

IMPLEMENTASI WEB TOOL ANALISIS SENTIMEN ULASAN APLIKASI GOOGLE PLAY STORE MENGGUNAKAN LOGISTIC REGRESSION

As'ad Nirot Ahmadi; Devi Afriyantari Puspa Putri,
Fakultas Komunikasi dan Informatika, Universitas Muhammadiyah Surakarta

Abstrak

Pesatnya pertumbuhan aplikasi di Google Play Store menghasilkan banyaknya volume ulasan pengguna, sehingga sulit dianalisis secara manual, padahal ulasan tersebut mengandung informasi penting untuk mengevaluasi kualitas aplikasi dan kepuasan pengguna. Penelitian ini bertujuan merancang sistem analisis sentimen berbasis website yang mampu mengklasifikasikan ulasan secara otomatis menjadi sentimen positif dan negatif. Sebanyak 20.000 ulasan dikumpulkan melalui teknik *scraping* menggunakan *Application ID*, kemudian dilanjutkan dengan pelabelan secara manual serta *preprocessing* yang mencakup *case folding*, normalisasi teks, dan *text cleaning*. Sistem dibangun menggunakan algoritma *Logistic Regression* serta menerapkan TF-IDF untuk mengekstraksi fitur, dengan rasio data latih dan data uji sebesar 90:10. Untuk meningkatkan ketepatan klasifikasi pada kelas positif, diterapkan *classification threshold* sebesar 0,6. Berdasarkan hasil pengujian, model memiliki performa dengan akurasi sebesar 97,70%, sedangkan evaluasi menggunakan *Repeated Stratified K-Fold Cross-Validation* menghasilkan rata-rata *F1-score* sebesar 89,47%, yang menunjukkan kestabilan kinerja model pada berbagai skema pembagian data. Model yang dihasilkan berhasil diimplementasikan ke dalam sistem berbasis web sehingga mampu melakukan analisis sentimen pada ulasan aplikasi Google Play Store secara otomatis.

Kata Kunci: analisis sentimen, google play store, *logistic regression*, TF-IDF

Abstract

The rapid growth of apps on the Google Play Store has resulted in a large volume of user reviews, making it difficult to analyze them manually—even though these reviews contain important information for evaluating app quality and user satisfaction. This study aims to design a web-based sentiment analysis system capable of automatically classifying reviews into positive and negative sentiments. A total of 20,000 reviews were collected via web scraping using Application IDs, followed by manual labeling and preprocessing that included case folding, text normalization, and text cleaning. The system was built using the Logistic Regression algorithm and applied TF-IDF for feature extraction, with a training-to-test data ratio of 90:10. To improve classification accuracy for the positive class, a classification threshold of 0.6 was applied. Based on the test results, the model achieved an accuracy of 97.70%, while evaluation using Repeated Stratified K-Fold Cross-Validation yielded an average F1-score of 89.47%, indicating the model's performance stability across various data partitioning schemes. The resulting model was successfully implemented into a web-based system, enabling it to automatically perform sentiment analysis on Google Play Store app reviews.

Keywords: google play store, sentiment analysis, logistic regression, TF-IDF.

1. PENDAHULUAN

Perkembangan teknologi perangkat seluler saat ini membawa dampak signifikan terhadap ekosistem digital, salah satunya adalah tingginya penggunaan aplikasi seluler. Berdasarkan data statistik, jumlah pengguna aktif *smartphone* di Indonesia pada tahun 2023 telah melonjak pesat hingga menembus 209,3 juta jiwa (Andalas, 2024). Pesatnya penyebaran perangkat ini berbanding lurus dengan tingginya konsumsi aplikasi digital di masyarakat. Google Play Store, sebagai platform utama distribusi aplikasi seluler di android, memfasilitasi pengguna untuk memberikan ulasan dan penilaian yang sangat berguna bagi pengembang. Namun, ulasan yang terus bertambah secara eksponensial menyulitkan pemantauan manual. Padahal, ulasan seringkali berisi keluhan teknis atau apresiasi krusial yang berdampak langsung pada reputasi kualitas aplikasi (Rullyana & Triandari, 2025). Ketidakselarasan antara penilaian bintang yang diberikan dengan isi teks yang justru bernada negatif kerap terjadi, sehingga analisis manual tidak lagi efisien (Kaburuan & Setiawan, 2023). Oleh karena itu, diperlukan otomatisasi komputasional untuk memproses opini publik ini secara masif, cepat, dan terukur.

Analisis sentiment merupakan salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut karena mampu mengekstraksi opini pengguna secara otomatis. Dalam konteks teks ulasan berbahasa Indonesia, tantangan utamanya adalah struktur kalimat yang tidak baku, penggunaan slang, serta *noise* pada teks. Hal ini menuntut adanya mekanisme *dynamic scraping* menggunakan *Application ID* serta tahap *preprocessing* yang komprehensif, meliputi *case folding*, *normalize text*, dan *text cleaning* sebelum data diolah lebih lanjut (Afif et al., 2025). Setelah teks distandarisasi, dilakukan konversi teks ke dalam representasi numerik menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*). TF-IDF terbukti esensial dalam menyoroti kata-kata berbobot tinggi yang paling mempengaruhi polaritas sentimen, menjadikannya standar ekstraksi fitur yang tangguh dalam pemrosesan bahasa alami pada skala global (Kaur & Sharma, 2023).

Sejumlah penelitian terdahulu telah mengkaji analisis sentimen menggunakan berbagai algoritma *machine learning*. Pada skala global, penelitian yang dilakukan oleh (He & Abisado, 2024) serta (Mao et al., 2024) menunjukkan tren peningkatan penggunaan kecerdasan buatan untuk mengklasifikasikan kepuasan pengguna layanan digital secara akurat di berbagai platform. Di ranah ulasan aplikasi nasional, banyak peneliti membuktikan bahwa algoritma seperti *Support Vector Machine* (SVM) maupun *Logistic Regression* memiliki kinerja yang baik dalam mendukung proses moderasi ulasan secara otomatis (Amalia et al., 2024; Lidinillah et al., 2023). Meskipun metode yang dieksplorasi sangat beragam, klasifikasi diskriminatif linear tetap menjadi pilihan utama, terutama karena performanya yang stabil dalam menangani data hasil ekstraksi TF-IDF. Namun, implementasi utuh mulai dari proses penarikan data dinamis hingga divisualisasikan ke dalam sebuah *dashboard*

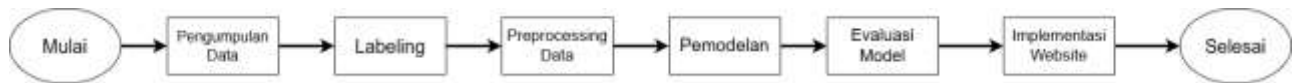
interaktif belum banyak dieksplorasi sebagai sebuah kesatuan sistem siap pakai.

Berdasarkan kajian terhadap penelitian terdahulu, pemilihan algoritma klasifikasi yang teruji menjadi landasan krusial. Pada klasifikasi teks, metode ekstraksi fitur TF-IDF sering kali menghasilkan matriks data berdimensi sangat tinggi dan *sparse*. Dalam kondisi tersebut, algoritma *Logistic Regression* menjadi pilihan yang sangat ideal karena efisiensi komputasi serta kemampuannya dalam menangani data berdimensi tinggi tanpa meningkatkan risiko *overfitting* (Y. Pratama et al., 2023). Beberapa studi menunjukkan bahwa *Logistic Regression* andal dalam menganalisis ulasan, termasuk pada sektor penerbangan. Berdasarkan penelitian sebelumnya, Rahmawati & Fitriani (2023) menunjukkan bahwa algoritma ini terbukti tangguh dengan mencatatkan akurasi sebesar 82% dan *F1-Score* 80% dalam mengklasifikasikan opini penumpang maskapai Lion Air, serta menunjukkan performa yang lebih baik dibandingkan *Naïve Bayes* dan *Random Forest* yang memperoleh akurasi di bawah 50%. Berdasarkan hasil penelitian sebelumnya, penelitian ini menitikberatkan pada pemanfaatan *Logistic Regression* sebagai algoritma pengklasifikasi terbaik yang akan dikembangkan secara penuh kinerjanya, alih-alih berfokus pada komparasi performa antar berbagai algoritma.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan mengembangkan *pipeline* analisis sentimen yang dinamis dalam mengevaluasi ulasan pada Google Play Store. Sistem ini memanfaatkan *Logistic Regression* sebagai inti metode klasifikasi untuk membedakan sentimen ulasan secara otomatis menjadi polaritas positif dan negatif. Agar hasil klasifikasi dan evaluasi dapat dimanfaatkan secara praktis, sistem ini diimplementasikan menjadi sebuah *dashboard* analitik interaktif berbasis *framework* streamlit. *Dashboard* ini akan menyajikan antarmuka *visual* seperti grafik perbandingan data sentimen positif dan negatif, *word cloud* kata terbanyak, serta wawasan (*insight*), seperti persentase keluhan pengguna pada versi aplikasi tertentu. Dengan demikian, luaran penelitian ini difokuskan sebagai alat bantu pemantauan analitik yang merangkum tren opini pengguna secara faktual dan dinamis.

2. METODE

Penelitian ini dilaksanakan melalui tahapan yang terstruktur, diawali dari tahap pengumpulan data ulasan secara dinamis, dilanjutkan dengan pembersihan data dan pelabelan sentimen. Selanjutnya, dilakukan *preprocecssing* data serta pemodelan menggunakan algoritma *Logistic Regression*. Kinerja model kemudian dievaluasi sebelum diimplementasikan ke dalam *dashboard* berbasis *website*. Seluruh tahap dikerjakan secara berurutan untuk memastikan alur penelitian yang terstruktur, mulai dari pengolahan data mentah Google Play Store hingga penerapan model dalam sistem interaktif. Seperti yang dapat dilihat secara lengkap pada gambar 1.



Gambar 1. Diagram Alur Penelitian

2.1 Pengumpulan Data

Tahap pengumpulan data bertujuan untuk memperoleh ulasan pengguna yang berasal dari Google Play Store, nantinya digunakan sebagai data pelatihan dan pengujian model. Mengacu pada penelitian sebelumnya terkait analisis ulasan aplikasi publik, proses pengambilan data dilakukan menggunakan teknik *scraping* (Kaburuan & Setiawan, 2023). Pada penelitian ini, pengambilan data dilakukan secara dinamis (*dynamic scraping*) memanfaatkan *library google-play-scraper* pada bahasa pemrograman Python. Sistem dirancang untuk menerima input berupa *Application ID* (id aplikasi) dari pengguna. Dari proses ini, data mentah yang diperoleh biasanya masih mengandung berbagai *noise* atau teks tidak standar, seperti komentar kosong, emoji, simbol, hingga bahasa slang.

2.2 Labeling

Tahap pelabelan bertujuan untuk menetapkan kategori sentimen pada data ulasan. Data yang telah diberi label ini merupakan bagian krusial dalam pendekatan *supervised learning* karena digunakan sebagai dasar dalam pelatihan model klasifikasi (Putri & Intan, 2025). Selain itu, (Zhdanovskaya et al., 2024) menekankan pentingnya pemahaman mendalam terhadap proses pelabelan sebagai kompetensi dasar dalam analisis sentimen, termasuk penerapan teknik pengendalian kualitas data.

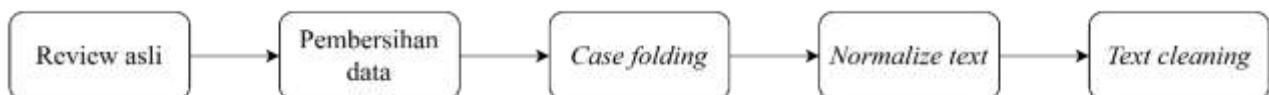
Pada penelitian ini, proses pemberian label dilakukan secara manual dengan menelaah setiap review. Penentuan kelas sentimen dikelompokkan ke dalam dua kategori, yaitu positif dengan label 1 serta negatif dengan label 0. Penelitian ini tidak menyertakan kategori sentimen netral karena ulasan yang bersifat netral umumnya tidak memberikan wawasan evaluasi yang spesifik dan konstruktif bagi pengembang. Selain itu, penghapusan kelas netral bertujuan untuk mengurangi ambiguitas (*noise*) pada data latih, sehingga algoritma *Logistic Regression* dapat mempelajari perbedaan antara sentimen positif dan negatif dengan lebih efektif serta menghasilkan proses klasifikasi yang optimal.

2.3 Preprocessing Data

Data ulasan selanjutnya melalui tahap *scraping* dan melalui proses pelabelan sentimen, data tersebut selanjutnya diproses pada tahap *preprocessing* untuk meningkatkan kualitas serta konsistensi teks sebelum dilakukan ekstraksi fitur dan klasifikasi (Fahri & Gunawan, 2025). Tahap ini diawali dengan pembersihan data (*data cleaning*) untuk menyeleksi kelayakan dataset, yang mencakup penghapusan data duplikat (*drop duplicates*), penghilangan data kosong (*missing values*), serta penyaringan ulasan yang tidak memiliki makna yang jelas (Afif et al., 2025).

Selanjutnya, dilakukan proses normalisasi dan transformasi teks untuk mengatasi variasi

penggunaan bahasa Indonesia yang tidak baku. Proses ini meliputi *case folding* untuk menyeragamkan teks dengan mengubah semua huruf menjadi huruf kecil, *normalize text* bertujuan menyeragamkan penggunaan bahasa dengan mengubah istilah tidak baku, singkatan, serta bahasa sehari-hari ke dalam kata yang lebih baku, dan terakhir *text cleaning* untuk menghilangkan tautan (URL), angka, tanda baca, dan emoji. Melalui rangkaian tahapan ini, data teks diubah menjadi bentuk yang lebih terstruktur sehingga model dapat mengekstraksi fitur kata dengan lebih efektif (Afif et al., 2025). Alur *preprocessing* data ditunjukkan pada diagram gambar 2.



Gambar 2. Alur *Preprocessing*

2.4 Pemodelan

Pada tahap pemodelan, dilakukan pelatihan model klasifikasi teks menggunakan algoritma *Logistic Regression*. Data ulasan yang telah diolah pada tahap *preprocessing* kemudian teks direpresentasikan sebagai vektor numerik melalui metode pembobotan TF-IDF guna mengekstraksi bobot kata-kata yang paling representatif pada teks ulasan (Y. Pratama et al., 2023). Dataset kemudian dipisahkan menjadi data latih dan data uji dengan proporsi pembagian sebesar 90% untuk pelatihan model serta 10% data untuk pengujian model. Melalui tahap ini, diperoleh model klasifikasi biner yang akan dimanfaatkan untuk mengidentifikasi polaritas sentimen ulasan, di mana ulasan dengan sentimen positif direpresentasikan sebagai label 1 dan sentimen negatif direpresentasikan sebagai label 0. Selain menghasilkan prediksi label kelas sentimen. Model terbaik yang telah dilatih selanjutnya diekspor dan disimpan dalam format *.joblib* atau *.pkl* agar dapat dipanggil dan diintegrasikan secara langsung ke dalam *dashboard* analitik berbasis *website* menggunakan *streamlit*.

2.5 Evaluasi Model

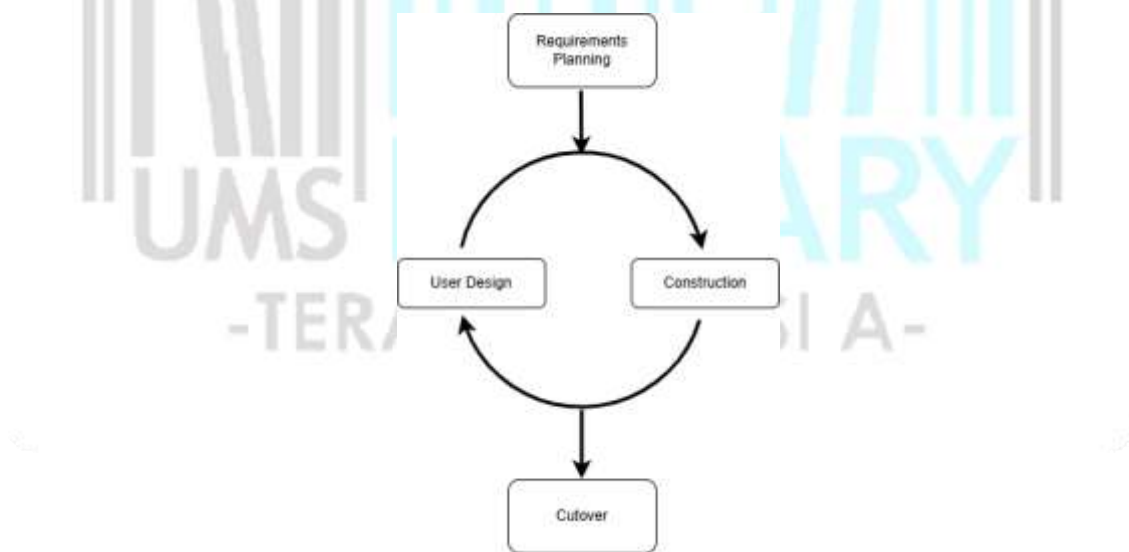
Evaluasi model ini dilakukan untuk menilai performa model *Logistic Regression* yang telah dilatih sebelumnya. Pengujian ini difokuskan pada kemampuan model dalam mengklasifikasikan ulasan ke dalam sentimen positif maupun negatif secara otomatis (Chandra, 2025). Evaluasi dilakukan memanfaatkan 10% data uji yang telah dipisahkan sebelumnya. Selain itu, teknik *cross-validation* (validasi silang) turut diterapkan untuk memperoleh hasil evaluasi yang lebih konsisten serta mengurangi potensi bias akibat pembagian data tertentu.

Kinerja model dianalisis menggunakan *confusion matrix*, yang menyajikan detail jumlah prediksi benar dan salah untuk setiap kategori kelas sentimen. Berdasarkan matriks tersebut, Diperoleh matriks evaluasi utama berupa akurasi, *precision*, *recall*, dan *F1-Score* (Amalia et al., 2024). Keempat matriks ini menggambarkan performa model secara menyeluruh, khususnya terkait

hal ketepatan, keseimbangan prediksi, serta kemampuannya dalam menggeneralisasi data ulasan baru pada kondisi nyata.

2.6 Implementasi Website

Implementasi *website* dalam penelitian ini bertujuan untuk menggabungkan model klasifikasi ke dalam sistem berbasis web, sehingga hasil penelitian dapat diakses dan dimanfaatkan secara langsung oleh pengguna akhir. Proses pengembangan perangkat lunak dilakukan berdasarkan kerangka *System Development Life Cycle* (SDLC) dengan menerapkan metode *Rapid Application Development* (RAD). Pada dasarnya metode RAD dipilih karena karakteristiknya yang menekankan efisiensi waktu pengerjaan sehingga proses pengembangan lebih cepat diselesaikan (Santoso & Amanullah, 2022), serta sifatnya yang iteratif atau dilakukan secara berulang (E. A. Pratama et al., 2023). Pendekatan ini dipilih karena kemampuannya dalam mempercepat pengembangan sistem dengan berfokus pada pembuatan sistem yang fungsional (Ramadhan et al., 2024), di mana integrasi model *machine learning* dan perancangan antarmuka dapat dilakukan secara simultan melalui siklus *prototyping* agar menghasilkan aplikasi yang memiliki tingkat stabilitas yang baik dan memenuhi kebutuhan analisis pengguna. Ilustrasi tahapan metode *Rapid Application Development* ditunjukkan pada gambar 3.



Gambar 3. Model *Rapid Application Development*

2.7.1 Requirements Planning

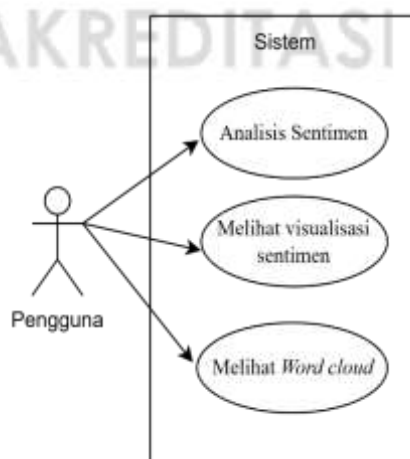
Pada tahap ini, pengembang mengidentifikasi tujuan aplikasi serta menjabarkan secara rinci fungsionalitas dan kebutuhan sistem yang akan dibangun (Santoso & Amanullah, 2022). Selain itu, pada tahap ini juga ditentukan kebutuhan perangkat lunak dan platform pengembangan yang mencakup sistem operasi, tools pemrograman, *framework*, serta *database* yang digunakan dalam implementasi sistem.

Dalam penelitian ini, sistem dikembangkan menggunakan Python sebagai bahasa pemrograman dan memanfaatkan *library scikit-learn* untuk melakukan klasifikasi sentimen menggunakan algoritma *Logistic Regression*. Proses pengolahan teks dilakukan melalui tahap *preprocessing* dan pembobotan kata yang menerapkan metode TF-IDF (*Term Frequency–Inverse Document Frequency*). Antarmuka sistem dibangun dengan *framework streamlit* sehingga memungkinkan pengguna untuk berinteraksi secara langsung melalui *browser*. Dengan pendekatan tersebut, sistem diharapkan mampu berjalan secara efisien dalam melakukan analisis sentimen terhadap data ulasan yang diperoleh secara dinamis.

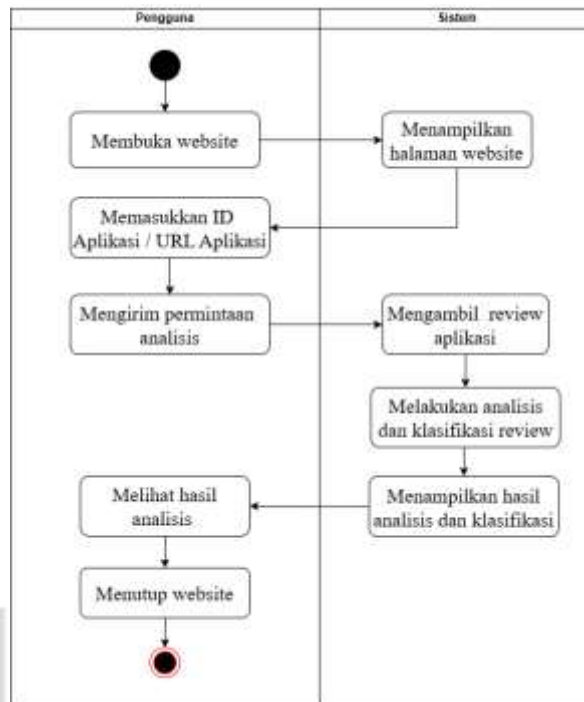
2.7.2 User Design

Setelah kebutuhan sistem teridentifikasi, tahap berikutnya adalah *user design* yang berfokus pada penerjemahan kebutuhan ke dalam rancangan visual dan arsitektur perangkat lunak. Proses ini bersifat iteratif, di mana *prototype* antarmuka dibuat, dievaluasi bersama pengguna, lalu diperbaiki berulang hingga sesuai kebutuhan (Iqbal, 2023). Hasil perancangan pada tahap *user design* kemudian divisualisasikan melalui gambar 4 berupa *use case diagram* yang menggambarkan hubungan antara aktor dengan sistem, serta gambar 5 berupa *activity diagram* yang menjelaskan alur kerja serta aktivitas yang terjadi dalam sistem.

Pada fase ini, dirancang sebuah *dashboard* analitik interaktif yang menampilkan fitur-fitur seperti kolom *input ID* aplikasi atau URL aplikasi, grafik persentase sentimen, dan visualisasi *word cloud*. Perancangan sistem dilakukan dengan mempertimbangkan aspek fitur, fungsionalitas, serta kebutuhan pengguna dalam pengembangan aplikasi berbasis web, sehingga hasil analisis dapat memberikan wawasan yang lebih mudah dipahami oleh pengguna (Al Irsyadi et al., 2025).



Gambar 4. *Use case diagram*



Gambar 5. Activity diagram

2.7.3 Construction

Tahap *construction* (konstruksi) adalah proses pengkodean (*coding*) di mana desain *visual* dan alur logika telah dirancang diimplementasikan menjadi aplikasi web yang fungsional (Iqbal, 2023). Fokus utama pada fase ini adalah membangun aplikasi web dengan mengintegrasikan model yang telah dilatih ke dalam sistem.

Hal ini sejalan dengan penelitian analisis sentimen sebelumnya, di mana pengembangan aplikasi web dilakukan sebagai implementasi langsung dari model klasifikasi yang telah dibangun dan dievaluasi, sehingga aplikasi dapat menerima *input* teks dari pengguna dan memberikan hasil prediksi secara instan (Fahri & Gunawan, 2025). Pendekatan integrasi serupa juga diterapkan dalam deteksi citra medis, di mana model *deep learning* diintegrasikan ke dalam aplikasi berbasis web menggunakan *Flask* sehingga sistem mampu menerima data baru dan menampilkan hasil klasifikasi secara *real-time* (Maheswari et al., 2025).

Dalam penelitian ini, seluruh tahapan pemrosesan data diintegrasikan ke dalam satu kesatuan sistem *back-end* yang terstruktur. Proses tersebut mencakup pengambilan data secara *dynamic scraping*, dilanjutkan dengan tahap *preprocessing* yang meliputi *case folding*, *normalize text*, *text cleaning*. Selanjutnya, dilakukan proses pembobotan kata menggunakan TF-IDF, lalu pada tahap akhir sentiment diklasifikasikan menggunakan model *Logistic Regression*. Dengan integrasi ini, sistem mampu mengeksekusi keseluruhan *pipeline* analisis secara otomatis dan memberikan hasil prediksi sentimen secara *real-time* berdasarkan *input* pengguna melalui antarmuka web.

2.7.4 Cutover

Tahap *Cutover* (peralihan) merupakan fase akhir dalam metode RAD yang mencakup pengujian akhir (*testing*), evaluasi, dan peluncuran (*deployment*) sistem ke lingkungan produksi. Pada fase ini, fungsionalitas sistem diuji secara menyeluruh guna memastikan *pipeline* analisis sentimen dapat berfungsi dengan baik dalam memproses ulasan secara dinamis tanpa adanya kesalahan (*bug*). Salah satu metode pengujian yang diterapkan adalah *black box*, yang digunakan untuk memeriksa kemampuan sistem dalam memproses *input* serta menghasilkan *output* yang tepat antara *frontend* dan *backend* (Ramadhan et al., 2024).

Selain itu, dilakukan untuk menilai kemudahan penggunaan dan kepuasan pengguna, sistem dapat dievaluasi menggunakan metode seperti *User Acceptance Testing* (UAT) atau *System Usability Scale* (SUS) (Iqbal, 2023). Dengan melalui tahapan pengujian dan evaluasi tersebut, sistem yang telah dikembangkan diharapkan berada dalam kondisi stabil, fungsional, dan siap untuk diimplementasikan serta digunakan secara praktis oleh pengguna.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pengumpulan data dilakukan menggunakan teknik *scraping* dengan memanfaatkan *library google-play-scraper* yang dijalankan pada google colab. Data diperoleh dari 20 aplikasi yang tersedia di Google Play Store dan dikelompokkan ke dalam 9 kategori label. Dari setiap aplikasi, diambil 1000 ulasan dengan 200 ulasan di setiap bintangnya. Hasil proses pengumpulan data tersebut disajikan pada tabel 1.

Tabel 1. Hasil proses pengumpulan data

Label	Aplikasi	Jumlah Data
Transportasi	Gojek - Transportasi & Makanan	1000
	Grab - Taxi & Food Delivery	1000
Keuangan	DANA Dompet Digital Indonesia	1000
	Livin' by Mandiri	1000
	Bank Jago/Jago Syariah	1000
E_Commerce	Shopee Indonesia	1000
	Tokopedia	1000
Layanan_Publik	PLN Mobile	1000
	MyTelkomsel – Beli Pulsa/Kuota	1000
	Access by KAI	1000

Label	Aplikasi	Jumlah Data
Game	Roblox	1000
	Mobile Legends: Bang Bang	1000
Streaming	Netflix	1000
	Youtube	1000
Sosial	TikTok	1000
	Instagram	1000
Ai	Google Gemini	1000
	ChatGPT	1000
Musik	Spotify: Music dan Podcast	1000
	SoundCloud - musik & audio	1000
Total		20.000

3.2 Labeling

Proses pelabelan dilakukan secara manual untuk mengidentifikasi sentimen yang terkandung dalam setiap ulasan. Hasil pelabelan ini berperan sebagai data latih pada tahap pengembangan model dan digunakan sebagai tolok ukur evaluasi terhadap performa prediksi model. Pada tabel 2 di tampilan hasil dari proses pelabelan manual.

Tabel 2. Hasil *labelling*

Kategori Sentimen	Jumlah Data
Positif	3489
Negatif	16511
Total	20000

3.3 Preprocessing Data

Proses *preprocessing* dimulai dengan pembersihan data (*data cleaning*) yang dilakukan untuk menjamin kualitas dan integritas dataset sebelum diproses lebih lanjut. Dari total 20.000 ulasan mentah yang digunakan, hasil dari pemeriksaan tidak ditemukan adanya data duplikat maupun nilai kosong (*missing values/NaN*). Dengan demikian, seluruh data dapat dipertahankan dan digunakan pada tahap berikutnya, sehingga jumlah data tetap sebanyak 20.000 ulasan yang terdiri atas 3.489 ulasan positif dan 16.511 ulasan negatif.

Tahapan dilanjutkan dengan menyeragamkan format teks dan membersihkan data dari *noise* sebelum masuk ke tahap ekstraksi fitur. Transformasi teks pada dataset dilakukan secara bertahap sesuai dengan alur yang telah dirancang agar menghasilkan representasi teks ulasan aplikasi yang lebih konsisten. Contoh perubahan teks ulasan disajikan pada tabel 3.

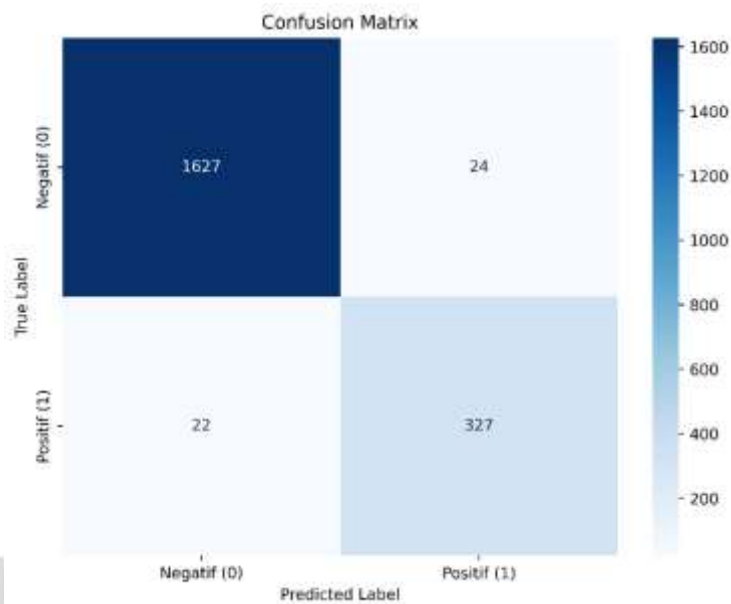
Tabel 3. Hasil *preprocessing* ulasan aplikasi

Tahapan	Hasil Teks Ulasan
Review asli	Awalnya bagus, tapi lama2 pas update malah jadi error & lemot parah gak bagus
<i>Case folding</i>	awalnya bagus, tapi lama2 pas update malah jadi error & lemot parah gak bagus
<i>Normalize text</i>	awalnya bagus, tapi lama2 pas update malah jadi error & lambat parah tidak bagus
<i>Text cleaning</i>	awalnya bagus tapi lama lama pas update malah jadi error lambat parah tidak bagus

tahap *preprocessing* ini memberikan dampak yang signifikan terhadap penyederhanaan variasi kata. Meskipun tidak terjadi pengurangan jumlah data, *preprocessing* berhasil menurunkan jumlah kosakata unik dari 60.086 kata menjadi 24.703 kata. Penurunan jumlah kosakata ini menunjukkan bahwa variasi kata yang bersifat *redundant*, seperti perbedaan huruf kapital, tanda baca, serta kata tidak baku berhasil dinormalisasi. Dengan demikian, representasi teks menjadi lebih konsisten, ringkas, dan lebih representatif untuk tahap klasifikasi.

3.4 Pemodelan dan Evaluasi Model

Penelitian ini menggunakan algoritma *Logistic Regression* dengan pembobotan fitur TF-IDF untuk membangun model klasifikasi sentimen. Dataset hasil *preprocessing* yang berjumlah 20.000 ulasan dibagi menjadi data latih dan data uji dengan rasio 90:10, sehingga diperoleh 18.000 ulasan sebagai data latih dan 2.000 ulasan sebagai data uji. Data latih digunakan dalam proses pelatihan model, sedangkan data uji digunakan untuk menguji performa model dalam melakukan klasifikasi sentimen pada data baru yang bersifat baru. Pada tahap prediksi, digunakan nilai *classification threshold* sebesar 0,6 sebagai batas penentuan kelas sentimen. Penggunaan *classification threshold* sebesar 0,6 dipilih guna menyeimbangkan nilai *precision* dan *recall* pada kelas positif, mengingat distribusi data yang tidak seimbang antara kelas positif dan negatif. Kinerja model kemudian dievaluasi menggunakan *confusion matrix*, sebagaimana yang ditunjukkan pada gambar 6.



Gambar 6. *Confusion matrix*

Berdasarkan hasil evaluasi confusion matrix, model mampu mengklasifikasikan 1.627 data *true negative* (TN) dan 327 data *true positive* (TP), sementara kesalahan klasifikasi hanya 24 data *false positive* (FP) dan 22 data *false negative* (FN). Tingginya nilai TN dan TP menunjukkan bahwa sebagian besar ulasan berhasil diklasifikasikan dengan tepat, sedangkan rendahnya nilai FP dan FN mengindikasikan tingkat kesalahan yang rendah. Hasil tersebut menunjukkan bahwa model memiliki performa yang baik dalam membedakan sentimen positif dan negatif.

Hasil evaluasi secara keseluruhan menunjukkan bahwa model memiliki performa yang sangat baik. Dengan penyesuaian *threshold* prediksi sebesar 0,6, model mampu mengklasifikasikan sentimen positif dan negatif dengan akurasi yang tinggi sebesar 97,70%. Keseimbangan antara *precision* dan *recall* pada kedua kelas tercermin dari nilai *F1-score* yang diperoleh, menunjukkan bahwa model memiliki kemampuan klasifikasi yang stabil dan dapat diandalkan. Ringkasan hasil evaluasi kinerja model disajikan pada tabel 4.

Tabel 4. Hasil evaluasi performa model

Kelas	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Negatif (0)	0.9770	0.9867	0.9855	0.9861
Positif (1)		0.9316	0.9370	0.9343

Agar performa model tidak hanya bergantung pada satu pembagian dataset, dilakukan evaluasi lanjutan menggunakan metode *Repeated Stratified K-Fold Cross-Validation* dengan 3 fold dan 5 repetisi ($n_splits = 3$ dan $n_repeats = 5$), sehingga diperoleh total 15 iterasi pengujian. Hasil evaluasi

menunjukkan nilai *F1-score* rata-rata sebesar 89,47% dengan standar deviasi yang relatif rendah, yaitu sebesar 0,57%. Selain itu, nilai *F1-score* pada setiap iterasi berada pada rentang 88,36% hingga 90,66%, yang menunjukkan bahwa model memiliki tingkat konsisten performa pada berbagai split data. Dengan demikian, model dapat dikatakan memiliki tingkat stabilitas dan mampu melakukan generalisasi dengan baik tanpa adanya indikasi *overfitting*. Ringkasan hasil pengujian *cross-validation* ditampilkan pada tabel 5.

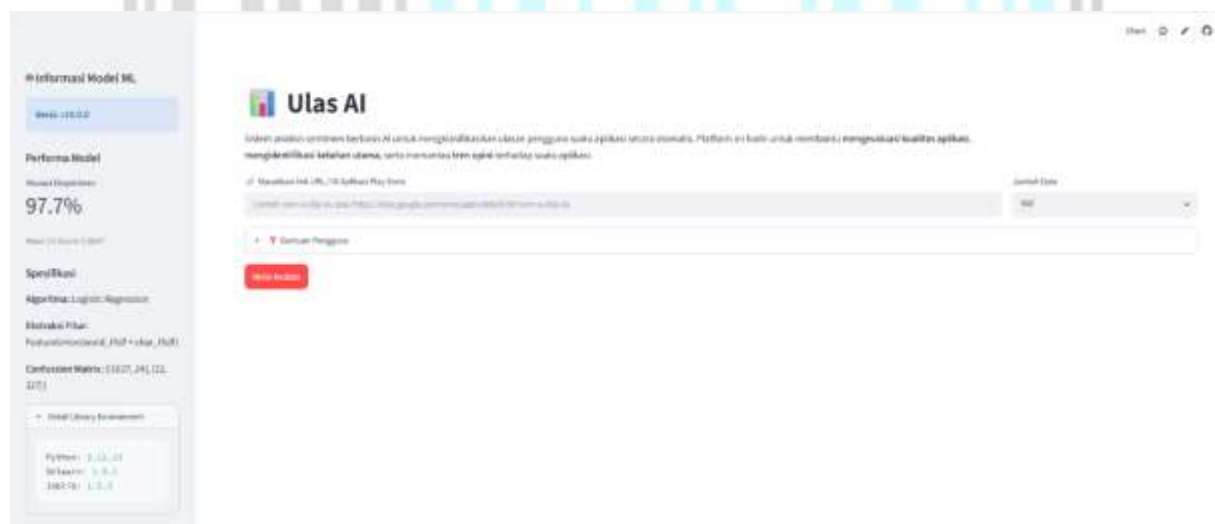
Tabel 5. Hasil pengujian *cross-validation* model (15 iterasi)

Metrik	Nilai terendah (<i>min</i>)	Nilai tertinggi (<i>max</i>)	Rata-rata	Standar deviasi sampel
<i>F1-score</i>	0.8836	0.9066	0.8947	0.0057

3.5 Implementasi Website (Construction)

3.5.1 Tampilan Halaman Utama

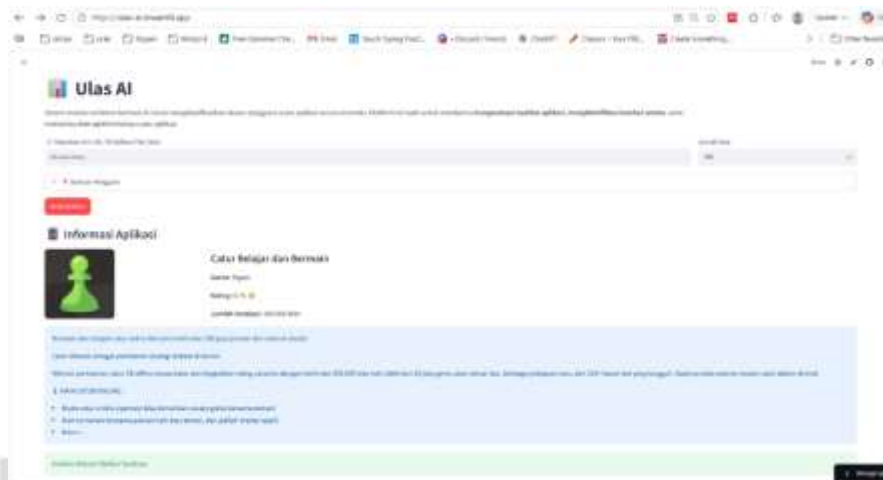
Halaman utama merupakan antarmuka yang digunakan pengguna untuk berinteraksi dengan sistem. Pada halaman ini pengguna dapat memasukkan tautan atau ID aplikasi yang ada di Google Play Store sebagai *input* untuk memulai proses analisis ulasan. Tampilan halaman utama *website* ditampilkan pada gambar 7.



Gambar 7. Tampilan halaman utama

3.5.2 Tampilan Hasil Klasifikasi Ulasan

Setelah proses analisis dijalankan, sistem menampilkan hasil klasifikasi ulasan dalam bentuk grafik persentase sentiment, grafik distribusi probabilitas, grafik persentase keluhan berdasarkan versi aplikasi, serta *word cloud*. Tampilan hasil klasifikasi ulasan ditunjukkan pada gambar berikut.



Gambar 8. Pengenalan singkat aplikasi



Gambar 9. Grafik persentase sentimen dan distribusi probabilitas



Gambar 10. Grafik persentase keluhan pada setiap versi aplikasi

Kata yang Paling Sering Muncul (Wordcloud)



Gambar 11. Word cloud kata

Top 10 Kata Dominan

Sentimen Positif

Term	Review
sangat bagus	24
bagus banget	13
game ini	14
game bagus	8
yang sangat	8
bagus untuk	8
saya suka	5
aplikasi yang	8
gamenya bagus	5
ini sangat	5

Sentimen Negatif

Term	Review
tidak bisa	19
tiba tiba	15
akun saya	10
game nya	7
bisa main	6
saya ingin	6
tidak ada	7
saya tidak	7
nya tidak	5
tidak jelas	6

Gambar 12. 10 kata teratas yang sering muncul

5 Ulasan Paling Representatif

Sentimen Positif

- Kepercayaan Model: 100.00%**
sangat membantu untuk banyak banyak belajar dan sangat memuaskan, termasuk banyak, dan semoga sukses selalu...
- Kepercayaan Model: 100.00%**
pemerian yg sangat bagus dan mengasah kemampuan, aku suka aku suka 🥳
- Kepercayaan Model: 100.00%**
sangat bagus untuk mengasah otak
- Kepercayaan Model: 91.99%**
game sangat baik bagus sekali
- Kepercayaan Model: 91.99%**
gamenya bagus banget bisanya aku pm bagus jadi aku punya teman main catur dimana bagus apalagi update terus 🥳🥳

Sentimen Negatif

- Kepercayaan Model: 100.00%**
game gk berguna, saat mau login ke akun bar ada basun "waktu habis" padahal simpel banget, game gampang, udah di korak korak 1 jam masih gk bisa login, kudu bisa bingung di perbaiki
- Kepercayaan Model: 100.00%**
kenapa ga bisa masuk ini, game tiba tiba di hentikan terus keluar dengan sendirinya bingung di perbaiki
- Kepercayaan Model: 100.00%**
gak masuknya saya ketika bermain waktu masuk telah habis tapi tiba tiba waktu saya yg malah habis jadinya kalah padahal jaringan saya sudah lancar
- Kepercayaan Model: 100.00%**
gak jelas tiba tiba di ban...
- Kepercayaan Model: 100.00%**
aplikasi nya seperti nya mengklam bug, karena ketika saya bermain streak bermain saya tidak menambah, dan beranda nya menampilkan aku saya tidak menambah ataupun mengurang, padahal saya kalah dan menang, dan juga ketika saya bermain 5 menit malah gangguan, spaly di mode lain, padahal jaringan sy tidak kenapa, mohon diperbaiki

Gambar 13. 5 ulasan paling representatif

3.6 Pengujian (Cutover)

3.6.1 Pengujian Black Box

Pada penelitian ini menerapkan metode pengujian *black box* untuk memastikan dan memverifikasi

setiap fitur pada sistem telah berfungsi sesuai standar kebutuhan fungsional yang ditentukan. Pengujian dilakukan pada seluruh fitur, mulai dari pengambilan data, klasifikasi sentimen, hingga visualisasi data dalam bentuk grafik dan word cloud. Hasil mengenai pengujian *black box* ini disajikan pada tabel 6.

Tabel 6. Hasil Pengujian *Black Box*

No.	Fitur yang Diuji	Skenario Pengujian	Hasil yang Diharapkan	Hasil Pengujian	Status
1.	Akses Halaman Utama	Pengguna membuka tautan aplikasi melalui <i>web browser</i> .	Sistem berhasil memuat antarmuka utama secara responsif	Halaman utama tampil sempurna.	Sesuai
2.	Analisis Sentimen (Input Valid)	Memasukkan ID aplikasi yang valid (contoh: com.duolingo) lalu menekan tombol analisis.	Sistem menampilkan pengenalan singkat aplikasi, menampilkan pesan notifikasi dan merujuk ke halaman hasil.	Data ulasan berhasil ditarik, sistem menampilkan pengenalan singkat aplikasi, validasi: "Analisis Selesai! Berikut hasilnya:", dan hasil klasifikasi	Sesuai
3.	Validasi Input Kosong	Mengosongkan kolom input ID aplikasi Play Store lalu langsung menekan tombol analisis.	Sistem membatalkan penarikan data dan menampilkan pesan peringatan form wajib diisi.	Muncul notifikasi validasi: "Jangan lupa masukkan Link atau ID Aplikasinya ya!".	Sesuai
4.	Validasi Input (Salah Format)	Memasukkan teks acak atau string yang bukan format ID Play Store yang sah.	Sistem mendeteksi format tidak valid dan memunculkan pesan kesalahan format.	Muncul notifikasi validasi format ID tidak sesuai	Sesuai

No.	Fitur yang Diuji	Skenario Pengujian	Hasil yang Diharapkan	Hasil Pengujian	Status
5	Visualisasi Distribusi Sentimen	Mengakses komponen visualisasi data sentimen setelah proses pada butir uji nomor 2 selesai.	Sistem menampilkan grafik yang memuat persentase data hasil klasifikasi.	Grafik representasi sentimen berhasil dimuat dengan pembagian data yang jelas.	Sesuai
6	Visualisasi Word Cloud	Memeriksa komponen visualisasi kata kunci (<i>Word Cloud</i>) pada antarmuka halaman hasil.	Sistem merender gambar <i>Word Cloud</i> berisi kata-kata yang paling sering muncul dari kumpulan teks ulasan.	Komponen Word Cloud berhasil di <i>generate</i> dan divisualisasikan dengan rapi.	Sesuai

3.6.2 Pengujian Usability

Pengujian *Usability* dilakukan untuk mengukur sejauh mana sistem dapat digunakan dengan mudah oleh pengguna. Sebanyak 42 responden berpartisipasi dengan mencoba seluruh fitur sistem, mulai dari *input* URL hingga peninjauan hasil klasifikasi. Responden diminta mengisi kuesioner *System Usability Scale* (SUS) setelah menggunakan sistem sebagai bentuk evaluasi. Hasil perhitungan skor setiap responden disajikan pada tabel 7.

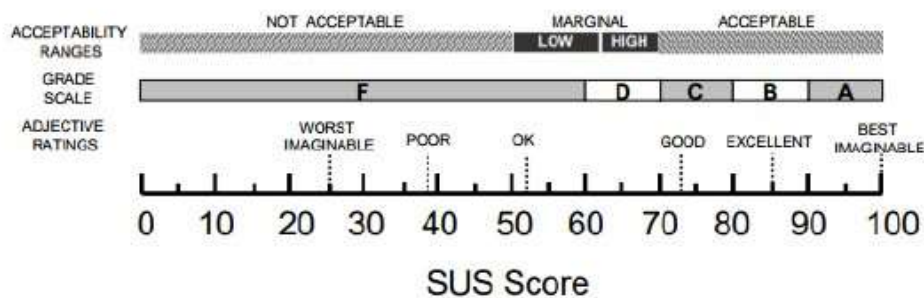
Tabel 7. Hasil perhitungan skor SUS

Responden	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Jumlah Skor	SUS Skor (x2.5)
R1	5	1	5	2	4	2	5	4	5	3	32	80
R2	5	1	4	1	4	2	5	1	5	3	35	87,5
R3	5	1	5	1	5	2	4	2	4	2	35	87,5
R4	5	2	4	2	5	1	5	2	5	3	34	85
R5	5	2	4	2	5	1	5	1	5	1	37	92,5
R6	5	1	5	1	5	1	5	1	4	2	38	95
R7	5	1	5	2	5	1	4	2	4	1	36	90
R8	5	1	5	2	5	1	5	1	4	1	38	95

Responden	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Jumlah Skor	SUS Skor (x2.5)
R9	4	3	4	2	4	1	4	2	4	3	29	72,5
R10	5	1	5	2	5	2	3	2	5	2	34	85
R11	5	1	5	1	5	1	5	1	5	1	40	100
R12	4	1	5	1	5	1	3	3	5	1	35	87,5
R13	5	1	5	2	5	1	5	1	4	1	38	95
R14	5	1	5	1	5	1	5	1	5	1	40	100
R15	5	3	4	5	5	1	1	3	5	5	23	57,5
R16	5	1	5	1	5	1	5	1	5	1	40	100
R17	5	2	5	1	5	1	5	1	5	1	39	97,5
R18	5	1	5	1	5	1	5	2	5	2	38	95
R19	5	2	4	1	5	1	4	1	4	3	34	85
R20	4	2	3	2	5	1	5	1	5	4	32	80
R21	4	1	3	4	3	3	2	1	5	4	24	60
R22	4	2	5	1	3	3	4	2	4	2	30	75
R23	4	2	5	1	4	2	5	2	5	1	35	87,5
R24	4	2	4	2	4	2	4	2	4	2	30	75
R25	4	1	5	3	5	1	5	2	4	2	34	85
R26	5	1	5	1	5	1	5	1	5	1	40	100
R27	4	1	5	1	5	1	5	1	5	1	39	97,5
R28	5	2	5	3	4	1	5	1	3	2	33	82,5
R29	4	2	5	2	4	1	4	2	4	2	32	80
R30	4	1	4	1	5	1	5	1	5	1	38	95
R31	4	2	4	1	4	3	4	1	4	2	31	77,5
R32	4	1	5	1	5	2	5	1	5	1	38	95
R33	5	2	5	2	5	3	5	2	5	3	33	82,5
R34	4	3	4	2	4	2	3	2	5	1	30	75
R35	5	1	5	2	5	2	5	1	5	2	37	92,5
R36	5	1	5	2	5	2	4	1	4	2	35	87,5
R37	4	2	5	1	5	2	4	2	4	2	33	82,5
R38	5	2	5	1	4	2	4	1	4	2	34	85

Responden	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Jumlah Skor	SUS Skor (x2.5)
R39	2	5	1	5	1	4	2	5	1	5	3	7,5
R40	4	1	5	1	4	2	4	1	5	1	36	90
R41	4	2	4	2	5	1	4	2	4	2	32	80
R42	4	2	4	3	4	2	4	2	4	2	29	72,5
Rata-rata											84,11	

Berdasarkan hasil pengujian SUS pada tabel 7, diperoleh rata-rata akhir sebesar 84,11. Mengacu pada pedoman penilaian SUS yang ada di gambar 14, skor tersebut masuk kedalam kategori *Good* dengan *Grade B*. Selain itu, nilai tersebut berada pada rentang *Acceptable*, yang mengindikasikan bahwa sistem memiliki tingkat *usability* yang baik. Dengan hasil tersebut, aplikasi yang dikembangkan dapat dikatakan telah memenuhi standar kelayakan *usability* serta mampu memberikan pengalaman pengguna yang memuaskan.



Gambar 14. Penilaian *System Usability Scale*

4. PENUTUP

Penelitian ini menyimpulkan bahwa integrasi model *Logistic Regression* dengan pembobotan fitur TF-IDF ke dalam sistem berbasis *website* terbukti efektif dalam melakukan klasifikasi sentimen ulasan aplikasi Google Play Store secara otomatis ke dalam kelas positif dan negatif. Mengacu pada tujuan penelitian, sistem yang dikembangkan mampu menyajikan hasil analisis sentimen melalui antarmuka yang intuitif dan mudah dipahami pengguna. Lebih dari itu, sistem juga menunjukkan performa klasifikasi yang baik, didukung oleh hasil *cross-validation* yang mengindikasikan tingkat stabilitas dan kemampuan generalisasi model yang tinggi. Signifikansi dari pencapaian ini terletak pada kemampuannya dalam membantu proses evaluasi aplikasi berdasarkan persepsi pengguna yang tercermin dalam ulasan, sehingga pengembang aplikasi maupun perusahaan dapat memperoleh informasi yang relevan untuk merumuskan strategi peningkatan kualitas layanan secara lebih cepat dan efisien. Dengan demikian, sistem yang dibangun dapat menjadi solusi praktis dalam mengolah

dan menganalisis ulasan pengguna dalam jumlah besar yang sulit dilakukan secara manual. Meskipun demikian, penelitian ini masih memiliki peluang untuk dikembangkan lebih lanjut melalui penambahan jumlah dan variasi data pelatihan, penerapan metode klasifikasi lain sebagai pembandingan performa, pengembangan fitur analisis yang lebih komprehensif, serta perluasan implementasi sistem pada platform digital atau media sosial lainnya guna meningkatkan ketangguhan dan cakupan penerapan model di masa mendatang.

DAFTAR PUSTAKA

- Afif, R., Supriyanto, A., Fitri Damaryanti, R., & Prasetya Adi, W. (2025). Analisis Sentimen Aplikasi Adiraku di Google Play Store Menggunakan Metode Support Vectore Machine. *JURNAL FASILKOM*, 15(1), 163–171.
- Al Irsyadi, F. Y., Rahman, A. A., & Gunawan, D. (2025). Pengembangan Website Informasi Sekolah di Madrasah Ibtidaiyah Muhammadiyah Plumbon, Eromoko, Wonogiri, Jawa Tengah. *Jurnal Pengabdian Masyarakat Terapan*, 1(3), 137–147. <https://doi.org/10.20884/1.jupiter.1.3.41>
- Amalia, N. A., Utami, I. T., & Wilandari, Y. (2024). ANALISIS SENTIMEN KEBIJAKAN PENYELENGGARA SISTEM ELEKTRONIK LINGKUP PRIVAT MENGGUNAKAN PENALIZED LOGISTIC REGRESSION DAN SUPPORT VECTOR MACHINE. *Jurnal Gaussian*, 12(4), 560–569. <https://doi.org/10.14710/j.gauss.12.4.560-569>
- Andalas, S. fadhilah tegar. (2024, June 12). 209,3 Juta Orang di Indonesia Menggunakan Smartphone pada Tahun 2023 - GoodStats Data. <https://data.goodstats.id/statistic/2093-juta-orang-di-indonesia-menggunakan-smartphone-pada-tahun-2023-cbha0>
- Chandra, R. R. D. (2025). ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI CAREFAST OPERATION DI GOOGLE PLAYSTORE MENGGUNAKAN METODE NAIVE BAYES DAN REGRESI LOGISTIK. <https://eskripsi.usm.ac.id/files/skripsi/G21A/2021/G.231.21.0132/G.231.21.0132-15-File-Komplit-20251231123508.pdf>
- Fahri, M. D. H., & Gunawan, D. (2025). Analisis Sentimen Pengguna X terhadap Perempuan di Lingkungan Kerja Menggunakan Algoritma Machine Learning. *Journal of Technology and Informatics (JoTI)*, 7(2), 134–146. <https://doi.org/10.37802/joti.v7i2.1087>
- He, A., & Abisado, M. (2024). Text Sentiment Analysis of Douban Film Short Comments Based on BERT-CNN-BiLSTM-Att Model. *IEEE Access*, 12, 45229–45237. <https://doi.org/10.1109/ACCESS.2024.3381515>
- Iqbal, D. M. (2023). PERANCANGAN SISTEM INFORMASI MANAJEMEN ADMINISTRASI BADAN USAHA MILIK DESA BERBASIS WEB (Studi Kasus: Bumdes Citeko Kaler Purwakarta). <http://repository.upi.edu>
- Kaburuan, E. R., & Setiawan, N. R. (2023). Sentimen Analisis Review Aplikasi Digital Korlantas Pada Google Play Store Menggunakan Metode SVM. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 12(1), 105–116. <https://doi.org/10.32736/sisfokom.v12i1.1614>
- Kaur, G., & Sharma, A. (2023). A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. *Journal of Big Data*, 10(5). <https://doi.org/10.1186/s40537-022-00680-6>
- Lidinillah, E. R., Rohana, T., & Juwita, A. R. (2023). Analisis sentimen twitter terhadap steam menggunakan algoritma logistic regression dan support vector machine. *TEKNOSAINS: Jurnal Sains, Teknologi Dan Informatika*, 10(2), 154–164. <https://doi.org/10.37373/tekno.v10i2.440>
- Maheswari, S., Gunawan, D., Yani, J. J. A., Pos, T., & Tengah -Indonesia, S.-J. (2025). DETEKSI DINI KANKER KULIT MENGGUNAKAN CNN, DNN, DAN EFFICIENTNET: PENDEKATAN DEEP LEARNING BERBASIS WEB. *RABIT: Jurnal Teknologi Dan Sistem Informasi Univrab*,

10(2), 932–944. <https://doi.org/10.36341/rabit.vvix.xxx>

- Mao, Y., Liu, Q., & Zhang, Y. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. In *Journal of King Saud University - Computer and Information Sciences* (Vol. 36, Number 4). King Saud bin Abdulaziz University. <https://doi.org/10.1016/j.jksuci.2024.102048>
- Pratama, E. A., Krisgianti, S., & Paramita, H. M. (2023). Penerapan Metode Rapid Application Development (RAD) pada Pengembangan Sistem Informasi Inventaris Aset Desa Rempoah. *Infotek: Jurnal Informatika Dan Teknologi*, 6(1), 49–59. <https://doi.org/10.29408/jit.v6i1.7313>
- Pratama, Y., Murdiansyah, D. T., & Lhaksana, K. M. (2023). Analisis Sentimen Kendaraan Listrik Pada Media Sosial Twitter Menggunakan Algoritma Logistic Regression dan Principal Component Analysis. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(1), 529–535. <https://doi.org/10.30865/mib.v7i1.5575>
- Putri, A. A., & Intan, B. (2025). Sentiment Analysis of User Reviews of the Job Search App “Kita Lulus” on the Google Play Store Using Machine Learning. *Bulletin of Information Technology (BIT)*, 6(3), 284–293. <https://doi.org/10.47065/bit.v5i2.2220>
- Rahmawati, I., & Fitriani, T. R. (2023). Analisis Sentimen Menggunakan Algoritma Logistic Regression Pada Penerbangan Lion Air berdasarkan Ulasan Pengguna Platform Online. *Jejaring Penelitian Dan Pengabdian Masyarakat (JPPM)*, 1(1), 11–16.
- Ramadhan, D. D., Mumpuni, R., & Sihananto, A. N. (2024). IMPLEMENTASI METODE RAPID APPLICATION DEVELOPMENT (RAD) DALAM PENGEMBANGAN SISTEM ENTERPRISE INDUSTRI TEKSTIL BERBASIS WEBSITE. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3S1). <https://doi.org/10.23960/jitet.v12i3s1.5222>
- Rullyana, G., & Triandari, R. (2025). Analisis Sentimen Ulasan Pengguna Indonesia terhadap Platform Pembelajaran Digital Ruangguru Menggunakan Algoritma Leksikal Multilingual. *Jurnal Komtika (Komputasi Dan Informatika)*, 9(2), 263–275. <https://doi.org/10.31603/komtika.v9i2.15110>
- Santoso, L., & Amanullah, J. (2022). PENGEMBANGAN SISTEM INFORMASI AKADEMIK BERBASIS WEBSITE MENGGUNAKAN METODE RAPID APPLICATION DEVELOPMENT (RAD). *JURNAL ILMIAH ELEKTRONIKA DAN KOMPUTER*, 15(2), 250–259. <http://journal.stekom.ac.id/index.php/elkom/page250>
- Zhdanovskaya, A., Baidakova, D., & Ustalov, D. (2024). Data Labeling for Machine Learning Engineers: Project-Based Curriculum and Data-Centric Competitions. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(13), 15886–15893. <https://doi.org/10.1609/aaai.v37i13.26886>

