

**ANALISIS SENTIMEN MULTI-ASPEK PADA TWITTER
TENTANG TIMNAS INDONESIA ERA PATRICK KLUIVERT
DENGAN *ENSEMBLE* LOGISTIC REGRESSION DAN
INDOBERT**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

RIO ANANDA

22.11.5134

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**ANALISIS SENTIMEN MULTI-ASPEK PADA TWITTER
TENTANG TIMNAS INDONESIA ERA PATRICK KLUIVERT
DENGAN *ENSEMBLE* LOGISTIC REGRESSION DAN
INDOBERT**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

RIO ANANDA

22.11.5134

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN MULTI-ASPEK PADA TWITTER TENTANG
TIMNAS INDONESIA ERA PATRICK KLUIVERT DENGAN *ENSEMBLE*
LOGISTIC REGRESSION DAN INDOBERT**

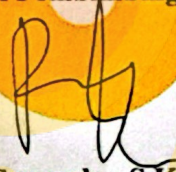
yang disusun dan diajukan oleh

Rio Ananda

22.11.5134

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 April 2026

Dosen Pembimbing,



Theopilus Bayu Sasongko, S.Kom., M.Eng.

NIK. 190302375

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN MULTI-ASPEK PADA TWITTER TENTANG TIMNAS INDONESIA ERA PATRICK KLUIVERT DENGAN *ENSEMBLE* LOGISTIC REGRESSION DAN INDOBERT

yang disusun dan diajukan oleh

Rio Ananda

22.11.5134

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 April 2026

Susunan Dewan Penguji

Nama Penguji

Arifiyanto Hadinegoro, S.Kom., M.T.
NIK. 190302289

Fauzia Anis Sekarningrum, S.T., M.T.
NIK. 190302683

Theopilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 April 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Rio Ananda
NIM : 22.11.5134

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Multi-Aspek pada Twitter tentang Timnas Indonesia Era Patrick Kluivert dengan *Ensemble Logistic Regression* dan IndoBERT

Dosen Pembimbing : Theopilus Bayu Sasongko, S.Kom., M.Eng.,

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 21 April 2026

Yang Menyatakan,



Rio Ananda

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah Swt. atas segala rahmat dan karunia-Nya, serta melalui proses yang panjang, penuh tantangan, dan pembelajaran, akhirnya karya ini dapat diselesaikan. Oleh karena itu, dengan kerendahan hati, karya ini saya persembahkan kepada:

1. Orang tua dan keluarga tercinta, yang senantiasa menjadi sumber kekuatan melalui doa, dukungan, dan kepercayaan yang tidak pernah berhenti mengiringi setiap langkah saya.
2. Dosen pembimbing, yang telah memberikan bimbingan, arahan, serta ilmu yang sangat berharga dengan penuh kesabaran selama proses penyusunan skripsi ini.
3. Universitas Amikom Yogyakarta, sebagai tempat saya belajar, berkembang, dan membentuk fondasi keilmuan yang akan menjadi bekal di masa depan.
4. Teman-teman kelas, yang telah menjadi bagian dari perjalanan ini, berbagi cerita, pengalaman, dan kebersamaan yang tidak terlupakan.
5. Teman-teman kontrakan, yang selalu hadir dalam keseharian, memberikan semangat, serta menciptakan suasana yang mendukung dalam menyelesaikan karya ini.
6. Diri saya sendiri, yang telah berproses, bertahan, dan terus melangkah hingga sampai pada titik ini.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah Swt. sehingga skripsi berjudul “Analisis Sentimen Multi-Aspek pada Twitter tentang Timnas Indonesia Era Patrick Kluivert dengan *Ensemble* Logistic Regression dan IndoBERT” dapat diselesaikan.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi S1-Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta. Dalam proses penyusunan skripsi ini, penulis menyampaikan terima kasih kepada:

1. Bapak Theopilus Bayu Sasongko, S.Kom., M.Eng., selaku dosen pembimbing yang telah memberikan bimbingan dan arahan.
2. Bapak Arifiyanto Hadinegoro, S.Kom., M.T. dan Ibu Fauzia Anis Sekarningrum, S.T., M.T., selaku dosen penguji yang telah memberikan saran dan masukan.
3. Seluruh dosen dan civitas akademika Program Studi S1-Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.
4. Orang tua dan keluarga penulis yang senantiasa memberikan doa dan dukungan.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun. Semoga skripsi ini dapat memberikan manfaat.

Yogyakarta, 21 April 2026

Penulis



Rio Ananda

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xv
DAFTAR LAMBANG DAN SINGKATAN	xvi
DAFTAR ISTILAH	xvii
INTISARI	xx
<i>ABSTRACT</i>	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5

2.2	Dasar Teori	13
2.2.1	Analisis Sentimen Berbasis Aspek	13
2.2.2	Analisis Sentimen Multi-Aspek	14
2.2.3	<i>Zero-Shot Labeling</i>	15
2.2.4	Logistic Regression	15
2.2.5	IndoBERT	17
2.2.6	<i>Ensemble Weighted Soft-Voting</i>	18
2.2.7	<i>Hyperparameter Tuning</i>	19
2.2.8	<i>Threshold Tuning</i>	20
2.2.9	Macro-F1 dan Micro-F1	21
2.2.10	<i>Area Under the Curve</i> (AUC)	23
2.2.11	<i>Confusion Matrix</i>	24
BAB III METODE PENELITIAN		26
3.1	Objek Penelitian	26
3.2	Alur Penelitian	27
3.2.1	Pengumpulan Data	28
3.2.2	Normalisasi Teks	32
3.2.3	<i>Case Folding</i>	32
3.2.4	Normalisasi Slang	32
3.2.5	Pembersihan Teks	33
3.2.6	Penyaringan <i>Noise</i>	34
3.2.7	Penentuan dan Validasi Aspek	35
3.2.8	Pelabelan Data	37
3.2.9	Visualisasi Data	39
3.2.10	Pembagian Dataset	39
3.2.11	Pemodelan dengan Logistic Regression	41
3.2.12	Pemodelan dengan IndoBERT	44
3.2.13	<i>Ensemble</i> dan Evaluasi Akhir	45
3.2.14	<i>Deployment</i>	46
3.3	Alat dan Bahan	47

3.3.1	Bahan Penelitian	47
3.3.2	Alat Penelitian	47
BAB IV HASIL DAN PEMBAHASAN		50
4.1	Hasil Pengumpulan Data	50
4.2	Hasil Normalisasi Teks	51
4.3	Hasil <i>Case Folding</i>	52
4.4	Hasil Normalisasi Slang	53
4.5	Hasil Pembersihan Teks	54
4.6	Hasil Penyaringan <i>Noise</i>	55
4.7	Hasil Validasi Aspek	58
4.8	Hasil Pelabelan Data	59
4.8.1	Hasil Pelabelan Aspek	59
4.8.2	Hasil Pelabelan Sentimen	63
4.9	Visualisasi Data	66
4.9.1	Distribusi Aspek	66
4.9.2	Distribusi Sentimen	67
4.9.3	Kemunculan Bersama Aspek	68
4.9.4	<i>Word Cloud</i>	69
4.10	Hasil Pembagian Dataset	70
4.11	Pemodelan Logistic Regression	72
4.11.1	Hasil Deteksi Aspek	74
4.11.1.1	Hasil <i>Hyperparameter Tuning</i>	74
4.11.1.2	Hasil <i>Threshold Tuning</i>	75
4.11.1.3	Evaluasi Kinerja Model	76
4.11.2	Hasil Klasifikasi Sentimen	79
4.11.2.1	Hasil <i>Hyperparameter Tuning</i>	80
4.11.2.2	Hasil <i>Threshold Tuning</i>	81
4.11.2.3	Evaluasi Kinerja Model	82

4.12	Pemodelan IndoBERT	84
4.12.1	Hasil Deteksi Aspek	84
4.12.1.1	Hasil <i>Hyperparameter Tuning</i>	84
4.12.1.2	Hasil <i>Threshold Tuning</i>	85
4.12.1.3	Evaluasi Kinerja Model	86
4.12.2	Hasil Klasifikasi Sentimen	88
4.12.2.1	Hasil <i>Hyperparameter Tuning</i>	88
4.12.2.2	Hasil <i>Threshold Tuning</i>	89
4.12.2.3	Evaluasi Kinerja Model	90
4.13	<i>Ensemble Weighted Soft-Voting</i>	92
4.13.1	Hasil Deteksi Aspek	92
4.13.1.1	Hasil Konfigurasi Bobot dan <i>Threshold</i>	92
4.13.1.2	Evaluasi Kinerja <i>Ensemble</i>	94
4.13.2	Hasil Klasifikasi Sentimen	96
4.13.2.1	Hasil Konfigurasi Bobot dan <i>Threshold</i>	96
4.13.2.2	Evaluasi Kinerja <i>Ensemble</i>	97
4.14	<i>Deployment</i>	99
4.15	Ringkasan Perbandingan Kinerja Model Tunggal dengan <i>Ensemble</i>	101
BAB V PENUTUP		105
5.1	Kesimpulan	105
5.2	Saran	106
REFERENSI		107

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	8
Tabel 3.1 Daftar Pertandingan Timnas Indonesia	26
Tabel 3.2 Contoh Kata Kunci dan Tagar	30
Tabel 3.3 Skema Penyaringan <i>Noise</i>	34
Tabel 3.4 Contoh Kata Kunci Validasi Aspek	35
Tabel 3.5 Skema Label Aspek	38
Tabel 3.6 Skema Label Sentimen	38
Tabel 3.7 Proporsi Pembagian Dataset	40
Tabel 3.8 Parameter <i>Hyperparameter Tuning</i> Logistic Regression	41
Tabel 3.9 Parameter <i>Hyperparameter Tuning</i> IndoBERT	44
Tabel 3.10 Spesifikasi Perangkat Keras	48
Tabel 4.1 Hasil Pengumpulan Data	50
Tabel 4.2 Contoh Data Hasil <i>Scraping</i>	51
Tabel 4.3 Contoh Hasil Normalisasi Teks	52
Tabel 4.4 Contoh Hasil <i>Case Folding</i>	52
Tabel 4.5 Contoh Hasil Normalisasi Slang	53
Tabel 4.6 Contoh Hasil Pembersihan Teks	54
Tabel 4.7 Contoh Data Hasil Penyaringan <i>Noise</i>	55
Tabel 4.8 Contoh Data Hasil Peninjauan Manual Penyaringan <i>Noise</i>	56
Tabel 4.9 Ringkasan Jumlah Data Hasil Penyaringan <i>Noise</i>	57
Tabel 4.10 Distribusi Aspek Berdasarkan Pemeriksaan Kata Kunci	58
Tabel 4.11 Contoh Data Hasil Pelabelan Aspek	60
Tabel 4.12 Contoh Data Hasil Peninjauan Manual Pelabelan Aspek	61
Tabel 4.13 Contoh Data Hasil Pelabelan Sentimen	63
Tabel 4.14 Contoh Data Hasil Peninjauan Manual Pelabelan Sentimen	64
Tabel 4.15 Hasil Pembagian Dataset	70
Tabel 4.16 Distribusi Aspek pada Tiga <i>Subset</i> Data	71
Tabel 4.17 Distribusi Sentimen pada Tiga <i>Subset</i> Data	71
Tabel 4.18 Contoh Hasil Tokenisasi	73
Tabel 4.19 Contoh Hasil Penghapusan <i>Stopwords</i>	73

Tabel 4.20 Contoh Hasil <i>Stemming</i>	74
Tabel 4.21 Hasil <i>Hyperparameter Tuning</i> Logistic Regression Deteksi Aspek	75
Tabel 4.22 Hasil <i>Threshold Tuning</i> Logistic Regression Deteksi Aspek	76
Tabel 4.23 Kinerja Logistic Regression Deteksi Aspek	77
Tabel 4.24 Hasil <i>Hyperparameter Tuning</i> Logistic Regression Klasifikasi Sentimen	80
Tabel 4.25 Hasil <i>Threshold Tuning</i> Logistic Regression Klasifikasi Sentimen	81
Tabel 4.26 Kinerja Logistic Regression Klasifikasi Sentimen	83
Tabel 4.27 Hasil <i>Hyperparameter Tuning</i> IndoBERT Deteksi Aspek	84
Tabel 4.28 Hasil <i>Threshold Tuning</i> IndoBERT Deteksi Aspek	85
Tabel 4.29 Kinerja Model IndoBERT Deteksi Aspek	87
Tabel 4.30 Hasil <i>Hyperparameter Tuning</i> IndoBERT Klasifikasi Sentimen	89
Tabel 4.31 Hasil <i>Threshold Tuning</i> IndoBERT Klasifikasi Sentimen	89
Tabel 4.32 Kinerja IndoBERT Klasifikasi Sentimen	91
Tabel 4.33 Konfigurasi Bobot dan <i>Threshold Ensemble</i> Deteksi Aspek	93
Tabel 4.34 Kinerja <i>Ensemble</i> Deteksi Aspek	94
Tabel 4.35 Konfigurasi Bobot dan <i>Threshold Ensemble</i> Klasifikasi Sentimen	97
Tabel 4.36 Kinerja <i>Ensemble</i> Klasifikasi Sentimen	98
Tabel 4.37 Perbandingan Performa Deteksi Aspek per Aspek	101
Tabel 4.38 Perbandingan Performa Deteksi Aspek	102
Tabel 4.39 Perbandingan Performa Klasifikasi Sentimen per Aspek	103
Tabel 4.40 Perbandingan Performa Klasifikasi Sentimen	104

DAFTAR GAMBAR

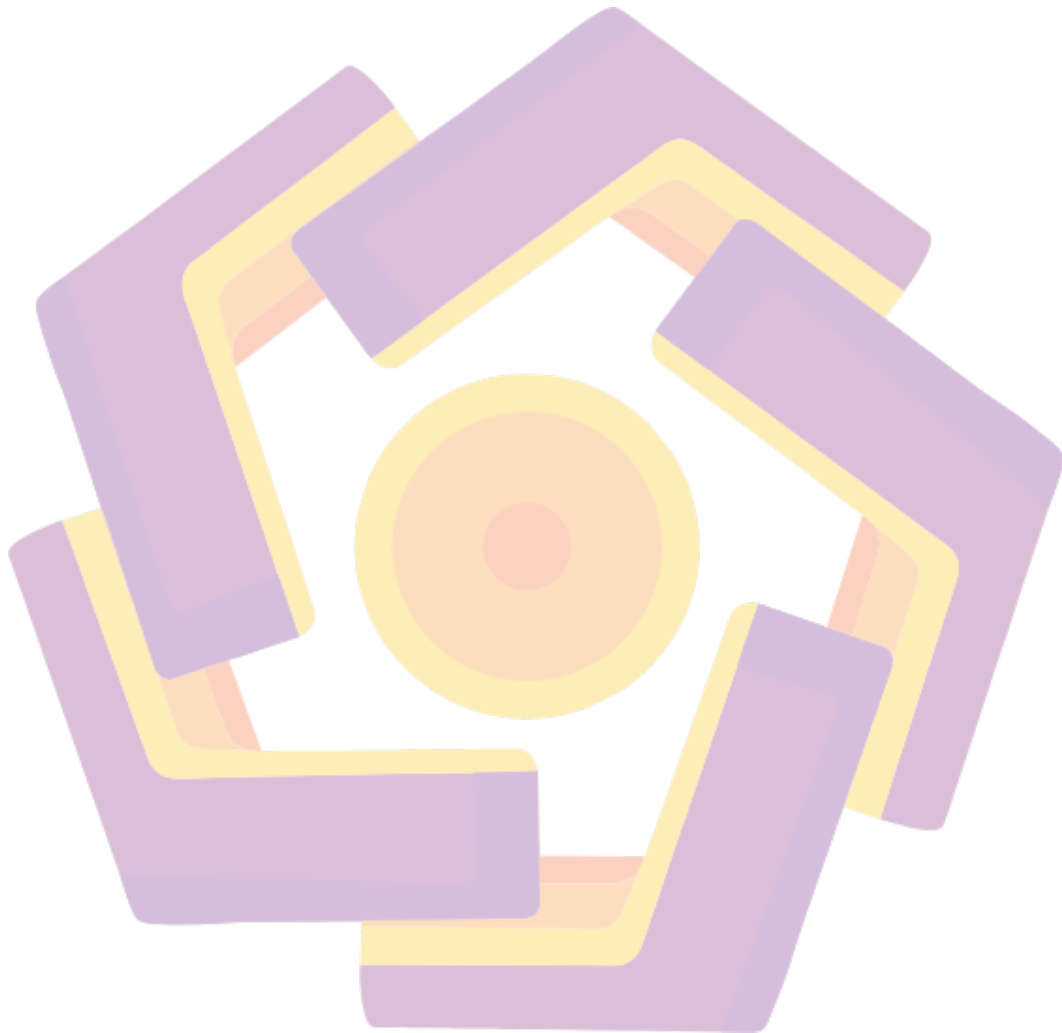
Gambar 2.1 Perbedaan ABSA <i>Single-label</i> dan Multi-label	14
Gambar 2.2 Arsitektur IndoBERT dalam ABSA	18
Gambar 2.3 <i>Hyperparameter</i> Tuning dengan <i>Grid Search</i>	20
Gambar 2.4 Struktur <i>Confusion Matrix</i>	25
Gambar 3.1 Alur Penelitian	28
Gambar 3.2 Topik yang sedang Tren	30
Gambar 3.3 Contoh <i>Query</i> Pengambilan Data	31
Gambar 4.1 Perubahan Distribusi Data <i>Relevan</i> akibat <i>Peninjauan Manual</i>	57
Gambar 4.2 Perubahan Distribusi <i>Aspek</i> akibat <i>Peninjauan Manual</i>	62
Gambar 4.3 Perubahan Distribusi Sentimen akibat <i>Peninjauan Manual</i>	66
Gambar 4.4 Distribusi <i>Aspek</i>	67
Gambar 4.5 Distribusi Sentimen	68
Gambar 4.6 Kemunculan Bersama <i>Aspek</i>	69
Gambar 4.7 <i>Word Cloud</i>	70
Gambar 4.8 Kinerja per <i>Aspek</i> Logistic Regression Deteksi <i>Aspek</i>	77
Gambar 4.9 <i>Confusion Matrix</i> Logistic Regression Deteksi <i>Aspek</i>	78
Gambar 4.10 Kurva ROC Logistic Regression Deteksi <i>Aspek</i>	79
Gambar 4.11 Contoh Hasil Format <i>Long</i>	79
Gambar 4.12 Kinerja per <i>Aspek</i> Logistic Regression Klasifikasi Sentimen	82
Gambar 4.13 <i>Confusion Matrix</i> Logistic Regression Klasifikasi Sentimen	83
Gambar 4.14 Kinerja per <i>Aspek</i> IndoBERT Deteksi <i>Aspek</i>	86
Gambar 4.15 <i>Confusion Matrix</i> IndoBERT Deteksi <i>Aspek</i>	87
Gambar 4.16 Kurva ROC IndoBERT Deteksi <i>Aspek</i>	88
Gambar 4.17 Kinerja per <i>Aspek</i> IndoBERT Klasifikasi Sentimen	90
Gambar 4.18 <i>Confusion Matrix</i> IndoBERT Klasifikasi Sentimen	91
Gambar 4.19 Kinerja per <i>Aspek</i> <i>Ensemble</i> Deteksi <i>Aspek</i>	94
Gambar 4.20 <i>Confusion Matrix</i> <i>Ensemble</i> Deteksi <i>Aspek</i>	95
Gambar 4.21 Kurva ROC <i>Ensemble</i> Deteksi <i>Aspek</i>	96
Gambar 4.22 Kinerja per <i>Aspek</i> <i>Ensemble</i> Klasifikasi Sentimen	98
Gambar 4.23 <i>Confusion Matrix</i> <i>Ensemble</i> Klasifikasi Sentimen	99

Gambar 4.24 Tampilan Halaman Awal Sistem *Deployment*

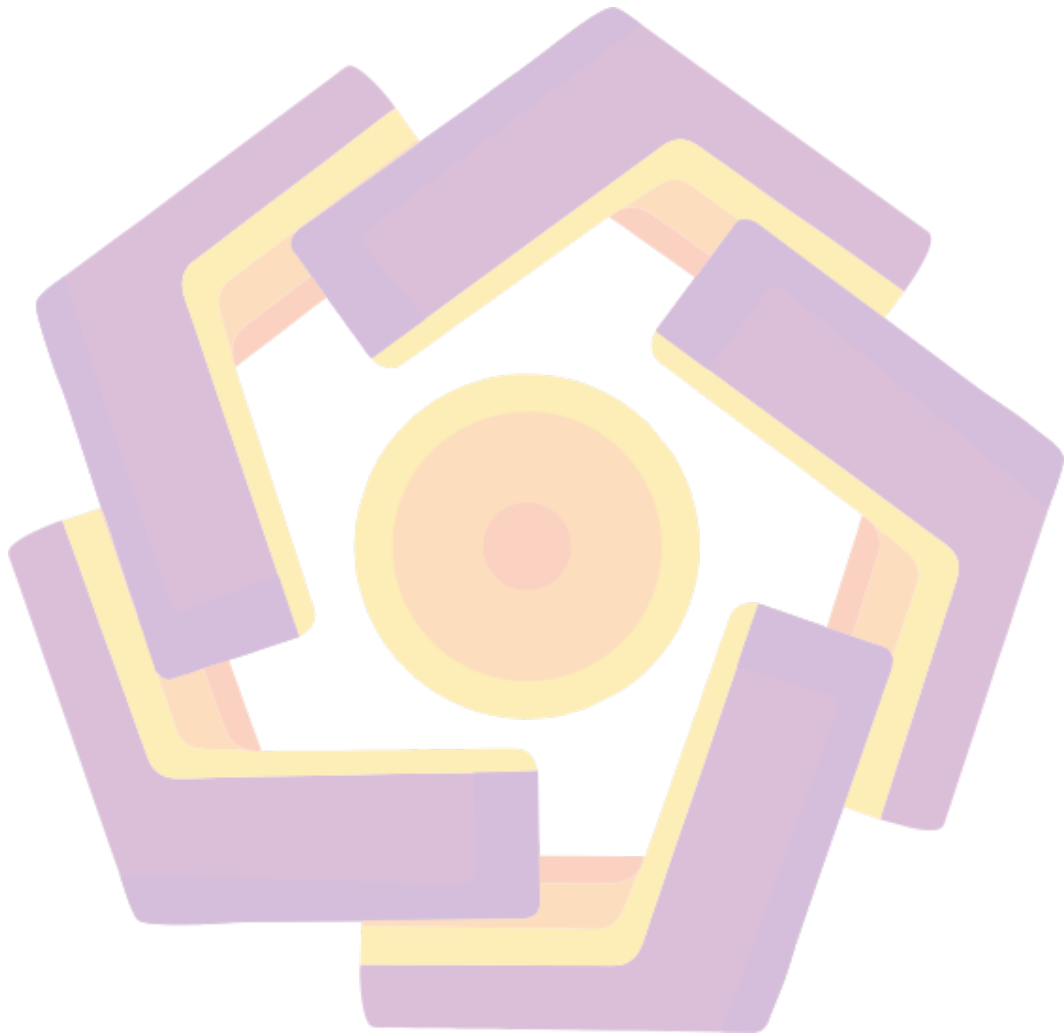
100

Gambar 4.25 Contoh Tampilan Hasil Prediksi

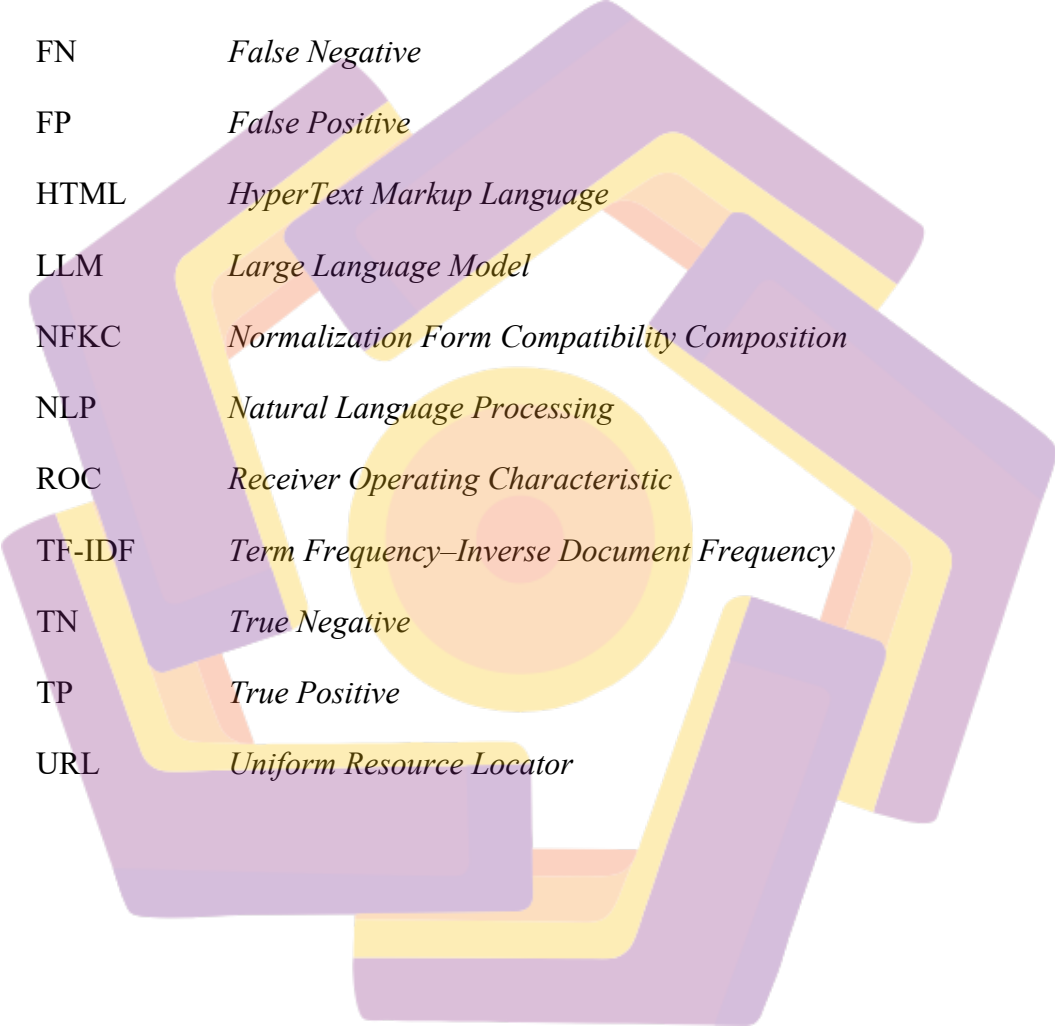
100



DAFTAR LAMPIRAN



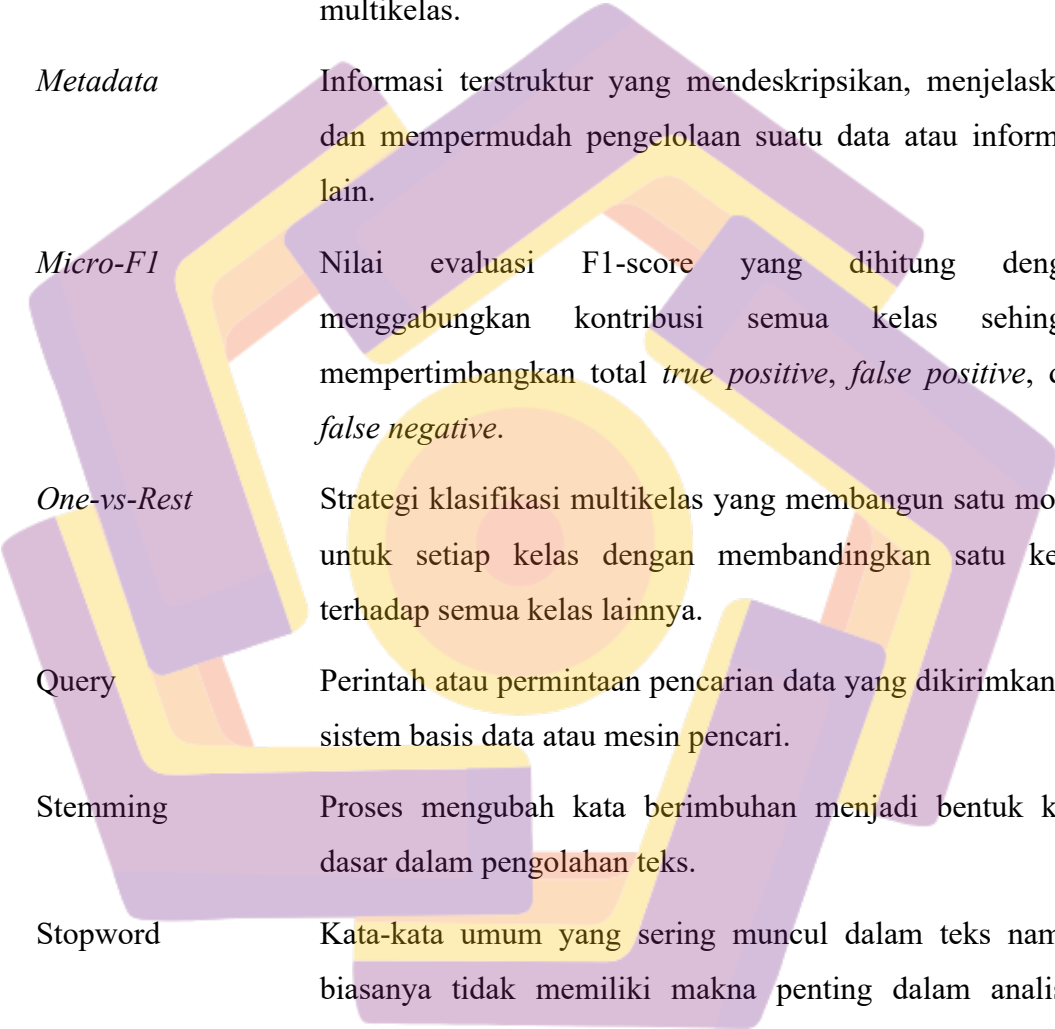
DAFTAR LAMBANG DAN SINGKATAN



ABSA	<i>Aspect Based Sentiment Analysis</i>
API	<i>Application Programming Interface</i>
AUC	<i>Area Under the Curve</i>
FN	<i>False Negative</i>
FP	<i>False Positive</i>
HTML	<i>HyperText Markup Language</i>
LLM	<i>Large Language Model</i>
NFKC	<i>Normalization Form Compatibility Composition</i>
NLP	<i>Natural Language Processing</i>
ROC	<i>Receiver Operating Characteristic</i>
TF-IDF	<i>Term Frequency–Inverse Document Frequency</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>
URL	<i>Uniform Resource Locator</i>

DAFTAR ISTILAH

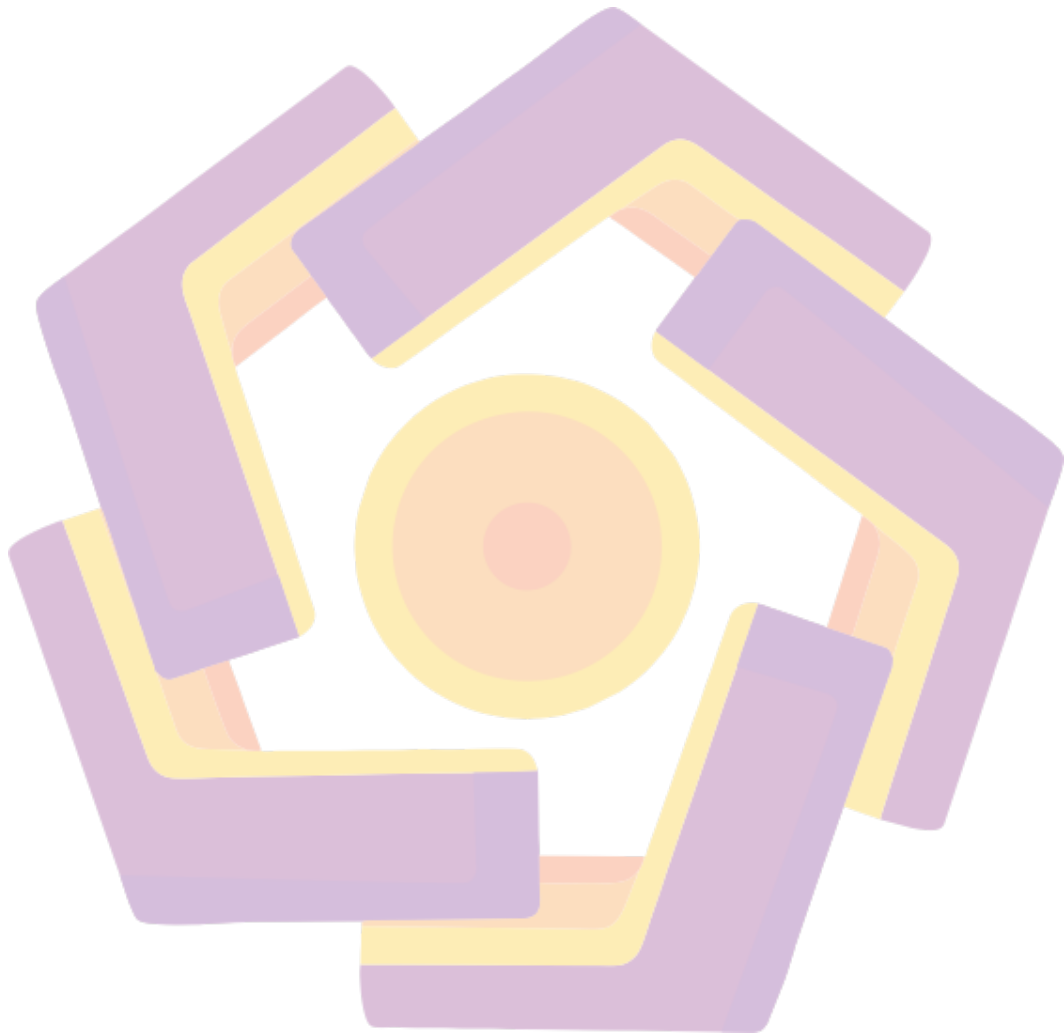
<i>Confusion Matrix</i>	Tabel evaluasi yang digunakan untuk menilai kinerja model klasifikasi dengan membandingkan hasil prediksi dengan data sebenarnya.
Cuitan	Pesan atau unggahan singkat yang dipublikasikan oleh pengguna di platform media sosial seperti Twitter/X.
Dataset	Kumpulan data terstruktur yang digunakan untuk berbagai keperluan, seperti analisis, penelitian, atau pengolahan informasi lainnya.
<i>ElasticNet</i>	Metode regularisasi dalam pembelajaran mesin yang menggabungkan penalti L1 (<i>Lasso</i>) dan L2 (<i>Ridge</i>) untuk meningkatkan performa model.
<i>Ensemble</i>	Teknik dalam pembelajaran mesin yang menggabungkan beberapa model untuk meningkatkan akurasi prediksi.
<i>Framework Flask</i>	Kerangka kerja (<i>framework</i>) berbasis bahasa pemrograman Python yang ringan dan digunakan untuk membangun aplikasi web.
Fitur leksikal	Karakteristik kata atau unsur kebahasaan dalam teks yang digunakan sebagai variabel atau <i>input</i> dalam analisis teks.
<i>Hyperparameter</i>	Parameter yang ditentukan sebelum proses pelatihan model dan digunakan untuk mengontrol proses pembelajaran algoritma.
IndoBERT	Model bahasa berbasis arsitektur BERT yang dilatih khusus menggunakan korpus bahasa Indonesia untuk tugas pemrosesan bahasa alami.



Logistic Regression	Algoritma klasifikasi statistik yang digunakan untuk memprediksi probabilitas suatu kelas berdasarkan variabel input.
Macro-AUC	Nilai rata-rata AUC yang dihitung dengan memberikan bobot yang sama pada setiap kelas dalam evaluasi model klasifikasi multikelas.
Metadata	Informasi terstruktur yang mendeskripsikan, menjelaskan, dan mempermudah pengelolaan suatu data atau informasi lain.
Micro-F1	Nilai evaluasi F1-score yang dihitung dengan menggabungkan kontribusi semua kelas sehingga mempertimbangkan total <i>true positive</i> , <i>false positive</i> , dan <i>false negative</i> .
One-vs-Rest	Strategi klasifikasi multikelas yang membangun satu model untuk setiap kelas dengan membandingkan satu kelas terhadap semua kelas lainnya.
Query	Perintah atau permintaan pencarian data yang dikirimkan ke sistem basis data atau mesin pencari.
Stemming	Proses mengubah kata berimbuhan menjadi bentuk kata dasar dalam pengolahan teks.
Stopword	Kata-kata umum yang sering muncul dalam teks namun biasanya tidak memiliki makna penting dalam analisis, sehingga sering dihapus dalam proses pra-pemrosesan teks.
Threshold	Nilai batas yang digunakan untuk menentukan keputusan klasifikasi atau pemisahan kategori dalam model.
Taktik	Cara atau strategi yang digunakan untuk mencapai tujuan tertentu dalam suatu konteks analisis atau penelitian.

Word Cloud

Visualisasi teks yang menampilkan kata-kata dengan ukuran berbeda berdasarkan frekuensi kemunculannya dalam suatu kumpulan data teks.



INTISARI

Perkembangan media sosial memungkinkan masyarakat menyampaikan opini secara terbuka terhadap berbagai isu, termasuk mengenai Tim Nasional Indonesia. Cuitan yang muncul di platform Twitter sering kali memuat beragam aspek sekaligus, seperti strategi pelatih, performa pemain, hasil pertandingan, dan opini umum. Oleh karena itu, analisis sentimen multi-aspek diperlukan untuk memahami sentimen publik secara lebih mendalam. Penelitian ini bertujuan untuk menerapkan metode *ensemble weighted soft-voting* dengan menggabungkan model Logistic Regression dan IndoBERT dalam tugas deteksi aspek multi-label dan klasifikasi sentimen per aspek terhadap cuitan Twitter berbahasa Indonesia mengenai Timnas Indonesia pada era kepelatihan Patrick Kluivert. Proses penelitian meliputi pengumpulan data cuitan, praproses teks, pemodelan menggunakan Logistic Regression berbasis TF-IDF dan IndoBERT, serta evaluasi kinerja model. Pada masing-masing model dilakukan *hyperparameter tuning* dan *threshold tuning* untuk memperoleh konfigurasi terbaik. Selanjutnya, kedua model digabungkan menggunakan metode *ensemble weighted soft-voting* dengan pencarian bobot dan ambang keputusan optimal pada setiap aspek. Evaluasi dilakukan menggunakan metrik Macro-F1, Micro-F1, dan *Area Under the Curve* (AUC). Hasil penelitian menunjukkan bahwa pendekatan *ensemble* mampu meningkatkan performa dibandingkan model tunggal. Pada tugas deteksi aspek, model *ensemble* menghasilkan nilai Macro-F1 keseluruhan yang lebih baik dibandingkan Logistic Regression maupun IndoBERT. Pada tugas klasifikasi sentimen per aspek, pendekatan *ensemble* juga menunjukkan kinerja yang lebih stabil dengan performa tertinggi pada aspek hasil pertandingan. Secara keseluruhan, kombinasi kedua model memberikan kontribusi dalam meningkatkan stabilitas dan akurasi prediksi pada analisis sentimen multi-aspek.

Kata kunci: analisis sentimen, multi-aspek, *ensemble*, Logistic Regression, IndoBERT, Twitter.

ABSTRACT

The rapid growth of social media enables the public to freely express their opinions on various issues, including discussions about the Indonesian National Football Team. Tweets posted on Twitter often contain multiple aspects simultaneously, such as coaching strategy, player performance, match results, and general opinions. Therefore, aspect-based sentiment analysis is required to better understand public sentiment in a more detailed manner. This study aims to implement an ensemble weighted soft-voting method that combines Logistic Regression and IndoBERT models for multi-label aspect detection and aspect-based sentiment classification on Indonesian-language tweets related to the Indonesian National Team during the Patrick Kluivert coaching era. The research process consists of tweet data collection, text preprocessing, model development using TF-IDF-based Logistic Regression and IndoBERT, and model evaluation. Each model undergoes hyperparameter tuning and threshold tuning to obtain optimal configurations. The predictions from both models are then combined using an ensemble weighted soft-voting approach by determining the optimal weights and decision thresholds for each aspect. Model performance is evaluated using Macro-F1, Micro-F1, and Area Under the Curve (AUC) metrics. The results show that the ensemble approach improves performance compared to individual models. In the aspect detection task, the ensemble produce better overall Macro-F1 values than both Logistic Regression and IndoBERT. In the aspect-based sentiment classification task, the ensemble approach also demonstrates more stable performance, with the highest results observed in the match result aspect. Overall, combining the two models contributes to improved prediction stability and accuracy in multi-aspect sentiment analysis.

Keywords: *sentiment analysis, aspect-based sentiment analysis, ensemble, Logistic Regression, IndoBERT, Twitter.*