

**ANALISIS SENTIMEN KOMENTAR INSTAGRAM TERHADAP
KEMENANGAN PSS SLEMAN PADA LIGA 2 PEGADAIAN
MENGUNAKAN ALGORITMA BERT**

SKRIPSI

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



disusun oleh

FIRMANSYAH EDLIS PUTRA

22.11.5356

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**ANALISIS SENTIMEN KOMENTAR INSTAGRAM TERHADAP
KEMENANGAN PSS SLEMAN PADA LIGA 2 PEGADAIAN
MENGUNAKAN ALGORITMA BERT**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

FIRMANSYAH EDLIS PUTRA

22.11.5356

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN KOMENTAR INSTAGRAM TERHADAP
KEMENANGAN PSS SLEMAN PADA LIGA 2 PEGADAIAN
MENGUNAKAN ALGORITMA BERT**

yang disusun dan diajukan oleh

Firmansyah Edlis Putra

22.11.5356

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 11 Juni 2026

Dosen Pembimbing,


Ahlihi Masruro, S. Kom., M. Kom.
NIK. 190302148

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS SENTIMEN KOMENTAR INSTAGRAM TERHADAP
KEMENANGAN PSS SLEMAN PADA LIGA 2 PEGADAIAN
MENGGUNAKAN ALGORITMA BERT

yang disusun dan diajukan oleh

Firmansyah Edlis Putra

22.11.5356

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Juni 2026

Susunan Dewan Penguji

Nama Penguji

Muhammad Rudvanto Arief, S.T., M.T.
NIK. 190302098

Yoga Pristyanto, S. Kom., M. Eng.
NIK. 190302412

Ahlihi Masruro, S. Kom., M. Kom.
NIK. 190302412

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Juni 2026

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Firmansyah Edlis Putra
NIM : 22.11.5356

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN KOMENTAR INSTAGRAM TERHADAP
KEMENANGAN PSS SLEMAN PADA LIGA 2 PEGADAIAN
MENGUNAKAN ALGORITMA BERT**

Dosen Pembimbing : Ahlihi Masruro, S.Kom., M.Kom.

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, 11 Juni 2026

Yang Menyatakan,



Firmansyah Edlis Putra

HALAMAN PERSEMBAHAN

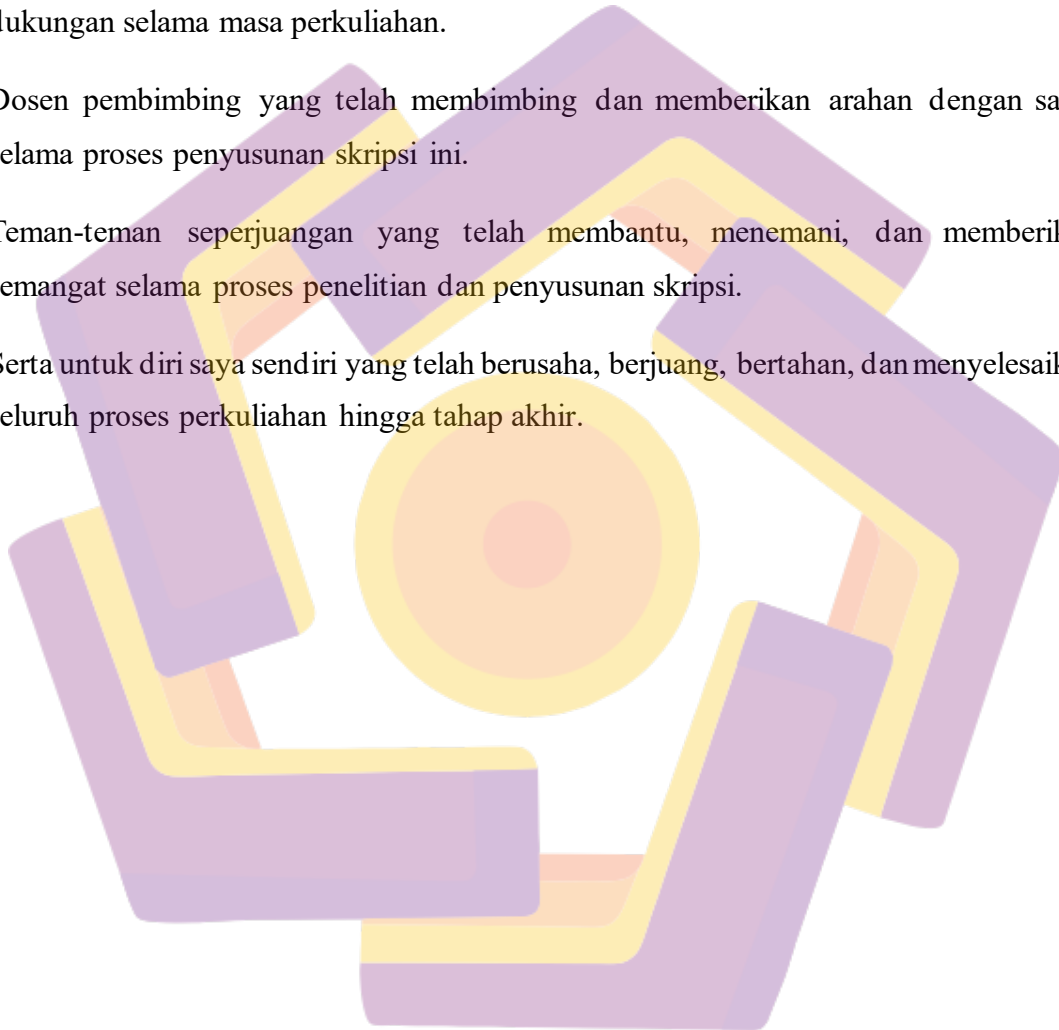
Kedua orang tua saya yang selalu memberikan doa, dukungan, dan semangat dalam setiap proses yang saya jalani hingga mampu menyelesaikan skripsi ini.

Keluarga dan orang-orang terdekat yang selalu memberikan motivasi, perhatian, serta dukungan selama masa perkuliahan.

Dosen pembimbing yang telah membimbing dan memberikan arahan dengan sabar selama proses penyusunan skripsi ini.

Teman-teman seperjuangan yang telah membantu, menemani, dan memberikan semangat selama proses penelitian dan penyusunan skripsi.

Serta untuk diri saya sendiri yang telah berusaha, berjuang, bertahan, dan menyelesaikan seluruh proses perkuliahan hingga tahap akhir.



KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena atas rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul "Analisis Sentimen Komentar Instagram Terhadap Kemenangan PSS Sleman Pada Liga 2 Pegadaian Menggunakan Algoritma BERT" dengan baik.

Penyusunan skripsi ini dilakukan sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta. Dalam proses penyusunan skripsi ini, saya menyadari bahwa banyak bantuan, dukungan, serta arahan dari berbagai pihak yang sangat berarti.

Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

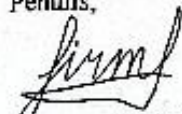
1. Bapak dan Ibu kedua orang tua saya yang selalu memberikan doa, dukungan, semangat, dan motivasi selama proses perkuliahan hingga penyusunan skripsi ini selesai.
2. Bapak Ahlihi Masruro, S.Kom., M.Kom. selaku dosen pembimbing yang telah memberikan arahan, masukan, dan bimbingan selama proses penelitian dan penyusunan skripsi.
3. Seluruh dosen Program Studi Informatika Universitas AMIKOM Yogyakarta yang telah memberikan ilmu dan pengalaman selama masa perkuliahan.
4. Keluarga dan teman-teman yang telah membantu serta memberikan dukungan selama proses penyusunan skripsi ini.

Saya menyadari bahwa skripsi ini masih memiliki kekurangan dan jauh dari kata sempurna. Oleh karena itu, saya terbuka terhadap kritik dan saran yang membangun agar penelitian ini dapat menjadi lebih baik.

Semoga skripsi ini dapat memberikan manfaat dan menambah wawasan bagi pembaca maupun peneliti selanjutnya.

Yogyakarta, 10 Juni 2026

Penulis,

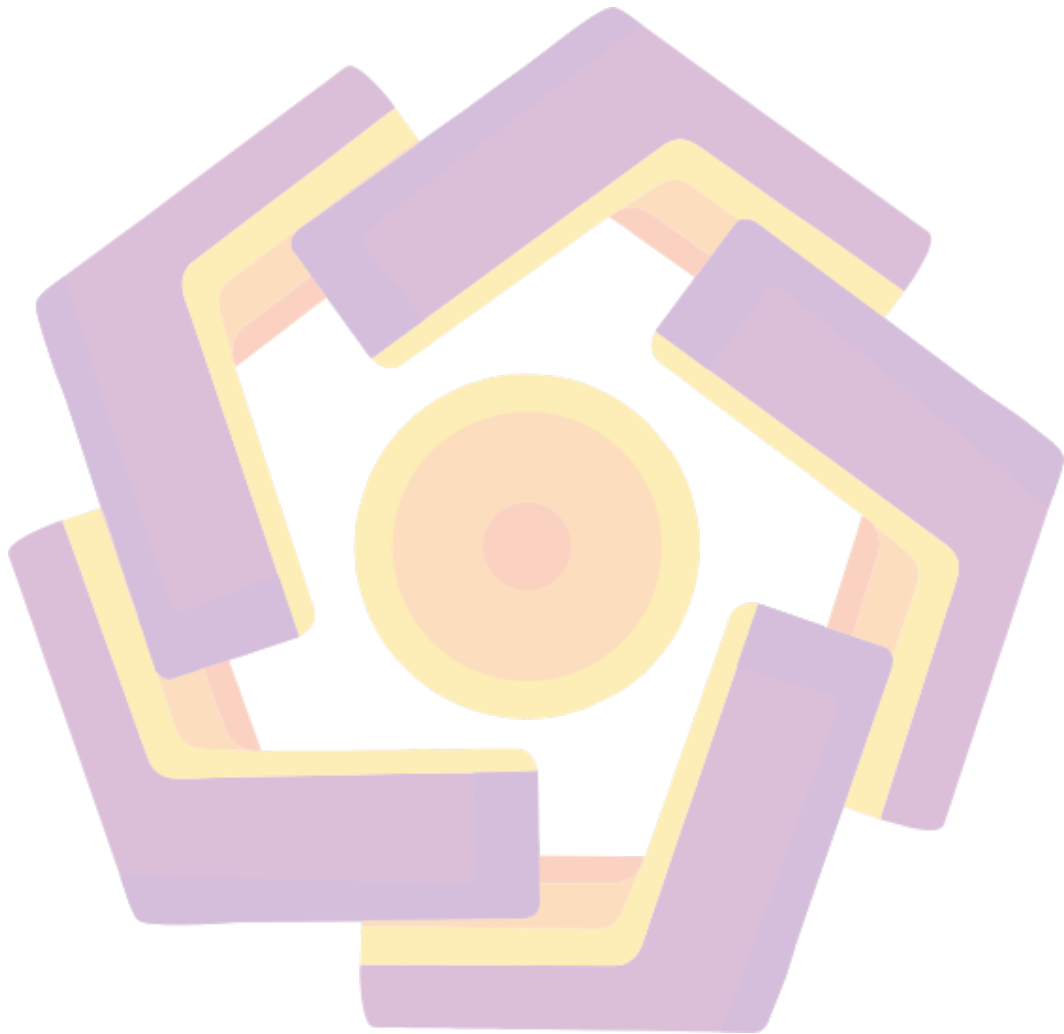


Firmansyah Edlis Putra

DAFTAR ISI

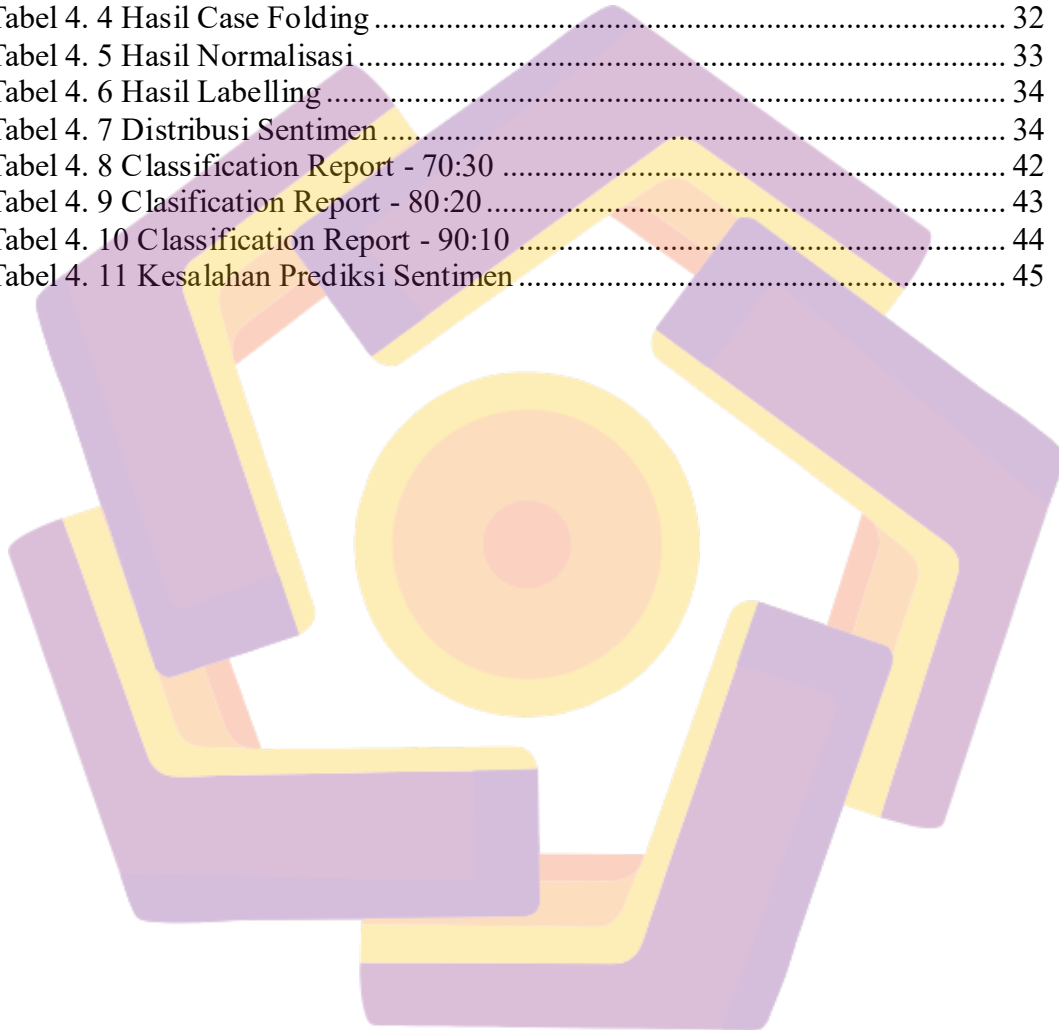
HALAMAN JUDUL.....	I
HALAMAN PERSETUJUAN.....	II
HALAMAN PENGESAHAN.....	III
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	IV
HALAMAN PERSEMBAHAN.....	V
KATA PENGANTAR.....	VI
DAFTAR ISI.....	VII
DAFTAR TABEL.....	IX
DAFTAR GAMBAR.....	X
DAFTAR LAMPIRAN.....	XI
DAFTAR LAMBANG DAN SINGKATAN.....	XII
DAFTAR ISTILAH	XIII
INTISARI.....	XVI
<i>ABSTRACT</i>	XVII
BAB I PENDAHULUAN.....	I
1.1 Latar Belakang	I
1.2 Rumusan Masalah	II
1.3 Batasan Masalah	II
1.4 Tujuan Penelitian	III
1.5 Manfaat Penelitian	III
1.6 Sistematika Penulisan	III
BAB II TINJAUAN PUSTAKA.....	V
2.1 Studi Literatur	V
2.2 Dasar Teori.....	XIII
BAB III METODE PENELITIAN.....	XXI
3.1 Objek Penelitian	XXI
3.2 Alur Penelitian	XXII
3.3 Alat dan Bahan.....	XXVII
BAB IV HASIL DAN PEMBAHASAN	XXVIII
BAB V PENUTUP.....	XLVII
5.1 Kesimpulan	XLVII
5.2 Saran.....	XLVIII
REFERENSI.....	XLIX

LAMPIRAN.....	LI
---------------	----



DAFTAR TABEL

Tabel 2.1. Perbandingan metode	10
Tabel 2.2. Rangkuman Tinjauan Pustaka	11
Tabel 4. 1 Postingan yang digunakan pada penelitian	28
Tabel 4. 2 Hasil Scraping	29
Tabel 4. 3 Hasil Cleaning	31
Tabel 4. 4 Hasil Case Folding	32
Tabel 4. 5 Hasil Normalisasi	33
Tabel 4. 6 Hasil Labelling	34
Tabel 4. 7 Distribusi Sentimen	34
Tabel 4. 8 Classification Report - 70:30	42
Tabel 4. 9 Clasification Report - 80:20	43
Tabel 4. 10 Classification Report - 90:10	44
Tabel 4. 11 Kesalahan Prediksi Sentimen	45

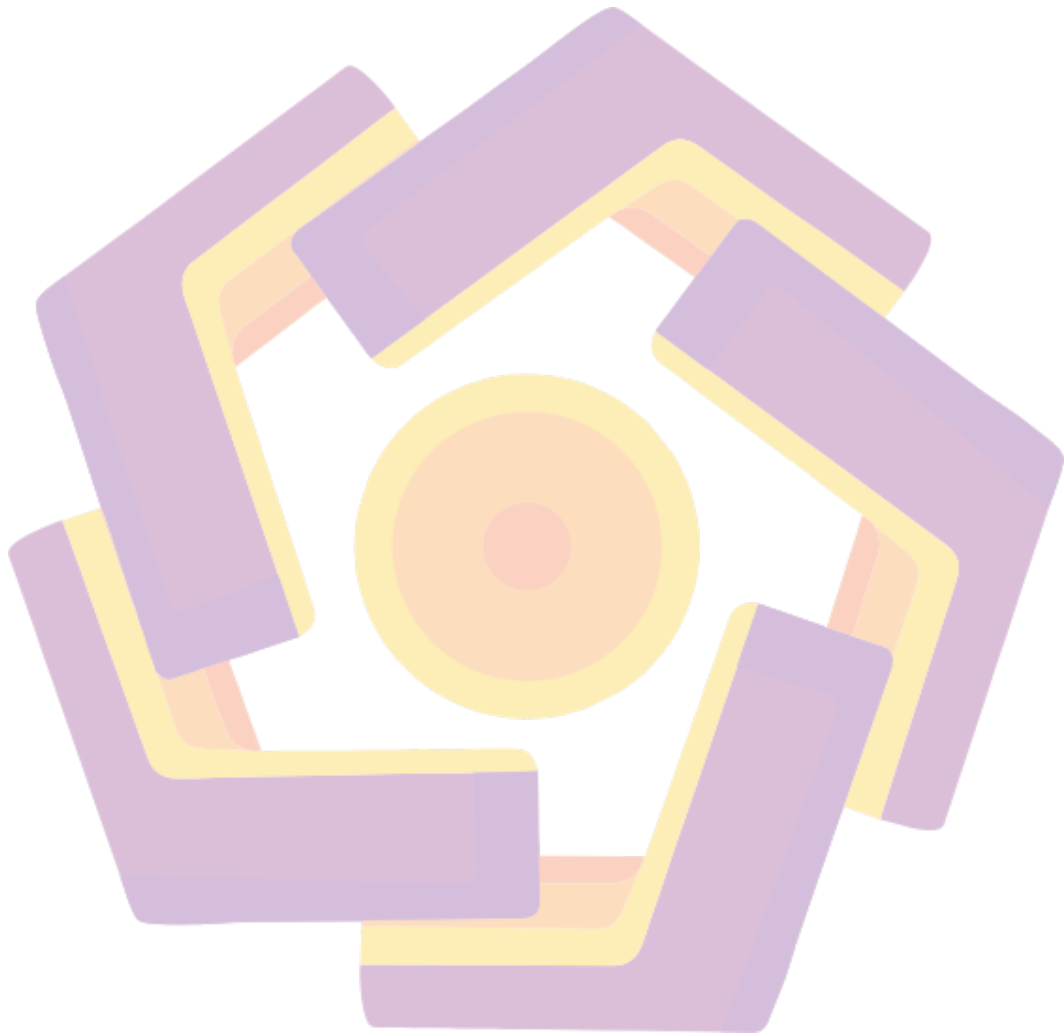


DAFTAR GAMBAR

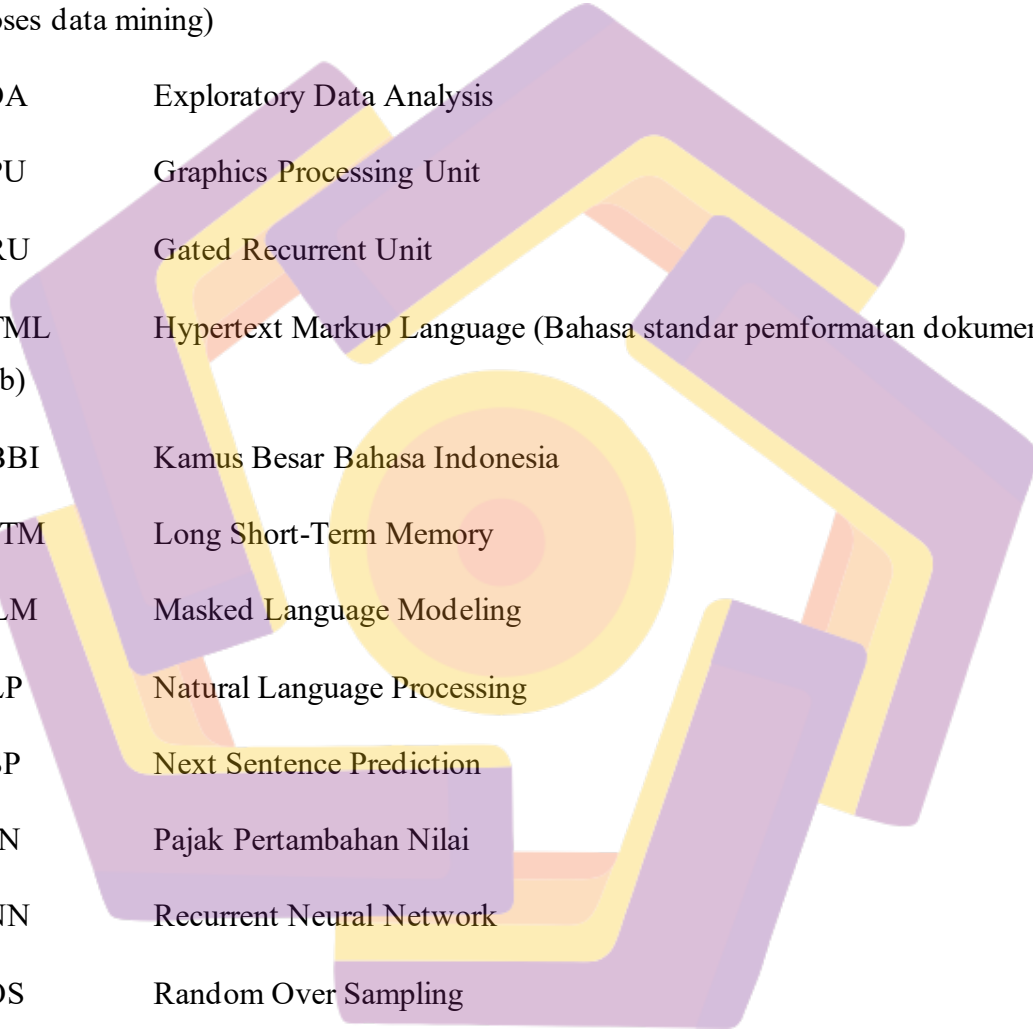
Gambar 2.1. Skema Diagram	10
Gambar 2.2. Skema Diagram Alir	11
Gambar 2. 1 Cara Kerja IndoBERT	19
Gambar 3. 1 Alur Penelitian	22
Gambar 4. 1 Kode Program Proses Cleaning Data	31
Gambar 4. 2 Kode Program Proses Case Folding	32
Gambar 4. 3 Kode Program Proses Normalisasi	32
Gambar 4. 4 Kode Program Proses Labelling	33
Gambar 4. 5 Kode Program Proses EDA	34
Gambar 4. 6 Diagram Batang	35
Gambar 4. 7 Kode Program Proses Visualisasi Word Cloud	36
Gambar 4. 8 Word Cloud Positif	36
Gambar 4. 9 Word Cloud Netral	36
Gambar 4. 10 Word Cloud Negatif	37
Gambar 4. 11 Fine-Tuning Model - Rasio 70:30	38
Gambar 4. 12 Fine-Tuning Model - Rasio 80:20	38
Gambar 4. 13 Fine-Tuning Model - Rasio 90:10	39
Gambar 4. 14 Kode Program Proses Fine-Tuning	40
Gambar 4. 15 Confusion Matrix - 70:30	42
Gambar 4. 16 Confusion Matrix - 80:20	43
Gambar 4. 17 Confusion Matrix - 90:10	44
Gambar 4. 18 Kode Program Proses Evaluasi	45

DAFTAR LAMPIRAN

Lampiran 1. Profil obyek Penelitian	10
Lampiran 2. Dokumentasi Penelitian	11



DAFTAR LAMBANG DAN SINGKATAN



BERT	Bidirectional Encoder Representations from Transformers
CSV	Comma-Separated Values (Format file teks terstruktur)
CRISP-DM	Cross Industry Standard Process for Data Mining (Metodologi standar proses data mining)
EDA	Exploratory Data Analysis
GPU	Graphics Processing Unit
GRU	Gated Recurrent Unit
HTML	Hypertext Markup Language (Bahasa standar pemformatan dokumen web)
KBBI	Kamus Besar Bahasa Indonesia
LSTM	Long Short-Term Memory
MLM	Masked Language Modeling
NLP	Natural Language Processing
NSP	Next Sentence Prediction
PPN	Pajak Pertambahan Nilai
RNN	Recurrent Neural Network
ROS	Random Over Sampling
SMOTE	Synthetic Minority Over-Sampling Technique
SVM	Support Vector Machine
URL	Uniform Resource Locator

DAFTAR ISTILAH

Aktivisme Digital	Gerakan atau aksi menyuarakan aspirasi dan protes yang dilakukan melalui platform digital
Backpropagation	Algoritma pelatihan jaringan saraf tiruan dengan cara menghitung balik nilai <i>error</i> untuk memperbarui bobot model
Batch Size	Jumlah sampel data yang diproses model dalam satu kali iterasi pelatihan
Bidirectional	Kemampuan model untuk memproses dan memahami informasi teks dari dua arah (kiri dan kanan) secara simultan
Case Folding	Tahap preprocessing teks yang dilakukan dengan mengubah seluruh huruf menjadi huruf kecil (<i>lowercase</i>)
Class Imbalance	Kondisi ketidakseimbangan jumlah data antar-kelas/kategori di dalam suatu <i>dataset</i>
Confusion Matrix	Tabel atau matriks evaluasi untuk mengukur performa model klasifikasi dengan membandingkan label asli dan hasil prediksi
Crawling	Proses pengumpulan data teks dari internet secara otomatis menggunakan kata kunci tertentu
Cross-Entropy Loss	Fungsi kerugian (<i>loss function</i>) yang digunakan untuk mengukur performa model klasifikasi pada output probabilitas
Data Loader	Fasilitas untuk mengemas <i>dataset</i> ke dalam kelompok-kelompok kecil (<i>batch</i>) agar proses pelatihan model berjalan lebih efisien
Deep Learning	Sub-bidang kecerdasan buatan berbasis jaringan saraf tiruan multi-layer yang mampu mempelajari pola rumit dari data
Deployment	Tahap pengintegrasian model kecerdasan buatan yang telah diuji ke dalam lingkungan produksi (seperti aplikasi web)
Epoch	Satu kali putaran penuh di mana seluruh <i>dataset</i> selesai dilewati dan dipelajari oleh model selama proses pelatihan
F1-Score	Metrik evaluasi yang menunjukkan nilai keseimbangan rata-rata antara tingkat <i>precision</i> dan <i>recall</i>

Fine-Tuning Proses penyesuaian kembali bobot dari model yang telah dilatih sebelumnya (*pre-trained*) menggunakan *dataset* baru yang lebih spesifik

Forward Propagation Proses penghitungan nilai input secara maju melalui lapisan-lapisan model untuk menghasilkan prediksi output

Hyperparameter Nilai parameter eksternal yang diatur sebelum proses pelatihan model dimulai untuk mengontrol jalannya pembelajaran (seperti *learning rate* dan *epoch*)

Learning Rate Nilai konstanta yang menentukan seberapa besar bobot model disesuaikan atau diubah pada setiap iterasi pelatihan

Lexicon (VADER / InSet) Pendekatan analisis sentimen berbasis kamus kata yang memiliki bobot nilai polaritas emosional tertentu

Noise Data atau elemen teks pengganggu yang tidak relevan dan dapat menurunkan akurasi model (seperti tanda baca acak, karakter non-ASCII, dan spasi berlebih)

Overfitting Kondisi di mana model terlalu fokus mempelajari detail data latih sehingga performanya buruk ketika diuji pada data baru

Precision Rasio prediksi positif model yang benar-benar akurat dibandingkan dengan keseluruhan data yang diprediksi positif

Pre-Training Tahap awal pelatihan model menggunakan korpus teks berskala masif untuk mempelajari struktur dasar bahasa secara umum

Recall Kemampuan model dalam menemukan dan menangkap kembali seluruh data yang termasuk dalam kategori target tertentu

Scraping Teknik mengekstrak atau mengambil data teks secara otomatis dari sebuah halaman situs web atau media sosial

Self-Attention Mekanisme pada arsitektur Transformer yang memungkinkan model menilai tingkat kepentingan dan hubungan antar-kata di dalam suatu kalimat

Stemming Tahap *preprocessing* teks yang bertujuan untuk mengubah kata berimbuhan menjadi kata dasarnya

Stopword Removal Proses eliminasi kata-kata umum yang sering muncul namun tidak memiliki makna penting dalam analisis teks (seperti “yang”, “dan”, “di”)

Stratified Split Teknik pembagian data yang mempertahankan proporsi atau persentase setiap kelas sentimen agar tetap seimbang di seluruh subset data

Text Cleaning Proses pembersihan teks dari karakter pengganggu seperti simbol, angka, emoji, tautan URL, maupun *mention*

Text Mining Teknik penambangan data untuk mengekstrak informasi atau pola berkualitas tinggi dari sekumpulan teks tidak terstruktur

Tokenisasi Proses pemecahan string teks atau kalimat menjadi potongan-potongan kata/sub-kata yang lebih kecil (*token*)

Transformer Arsitektur *deep learning* modern berbasis mekanisme *attention* yang mampu memproses urutan teks secara paralel

Weight Decay Mekanisme regularisasi pada optimizer (seperti AdamW) yang bekerja dengan memberikan penalti pada bobot model demi mencegah risiko *overfitting*



INTISARI

Kemenangan PSS Sleman pada kompetisi Liga 2 Pegadaian memicu respons masif dari para pendukung dan pencinta sepak bola di media sosial, khususnya Instagram. Komentar-komentar tersebut merefleksikan berbagai opini, emosi, dan sentimen publik yang penting untuk dianalisis sebagai bahan evaluasi manajemen klub maupun persepsi masyarakat. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap komentar Instagram mengenai kemenangan PSS Sleman dengan menerapkan algoritma *Bidirectional Encoder Representations from Transformers* (BERT). Metode BERT dipilih karena kemampuannya yang unggul dalam memahami konteks dan semantik kalimat secara dua arah (*bidirectional*). Data penelitian berupa komentar publik di akun Instagram resmi PSS Sleman yang dikumpulkan melalui teknik *scrapping*. Data tersebut melewati tahap *preprocessing* sebelum diklasifikasikan ke dalam kategori sentimen positif dan negatif. Hasil penelitian menunjukkan bahwa model BERT mampu mengklasifikasikan sentimen dengan performa yang sangat baik, menghasilkan akurasi sebesar 92%, *precision* 91%, dan *recall* 93%. Analisis ini membuktikan bahwa algoritma BERT sangat efektif dalam menangani karakteristik teks media sosial yang informal dan kontekstual, sekaligus memberikan gambaran dominasi sentimen positif dari suporter atas kemenangan PSS Sleman.

Kata kunci: Analisis Sentimen, PSS Sleman, Liga 2 Pegadaian, Instagram, BERT.

ABSTRACT

PSS Sleman's victory in the Pegadaian League 2 competition sparked a massive response from supporters and football fans on social media, particularly Instagram. These comments reflect various opinions, emotions, and public sentiments that are important to analyze as material for evaluating club management and public perception. This study aims to conduct sentiment analysis on Instagram comments regarding PSS Sleman's victory by applying the BiDirectional Encoder Representations from Transformers (BERT) algorithm. The BERT method was chosen because of its superior ability to understand the context and semantics of sentences in two directions (bidirectional). The research data consisted of public comments on PSS Sleman's official Instagram account collected through a scraping technique. The data underwent a preprocessing stage before being classified into positive and negative sentiment categories. The results showed that the BERT model was able to classify sentiment with excellent performance, achieving 92% accuracy, 91% precision, and 93% recall. This analysis proves that the BERT algorithm is very effective in handling the informal and contextual characteristics of social media text, while also providing an overview of the dominance of positive sentiment from supporters regarding PSS Sleman's victory.

Keyword: Sentiment Analysis, PSS Sleman, Pegadaian League 2, Instagram, BERT.