



SKRIPSI

**ASPECT-BASED SENTIMENT ANALYSIS PADA
DATA MULTI MODAL TWITTER UNTUK
BENCHMARK PERFORMA PRODUK STUDI
KASUS SMARTPHONE**

TEGAR OKTAVIANTO SIMBOLON

NPM 21082010140

DOSEN PEMBIMBING

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

M. Irwan Afandi, S.T., M.Sc.

**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



SKRIPSI

ASPECT-BASED SENTIMENT ANALYSIS PADA DATA MULTI MODAL TWITTER UNTUK BENCHMARK PERFORMA PRODUK STUDI KASUS SMARTPHONE

TEGAR OKTAVIANTO SIMBOLON

NPM 21082010140

DOSEN PEMBIMBING

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

M. Irwan Afandi, S.T., M.Sc.

**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

ASPECT-BASED SENTIMENT ANALYSIS PADA DATA MULTI MODAL TWITTER UNTUK BENCHMARK PERFORMA PRODUK STUDI KASUS SMARTPHONE

Oleh:

TEGAR OKTAVIANTO SIMBOLON
NPM. 21082010140

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Sistem Informasi
Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 4
Juni 2025

Menyetujui

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

(Pembimbing I)

M. Irwan Afandi, S.T., M.Sc.
NIP. 19760718 2021211 003

(Pembimbing II)

Nur Cahyo Wibowo, S.Kom, M.Kom
NIP. 19790317 2021211 002

(Ketua Penguji)

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

(Penguji II)

Prasasti Karunia F.A., S.Kom., M.Kom
NIP. 19970704 2024062 001

(Penguji III)

Mengetahui,

Dekan Fakultas Ilmu Komputer

Prof. Dr. Ir. Novirina Hendrasarie, MT

NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**ASPECT-BASED SENTIMENT ANALYSIS PADA DATA MULTI MODAL
TWITTER UNTUK BENCHMARK PERFORMA PRODUK STUDI KASUS
SMARTPHONE**

Oleh:

TEGAR OKTAVIANTO SIMBOLON

NPM. 21082010140



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

Yang bertandatangan di bawah ini:

Nama : TEGAR OKATAVIAN TO SIMBOLON
NPM : 21082010181
Program : Sarjana (S1) Program Studi : Sistem
Informasi 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, 4 Juni 2025

Yang Membuat Pernyataan,



TEGAR OKTAVIAN TO SIMBOLON

NPM. 21082010140

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN ORISINALITAS

Yang bertandatangan di bawah ini:

Nama Mahasiswa : TEGAR OKTAVIANTO SIMBOLON
Program Studi : SISTEM INFORMASI
Dosen Pembimbing : 1. Eka Dyar Wahyuni, S.Kom., M.Kom.
2. M. Irwan Afandi, S.T., M.Sc.

dengan ini menyatakan bahwa isi sebagian maupun keseluruhan disertai dengan judul:

ASPECT-BASED SENTIMENT ANALYSIS PADA DATA MULTI MODAL TWITTER UNTUK BENCHMARK PERFORMA PRODUK STUDI KASUS SMARTPHONE

adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diizinkan dan bukan merupakan karya pihak lain yang saya akui sebagai karya sendiri. Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pustaka. Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.

Surabaya, 4 Juni 2025

Yang Membuat Pernyataan,



TEGAR OKTAVIANTO SIMBOLON

NPM. 21082010140

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Tegar Oktavianto Simbolon / 21082010140
Judul Skripsi : Aspect-Based Sentiment Analysis Pada Data Multi
Modal Twitter Untuk Benchmark Performa Produk
Studi Kasus Smartphone
Dosen Pembimbing : 1. Eka Dyar Wahyuni, S.Kom., M.Kom.
2. M. Irwan Afandi, S.T., M.Sc..

Penelitian ini mengembangkan dashboard benchmark smartphone berbasis analisis sentimen aspek (Aspect-Based Sentiment Analysis/ABSA) menggunakan data multimodal dari teks tweet dan gambar yang diekstraksi dengan Optical Character Recognition (OCR). Metode yang digunakan mencakup preprocessing teks (case folding, tokenisasi, stopwords removal) dan klasifikasi sentimen menggunakan Support Vector Machine (SVM) pada aspek produk seperti kamera, performa, dan lainnya. Dataset yang digunakan terdiri dari teks tweet dan URL gambar yang tercantum dalam file CSV. Gambar yang ada di URL tersebut diproses menggunakan OCR untuk diubah menjadi teks, kemudian digabungkan dengan teks tweet sebelum diprediksi oleh model untuk menentukan sentimen. Fitur OCR ini memungkinkan gambar yang tercantum di CSV digunakan untuk memperkaya informasi sentimen dalam data. Hasil penelitian menunjukkan bahwa model SVM dapat memprediksi sentimen yang berguna bagi konsumen dalam membandingkan produk, serta bagi produsen untuk meningkatkan kualitas produk dan strategi pemasaran. Penelitian ini membuktikan potensi pemanfaatan data multimodal dalam menghasilkan analisis sentimen yang lebih komprehensif terhadap produk smartphone.

Kata Kunci : Data Mining, Analisis Sentimen Berbasis Aspek, Support Vector Machine, Twitter, Data Multimodal, Optical Character Recognition, Smartphone, Dashboard Benchmark

Halaman ini sengaja dikosongkan

ABSTRACT

Nama Mahasiswa / NPM : Tegar Oktavianto Simbolon / 21082010140
Judul Skripsi : Aspect-Based Sentiment Analysis Pada Data Multi
Modal Twitter Untuk Benchmark Performa
Produk : Studi Kasus Smartphone
Dosen Pembimbing : 1. Eka Dyar Wahyuni, S.Kom., M.Kom.
2. M. Irwan Afandi, S.T., M.Sc.

This study develops a smartphone benchmarking dashboard based on Aspect-Based Sentiment Analysis (ABSA) using multimodal data from tweet texts and images extracted through Optical Character Recognition (OCR). The methods employed include text preprocessing (case folding, tokenization, and stopword removal) and sentiment classification using the Support Vector Machine (SVM) algorithm on product aspects such as camera, performance, and others. The dataset consists of tweet texts and image URLs listed in a CSV file. Images from these URLs are processed with OCR to convert them into text, which is then combined with the tweet text before being predicted by the model to determine sentiment. This OCR feature allows images listed in the CSV to enrich sentiment information within the data. The results show that the SVM model is capable of predicting sentiments that are useful for consumers in comparing products and for manufacturers in improving product quality and marketing strategies. This study demonstrates the potential of utilizing multimodal data to produce more comprehensive sentiment analysis for smartphone products.

Keywords: Data Mining, Aspect-Based Sentiment Analysis, Support Vector Machine, Twitter, Multimodal Data, Optical Character Recognition, Smartphone, Benchmark Dashboard

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT, Sang Maha Pemberi Ilmu, atas limpahan rahmat, taufik, dan hidayah-Nya yang tak terhingga. Hanya dengan izin-Nya, penulis dapat menyelesaikan skripsi ini dengan judul “Aspect-Based Sentiment Analysis Pada Data MultiModal Twitter Untuk Benchmark Performa Produk Studi Kasus Smartphone” sebagai salah satu syarat meraih gelar Sarjana Komputer di Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Proses perjalanan akademik ini terukir dalam lembaran waktu yang penuh arti. Warkop CN, dengan aroma kopi yang membangkitkan semangat dan perbincangan hangat di antara deru mesin espresso, menjadi saksi bisu perjuangan penulis menyusun setiap bab, menuangkan gagasan, dan merangkai analisis. Di sudut berbeda, Warkop Joker hadir sebagai ruang akhir yang penuh ketenangan, tempat penulis merapikan ide, menyempurnakan kalimat, dan meniupkan roh kesempurnaan pada skripsi ini.

Sebuah karya tak mungkin lahir dari satu tangan. Oleh karena itu, dengan hati yang tulus, penulis ingin menyampaikan rasa terima kasih yang sedalam-dalamnya kepada:

1. Kedua orang tua tercinta, yang menjadi pelita hidup penulis. Doa, pengorbanan, serta kasih sayang yang tak pernah lekang oleh waktu menjadi energi terbesar untuk melangkah, bahkan di saat langit terasa kelam. Setiap tetes keringat dan doa Bapak-Ibu adalah mahkota yang menghiasi perjalanan ini.
2. Ibu Eka Dyar Wahyuni, S.Kom., M.Kom., Dosen Pembimbing Pertama, yang dengan sabar membimbing penulis layaknya seorang ibu mengajarkan anaknya berjalan. Saran, kritik, dan arahan Ibu menjadi kompas yang mengarahkan setiap langkah penyusunan skripsi.
3. Bapak M. Irwan Afandi, S.T., M.Sc., Dosen Pembimbing Kedua, yang dengan ketelitiannya membantu penulis mempertajam analisis dan memberikan solusi sangat detail hingga skripsi ini dapat selesai.

4. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T., selaku Dekan Fakultas Ilmu Komputer, serta Bapak Agung Brastama Putra, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi, yang telah membuka gerbang ilmu dan memfasilitasi atmosfer akademik yang kondusif.
5. Seluruh dosen dan staf Program Studi Sistem Informasi, yang telah menyalakan lentera pengetahuan, mengasah keterampilan, dan membentuk penulis menjadi pribadi yang siap berkontribusi bagi negeri.
6. Abrilla Fatzali, Nizar Maulana, Maulana Khariska, Miftah Ramadhani, M. Aghis Sufiantoro, Dhavina Oxca, Radhyana Gayatri dan M. Nur Rachman, sahabat-sahabat terbaik yang selalu menjadi penolong dalam mengerjakan skripsi ini .

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Namun, di balik setiap kekurangan, terdapat perjuangan, usaha, dan doa yang telah menyertai setiap langkahnya. Penulis percaya bahwa tidak ada yang sempurna, dan yang terpenting adalah menyelesaikan tugas ini dengan baik. Skripsi yang baik bukanlah yang sempurna, melainkan yang selesai. Dengan kerendahan hati, penulis siap menerima kritik dan saran yang konstruktif untuk perbaikan di masa yang akan datang.

Akhir kata, semoga skripsi ini dapat menjadi setitik cahaya yang bermanfaat bagi perkembangan ilmu pengetahuan, khususnya dalam bidang sistem rekomendasi dan analisis data. Seluruh dedikasi ini penulis persembahkan untuk Indonesia, tanah air tercinta, yang terus menginspirasi generasi muda untuk berkarya.

Surabaya, 1 Juni 2025

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERSETUJUAN.....	ii
SURAT PERNYATAAN BEBAS PLAGIASI.....	iii
SURAT PERNYATAAN ORISINALITAS	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xvi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah.....	5
1.4. Tujuan	5
1.5. Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	9
2.1. Penelitian Terdahulu	9
2.2. Benchmark Smartphone	12
2.3. Analisis Sentimen Berbasis Aspek (ABSA)	12
2.4. Analisis Sentimen Text Mining	13
2.5. Inter Rater Agreement.....	14
2.6. Preprocessing data.....	16
2.6.1 Case-Folding	16
2.6.2 Cleaning	17
2.6.3 Normalization.....	17
2.6.4 Tokenizing.....	17
2.6.5 Stopword Removal.....	18
2.6.6 Stemming	18
2.7. Data Augmentation	19

2.8 Pembobotan TF-IDF	20
2.9 Support Vector Machine (SVM)	22
2.10 GridSearch	24
2.11 K-Fold Cross Validation.....	25
2.12 Confusion Matrix.....	25
2.13 Accuracy	27
2.14 Precision	28
2.15 Recall	28
2.16 F1-Score.....	29
2.17 Classification Report	29
BAB III METODOLOGI PENELITIAN	31
3.1 Studi Literatur.....	31
3.2 Pengumpulan Data.....	32
3.2.1 Source code scraping	33
3.2.2 Data Hasil Scraping	35
3.3 Word Cloud	35
3.4 Pelabelan data	36
3.5 Preprocessing Data	41
3.5.1 Case-folding.....	41
3.5.2 Cleaning.....	43
3.5.3 Normalization	45
3.5.4 Tokenizing	46
3.5.5 Stopword Removal	48
3.5.6 Stemming.....	49
3.6 Pelatihan Model.....	51
3.6.1 Konfigurasi Penelitian Dengan Balancing Proses	52
3.6.2 Konfigurasi Penelitian Dengan Pembagian data train dan data testing	53
3.6.3 Transformasi Data dengan TF-IDF	55
3.6.4 Penerapan Algoritma Support Vector Machine (SVM)	55
3.6.5 Data Augmentation.....	58
3.7 Evaluasi	59
3.7.1 Perhitungan Performa Algoritma.....	59

3.7.2 Processing Time	60
3.7.2 CPU Usage	61
3.8 Pembuatan Website	61
3.8.1 WireFrame Dashboard Website	62
3.8.2 Image Extraction	63
3.8.3 Data hasil image processing.....	63
3.9 Pembuatan Laporan Skripsi	66
3.10 Pembuatan Jurnal	66
BAB IV HASIL DAN PEMBAHASAN	67
4.1 Pengumpulan Data	67
4.1.1 Source code scraping.....	68
4.1.2 Data Hasil Scraping.....	70
4.2 Word Cloud.....	71
4.3 Pelabelan data.....	72
4.3.1 Pengecekan Sebaran data	72
4.3.2 Pelabelan Anotator	85
4.3.3 Penghitungan Inter rater Agreement	90
4.4 Preprocessing Data.....	95
4.4.1 Case Folding	96
4.4.2 Cleaning	97
4.4.3 Normalization.....	99
4.4.4 Tokenize.....	100
4.4.5 Stopword Removal.....	101
4.4.6 Stemming	108
4.4.7 Hasil Preprocessing Data	110
4.5 Pelatihan Model	111
4.5.1 Konfigurasi Eksperimen Awal.....	112
4.5.2 Hasil Pelatihan Awal.....	157
4.5.3 Data Augmentation	165
4.5.4 Output Data Augmentation	171
4.5.5 Pelatihan Model Setelah Dilakukan Data Augmentation.....	172
4.6 Evaluasi	174

4.6.1 Perhitungan Performa	175
4.6.2 Perbandingan F1 Macro Avg dan Accuracy per aspek	195
4.6.3 Perbandingan Time Processing Test dan Core CPU yang Digunakan	199
4.7 Pembuatan Website	211
4.7.1 Halaman Awal Website	212
4.7.2 Halaman Hasil Prediksi dari file CSV	214
4.7.3 Prediksi Text Komentar Handphone	231
BAB V PENUTUP	239
5.1 Kesimpulan.....	239
5.2 Saran	240
DAFTAR PUSTAKA.....	241
LAMPIRAN	245
Lampiran 1. Tabel Hasil Pelatihan Awal Lengkap.....	245
Lampiran 2. Tabel Hasil Pelatihan Setelah Data Augmentation Lengkap	298

DAFTAR GAMBAR

Gambar 2.1. pemisahkan dua kelompok data (hyperlane).....	23
Gambar 2.2. Gambar K-Fold Cross Validation	25
Gambar 2.3. Gambar Contoh Confusion Matrix.....	26
Gambar 2.4. Gambar Contoh Tabel Classification report.....	30
Gambar 3.1 Alur Penelitian.....	31
Gambar 3.2 Contoh Tweet yang akan di scraping	33
Gambar 3.3 Penerapan Scraping Data Twitter.....	34
Gambar 3.4 Contoh data hasil Scraping.....	35
Gambar 3.5 Source code World Cloud	36
Gambar 3.6 Alur Pelabelan Oleh 5 Anotator	40
Gambar 3.7 Alur Preprocessing Data.....	41
Gambar 3.8 Source code case folding.....	42
Gambar 3.9 Penerapan Cleaning Data	44
Gambar 3.10 Penerapan Normalization	45
Gambar 3.11 Contoh Penerapan Source code.....	47
Gambar 3.12 Contoh penerapan source code.....	48
Gambar 3.13 Contoh Penerapan Source Code Tahap Stemming	50
Gambar 3.14 Tahap Pelatihan Model.....	51
Gambar 3.15 Tahap Pembuatan Website	62
Gambar 3.16 Tahap Klasifikasi Data	62
Gambar 4.1 Contoh Tweet yang akan di scraping	68
Gambar 4.2 Penerapan Scraping Data Twitter.....	69
Gambar 4.3 Contoh data hasil Scraping.....	70
Gambar 4.4 Dataset Setelah Digabungkan.....	71
Gambar 4.5 Source code World Cloud	72
Gambar 4.5 hasil pelabelan awal	73
Gambar 4.6 Source Code Pengecekan	73
Gambar 4.7 Sebaran Data Tiap Aspek.....	74
Gambar 4.8 Data Setelah Scraping Ulang dan Pelabelan Awal.....	75
Gambar 4.9 Sebaran Data Setelah Scraping Ulang.....	76
Gambar 4.10 Source Code Untuk Menyaring Data	77

Gambar 4.11 Source Code Untuk Mengurutkan Data.....	78
Gambar 4.12 Hasil Pengurutan Data	80
Gambar 4.13 Hasil Sebaran Data Setelah BackTranslation	84
Gambar 4.14 Source Code Pelabelan IndoBert	86
Gambar 4.15 Source Code Pelabelan NLTK.....	88
Gambar 4.16 Source Code Penghitungan Inter Rater Aggrement.....	92
Gambar 4.17 Hasil Penghitungan Setelah Pelabelan Ulang.....	95
Gambar 4.18 Source Code Case Folding.....	96
Gambar 4.19 Source Code Cleaning	98
Gambar 4.20 Source Code Normalization.....	100
Gambar 4.21 Source Code Tokenize	101
Gambar 4.22 Source Code Stopword Removal.....	106
Gambar 4.23 Source Code WordCloud.....	108
Gambar 4.24 Source Code Stemming	109
Gambar 4.25 Source Code Pembagian Data Train dan Testing	114
Gambar 4.26 Source Code Skenario 1.....	116
Gambar 4.27 Source Code Skenario 2.....	119
Gambar 4.28 Source Code Skenario 3.....	123
Gambar 4.29 Source Code Skenario 4.....	126
Gambar 4.30 Source Code Skenario 5.....	130
Gambar 4.31 Source Code Skenario 6.....	133
Gambar 4.32 Source Code Skenario 7.....	137
Gambar 4.33 Source Code Skenario 8.....	140
Gambar 4.34 Source Code Skenario 9.....	144
Gambar 4.35 Source Code Skenario 10.....	147
Gambar 4.36 Source Code Skenario 11.....	151
Gambar 4.37 Source Code Skenario 12.....	154
Gambar 4.38 Source Code Visualisasi Semua Skenario	158
Gambar 4.39 Source Code Pengecekan Underfit dan Overfit.....	162
Gambar 4.40 Source Code Postaging.....	166
Gambar 4.41 Source Code Synonim Replacement	170
Gambar 4.42 Source Code Confusion Matrix Kedua Skenario.....	176

Gambar 4.43 Perbandingan Confusion Matrix Aspek Network	178
Gambar 4.44 Perbandingan Confusion Matrix Aspek body	179
Gambar 4.45 Perbandingan Confusion Matrix Aspek Display.....	181
Gambar 4.46 Perbandingan Confusion Matrix Aspek Platform	182
Gambar 4.47 Perbandingan Confusion Matrix Aspek performance	184
Gambar 4.48 Perbandingan Confusion Matrix Aspek Memory	185
Gambar 4.49 Perbandingan Confusion Matrix Aspek Camera.....	187
Gambar 4.50 Perbandingan Confusion Matrix Aspek Sound.....	188
Gambar 4.51 Perbandingan Confusion Matrix Aspek Communication	189
Gambar 4.52 Perbandingan Confusion Matrix Aspek features	191
Gambar 4.53 Perbandingan Confusion Matrix Aspek Battery	192
Gambar 4.54 Perbandingan Confusion Matrix Aspek Accessories	193
Gambar 4.55 Source Code Perbandingan F1 Score dan Accuracy	195
Gambar 4.56 Source Code Perbandingan Time Processing dan Core Usage	201
Gambar 4.57 Grafik Perbandingan Processing Time.....	202
Gambar 4.58 Grafik Perbandingan Core CPU yang Terpakai	203
Gambar 4.59 Potongan Source Code Halaman Awal Index.html.....	212
Gambar 4.60 Load Model dan Vectorizer Backend app.py	213
Gambar 4.61 Dashboard Halaman Awal	213
Gambar 4.62 Potongan Source Code Halaman Result	216
Gambar 4.63 Potongan Source Code Backend Proses File CSV	223
Gambar 4.64 Potongan Source Code Backend Download CSV	227
Gambar 4.65 Tampilan Upload File di Halaman Utama	229
Gambar 4.66 Tampilan Hasil Proses Benchmark	230
Gambar 4.67 Potongan Source Code Halaman Prediksi Text	232
Gambar 4.68 Source Code Backend Halaman text_prediction.....	234

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 3.1 Keyword untuk melengkapi kekurangan class pada aspek	38
Tabel 3.2 contoh case-Folding	42
Tabel 3.3 contoh Cleaning Data	44
Tabel 3.4 Contoh Normalization	46
Tabel 3.5 Contoh Tokenizing	47
Tabel 3.6 Contoh Stopword Removal	48
Tabel 3.7 Contoh Steaming	50
Tabel 3.8 Skenario Model yang dihasilkan	56
Tabel 3.9 Proses Confusion Matrix	60
Tabel 3.10 contoh setelah gambar berhasil diproses	64
Tabel 4.1 Kolom Tambahan Hasil Preprocessing	81
Tabel 4.2 Proses Backtranslation Googlesheet	83
Tabel 4.3 Kolom Tambahan Setelah Preprocessing Data	110
Tabel 4.4 Skenario Model	112
Tabel 4.6 Hasil Pelatihan Model	159
Tabel 4.7 Hasil Pengecekan	163
Tabel 4.8 Hasil Postaging	167
Tabel 4.9 Hasil Synonym Replacement	172
Tabel 4.10 Hasil Pelatihan Setelah Data Augmentation	173
Tabel 4.11 Hasil Pengecekan Ulang Underfit dan Overfit	174
Tabel 4.12 Hasil Synonym Replacement	198
Tabel 4.13 Classification Report Skenario 12	204

Halaman ini sengaja dikosongkan