 Kode Implementasi Proses <i>Case Folding</i>	87
Gambar 4. 26 Kode Implementasi Proses <i>Normalization</i>	88
Gambar 4. 27 Kode Normalisasi Frasa Hotel.....	89
Gambar 4. 28 Kode <i>Custom Normalization Dictionary</i>	89
Gambar 4. 29 Kode Implementasi dan load <i>new_kamusalay.csv</i>	90
Gambar 4. 30 Kode Implementasi Proses <i>Stopword Removal</i> Tahap 1	91
Gambar 4. 31 Kode Implementasi Proses <i>Stopword Removal</i> Tahap 2	91
Gambar 4. 32 Kode Implementasi Proses <i>Stemming</i>	93
Gambar 4. 33 Contoh Konversi Label ke Bentuk Biner.....	94
Gambar 4. 34 Kode Implementasi Proses Pembagian Dataset.....	95
Gambar 4. 35 Kode Persiapan Implementasi <i>Word Embedding</i>	98
Gambar 4. 36 Kode Pelatihan <i>Word embedding</i> Word2Vec.....	99
Gambar 4. 37 Kode Pelatihan <i>Word embedding</i> FastText	99
Gambar 4. 38 Kode Pembuatan Tokenizer.....	103
Gambar 4. 39 Kode Pembuatan Embedding Matrix	104
Gambar 4. 40 Diagram Arsitektur CNN.....	105
Gambar 4. 41 Kode Implementasi Arsitektur CNN	106
Gambar 4. 42 Diagram Arsitektur BiLSTM.....	107
Gambar 4. 43 Kode Implementasi Arsitektur BiLSTM	109
Gambar 4. 44 Kode Implementasi Training Model.....	110
Gambar 4. 45 Classification Report Model Terbaik Klasifikasi Aspek	131

Gambar 4. 46 Confusion Matrix Model Terbaik Klasifikasi Aspek Multilabel .	132
Gambar 4. 47 PR Curve Model Terbaik Klasifikasi Aspek Multilabel	133
Gambar 4. 48 Classification Report Model Terbaik Sentimen <i>Room</i>	134
Gambar 4. 49 Confusion Matrix Model Terbaik Klasifikasi Sentimen <i>Room</i>	135
Gambar 4. 50 PR Curve Model Terbaik Klasifikasi Sentimen <i>Room</i>	135
Gambar 4. 51 Classification Report Model Terbaik Sentimen Hotel	137
Gambar 4. 52 Confusion Matrix Model Terbaik Klasifikasi Sentimen Hotel	138
Gambar 4. 53 PR Curve Model Terbaik Klasifikasi Sentimen Hotel	138
Gambar 4. 54 Classification Report Model Terbaik Sentimen <i>Location</i>	140
Gambar 4. 55 Confusion Matrix Model Terbaik Klasifikasi Sentimen <i>Location</i>	141
Gambar 4. 56 PR Curve Model Terbaik Klasifikasi Sentimen <i>Location</i>	141
Gambar 4. 57 Classification Report Model Terbaik Sentimen <i>Service</i>	143
Gambar 4. 58 Confusion Matrix Model Terbaik Klasifikasi Sentimen <i>Service</i> .	144
Gambar 4. 59 PR Curve Model Terbaik Klasifikasi Sentimen <i>Service</i>	144
Gambar 4. 60 Tampilan Halaman Analisis Teks	147
Gambar 4. 61 Tampilan Hasil Prediksi Analisis Teks	147
Gambar 4. 62 Tampilan Halaman Analisis File.....	148
Gambar 4. 63 Tampilan Tabel Hasil Analisis File.....	148
Gambar 4. 64 Tampilan Dashboard Visualisasi Hasil Analisis File	149
Gambar 4. 65 Tampilan Fitur Deep Dive per Aspek Hasil Analisis File	150
Gambar 4. 66 Confusion Matrix Skenario 5 Model Klasifikasi Aspek	151
Gambar 4. 67 Confusion Matrix Skenario 8 Model Klasifikasi Aspek	151
Gambar 4. 68 Confusion Matrix 2 Skenario Terbaik Model Sentimen <i>Room</i>	152
Gambar 4. 69 Confusion Matrix 2 Skenario Terbaik Model Sentimen Hotel	153
Gambar 4. 70 Confusion Matrix 2 Skenario Terbaik Model Sentimen <i>Location</i>	153
Gambar 4. 71 Confusion Matrix 2 Skenario Terbaik Model Sentimen <i>Service</i> .	154

DAFTAR TABEL

Tabel 2. 1 Interpretasi Nilai Krippendorff's Alpha [53]	20
Tabel 2. 2 Penelitian Terdahulu	35
Tabel 3. 1 Struktur Data Platform Tripadvisor	44
Tabel 3. 2 Daftar Item Setiap Aspek dalam Ulasan Hotel [28]	48
Tabel 3. 3 Skenario Penelitian	56
Tabel 4. 1 Hasil Analisis Deskriptif Dataset.....	68
Tabel 4. 2 Hasil Statistik Kolom Numerik Dataset.....	69
Tabel 4. 3 Jumlah Total Data Setelah Pembersihan Data	74
Tabel 4. 4 Contoh Hasil Penerjemahan Data	75
Tabel 4. 5 Daftar Item Setiap Aspek dalam Ulasan Hotel [28]	76
Tabel 4. 6 Hasil Perhitungan Krippendorff's Alpha.....	80
Tabel 4. 7 Contoh Hasil Implementasi Proses <i>Text Cleaning</i>	86
Tabel 4. 8 Contoh Hasil Implementasi Proses <i>Case Folding</i>	87
Tabel 4. 9 Contoh Hasil Implementasi Proses <i>Normalization</i>	90
Tabel 4. 10 Contoh Hasil Implementasi Proses <i>Stopword Removal</i>	92
Tabel 4. 11 Contoh Hasil Implementasi Proses <i>Stemming</i>	93
Tabel 4. 12 Distribusi Dataset Model Task 1 (Klasifikasi Aspek Multilabel).....	96
Tabel 4. 13 Distribusi Data Model Task 2-5 (Klasifikasi Sentimen per Aspek) ..	97
Tabel 4. 14 Pembagian Data Model Task 2-5 (Klasifikasi Sentimen Aspek)	98
Tabel 4. 15 Hasil Most Similar <i>Word embedding</i> pada Data Non-Stemmed	100
Tabel 4. 16 Hasil Most Similar <i>Word embedding</i> pada Data Stemmed.....	101
Tabel 4. 17 Parameter Training Model	111
Tabel 4. 18 Hasil Training dan Validation Model Klasifikasi Aspek Multilabel	112
Tabel 4. 19 <i>F1 score</i> Aspek pada Model Klasifikasi Aspek Multilabel	114
Tabel 4. 20 Training & Validation Model Klasifikasi Sentimen Aspek <i>Room</i> ..	117
Tabel 4. 21 Training & Validation Model Klasifikasi Sentimen Aspek Hotel...	120
Tabel 4. 22 Training & Validation Model Klasifikasi Sentimen <i>Location</i>	123
Tabel 4. 23 Training & Validation Model Klasifikasi Sentimen Aspek <i>Service</i>	126
Tabel 4. 24 Daftar Model dengan Skenario Terbaik pada Setiap Task	129

Tabel 4. 25 Evaluasi Generalisasi Model Klasifikasi Aspek Multilabel.....	133
Tabel 4. 26 Evaluasi Generalisasi Model Klasifikasi Sentimen <i>Room</i>	136
Tabel 4. 27 Evaluasi Generalisasi Model Klasifikasi Sentimen Hotel.....	139
Tabel 4. 28 Evaluasi Generalisasi Model Klasifikasi Sentimen <i>Location</i>	142
Tabel 4. 29 Evaluasi Generalisasi Model Klasifikasi Sentimen <i>Service</i>	145
Tabel 4. 30 Perbandingan <i>F1 Macro</i> Training & Validation dengan Data Test .	159
Tabel 4. 31 Contoh Ulasan Salah Prediksi Aspek.....	160
Tabel 4. 32 Contoh Ulasan Salah Prediksi Sentimen	161

DAFTAR LAMPIRAN

Lampiran 1. Daftar Hotel	181
Lampiran 2. Profil Anotator	183
Lampiran 3. Link Source Code	184

DAFTAR SINGKATAN, ISTILAH DAN SIMBOL

SINGKATAN/ISTILAH/SIMBOL	ARTI DAN KETERANGAN
α	: Nilai Krippendorff's Alpha (ukuran reliabilitas interrater)
Do	: Observed disagreement (ketidaksepakatan yang diamati)
De	: Expected disagreement (ketidaksepakatan yang diharapkan secara acak)
$w_{e_p} \in \mathbb{R}_z$: Vektor untuk kata ke-p
z	: Vektor pencocokan kata
$E_{we} \in \mathbb{R}^{n \times z}$: Matriks kata yang setiap barisnya menyatakan vektor untuk kata tertentu
C_p	: Penggunaan window
$x^{(t)}$: <i>Current input</i> dan <i>output</i>
$y^{(t-1)}$: Iterasi
W_z dan R_z	: Bobot
b_z	: Vektor Bobot
\odot	: Perkalian titik antara dua vektor
b_i	: Vektor bias
$c^{(t)}$: Lapisan LSTM
$z^{(t)}$: <i>Candidate Value</i>
t	: <i>Time Step</i>
$c^{(t-1)}$: <i>Memory Cell</i>
W_f, R_f, p_f	: Bobot terkait
b_f	: Vektor bias bobot
$z^{(t)}$: Blok Input
$i^{(t)}$: <i>Input Gate</i>
$f^{(t)}$: <i>Forgot Gate</i>
$y^{(t-1)}$: Output unit LSTM