



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



**ANALISIS SENTIMEN TEMPORAL
MENGUNAKAN PENDEKATAN HYBRID
INDOBERT-XGBOOST PADA ULASAN APLIKASI
GOJEK**

SKRIPSI

TARISSA AZZAHRA DANANTYA

2207411037

**POLITEKNIK
NEGERI
JAKARTA**

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN
KOMPUTER**

POLITEKNIK NEGERI JAKARTA

2026



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



**ANALISIS SENTIMEN TEMPORAL
MENGUNAKAN PENDEKATAN HYBRID
INDOBERT-XGBOOST PADA ULASAN APLIKASI
GOJEK**

SKRIPSI

**Dibuat untuk Melengkapi Syarat-Syarat yang Diperlukan untuk
Memperoleh Diploma Empat Politeknik**

**TARISSA AZZAHRA DANANTYA
2207411037**

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN
KOMPUTER**

POLITEKNIK NEGERI JAKARTA

2026



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

SURAT PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan di bawah ini:

Nama : Tarissa Azzahra Danantya
NIM : 2207411037
Jurusan/Program Studi : T. Informatika dan Komputer / Teknik Informatika
Judul Skripsi : Analisis Sentimen Temporal Menggunakan Pendekatan Hybrid IndoBERT–XGBoost Pada Ulasan Aplikasi Gojek

Menyatakan dengan sebenarnya bahwa skripsi ini benar-benar merupakan hasil karya saya sendiri, bebas dari peniruan terhadap karya dari orang lain. Kutipan pendapat dan tulisan orang lain ditunjuk sesuai dengan cara-cara penulisan karya ilmiah yang berlaku.

Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa dalam skripsi ini terkandung ciri-ciri plagiat dan bentuk-bentuk peniruan lain yang dianggap melanggar peraturan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Depok, 8 Juni 2026

Tarissa Azzahra Danantya
NIM 2207411037



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

LEMBAR PENGESAHAN

Skripsi diajukan oleh:

Nama : Tarissa Azzahra Danantya
NIM : 2207411037
Program Studi : Teknik Informatika
Judul Skripsi : Analisis Sentimen Temporal Menggunakan Pendekatan Hybrid IndoBERT-XGBoost Pada Ulasan Aplikasi Gojek

Telah diuji oleh tim penguji dalam Sidang Skripsi pada hari Senin, Tanggal 22, Bulan Juni, Tahun 2026 dan dinyatakan LULUS.

Disahkan oleh

Pembimbing I : Rizki Elisa Nalawati, S.T., M.T. ()

Penguji I : Dr. Dewi Yanti Liliana, S.Kom., M.Kom. ()

Penguji II : Iik Muhamad Malik Matin, S.Kom. M.T. ()

Penguji III : Asep Kurniawan, S.Pd., M.Kom. ()

Mengetahui:

Jurusan Teknik Informatika dan Komputer

Ketua



Dr. Anita Hidayati, S.Kom., M.Kom

NIP 197908032003122003



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Sentimen Temporal Menggunakan Pendekatan Hybrid IndoBERT-XGBoost Pada Ulasan Aplikasi Gojek" dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika.

Penyelesaian penelitian ini memperoleh banyak bantuan, arahan, serta dukungan dari berbagai pihak yang berkontribusi selama proses penelitian berlangsung. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua dan keluarga penulis yang senantiasa memberikan doa, dukungan moral, dan semangat yang tak ternilai selama penulis menempuh pendidikan hingga menyelesaikan skripsi ini.
2. Ibu Dr. Anita Hidayati, S.Kom.,M.Kom., selaku Ketua Jurusan Teknik Informatika dan Komputer.
3. Ibu Euis Oktavianti, S.Si., M.T.I., selaku Kepala Program Studi Teknik Informatika sekaligus Dosen Pembimbing Akademik TI B selama 4 tahun masa studi.
4. Ibu Rizki Elisa Nalawati, S.T., M.T., selaku dosen pembimbing yang telah meluangkan waktu, tenaga, dan pikiran untuk memberika bimbingan, arahan, serta masukan yang sangat berharga selama proses pnyusunan skripsi ini.
5. Seluruh Bapak/Ibu Dosen dan staf pengajar yang telah memberikan ilmu pengetahuan dan pengalaman berharga selama penulis menempuh studi.
6. Teman-teman dan semua pihak yang telah membantu serta memberikan dukungan dalam penyelesaian skripsi ini.

Penelitian ini masih memiliki keterbatasan sehingga kritik dan saran yang konstruktif sangat diharapkan untuk pengembangan penelitian selanjutnya Penulis berharap skripsi ini dapat memberikan manfaat serta menjadi referensi bagi



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

penelitian di bidang analisis sentimen, pemrosesan bahasa alami atau *Natural Language Processing*, dan pengembangan sistem berbasis web.

Jakarta, 8 Juni 2026

Tarissa Azzahra Danantya





© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Politeknik Negeri Jakarta, saya bertanda tangan dibawah ini:

Nama : Tarissa Azzahra Danantya

NIM : 2207411037

Jurusan/Program Studi : T. Informatika dan Komputer / Teknik Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Politeknik Negeri Jakarta Hak Bebas Royalti Non-Eksklusif atas karya ilmiah saya yang berjudul:

Analisis Sentimen Temporal Menggunakan Pendekatan Hybrid IndoBERT-XGBoost Pada Ulasan Aplikasi Gojek

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksklusif ini Politeknik Negeri Jakarta Berhak menyimpan, mengalih mediakan/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 8 Juni 2026

Yang Menyatakan

Tarissa Azzahra Danantya

NIM 2207411037



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

ABSTRAK

Ulasan pengguna pada aplikasi Gojek di Google Play Store dapat digunakan untuk memahami persepsi dan pengalaman pengguna terhadap layanan yang diberikan. Penelitian ini bertujuan mengembangkan sistem analisis sentimen temporal berbasis web menggunakan pendekatan *hybrid* IndoBERT-XGBoost pada ulasan aplikasi Gojek. Data diperoleh melalui proses *scraping* sebanyak 53.093 ulasan, kemudian dilakukan *stratified sampling* sebanyak 600 ulasan yang diberi label ke dalam tiga kategori, yaitu positif, netral, dan negatif. Model memanfaatkan kombinasi fitur *embedding* IndoBERT, TF-IDF, fitur linguistik, fitur sentimen berbasis leksikon, dan fitur layanan hasil *Named Entity Recognition* (NER) yang digabungkan melalui *feature concatenation* dan diklasifikasikan menggunakan XGBoost. Optimasi model dilakukan dengan GridSearchCV dan *Stratified K-Fold Cross Validation*. Model terbaik dengan augmentasi teks menghasilkan *accuracy* 0.82, *precision* 0.81, *recall* 0.79, dan *F1-score* 0.80. Sistem diimplementasikan menggunakan ReactJS, Flask, dan PostgreSQL dilengkapi dashboard visualisasi temporal yang menampilkan tren sentimen, *streamgraph*, distribusi layanan, dan *word cloud*. Pengujian *black box* menunjukkan keberhasilan 100%, sedangkan evaluasi *System Usability Scale* (SUS) memperoleh skor rata-rata 78,3 (kategori *Good*). Hasil penelitian menunjukkan bahwa pendekatan *hybrid* IndoBERT-XGBoost mampu mengklasifikasikan sentimen ulasan berbahasa Indonesia, sekaligus efektif dalam mengungkap dinamika perubahan sentimen pengguna dari waktu ke waktu, sehingga dapat menjadi alat bantu yang digunakan untuk memantau persepsi pengguna terhadap layanan digital.

Kata kunci: analisis sentimen temporal, IndoBERT, XGBoost, NER, sistem berbasis web.

POLITEKNIK
NEGERI
JAKARTA



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

DAFTAR ISI

SURAT PERNYATAAN BEBAS PLAGIARISME	iii
LEMBAR PENGESAHAN	iv
KATA PENGANTAR.....	v
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	vii
ABSTRAK	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiv
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan dan Manfaat.....	3
1.4.1 Tujuan Penelitian	3
1.4.2 Manfaat Penelitian	3
1.5 Sistematika Penulisan	3
BAB II	5
TINJAUAN PUSTAKA	5
2.1 Natural Language Processing (NLP).....	5
2.2 Analisis Sentimen	6
2.3 Analisis Sentimen Temporal.....	6
2.4 Model Transformer dan IndoBERT	7
2.5 Named Entity Recognition (NER).....	8
2.6 Feature Engineering pada Analisis Sentimen	8
2.7 Algoritma XGBoost	9
2.8 Pendekatan Hybrid IndoBERT dan XGBoost.....	10
2.9 Sistem Berbasis Web Untuk Analisis Data	10
2.10 Confusion Matrix.....	11



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2.10.1 Accuracy	11
2.10.2 Precision.....	12
2.10.3 Recall	12
2.10.4 F1-Score	12
2.11 Penelitian Terdahulu.....	12
BAB III	16
METODE PENELITIAN.....	16
3.1 Rancangan Penelitian	16
3.1.1 Pendekatan Penelitian	16
3.1.2 Jenis Penelitian.....	16
3.1.3 Teknik Pengumpulan.....	17
3.1.4 Teknik Analisis Data	17
3.2 Tahap Penelitian.....	18
3.2.1 Identifikasi Masalah dan Studi Literatur.....	19
3.2.2 Pengumpulan Data	19
3.2.3 Pelabelan Data (Expert Labeling).....	21
3.2.4 Pengembangan Model.....	22
3.2.5 Pengembangan Sistem Berbasis Web	23
3.2.6 Pengujian Sistem	24
3.2.7 Penulisan Laporan	25
3.3 Objek Penelitian.....	25
3.3.1 Objek dan Subjek Penelitian	25
3.3.2 Ruang Lingkup Penelitian	25
BAB IV	27
HASIL DAN PEMBAHASAN.....	27
4.1 Analisis Kebutuhan	27
4.1.1 Analisis Kebutuhan Model.....	27
4.1.2 Kebutuhan Sistem Berbasis Web	28
4.2 Perancangan Sistem.....	29
4.2.1 Arsitektur Sistem	29
4.2.2 Diagram UML.....	30
4.2.3 Alur Sistem	40
4.2.4 Perancangan Database.....	42
4.2.5 Perancangan API	45



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.2.6 Perancangan Dashboard Temporal	46
4.3 Implementasi Sistem	51
4.3.1 Implementasi Model	51
4.3.2 Implementasi Sistem Berbasis Web	69
4.4 Pengujian Sistem dan Model	83
4.4.1 Deskripsi Pengujian	83
4.4.2 Prosedur Pengujian Model	84
4.4.3 Data Hasil Pengujian	90
4.4.4 Analisis Hasil Pengujian	97
BAB V	101
PENUTUP	101
5.1 Simpulan	101
5.2 Saran	102
DAFTAR PUSTAKA	104
DAFTAR RIWAYAT HIDUP	107
LAMPIRAN	108

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian.....	19
Gambar 3.2 Dataset Hasil <i>Scraping</i>	20
Gambar 3.3 Dataset Hasil Stratified Sampling	20
Gambar 3.4 Dataset Hasil Pelabelan Manual.....	22
Gambar 4.1 Arsitektur Sistem.....	30
Gambar 4.2 Use Case Diagram	31
Gambar 4.3 Activity Diagram Registrasi.....	32
Gambar 4.4 Activity Diagram Login.....	33
Gambar 4.5 Activity Diagram Reset Password.....	34
Gambar 4.6 Activity Diagram Melihat Dashboard.....	35
Gambar 4.7 Activity Diagram Klasifikasi Sentimen Teks Input	36
Gambar 4.8 Activity Diagram Klasifikasi Sentimen Dataset.....	37
Gambar 4.9 Activity Diagram Klasifikasi Sentimen Scraping Data.....	39
Gambar 4.10 Activity Diagram Melihat Historis Analisis.....	40
Gambar 4.11 Flowchart Sistem	41
Gambar 4.12 Relasi Antar Tabel	44
Gambar 4.13 Wireframe Halaman Register, Login, dan Reset Password	47
Gambar 4.14 Wireframe Halaman Dashboard.....	48
Gambar 4.15 Wireframe Halaman Analisis.....	49
Gambar 4.16 Wireframe Halaman Histori	50
Gambar 4.17 SHAP Feature Importance Netral.....	62
Gambar 4.18 SHAP Feature Importance Positif	63
Gambar 4.19 SHAP Feature Importance Negatif.....	64
Gambar 4.20 SHAP Force Plot	65
Gambar 4.21 Attention Saliency pada Prediksi Sentimen Negatif.....	66
Gambar 4.22 Attention Saliency pada Prediksi Sentimen Positif	67
Gambar 4.23 Attention Saliency Kesalahan Klasifikasi Positif Menjadi Negatif.....	67
Gambar 4.24 Attention Saliency Kesalahan Klasifikasi Netral Menjadi Negatif.....	68
Gambar 4.25 Attention Saliency Kesalahan Klasifikasi Netral Menjadi Positif	68
Gambar 4.26 Tampilan Halaman Registrasi.....	70
Gambar 4.27 Tampilan Halaman Login	71



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Gambar 4.28 Tampilan Halaman Reset Password.....	72
Gambar 4.29 Halaman Dashboard	73
Gambar 4.30 Ringkasan Temporal	74
Gambar 4.31 Grafik Tren Sentimen	75
Gambar 4.32 Streamgraph Sentimen.....	76
Gambar 4.33 Distribusi Sentimen Layanan Gojek	77
Gambar 4.34 Word cloud Seluruh Ulasan.....	78
Gambar 4.35 Word cloud Sentimen Positif.....	79
Gambar 4.36 Word cloud Sentimen Netral	79
Gambar 4.37 Word cloud Sentimen Negatif	80
Gambar 4.38 Tampilan Halaman Analisis Teks Manual	80
Gambar 4.39 Tampilan Halaman Analisis Upload Dataset.....	81
Gambar 4.40 Tampilan Halaman Analisis Scraping.....	82
Gambar 4.41 Tampilan Halaman Historis.....	83
Gambar 4.42 Confusion Matrix Model Akhir.....	92

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

DAFTAR TABEL

Tabel 4.1 Struktur Tabel Users.....	42
Tabel 4.2 Struktur Tabel Reviews	43
Tabel 4.3 Endpoint Sistem.....	45
Tabel 4.4 Contoh Hasil Preprocessing Teks.....	52
Tabel 4.5 Hasil Named Entity Recognition (NER).....	56
Tabel 4.6 Fitur Sentimen Berbasis Leksikon	58
Tabel 4.7 Skenario Black Box Testing	85
Tabel 4.8 Daftar Pernyataan Kuesioner SUS.....	88
Tabel 4.9 Hasil Eksperimen Konfigurasi Arsitektur Model.....	90
Tabel 4.10 Hasil Eksperimen Teknik Peningkatan Data	91
Tabel 4.11 Hasil Perhitungan Per Kelas	94
Tabel 4.12 Hasil Pengujian Sistem Berbasis Web.....	94
Tabel 4.13 Skor System Usablity Scale (SUS) PerResponden.....	96

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB I PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital telah mengubah cara perusahaan berinteraksi dengan pengguna, di mana aplikasi *mobile* menjadi sarana utama dalam menyediakan layanan dan membangun pengalaman pengguna. Di Indonesia, aplikasi layanan berbasis digital seperti Gojek memainkan peran penting dalam mendukung mobilitas, transaksi keuangan, dan kebutuhan harian masyarakat. Sebagai aplikasi *super-app* dengan berbagai layanan terintegrasi, gojek secara berkala melakukan pembaruan fitur, penyesuaian kebijakan, serta peningkatan sistem untuk menjaga daya saing dan kepuasan pengguna.

Perubahan fitur, pembaruan sistem, maupun kebijakan layanan sering kali memicu beragam respons dari pengguna yang tercermin dalam komentar dan penilaian pengguna yang dipublikasikan melalui Google Play Store. Informasi yang terkandung dalam ulasan tersebut dapat mengalami fluktuasi seiring waktu, terutama setelah adanya pembaruan aplikasi, gangguan sistem, atau perubahan layanan tertentu. Pola perubahan sentimen ini menjadi penting untuk dianalisis karena dapat menggambarkan persepsi dan tingkat kepuasan pengguna secara dinamis.

Meskipun demikian, berbagai studi terdahulu analisis sentimen lebih banyak diarahkan pada klasifikasi statis dengan mengabaikan perkembangan pola dinamika perubahan sentimen dari waktu ke waktu. Di sisi lain, integrasi antara analisis sentimen dengan identifikasi fitur layanan spesifik menggunakan *Named Entity Recognition* (NER) untuk mengetahui faktor penyebab perubahan sentimen pengguna masih relatif terbatas dalam penelitian terdahulu.

Analisis sentimen tradisional sering kali hanya berfokus pada klasifikasi positif atau negatif tanpa mempertimbangkan aspek temporal dan faktor kontekstual spesifik. Padahal, evaluasi sentimen dari waktu ke waktu, mengidentifikasi pemicu perubahan, dan memahami nuansa linguistik dalam ekspresi pengguna (Tan *et al.*,

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2022). Ulasan pengguna memiliki karakteristik unik, seperti penggunaan bahasa gaul, campuran kode (*code-mixing*), dan referensi budaya lokal yang memerlukan metode yang mampu menyesuaikan diri terhadap karakteristik data. Berbagai metode pembelajaran mesin telah diterapkan untuk menyelesaikan permasalahan ini, khususnya model berbasis sekuens yang mampu memodelkan hubungan konteks dalam teks. Meskipun demikian, pendekatan tersebut memiliki keterbatasan dari segi efisiensi komputasi akibat proses pemodelan yang bersifat berurutan. Model *transformer* hadir sebagai solusi dengan memanfaatkan mekanisme pemrosesan secara bersamaan (*parallel processing*) yang memungkinkan waktu komputasi menjadi lebih efisien, sekaligus memperbaiki kemampuan dalam merepresentasikan hubungan kontekstual antar kata. (Tan *et al.*, 2022).

Berdasarkan hal tersebut, penelitian ini mengembangkan sistem analisis sentimen temporal menggunakan pendekatan *hybrid* IndoBERT-XGBoost pada ulasan aplikasi mobile berbahasa Indonesia. Studi kasus dilakukan pada salah satu aplikasi layanan digital populer di Indonesia guna mengevaluasi kemampuan sistem dalam mengidentifikasi pola pergeseran sentimen pengguna dari waktu ke waktu.

1.2 Perumusan Masalah

Mengacu pada permasalahan yang telah diuraikan sebelumnya, penelitian ini berupaya menjawab pertanyaan bagaimana merancang dan mengimplementasikan sistem analisis sentimen temporal menggunakan pendekatan *hybrid* IndoBERT dan XGBoost pada ulasan aplikasi Gojek berbahasa Indonesia di Google Play Store untuk mengidentifikasi dinamika perubahan sentimen pengguna dari waktu ke waktu?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Dataset penelitian berasal dari hasil scraping ulasan pengguna aplikasi Gojek melalui Google Play Store sebanyak 53.093 ulasan. Namun,



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

penelitian membatasi proses pelabelan manual dan pelatihan model hanya menggunakan 600 ulasan yang dipilih melalui Teknik sampling.

2. Analisis yang dilakukan terhadap data ulasan yang ditulis dalam bahasa Indonesia.
3. Metode yang digunakan adalah pendekatan *hybrid* antara IndoBERT sebagai ekstraksi fitur dan XGBoost sebagai *classifier*.
4. Analisis temporal dilakukan berdasarkan agregasi sentimen dalam rentang waktu 12 bulan, dari bulan februari 2025 sampai januari 2026.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan Penelitian

Penelitian ini bertujuan untuk membangun dan menerapkan sebuah sistem analisis sentimen temporal menggunakan pendekatan *hybrid* IndoBERT dan XGBoost berbasis web pada ulasan aplikasi *mobile* berbahasa Indonesia guna mengidentifikasi pola dan dinamika perubahan sentimen pengguna secara temporal.

1.4.2 Manfaat Penelitian

Adapun manfaat yang diharapkan dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Menghasilkan sebuah aplikasi berbasis web yang memiliki kemampuan dalam menganalisis dan menampilkan sentimen dari ulasan pengguna secara temporal dan visual.
2. Memberikan *insight* mengenai tren persepsi pengguna terhadap aplikasi dari waktu ke waktu.
3. Memberikan informasi dan referensi bagi pengembang aplikasi, tim data analyst atau perusahaan dalam memantau kepuasan pengguna dan mengevaluasi dampak pembaruan layanan.

1.5 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun sebagai berikut:



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

1. Bab I - Pendahuluan

Bab ini memaparkan latar belakang penelitian, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan.

2. Bab II - Tinjauan Pustaka

Bab ini berisi landasan teori, penelitian terkait, konsep analisis sentimen, analisis sentimen temporal, *feature engineering*, IndoBERT, XGBoost, serta pengembangan sistem web.

3. Bab III - Metodologi Penelitian

Bab ini menjelaskan rancangan penelitian yang berisi pendekatan penelitian, jenis penelitian, teknik pengumpulan dan analisis data. Selain itu menjelaskan tahapan penelitian, serta objek penelitian.

4. Bab IV – Hasil dan Pembahasan

Bab ini menyajikan hasil implementasi sistem,, analisis sentimen temporal, hasil pengujian sistem, serta pembahasan hasil penelitian.

5. Bab V – Penutup

Bab ini berisi Kesimpulan dari hasil penelitian dan saran untuk pengembangan penelitian selanjutnya.

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB II TINJAUAN PUSTAKA

2.1 Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan salah satu cabang ilmu dalam bidang kecerdasan buatan (AI) yang berfokus pada interaksi antara sistem komputer dengan bahasa yang digunakan manusia manusia. NLP memegang peranan penting yang signifikan karena memungkinkan komputer untuk memproses, memahami, serta menghasilkan bahasa alami manusia, baik yang berbentuk tulisan maupun ucapan. Secara sederhana, NLP bertujuan agar mesin dapat memahami bahasa manusia sehari-hari sehingga interaksi antara manusia dan komputer menjadi lebih mudah, alami, dan efektif (Tarumingkeng, 2024). Tujuan utama dari bidang ini adalah agar mesin dapat berinteraksi secara lebih efisien dengan teks dan data dalam bahasa alami. NLP digunakan untuk menangani berbagai aspek bahasa seperti struktur (sintaksis), bentuk kata (morfologi), dan makna (semantik). Hal ini sangat penting karena setiap bahasa, termasuk bahasa Indonesia, memiliki karakteristik unik yang memerlukan pendekatan khusus dalam pengembangannya (Amien, 2023).

Menurut (Amien, 2023), teknologi dasar yang dikembangkan dalam NLP ialah, stemming yang merupakan proses penghilangan imbuhan (awalan, sisipan, dan akhiran) untuk mendapatkan kata dasar, Part-of-Speech (POS) Tagging yaitu penandaan kelas kata dalam suatu kalimat, seperti kata benda atau kata kerja, *Named Entity Recognition* (NER) yaitu pengenalan entitas bernama seperti nama orang, organisasi, atau lokasi, dan *Sentiment Analysis* yaitu analisis untuk mengidentifikasi opini atau emosi dalam teks. Pengembangan NLP menggunakan berbagai metode, mulai dari pembelajaran mesin (*machine learning*), terjemahan mesin berbasis statistik (SMT), hingga pendekatan terbaru seperti *Neural Machine Translation* (NMT) yang menggunakan jaringan saraf untuk hasil yang lebih akurat. NLP juga diterapkan dalam berbagai sistem, seperti pencarian informasi lintas bahasa, ekstraksi informasi dari teks yang tidak terstruktur, sistem pengecekan ejaan,

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

serta pemantauan opini publik di media sosial melalui analisis sentimen (Amien, 2023).

2.2 Analisis Sentimen

Analisis sentimen dapat di definisikan sebagai suatu metode berbasis komputasi yang dimanfaatkan untuk mengidentifikasi dan mengelompokkan pandangan, sikap, serta emosi yang disampaikan oleh seseorang terhadap objek, peristiwa atau isu tertentu. Proses ini bertujuan untuk menentukan kecenderungan sentimen dalam sebuah kalimat atau dokumen, apakah condong kearah positif atau negatif (Hakim and Kinasih, 2023). Secara garis besar, terdapat empat pendekatan yang digunakan dalam analisis sentimen, yaitu pendekatan berbasis leksikon, *machine learning*, *deep learning*, dan *hybrid*. Pendekatan berbasis leksikon bekerja dengan mengacu pada kamus sentimen tanpa memerlukan data latih, namun kurang fleksibel karena sangat bergantung pada konteks dan domain. Pendekatan *machine learning* memandang analisis sentimen sebagai tugas klasifikasi teks berbasis data berlabel dan umumnya menghasilkan akurasi lebih baik, meskipun membutuhkan dataset yang cukup besar. Pendekatan *deep learning* mengembangkan *machine learning* dengan jaringan saraf berlapis, seperti RNN, LSTM, hingga *Transformer*, yang mampu menangkap konteks secara lebih mendalam dan efisien melalui mekanisme *self-attention*, terutama untuk teks pendek dan informal seperti ulasan berbahasa Indonesia. Sementara itu, pendekatan *hybrid* menggabungkan keunggulan leksikon dan *machine learning* atau *deep learning* untuk meningkatkan performa analisis sentimen dengan memanfaatkan pengetahuan linguistik sekaligus kemampuan pembelajaran statistik dalam menangkap pola sentimen yang kompleks (Wankhade, Rao and Kulkarn, 2022).

2.3 Analisis Sentimen Temporal

Analisis deret waktu adalah metode analisis data yang digunakan untuk mempelajari pola, tren, dan perubahan suatu fenomena berdasarkan urutan waktu tertentu. Dalam konteks media sosial, analisis deret waktu memungkinkan peneliti untuk mengamati dinamika aktivitas pengguna dan perubahan sentimen dari waktu



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

ke waktu (Farizah and Wang, 2019). Analisis sentimen statis memperlakukan teks sebagai dokumen yang berdiri sendiri tanpa mempertimbangkan dimensi waktu, sehingga hanya menghasilkan gambaran agregat distribusi sentimen dan kurang mampu menangkap dinamika perubahan opini. Sebaliknya, analisis sentimen temporal mengaitkan hasil klasifikasi dengan informasi waktu dan mengagregasikannya dalam unit tertentu (harian, mingguan, atau bulanan), sehingga membentuk deret waktu sentimen yang memungkinkan pengamatan tren, fluktuasi, serta perubahan sentimen akibat peristiwa tertentu, menjawab tidak hanya apa sentimennya tetapi juga kapan dan bagaimana sentimen tersebut berubah (Harris, Jacobson and Proveti, 2024).

2.4 Model Transformer dan IndoBERT

Salah satu model *deep learning* ialah berbasis *transformer* seperti BERT (*Bidirectional Encoder Representations from Transformers*) telah menjadi standar dalam pembuatan NLP, termasuk analisis sentimen. BERT mampu memahami konteks kata secara dua arah, sehingga lebih akurat dalam menangani ambiguitas makna, bahasa informal, serta variasi ekspresi opini di media sosial. Selain itu, pendekatan *transfer learning* pada BERT memungkinkan model yang telah dilatih pada korpus besar untuk disesuaikan (*fine-tuning*) pada tugas analisis sentimen dengan data yang relatif lebih kecil (Wankhade, Rao and Kulkarn, 2022). Model BERT multibahasa dinilai kurang optimal untuk bahasa Indonesia karena keterbatasannya dalam menangkap karakteristik linguistik lokal seperti struktur kalimat, bahasa informal, dan slang, sehingga dikembangkan IndoBERT yang dilatih khusus pada korpus berbahasa Indonesia. Pelatihan berbasis data lokal ini membuat IndoBERT lebih mampu memahami konteks, tata bahasa, dan variasi bahasa Indonesia, dan berbagai penelitian menunjukkan bahwa *fine-tuning* IndoBERT pada tugas analisis sentimen menghasilkan kinerja yang lebih unggul dibandingkan model multibahasa maupun metode *machine learning* tradisional, sehingga sangat relevan untuk melakukan analisis teks ulasan dan media sosial berbahasa Indonesia yang bersifat tidak terstruktur dan informal (Widyananda, Fauzi and Maskur, 2025).

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2.5 Named Entity Recognition (NER)

Named Entity Recognition atau pengenalan entitas bernama, merupakan salah satu komponen dalam NLP yang digunakan untuk mengidentifikasi dan menandai entitas tertentu dalam suatu teks. Entitas yang dimaksud mencakup kata atau frasa yang merepresentasikan informasi khusus seperti nama orang, lokasi, organisasi, tanggal, waktu, maupun angka. NER memiliki peran penting karna membantu sistem komputer memahami informasi yang bersifat spesifik dalam teks serta mengelompokkannya ke dalam kategori yang sesuai. Dengan kemampuan ini, sistem dapat mendukung berbagai tugas seperti ekstraksi informasi, klasifikasi data, hingga peringkasan teks secara lebih akurat dan efisien (Tarumingkeng, 2024). Proses ekstraksi informasi melibatkan identifikasi penggalan informasi relevan dari teks yang tidak terstruktur, termasuk entitas yang diberi nama, hubungan antar entitas, dan peristiwa tertentu. Dalam konteks perkembangan NLP di Indonesia, NER dikategorikan sebagai salah satu teknologi dasar yang menjadi landasan bagi pengembang aplikasi NLP yang lebih kompleks. Teknologi pengenalan entitas bernama ini telah diimplementasikan dalam berbagai sektor industri, seperti perbankan, pemerintahan, dan media. Selain itu, NER juga digunakan dalam pengembangan sistem pencarian informasi dan sistem rekomendasi (Amien, 2023)'

2.6 Feature Engineering pada Analisis Sentimen

Feature engineering berperan penting dalam analisis sentimen karena menentukan kemampuan model dalam merepresentasikan makna, emosi, dan kecenderungan opini dari teks ke dalam bentuk numerik. Secara umum, fitur linguistik dalam analisis sentimen dikelompokkan menjadi fitur leksikal, sintaktik, dan pragmatik atau *domain-specific*. Fitur leksikal berfokus pada kata atau unit teks, seperti *bag of words* dan TF-IDF yang merepresentasikan frekuensi kata, serta *word embeddings* (Word2Vec, FastText) yang mampu menangkap hubungan semantik antar kata dan efektif dalam menangani variasi bahasa informal pada ulasan berbahasa Indonesia. Fitur sintaktik mencakup informasi struktur bahasa seperti *part-of-speech*, bentuk kata kerja, dan negasi yang membantu memperjelas orientasi sentimen. Sementara itu, fitur pragmatik atau *domain-specific*



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

mempertimbangkan konteks dan karakteristik topik tertentu, seperti penggunaan kamus sentimen, emosi, istilah populer, dan bahasa gaul, yang penting untuk memahami makna sentimen yang bergantung pada konteks penggunaan bahasa, khususnya pada teks ulasan dan media sosial (Plaza and Carrillo-de-albornoz, 2018).

2.7 Algoritma XGBoost

XGBoost atau *Extreme Gradient Boosting* adalah algoritma pembelajaran mesin (*machine learning*) yang merupakan implementasi dari *gradient boosted decision trees*. XGBoost merupakan algoritma berbasis ensemble yang menggabungkan kekuatan prediksi dari beberapa model "learner". Algoritma ini menggabungkan banyak pembelajar lemah (*weak learner*) untuk membangun satu pembelajar yang kuat (*strong learner*). Dalam teknik *boosting*, pohon-pohon keputusan (*trees*) dibangun secara berurutan atau sekuensial, yang artinya setiap pohon baru yang dibuat bertujuan untuk mengurangi kesalahan (*residual errors*) dari pohon sebelumnya. Pohon berikutnya akan belajar dari sisa kesalahan yang diperbarui tersebut. Keunggulan utamanya, XGBoost dikenal dalam mengoptimalkan kecepatan dan penggunaan sumber daya teori dan terbukti sangat baik dalam menangani data yang tidak seimbang ((Afifah, Yulita and Sarathan, 2021).

XGBoost digunakan sebagai metode klasifikasi sentimen untuk memprediksi sentimen dengan kategori positif, netral, atau negatif, dengan keunggulan dalam menangani dataset besar dan fitur kompleks dibandingkan Naïve Bayes maupun SVM, serta dilengkapi mekanisme regularisasi bawaan untuk mengurangi *overfitting*. Model ini bekerja melalui teknik *boosting* dengan meminimalkan fungsi kerugian secara bertahap menggunakan informasi gradien, sementara optimasi *hyperparameter* dilakukan menggunakan GridSearchCV. Dalam analisis sentimen, XGBoost memanfaatkan fitur *embedding* dan fitur linguistik, serta menyediakan *feature importance* yang mendukung interpretasi kontribusi setiap fitur terhadap hasil klasifikasi (Nur and Handoko, 2025).

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2.8 Pendekatan Hybrid IndoBERT dan XGBoost

Pendekatan *hybrid* dalam analisis sentimen mengombinasikan *embedding* kontekstual berbasis *transformer* dengan model *machine learning* klasik untuk meningkatkan kinerja klasifikasi. Pada pendekatan ini, IndoBERT digunakan sebagai *text embedding* yang merepresentasikan makna semantik teks secara kontekstual dan dua arah, sementara XGBoost berperan sebagai *classifier* yang memanfaatkan *embedding* tersebut bersama fitur linguistik hasil *feature engineering* (Moeslim, Firmasyah and Sutara, 2025). Parameter IndoBERT hasil *pre-training* dipertahankan sehingga tidak diperbarui selama proses pelatihan. Representasi yang digunakan sebagai masukan bagi XGBoost diperoleh dari *embedding* token [CLS], yang merepresentasikan makna keseluruhan kalimat (Putri *et al.*, 2026)

Kombinasi ini memungkinkan model menangkap konteks bahasa Indonesia, termasuk variasi informal dan istilah *domain-specific*, sekaligus memanfaatkan kekuatan XGBoost dalam mengolah fitur kompleks. Selain itu, mekanisme regularisasi pada XGBoost membantu mengurangi risiko *overfitting*, terutama pada dataset berukuran terbatas. Pemilihan pendekatan *hybrid* didasarkan pada pertimbangan *trade-off* antara akurasi dan efisiensi komputasi. Meskipun model *transformer* menawarkan performa tinggi, kebutuhan komputasinya relatif besar, sedangkan XGBoost lebih efisien dan stabil. Dengan menggunakan *embedding* IndoBERT sebagai fitur tanpa *fine-tuning* penuh, pendekatan *hybrid* mampu mempertahankan pemahaman konteks yang kuat sekaligus mengurangi biaya komputasi (Moeslim, Firmasyah and Sutara, 2025).

2.9 Sistem Berbasis Web Untuk Analisis Data

Web sistem dapat diartikan sebagai suatu *platform* berbasis web yang memungkinkan pengelolaan dan integrasi data secara terpusat dan *real-time* guna mendukung pengambilan keputusan. Keunggulan utama aplikasi web meliputi kemudahan akses lintas perangkat, pemeliharaan yang sederhana, serta integrasi data yang efisien tanpa memerlukan instalasi khusus. Dalam konteks analisis data, sistem berbasis web efektif digunakan untuk menampilkan dashboard analitik yang



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

menyajikan ringkasan informasi, monitoring, serta visualisasi data temporal secara interaktif dan mudah dipahami (Godang *et al.*, 2025). Visualisasi data digunakan untuk menyajikan informasi secara ringkas dan mudah dipahami melalui grafik. Dalam analisis temporal, data disusun menurut urutan waktu tertentu sehingga perubahan, tren, dan fluktuasi dapat diamati secara jelas. Penyajian ini membantu pengguna memahami dinamika data dari waktu ke waktu secara cepat dan informatif (Jannah, 2024).

2.10 Confusion Matrix

Confusion matrix adalah sebuah tabel kontingensi berukuran 2x2 yang digunakan sebagai prosedur standar untuk mengevaluasi hasil dari klasifikasi biner. Matrix ini berfungsi untuk memetakan hasil prediksi model pembelajaran mesin terhadap kategori data yang sebenarnya (Chicco and Jurman, 2022).

Tabel ini membagi hasil prediksi ke dalam empat kategori dasar:

1. True Positive (TP): Data positif yang diprediksi dengan benar sebagai positif.
2. True Negative (TN): Data negatif yang diprediksi dengan benar sebagai negatif.
3. False Negative (FN): Data positif yang salah diberi label sebagai negatif.
4. False Positive (FP): Data negatif yang salah diberi label sebagai positif.

Untuk permasalahan dalam klasifikasi, pengukuran yang biasa digunakan adalah akurasi, precision dan recall yang dapat dihitung dengan rumus berikut (Obi, 2023):

Predicted Values	Actual Values	
	Positive (P)	Negative (N)
Positive (P)	TP	FP
Negative (N)	FN	TN

2.10.1 Accuracy

Akurasi (*Accuracy*) merupakan metrik yang menunjukkan proporsi prediksi yang benar terhadap seluruh data uji. Dengan kata lain, akurasi mengukur seberapa banyak data yang berhasil diklasifikasikan dengan benar oleh model. Nilai akurasi berbanding terbalik dengan tingkat kesalahan (*error rate*), sehingga keduanya mencerminkan informasi yang sama mengenai performa suatu model klasifikasi, secara matematis, akurasi dirumuskan sebagai berikut:

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (1)$$

2.10.2 Precision

Presisi (*Precision*) merupakan rasio antara jumlah *true positive* (TP) terhadap seluruh data yang diprediksi sebagai positif. Seluruh prediksi positif tersebut terdiri atas *true positive* (TP) dan *false positive* (FP), yaitu data yang sebenarnya negatif tetapi salah diprediksi sebagai positif. Presisi dirumuskan sebagai berikut:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (2)$$

2.10.3 Recall

Recall dikenal juga sebagai sensitivitas (*sensitivity*) atau *true positive rate*, merupakan kemampuan suatu model dalam mengidentifikasi data yang benar-benar positif dengan tepat. *Recall* dirumuskan sebagai berikut:

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (3)$$

2.10.4 F1-Score

F1-Score merupakan rata-rata harmonik atau rata-rata tertimbang dari nilai *precision* dan *recall*. Secara umum, *F1-score* digunakan sebagai ukuran untuk mengevaluasi tingkat akurasi suatu model klasifikasi dan dihitung berdasarkan nilai *precision* serta *recall* yang diperoleh dari hasil pengujian. *F1-score* digunakan untuk ketika kesalahan klasifikasi menjadi pertimbangan atau ketika distribusi kelas tidak seimbang. *F1-score* dirumuskan sebagai berikut:

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

2.11 Penelitian Terdahulu

Tabel 2.1. Penelitian Terdahulu

Penulis	Judul	Metode	Hasil	Persamaan dan Perbedaan
(Putri <i>et al.</i> , 2026)	IndoBERT Embeddings vs. Fine-Tuning: An	IndoBERT Fine-Tune and	Hasil penelitian menunjukkan bahwa kombinasi IndoBERT-XGBoost	Persamaan: menggunakan indobert sebagai



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Penulis	Judul	Metode	Hasil	Persamaan dan Perbedaan
	Optimized Approach for Sentiment Analysis of Indonesian Public Service Reviews	IndoBERT XGBoost	memperoleh tingkat akurasi tertinggi, yaitu sebesar 98,2%, sedikit lebih tinggi dibandingkan model IndoBERT yang sepenuhnya di-fine-tune.	embedding dan XGBoost sebagai classifier Perbedaan: fitur-fitur tambahan yang digunakan berbeda
(Widyananda, Fauzi and Maskur, 2025)	Machine Learning and Transformer-based Model for Sentiment Analysis of Indonesian E-Commerce Reviews	Naïve Bayes, Support Vector Machine (SVM), dan IndoBERT	Penelitian tersebut menyimpulkan bahwa penggunaan model IndoBERT memberikan performa klasifikasi sentimen yang lebih unggul dibandingkan metode pembandingan yang digunakan. Pada dataset yang digunakan, metode ini memperoleh akurasi sebesar 94,1%, mengungguli SVM yang memiliki akurasi 89,5% dan Multinomial Naïve	Persamaan: Menggunakan Indobert Perbedaan: Tidak menggunakan model <i>hybrid</i>

© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Penulis	Judul	Metode	Hasil	Persamaan dan Perbedaan
			Bayes yang memiliki akurasi 84,2%.	
(Imadudin, A and Nugroho, 2023)	Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach	IndoBERT	Hasil penelitian menunjukkan performa yang sangat baik, model memperoleh nilai akurasi 96%, presisi 95%, recall 96%, dan F1-score 95%.	Persamaan: Menggunakan model Indobert Perbedaan: Tidak menggunakan <i>hybrid</i> sebagai metodenya.
(Adhim and Cahyono, 2025)	Optimization of IndoBERT for Sentiment Analysis of FOMO on Social Media Through Fine-Tuning and Hybrid Labeling	IndoBert dan BiLSTM, SVM, Naive Bayes	Hasil pengujian menunjukkan bahwa model IndoBERT yang telah melalui proses optimasi mampu mencapai performa baik dengan nilai akurasi 94,50%, F1-score 94,52%, serta Macro AUC sebesar 0,987. Hasil ini secara signifikan melampaui model pembandingan, termasuk BiLSTM (akurasi 86,60%), Support Vector	Persamaan: Menggunakan model IndoBERT sebagai model analisis sentimen serta menggunakan fitur linguisitik leksikon Perbedaan: tidak menggunakan metode <i>hybrid</i>



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Penulis	Judul	Metode	Hasil	Persamaan dan Perbedaan
			Machine (88,06%), dan Naive Bayes (80,73%).	
(Faqih, Aji and Suseno, 2025)	Studi Komparatif Metode Naive Bayes dan Support Vector Machine dalam Analisis Sentimen Ulasan Ask-AI	Naive Bayes dan Support Vector Machine	Hasil pengujian memperlihatkan bahwa algoritma SVM unggul dibandingkan Naive Bayes dalam hal performa klasifikasi. SVM mencapai accuracy 94,08%, precision 96,23%, recall 92,31%, dan AUC sebesar 0,981. Sementara itu, Naive Bayes hanya mencapai accuracy 78%, precision 85,23%, recall 68,37%, dengan AUC 0,801, yang menunjukkan performa jauh lebih rendah dibandingkan SVM pada keseluruhan metrik evaluasi.	Persamaan: Analisis sentimen Ulasan pengguna dan klasifikasi ulasan ke dalam kategori sentimen positif dan negatif. Perbedaan: Menggunakan pendekatan model yang berbeda



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB III

METODE PENELITIAN

3.1 Rancangan Penelitian

3.1.1 Pendekatan Penelitian

Pendekatan yang digunakan dalam penelitian ini ialah kuantitatif berbasis rekayasa sistem (*engineering research*). Pendekatan ini digunakan karena penelitian tidak hanya berfokus pada analisis data, tetapi juga membangun sebuah sistem berbasis teknologi *Natural Language Processing* (NLP) dan *machine learning* yang mampu mengolah data ulasan pengguna secara otomatis dan menyajikan hasil analisis dalam bentuk tren sentimen temporal. Tujuan utama penelitian ini adalah merancang dan mengimplementasi sistem analisis sentimen temporal menggunakan pendekatan *hybrid* antara IndoBERT dan XGBoost serta fitur tambahan berupa TF-IDF, fitur linguistik, dan *Named Entity Recognition* (NER).

Sistem yang dibangun memiliki kemampuan untuk:

- Mengolah data ulasan pengguna secara otomatis
- Mengidentifikasi layanan aplikasi yang dibahas pengguna
- Mengklasifikasikan sentimen menjadi positif, netral, dan negatif
- Menganalisis perubahan sentimen secara temporal
- Menyajikan hasil analisis dalam bentuk visualisasi tren waktu pada dashboard berbasis web

Penelitian ini bersifat deskriptif-komputasional yang bertujuan untuk menganalisis perubahan sentimen pengguna dari waktu ke waktu tanpa melakukan peramalan sentimen di masa mendatang.

3.1.2 Jenis Penelitian

Ditinjau dari jenis penelitiannya, penelitian eksperimental berbasis rekayasa (*experimental engineering research*). Jenis ini dipilih karena penelitian melibatkan perancangan, pembangunan, dan pengujian sistem secara langsung. Model klasifikasi sentimen dibangun, dilatih, dan dievaluasi menggunakan data ulasan nyata, kemudian diintegrasikan ke dalam sebuah sistem berbasis web yang mampu



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

melakukan analisis sentimen secara *real-time*. Proses pengumpulan data dilakukan melalui teknik *web scraping* pada Google Play Store untuk memperoleh ulasan pengguna aplikasi Gojek. Sementara itu, teknik analisis data mencakup *preprocessing* teks, pelabelan data (*expert labeling*), augmentasi teks, *feature extraction* dan *feature engineering*, klasifikasi sentimen menggunakan model hybrid IndoBERT-XGBoost, evaluasi model, serta analisis temporal.

3.1.3 Teknik Pengumpulan

Pengumpulan data dalam penelitian ini dilakukan dengan menerapkan teknik *web scraping* pada ulasan pengguna aplikasi Gojek yang tersedia di Google Play Store. Teknik tersebut digunakan untuk memperoleh data ulasan pengguna beserta metadata pendukung seperti tanggal ulasan, *rating* aplikasi, versi aplikasi dan metadata lainnya. Data yang dikumpulkan merupakan ulasan berbahasa Indonesia dalam rentang waktu tertentu sehingga dapat digunakan untuk analisis sentimen temporal. Hasil *scraping* menghasilkan sekitar 53.093 data ulasan pengguna aplikasi Gojek. Selain data utama hasil *scraping*, penelitian ini juga menggunakan data hasil pelabelan manual oleh *expert linguistic* sebagai *ground truth* pada proses pelatihan model klasifikasi sentimen.

3.1.4 Teknik Analisis Data

Analisis data pada penelitian ini dilakukan melalui beberapa tahapan yang saling berkaitan, meliputi *preprocessing* teks, augmentasi teks, *feature extration* dan *feature engineering*, klasifikasi sentimen, evaluasi model, analisis temporal, serta pengujian sistem berbasis web. Tahap awal dilakukan *preprocessing* teks dengan tujuan membersihkan dan menyiapkan data ulasan agar sesuai untuk digunakan dalam proses model *machine learning*. Setelah itu dilakukan augmentasi teks pada data *training* dan mengurangi ketidakseimbangan kelas.

Tahap *feature extraction* dan *feature engineering* dilakukan untuk memperoleh representasi data yang lebih informatif. Fitur yang digunakan menggunakan kombinasi fitur semantik IndoBERT, fitur leksikal *TF-IDF*, fitur linguistik, fitur sentimen berbasis leksikon, dan fitur layanan hasil *Named Entity Recognition*



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

(NER). Seluruh fitur kemudian digabungkan menggunakan teknik *feature concatenation* sehingga menghasilkan satu vektor fitur yang dapat digunakan sebagai masukan bagi model klasifikasi. Proses klasifikasi sentimen dilakukan menggunakan pendekatan *hybrid* IndoBERT-XGBoost yang dioptimasi menggunakan GridSearchCV dan *Stratified K-Fold Cross Validation*. Pada pendekatan ini, IndoBERT berperan dalam menghasilkan representasi teks yang mampu menangkap makna dan konteks kalimat, sedangkan XGBoost digunakan sebagai algoritma klasifikasi multikelas untuk menentukan kategori sentimen positif, netral, dan negatif.

Kinerja model dievaluasi menggunakan *accuracy, precision, recall, F1-score*, dan *confusion matrix*. Hasil klasifikasi kemudian dianalisis secara temporal untuk mengamati perubahan kecenderungan sentimen pengguna dari waktu ke waktu. Tahap akhir dilakukan pengujian sistem berbasis web menggunakan metode *Black Box Testing* untuk memastikan seluruh fungsi sistem berjalan sesuai kebutuhan. Evaluasi terhadap pengalaman pengguna dilakukan dengan *System Usability Scale* (SUS) untuk mengukur Tingkat *usability* dan penerimaan sistem oleh pengguna.

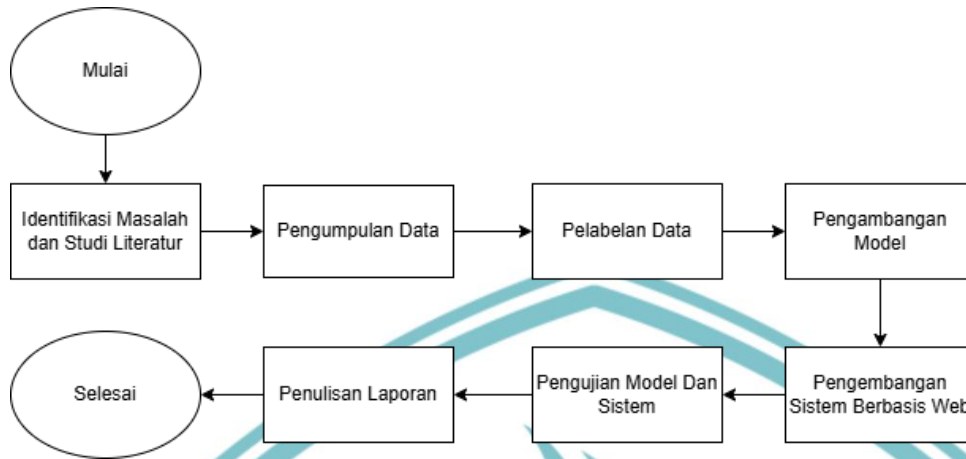
3.2 Tahap Penelitian

Penelitian ini dilakukan melalui serangkaian tahapan sistematis yang saling berkaitan. Secara umum, alur penelitian terdiri dari tujuh tahapan utama yang dimulai dari pengumpulan data hingga visualisasi hasil dalam bentuk dashboard berbasis web. Diagram alur tahapan penelitian dapat dilihat pada Gambar 3.1



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 3.1 Tahapan Penelitian

Berikut adalah uraian masing-masing tahapan penelitian:

3.2.1 Identifikasi Masalah dan Studi Literatur

Langkah pertama dalam pelaksanaan penelitian ini adalah melakukan identifikasi dan analisis permasalahan serta pencarian informasi yang berkaitan dengan analisis sentimen pada ulasan aplikasi Gojek, khususnya terkait perubahan kecenderungan sentimen dari waktu ke waktu. Informasi diperoleh melalui kajian literatur terhadap berbagai artikel ilmiah dan penelitian terdahulu yang memiliki relevansi dengan topik penelitian yang dilakukan.

3.2.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data ulasan yang diberikan oleh pengguna aplikasi Gojek melalui *platform* Google Play Store. Proses pengambilan data dilakukan menggunakan teknik *scraping* dalam rentang waktu tertentu. Adapun kriteria data yang digunakan dalam seleksi data meliputi, data ulasan berbahasa Indonesia, memiliki tanggal ulasan, memiliki isi teks ulasan pengguna, dan berasal dari aplikasi Gojek di Google Play. Dari proses *scraping* tersebut diperoleh sekitar 53.093 data ulasan pengguna. Berikut tampilan hasil pengumpulan data yang didapat:

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

reviewId	userName	userImage	content	score	thumbsUp	reviewCreatedVersion	at	replyContent	repliedAt	appVersion
2	a3570260-Panji	â€œPanji	https://play-lh.g.udah install uli	2	1	4.87.2	30/01/2026 23:33	Hai Kak Panji,	31/01/2026 02:33	4.87.2
3	63a6207c-Ani	Irfan	https://play-lh.g.Alhamdulillah :	5	0	5.46.2	30/01/2026 22:11			5.46.2
4	b87659dc-Soni	Sonjaya	https://play-lh.g.sangat mamba	5	0		30/01/2026 18:54			
5	c6e6152a-ririn	fatmawa	https://play-lh.g.pemesanan m	1	7	5.46.2	30/01/2026 18:17	Hai Kak Ririn,	31/01/2026 02:30	5.46.2
6	3ab427d2-MACOB	BOY Ch	https://play-lh.g.baik	5	0	4.49.2	30/01/2026 17:27			4.49.2
7	09723dc8-Azis	maulana	https://play-lh.g.top bgt	5	0	5.47.1	30/01/2026 17:12			5.47.1
8	ac774bb9-Nawla	Lala	https://play-lh.g.bintang 3 dulu	3	2	5.40.1	30/01/2026 16:46	Mohon maaf :	31/01/2026 00:30	5.40.1
9	7926b35b-Aping	Aja	https://play-lh.g.Makin ga jelas	1	1	5.40.1	30/01/2026 16:24	Hai, Kak @Api	30/01/2026 16:44	5.40.1
10	97b7496c-bintang	Pamu	https://play-lh.g.sangat bagus	5	0	5.44.1	30/01/2026 16:22			5.44.1
11	8f40d045-Metal	Head	https://play-lh.g.CUSTOMER HE	1	1	5.47.1	30/01/2026 16:12	Hai Kak @Met	30/01/2026 16:39	5.47.1
12	b4193645-rizky	adzano	https://play-lh.g.saya rela ikhla:	1	1		30/01/2026 16:07			
13	2c6ad8ee-Andriani	Santri	https://play-lh.g.mudah diguna	5	0	5.47.1	30/01/2026 16:03			5.47.1
14	20b1d4ef-Aditya	Harsya	https://play-lh.g.bagus sekali	5	0		30/01/2026 15:53			
15	4a489f8a-Nur	alfiah Ade	https://play-lh.g.tolong buat pa	1	4	5.47.1	30/01/2026 15:32	Hai Kak Nur,	01/12/2024 23:13	5.47.1

Gambar 3.2 Dataset Hasil *Scraping*

Dari keseluruhan data *scraping*, dilakukan pemilihan sampel menggunakan teknik *Stratified Sampling*. Teknik ini dipilih agar sampel yang diperoleh dapat mempresentasikan distribusi keseluruhan data secara proporsional. *Stratified sampling* dilakukan berdasarkan dua dimensi, yaitu distribusi waktu ulasan dan *rating* aplikasi. Hasil dari stratfied sampling menghasilkan 600 data ulasan. Berikut hasil *stratified sampling* dataset:

1	reviewId	content	rating	at	label_rating	appversion
2	0e5cbb53-7dd	Drivernya me	2	18/11/2025 15:15	negatif	5.39.1
3	83a39d06-16c	mohon maaf :	2	08/05/2025 03:12	negatif	5.18.2
4	f5c0bb4a-d15	gojek ku knp y	2	15/02/2025 09:22	negatif	5.10.1
5	9dd3c0b3-2df	skrg mh maha	3	07/12/2025 08:44	netral	5.42.1
6	ebaa03ab-6b7	lumayan mem	4	27/10/2025 20:51	positif	0
7	1a8edbd1-13d	good, tp serin	4	24/08/2025 12:01	positif	5.28.2
8	8e1ae61f-d66	ceror lama bar	2	13/01/2026 08:19	negatif	0
9	bb11e418-f7d	akhir akhir ini	2	06/04/2025 08:02	negatif	5.14.2
10	a450fe14-9faa	aplikasinya se	3	03/12/2025 05:24	netral	5.41.1
11	8cdefb7a-df35	gotek tetap yg	5	18/07/2025 06:31	positif	5.23.1
12	ce042dae-5b6	Saat hujan tar	1	29/12/2025 06:16	negatif	4.14.2
13	cd7af256-f7d6	baru pake goj	2	20/11/2025 09:49	negatif	5.38.1
14	c0ffdd14-c39a	saldo sy 38rb,	1	02/11/2025 03:37	negatif	5.28.1
15	a502ec3b-74f7	Bagus sih, cm	4	14/02/2025 08:47	positif	5.10.1

Gambar 3.3 Dataset Hasil *Stratified Sampling*

Dari keseluruhan data hasil *scraping*, dilakukan proses *stratified sampling* untuk memperoleh sampel data yang representatif terhadap distribusi data asli. Stratifikasi dilakukan berdasarkan dua aspek, yaitu distribusi waktu ulasan dan distribusi *rating* aplikasi. Pengambilan sampel dilakukan pada rentang 12 bulan ulasan, dengan jumlah 50 data ulasan pada setiap bulan. Pada masing-masing bulan, data dipilih secara seimbang berdasarkan *rating* aplikasi, yaitu terdiri dari 10 ulasan *rating* 1,



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

10 ulasan *rating* 2, 10 ulasan *rating* 3, 10 ulasan *rating* 4, dan 10 ulasan *rating* 5. Dengan demikian, total sampel yang digunakan dalam penelitian berjumlah 600 data ulasan yang selanjutnya digunakan sebagai data pelabelan oleh *expert*. Pendekatan ini dilakukan agar sampel penelitian mampu merepresentasikan berbagai periode waktu dan variasi tingkat penilaian pengguna sehingga proses pelabelan dan pelatihan model dapat dilakukan secara lebih objektif dan representatif.

3.2.3 Pelabelan Data (Expert Labeling)

Data hasil *stratified sampling* kemudian dilabeli secara manual oleh seorang *expert linguistic* menjadi tiga kategori sentimen, yaitu positif, netral, dan negatif. Label ini berfungsi sebagai *ground truth* dalam proses pelatihan model klasifikasi sentimen. *Expert* yang terlibat dalam proses pelabelan merupakan seorang ahli linguistik yang berprofesi sebagai editor bahasa Indonesia di perusahaan media nasional Kompas. Keterlibatan *expert* berlatar belakang linguistik dan jurnalistik ini dipilih untuk memastikan bahwa proses pelabelan dilakukan dengan pemahaman mendalam terhadap struktur dan nuansa bahasa Indonesia, termasuk ragam bahasa informal dan slang yang lazim ditemukan dalam ulasan digital.

Dalam proses pelabelan, *expert* memberikan catatan panduan validasi yang menjadi acuan pemberian label sentimen. Prinsip utama yang ditekankan adalah bahwa penentuan sentimen harus didasarkan pada konteks kalimat secara keseluruhan, bukan sekadar penilaian kata per kata. Berdasarkan catatan tersebut, kriteria pelabelan yang digunakan adalah sebagai berikut:

1. Label Positif diberikan apabila ulasan secara keseluruhan menggunakan kata-kata yang bermakna positif, seperti ungkapan puas, pujian terhadap layanan, atau apresiasi terhadap kinerja driver/aplikasi. Contoh: "sangat membantu saat urgent", "sangat profesional, cepat dan aman", "orang nya ramah dan cepet".
2. Label Negatif diberikan apabila ulasan secara keseluruhan menggunakan kata-kata yang bermakna negatif, mencerminkan kekecewaan, keluhan, atau ketidakpuasan pengguna terhadap layanan. Contoh: "aplikasi jelek, pas

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

order Gojek malah ngeleg”, ”saya kesal, kenapa saat pakai GoTransit tidak bisa di refund saldonya”, ”kecewa sama driver, kalau jarak dekat tidak mau antar”.

3. Label Netral diberikan apabila kalimat ulasan tidak memiliki muatan sentimen yang jelas, misalnya karena kalimatnya tidak lengkap, lepas konteks, atau tidak didukung latar belakang yang memadai. Contoh: ”habis update lokasi error.. tidak bisa..”, ”aplikasinya menjadi berat”, ”isi saldo tidak bertambah, padahal notifikasinya sudah masuk”.

Berikut adalah dataset yang sudah diberikan label manual oleh ahli linguistik bahasa indonesia:

1	reviewId	content	rating	at	label_rating	validation	label_expert	appversion
2	0e5cbb53-	Drivernya mer	2	18/11/2025 15:15	negatif	salah	netral	5.39.1
3	83a39d06-	mohon maaf t	2	08/05/2025 03:12	negatif	salah	netral	5.18.2
4	f5c0bb4a-	gojek ku knp y	2	15/02/2025 09:22	negatif	benar	negatif	5.10.1
5	9dd3c0b3-	skrg mh maha	3	07/12/2025 08:44	netral	salah	negatif	5.42.1
6	ebaa03ab-	lumayan mem	4	27/10/2025 20:51	positif	benar	positif	0
7	1a8edbd1-	good, tp sering	4	24/08/2025 12:01	positif	benar	positif	5.28.2
8	8e1ae61f-	eror lama ban	2	13/01/2026 08:19	negatif	benar	negatif	0
9	bb11e418-	akhir akhir ini	2	06/04/2025 08:02	negatif	benar	negatif	5.14.2
10	a450fe14-	aplikasinya se	3	03/12/2025 05:24	netral	benar	netral	5.41.1
11	8cdefb7a-	gotek tetap yg	5	18/07/2025 06:31	positif	benar	positif	5.23.1
12	ce042dae-	Saat hujan tar	1	29/12/2025 06:16	negatif	benar	negatif	4.14.2
13	cd7af256-	baru pake goje	2	20/11/2025 09:49	negatif	benar	negatif	5.38.1
14	c0ffdd14-	saldo sy 38rb,	1	02/11/2025 03:37	negatif	salah	netral	5.28.1
15	a502ec3b-	Bagus sih, cm	4	14/02/2025 08:47	positif	salah	netral	5.10.1

Gambar 3.4 Dataset Hasil Pelabelan Manual

Berdasarkan hasil pelabelan manual yang dilakukan oleh *expert*, diperoleh distribusi data sentimen yang terdiri dari 259 ulasan berlabel negatif, 184 ulasan berlabel positif, dan 157 ulasan berlabel netral.

3.2.4 Pengembangan Model

Tahap pengembangan model dilakukan untuk membangun sebuah model klasifikasi dengan menggunakan pendekatan *hybrid* IndoBERT-XGBoost. Pada tahap ini dilakukan beberapa proses, yaitu *preprocessing* teks, pembagian data latih dan data uji, augmentasi teks, *feature extraction* dan *feature engineering*, *feature concatenation*, pelatihan model, *hyperparameter tuning*, serta evaluasi model. *Preprocessing* teks dilakukan untuk membersihkan teks ulasan agar dapat diproses



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

model *machine learning*. Selanjutnya dilakukan augmentasi teks pada data *training* untuk meningkatkan variasi data dan mengurangi ketidakseimbangan kelas sentimen.

Tahap *feature extraction* dan *feature engineering* dilakukan dengan mengekstraksi beberapa jenis fitur, yaitu *embedding* semantik IndoBERT, TF-IDF, fitur linguistik, fitur sentimen berbasis leksikon, serta fitur layanan hasil *Named Entity Recognition* (NER). Seluruh fitur kemudian digabungkan menggunakan teknik *feature concatenation* dan digunakan sebagai input model XGBoost untuk melakukan klasifikasi sentimen ke dalam tiga kategori, yaitu positif, netral, dan negatif. Model yang dibangun selanjutnya dioptimasi menggunakan GridSearchCV dan validasi *Stratified K-Fold Cross* untuk memperoleh konfigurasi parameter terbaik sebelum dilakukan evaluasi performa model menggunakan *accuracy*, *precision*, *recall*, *F1-score*, dan *confussion matrix*.

3.2.5 Pengembangan Sistem Berbasis Web

Tahap selanjutnya adalah pengembangan sistem analisis sentimen berbasis web menggunakan arsitektur *client-server* dengan frontend dan backend yang terpisah. Sistem yang dikembangkan memungkinkan pengguna melakukan analisis sentimen secara *real-time* serta memantau perubahan sentimen dari waktu ke waktu melalui dashboard visualisasi temporal. Sistem yang dikembangkan menggunakan reactJS untuk membuat frontend, Flask untuk membuat backend, serta PostgreSQL untuk penyimpanan database. Tahapan pengembangan yang dilakukan meliputi:

1. Pengembangan backend untuk proses klasifikasi sentimen
2. Integrasi model IndoBERT-XGBoost ke backend
3. Pengembangan frontend menggunakan ReactJS berbasis web
4. Integrasi frontend dan backend melalui REST API
5. Implementasi dashboard visualisasi temporal
6. Penyimpanan data historis untuk mendukung analisis temporal secara berkelanjutan

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Sistem memungkinkan pengguna melakukan analisis sentimen secara *real-time* melalui input teks manual, *upload* dataset CSV/XLSX, atau *scraping* langsung Google Play Store.

3.2.6 Pengujian Sistem

3.2.6.1 Pengujian Model

Evaluasi model dilakukan untuk menilai kemampuan klasifikasi sentimen yang dihasilkan model *hybrid* IndoBERT-XGBoost. Penilaian kinerja dilakukan menggunakan lima metrik standar, yaitu *accuracy*, *precision*, *recall* *f1-score* dan *confusion matrix*. Nilai *accuracy* digunakan untuk menunjukkan proporsi prediksi yang benar terhadap seluruh data yang diuji. *Recall* berfungsi untuk mengetahui seberapa baik model mampu mengenali seluruh data yang termasuk dalam suatu kelas. *F1-score* digunakan sebagai indikator keseimbangan antara *precision* dan *recall*. Sedangkan *confusion matrix* digunakan untuk melihat distribusi kesalahan klasifikasi pada setiap kelas. Melalui kombinasi kelima metrik tersebut, performa model dapat dievaluasi secara lebih menyeluruh, termasuk kemampuannya dalam menangani distribusi kelas yang tidak seimbang.

3.2.6.2 Pengujian Sistem Berbasis Web

Pengujian sistem berbasis web dilakukan untuk memastikan seluruh fitur pada aplikasi analisis sentimen temporal berfungsi sesuai dengan kebutuhan yang telah ditentukan. Pengujian mencakup proses input data, komunikasi frontend dan backend, klasifikasi sentimen, penyimpanan database, serta visualisasi dashboard temporal. Evaluasi sistem menggunakan dua pendekatan, yaitu *black box testing* dan *system usability scale* (SUS). Metode *Black box testing* diterapkan untuk memverifikasi kesesuaian keluaran sistem terhadap masukan yang diberikan tanpa meninjau implementasi kode program. Sementara *system usability scale* (SUS) digunakan untuk mengukur tingkat kemudahan pengguna, penerimaan, serta pengalaman pengguna terhadap sistem yang telah dikembangkan.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

3.2.7 Penulisan Laporan

Tahapan terakhir dalam penelitian ini adalah penyusunan laporan skripsi yang memuat dokumentasi seluruh proses penelitian mulai dari identifikasi masalah, pengembangan model, implementasi sistem, pengujian, hingga hasil analisis penelitian ke dalam bentuk laporan tertulis.

3.3 Objek Penelitian

3.3.1 Objek dan Subjek Penelitian

Objek penelitian dalam penelitian ini adalah ulasan pengguna (*user reviews*) aplikasi Gojek yang dipublikasikan di Google Play Store. Ulasan tersebut merupakan teks berbahasa Indonesia yang ditulis oleh pengguna nyata aplikasi Gojek sebagai bentuk penilaian dan ekspresi pengalaman mereka terhadap layanan yang digunakan. Objek penelitian ini mencakup empat dimensi data, yaitu:

1. Konten teks ulasan sebagai data utama analisis sentimen
2. Nilai *rating* yang diberikan pengguna (skala 1-5) sebagai informasi pendukung *stratified sampling*
3. Tanggal ulasan sebagai dimensi temporal dalam analisis perubahan sentimen
4. Entitas layanan Gojek yang disebutkan dalam ulasan sebagai dimensi layanan dalam analisis.

Subjek penelitian adalah pengguna aplikasi Gojek yang secara aktif memberikan ulasan di Google Play Store. Pengguna ini merupakan konsumen dari berbagai layanan yang tersedia dalam ekosistem aplikasi Gojek, di antaranya GoFood, GoRide, GoCar, GoSend, GoPay, GoMart GoShop, GoTransit, GoBluebird, GoBox, dan GoDineIn.

3.3.2 Ruang Lingkup Penelitian

Penelitian ini dibatasi pada ulasan pengguna berbahasa Indonesia yang diunggah pada aplikasi Gojek di Google Play Store dalam rentang waktu 12 bulan. Analisis yang dilakukan bersifat deskriptif-komputasional, yaitu menganalisis dan mendeskripsikan pola sentimen serta perubahannya dari waktu ke waktu berdasarkan data yang ada, tanpa melakukan prediksi atau peramalan sentimen di



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

masa mendatang. Sistem yang dihasilkan hanya mencakup ulasan dari satu *platform* (Google Play Store) dan satu aplikasi (Gojek), sehingga hasil analisis tidak secara langsung dapat digeneralisasi ke platform atau aplikasi lain.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB IV HASIL DAN PEMBAHASAN

4.1 Analisis Kebutuhan

Analisis kebutuhan dilakukan untuk mengidentifikasi kebutuhan yang diperlukan dalam pengembangan sistem analisis temporal pada ulasan aplikasi Gojek. Tahap ini bertujuan memastikan sistem yang dikembangkan mampu melakukan proses pengolahan data, klasifikasi sentimen, analisis temporal, serta visualisasi hasil analisis secara efektif. Analisis kebutuhan mencakup kebutuhan model *machine learning* serta kebutuhan sistem berbasis web yang digunakan untuk melakukan klasifikasi sentimen dan visualisasi temporal.

4.1.1 Analisis Kebutuhan Model

Analisis kebutuhan model dilakukan untuk menentukan kebutuhan data, proses pengolahan teks, *feature extraction*, serta metode klasifikasi yang digunakan dalam penelitian. Model yang dikembangkan menggunakan pendekatan *hybrid* IndoBERT-XGBoost dengan kombinasi beberapa fitur tambahan untuk meningkatkan performa klasifikasi sentimen.

4.1.1.1 Kebutuhan Dataset

Dataset yang digunakan dalam penelitian ini berupa kumpulan ulasan pengguna aplikasi Gojek yang diperoleh dari Google Play Store melalui teknik *web scraping*. Data yang dikumpulkan harus memenuhi beberapa persyaratan, yaitu:

1. Data berupa ulasan berbahasa Indonesia
2. Memiliki isi teks ulasan
3. Memiliki tanggal ulasan
4. Memiliki *rating* pengguna
5. Berasal dari aplikasi Gojek pada Google Play Store

Hasil *scraping* menghasilkan 53.093 data ulasan pengguna

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.1.2 Kebutuhan Sistem Berbasis Web

Kebutuhan sistem web dalam penelitian ini dikelompokkan ke dalam dua kategori, yaitu kebutuhan fungsional dan kebutuhan non-fungsional.

4.1.2.1 Kebutuhan Fungsional

Kebutuhan fungsional menjelaskan berbagai layanan dan fungsi yang harus disediakan oleh sistem. Berdasarkan hasil analisis, sistem yang dirancang pada penelitian ini harus mampu memenuhi beberapa fungsi utama sebagai berikut:

1. Menerima input ulasan melalui tiga mode: input teks manual, unggah dataset (CSV/XLSX), dan *scraping* langsung dari Google Play Store.
2. Mampu melakukan *preprocessing* teks secara otomatis, meliputi *case folding*, penghapusan URL, penghapusan karakter khusus, penghapusan angka, dan penghapusan spasi berlebih.
3. Mengidentifikasi layanan Gojek yang disebutkan dalam ulasan pengguna.
4. Melakukan ekstraksi berbagai fitur yang digunakan dalam klasifikasi sentimen, seperti *embedding* semantik IndoBERT, fitur TF-IDF, fitur linguistik, fitur sentimen berbasis leksikon, dan fitur layanan hasil NER.
5. Melakukan klasifikasi sentimen secara otomatis terhadap ulasan yang dimasukkan dalam tiga kategori, yaitu positif, netral, dan negatif.
6. Menampilkan ringkasan hasil analisis sentimen berupa total ulasan, jumlah per kelas sentimen, dan distribusi persentase.
7. Menampilkan visualisasi tren sentimen secara temporal dalam berbagai granularitas waktu (harian, mingguan, bulanan, kuartalan, tahunan) berdasarkan periode waktu tertentu untuk melihat tren sentimen pengguna.
8. Menyimpan riwayat analisis sehingga pengguna dapat mengakses dan mengunduh hasil analisis sebelumnya.

4.1.2.2 Kebutuhan Non-Fungsional

Kebutuhan non-fungsional berkaitan dengan aspek kualitas, performa, dan karakteristik sistem yang dibangun. Oleh karena itu, sistem yang dikembangkan perlu memenuhi beberapa kebutuhan non-fungsional sebagai berikut:



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

1. Antarmuka berbasis web yang dapat diakses langsung melalui browser tanpa memerlukan proses instalasi.
2. Antarmuka sistem harus mudah digunakan dan dipahami oleh pengguna melalui dashboard yang sederhana namun informatif.
3. Arsitektur *client-server* dengan pemisahan komponen frontend dan backend untuk mempermudah proses pengembangan maupun pemeliharaan.
4. Kemampuan memproses data ulasan dalam jumlah besar untuk mendukung analisis temporal secara berkelanjutan.
5. Dapat diakses melalui berbagai browser modern pada perangkat dekstop maupun laptop.

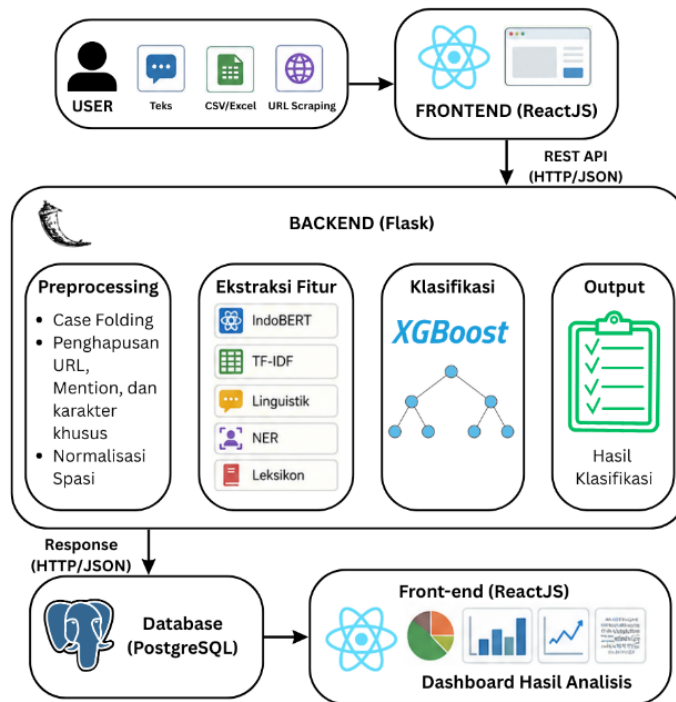
4.2 Perancangan Sistem

4.2.1 Arsitektur Sistem

Sistem pada penelitian ini dirancang menggunakan *client-server* sebagai arsitekturnya dengan memisahkan komponen frontend dan backend. Frontend dibangun menggunakan ReactJS yang berfungsi sebagai antarmuka pengguna dan terhubung dengan backend melalui REST API. Backend dibangun menggunakan Flask (Python) yang menangani logika klasifikasi sentimen dan pengelolaan data. Data hasil analisis disimpan dalam database PostgreSQL. Gambar 4.1 berikut mengilustrasikan sistem secara keseluruhan:

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.1 Arsitektur Sistem

Berdasarkan gambar 4.1 pengguna dapat memasukkan data ulasan melalui beberapa sumber, yaitu input teks, file CSV/Excel, maupun URL *scraping* Google Play Store melalui frontend ReactJS. Data tersebut kemudian dikirim ke backend Flask menggunakan REST API dengan format pertukaran data JSON. Pada backend, sistem melakukan tahapan *preprocessing* dan ekstraksi fitur yang akan diproses oleh model klasifikasi XGBoost untuk menentukan sentimen. Output hasil klasifikasi disimpan ke database PostgreSQL dan ditampilkan kembali pada dashboard frontend dalam bentuk tabel dan visualisasi sentimen.

4.2.2 Diagram UML

4.2.2.1. Use Case

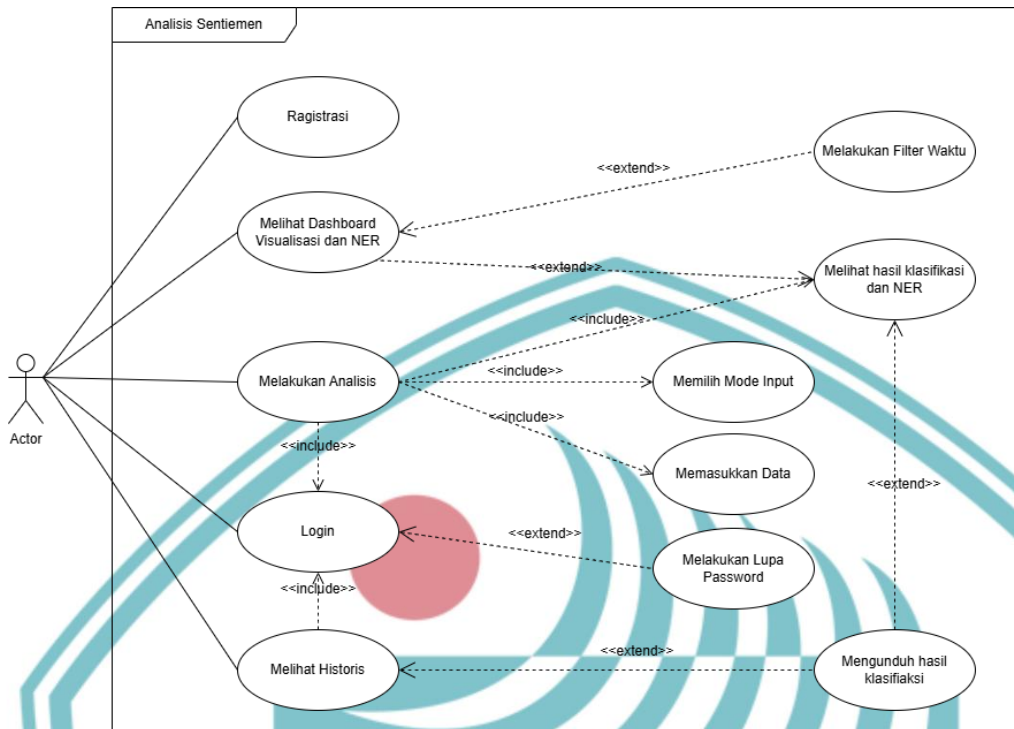
Use case diagram digunakan untuk menggambarkan interaksi antara pengguna dengan sistem analisis sentimen yang dikembangkan. Diagram ini menunjukkan fungsi-fungsi utama yang dapat diakses oleh pengguna dalam sistem.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.2 Use Case Diagram

Pada gambar 4.2, pengguna (*actor*) dapat melakukan registrasi untuk membuat akun dan login untuk mengakses fitur sistem. Sistem juga menyediakan fitur lupa *password* yang digunakan pengguna ketika mengalami kendala saat login atau lupa kata sandi akun. Setelah proses login berhasil, pengguna dapat mengakses fitur utama pada aplikasi. Pada proses analisis, pengguna dapat melakukan analisis sentimen melalui menu analisis yang tersedia pada aplikasi. Pengguna terlebih dahulu memilih input data yang akan digunakan, yaitu input teks manual, *upload* dataset, atau *scraping* Google Play Store. Setelah mode input dipilih, pengguna menampilkan hasil klasifikasi sentimen beserta hasil *Named Entity Recognition* (NER).

Dari halaman hasil klasifikasi, pengguna dapat melakukan beberapa aksi tambahan seperti mengunduh hasil klasifikasi dan melihat dashboard visualisasi sentimen. Dashboard digunakan untuk menampilkan hasil analisis dalam bentuk visualisasi statistik dan grafik sentimen. Selain itu, pengguna juga dapat melakukan filter waktu untuk menampilkan data analisis berdasarkan rentang periode tertentu. Sistem juga menyediakan fitur historis analisis yang menampilkan riwayat hasil

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

klasifikasi yang telah dilakukan sebelumnya. Pada halaman historis, pengguna dapat melihat kembali hasil analisis dan mengunduh hasil klasifikasi yang tersimpan di dalam sistem.

4.2.2.2 Activity Diagram

Activity diagram digunakan untuk merepresentasikan urutan aktivitas atau proses yang terjadi dalam sistem, mulaidari proses awal hingga proses selesai, termasuk interaksi antara pengguna dan sistem selama menjalankan suatu fungsi.

1. Registrasi

Activity diagram registrasi atau pendaftaran akun menggambarkan rangkaian proses pembuatan akun pengguna pada aplikasi analisis sentimen. Diagram ini menunjukkan alur interaksi antara pengguna dan sistem selama proses registrasi akun baru.



Gambar 4.3 Activity Diagram Registrasi

Berdasarkan gambar 4.3, proses registrasi diawali ketika pengguna mengisi formulir registrasi yang terdiri atas *username*, email, dan *password*. Setelah data dikirimkan, sistem melakukan proses penyimpanan akun ke dalam database. Selanjutnya, sistem melakukan pengecekan validitas data registrasi. Apabila ditemukan data yang tidak sesuai atau terjadi kegagalan pada proses validasi, sistem



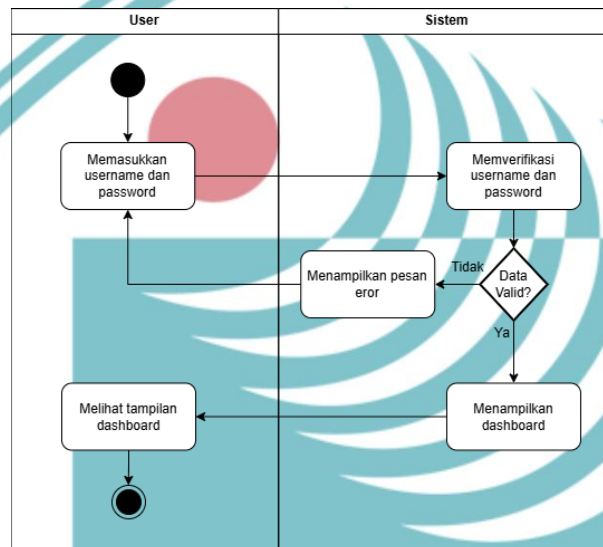
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

akan memberikan pesan eror kepada pengguna. Jika seluruh data memenuhi ketentuan dan proses registrasi berhasil dilakukan, sistem akan menampilkan pesan registrasi telah berhasil. Setelah proses selesai, aktivitas berakhir.

2. Login

Activity diagram login menjelaskan alur proses autentikasi pengguna saat mengakses aplikasi analisis sentimen. Diagram ini memperlihatkan bagaimana pengguna dan sistem berinteraksi selama proses login dilakukan.



Gambar 4.4 Activity Diagram Login

Berdasarkan gambar 4.4, proses dimulai ketika pengguna memasukkan *username* dan *password* pada halaman login. Data yang dikirimkan kemudian diperiksa oleh sistem dengan melakukan validasi akun untuk memeriksa kesesuaian data login dengan data akun yang tersimpan di database. Selanjutnya, sistem melakukan validasi kredensial untuk memastikan kebenaran informasi yang diberikan. Jika *username* atau *password* tidak sesuai, sistem akan menampilkan pesan eror. Namun, apabila data yang dimasukkan valid, pengguna akan diarahkan ke halaman dashboard dan dapat mengakses halaman utama aplikasi. Setelah proses selesai, aktivitas berakhir.

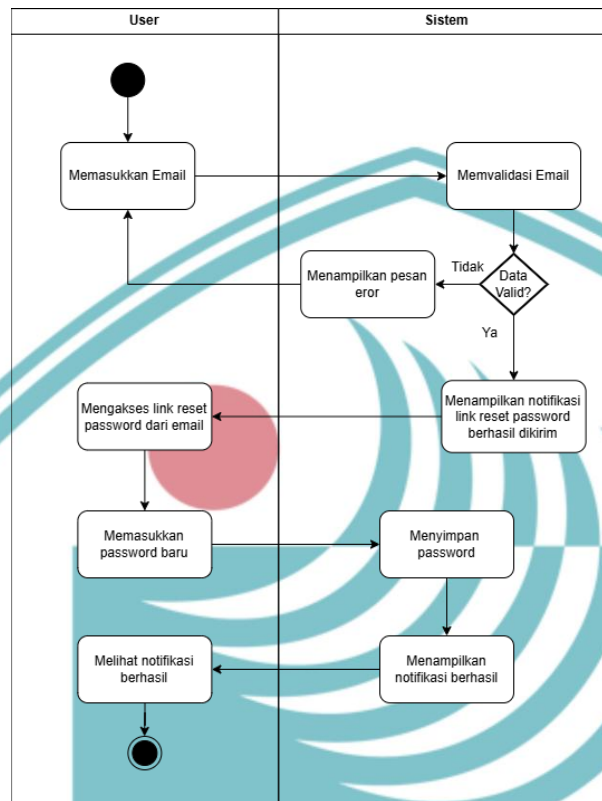
3. Reset Password

Activity diagram reset password menggambarkan alur proses pemulihan akun ketika pengguna ingin mereset password pada sistem analisis sentimen. Diagram

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

ini menunjukkan interaksi antara pengguna dan sistem dalam proses penggantian password akun.



Gambar 4.5 Activity Diagram Reset Password

Berdasarkan gambar 4.5, proses dimulai ketika pengguna memasukkan alamat email pada halaman reset password. Setelah email dikirimkan, sistem melakukan validasi email untuk memeriksa apakah email tersebut terdaftar pada sistem. Selanjutnya, sistem melakukan pengecekan validitas email. Jika email tidak valid atau tidak terdaftar, sistem akan menampilkan pesan eror kepada pengguna. Namun, apabila email valid, sistem akan menampilkan notifikasi bahwa *link* reset password berhasil dikirimkan ke email pengguna. Pengguna kemudian mengakses *link* reset password yang dikirimkan melalui email dan memasukkan *password* baru. Setelah itu, sistem menyimpan *password* baru ke dalam database dan menampilkan notifikasi bahwa proses penggantian *password* berhasil dilakukan. Setelah proses selesai, aktivitas berakhir.

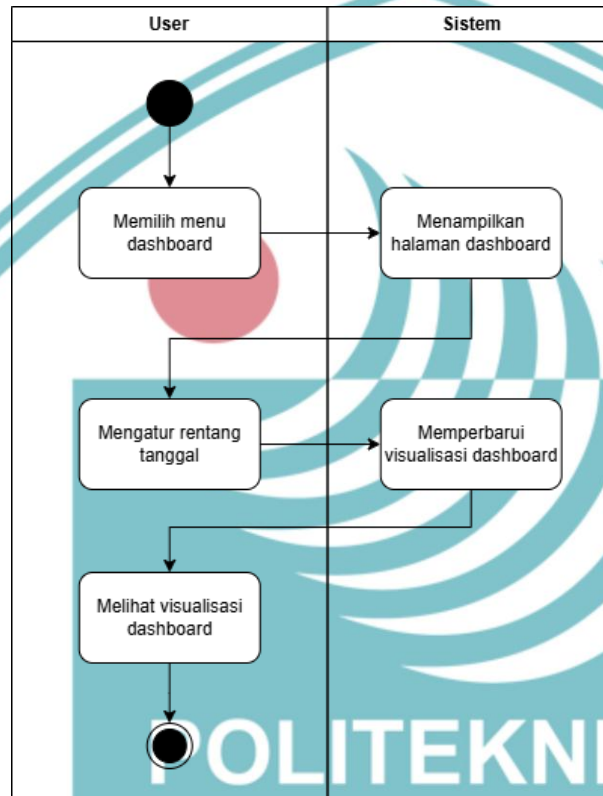


Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4. Melihat Halaman Dashboard

Activity diagram melihat dashboard menggambarkan langkah-langkah yang dilakukan pengguna ketika mengakses halaman utama aplikasi analisis sentimen. Dashboard merupakan halaman pertama yang ditampilkan setelah pengguna berhasil melakukan login.



Gambar 4.6 Activity Diagram Melihat Dashboard

Berdasarkan gambar 4.6, proses diawali saat pengguna memilih menu dashboard atau mengakses halaman dashboard aplikasi. Sistem kemudian menampilkan halaman dashboard yang berisi visualisasi hasil analisis sentimen dan *Named Entity Recognition* (NER). Selanjutnya, pengguna dapat mengatur rentang tanggal untuk memfilter data analisis sesuai periode yang diinginkan. Setelah rentang tanggal dipilih, sistem memperbarui visualisasi dashboard yang telah diperbarui, kemudian proses berakhir.

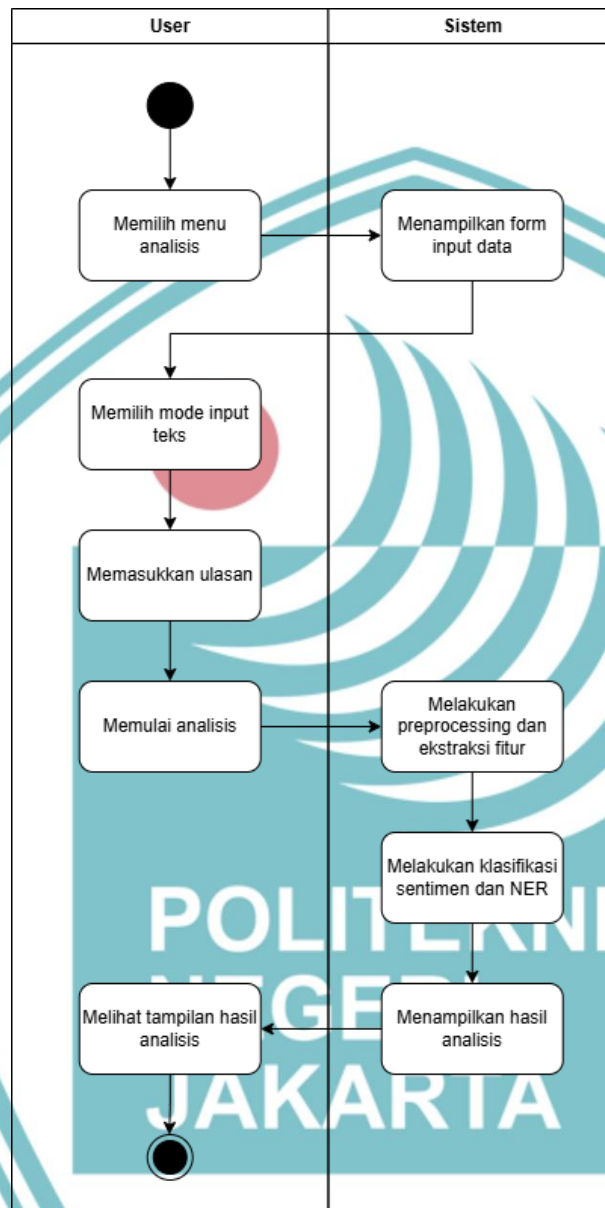
5. Melakukan Klasifikasi Sentimen Teks Input

Activity diagram melakukan sentimen teks input menggambarkan alur proses analisis sentimen terhadap ulasan yang dimasukkan secara langsung oleh pengguna

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

ke dalam sistem. Diagram ini menunjukkan interaksi antara pengguna dan sistem selama proses analisis dilakukan.



Gambar 4.7 Activity Diagram Klasifikasi Sentimen Teks Input

Berdasarkan gambar 4.7, proses dimulai ketika pengguna memilih menu analisis pada aplikasi. Sistem kemudian menampilkan *form* input data yang digunakan untuk memasukkan ulasan. Selanjutnya, pengguna memilih mode input teks dan memasukkan ulasan yang akan dianalisis. Setelah data dimasukkan, pengguna memulai proses analisis sentimen. Pada sisi sistem, data ulasan terlebih dahulu melalui tahap *preprocessing* dan ekstraksi fitur untuk membersihkan teks serta



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

menghasilkan representasi fitur yang digunakan dalam proses klasifikasi. Setelah itu, sistem melakukan klasifikasi sentimen menggunakan model XGBoost dan proses *Named Entity Recognition* (NER) untuk mendeteksi entitas terkait layanan Gojek. Hasil analisis kemudian ditampilkan kepada pengguna, dan proses berakhir.

6. Melakukan Klasifikasi Sentimen Dataset

Activity diagram melakukan klasifikasi sentimen dengan memasukkan dataset menggambarkan alur proses analisis sentimen terhadap data ulasan yang diunggah dalam bentuk dataset. Diagram ini menunjukkan interaksi antara pengguna dan sistem selama proses analisis dataset dilakukan.



Gambar 4.8 Activity Diagram Klasifikasi Sentimen Dataset

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Berdasarkan gambar 4.8, proses dimulai ketika pengguna memilih menu analisis pada aplikasi. Sistem kemudian menampilkan *form* input data yang digunakan untuk proses analisis. Selanjutnya, pengguna memilih mode *upload* dataset dan memasukkan dataset ulasan ke dalam sistem. Setelah dataset berhasil diunggah, pengguna memulai proses analisis sentimen. Pada sisi sistem, dataset yang diterima akan melalui tahap *preprocessing* dan ekstraksi fitur untuk membersihkan data tes dan menghasilkan fitur yang digunakan dalam proses klasifikasi. Setelah itu, sistem melakukan klasifikasi sentimen menggunakan model XGBoost serta proses *Named Entity Recognition* (NER) untuk mendeteksi entitas layanan terkait pada ulasan. Hasil analisis kemudian ditampilkan kepada pengguna. Selain melihat hasil analisis, pengguna juga dapat mengunduh hasil klasifikasi yang telah diproses oleh sistem.

7. Melakukan Klasifikasi Sentimen *Scraping Data*

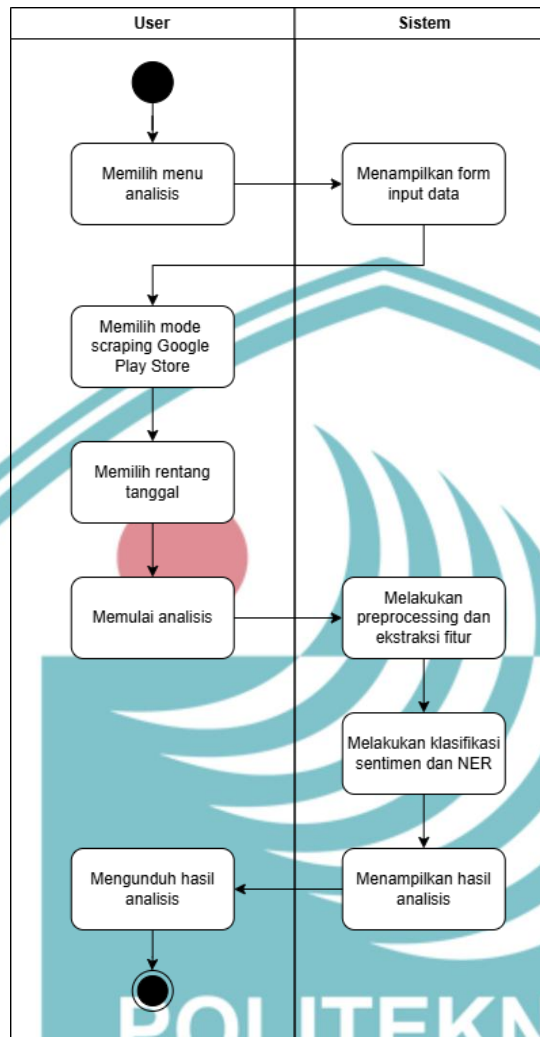
Activity diagram melakukan klasifikasi sentimen scraping data menggambarkan alur proses analisis sentimen terhadap data ulasan yang di *scraping* dari google play store. Diagram ini menunjukkan interaksi antara pengguna dan sistem selama proses analisis dataset dilakukan.

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.9 Activity Diagram Klasifikasi Sentimen Scraping Data

Berdasarkan gambar 4.9, proses dimulai ketika pengguna memilih menu analisis pada aplikasi. Sistem kemudian menampilkan *form* input data yang digunakan untuk proses analisis. Selanjutnya, pengguna memilih mode *scraping* Google Play Store dan menentukan rentang tanggal ulasan yang akan dianalisis. Setelah parameter *scraping* ditentukan, pengguna memulai proses analisis.

Pada sisi sistem, data ulasan hasil scraping terlebih dahulu melalui tahap *preprocessing* dan ekstraksi fitur untuk membersihkan teks dan menghasilkan fitur analisis. Setelah itu, sistem melakukan klasifikasi sentimen menggunakan model XGBoost serta proses *Named Entity Recognition* (NER) untuk mendeteksi entitas layanan yang terdapat pada ulasan. Hasil analisis kemudian ditampilkan kepada

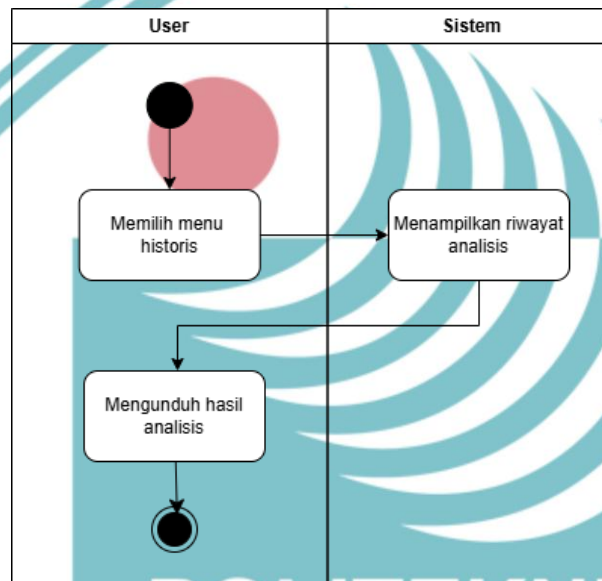
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

pengguna dalam bentuk visualisasi dan data hasil klasifikasi. Selain itu, pengguna juga dapat mengunduh hasil analisis yang telah diproses oleh sistem.

8. Melihat Historis Analisis

Activity diagram melihat historis atau riwayat analisis menggambarkan alur proses pengguna dalam melihat riwayat hasil analisis sentimen yang telah dilakukan sebelumnya pada sistem. Diagram ini menunjukkan interaksi antara pengguna dan sistem saat menampilkan data historis analisis.



Gambar 4.10 Activity Diagram Melihat Historis Analisis

Berdasarkan gambar 4.10, proses dimulai ketika pengguna memilih menu historis pada aplikasi. Setelah itu, sistem menampilkan riwayat analisis yang berisi data hasil klasifikasi sentimen yang telah tersimpan sebelumnya. Selanjutnya, pengguna dapat melihat informasi historis analisis yang tersedia pada sistem. Selain melihat riwayat analisis, pengguna juga dapat mengunduh hasil analisis yang telah tersimpan. Setelah proses selesai, aktivitas berakhir.

4.2.3 Alur Sistem

Alur proses sistem dimulai dari pengambilan data ulasan pengguna, kemudian dilakukan *preprocessing* teks untuk membersihkan data. Setelah itu sistem melakukan ekstraksi fitur menggunakan IndoBERT, TF-IDF, fitur linguistik, fitur sentimen berbasis leksikon, dan fitur layanan hasil NER. Seluruh fitur kemudian



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

digabungkan melalui proses *feature concatenation* dan digunakan sebagai input model XGBoost untuk melakukan klasifikasi sentimen. Hasil klasifikasi selanjutnya disimpan ke database dan divisualisasikan pada dashboard temporal.

Flowchart sistem atau alur kerja aplikasi dapat digambarkan sebagai berikut:



Gambar 4.11 Flowchart Sistem

Alur kerja sistem secara umum adalah sebagai berikut:

1. Pengguna mengakses sistem melalui *browser* dan mengirimkan input (teks, file CSV/Excel, atau URL scraping)
2. Frontend ReactJS mengirimkan *request* ke backend Flask melalui REST API

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

3. Backend melakukan *preprocessing*, ekstraksi fitur (IndoBERT + TF-IDF + linguistik + NER + Leksikon), dan klasifikasi menggunakan model XGBoost
4. Hasil klasifikasi disimpan ke database PostgreSQL dan dikembalikan ke frontend
5. Frontend menampilkan hasil analisis dalam bentuk visualisasi dashboard

4.2.4 Perancangan Database

Database digunakan untuk menyimpan data ulasan, hasil klasifikasi sentimen, informasi layanan, serta histori analisis temporal. Sistem menggunakan PostgreSQL sebagai database utama karena memiliki kemampuan pengelolaan data relasional yang stabil dan mendukung penyimpanan data dalam jumlah besar. Pada penelitian ini, database diirancang menggunakan dua tabel utama, yaitu tabel *users* dan *reviews*. Kedua tabel tersebut saling berhubungan untuk mendukung proses autentikasi pengguna serta penyimpanan hasil analisis sentimen.

1. Tabel Users

Tabel users digunakan untuk menyimpan data akun pengguna yang dapat mengakses sistem analisis sentimen berbasis web. Tabel ini mendukung fitur autentikasi seperti registrasi, login, dan manajemen pengguna.

Tabel 4.1 Struktur Tabel Users

Field	Tipe Data	Keterangan
Id	Integer	Primary key
Username	Varchar(50)	Nama pengguna
Email	Varchar(255)	Email pengguna
Password	Text	Password pengguna
Created_at	Timestamp	Waktu pembuatan akun

Tabel users digunakan untuk mengidentifikasi pengguna yang melakukan proses analisis sentimen pada sistem.



2. Tabel Reviews

Tabel reviews digunakan untuk menyimpan seluruh data hasil *scraping* Google Play Store beserta hasil analisis sentimen yang dihasilkan oleh sistem. Tabel ini menjadi pusat penyimpanan data utama pada sistem karena seluruh proses analisis temporal dilakukan berdasarkan data yang tersimpan pada tabel ini. Beberapa *field* pada tabel berasal langsung dari hasil *scraping* Google Play Store, sedangkan beberapa *field* lain merupakan hasil proses analisis sistem, seperti hasil klasifikasi sentimen dan deteksi layanan Gojek. Struktur tabel *reviews* dapat dilihat pada tabel 4.2 berikut.

Tabel 4.2 Struktur Tabel Reviews

Field	Tipe Data	Keterangan
Id	Integer	Primary key
Content	Text	Isi ulasan pengguna
Sentimen	Varchar	Hasil klasifikasi sentimen
Sumber	Varchar	Sumber data
Tanggal	Timestamp	Tanggal ulasan pengguna
Batch_id	Text	Identitas batch proses analisis
Uploaded_at	Timestamp	Waktu data disimpan ke database
Serveies	Text	Layanan gojek yang terdeteksi
Review_id	Text	Id ulasan dari Google Play Store
User_name	Text	Nama pengguna
Score	Integer	Rating pengguna
Thumbs_up	Integer	Jumlah like pada ulasan
Review_version	Text	Versi aplikasi saat ulasan dibuat
Reply_content	Text	Balasan resmi dari Gojek
Replied_at	Timestamp	Waktu balasan diberikan
App_version	Text	Versi aplikasi
User_id	Integer	Relasi ke tabel users

Field content digunakan sebagai data utama dalam proses analisis sentimen. Data pada *field* ini akan diproses melalui tahapan *preprocessing*, ekstraksi fitur, dan klasifikasi sentimen menggunakan model *hybrid* IndoBERT-XGBoost. *Field*

Hak Cipta :

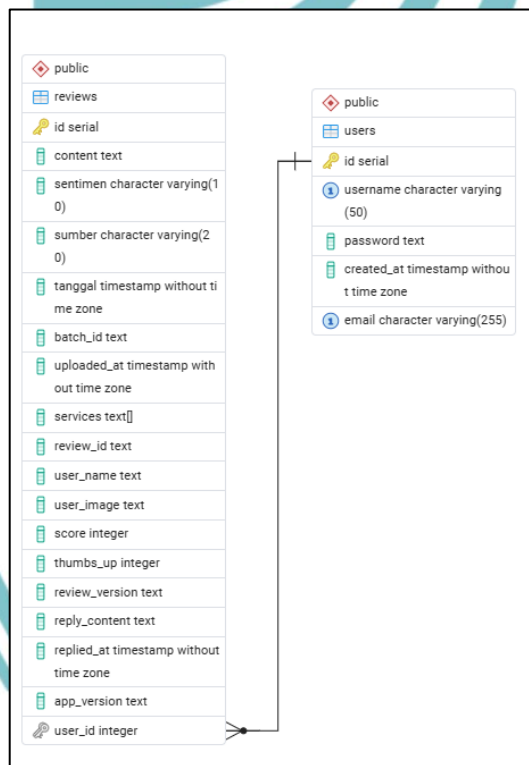
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

sentimen merupakan hasil klasifikasi sentimen yang dihasilkan oleh model *machine learning*, sedangkan *field services* merupakan hasil deteksi layanan Gojek menggunakan pendekatan *Named Entity Recognition* (NER).

Sementara itu, *field* seperti *review_id*, *user_name*, *score*, *thumbs_up*, *review version*, *reply_content*, *replied_at*, dan *app_version* merupakan metadata hasil *scraping* Google Play Store yang disimpan tanpa melalui proses pengolahan lebih lanjut. *Field* *user_id* digunakan sebagai foreign key yang menghubungkan data ulasan dengan pengguna sistem pada tabel *users*. Relasi ini memungkinkan sistem menyimpan histori analisis berdasarkan pengguna yang melakukan proses analisis. Analisis temporal dilakukan berdasarkan *field* tanggal sehingga sistem dapat menampilkan perubahan sentimen dari waktu ke waktu secara dinamis melalui dashboard visualisasi.



Gambar 4.12 Relasi Antar Tabel

Relasi database pada sistem ini menggunakan hubungan one-to-many, di mana satu pengguna pada tabel *users* dapat memiliki banyak data hasil analisis pada tabel *reviews*. Relasi tersebut ditunjukkan melalui *field* *user_id* pada tabel *reviews* yang mereferensikan *field* *id* pada tabel *users*.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.2.5 Perancangan API

Backend sistem dikembangkan menggunakan framework Flask yang menerapkan konsep REST API sebagai penghubung antara frontend berbasis ReactJS dengan model klasifikasi sentimen IndoBERT–XGBoost. API berfungsi sebagai media komunikasi data antara *client* dan *server*, sehingga proses analisis sentimen, pengelolaan data, autentikasi pengguna, hingga visualisasi dashboard dapat dilakukan secara terintegrasi. Arsitektur API yang digunakan menerapkan pendekatan *client-server*, Dimana frontend bertugas sebagai antarmuka pengguna, sedangkan backend bertanggung jawab terhadap proses bisnis sistem, pengolahan data, komunikasi database, serta integrasi model *machine learning*. Berikut tabel 4.3 merupakan endpoint utama yang digunakan pada sistem.

Tabel 4.3 Endpoint Sistem

Endpoint	Fungsi
/predict	Melakukan klasifikasi sentimen input teks Tunggal
/predict_batch	Melakukan klasifikasi sentimen dataset CSV/XLSX
/scrape_predict	<i>Scraping</i> ulasan Google Play Store dan klasifikasi otomatis
/auth/register	Registrasi akun pengguna
/auth/login	Login pengguna
/auth/forgot-password	Mengirim <i>link</i> reset password melalui email
/auth/reset-password	Mengubah <i>password</i> pengguna

a. Endpoint Prediksi Sentimen

Endpoint /predict digunakan untuk melakukan klasifikasi sentimen pada input teks tunggal. Frontend mengirimkan teks ulasan ke backend menggunakan metode POST dalam format JSON. Backend kemudian melakukan *preprocessing* teks, ekstraksi fitur *hybrid*, prediksi menggunakan model XGBoost, dan mengembalikan hasil sentimen beserta *confidence score*.

b. Endpoint Prediksi Batch Dataset

Endpoint /predict_batch digunakan untuk memproses file dataset dalam format CSV maupun XLSX. Sistem akan membaca seluruh data ulasan, melakukan

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

klasifikasi secara batch menggunakan IndoBERT *embedding*, kemudian menyimpan hasil klasifikasi ke database PostgreSQL. Hasil akhir dapat diunduh kembali oleh pengguna dalam bentuk file CSV.

c. Endpoint *Scraping* dan Analisis Otomatis

Endpoint `/scrape_predict` digunakan untuk melakukan *scraping* ulasan aplikasi Gojek secara langsung dari Google Play Store menggunakan *library* google-play-scraper. Data hasil *scraping* selanjutnya diproses menggunakan model klasifikasi sentimen dan disimpan ke database untuk kebutuhan dashboard temporal.

d. Endpoint Autentikasi Pengguna

Sistem menyediakan fitur autentikasi menggunakan JSON Web Token (JWT). Endpoint `/auth/register` digunakan untuk registrasi akun, sedangkan `/auth/login` digunakan untuk proses login dan pembuatan token autentikasi. Endpoint `/auth/forgot-password` dan `/auth/reset-password` digunakan untuk mendukung fitur reset password melalui email.

4.2.6 Perancangan Dashboard Temporal

Dashboard temporal dirancang sebagai media visualisasi untuk menampilkan perubahan sentimen pengguna aplikasi Gojek berdasarkan periode waktu tertentu. Dashboard ini bertujuan membantu pengguna dalam memantau dinamika sentimen secara interaktif serta memahami kecenderungan opini pengguna terhadap layanan Gojek. Dashboard menerima data hasil klasifikasi sentimen yang tersimpan pada database PostgreSQL, kemudian menampilkan hasil analisis dalam bentuk visualisasi grafik dan ringkasan statistik. Visualisasi dirancang agar pengguna dapat melihat distribusi sentimen positif, netral, dan negatif secara lebih mudah.

Fitur utama dashboard meliputi:

1. Ringkasan jumlah sentimen positif, netral, dan negatif
2. Grafik tren sentimen berdasarkan waktu
3. Distribusi layanan Gojek yang paling banyak dibahas
4. *Word cloud* kata yang sering muncul pada ulasan
5. Filter rentang tanggal analisis



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Dashboard temporal menggunakan pendekatan visual analitik sehingga informasi perubahan sentimen dapat dipahami secara cepat melalui grafik interaktif.

4.2.6.1 Halaman Registrasi, Login dan Reset Password

Halaman registrasi, login, dan reset password merupakan halaman autentikasi yang digunakan untuk mengelola akses pengguna ke dalam sistem. Fitur registrasi memungkinkan pengguna membuat akun baru dengan memasukkan informasi yang diperlukan. Fitur login digunakan untuk melakukan autentikasi pengguna sebelum mengakses fitur sistem, sedangkan fitur reset password digunakan untuk membantu pengguna mengatur ulang kata sandi apabila lupa *password*. Halaman ini dirancang dengan antarmuka yang sederhana agar pengguna dapat melakukan proses autentikasi dengan mudah dan cepat. Data akun yang dimasukkan pengguna akan divalidasi oleh sistem sebelum diberikan akses ke halaman utama aplikasi.



Gambar 4.13 Wireframe Halaman Register, Login, dan Reset Password

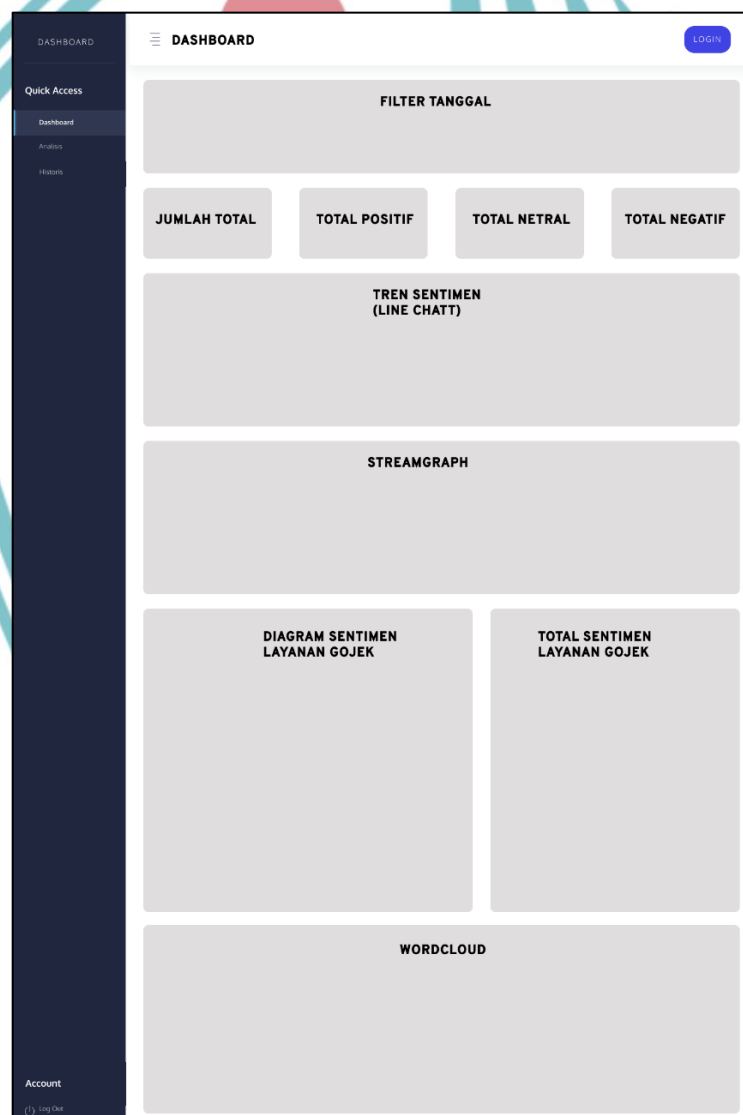
Berdasarkan gambar 4.13, menunjukkan *wireframe* halaman registrasi, login, dan reset password. Pada bagian kiri terdapat area informasi aplikasi yang berisi identitas sistem analisis sentimen Gojek. Pada bagian kanan terdapat area formula autentikasi yang digunakan untuk proses registrasi akun, *login* pengguna, maupun *reset password*. Desain halaman dibuat sederhana untuk memudahkan pengguna dalam mengakses sistem.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.2.6.2 Halaman Dashboard

Halaman dashboard merupakan halaman utama sistem yang berfungsi sebagai media visualisasi hasil analisis sentimen temporal pengguna aplikasi Gojek. Dashboard menyajikan berbagai informasi dalam bentuk grafik dan ringkasan statistik sehingga pengguna dapat memahami perkembangan sentimen secara lebih mudah. Data yang ditampilkan pada dashboard berasal dari hasil klasifikasi sentimen yang tersimpan pada database PostgreSQL. Pengguna juga dapat menggunakan fitur filter tanggal untuk melihat perubahan sentimen pada rentang waktu tertentu.



Gambar 4.14 Wireframe Halaman Dashboard



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Berdasarkan gambar 4.14, menunjukkan *wireframe* halaman dashboard sistem. Halaman ini terdiri dari menu navigasi pada sisi kiri dan area visualisasi pada bagian utama. Dashboard menampilkan filter tanggal, ringkasan jumlah positif, netral, dan negatif. Grafik tren sentimen berdasarkan waktu visualisasi *streamgraph*, diagram distribusi sentimen berdasarkan layanan Gojek, tabel atau grafik jumlah sentimen per layanan, serta *word cloud* yang menampilkan kata-kata yang paling sering muncul dalam ulasan pengguna. Seluruh komponen visualisasi dirancang untuk membantu pengguna memahami pola sentimen secara cepat dan interaktif.

4.2.6.3 Halaman Analisis

Halaman analisis digunakan sebagai media utama untuk melakukan proses klasifikasi sentimen terhadap data ulasan, pengguna dapat memilih berbagai metode input data yang disediakan sistem, yaitu input teks secara langsung, *upload* dataset dalam format CSV/XLSX, maupun *scraping* ulasan aplikasi Gojek dari Google Play Store. Setelah data dimasukkan dan proses analisis dijalankan, sistem akan melakukan *preprocessing*, ekstraksi fitur, klasifikasi sentimen, serta proses *Named Entity Recognition* (NER). Hasil analisis kemudian ditampilkan kepada pengguna dan dapat diunduh dalam bentuk file.



Gambar 4.15 Wireframe Halaman Analisis

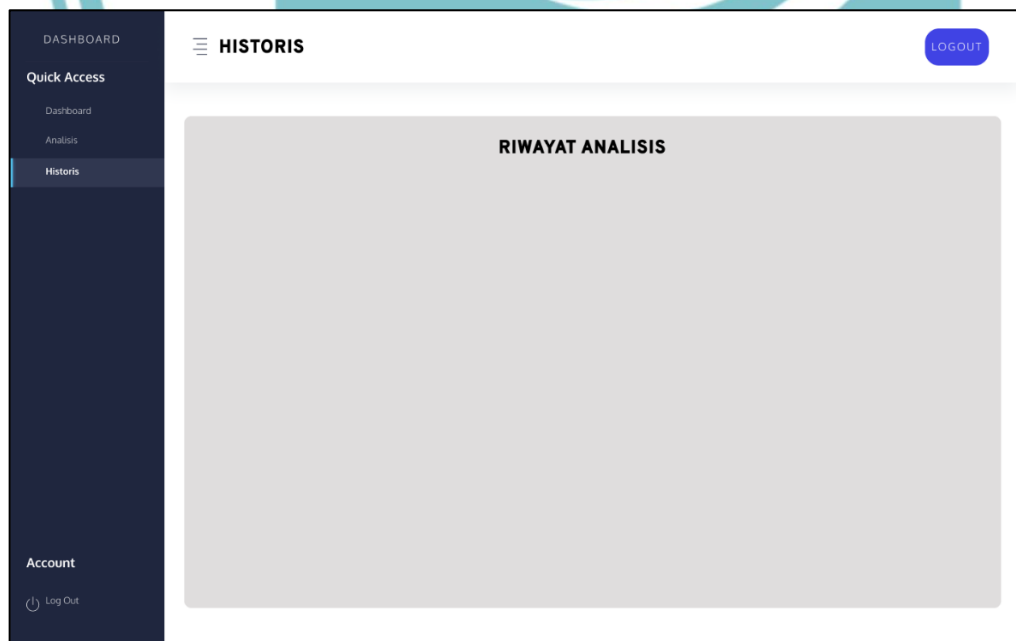
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Berdasarkan gambar 4.15, menunjukkan *wireframe* halaman analisis sistem. Pada bagian atas terdapat area pemilihan metode input data yang digunakan pengguna untuk memasukkan data ulasan. Setelah proses analisis dilakukan, hasil klasifikasi ditampilkan pada bagian hasil klasifikasi yang berisi informasi sentimen serta hasil deteksi entitas layanan Gojek. Halaman ini juga menyediakan fitur untuk mengunduh hasil analisis dan mengakses visualisasi dashboard berdasarkan data yang telah dianalisis.

4.2.6.4 Halaman Histori

Halaman histori digunakan untuk menyimpan dan menampilkan riwayat analisis yang pernah dilakukan pengguna, fitur ini memungkinkan pengguna untuk melihat kembali hasil analisis sebelumnya tanpa harus melakukan proses klasifikasi ulang. Data histori disimpan pada database PostgreSQL dan dapat diakses kapan saja selama pengguna memiliki hak akses ke sistem. Selain menampilkan informasi hasil analisis, halaman ini juga menyediakan fasilitas untuk mengunduh kembali hasil klasifikasi yang telah tersimpan.



Gambar 4.16 Wireframe Halaman Histori

Berdasarkan gambar 4.16, menunjukkan *wireframe* halaman histori analisis. Halaman ini menampilkan daftar riwayat hasil analisis yang telah dilakukan



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

pengguna sebelumnya. Informasi yang ditampilkan meliputi hasil analisis yang telah dilakukan pengguna sebelumnya. Informasi yang ditampilkan meliputi data hasil klasifikasi sentimen, waktu analisis, sumber data, serta fasilitas untuk mengunduh kembali hasil analisis. Dengan adanya halaman historis, pengguna dapat melakukan penelusuran dan evaluasi terhadap hasil analisis yang telah dilakukan sebelumnya.

4.3 Implementasi Sistem

Tahap implementasi sistem merupakan proses penerapan hasil perancangan ke dalam bentuk sistem yang dapat dijalankan. Implementasi pada penelitian ini mencakup pengembangan backend, frontend, integrasi model *hybrid* IndoBERT-XGBoost, implementasi analisis temporal, serta penyimpanan data menggunakan database PostgreSQL. Sistem dikembangkan menggunakan arsitektur *client-server* dengan pemisahan antara frontend dan backend agar proses pengolahan data dan visualisasi dapat berjalan secara terstruktur dan fleksibel.

4.3.1 Implementasi Model

4.3.1.1 Preprocessing Data

a. Preprocessing Teks

Tahap *preprocessing* diimplementasikan untuk membersihkan teks ulasan sebelum dilakukan ekstraksi fitur dan klasifikasi sentimen. Proses ini bertujuan meningkatkan kualitas data. Proses *preprocessing* yang diimplementasikan sebagai berikut:

```
def clean_text(text):
    text = str(text).lower()
    text = re.sub(r"http\S+", "", text)
    text = re.sub(r"@w+", "", text)
    text = re.sub(r"^[a-zA-Z\s]", " ", text)
    text = re.sub(r"\s+", " ", text)
    return text.strip()
```

Proses *preprocessing* yang diimplementasikan meliputi:

1. *Case Folding*: seluruh teks diubah menjadi huruf kecil menggunakan `str.lower()` untuk menyeragamkan representasi kata

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2. Penghapusan URL: tautan yang diawali http atau https dihapus menggunakan polaregex `http\S+` karena tidak mengandung informasi sentimen
3. Penghapusan mention: karakter mention `@username` dihapus menggunakan pola regex `\w+` karena tidak relevan dengan isi ulasan
4. Penghapusan Karakter Khusus dan Angka: seluruh karakter selain huruf dan spasi dihapus menggunakan pola regex `[^a-zA-Z\s]`
5. Normalisasi spasi: Spasi berlebih atau ganda diubah menjadi satu spasi tunggal menggunakan pola regex `\s+`

Berikut tabel 4.4 yang menyajikan contoh beberapa hasil *preprocessing* pada beberapa ulasan.

Tabel 4.4 Contoh Hasil Preprocessing Teks

No	Teks Asli	Hasil Preprocessing
1	"Aplikasi paling gk amanah, saya top up dari se@bank sudah berhasil, tapi saldo tidak berubah"	"aplikasi paling gk amanah sya top up dari sebank sudah berhasil tapi saldo tidak berubah:
2	"sangat mambantu sekali ðŸ˜ŠðŸ™"	"Sangat membantu sekali"
3	"TOLONG AJARIN DRIVERNYA BACA MAPS!!!"	"tolong ajarin drivernya baca maps"

Berdasarkan tabel 4.4 menunjukkan bahwa tahapan *preprocessing* mampu membersihkan berbagai elemen yang tidak relevan terhadap proses analisis sentimen. Pada ulasan pertama, karakter khusus seperti simbol "@" dihapus sehingga teks menjadi lebih seragam. Pada ulasan kedua, emoji dihilangkan karena tidak termasuk karakter alfabet yang digunakan dalam proses ekstraksi fitur. Pada ulasan ketiga dilakukan konversi terhadap seluruh huruf kapital menjadi huruf kecil serta penghapusan tanda baca. Hasil *preprocessing* menghasilkan teks yang lebih bersih dan terstandarisasi sehingga representasi fitur yang dihasilkan dapat mendukung peningkatan performa model klasifikasi sentimen.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

b. Label Encoding

Label *encoding* dilakukan untuk mengubah label kategorikal seperti positif, netral, dan negatif, menjadi representasi numerik agar dapat diproses oleh algoritma *machine learning*. label *encoding* menggunakan *library* yang disediakan dari *from sklearn.preprocessing* yaitu `LabelEncoder`. Label encoder akan melakukan *mapping* label dengan nilai 0 untuk label negatif, nilai 1 untuk label netral, dan nilai 2 untuk label positif dengan implementasi sebagai berikut:

```
from sklearn.preprocessing import LabelEncoder

label_encoder = LabelEncoder()
df["label_encoded"] = label_encoder.fit_transform(
    df["label_expert"].str.lower()
)
```

4.3.1.2 Pembagian Data *Training* dan *Testing* (*split data*)

Setelah proses pelabelan selesai, dataset dibagi menjadi dua bagian, yaitu data latih dan data uji, dengan perbandingan 80:20 menggunakan metode *stratified split*. Teknik ini digunakan untuk menjaga proporsi distribusi kelas sentimen tetap proporsional pada kedua kelompok data. Data latih digunakan untuk proses pembelajaran model serta augmentasi teks, sedangkan data uji dimanfaatkan untuk mengevaluasi kinerja model setelah proses pelatihan selesai. Pemisahan kedua kelompok data tersebut bertujuan untuk menghindari terjadinya data *leakage*, sehingga hasil evaluasi dapat merepresentasikan performa model secara objektif.

4.3.1.3 Augmentasi Teks

Augmentasi teks dilakukan pada data *training* untuk meningkatkan variasi data *training* sekaligus mengatasi ketidakseimbangan kelas sentimen (*class imbalance*). Proses augmentasi hanya diterapkan pada data *training* untuk mencegah kebocoran data (*data leakage*) pada proses evaluasi model, sehingga data *testing* tetap mencerminkan distribusi data asli dan dapat digunakan sebagai evaluasi yang objektif. Teknik augmentasi diterapkan secara sekuensial menggunakan *library* `nplaug` melalui tiga tahapan yang dijalankan berurutan yaitu:

1. *Synonym Replacement*: mengganti kata dengan sinonimnya menggunakan WordNet untuk memperkenalkan variasi leksikal

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumuk dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2. *Random Word Swap*: menukar posisi dua kata secara acak untuk memperkenalkan variasi urutan kata
3. *Random Word Delete*: menghapus kata secara acak untuk mensimulasikan variasi penulisan informal

Penentuan jumlah target augmentasi per kelas dilakukan secara adaptif berdasarkan distribusi data *training*, sehingga augmentasi tidak berlebihan pada kelas mayoritas namun tetap mencukupi untuk kelas minoritas dengan proses implementasi sebagai berikut:

```
aug = naf.Sequential([
    naw.SynonymAug(aug_src="wordnet"),
    naw.RandomWordAug(action="swap"),
    naw.RandomWordAug(action="delete")
])
```

Penentuan jumlah target augmentasi dilakukan secara adaptif berdasarkan distribusi jumlah data pada setiap kelas sentimen. Sebelum augmentasi, distribusi data *training* terdiri dari 205 data negatif, 147 data positif, dan 128 data netral. Setelah proses augmentasi dilakukan, seluruh kelas memiliki jumlah data yang seimbang, yaitu masing-masing sebanyak 205 data mengikuti jumlah kelas mayoritas.

4.3.1.4 Feature Extraction dan Feature Engineering

Tahap *feature extraction* dan *feature engineering* diimplementasikan untuk menghasilkan representasi data yang lebih komprehensif sebelum dilakukan klasifikasi sentimen. Penelitian ini menggunakan kombinasi lima kelompok fitur utama, yaitu fitur semantik IndoBERT, TF-IDF, fitur linguistik, fitur layanan hasil *Named Entity Recognition* (NER), serta fitur sentimen berbasis leksikon.

1. TF-IDF

Fitur TF-IDF digunakan untuk menangkap pola kata dan frasa yang sering muncul pada ulasan pengguna. Implementasi TF-IDF dilakukan menggunakan *TfidfVectorizer* dengan konfigurasi *maximum features*: 2000 dan *N-gram range*: unigram dan bigram. Dengan proses implementasi sebagai berikut:

```
tfidf = TfidfVectorizer(
    max_features=2000,
    ngram_range=(1,2)
)
```



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

```
X_tfidf_train = tfidf.fit_transform(train_df["clean_text"]).toarray()
X_tfidf_test = tfidf.transform(test_df["clean_text"]).toarray()
```

5. Fitur semantik IndoBERT

Fitur semantik dihasilkan menggunakan model IndoBERT (indobenchmark/indobert-base-p1) dari *library* Transformers, yaitu model *transformer* berbasis BERT yang telah dilatih pada korpus bahasa Indonesia. Pada penelitian ini, IndoBERT digunakan sebagai *pretrained language model* untuk menghasilkan representasi semantik teks ulasan pengguna dengan konfigurasi IndoBERT sebagai berikut:

```
inputs = tokenizer(
    batch,
    return_tensors="pt",
    truncation=True,
    padding=True,
    max_length=128
)
with torch.no_grad():
    outputs = model(**inputs)
embeddings = outputs.last_hidden_state[:,0,:].cpu().numpy()
```

Proses *embedding* dilakukan dengan melakukan tokenisasi teks menggunakan AutoTokenizer, kemudian teks diproses menggunakan AutoModel untuk memperoleh representasi vektor dari token [CLS] pada *last hidden state*. Representasi tersebut digunakan sebagai fitur semantik utama pada proses klasifikasi sentimen.

6. Fitur linguistik

Fitur linguistik digunakan untuk menangkap karakteristik struktur teks ulasan pengguna. Fitur linguistik yang digunakan meliputi: jumlah karakter, jumlah kata, jumlah tanda seru. Implementasi fitur linguistik dilakukan menggunakan fungsi sederhana pada Python sebagai berikut:

```
df["text_length"] = df["clean_text"].apply(len)
df["word_count"] = df["clean_text"].apply(lambda x:
    len(x.split()))
```

Menunjukkan implementasi ekstraksi fitur linguistik. Fitur yang dihasilkan meliputi jumlah karakter (text_length) dan jumlah kata (word_count) Fitur-fitur tersebut digunakan untuk menangkap karakteristik struktur teks ulasan pengguna yang dapat membantu proses klasifikasi sentimen.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

7. Named Entity Recognition (NER)

Implementasi *Named Entity Recognition* (NER) dilakukan menggunakan pendekatan *rule-based dictionary matching*. Pendekatan ini digunakan untuk mendeteksi layanan Gojek yang disebutkan dalam ulasan pengguna. Setiap layanan didefinisikan dengan beberapa kata kunci yang mencakup variasi penulisan yang umum ditemukan dalam ulasan pengguna, seperti penulisan tanpa spasi ("gofood"), dengan spasi ("go food"), maupun istilah deskriptif yang merujuk pada layanan tersebut ("pesen makan"). Deteksi dilakukan pada kolom `clean_text` hasil *preprocessing* sehingga pencocokan bersifat *case-insensitive*. Tabel 4.7 berikut menyajikan daftar layanan beserta kata kunci yang digunakan:

Tabel 4.5 Hasil Named Entity Recognition (NER)

Layanan	Kata Kunci
GoFood	"gofood", "go food", "pesan makanan", "order makanan", "pesan makan", "mesan makanan", "driver makanan", "antar makanan", "orderan makanan", "kurir makanan"
GoRide	"goride", "go ride", "ojek online", "naik ojek", "pesen ojek", "order ojek", "motor online", "promo goride"
GoCar	"gocar", "go car", "pesan mobil", "order mobil", "driver mobil", "pesen mobil", "taksi online", "prioritas"
GoTransit	"gotransit", "go transit", "tiket kereta", "tiket bus", "transportasi umum", "krl", "commuter"
GoSend	"gosend", "go send", "kirim barang", "kirim paket", "antar barang", "kurir", "order gosend", "driver gosend"
GoBluebird	"gobluebird", "go bluebird", "bluebird", "taxi bluebird", "taksi biru", "blue bird"
GoBox	"gobox", "go box", "angkut barang", "pindahan", "kirim barang besar", "pickup truck", "mobil bak"
GoPay	"gopay", "go pay", "saldo", "dompet digital", "bayar pakai gopay", "top up", "transfer gopay", "gopaynya", "cashback gopay"



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Layanan	Kata Kunci
GoMart	"gomart", "go mart", "belanja kebutuhan", "beli kebutuhan", "belanja harian", "minimart", "convenience store"
GoShop	"goshop", "go shop", "belanja online", "titip beli", "jastip", "beliin", "cariin barang"
GoDineIn	"godinein", "go dine in", "makan di tempat", "reservasi makan", "booking meja", "makan di restoran"

Implementasi deteksi layanan dilakukan dengan melakukan pencocokan kata kunci pada teks hasil *preprocessing*. Berikut implementasi kode program *preprocessing*:

```
def detect_services(text):
    result = {}
    for service, keywords in services.items():
        result[service] = int(any(k in text for k in keywords))
    return pd.Series(result)
```

Proses tersebut mendeteksi layanan menggunakan pencocokan kata kunci. Sistem akan memeriksa apakah salah satu kata kunci suatu layanan ditemukan pada teks ulasan. Jika ditemukan, maka nilai fitur layanan tersebut diberikan nilai 1, sedangkan jika tidak ditemukan diberikan nilai 0. Hasil deteksi disimpan dalam bentuk fitur biner yang kemudian digunakan sebagai masukan pada model klasifikasi sentimen.

8. Fitur sentimen berbasis leksikon

Fitur sentimen berbasis leksikon diimplementasikan menggunakan daftar kata positif dan negatif yang umum digunakan pada ulasan aplikasi digital. Fitur sentimen dihitung berdasarkan kamus leksikon yang terdiri dari 30 kata positif dan 30 kata negatif yang relevan dengan konteks ulasan aplikasi berbahasa Indonesia. Untuk setiap ulasan, dihitung jumlah kemunculan kata positif (*sent_pos*) dan kata negatif (*sent_neg*) menggunakan pencocokan *substring* pada *clean_text*, sehingga menghasilkan vektor berdimensi 2. Kamus ini mencakup variasi ekspresi formal maupun informal yang umum digunakan pengguna Indonesia, termasuk frasa seperti "tidak bisa" dan "tidak responsif" yang merupakan ekspresi negatif multi-kata".

Tabel 4.6 Fitur Sentimen Berbasis Leksikon

Sentimen	Kata Kunci
Positif	"bagus", "cepat", "mantap", "baik", "ramah", "murah", "recommended", "nyaman", "top", "memuaskan", "praktis", "mudah", "suka", "puas", "keren", "helpful", "terbaik", "lancar", "responsif", "aman", "terpercaya", "canggih", "efisien", "hemat", "tepat", "profesional", "cukup", "senang", "terima kasih", "luar biasa", "sempurna"
Negatif	"lama", "buruk", "jelek", "error", "mahal", "kecewa", "parah", "lemot", "cancel", "dibatalkan", "susah", "ribet", "lambat", "telat", "gagal", "payah", "zonk", "mengecewakan", "tidak bisa", "tidak jelas", "hilang", "bohong", "penipuan", "rugi", "percuma", "tidak responsif", "bermasalah", "gangguan", "down"

Seluruh fitur hasil *feature engineering* digabungkan menggunakan Teknik *feature concatenation* untuk membentuk satu vektor fitur tunggal. Fitur yang digabungkan meliputi *embedding* IndoBERT, fitur TF-IDF, fitur linguistik, fitur layanan hasil NER, serta fitur sentimen berbasis leksikon. Penggabungan fitur ini bertujuan agar model dapat memanfaatkan berbagai jenis informasi secara bersamaan, baik informasi semantik, statistik, maupun kontekstual. Implementasi *feature concatenation* dapat dilihat pada kode program berikut:

```
X_train = np.concatenate([
    emb_train,
    X_tfidf_train,
    stat_train,
    service_train,
    sentiment_train
], axis=1)
```

Seluruh fitur digabungkan secara horizontal menggunakan fungsi `np.concatenate()` dengan parameter `axis=1`. Proses ini menghasilkan representasi fitur berdimensi tinggi yang lebih komprehensif karena menggabungkan berbagai karakteristik teks ulasan dalam satu vektor input. Hasil *concatenation* kemudian digunakan sebagai masukan utama pada model XGBoost untuk proses klasifikasi sentimen.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.3.1.5 Implementasi Model XGBoost

Model klasifikasi sentimen pada penelitian ini dirancang menggunakan pendekatan *hybrid* antara IndoBERT dan XGBoost. IndoBERT digunakan untuk menghasilkan representasi semantik teks ulasan pengguna, sedangkan XGBoost digunakan sebagai algoritma klasifikasi sentimen. Pendekatan *hybrid* dipilih karena mampu menggabungkan pemahaman konteks bahasa Indonesia dari IndoBERT dengan kemampuan klasifikasi XGBoost dalam mengolah data fitur numerik secara efektif.

Model dikonfigurasi untuk melakukan klasifikasi multikelas (multiclass classification) terhadap tiga kategori sentimen, yaitu positif, netral, dan negatif. Proses klasifikasi dilakukan setelah seluruh fitur hasil ekstraksi digabungkan melalui proses feature concatenation. Vektor fitur hasil penggabungan kemudian digunakan sebagai input pada model XGBoost untuk menghasilkan prediksi sentimen. Konfigurasi XGBoost dapat diimplementasikan sebagai berikut dengan parameternya:

```
xgb = XGBClassifier(
    objective="multi:softprob",
    num_class=3,
    eval_metric="mlogloss",
    random_state=42,
    tree_method="hist"
)
```

Konfigurasi dasar model XGBoost yang digunakan dalam penelitian. Parameter `objective="multi:softprob"` digunakan untuk klasifikasi multikelas, sedangkan `num_class=3` menunjukkan jumlah kelas sentimen yang diprediksi. Parameter `eval_metric="mlogloss"` digunakan sebagai metrik evaluasi selama pelatihan, `tree_method="hist"` dipilih untuk efisiensi komputasi, dan `random_state=42` digunakan untuk menjaga reproduksibilitas hasil eksperimen.

4.3.1.6 Implementasi *Hyperparameter Tuning*

Tahap *hyperparameter tuning* diimplementasikan untuk memperoleh konfigurasi parameter XGBoost yang menghasilkan performa klasifikasi sentimen terbaik. Proses *tuning* dilakukan menggunakan GridSearchCV yang dikombinasikan dengan *Stratified K-Fold Cross Validation*. Ruang pencarian *hyperparameter* yang diuji didefinisikan sebagai berikut:

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

```
param_grid = {  
    "n_estimators": [150, 200],  
    "max_depth": [3, 4],  
    "learning_rate": [0.03, 0.05],  
    "subsample": [0.8],  
    "colsample_bytree": [0.8],  
    "reg_alpha": [0, 0.3],  
    "reg_lambda": [1, 1.3]  
}
```

Kombinasi *hyperparameter* yang diuji selama proses pencarian parameter terbaik mencakup tujuh parameter dengan total 32 kombinasi, yang masing-masing dievaluasi menggunakan 3-fold cross validation sehingga menghasilkan 96 iterasi pelatihan secara keseluruhan.

Metode GridSearchCV digunakan untuk menguji beberapa kombinasi parameter secara otomatis, sedangkan *Stratified K-Fold* digunakan untuk menjaga distribusi kelas sentimen pada setiap *fold* pelatihan dan validasi tetap seimbang. Berikut implementasi Stratified K-Fold dilakukan dengan konfigurasi tiga *fold* dan proses *shuffle data* menggunakan *random_state* bernilai 42.

```
skf = StratifiedKFold(  
    n_splits=3,  
    shuffle=True,  
    random_state=42  
)
```

Kombinasi *hyperparameter* yang diuji selama proses pencarian parameter terbaik menggunakan GridSearchCV. Untuk evaluasi *tuning* menggunakan *f1_macro* untuk menghasilkan performa yang seimbang pada seluruh kelas sentimen. Implementasi GridSearchCV dapat dilihat pada potongan kode berikut:

```
grid_search = GridSearchCV(  
    estimator=xgb,  
    param_grid=param_grid,  
    cv=skf,  
    scoring="f1_macro",  
    n_jobs=-1,  
    verbose=2,  
    error_score="raise"  
)
```

Kode program ini merupakan proses pencarian kombinasi *hyperparameter* terbaik menggunakan GridSearchCV dengan metrik evaluasi *F1-score Macro*. Penggunaan



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

metrik ini bertujuan untuk menghasilkan performa yang seimbang pada seluruh kelas sentimen. Model terbaik hasil *tuning* kemudian digunakan kembali untuk proses pelatihan akhir sebelum dilakukan evaluasi performa model.

Berdasarkan GridSearch CV, kombinasi *hyperparameter* terbaik yang diperoleh adalah:

```
Best Parameters:
{'colsample_bytree': 0.8, 'learning_rate': 0.05, 'max_depth': 3,
'n_estimators': 200, 'reg_alpha': 0, 'reg_lambda': 1,
'subsample': 0.8}
```

Model terbaik menggunakan *learning_rate* sebesar 0,05 yang relatif kecil, dikombinasikan dengan *n_estimator* sebesar 200 pohon, sehingga model belajar secara bertahap dengan langkah kecil namun jumlah iterasi yang cukup banyak untuk mencapai parameter yang optimal. Parameter *max_depth* sebesar 3 menunjukkan pohon keputusan yang relatif dangkal membantu mencegah *overfitting* pada data *training* yang terbatas. Parameter *reg_alpha* bernilai 0 mengindikasikan regularisasi L1 tidak diperlukan, sedangkan *reg_lambda* sebesar 1 menunjukkan L2 default dipertahankan. Kombinasi *subsample* dan *colsample_bytree* masing-masing 0,8 berarti setiap pohon dilatih menggunakan 80% sampel data dan 80% fitur secara acak untuk meningkatkan generalisasi model.

4.3.1.7 Implementasi Explainable AI (XAI)

Implementasi *Explainable Artificial Intelligence* (XAI) dilakukan untuk meningkatkan interpretabilitas model *hybrid* IndoBERT-XGBoost dalam proses klasifikasi sentimen. Pendekatan XAI digunakan agar mampu menjelaskan faktor-faktor yang memengaruhi keputusan model. Pada penelitian ini, implementasi XAI dilakukan menggunakan dua pendekatan utama, yaitu SHAP (*Shapley Additive exPlanations*) dan *Attention Saliency*. Kombinasi kedua metode tersebut digunakan karena model yang dibangun menggunakan pendekatan *hybrid* antara IndoBERT dan XGBoost. SHAP digunakan untuk menjelaskan kontribusi fitur pada model XGBoost, sedangkan *Attention Saliency* digunakan untuk menganalisis token atau

kata yang memperoleh perhatian terbesar dari model IndoBERT saat memproses teks ulasan.

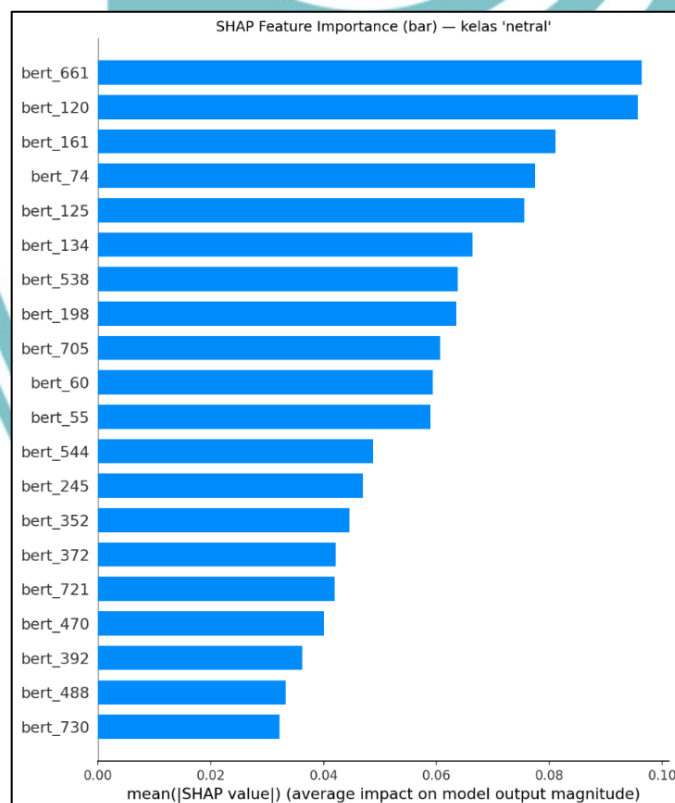
a. Implementasi SHAP

SHAP diimplementasikan untuk menganalisis kontribusi setiap fitur terhadap hasil prediksi model XGBoost. Metode ini menggunakan pendekatan berbasis nilai *Shapley* dari teori permainan untuk menghitung pengaruh masing-masing terhadap output model.

1. SHAP Feature Importance

SHAP *Feature Importance* digunakan untuk mengetahui fitur yang memiliki pengaruh terbesar terhadap hasil prediksi model secara keseluruhan. Nilai *importance* dihitung berdasarkan rata-rata nilai absolut SHAP pada seluruh data pengujian.

SHAP Netral



Gambar 4.17 SHAP Feature Importance Netral

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

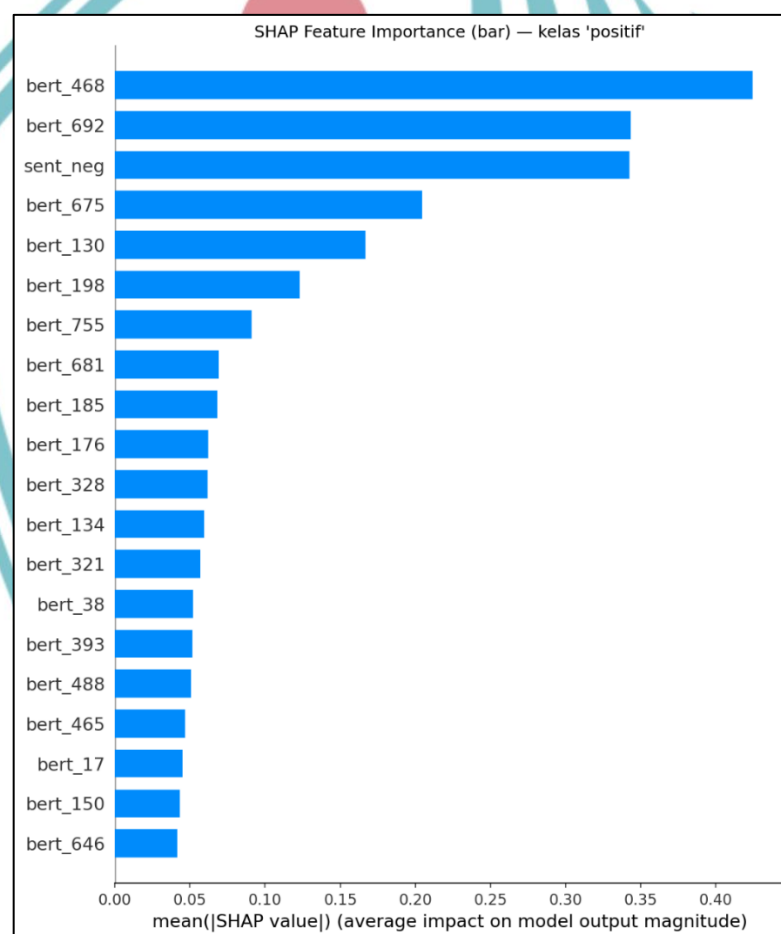


Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Berdasarkan gambar 4.17, fitur-fitur yang memberikan kontribusi terbesar terhadap prediksi sentimen netral didominasi oleh fitur *embedding* hasil representasi IndoBERT, seperti bert_661, bert_120, dan bert_161. Nilai mean absolute SHAP yang relatif tinggi menunjukkan bahwa dimensi *embedding* tersebut memiliki pengaruh yang lebih besar dibandingkan fitur lainnya dalam membedakan ulasan yang termasuk kategori netral. Hal ini menunjukkan bahwa model lebih mengandalkan representasi semantik yang dihasilkan oleh IndoBERT untuk mengenali karakteristik ulasan netral.

SHAP Positif



Gambar 4.18 SHAP Feature Importance Positif

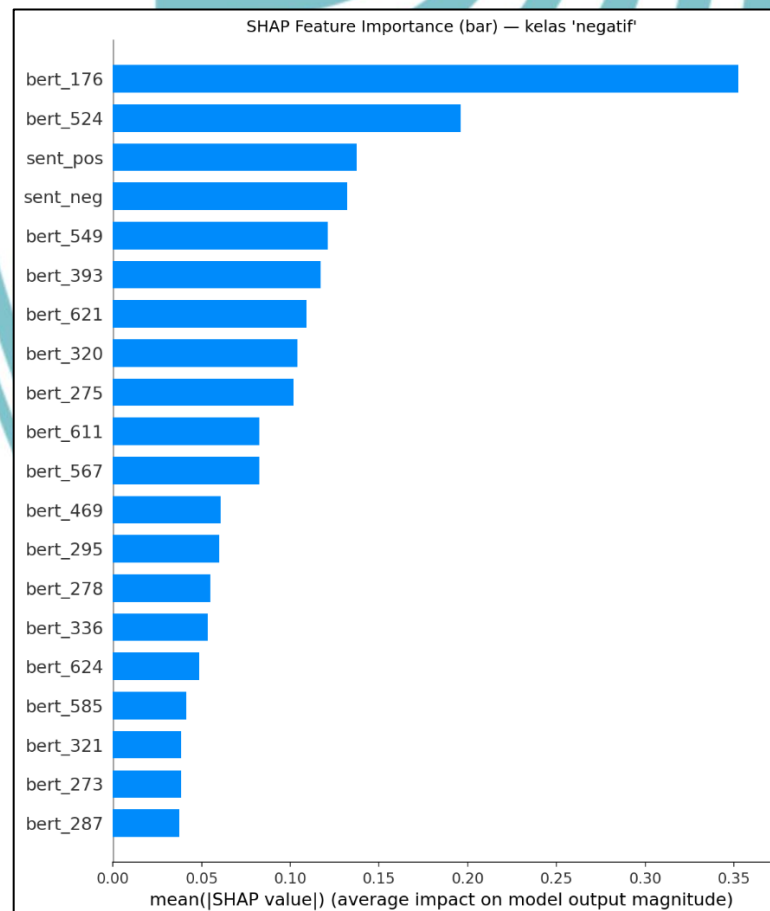
Berdasarkan gambar 4.18, fitur yang memiliki pengaruh terbesar terhadap prediksi sentimen positif adalah bert_468, dan bert_692, yang merupakan bagian dari *embedding* IndoBERT. Selain itu, terdapat fitur hasil rekayasa berupa sent_neg yang juga muncul sebagai salah satu yang juga muncul sebagai salah satu fitur

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

penting dalam proses prediksi. Dominasi fitur *embedding* menunjukkan bahwa model memanfaatkan informasi semantik yang diperoleh dari IndoBERT untuk memahami konteks ulasan positif secara lebih mendalam. Munculnya fitur *sent_neg* sebagai fitur penting menunjukkan bahwa keberadaan atau ketiadaan kata-kata bernuansa negatif turut menjadi pertimbangan model dalam menetikan suatu ulasan sebagai sentimen positif. Dengan kata lain, model tidak hanya mencari indikasi kata-kata positif, tetapi juga memperhatikan rendahnya intensitas kata negatif dalam ulasan tersebut. Hal ini menunjukkan bahwa kombinasi antara representasi semantik IndoBERT dan fitur sentimen berbasis leksikon mampu memberikan kontribusi dalam meningkatkan kemampuan model untuk mengidentifikasi sentimen positif.

SHAP Negatif



Gambar 4.19 SHAP Feature Importance Negatif



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

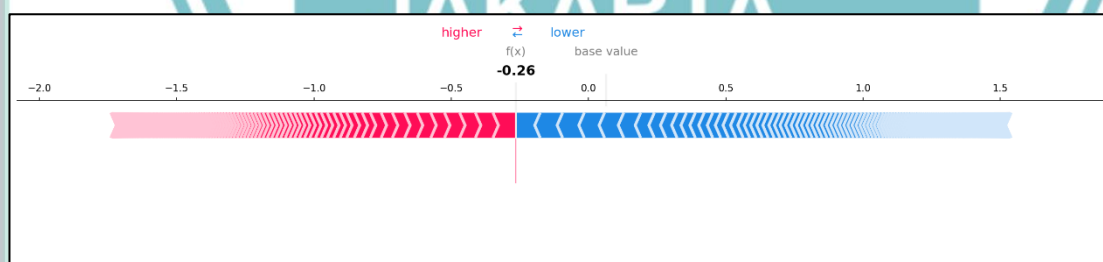
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Berdasarkan gambar 4.19, fitur yang paling berpengaruh dalam prediksi sentimen negatif adalah bert_176 dan bert_524, dengan nilai SHAP yang jauh lebih tinggi dibandingkan fitur lainnya. Selain fitur *embedding* IndoBERT, terlihat bahwa fitur sentimen berbasis leksikon yaitu sent_pos dan sent_neg juga termasuk dalam daftar fitur dengan kontribusi terbesar. Kehadiran fitur sent_neg menunjukkan bahwa jumlah kata-kata negatif yang terdeteksi dalam ulasan memiliki pengaruh signifikan terhadap keputusan model dalam mengklasifikasikan sentimen negatif. Sementara itu, fitur sent_pos juga memberikan kontribusi karena model mempertimbangkan keseimbangan antara kata-kata positif dan negatif yang muncul dalam suatu ulasan.

Hal ini menunjukkan bahwa kelas negatif, model tidak hanya bergantung pada representasi semantik yang dihasilkan oleh IndoBERT, tetapi juga memanfaatkan fitur sentimen berbasis leksikon sebagai indikator tambahan. Dengan demikian, kombinasi fitur hybrid yang digunakan dalam penelitian ini terbukti membantu model dalam mengenali ulasan bernada negatif secara lebih akurat.

2. SHAP Force Plot

SHAP *Force Plot* digunakan untuk menjelaskan kontribusi fitur terhadap prediksi pada satu sampel data tertentu. Pada visualisasi *force plot*, warna merah menunjukkan fitur yang meningkatkan nilai prediksi model, sedangkan warna biru menunjukkan fitur yang menurunkan nilai prediksi. Kombinasi kontribusi antar fitur menghasilkan nilai prediksi akhir pada sampel yang dianalisis.



Gambar 4.20 SHAP Force Plot

Berdasarkan gambar 4.20, beberapa fitur *embedding* IndoBERT memberikan kontribusi dominan terhadap hasil prediksi sentimen. Force plot pada sampel 0 menunjukkan bahwa fitur-fitur yang mendorong prediksi menjauhi kelas netral (merah).



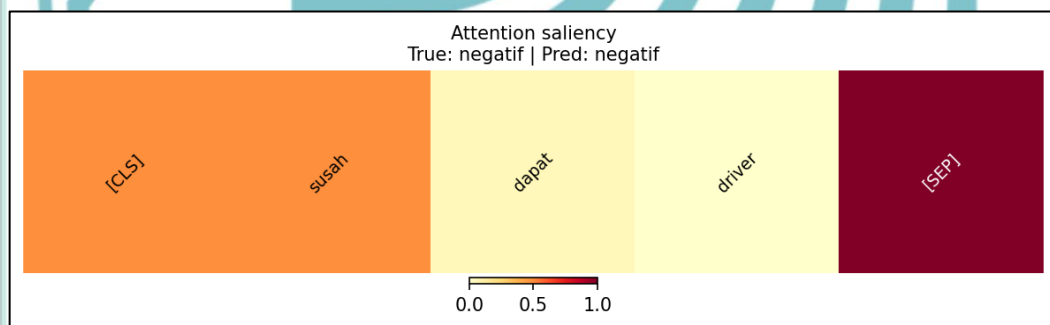
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

b. Implementasi Attention Saliency

Attention saliency diimplementasikan untuk menganalisis token atau kata yang memperoleh perhatian terbesar dari model IndoBERT saat melakukan klasifikasi sentimen. Pendekatan ini memanfaatkan *attention weights* dari *transformer* IndoBERT untuk melihat bagian teks yang dianggap paling penting oleh model dalam memahami konteks kalimat. Pada visualisasi *attention saliency*, warna yang lebih gelap (merah tua) menunjukkan tingkat perhatian model yang lebih tinggi terhadap token tertentu, sedangkan warna yang lebih terang (kuning) menunjukkan perhatian yang lebih rendah. Token yang memperoleh *attention* tinggi dianggap memiliki kontribusi lebih besar dalam proses pembentukan representasi semantik teks.

1. Attention saliency pada prediksi sentimen negatif



Gambar 4.21 Attention Saliency pada Prediksi Sentimen Negatif

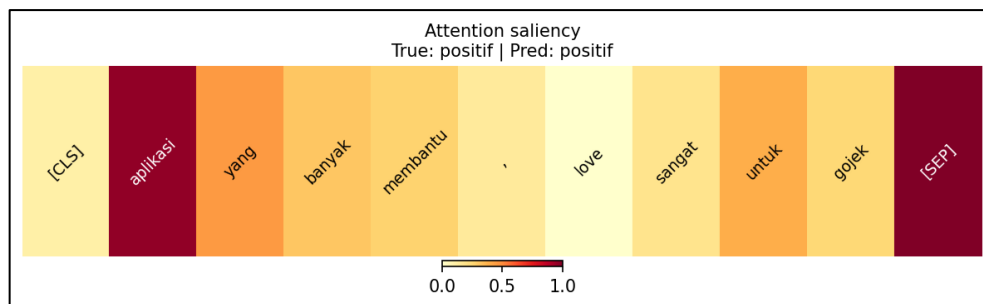
Berdasarkan gambar 4.21, model berhasil mengklasifikasikan ulasan dengan label negatif secara benar. Token [CLS] dan kata "susah" memperoleh perhatian cukup tinggi dari model. Hal ini menunjukkan bahwa model menganggap kata "susah" sebagai indikator sentimen negatif yang kuat pada ulasan pengguna. Selain itu, token [SEP] juga memperoleh perhatian tinggi. Kondisi tersebut merupakan hal yang normal pada arsitektur BERT karena token tersebut digunakan sebagai penanda akhir kalimat pada proses pemrosesan teks. Hasil ini menunjukkan bahwa IndoBERT mampu mengenali kata-kata bernuansa keluhan dalam ulasan pengguna sehingga proses klasifikasi sentimen negatif dapat dilakukan dengan baik.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

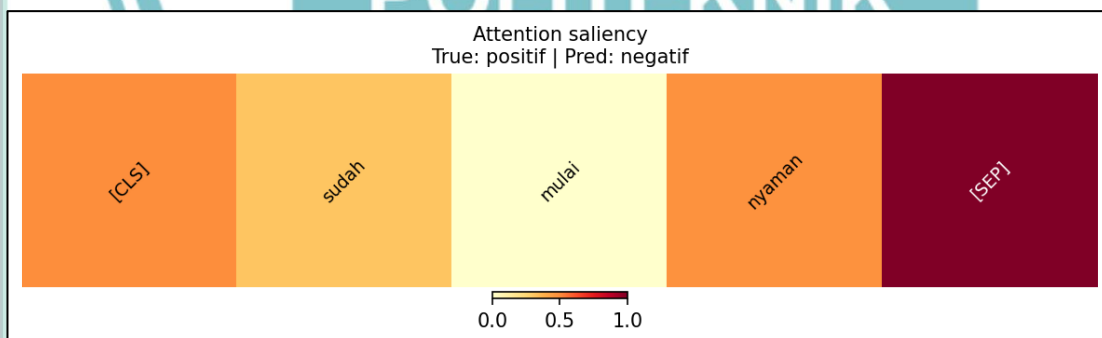
2. Attention saliency pada prediksi sentimen positif



Gambar 4.22 Attention Saliency pada Prediksi Sentimen Positif

Berdasarkan gambar 4.22, model berhasil memprediksi sentimen positif dengan benar. Token seperti "aplikasi", "membantu", dan "sangat" memperoleh perhatian cukup tinggi dari model. Kata "membantu" menunjukkan ekspresi positif terhadap layanan aplikasi, sedangkan kata "sangat" berfungsi memperkuat makna positif dalam kalimat. Hasil ini menunjukkan bahwa model mampu memahami konteks pujian yang diberikan pengguna terhadap aplikasi gojek. Selain itu, perhatian model yang cukup tinggi terhadap token tertentu menunjukkan bahwa IndoBERT dapat menangkap hubungan kontekstual antar kata dalam kalimat berbahasa Indonesia.

3. Attention saliency pada kesalahan klasifikasi sentimen positif menjadi negatif



Gambar 4.23 Attention Saliency Kesalahan Klasifikasi Positif Menjadi Negatif

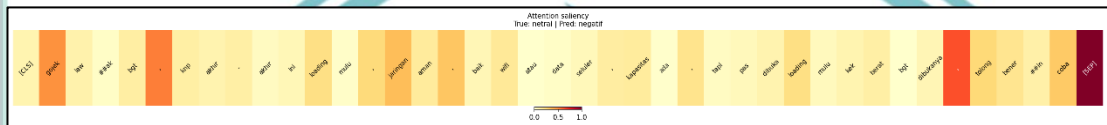
Berdasarkan gambar 4.23, model memprediksi ulasan dengan label sebenarnya positif sebagai sentimen negatif, token seperti "nyaman" memperoleh perhatian cukup tinggi, namun model tetap menghasilkan prediksi negatif. Kesalahan klasifikasi ini menunjukkan bahwa model kemungkinan mengalami kebingungan terhadap konteks kalimat secara keseluruhan. Beberapa kata lain dalam kalimat,

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

seperti ”sudah” atau struktur kalimat yang ambigu, diduga memengaruhi distribusi *attention* sehingga model gagal menangkap makna positif secara utuh. Hasil ini menunjukkan bahwa *attention* tinggi pada suatu token positif belum tentu menghasilkan prediksi yang benar apabila konteks keseluruhan kalimat masih ambigu.

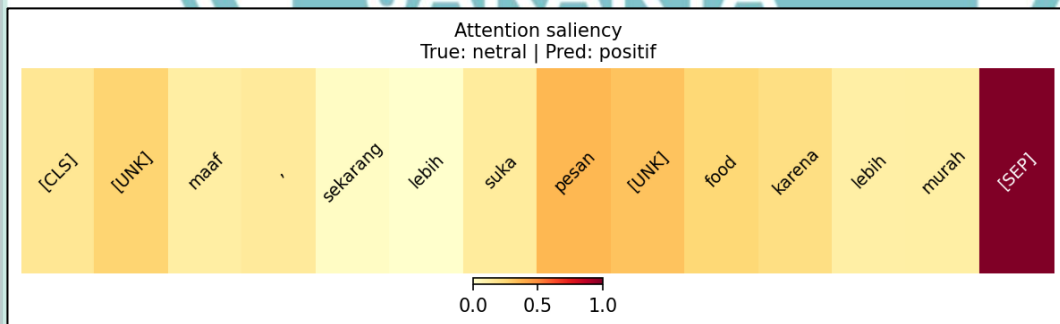
4. Attention saliency pada kesalahan klasifikasi sentimen netral menjadi negatif



Gambar 4.24 Attention Saliency Kesalahan Klasifikasi Netral Menjadi Negatif

Berdasarkan gambar 4.24, model memprediksi ulasan dengan label sebenarnya netral sebagai sentimen negatif. Token seperti ”gojek”, ”bolong”, ”loading”, dan ”jaringan” memperoleh perhatian cukup tinggi dari model. Kata-kata tersebut berkaitan dengan kendala teknis atau performa aplikasi sehingga model menganggap ulasan memiliki kecenderungan sentimen negatif. Padahal secara keseluruhan ulasan tidak secara eksplisit menunjukkan ekspresi negatif yang kuat sehingga termasuk kategori netral. Hasil ini menunjukkan bahwa model masih mengalami kesulitan dalam membedakan sentimen netral dan negatif, khususnya pada ulasan yang mengandung kata-kata bernuansa masalah teknis.

5. Attention saliency pada kesalahan klasifikasi netral menjadi positif



Gambar 4.25 Attention Saliency Kesalahan Klasifikasi Netral Menjadi Positif

Berdasarkan gambar 4.25, model memprediksi ulasan dengan label sebenarnya netral sebagai sentimen positif. Token seperti ”pesan”, ”suka”, dan ”murah”



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

memperoleh perhatian tinggi dari model. Kata-kata tersebut memiliki nuansa positif sehingga model cenderung mengklasifikasikan ulasan ke dalam kategori positif. Namun, secara keseluruhan isi ulasan tidak menunjukkan ekspresi yang dominan sehingga label sebenarnya termasuk kategori netral. Hasil ini menunjukkan bahwa model masih mengalami tumpang tindih antara sentimen netral dan positif pada ulasan yang mengandung beberapa kata positif tetapi tidak merepresentasikan opini positif secara penuh.

4.3.2 Implementasi Sistem Berbasis Web

Implementasi sistem berbasis web dilakukan untuk membangun antarmuka pengguna yang terintegrasi dengan backend Flask dan model klasifikasi sentimen hybrid IndoBERT-XGBoost. Sistem dikembangkan menggunakan ReactJS sebagai frontend sehingga mampu menyediakan tampilan yang interaktif, responsif, dan mendukung konsep *Single Page Application* (SPA).

Pada implementasinya, sistem berbasis web menyediakan beberapa halaman utama yang mendukung proses analisis sentimen secara menyeluruh, mulai dari autentikasi pengguna, proses analisis sentimen, visualisasi temporal, hingga penyimpanan histori analisis, seluruh data hasil analisis dikirim dan diterima melalui REST API yang telah diimplementasikan pada backend Flask.

Berikut merupakan implementasi halaman utama pada sistem berbasis web.

4.3.2.1 Halaman Registrasi

Halaman pendaftaran akun atau register diimplementasikan untuk memungkinkan pengguna membuat akun baru pada sistem analisis sentimen berbasis web. Pengguna diminta mengisi *username*, email, dan *password* melalui *form* registrasi yang tersedia pada frontend ReactJS. Data registrasi kemudiann dikirim ke backend Flask melalui endpoint `/auth/register` menggunakan metode POST. *Password* pengguna dienkripsi menggunakan *bcrypt* sebelum disimpan ke database PostgreSQL untuk meningkatkan keamanan data pengguna. Halaman register juga menerapkan validasi input seperti pengecekan email, *username*, dan *password* sebelum diproses registrasi dilakukan.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Gambar 4.26 Tampilan Halaman Registrasi

Pada gambar 4.26, ditampilkan halaman registrasi yang digunakan pengguna untuk membuat akun baru. Formulir registrasi terdiri dari kolom *username*, email, *password*, dan konfirmasi *password*. Setelah seluruh data valid, pengguna dapat menekan tombol Daftar untuk menyelesaikan proses pendaftaran akun.

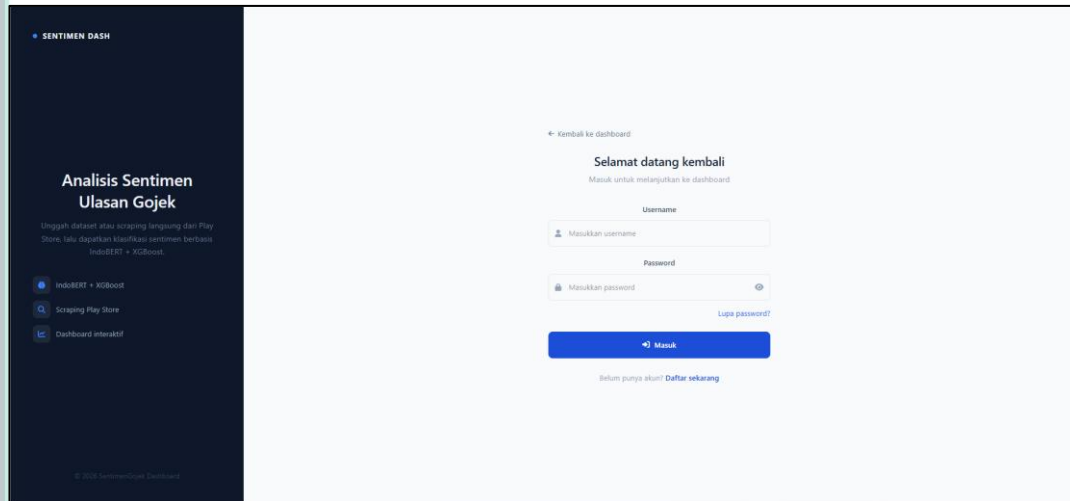
4.3.2.2 Halaman Login

Halaman login diimplementasikan sebagai mekanisme autentikasi pengguna sebelum mengakses fitur tertentu pada sistem, seperti halaman historis analisis. Pengguna melakukan login menggunakan *username* dan *password* yang telah terdaftar sebelumnya. Data login dikirimkan ke backend flask melalui endpoint `/auth/login`. Setelah proses autentikasi berhasil, sistem menghasilkan JSON web token (JWT) yang digunakan sebagai token autentikasi pada setiap proses komunikasi API berikutnya. Implementasi JWT memungkinkan sistem menjaga keamanan akses pengguna serta membatasi akses data histori hanya kepada pengguna yang bersangkutan.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.27 Tampilan Halaman Login

Pada gambar 4.27, ditampilkan halaman login yang digunakan pengguna untuk masuk ke dalam sistem. Halaman ini menyediakan kolom *username* dan *password*, tombol masuk, serta tautan menuju halaman registrasi dan fitur reset password.

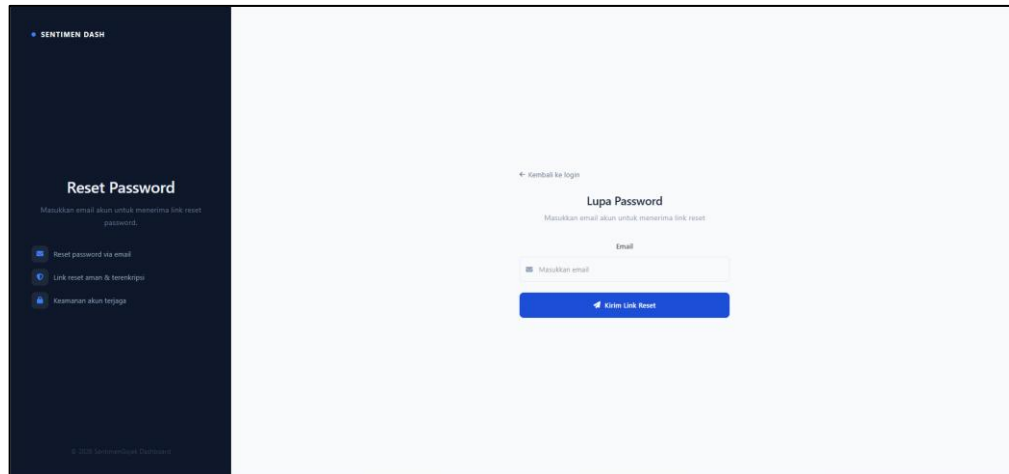
4.3.2.3 Halaman Reset Password

Halaman reset password diimplementasikan untuk membantu pengguna yang lupa *password* akun. Sistem menyediakan fitur reset password yang memungkinkan pengguna meminta tautan reset password melalui email yang terdaftar pada sistem. Proses *reset password* dilakukan melalui *endpoint* `/auth/forgot-password` dan `/auth/reset-password`. Pada backend flask, Sistem menggunakan Flask-Mail untuk mengirim email reset password serta token autentikasi sementara menggunakan `URLSafeTimedSerializer`. Tautan reset password memiliki batas waktu tertentu untuk meningkatkan keamanan sistem. Setelah pengguna memasukkan *password* baru, sistem akan memperbarui *password* pada database PostgreSQL dalam bentuk terenkripsi menggunakan *bcrypt*.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.28 Tampilan Halaman Reset Password

Pada gambar 4.28, ditampilkan halaman reset password yang digunakan pengguna untuk mengirim permintaan perubahan *password*. Pengguna cukup memasukkan alamat email yang telah terdaftar, kemudian sistem akan mengirimkan tautan reset password ke email tersebut.

4.3.2.4 Halaman Dashboard

Halaman dashboard diimplementasikan sebagai pusat visualisasi hasil analisis sentimen temporal pada aplikasi Gojek. Dashboard berfungsi untuk menampilkan informasi dalam bentuk grafik dan statistik sehingga pengguna dapat memantau perubahan sentimen dari waktu ke waktu secara interaktif. Data yang ditampilkan pada dashboard diperoleh dari hasil analisis sentimen yang tersimpan pada database PostgreSQL. Seluruh visualisasi diintegrasikan dengan backend Flask melalui REST API dan ditampilkan pada frontend ReactJS secara dinamis tanpa melakukan *reload* halaman. Visualisasi utama yang diimplementasikan pada dashboard meliputi grafik distribusi sentimen, *streamgraph* temporal, distribusi layanan Gojek yang paling sering dibahas pengguna, serta *word cloud*.

a. Implementasi Dashboard

Grafik distribusi sentimen diimplementasikan untuk menampilkan jumlah ulasan berdasarkan kategori sentimen positif, netral, dan negatif. Visualisasi ini digunakan untuk memberikan gambaran umum mengenai persebaran sentimen pengguna terhadap aplikasi Gojek. Data distribusi sentimen diperoleh dari hasil agregasi data

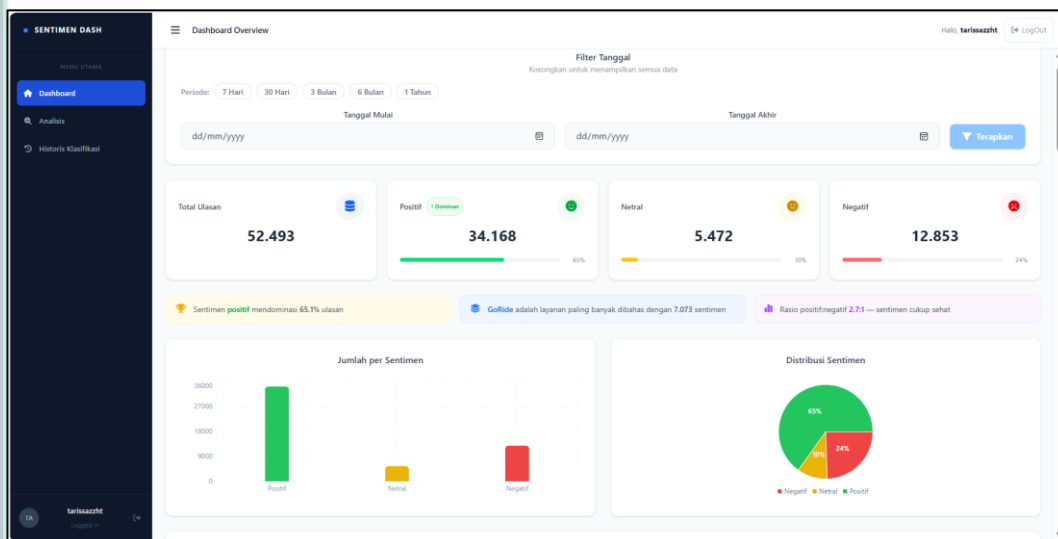


© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

pada database PostgreSQL melalui endpoint /dashboard. Hasil agregasi kemudian divisualisasikan dalam bentuk grafik batang pada frontend ReactJS.



Gambar 4.29 Halaman Dashboard

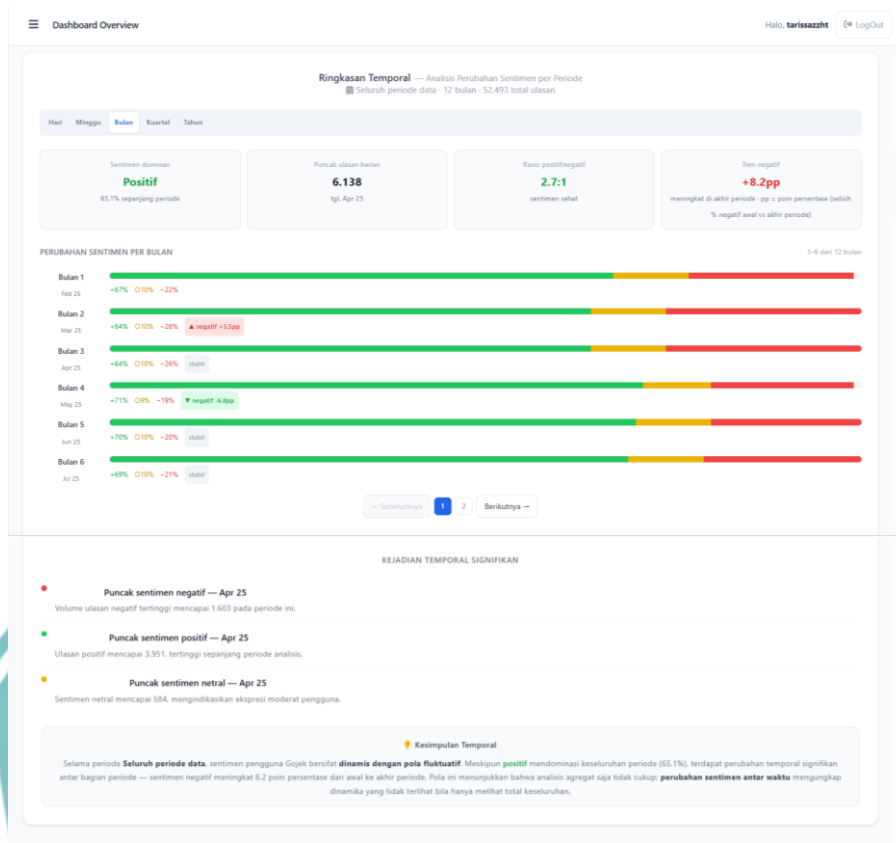
Gambar 4.29 menunjukkan tampilan utama halaman dashboard yang menampilkan ringkasan hasil analisis sentimen ulasan aplikasi Gojek. Dashboard menyajikan informasi berupa jumlah ulasan, jumlah sentimen positif, netral, dan negatif, serta beberapa visualisasi yang mendukung analisis sentimen temporal. Berdasarkan hasil yang ditampilkan, terdapat 52.493 ulasan selama 12 bulan dari tanggal 1 februari 2025 sampai dengan tanggal 31 januari 2026. Jumlah ulasan yang digunakan diambil dari selain dari 600 data ulasan berlabel.

Hasil ulasan yang dianalisis dengan dominasi sentimen positif sebanyak 34.168 ulasan (65%), diikuti sentimen negatif sebanyak 12.853 ulasan (24%) dan sentimen netral sebanyak 5.472 ulasan (10%). Hasil tersebut menunjukkan bahwa sebagian besar pengguna memberikan penilaian positif terhadap aplikasi Gojek selama periode pengamatan.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.30 Ringkasan Temporal

Berdasarkan gambar 4.32, menunjukkan ringkasan temporal yang terjadi selama 12 bulan pengamatan. Selama dalam waktu tersebut sentimen dominan ialah kategori positif sebesar 65,1% mayoritas pengguna memberikan ulasan positif yang artinya secara keseluruhan kepuasan pengguna Gojek cukup baik. Puncak ulasan atau periode yang paling ramai diulas ialah pada bulan April 2025 sebanyak 6.138 ulasan, Dimana kemungkinan ada *event* atau *update* besar di bulan itu. Rasio positif negatif sebesar 2.7:1, Dimana untuk setiap 1 ulasan negatif, ada 2.6 ulasan positif dan ini dikategorikan sebagai “sentimen sehat”. Untuk tren negatifnya, keluhan naik sebesar 8,2% (+8,2 Poin Persentase, yaitu selisih % negatif awal dengan akhir periode), hal ini perlu perhatian, karena dibanding bulan-bulan awal, keluhan di bulan-bulan akhir periode meningkat sebesar 8,2%. Walaupun positif masih dominan, namun indikasi kepuasan pengguna mulai turun menjelang akhir periode.

b. Implementasi Tren Sentimen Temporal

Visualisasi distribusi sentimen diimplementasikan untuk menampilkan persebaran jumlah ulasan berdasarkan kategori sentimen positif, netral, dan negatif. Visualisasi



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

ini bertujuan membantu pengguna memahami dinamika opini pengguna terhadap aplikasi Gojek dari waktu ke waktu. Sistem menyediakan beberapa tingkat agregasi waktu, yaitu harian, mingguan, bulanan, kuartalan, dan tahunan. Namun, pada penelitian ini visualisasi yang digunakan untuk analisis adalah periode bulanan karena mampu memberikan keseimbangan antara detail informasi dan kemudahan interpretasi data.



Gambar 4.31 Grafik Tren Sentimen

Gambar 4.31, menunjukkan visualisasi tren sentimen dalam bentuk *line chart* yang menampilkan perubahan jumlah ulasan positif, netral, dan negatif berdasarkan periode waktu tertentu. Pada visualisasi ini terlihat bahwa sentimen positif secara konsisten memiliki jumlah tertinggi dibandingkan kategori lainnya sepanjang periode pengamatan. Puncak sentimen positif terjadi pada bulan April 2025 dengan jumlah sekitar 3.951 ulasan. Untuk sentimen negatif mengalami peningkatan hingga sekitar 1.603 ulasan pada bulan April 2025, sedangkan sentimen netral mencapai sekitar 584 ulasan di periode waktu yang sama. Kondisi ini menunjukkan adanya peningkatan aktivitas pengguna yang menghasilkan lebih banyak ulasan, baik yang bersifat positif maupun negatif. Setelah bulan April 2025, jumlah ulasan pada ketiga kategori sentimen cenderung mengalami penurunan dan kemudian relatif stabil hingga akhir periode pengamatan. Meskipun terjadi fluktuasi, sentimen positif tetap mendominasi dibandingkan sentimen netral dan negatif.

c. Implementasi Streamgraph Temporal

Streamgraph temporal diimplementasikan untuk memvisualisasikan perubahan sentimen pengguna dari waktu ke waktu secara dinamis. Visualisasi ini

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

menampilkan tren sentimen positif, netral, dan negatif berdasarkan periode tertentu seperti harian, mingguan, bulanan, maupun tahunan dalam bentuk area bertumpuk sehingga perubahan volume sentimen terhadap setiap periode dapat diamati dengan lebih jelas. Data yang digunakan berasal dari hasil agregasi sentimen berdasarkan waktu yang tersimpan pada database PostgreSQL.



Gambar 4.32 Streamgraph Sentimen

Berdasarkan gambar 4.32, menunjukkan visualisasi *streamgraph* sentimen berdasarkan periode bulanan. Luas area pada masing-masing warna menunjukkan volume ulasan pada periode tertentu. Area berwarna hijau mempresentasikan sentimen positif, area kuning menunjukkan sentimen netral, dan area merah menunjukkan sentimen negatif. Berdasarkan visualisasi, area sentimen positif terlihat paling dominan sepanjang periode pengamatan. Volume ulasan mencapai puncak pada bulan April 2025, kemudian mengalami penurunan pada pertengahan tahun sebelum kembali stabil pada bulan-bulan berikutnya. Sentimen negatif dan netral juga mengikuti pola yang relatif serupa, namun dengan volume yang lebih rendah.

d. Impelementasi Distribusi Layanan Gojek

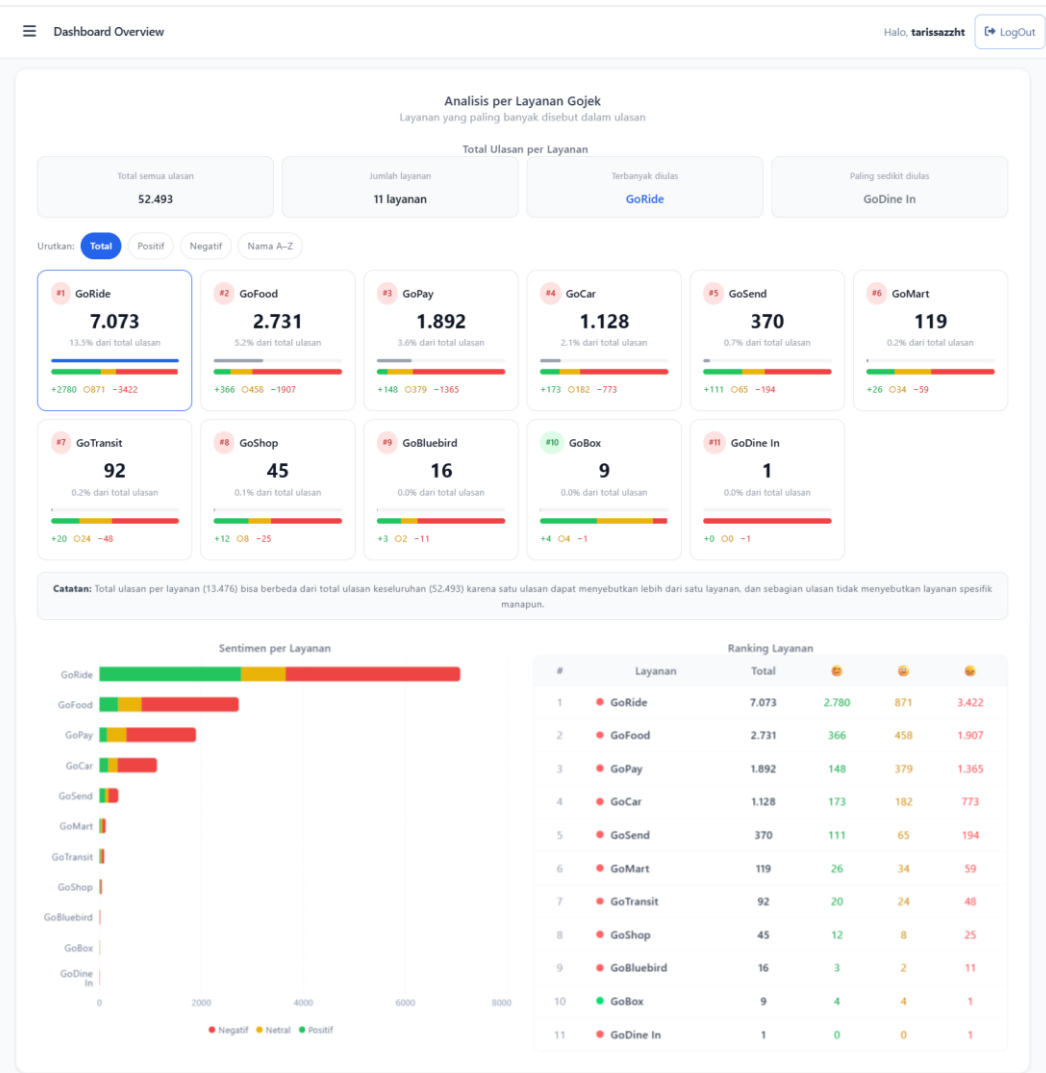
Visualisasi distribusi layanan gojek diimplementasikan untuk menampilkan layanan yang paling sering dibahas oleh pengguna pada ulasan aplikasi. Data layanan diperoleh melalui proses deteksi layanan menggunakan pendekatan *Named Entity Recognition* (NER) berbasis *dictionary matching*. Setiap ulasan dianalisis untuk mendeteksi penyebutan layanan seperti GoFood, GoRide, GoPay, GoCar, dan layanan lainnya. Hasil deteksi kemudian disimpan pada database PostgreSQL



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

dan divisualisasikan dalam bentuk grafik distribusi layanan pada dashboard. Implementasi visualisasi layanan membantu pengguna mengetahui layanan Gojek yang paling banyak memperoleh perhatian maupun keluhan dari pengguna berdasarkan hasil analisis sentimen yang dilakukan sistem.



Gambar 4.33 Distribusi Sentimen Layanan Gojek

Berdasarkan gambar 4.33, menunjukkan hasil visualisasi distribusi layanan Gojek berdasarkan jumlah ulasan yang terdeteksi. Berdasarkan hasil visualisasi, layanan GoRide menjadi layanan yang paling banyak dibahas pengguna dengan total 7.073 ulasan, sedangkan sentimen negatif mencapai 3.422 ulasan dan sentimen netral sebanyak 871 ulasan. Selain GoRide, layanan yang juga sering dibahas adalah GoFood, GoPay, dan GoCar. Temuan ini menunjukkan bahwa layanan transportasi dan pembayaran digital merupakan fitur yang paling sering digunakan sehingga

- 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :**

Wordcloud Ulasan

Kata yang paling sering muncul dalam ulasan

Semua Positif Netral Negatif

Jurusan Teknik Informatika dan Komputer – Politeknik Negeri Jakarta

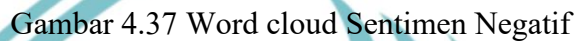
[illegible]

Berdasarkan gambar 4.35, menunjukkan *word cloud* khusus untuk ulasan dengan sentimen positif. Kata-kata seperti “mantap”, “bantu”, “bagus”, “ramah”, dan “puas” muncul dengan ukuran lebih besar, yang menunjukkan bahwa pengguna cenderung memberikan apresiasi terhadap kemudahan layanan, kecepatan pelayanan, dan kualitas aplikasi Gojek.



Berdasarkan gambar 4.36, menunjukkan *word cloud* yang dibentuk dari ulasan dengan kategori sentimen netral. Kata-kata yang muncul umumnya bersifat informatif dan deskriptif, seperti “driver”, “aplikasi”, “bagus”, “pesan”, “dan”, “mahal”. Hal ini menunjukkan bahwa ulasan netral lebih banyak berisi pengalaman penggunaan atau penyampaian informasi tanpa menunjukkan kecenderungan sentimen yang kuat.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



4.3.2.5 Halaman Analisis Sentimen

Halaman analisis sentimen diimplementasikan sebagai fitur utama sistem untuk melakukan kaslfikasi sentimen terhadap ulasan aplikasi Gojek. Sistem menyediakan tiga metode analisis, yaitu input teks manual, *upload* dataset CSV/XLSX, dan *scraping* ulasan langsung dari Google Play Store.

1. Metode Input Teks Manual

Gambar 4.38 Tampilan Halaman Analisis Teks Manual

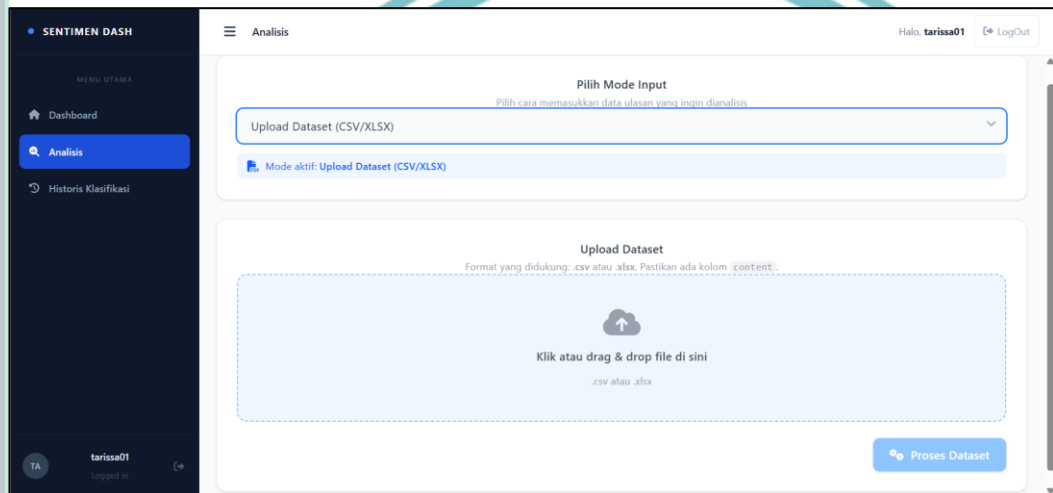


Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Pada gambar 4.38, ditampilkan fitur analisis sentimen menggunakan input teks manual. Pengguna dapat mengetik ulasan secara langsung pada area teks input teks yang tersedia sebelum menjalankan proses klasifikasi sentimen. Pada analisis teks manual, pengguna dapat memasukkan satu ulasan secara langsung dan sistem akan menampilkan hasil klasifikasi beserta layanan Gojek yang terdeteksi.

2. Metode Upload Dataset



Gambar 4.39 Tampilan Halaman Analisis Upload Dataset

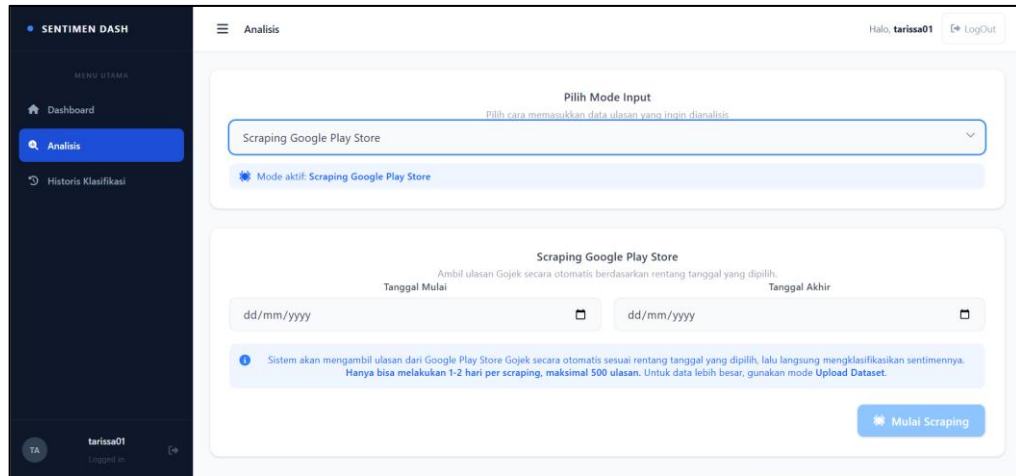
Pada Gambar 4.39, ditampilkan fitur upload dataset yang mendukung format CSV dan XLSX. Pengguna dapat mengunggah file ulasan melalui area *drag and drop* atau memilih file secara langsung, kemudian menjalankan proses analisis dengan menekan tombol Proses Dataset.

3. Metode Scraping

Sistem juga menyediakan fitur scraping otomatis ulasan Google Play Store berdasarkan rentang tanggal tertentu menggunakan library google-plat-scrapers. Halaman ini terintegrasi langsung dengan backend Flask melalui endpoint API seperti /predict, /predict_batch, dan /scrape_predict.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.40 Tampilan Halaman Analisis Scraping

Pada gambar 4.40, ditampilkan fitur scraping ulasan secara otomatis, pengguna dapat menentukan rentang tanggal pengambilan ulasan. Seluruh data yang berhasil diperoleh akan diproses menggunakan model *hybrid* IndoBERT-XGBoost untuk menghasilkan klasifikasi sentimen. Hasil analisis dapat disimpan ke database dan diunduh kembali dalam format CSV untuk kebutuhan analisis lanjutan.

4.3.2.6 Halaman Historis

Halaman historis diimplementasikan untuk menampilkan riwayat hasil analisis sentimen yang telah dilakukan oleh pengguna. Data historis disimpan berdasarkan identitas pengguna melalui relasi antara tabel *users* dan *reviews* pada basis data PostgreSQL. Setiap riwayat analisis menampilkan informasi berupa tanggal dan waktu analisis, rentang data, sumber data, jumlah ulasan yang dianalisis, distribusi sentimen positif, netral, dan negatif, serta sentimen dominan dari setiap batch analisis. Selain itu, sistem menyediakan fitur untuk mengunduh kembali hasil klasifikasi dalam format CSV berdasarkan *batch_id* tertentu sehingga pengguna dapat mengakses kembali hasil analisis yang telah dilakukan.

Halaman historis terintegrasi dengan mekanisme autentikasi JWT (JSON Web Token) sehingga setiap pengguna hanya dapat mengakses data historis miliknya sendiri. Implementasi ini membantu pengguna dalam melakukan pelacakan, evaluasi, dan pengelolaan hasil analisis sentimen secara lebih terstruktur.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

NO	TGL UPLOAD	RENTANG DATA	SUMBER	TOTAL	POSITIF	NETRAL	NEGATIF	DOMINAN	UNDUH
1	28 Mei 2026, 22:16	—	Teks	1	1	0	0	positif	—
2	26 Mei 2026, 11:42	01 Jan 2026	Scraping	164	92	20	52	positif	Unduh
3	25 Mei 2026, 23:56	03 Feb 2025 — 04 Feb 2025	Dataset	100	73	7	20	positif	Unduh
4	25 Mei 2026, 15:29	09 Apr 2025 — 10 Apr 2025	Dataset	100	63	11	26	positif	Unduh
5	25 Mei 2026, 13:30	13 Mar 2025 — 14 Mar 2025	Dataset	100	71	2	27	positif	Unduh

Gambar 4.41 Tampilan Halaman Historis

Pada gambar 4.41, menampilkan halaman histori yang berisi daftar riwayat analisis sentimen pengguna. Pada halaman ini ditampilkan informasi waktu unggah, rentang data, sumber data, jumlah ulasan, distribusi sentimen, sentimen dominan, serta tombol untuk mengunduh kembali hasil klasifikasi. Selain itu, tersedia tombol *refresh* yang digunakan untuk memperbarui data histori yang ditampilkan pada sistem.

4.4 Pengujian Sistem dan Model

4.4.1 Deskripsi Pengujian

Pengujian pada penelitian ini bertujuan untuk menilai dua aspek utama, yaitu performa model klasifikasi sentimen dan fungsionalitas sistem analisis sentimen temporal berbasis web yang dikembangkan.

4.4.1.1 Pengujian Model

Evaluasi model dilakukan untuk mengukur kemampuan model klasifikasi sentimen dan sistem berbasis web yang dikembangkan. Penilaian kinerja menggunakan metrik *accuracy*, *precision*, *recall*, *f1-score*, dan *confusion matrix*. Selain itu, beberapa skenario eksperimen diterapkan untuk menganalisis pengaruh kombinasi Pengujian teknik *feature engineering* serta teknik peningkatan data terhadap performa model.

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.4.1.2 Pengujian Sistem Berbasis Web

Pengujian sistem berbasis web dilakukan untuk memastikan seluruh fungsi pada aplikasi analisis sentimen temporal dapat beroperasi sesuai dengan kebutuhan fungsional yang telah ditetapkan. Pengujian mencakup proses input data, komunikasi frontend dan backend, klasifikasi sentimen, penyimpanan database, serta visualisasi dashboard temporal. Evaluasi sistem menggunakan dua pendekatan, yaitu *black box testing* dan *System Usability Scale (SUS)*. *Black box testing* diterapkan untuk memverifikasi kesesuaian fungsi sistem berdasarkan masukan dan keluaran yang dihasilkan tanpa meninjau implementasi kode program. Pengujian ini dilakukan pada frontend ReactJS dan backend Flask yang telah terintegrasi melalui REST API. Sedangkan *System Usability Scale (SUS)* digunakan untuk menilai tingkat penerimaan dan kebergunaan sistem yang telah dikembangkan. Metode ini dipilih karena mampu memberikan ukuran yang sederhana namun efektif dalam mengevaluasi kenyamanan, kemudahan penggunaan, dan tingkat penerimaan terhadap aplikasi sistem berbasis web. Proses pengujian dilakukan dengan melibatkan sejumlah responden yang menggunakan sistem secara langsung, kemudian memberikan penilaian melalui kuesioner SUS berdasarkan pengalaman penggunaan sistem.

4.4.2 Prosedur Pengujian Model

4.4.2.1 Prosedur Pengujian Model

Tahapan pengujian model dilakukan sebagai berikut:

1. Membagi dataset menjadi data *training* dan *testing* dengan rasio 80:20.
2. Melakukan teknik peningkatan data pada data training.
3. Melakukan *preprocessing* dan ekstraksi fitur.
4. Melakukan *training* model menggunakan GridSearchCV.
5. Melakukan evaluasi menggunakan data *testing*.
6. Membandingkan performa beberapa eksperimen model.

4.4.2.2 Prosedur Pengujian Sistem Berbasis Web

1. Blackbox Testing



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Skenario pengujian dilakukan terhadap fitur-fitur utama sistem yang meliputi autentikasi pengguna, input data, klasifikasi sentimen, dashboard temporal, dan histori analisis. Setiap fitur diuji berdasarkan skenario input tertentu dan dibandingkan hasilnya dengan output yang diharapkan.

Tabel 4.7 Skenario Black Box Testing

Kode	Kasus	Skenario	Hasil yang diharapkan
1-1	Registrasi	Pengguna mendaftarkan akun dengan memasukkan <i>username</i> , email dan <i>password</i>	Sistem dapat menampilkan akun berhasil dibuat
1-2		Pengguna gagal mendaftar karena <i>username</i> atau email sudah terdaftar	Sistem menampilkan pesan eror <i>username</i> atau email sudah digunakan
1-3		Pengguna gagal mendaftar karena tidak mengisi salah satu <i>field</i>	Sistem menampilkan pesan eror untuk melengkapi semua <i>field</i>
2-1	Login	Pengguna memasukkan <i>username</i> dan <i>password</i>	Sistem menampilkan halaman dashboard
2-2		Pengguna gagal login dengan informasi akun yang tidak valid	Sistem menampilkan pesan eror kredensial tidak valid
2-3		Pengguna tidak mengisi salah satu <i>field</i>	Sistem menampilkan pesan eror untuk melengkapi semua <i>field</i>
3-1	Reset password	Pengguna memasukkan email dan mengubah <i>password</i>	Sistem mengirimkan tautan reset password ke email pengguna
3-2		Pengguna memasukkan email yang tidak terdaftar	Sistem menampilkan pesan eror bahwa email tidak ditemukan

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Kode	Kasus	Skenario	Hasil yang diharapkan
4	Input teks manual	Pengguna memasukkan ulasan	Sistem menampilkan hasil sentimen
5-1	Upload dataset CSV/XLSX	Pengguna mengunggah file dataset	Sistem memproses seluruh data dan menampilkan hasil analisis sentimen
5-2		Pengguna mengunggah file CSV/XLSX dengan struktur kolom yang tidak sesuai	Sistem menampilkan pesan eror gagal memproses file
6-1	Scraping Google Play Store	Pengguna melakukan <i>scraping</i>	Sistem berhasil mengambil ulasan dari Google Play Store dan memprosesnya dan menampilkan hasil analisis sentimen
6-2		Pengguna melakukan <i>scraping</i> dengan melewati batas tanggal yang ditentukan	Sistem menampilkan pesan eror batas rentang tanggal yang dapat dianalisis
7-1	Deteksi layanan Gojek	Pengguna memasukkan ulasan yang mengandung nama layanan Gojek	Sistem berhasil mendeteksi dan menampilkan layanan yang teridentifikasi
7-2		Pengguna memasukkan ulasan yang tidak mengandung nama layanan Gojek	Sistem menampilkan hasil tanpa deteksi layanan atau menandai sebagai tidak terdeteksi



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Kode	Kasus	Skenario	Hasil yang diharapkan
8-1	Dashboard temporal	Pengguna membuka halaman dashboard dengan data yang telah tersedia	Grafik temporal berhasil ditampilkan sesuai data yang ada
8-2		Pengguna membuka halaman dashboard saat belum ada data analisis	Sistem menampilkan halaman kosong dashboard dengan kondisi data kosong
9-1	Filter tanggal dashboard	Pengguna mengatur rentang tanggal yang valid pada dashboard	Grafik temporal diperbarui sesuai rentang tanggal yang dipilih
9-2		Pengguna mengatur rentang tanggal yang tidak valid	Sistem akan menampilkan grafik kosong
10	Grafik tren sentimen dan streamgrapgh	Pengguna mengubah filter periode waktu pada grafik	Sistem menampilkan grafik sesuai dengan periode waktu yang dipilih
11-1	Word cloud	Pengguna memfilter <i>word cloud</i> berdasarkan sentimen positif, netral, dan negatif	<i>Word cloud</i> diperbarui dan menampilkan kata-kata sesuai sentimen yang dipilih
11-2		Pengguna membuka <i>word cloud</i> saat belum ada data analisis	Sistem menampilkan kondisi kosong
12	Download hasil klasifikasi	Pengguna mengklik tombol unduh setelah hasil klasifikasi tersedia	File csv berhasil diunduh

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Kode	Kasus	Skenario	Hasil yang diharapkan
13-1	Histori analisis	Pengguna membuka halaman histori setelah sebelumnya melakukan analisis	Riwayat analisis berhasil ditampilkan beserta detail waktu dan hasilnya
13-2		Pengguna membuka halaman histori saat belum pernah melakukan analisis	Sistem menampilkan halaman histori dengan kondisi data kosong

2. System Usability Scale (SUS)

Pengujian *usability* dilakukan terhadap pengguna yang telah mencoba sistem analisis sentimen temporal berbasis web secara langsung. Responden terdiri dari pengguna umum yang memiliki pengalaman menggunakan aplikasi Gojek dan aktif memberikan ulasan pada aplikasi Gojek di Google Play Store. Kuesioner terdiri dari 10 pernyataan dengan skala likert 1-5, daftar pernyataan SUS yang digunakan adalah sebagai berikut:

Tabel 4.8 Daftar Pernyataan Kuesioner SUS

No	Pertanyaan	Sangat Tidak Setuju	Tidak Setuju	Ragu-Ragu	Setuju	Sangat Setuju
1	Saya merasa ingin menggunakan sistem ini lagi	1	2	3	4	5
2	Sistem terasa rumit untuk digunakan	1	2	3	4	5
3	Sistem mudah digunakan	1	2	3	4	5
4	Saya membutuhkan bantuan teknis untuk menggunakan sistem ini	1	2	3	4	5



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

No	Pertanyaan	Sangat Tidak Setuju	Tidak Setuju	Ragu-Ragu	Setuju	Sangat Setuju
5	Fitur-fitur pada sistem berjalan dengan baik	1	2	3	4	5
6	Sistem terasa membingungkan	1	2	3	4	5
7	Sistem mudah dipahami oleh pengguna	1	2	3	4	5
8	Sistem terasa sulit digunakan	1	2	3	4	5
9	Saya merasa percaya diri menggunakan sistem ini	1	2	3	4	5
10	Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem	1	2	3	4	5

Perhitungan skor SUS dilakukan menggunakan aturan berikut:

1. Untuk pernyataan bernomor ganjil (1,3,5,7,9): Skor kontribusi = nilai jawaban -1
2. Untuk pernyataan bernomor genap (2,4,6,8,10): Skor kontribusi = 5 – nilai jawaban
3. Seluruh skor kontribusi dijumlahkan kemudian dikalikan 2,5.

Rumus SUS:

$$SUS = (\sum \text{skor kontribusi}) \times 2.5 \quad (5)$$

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Nilai SUS berada pada rentang 1-100 dan diinterpretasikan berdasarkan kategori berikut:

Rentang Skor	Kategori	Grade
> 80,3	Excellent	A
68 – 80,3	Good	B
51 – 68	OK	C
25 -51	Poor	D
< 25	Awful	E

4.4.3 Data Hasil Pengujian

4.4.3.1. Hasil Pengujian Model

Pengujian dilakukan dalam dua tahap eksperimen. Tahap pertama membandingkan berbagai konfigurasi arsitektur model, dan tahap kedua membandingkan pengaruh teknik peningkatan data terhadap performa model.

Tabel 4.9 Hasil Eksperimen Konfigurasi Arsitektur Model

No	Eksperimen	Modifikasi	Accuracy	Precision	Recall	F1-Score
1	Baseline	IndoBERT + XGBoost	0.79	0.79	0.76	0.77
2	Baseline + Fine-Tune	IndoBERT Fine-Tune + XGBoost	0.78	0.76	0.75	0.75
3	IndoBERT Fine-Tune	IndoBERT Only Fine- Tune	0.79	0.78	0.77	0.77
4	Tanpa TF- IDF	IndoBERT + XGBoost	0.78	0.78	0.75	0.76

Berikut tabel hasil pengujian dengan melakukan beberapa teknik peningkatan data:



Tabel 4.10 Hasil Eksperimen Teknik Peningkatan Data

No	Eksperimen	Modifikasi	Accuracy	Precision	Recall	F1-Score
1	Augmentasi Teks	wordnet + synonym replacement, random swap&delete	0.82	0.81	0.79	0.80
2	Data Sintetis	+ 750 ulasan AI Generated Gemini	0.78	0.77	0.75	0.76
3	Data Sintetis 2	+ 750 ulasan AI Generated Groq	0.79	0.79	0.76	0.77
4	Pseudo Labeling	+ 1.200 pseudo label dari 53.093 data	0.74	0.75	0.73	0.73
5	Pseudo + Aug	+1.200 pseudo label dari 53.093 data	0.81	0.81	0.78	0.79

Berdasarkan hasil eksperimen, model terbaik adalah eksperimen Augmentasi Teks dengan konfigurasi IndoBERT + XGBoost yang dikombinasikan dengan teknik augmentasi teks menggunakan *WordNet synonym replacement*, *random swap*, dan *random delete*. Model ini menghasilkan *accuracy* sebesar 0,82, *precision* 0,81, *recall* 0,79 dan *F1-score* 0,80.

4.4.3.1.1 Confusion Matrix Model Terbaik

Evaluasi akhir model terbaik yaitu eksperimen yang menggunakan Augmentasi Teks dilakukan pada *data testing*, yaitu 20% dari total 600 data ulasan yang sudah dipisahkan sejak awal dan tidak pernah dilihat oleh model selama proses *training*

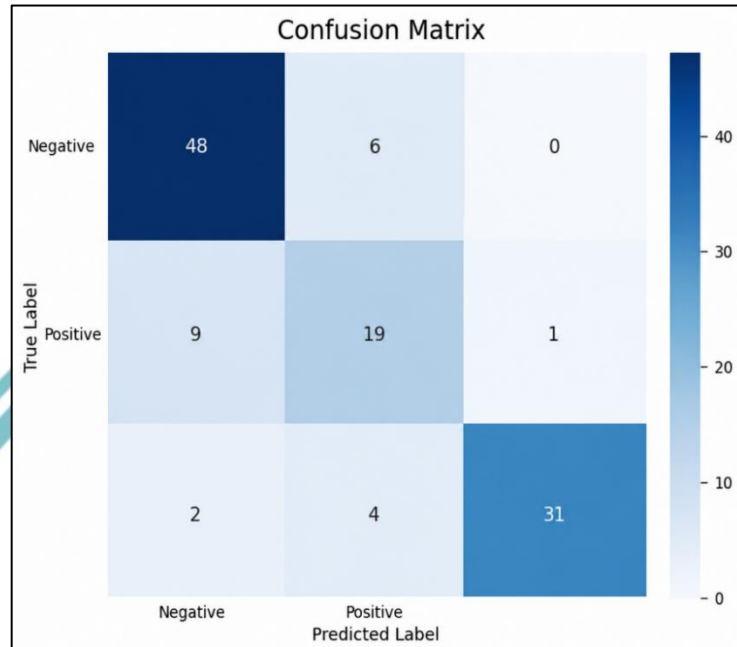
- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

maupun GridSearchCV. Hal ini sesuai dengan prosedur pengujian poin 5 (Bagian 4.4.2.1), yaitu melakukan evaluasi menggunakan *data testing*. Hasil prediksi model pada *data testing* disajikan dalam gambar 4.42 berikut.



Gambar 4.42 Confusion Matrix Model Akhir

Dari gambar 4.42, didapatkan total keseluruhan *data testing* adalah $48+6+0+9+19+1+2+4+31 = 120$ data, terdiri atas 54 data kelas negatif, 29 data kelas positif dan 37 data kelas netral.

Perhitungan Manual Metrik Evaluasi

Nilai *accuracy* sebesar 0,82 dihitung langsung dari confusion matrix pada gambar 4.42, dengan rincian perhitungan sebagai berikut.

a. Accuracy

Accuracy dihitung dengan membagi jumlah seluruh prediksi yang benar (nilai pada diagonal utama confusion matrix) dengan jumlah seluruh data.

$$\text{Accuracy} = \text{TP Negative} + \text{TP Positive} + \text{TP Netral} / \text{Total Data}$$

$$\text{Accuracy} = (48 + 19 + 31) / 120$$

$$\text{Accuracy} = 0,8167 \sim 0,82$$



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Nilai 98 diperoleh dari penjumlahan data yang diprediksi tepat pada masing-masing kelas, yaitu 48 data negatif, 19 data positif dan 31 data netral. Angka inilah yang menjadi sumber nilai akurasi.

b. Precision, Recall dan F1-Score

Precision mengukur proporsi prediksi positif suatu kelas yang benar-benar tepat, sedangkan *recall* mengukur proporsi data aktual suatu kelas yang berhasil dikenali model.

Perhitungan untuk kelas Negatif

$$TP = 48, FP = 9+2 = 11, FN = 6+0 = 6$$

$$Precision = 48 / (48 + 11) = 48 / 59 = 0,8136$$

$$Recall = 48 / (48 + 6) = 48 / 54 = 0,8889$$

$$F1-Score = 2 \times (0,8136 \times 0,8889) / (0,8136 + 0,8889) = 0,6552$$

Perhitungan untuk kelas Positif

$$TP = 19, FP = 6+4 = 10, FN = 9+1 = 10$$

$$Precision = 19 / (19 + 10) = 19 / 29 = 0,6552$$

$$Recall = 19 / (19 + 10) = 19 / 29 = 0,6552$$

$$F1-Score = 2 \times (0,6552 \times 0,6552) / (0,6552 + 0,6552) = 0,6552$$

Perhitungan untuk kelas Netral

$$TP = 31, FP = 0+1 = 1, FN = 2+4 = 6$$

$$Precision = 31 / (31 + 1) = 31 / 32 = 0,9688$$

$$Recall = 31 / (31 + 6) = 31 / 37 = 0,8378$$

$$F1-Score = 2 \times (0,9688 \times 0,8378) / (0,9688 + 0,8378) = 0,8986$$

Ringkasan hasil perhitungan per kelas disajikan pada tabel 4.11. nilai *macro average* dari *precision* (0,8125), *recall* (0,7940), dan *F1-Score* (0,8011) diperoleh dari rata-rata sederhana ketiga kelas, dan nilai ini sejalan dengan *precision* 0,81, *recall* 0,79, dan *F1-score* 0,80 yang dilaporkan pada tabel 4.11.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Tabel 4.11 Hasil Perhitungan Per Kelas

Kelas	Precision	Recall	F1-Score	Support
Negative	0.8136	0,8889	0,8496	54
Positive	0,6552	0,6552	0,6552	29
Netral	0,9688	0,8378	0,8986	37
Macro Average	0,8125	0,7940	0,8011	120

4.4.3.2. Hasil Pengujian Sistem

1. Hasil Black Box Testing

Berdasarkan hasil pengujian yang telah dilakukan, seluruh fitur utama pada sistem dapat berjalan dengan baik. Sistem mampu menerima input pengguna, menjalankan model klasifikasi sentimen, serta menampilkan hasil analisis dalam bentuk dashboard visualisasi temporal.

Tabel 4.12 Hasil Pengujian Sistem Berbasis Web

Kode	Kasus	Skenario	Status
1-1	Registrasi	Pengguna mendaftarkan akun dengan memasukkan <i>username</i> , email dan <i>password</i>	Berhasil
1-2		Pengguna gagal mendaftar karena <i>username</i> atau email sudah terdaftar	Berhasil
1-3		Pengguna gagal mendaftar karena tidak mengisi salah satu <i>field</i>	Berhasil
2-1	Login	Pengguna memasukkan <i>username</i> dan <i>password</i>	Berhasil
2-2		Pengguna gagal login dengan informasi akun yang tidak valid	Berhasil
2-3		Pengguna tidak mengisi salah satu <i>field</i>	Berhasil
3-1	Reset password	Pengguna memasukkan email dan mengubah <i>password</i>	Berhasil
3-2		Pengguna memasukkan email yang tidak terdaftar	Berhasil



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Kode	Kasus	Skenario	Status
4	Input teks manual	Pengguna memasukkan ulasan	Berhasil
5-1	Upload dataset CSV/XLSX	Pengguna mengunggah file dataset	Berhasil
5-2		Pengguna mengunggah file CSV/XLSX dengan struktur kolom yang tidak sesuai	Berhasil
6-1	Scraping Google Play Store	Pengguna melakukan <i>scraping</i>	Berhasil
6-2		Pengguna melakukan <i>scraping</i> dengan melewati batas tanggal yang ditentukan	Berhasil
7-1	Deteksi layanan Gojek	Pengguna memasukkan ulasan yang mengandung nama layanan Gojek	Berhasil
7-2		Pengguna memasukkan ulasan yang tidak mengandung nama layanan Gojek	Berhasil
8-1	Dashboard temporal	Pengguna membuka halaman dashboard dengan data yang telah tersedia	Berhasil
8-2		Pengguna membuka halaman dashboard saat belum ada data analisis	Berhasil
9-1	Filter tanggal dashboard	Pengguna mengatur rentang tanggal yang valid pada dashboard	Berhasil
9-2		Pengguna mengatur rentang tanggal yang tidak valid	Berhasil
10	Grafik tren sentimen dan streamgrapgh	Pengguna mengubah filter periodewaktu pada grafik	Berhasil
11-1	Word cloud	Pengguna memfilter <i>word cloud</i> berdasarkan sentimen positif, netral, dan negatif	Berhasil

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Kode	Kasus	Skenario	Status
11-2		Pengguna membuka <i>word cloud</i> saat belum ada data analisis	Berhasil
12	Download hasil klasifikasi	Pengguna mengklik tombol unduh setelah hasil klasifikasi tersedia	Berhasil
13-1	Histori analisis	Pengguna membuka halaman histori setelah sebelumnya melakukan analisis	Berhasil
13-2		Pengguna membuka halaman histori saat belum pernah melakukan analisis	Berhasil

2. Hasil System Usability Scale (SUS)

Pengujian SUS melibatkan 30 responden yang berasal dari para pengguna aplikasi Gojek yang aktif memberikan ulasan pada aplikasi Gojek di Google Play Store.

Perhitungan skor SUS per responden ditampilkan pada tabel berikut:

Tabel 4.13 Skor *System Usability Scale* (SUS) PerResponden

Responden	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Skor
R1	4	1	5	1	4	1	5	1	4	2	90
R2	5	1	5	3	4	2	5	1	5	2	87.5
R3	5	1	5	1	5	1	5	1	5	2	97.5
R4	3	1	5	1	4	1	4	1	5	1	90
R5	3	2	4	1	4	2	4	1	3	5	67.5
R6	3	2	5	1	5	2	5	1	5	1	90
R7	4	2	5	2	3	1	5	1	5	1	87.5
R8	3	2	4	3	4	3	4	3	3	4	57.5
R9	4	1	5	1	5	1	5	1	5	1	97.5
R10	4	3	3	3	3	3	3	3	3	3	52.5
R11	3	3	3	3	3	3	3	3	3	3	50
R12	4	2	5	2	4	2	5	2	5	2	82.5
R13	3	2	4	2	4	2	4	1	3	4	67.5
R14	3	3	4	3	3	3	3	3	3	3	52.5
R15	4	2	5	3	4	1	4	2	4	3	75



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Responden	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Skor
R16	5	1	5	1	5	1	5	1	5	1	100
R17	3	2	5	1	4	2	4	2	5	3	77.5
R18	5	2	5	1	5	1	4	1	4	2	90
R19	3	2	4	3	3	2	4	2	4	3	65
R20	4	2	4	3	4	2	4	2	4	4	67.5
R21	5	1	5	1	5	1	5	1	5	2	97.5
R22	4	2	4	3	4	2	4	1	5	4	72.5
R23	5	4	4	4	5	2	4	2	4	3	67.5
R24	4	2	4	2	4	2	4	3	4	2	72.5
R25	3	2	4	3	5	2	4	2	4	3	70
R26	3	2	5	3	5	2	4	2	5	4	72.5
R27	4	2	3	2	4	2	4	2	4	4	67.5
R28	4	1	5	2	5	2	4	1	5	3	85
R29	5	1	5	1	5	1	5	1	5	1	100
R30	5	1	5	1	5	1	5	1	5	1	100
Rata-rata											78.3

4.4.4 Analisis Hasil Pengujian

4.4.4.1. Analisis Hasil Pengujian Model

Berdasarkan hasil pengujian pada tabel 4.10, dapat dianalisis bahwa teknik augmentasi teks memberikan pengaruh paling signifikan terhadap peningkatan performa model. Model dengan augmentasi teks menghasilkan akurasi tertinggi sebesar 0,82 dibandingkan seluruh eksperimen lainnya.

Beberapa temuan penting dari hasil eksperimen adalah sebagai berikut:

1. *Baseline* (IndoBERT + XGBoost) menghasilkan akurasi 0,79 dan menjadi acuan perbandingan bagi seluruh eksperimen.
2. *Fine-tuning* IndoBERT tidak memberikan peningkatan yang signifikan bahkan sedikit menurunkan performa menjadi 0,78, yang kemungkinan disebabkan oleh keterbatasan ukuran dataset pelatihan sehingga proses *fine-tuning* sebelum optimal.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

3. Augmentasi teks menghasilkan performa terbaik dengan *accuracy* 0,82 dan *F1-score* 0,80. Hal ini menunjukkan bahwa penambahan variasi *data training* melalui augmentasi teks efektif dalam meningkatkan generalisasi model.
4. Data sintesis dari model AI Generatif (Gemini dan Groq) tidak memberikan peningkatan berarti, bahkan data sintesis Gemini menurunkan performa menjadi 0,78. Hal ini mengindikasikan bahwa data sintesis yang dihasilkan AI memiliki distribusi yang kurang sesuai dengan karakteristik data asli.
5. Pseudo labeling menurunkan performa menjadi 0,75, namun kombinasinya dengan augmentasi (pseudo + aug) menghasilkan performa 0,81, mendekati model terbaik.

Berdasarkan analisis di atas, model Augmentasi Teks dipilih sebagai model final karena menghasilkan keseimbangan terbaik antara *accuracy*, *precision*, *recall*, dan *F1-score* dengan nilai masing-masing sebesar 0,82, 0,81, 0,79, dan 0,80.

Berdasarkan hasil *confusion matrix* pada gambar 4.42, model menunjukkan kemampuan yang baik dalam mengelompokkan data ke dalam tiga kategori sentimen. Pada kelas negatif, sebanyak 48 data berhasil diklasifikasikan secara tepat, sedangkan 6 data lainnya keliru diprediksi sebagai sentimen positif. Untuk kelas positif, terdapat 19 data yang teridentifikasi dengan benar, sementara 9 data diprediksi sebagai negatif dan 1 data masuk ke kategori lain. Adapun pada kelas netral, model mampu mengklasifikasikan 31 data secara akurat, dengan 2 data salah diprediksi sebagai negatif dan 4 data sebagai positif. Hasil tersebut mengindikasikan bahwa model telah mampu membedakan ketiga kelas sentimen dengan cukup baik. Namun, masih ditemukan beberapa kesalahan prediksi, terutama antara kelas positif dan negatif, yang mengindikasikan adanya kemiripan karakteristik linguistik pada sebagian ulasan sehingga sulit dibedakan oleh model.

Walaupun mode terbaik yang dihasilkan memperoleh nilai akurasi sebesar 0,82 performanya masih memiliki ruang untuk ditingkatkan. Hal ini dapat dijelaskan oleh beberapa faktor. Pertama, keterbatasan ukuran dataset menjadi faktor utama, di mana jumlah data training hanya sebesar 600 sampel dan hanya bertambah



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

menjadi 615 setelah augmentasi. Jumlah ini tergolong sangat kecil untuk melatih model berbasis *transformer* seperti IndoBERT yang memiliki parameter dalam jumlah besar, sehingga model belum bisa mempelajari pola bahasa secara optimal dan rentan terhadap kondisi *underfitting*. Kedua, meskipun augmentasi teks dengan teknik *synonym replacement*, *random swape*, dan *random delete* berhasil menambah variasi data, penambahan jumlah sampel yang dihasilak relatif sedikit, yaiut hanya 15 data. Penambahan tersebut belum cukup untuk memberikan peningkatan performa yang lebih signifikan. Ketiga, distribusi kelas pada dataset yang digunakan masih belum sepenuhnya seimbang. Kondisi ini dapat menyebabkan model lebih mudah mempelajari pola pada kelas tertentu dibandingkan kelas lainnya, sehingga memengaruhi kemampuan model dalam melakukan klasifikasi secara merata pada seluruh kategori sentimen.

Berdasarkan hasil tersebut, dapat disimpulkan bahwa penggunaan augmentasi teks berhasil meningkatkan performa model dibandingkan pendekatan lainnya. Namun, peningkatan yang diperoleh masih dibatasi oleh jumlah dan kualitas data yang tersedia. Oleh karena itu, penelitian selanjutnya dapat difokuskan pada pengumpulan dataset yang lebih besar dan lebih beragam agar performa model klasifikasi sentimen dapat ditingkatkan lebih lanjut.

4.4.4.2 Analisis Hasil Pengujian Sistem Berbasis Web

1. Analisis Black Box Testing

Hasil *black box testing* yang dilakukan terhadap 13 skenario pengujian menunjukkan bahwa seluruh fungsi yang diuji dapat berjalan sesuai dengan kebutuhan sistem. Seluruh scenario memperoleh hasil yang sesuai dengan keluaran yang diharapkan sehingga tingkat keberhasilan pengujian mencapai 100% (13 dari 13 fitur berhasil). Hal ini mengindikasikan bahwa implementasi sistem telah berjalan dengan baik, termasuk proses pertukaran data frontend ReactJS dan backend Flask berjalan lancar melalui REST API. Selain itu, proses integrasi model klasifikasi sentimen ke dalam aplikasi sistem berbasis web dapat beroperasi secara stabil tanpa ditemukan kendala fungsional yang signifikan selama proses pengujian.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2. System Usability Scale (SUS)

Berdasarkan hasil perhitungan kuesioner SUS yang diberikan kepada 30 responden, diperoleh rata-rata skor SUS sebesar 78,3. Mengacu pada interpretasi skor SUS, nilai tersebut berada pada rentang 68-80,3, yang masuk ke dalam kategori “Good” atau baik dengan *grade* B. Hasil ini mengindikasikan bahwa sistem yang dikembangkan memiliki tingkat kegunaan yang baik, mudah digunakan oleh pengguna, serta memperoleh penerimaan yang positif dari pengguna.

Dari data per responden pada tabel 4.12, terdapat variasi skor antara 50 (R10) hingga 100 (R16, R29, dan R30). Responden dengan skor tinggi memberikan nilai tinggi pada pernyataan positif dan nilai terendah pada pernyataan negatif, mencerminkan pengalaman penggunaan yang sangat memuaskan. Responden dengan skor lebih rendah cenderung memberikan jawaban netral (nilai 3) pada seluruh pernyataan, yang dapat mengindikasikan kurangnya familiaritas dengan sistem analitik berbasis web. Secara keseluruhan, rata-rata skor SUS 78,3 menunjukkan bahwa sistem analisis sentimen temporal berbasis web yang dikembangkan memiliki *usability* yang baik dan layak digunakan sebagai alat analisis sentimen oleh pengguna umum maupun peneliti.

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB V PENUTUP

5.1 Simpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat ditarik kesimpulan sebagai berikut:

1. Sistem analisis sentimen temporal berbasis web berhasil dibangun dan diimplementasikan menggunakan pendekatan *hybrid* IndoBERT-XGBoost pada ulasan aplikasi Gojek berbahasa Indonesia di Google Play Store. Sistem mampu mengklasifikasikan sentimen pengguna ke dalam tiga kategori, yaitu positif, netral, dan negatif, serta menampilkan dinamika perubahan sentimen dari waktu ke waktu melalui dashboard visualisasi interaktif.
2. Model *hybrid* IndoBERT-XGBoost dengan augmentasi teks menghasilkan performa terbaik dibandingkan seluruh eksperimen yang dilakukan, dengan nilai *accuracy* 0,82, *precision* 0,81, *recall* 0,79, dan *F1-score* 0,80. Hasil ini menunjukkan bahwa kombinasi *embedding* semantik IndoBERT dengan kemampuan klasifikasi XGBoost, yang diperkuat dengan teknik augmentasi teks *synonym replacement*, *random swap*, dan *random delete*, efektif dalam meningkatkan generalisasi model pada data ulasan berbahasa Indonesia yang bersifat informal.
3. Penggunaan *feature concatenation* dari lima kelompok fitur, yaitu *embedding* semantik IndoBERT, TF-IDF, fitur linguistik, fitur sentimen berbasis leksikon, dan layanan hasil *Named Entity Recognition* (NER), terbukti memberikan representasi data yang komprehensif dibandingkan penggunaan fitur tunggal. Analisis SHAP menunjukkan bahwa fitur *embedding* IndoBERT mendominasi kontribusi prediksi pada seluruh kelas sentimen, sementara fitur sentimen berbasis leksikon memberikan kontribusi tambahan yang signifikan terutama pada kelas negatif.
4. Implementasi *Named Entity Recognition* (NER) berbasis *dictionary matching* berhasil mengidentifikasi layanan Gojek yang paling banyak dibahas pengguna. Berdasarkan hasil analisis terhadap 53.093 ulasan selama 12 bulan, layanan GoRide menjadi layanan yang paling banyak mendapat ulasan dengan total

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

7.073 ulasan, diikuti oleh GoFood, GoPay, dan GoCar. Temuan ini menunjukkan bahwa layanan transportasi dan pembayaran digital merupakan fitur yang paling sering digunakan sekaligus paling banyak mendapat perhatian dari pengguna.

5. Hasil analisis temporal menunjukkan bahwa sentimen positif mendominasi sepanjang periode pengamatan proporsi 65% dari total ulasan, diikuti sentimen negatif sebesar 24% dan sentimen netral sebesar 10%. Puncak aktivitas ulasan terjadi pada bulan April hingga Mei 2025, yang tercermin pada peningkatan jumlah ulasan positif maupun negatif secara bersamaan pada periode tersebut.
6. Pengujian sistem berbasis web menggunakan *black box testing* menunjukkan tingkat keberhasilan fungsional 100% pada seluruh 13 skenario pengujian, yang mengindikasikan bahwa integrasi antara frontend ReactJS, backend Flask, dan model klasifikasi sentimen berjalan dengan baik. Evaluasi *usability* menggunakan *System Usability Scale* (SUS) terhadap 30 responden menghasilkan rata-rata skor 78,3, yang termasuk dalam kategori “Good” dengan gradi B, menunjukkan bahwa sistem dapat diterima dan digunakan dengan baik oleh pengguna.

5.2 Saran

Berdasarkan hasil penelitian dan keterbatasan yang ditemukan, beberapa saran untuk pengembangan penelitian selanjutnya adalah sebagai berikut:

1. Perluasan dataset pelatihan perlu dilakukan untuk meningkatkan performa model lebih lanjut. Dataset yang digunakan dalam penelitian ini hanya berjumlah 600 sampel berlabel, yang tergolong kecil untuk melatih model berbasis *transformer*. Penambahan data berlabel yang lebih besar dan beragam diharapkan dapat meningkatkan kemampuan generalisasi model secara signifikan.
2. Pengembangan kamus leksikon dan kata kunci (NER) yang lebih komprehensif perlu dilakukan untuk meningkatkan akurasi deteksi sentimen dan identifikasi layanan. Kamus yang digunakan saat ini masih terbatas dan perlu diperluas untuk mencakup variasi bahasa informal, slang baru, dan istilah layanan yang terus berkembang.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

3. Penambahan fitur analisis aspek (*aspect-based sentiment analysis*) dapat dipertimbangkan agar sistem tidak hanya mampu mengklasifikasikan sentimen secara keseluruhan, tetapi juga mengidentifikasi aspek spesifik dari layanan yang mendapat sentimen positif maupun negatif, sehingga memberikan informasi yang lebih detail bagi pengembang aplikasi.
4. Perluasan cakupan *platform* dan aplikasi dapat dilakukan pada penelitian selanjutnya. Sistem yang dikembangkan saat ini hanya mencakup ulasan dari Google Play Store dan aplikasi Gojek. Integrasi dengan *platform* ulasan lain seperti App Store maupun perluasan ke aplikasi *super-app* lainnya dapat memberikan gambaran perbandingan yang lebih komprehensif.





DAFTAR PUSTAKA

- Adhim, N.F. and Cahyono, N. (2025) "Optimization of IndoBERT for Sentiment Analysis of FOMO on Social Media Through Fine-Tuning and Hybrid Labeling," *Journal of Applied Informatics and Computing (JAIC)*, 9(6), pp. 3786–3797.
- Afifah, K., Yulita, I.N. and Sarathan, I. (2021) "Sentiment Analysis on Telemedicine App Reviews using XGBoost Classifier," *2021 International Conference on Artificial Intelligence and Big Data Analytics*, pp. 22–27. Available at: <https://doi.org/10.1109/ICAIBDA53487.2021.9689735>.
- Amien, M. (2023) "SEJARAH DAN PERKEMBANGAN TEKNIK NATURAL LANGUAGE PROCESSING (NLP) BAHASA INDONESIA: TINJAUAN TENTANG SEJARAH, PERKEMBANGAN TEKNOLOGI, DAN APLIKASI NLP DALAM BAHASA INDONESIA," pp. 1–7. Available at: <https://arxiv.org/abs/2304.02746>.
- Chicco, D. and Jurman, G. (2022) "An Invitation to Greater Use of Matthews Correlation Coefficient in Robotics and Artificial Intelligence," 9(March), pp. 1–4. Available at: <https://doi.org/10.3389/frobt.2022.876814>.
- Faqih, H., Aji, S. and Suseno, K.A. (2025) "Studi Komparatif Metode Naive Bayes dan Support Vector Machine dalam Analisis Sentimen," *Multinetics*, 11(1), pp. 94–102. Available at: <https://jurnal.pnj.ac.id/index.php/multinetics/article/view/7534/3595>.
- Farizah, N. and Wang, X. (2019) "Computers in Human Behavior Decoding the sentiment dynamics of online retailing customers : Time series analysis of social media," *Computers in Human Behavior*, 96(February), pp. 32–45. Available at: <https://doi.org/10.1016/j.chb.2019.02.004>.
- Godang, R.H. *et al.* (2025) "Sistem Informasi Lembaga Penjamin Mutu (LPM) pada AMIK Luwuk Banggai Berbasis Web Program Studi Keamanan Sistem Informasi , Politeknik Banggai Industri , Indonesia," (September), pp. 210–225.
- Hakim, B. and Kinasih, P. (2023) "Sentiment Analysis of Indonesian Citizen Tweets Using Support Vector Machine on the Rebranding of Twitter to X," 18(2), pp. 468–479.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

- Harris, M., Jacobson, J. and Provetti, A. (2024) "Sentiment and time-series analysis of direct-message conversations.pdf," *Forensic Science International*, p. 14.
- Imaduddin, H., A, F.Y. and Nugroho, Y.S. (2023) "Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach," *International Journal of Advanced Computer Science and Applications*, 14(8), pp. 113–117.
- Jannah, U.N. (2024) "SISTEM INFORMASI ANALISIS AKUNTANSI PENJUALAN BERBASIS WEB PADA DJAROT CELL MENGGUNAKAN METODE TIME SERIES," *Jurnal Ilmiah Sistem Informasi*, 3, pp. 34–42. Available at: <http://ejurnal.provisi.ac.id/index.php/JUISI>.
- Moeslim, A.G.S., Firmasyah, E. and Sutara, B. (2025) "Perbandingan Kinerja XGBoost dan IndoBERT untuk Klasifikasi Teks Kesehatan Bahasa Indonesia," p. 11.
- Nur, L. and Handoko, L.B. (2025) "Optimasi Analisis Sentimen Lowongan Kerja di Twitter Dengan XGBoost-Vader dan Evaluasi SMOTE Borderline," *Jurnal Pendidikan dan Teknologi Indonesia*, 5(4), pp. 999–1010.
- Obi, J.C. (2023) "A Comparative Study of Several Classification Metrics and Their Performances on Data," 08(01), pp. 308–314. Available at: <https://doi.org/10.30574/wjaets.2023.8.1.0054>.
- Plaza, L. and Carrillo-de-albornoz, J. (2018) "Feature engineering for sentiment analysis in e-health forums," pp. 1–25.
- Putri, N.A. *et al.* (2026) "IndoBERT Embeddings vs . Fine-Tuning : An Optimized Approach for Sentiment Analysis of Indonesian Public Service Reviews," *2026 International Conference on Current Research in Artificial Intelligence and Data Science (ICCRAIDS)*, 1, pp. 1–6. Available at: <https://doi.org/10.1109/ICCRAIDS67816.2026.11519745>.
- Tan, K.L. *et al.* (2022) "RoBERTa-LSTM : A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network," *IEEE Access*, 10, pp. 21517–21525. Available at: <https://doi.org/10.1109/ACCESS.2022.3152828>.
- Tarumingkeng, R.C. (2024) *Natural Language Processing (NLP)*. Bogor.
- Wankhade, M., Rao, A.C.S. and Kulkarn, C. (2022) "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

(2022) 55:5731–5780, 55(October 2022), pp. 5731–5780. Available at:
<https://doi.org/https://doi.org/10.1007/s10462-022-10144-1>.

Widyananda, W., Fauzi, A. and Maskur (2025) “Machine Learning and Transformer-based Model for Sentiment Analysis of Indonesian E-Commerce Reviews,” 14(4), pp. 6262–6271. Available at:
<https://doi.org/https://doi.org/10.33022/ijcs.v14i4.4980>.





Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

DAFTAR RIWAYAT HIDUP



Tarissa Azzahra Danantya

Lahir di Jakarta, 14 Februari 2003. Lulus dari MI Rabiah Al-Adawiyah pada tahun 2015, MTs Negeri 20 Jakarta pada tahun 2018, dan MA Negeri 8 Jakarta pada tahun 2021. Saat ini sedang menempuh Pendidikan Sarjana Terapan di Politeknik Negeri Jakarta dengan jurusan Teknik Informatika dan Komputer program studi Teknik Informatika. Tertarik pada bidang *Data Analyst*, *Data Science* serta *Machine Learning*.

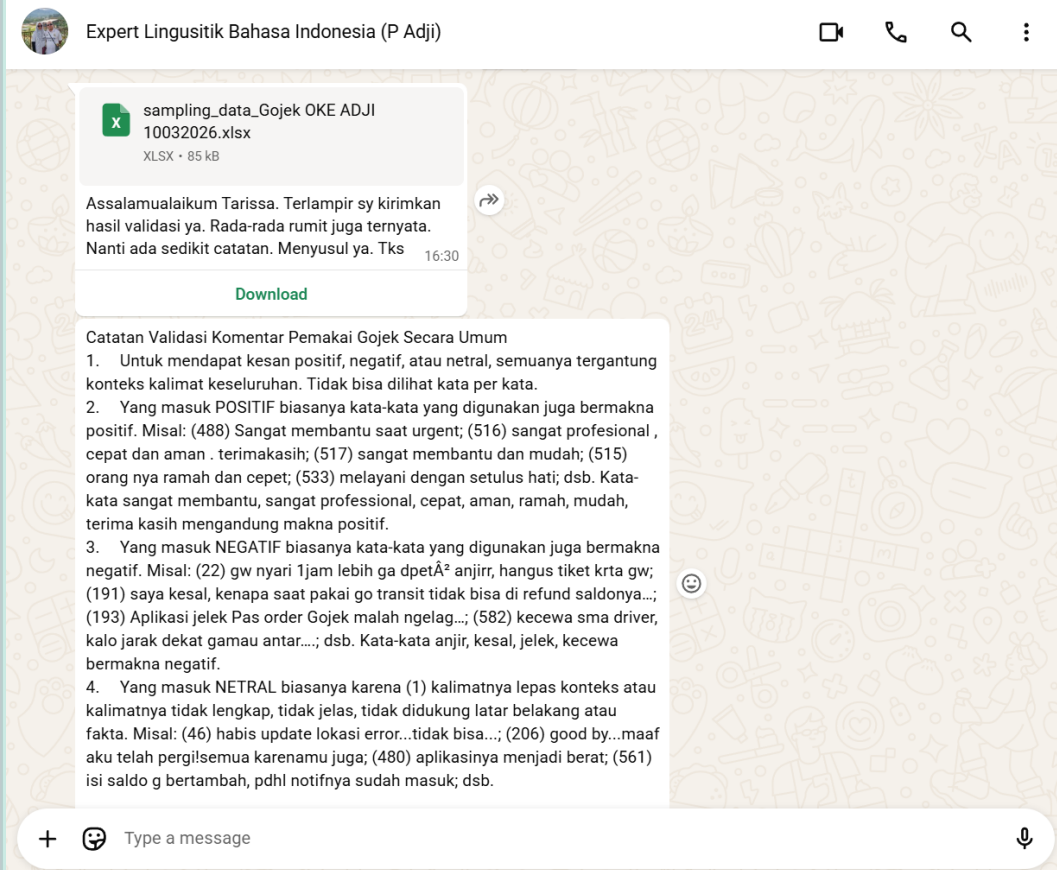
**POLITEKNIK
NEGERI
JAKARTA**

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

LAMPIRAN

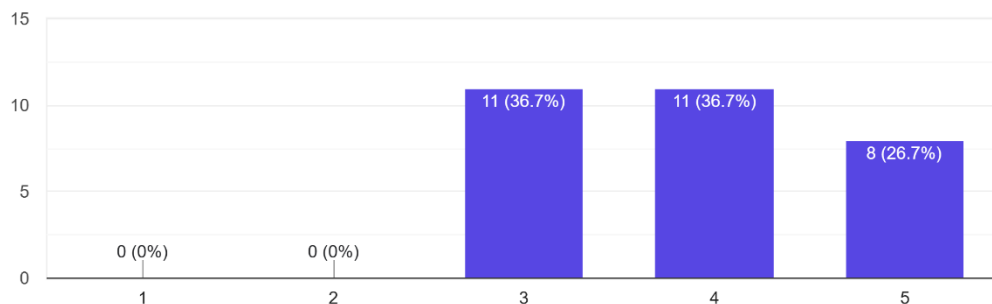
Lampiran 1 Catatan Expert Linguistik Bahasa Indonesia



Lampiran 2 Hasil Kuesioner System Usability Scale (SUS)

1. Saya merasa ingin menggunakan sistem ini lagi

30 responses



Lanjutan



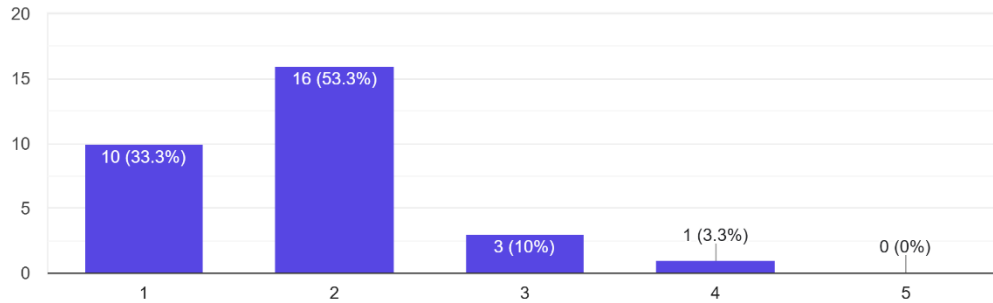
© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

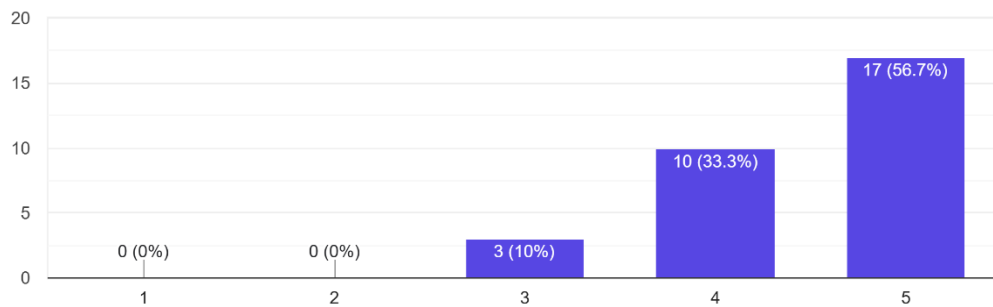
2. Sistem terasa rumit untuk digunakan

30 responses



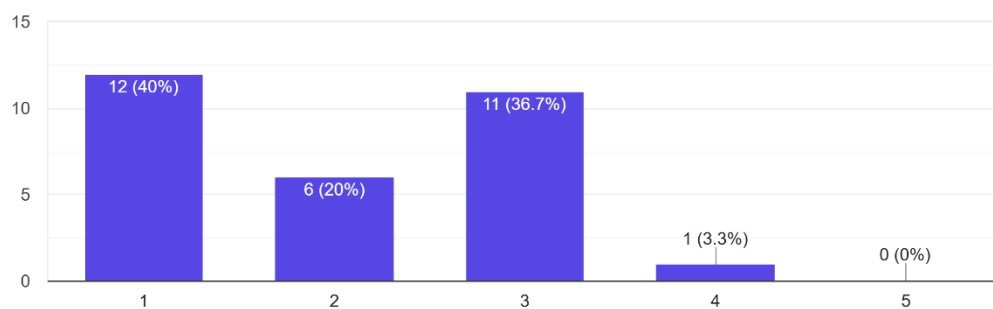
3. Sistem mudah digunakan

30 responses



4. Saya membutuhkan bantuan teknis untuk menggunakan sistem ini

30 responses



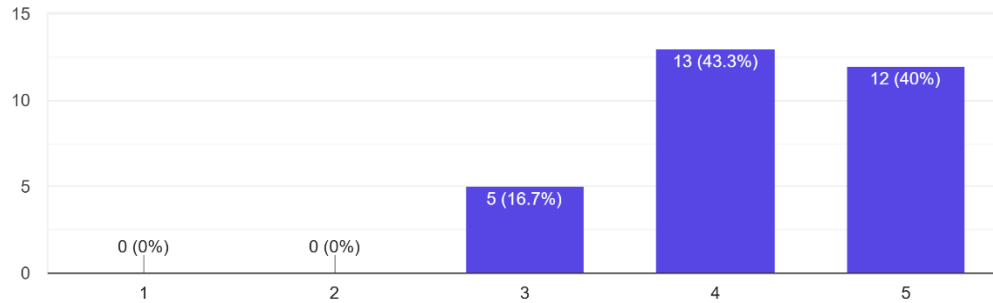


Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

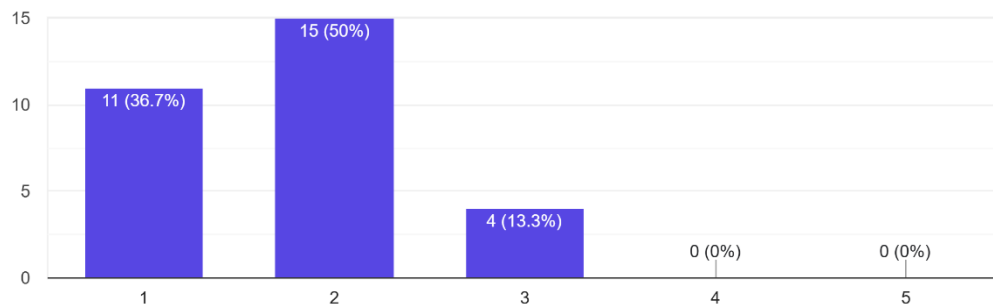
5. Fitur-fitur pada sistem berjalan dengan baik

30 responses



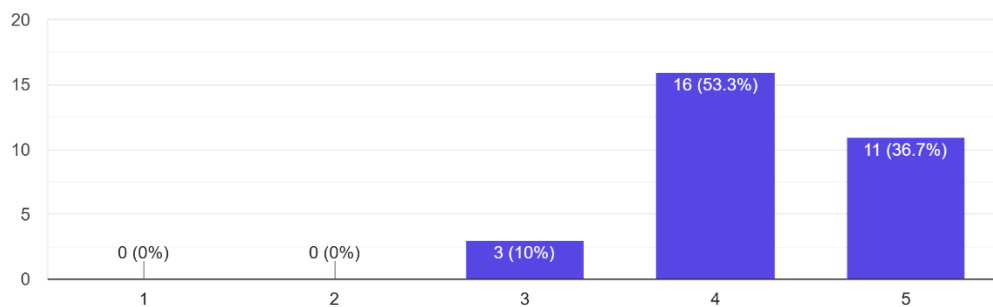
6. Sistem terasa membingungkan

30 responses



7. Sistem mudah dipahami oleh pengguna

30 responses





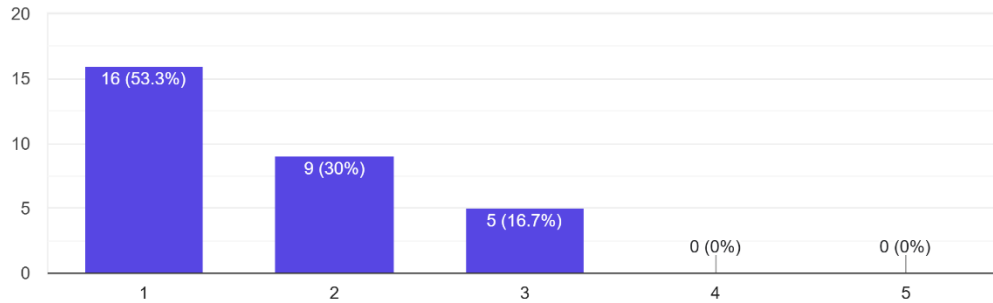
© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

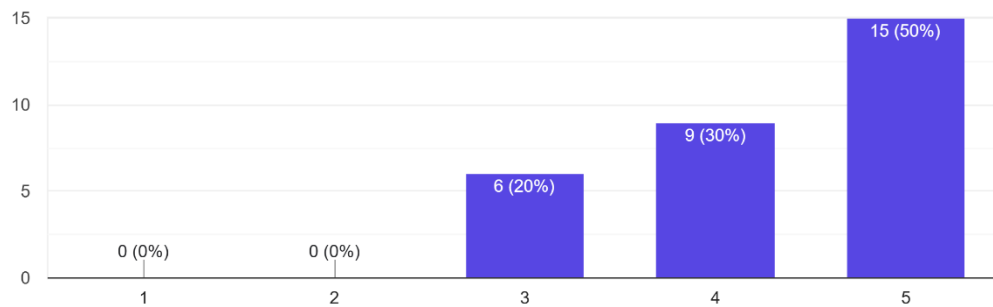
8. Sistem terasa sulit digunakan

30 responses



9. Saya merasa percaya diri menggunakan sistem ini

30 responses



10. Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem

30 responses

