Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 5 No. 2 (2026) pp: 11559-11569
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Sentimen pada Google Play Store Pengguna Aplikasi TikTok dan Instagram Menggunakan Algoritma XGBoost Berbasis Website

Afriyanto, Usman
Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Islam Indragiri
apriyantotbh@gmail.com, usmanovsky13411@gmail.com

Abstrak

Pertumbuhan pengguna media sosial di Indonesia, khususnya TikTok dan Instagram, mendorong meningkatnya jumlah ulasan aplikasi pada Google Play Store yang memuat berbagai opini dan pengalaman pengguna. Banyaknya ulasan tersebut menyulitkan proses analisis secara manual, sehingga diperlukan pendekatan otomatis melalui analisis sentimen. Penelitian ini bertujuan mengembangkan sistem analisis sentimen berbasis website untuk mengklasifikasikan ulasan pengguna aplikasi TikTok dan Instagram ke dalam kategori positif dan negatif menggunakan algoritma Extreme Gradient Boosting (XGBoost). Dataset yang digunakan terdiri dari 8.742 ulasan hasil scraping Google Play Store menggunakan google-play-scraper dan diberi label berdasarkan rating bintang pengguna. Tahap pengolahan data meliputi preprocessing teks, seperti case folding, penghapusan noise, dan stopword removal, kemudian dilakukan ekstraksi fitur menggunakan TF-IDF. Model dievaluasi dengan pembagian data stratified 80:20. Hasil penelitian menunjukkan bahwa model XGBoost mampu mencapai akurasi sebesar 87,59%, dengan nilai presisi 88,10%, recall 87,59%, dan F1-score 87,41%, sehingga menunjukkan kinerja yang baik dalam mengklasifikasikan sentimen ulasan berbahasa Indonesia. Sistem yang dibangun berhasil diimplementasikan menggunakan Flask dan Bootstrap 5 serta dilengkapi fitur prediksi manual, pemrosesan data melalui unggahan file CSV/XLSX, dan penyimpanan riwayat prediksi. Hasil penelitian ini menunjukkan bahwa kombinasi XGBoost dan TF-IDF dapat menjadi solusi yang efektif untuk membantu analisis sentimen ulasan pengguna secara otomatis dan mendukung evaluasi kualitas layanan aplikasi berdasarkan masukan pengguna.

Kata kunci: Analisis Sentimen, XGBoost, Google Play Store, TikTok, Instagram, TF-IDF, Flask

1. Latar Belakang

Perkembangan teknologi informasi yang pesat dalam beberapa tahun terakhir telah mengubah cara masyarakat berkomunikasi dan berinteraksi satu sama lain. Salah satu perubahan yang paling terasa adalah meningkatnya penggunaan media sosial, terutama di Indonesia. Berdasarkan laporan DataReportal, pengguna aktif media sosial di Indonesia telah mencapai 167 juta jiwa atau sekitar 60,4% dari total populasi [1]. Dari berbagai platform yang ada, TikTok dan Instagram menjadi dua aplikasi yang paling banyak digunakan, khususnya di kalangan generasi muda.

Tingginya jumlah pengguna kedua platform tersebut turut mendorong bertambahnya ulasan yang diunggah melalui Google Play Store [2]. Ulasan-ulasan ini merupakan bentuk umpan balik langsung dari pengguna yang mencerminkan pengalaman nyata mereka, mulai dari kepuasan terhadap fitur, kritik atas kekurangan, hingga keluhan terkait performa aplikasi. Bagi pengembang, ulasan semacam ini tentu sangat berharga sebagai bahan evaluasi. Namun, volume ulasan yang terus bertambah membuat proses analisis secara manual menjadi tidak praktis dan memakan waktu yang cukup banyak.

Untuk mengatasi permasalahan tersebut, diperlukan pendekatan yang mampu memproses dan mengklasifikasikan ulasan secara otomatis dan efisien. Salah satu pendekatan yang relevan adalah analisis sentimen, yakni cabang dari Natural Language Processing (NLP) yang berfokus pada identifikasi dan ekstraksi opini atau kecenderungan emosional dari sebuah teks [3]. Dengan analisis sentimen, ribuan ulasan dapat diproses secara sistematis sehingga menghasilkan informasi yang berguna untuk mendukung evaluasi maupun pengambilan keputusan.

Berbagai algoritma machine learning telah digunakan dalam penelitian analisis sentimen, seperti Naive Bayes, Support Vector Machine (SVM), Random Forest, dan Extreme Gradient Boosting (XGBoost) [4]. Di antara

algoritma tersebut, XGBoost menonjol karena kemampuannya menghasilkan klasifikasi yang akurat, efisiensi komputasinya yang tinggi, serta ketangguhannya dalam menangani data yang kompleks dan tidak seimbang [5]. Performa XGBoost juga telah terbukti kompetitif dalam berbagai penelitian dan kompetisi machine learning, termasuk untuk tugas klasifikasi teks.

Sejumlah penelitian sebelumnya telah mengeksplorasi analisis sentimen pada ulasan aplikasi mobile menggunakan berbagai metode dan platform. Budianto et al. menerapkan XGBoost pada ulasan aplikasi M-Pajak di Google Play Store dan berhasil mencapai akurasi klasifikasi yang kompetitif [6]. Indriyani et al. membandingkan Naïve Bayes dan SVM pada ulasan TikTok, di mana SVM menunjukkan keunggulan pada kondisi data tidak seimbang [7]. Tangke et al. menggunakan SVM dan Random Forest untuk sentimen TikTok dan memperoleh akurasi terbaik 93% menggunakan Random Forest [8]. Pradhisa dan Fajriyah mengimplementasikan IndoBERT pada ulasan e-commerce di Google Play Store dengan pendekatan deep learning [9]. Husain et al. menggabungkan TF-IDF dengan SVM untuk mengklasifikasikan ulasan TikTok dan membuktikan efektivitas representasi fitur berbasis frekuensi kata [10]. Sementara itu, Hendrawan et al. membandingkan TF-IDF dan Word2Vec dalam klasifikasi sentimen produk lokal menggunakan XGBoost, dengan hasil menunjukkan bahwa kombinasi Word2Vec dan XGBoost menghasilkan F1-Score yang sedikit lebih tinggi [11].

Berdasarkan kajian terhadap penelitian terdahulu, berbagai algoritma seperti Naïve Bayes, Support Vector Machine (SVM), Random Forest, IndoBERT, dan XGBoost telah diterapkan untuk analisis sentimen dengan fokus, dataset, serta tujuan implementasi yang beragam [12]. Sebagian besar penelitian tersebut berfokus pada analisis satu aplikasi atau platform tertentu maupun pada evaluasi performa model klasifikasi. Oleh karena itu, penelitian ini mengembangkan analisis sentimen pada ulasan aplikasi TikTok dan Instagram yang diperoleh dari Google Play Store menggunakan algoritma XGBoost dengan ekstraksi fitur TF-IDF, serta mengimplementasikannya dalam bentuk sistem berbasis website yang mendukung prediksi manual, pemrosesan batch, dan visualisasi hasil analisis.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komputasional [13]. Alur penelitian secara umum dimulai dari pengumpulan data, preprocessing teks, ekstraksi fitur, pelatihan model, evaluasi performa, hingga implementasi sistem berbasis website. Alur lengkap penelitian dapat dilihat pada Gambar 1.



Gambar 1. Flowchart Kerangka Penelitian

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

2.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan ulasan pengguna aplikasi TikTok dan Instagram yang diperoleh dari Google Play Store menggunakan library google-play-scraper berbasis Python [6]. Dari proses scraping tersebut, diperoleh dua atribut utama, yaitu teks ulasan dan rating bintang pengguna dengan skala 1 hingga 5. Total data mentah yang berhasil dikumpulkan dari kedua aplikasi sebanyak 10.000 ulasan. Setelah melalui proses pembersihan awal yang mencakup penghapusan data duplikat, ulasan kosong, dan koreksi mismatch rating, jumlah data bersih yang digunakan dalam penelitian ini adalah 8.742 ulasan.

Pelabelan data dilakukan secara otomatis berdasarkan rating bintang [14]. Ulasan dengan rating 4 dan 5 dikategorikan sebagai sentimen positif, sedangkan ulasan dengan rating 1 dan 2 dikategorikan sebagai sentimen negatif. Ulasan dengan rating 3 dikeluarkan dari dataset karena dianggap netral dan dapat mengaburkan batas klasifikasi biner. Distribusi data yang digunakan disajikan pada Tabel 1.

Tabel 1. Distribusi Dataset

Kelas Sentimen	Data Training (80%)	Data Testing (20%)	Total
Positif	3.969	993	4.961 (56,75%)
Negatif	3.024	756	3.781 (43,25%)
Total	6.993	1.749	8.742

2.2. Preprocessing Teks

Tahap *preprocessing* bertujuan untuk membersihkan dan menormalisasi teks ulasan sebelum dilakukan proses ekstraksi fitur [15]. Pipeline *preprocessing* yang diterapkan dalam penelitian ini terdiri atas delapan tahapan yang dilakukan secara berurutan. Tahap pertama adalah mengubah seluruh teks menjadi huruf kecil (*case folding*) [15]. Tahap kedua yaitu menghapus URL yang diawali dengan http atau www. Tahap ketiga adalah menghapus *mention* (@username) dan *hashtag* (#topik). Selanjutnya, pada tahap keempat dilakukan penghapusan karakter angka, kemudian pada tahap kelima dilakukan penghapusan tanda baca serta karakter khusus menggunakan ekspresi regular [15]. Tahap keenam bertujuan untuk menormalkan spasi yang berlebih, sedangkan tahap ketujuh adalah melakukan *stopword removal* menggunakan kamus *stopword* bahasa Indonesia yang diperkaya dengan *custom stopwords* [16]. Terakhir, pada tahap kedelapan dilakukan penghapusan token yang memiliki panjang kurang dari tiga karakter. Seluruh tahapan tersebut diterapkan secara konsisten, baik pada proses pelatihan model maupun pada saat sistem melakukan prediksi secara *real-time*.

2.3. Ekstraksi Fitur TF-IDF

Representasi numerik teks dilakukan menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Metode ini mengukur tingkat kepentingan suatu kata dalam sebuah dokumen relatif terhadap keseluruhan korpus, sehingga kata yang terlalu umum akan mendapat bobot lebih rendah dibandingkan kata yang lebih spesifik dan informatif [17]. Nilai TF-IDF dihitung menggunakan persamaan berikut:

$$TF\text{-}IDF(t,d) = TF(t,d) \times log\left(\frac{N}{DF(t)}\right) \quad \ldots (1)$$

Keterangan: TF(t,d) adalah frekuensi kemunculan term t dalam dokumen d, N adalah jumlah total dokumen dalam korpus, dan DF(t) adalah jumlah dokumen yang mengandung term t. Konfigurasi TF-IDF yang digunakan dalam penelitian ini meliputi jumlah fitur maksimum sebesar 4.070, n-gram range (1,2) untuk menghasilkan fitur unigram dan bigram, minimum document frequency (min_df) sebesar 2, serta sublinear_tf=True untuk mengurangi dominasi kata yang sangat sering muncul [11].

2.4. Algoritma XGBoost

XGBoost (Extreme Gradient Boosting) merupakan implementasi dari teknik gradient boosting yang dioptimalkan untuk efisiensi komputasi dan performa prediksi [18]. Algoritma ini membangun model secara sekuensial, di mana setiap pohon baru memperbaiki kesalahan prediksi pohon sebelumnya dengan meminimalkan fungsi loss:

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

$$L(\varphi) = \Sigma\, l(\hat{y}_i, y_i) + \Sigma\, \Omega(f_k) \ldots (2)$$

Di mana l adalah fungsi loss yang mengukur selisih antara nilai prediksi dan nilai aktual, $\hat{y}_i$adalah nilai prediksi, $y_i$adalah nilai aktual, dan $\Omega(f_k)$adalah term regularisasi yang berfungsi mencegah overfitting. Pembagian data dilakukan menggunakan stratified split dengan rasio 80:20 untuk memastikan distribusi kelas yang proporsional pada data training maupun data testing [18]. Parameter model XGBoost yang digunakan dalam penelitian ini disajikan pada Tabel 2.

Tabel 2. Parameter Model XGBoost

Parameter	Nilai	Keterangan
n_estimators	200	Jumlah pohon
max_depth	6	Kedalaman pohon maksimum
learning_rate	0,1	Laju pembelajaran
subsample	0,8	Proporsi subsampel data
scale_pos_weight	Otomatis	Penanganan class imbalance
eval_metric	logloss	Fungsi evaluasi training

Kinerja model dievaluasi menggunakan empat metrik utama, yaitu akurasi, presisi, recall, dan F1-score yang dihitung berdasarkan confusion matrix [18]. Rumus setiap metrik evaluasi adalah sebagai berikut:

Akurasi: $$(TP + TN) / (TP + TN + FP + FN) \ldots (3)$$

Presisi: $$TP / (TP + FP) \ldots (4)$$

Recall: $$TP - (TP + FN) \ldots (5)$$

F1-Scorel: $$2 \times (Presisi \times Recall) / (Presisi + Recall) \ldots (6)$$

Keterangan: TP (True Positive), TN (True Negative), FP (False Positive), dan FN (False Negative) adalah komponen dari confusion matrix.

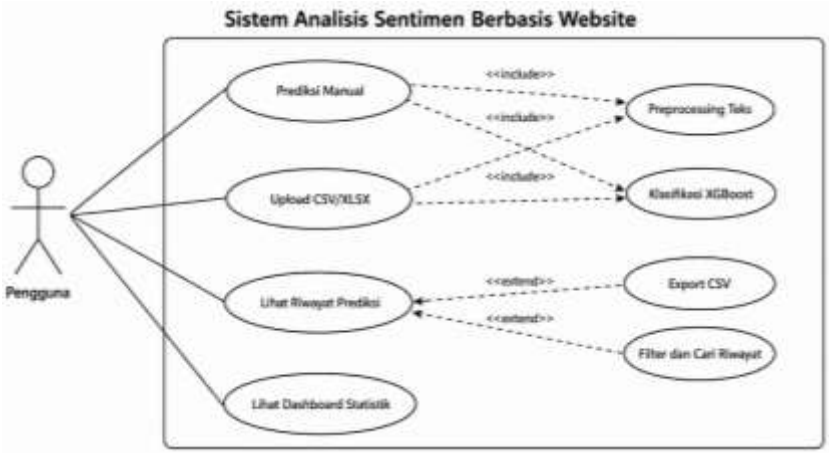
2.5. Perancangan Sistem

Perancangan sistem dilakukan untuk menggambarkan struktur dan alur kerja sistem analisis sentimen sebelum tahap implementasi. Pada tahap ini digunakan dua diagram utama, yaitu use case diagram untuk memodelkan interaksi antara pengguna dengan sistem, serta activity diagram untuk menggambarkan alur proses secara keseluruhan [19].

2.5.1. Use Case Diagram

Use case diagram pada Gambar 2 menggambarkan interaksi antara pengguna dengan sistem analisis sentimen berbasis website. Sistem memiliki satu aktor utama, yaitu Pengguna, yang dapat mengakses empat fungsi utama meliputi Prediksi Manual, Upload CSV/XLSX, Lihat Riwayat Prediksi, dan Lihat Dashboard Statistik.

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

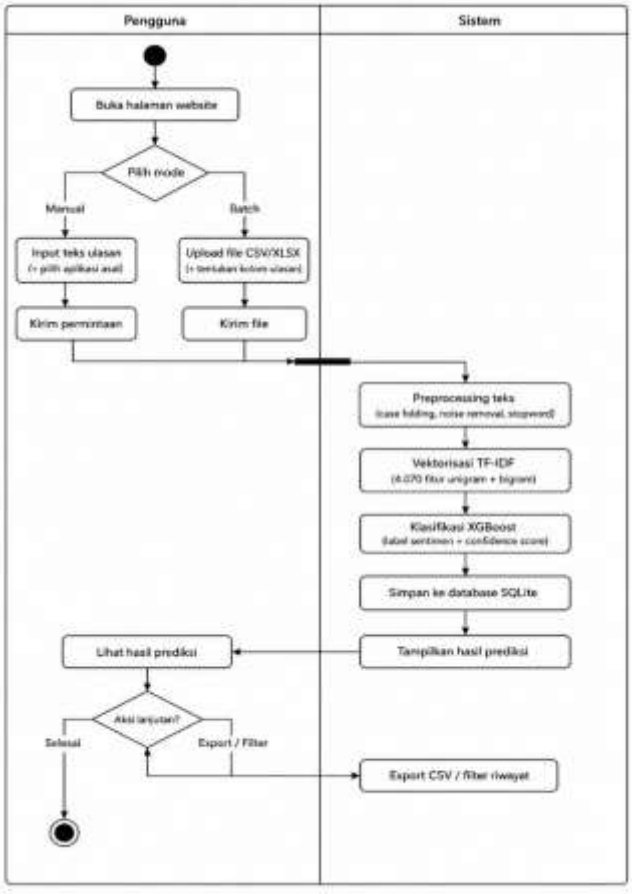


Gambar 2. Use Case Diagram

Berdasarkan Gambar 2, sistem dirancang agar pengguna dapat melakukan analisis sentimen baik secara individual maupun dalam jumlah besar sekaligus. Pada mode prediksi manual, pengguna memasukkan satu teks ulasan dan sistem akan menghasilkan label sentimen beserta confidence score. Pada mode batch, pengguna mengunggah file CSV atau XLSX yang berisi kumpulan ulasan untuk diproses sekaligus. Selain itu, pengguna juga dapat melihat riwayat prediksi yang tersimpan dan memantau statistik keseluruhan melalui halaman dashboard.

2.5.2. Activity Diagram

Activity diagram pada Gambar 3 menggambarkan alur aktivitas pengguna saat menggunakan sistem dari awal hingga mendapatkan hasil prediksi.



Gambar 3. Activity Diagram

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Berdasarkan Gambar 3, alur sistem dimulai ketika pengguna membuka halaman website dan memilih mode prediksi yang diinginkan. Pada mode prediksi manual, pengguna memasukkan teks ulasan dan memilih aplikasi asal, kemudian sistem secara otomatis menjalankan pipeline preprocessing, melakukan vektorisasi TF-IDF, dan meneruskan hasilnya ke model XGBoost untuk menghasilkan label sentimen beserta confidence score. Pada mode batch, pengguna mengunggah file CSV atau XLSX yang berisi kolom ulasan, lalu sistem memproses seluruh baris data secara iteratif menggunakan pipeline yang identik dengan prediksi manual. Semua hasil prediksi disimpan ke database SQLite dan dapat diakses kembali melalui halaman riwayat maupun diekspor dalam format CSV.

2.6. Pengembangan Sistem

Sistem dibangun menggunakan framework Flask berbasis Python dengan antarmuka web yang dikembangkan menggunakan Bootstrap 5 agar tampilannya responsif di berbagai ukuran perangkat [20]. Untuk penyimpanan riwayat prediksi digunakan database SQLite yang ringan namun cukup memadai untuk kebutuhan sistem ini. Model XGBoost yang telah dilatih disimpan dalam format pickle (.pkl) bersama dengan TF-IDF vectorizer dan kamus stopwords, kemudian dimuat secara otomatis saat aplikasi dijalankan. Selain itu, sistem juga menyediakan tiga endpoint RESTful API yang memungkinkan integrasi dengan sistem eksternal untuk keperluan prediksi otomatis [21].

3. Hasil dan Diskusi

Bab ini menyajikan hasil yang diperoleh dari setiap tahapan penelitian, mulai dari proses preprocessing dan ekstraksi fitur, evaluasi kinerja model XGBoost, implementasi sistem berbasis website, hingga pembahasan terhadap hasil yang diperoleh

3.1. Hasil Preprocessing dan Ekstraksi Fitur

Tahap *preprocessing* berhasil mengurangi berbagai *noise* pada data ulasan sehingga teks menjadi lebih bersih dan siap untuk diproses pada tahap selanjutnya. Dari total 8.742 ulasan yang digunakan, rata-rata panjang teks sebelum *preprocessing* mencapai 47,3 karakter, sedangkan setelah dilakukan pembersihan berkurang menjadi 28,6 karakter.

Proses *stopword removal* juga berjalan dengan baik dengan menghilangkan kata-kata yang kurang memberikan informasi penting, seperti *"yang"*, *"dan"*, serta *"di"*. Selain itu, frasa umum yang tidak berkontribusi dalam membedakan sentimen turut dihapus dari dataset. Setelah melalui tahapan tersebut, proses ekstraksi menggunakan metode TF-IDF menghasilkan 4.070 fitur yang terdiri atas unigram dan bigram. Beberapa contoh bigram yang memiliki nilai informatif antara lain *"sangat bagus"*, *"tidak bisa"*, *"aplikasi lambat"*, dan *"sering error"*, sehingga mampu memberikan representasi konteks yang lebih baik dibandingkan penggunaan unigram saja.

3.2. Hasil Evaluasi Model

Evaluasi model XGBoost dilakukan menggunakan data pengujian (testing) sebanyak 1.749 sampel dengan mengacu pada empat metrik utama, yaitu precision, recall, F1-score, dan akurasi. Hasil evaluasi model ditampilkan pada Tabel 3.

Tabel 3. Hasil Evaluasi Model XGBoost

Kelas	Precision	Recall	F1-Score	Support
Negatif	0,8951	0,7566	0,8201	756
Positif	0,8341	0,9325	0,8805	993
Macro avg	0,8646	0,8446	0,8503	1.749
Weighted avg	0,8810	0,8759	0,8741	1.749

Berdasarkan Tabel 3, model XGBoost memperoleh tingkat akurasi sebesar 87,59% dengan nilai *weighted average precision* sebesar 88,10%, *recall* sebesar 87,59%, dan *F1-score* sebesar 87,41%. Pada kelas negatif, nilai *precision* mencapai 89,51%, lebih tinggi dibandingkan kelas positif yang memperoleh 83,41%. Hal ini menunjukkan bahwa model cukup baik dalam mengidentifikasi ulasan negatif dengan tingkat kesalahan prediksi yang relatif rendah.

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Di sisi lain, kelas positif memiliki nilai *recall* sebesar 93,25%, yang mengindikasikan bahwa sebagian besar ulasan positif berhasil dikenali dengan benar oleh model. Untuk memberikan gambaran yang lebih rinci mengenai hasil klasifikasi, *confusion matrix* disajikan pada Tabel 4.

Tabel 4. Confusion Matrix Model XGBoost (Data Test = 1.749 Sampel)

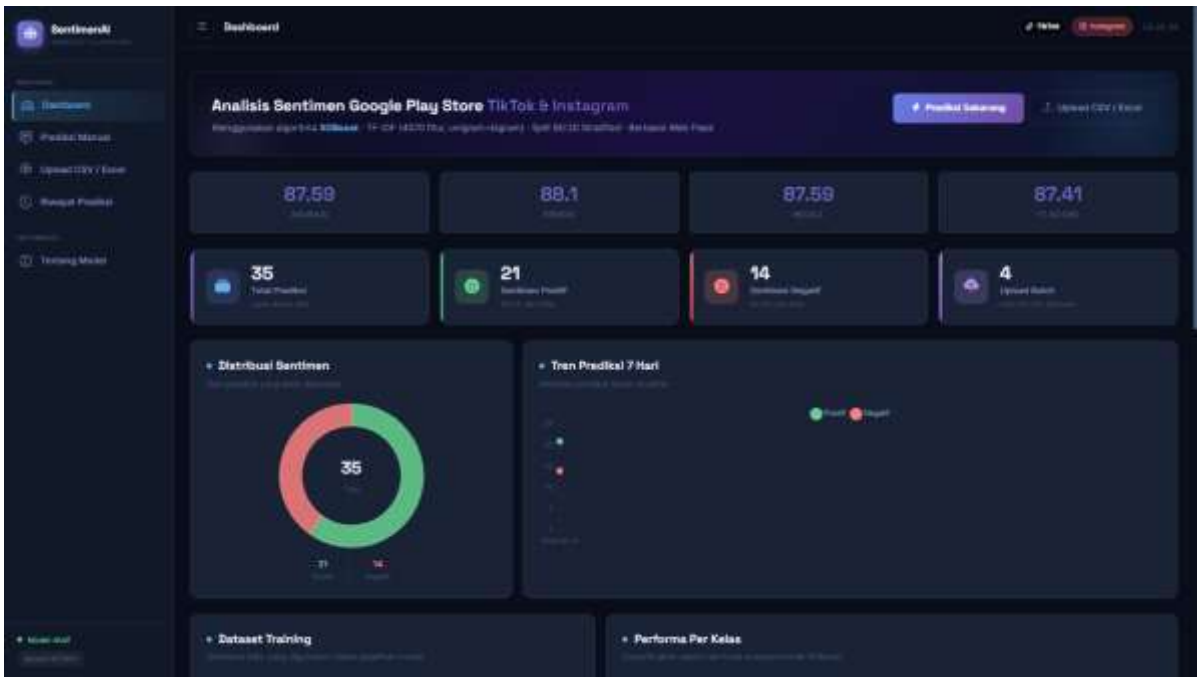
Aktual	Prediksi Negatif	Prediksi Positif
Aktual Negatif	TN = 572	FP = 184
Aktual Positif	FN = 67	TP = 925

Berdasarkan Tabel 4, dari 756 data yang termasuk ke dalam kelas negatif, sebanyak 572 ulasan berhasil diprediksi dengan benar sebagai negatif (*True Negative*), sedangkan 184 ulasan salah diklasifikasikan sebagai positif (*False Positive*). Sementara itu, dari 993 data pada kelas positif, sebanyak 925 ulasan berhasil diprediksi dengan benar (*True Positive*) dan hanya 67 ulasan yang salah diklasifikasikan sebagai negatif (*False Negative*).

Masih ditemukannya nilai *False Positive* yang cukup tinggi menunjukkan bahwa beberapa ulasan negatif memiliki pola bahasa yang menyerupai ulasan positif. Kondisi ini diduga dipengaruhi oleh penggunaan ungkapan sarkastik, ironi, atau pujian yang disertai kritik, sehingga model mengalami kesulitan dalam membedakan makna sebenarnya dari teks ulasan tersebut.

3.3. Hasil Implementasi Sistem

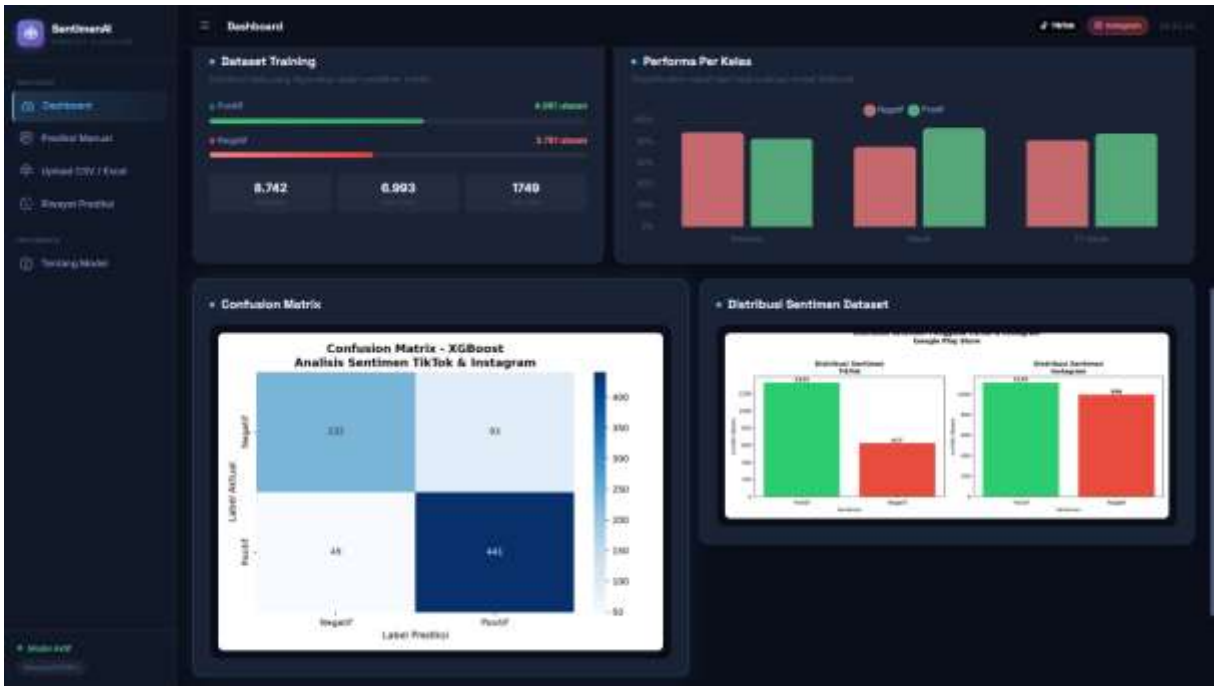
Sistem analisis sentimen pada penelitian ini berhasil diimplementasikan dalam bentuk aplikasi berbasis web menggunakan framework Flask dengan antarmuka yang dibangun menggunakan Bootstrap 5. Aplikasi terdiri atas empat halaman utama, yaitu Dashboard, Prediksi Manual, Upload CSV, dan Riwayat Prediksi. Tampilan masing-masing halaman ditunjukkan pada Gambar 4 hingga Gambar 8.



Gambar 4. Halaman Dashboard

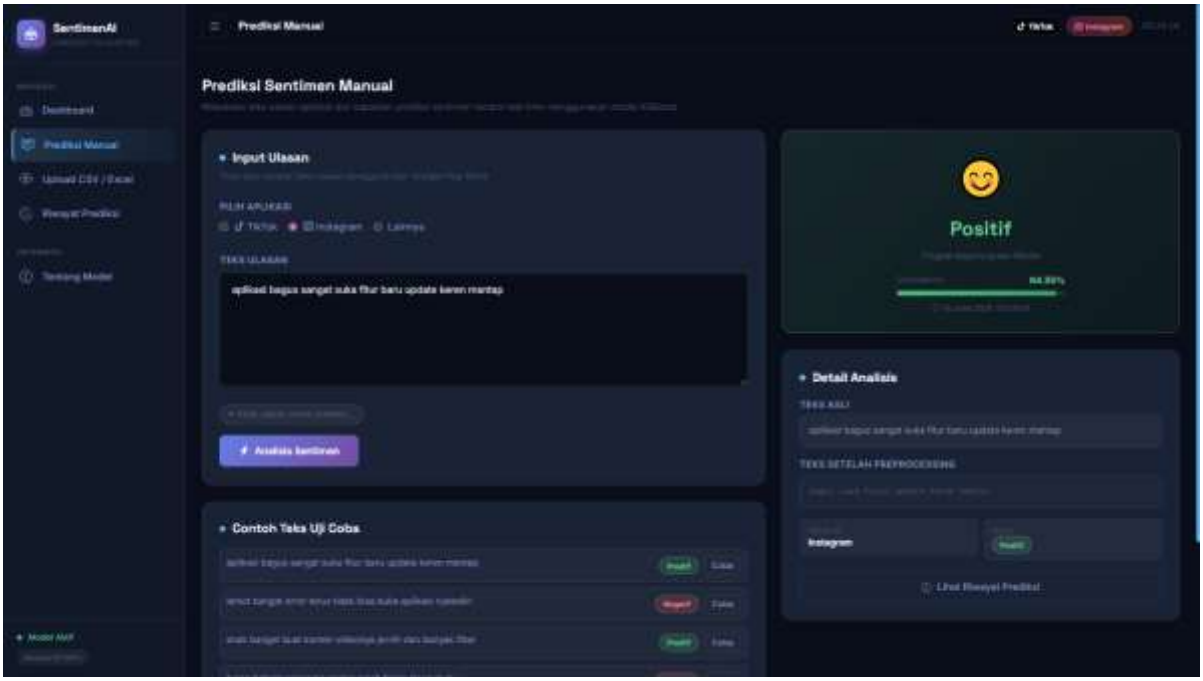
Gambar 4 merupakan halaman dashboard yang menyajikan ringkasan informasi mengenai hasil prediksi yang telah dilakukan, seperti jumlah total data yang diproses, distribusi sentimen positif dan negatif, serta visualisasi tren prediksi harian dalam bentuk grafik interaktif menggunakan Chart.js. Informasi tersebut membantu pengguna memperoleh gambaran umum mengenai hasil analisis secara cepat.

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



Gambar 5. Dashboard Visualisasi Hasil Analisis Sentimen

Gambar 5 menampilkan halaman dashboard utama sistem analisis sentimen yang menyajikan ringkasan informasi hasil pemodelan dan evaluasi secara visual. Dashboard ini memuat distribusi dataset pelatihan, perbandingan metrik performa model untuk setiap kelas sentimen, *confusion matrix*, serta distribusi sentimen pada ulasan aplikasi TikTok dan Instagram. Penyajian informasi dalam bentuk grafik dan diagram memudahkan pengguna dalam memantau hasil klasifikasi serta mengevaluasi kinerja model XGBoost secara lebih cepat dan informatif.

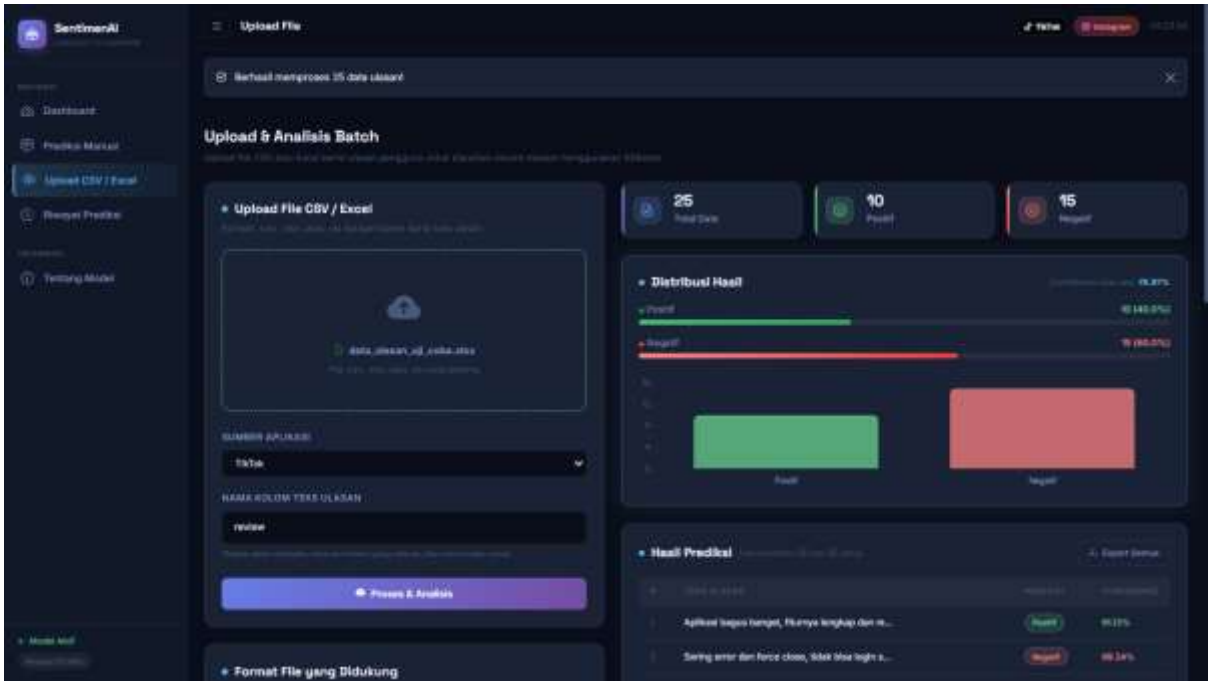


Gambar 6. Halaman Prediksi Manual

Gambar 6 merupakan halaman prediksi manual yang memungkinkan pengguna memasukkan satu ulasan secara langsung serta memilih sumber aplikasi, yaitu TikTok atau Instagram. Setelah data dikirim, sistem secara otomatis menjalankan proses *preprocessing*, melakukan ekstraksi fitur menggunakan TF-IDF, dan mengklasifikasikan

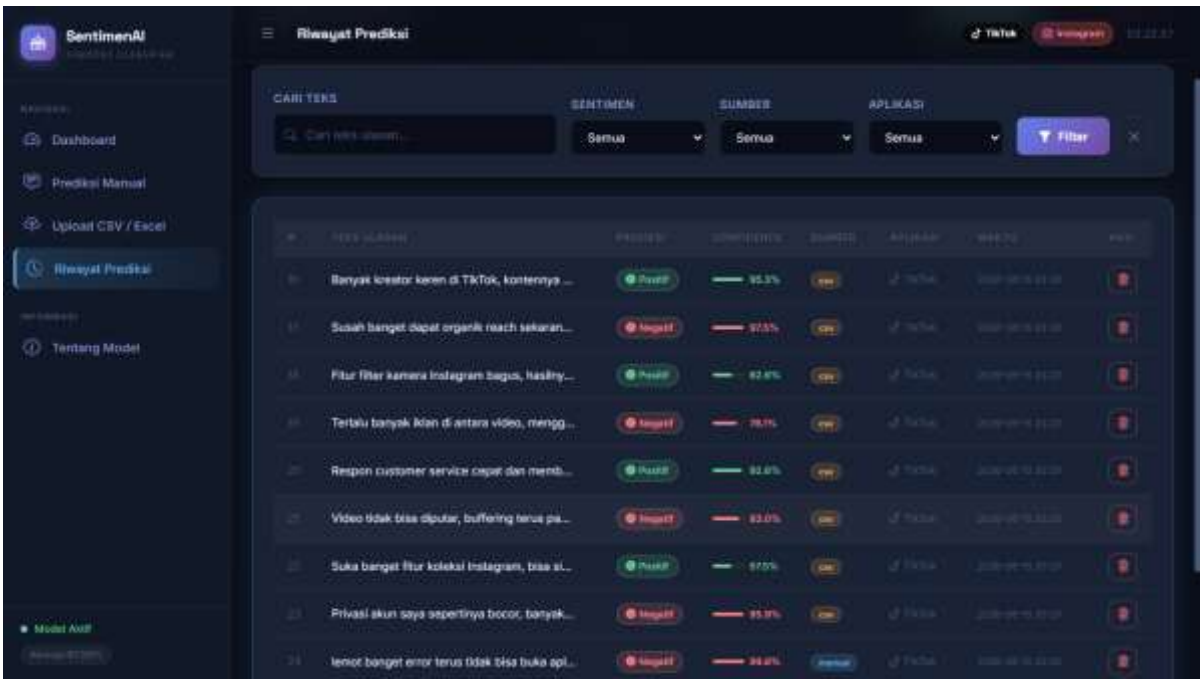
DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

sentimen menggunakan model XGBoost. Hasil prediksi ditampilkan dalam bentuk label sentimen beserta *confidence score* dan probabilitas masing-masing kelas dengan waktu pemrosesan yang relatif singkat.



Gambar 7. Halaman Upload CSV/XLSX

Gambar 7 merupakan halaman upload CSV/XLSX yang dirancang untuk memfasilitasi prediksi dalam jumlah besar (*batch processing*). Pengguna dapat mengunggah file berformat CSV, XLSX, maupun XLS yang berisi kumpulan ulasan, kemudian menentukan kolom yang akan diproses serta memilih asal aplikasi. Setelah proses selesai, sistem menampilkan ringkasan hasil prediksi beserta pratinjau data dan menyimpan seluruh hasil ke dalam basis data.



Gambar 8. Halaman Riwayat Prediksi

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Gambar 8 merupakan halaman riwayat prediksi yang menampilkan seluruh hasil analisis yang telah tersimpan di dalam database. Informasi yang disajikan meliputi teks ulasan, label sentimen, *confidence score*, sumber prediksi, dan asal aplikasi. Selain itu, pengguna juga dapat melakukan pencarian, penyaringan berdasarkan kategori tertentu, serta mengekspor data ke dalam format CSV untuk kebutuhan analisis lebih lanjut.

3.4. Diskusi

Hasil penelitian menunjukkan bahwa algoritma XGBoost mampu mengklasifikasikan sentimen ulasan pengguna TikTok dan Instagram dari Google Play Store dengan akurasi sebesar 87,59%. Hasil tersebut sejalan dengan penelitian Budianto dkk. [6] yang menunjukkan bahwa XGBoost memiliki performa yang baik dalam tugas klasifikasi sentimen pada ulasan aplikasi dan mampu menghasilkan prediksi yang kompetitif.

Performa model yang diperoleh didukung oleh kemampuan XGBoost sebagai algoritma *ensemble learning* yang membangun model secara bertahap untuk memperbaiki kesalahan prediksi sebelumnya, sehingga efektif dalam mempelajari pola-pola kompleks pada data teks [18]. Selain itu, penggunaan metode TF-IDF dengan kombinasi unigram dan bigram mampu menghasilkan representasi fitur yang lebih informatif karena tidak hanya mempertimbangkan frekuensi kemunculan kata tunggal, tetapi juga hubungan antarkata yang sering muncul secara bersamaan. Kondisi ini berkontribusi terhadap peningkatan kualitas proses klasifikasi sentimen [17]. Temuan ini juga didukung oleh penelitian Thomas dan Rumaisa [22] yang menunjukkan bahwa pemanfaatan TF-IDF dapat meningkatkan efektivitas analisis sentimen pada data teks berbahasa Indonesia.

Meskipun demikian, nilai *recall* pada kelas negatif masih lebih rendah dibandingkan kelas positif, yang mengindikasikan bahwa sebagian ulasan negatif belum berhasil dikenali secara tepat oleh model. Kondisi tersebut diduga dipengaruhi oleh penggunaan bahasa yang ambigu, sarkasme, atau campuran bahasa formal dan informal yang sering ditemukan pada ulasan pengguna. Di samping itu, tahapan *preprocessing* seperti *stemming* dan *stopword removal* juga berpengaruh terhadap kualitas representasi teks yang digunakan dalam proses klasifikasi [15].

Oleh karena itu, penelitian selanjutnya dapat mempertimbangkan penambahan leksikon sentimen bahasa Indonesia maupun penerapan metode yang lebih canggih untuk meningkatkan kemampuan model dalam mengenali variasi bahasa dan konteks sentimen yang lebih kompleks. Optimalisasi proses *preprocessing*, termasuk pemilihan *stopword*, juga dapat menjadi salah satu strategi untuk meningkatkan performa model [16].

4. Kesimpulan

Penelitian ini berhasil merancang dan mengimplementasikan sistem analisis sentimen berbasis website untuk mengklasifikasikan ulasan pengguna aplikasi TikTok dan Instagram dari Google Play Store menggunakan algoritma XGBoost dengan ekstraksi fitur TF-IDF. Berdasarkan hasil pengujian, model yang dibangun mampu mencapai akurasi sebesar 87,59%, dengan nilai precision sebesar 88,10%, recall sebesar 87,59%, dan F1-score sebesar 87,41%, sehingga menunjukkan kinerja yang baik dalam mengklasifikasikan sentimen positif dan negatif pada ulasan pengguna. Selain menghasilkan performa klasifikasi yang baik, penelitian ini juga berhasil mengimplementasikan model ke dalam sistem berbasis website yang menyediakan fitur prediksi manual, pemrosesan data melalui unggahan file CSV/XLSX, visualisasi hasil analisis, dan penyimpanan riwayat prediksi. Implementasi tersebut diharapkan dapat membantu pengguna maupun peneliti dalam melakukan analisis sentimen secara lebih praktis dan efisien. Meskipun demikian, penelitian ini masih memiliki beberapa keterbatasan, seperti pelabelan sentimen yang didasarkan pada rating pengguna dan penggunaan data yang hanya berasal dari Google Play Store. Oleh karena itu, penelitian selanjutnya dapat mempertimbangkan penggunaan teknik pelabelan yang lebih komprehensif, penambahan sumber data dari platform lain, serta eksplorasi metode berbasis *deep learning* atau optimasi model untuk meningkatkan performa klasifikasi.

Referensi

[1] S. Kemp, “Digital 2024: Indonesia,” DataReportal. Accessed: Jun. 17, 2026. [Online]. Available: https://datareportal.com/reports/digital-2024-indonesia
[2] D. S. A. Rizki, M. S. Khabib, N. Rahmayuna, and V. G. Utomo, “Klasifikasi Sentimen Ulasan Pengguna Aplikasi Layanan Publik Google Play Store Menggunakan NLP dan ML,” *J. Tempo Kompak*, vol. 20, no. 1, pp. 51–64, 2025, doi: 10.33365/jtk.v20i1.586.
[3] J. P. Hidayat and I. Nurhaida, “Analisis Sentimen pada Ulasan LMS Pembelajaran Menggunakan Metode Natural Language Processing,” *J. Ilm. Penelit. dan Pembelajaran Inform.*, vol. 10, no. 1, pp. 565–575, 2025, doi: 10.29100/jipi.v10i1.5899.
[4] O. S. D. Fadhillah, J. H. Jaman, and Carudin, “Perbandingan Naive Bayes, Support Vector Machine, Logistic Regression, dan Random

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Forest dalam Menganalisis Sentimen Mengenai TikTok Shop," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 1, pp. 840–847, 2025, doi: 10.23960/jitet.v13i1.5746.

[5] I. R. Hendrawan, E. Utami, and A. D. Hartanto, "Comparison of Naïve Bayes Algorithm and XGBoost on Local Product Review Text Classification," *J. Pendidik. Inform.*, vol. 6, no. 1, pp. 143–149, 2022, doi: 10.29408/edumatic.v6i1.5613.

[6] T. A. C. Budianto, B. A. Dermawan, and M. Jajuli, "Analisis Sentimen Ulasan Aplikasi M-Pajak pada Google Play Store Menggunakan XGBoost," *J. Sist. Inf. Apl.*, vol. 3, no. 1, pp. 11–25, 2025, doi: 10.52958/jsia.v3i1.12440.

[7] F. A. Indriyani, A. Fauzi, and S. Faisal, "Analisis Sentimen Aplikasi TikTok Menggunakan Algoritma Naïve Bayes dan Support Vector Machine," *J. Sains, Teknol. dan Inform.*, vol. 10, no. 2, pp. 176–184, 2023, doi: 10.37373/tekno.v10i2.419.

[8] R. Tangke, D. T. Salaki, W. W. Kalengkongan, and E. Ketaren, "Analisis Sentimen Aplikasi TikTok Menggunakan Algoritma Support Vector Machine (SVM) dan Random Forest," *J. Times*, vol. 13, no. 2, pp. 53–62, 2024, doi: 10.51351/jtm.13.2.2024762.

[9] K. C. Pradhisa and R. Fajriyah, "Analisis Sentimen Ulasan Pengguna E-commerce di Google Play Store Menggunakan Metode IndoBERT," *Build. Informatics, Technol. Sci.*, vol. 6, no. 1, pp. 92–104, 2024, doi: 10.47065/bits.v6i1.5247.

[10] Sukirman, Sajiah, N. P. Husain, A. F. Syam, and Ragil, "Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine," *J. Syst. Comput. Eng.*, vol. 5, no. 1, pp. 91–102, 2024, doi: 10.61628/jsce.v5i1.1105.

[11] I. R. Hendrawan, E. Utami, and A. D. Hartanto, "Analisis Perbandingan Metode Tf-Idf dan Word2vec pada Klasifikasi Teks Sentimen Masyarakat Terhadap Produk Lokal di Indonesia," *Smart Comp*, vol. 11, no. 3, pp. 497–503, 2022, doi: 10.30591/smartcomp.v11i3.3902.

[12] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes," *J. Tek. Inform. dan Sist. Inf.*, vol. 9, no. 2, pp. 785–795, 2022, doi: 10.35957/jatisi.v9i2.1835.

[13] Juliana, S. Ondeng, and U. Rahman, "Konsep Dasar Penelitian Kuantitatif," *J. Pendidik. Tambusai*, vol. 10, no. 1, pp. 140–146, 2026, doi: 10.31004/jptam.v10i1.35817.

[14] N. D. Kurniawan, P. R. Ferdian, and N. Hidayati, "Analisis Sentimen Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest pada Ulasan Aplikasi Ajaib," *J. Nas. Teknol. dan Sist. Inf.*, vol. 11, no. 1, pp. 87–97, 2025, doi: 10.25077/TEKNOSI.v11i1.2025.87-97.

[15] M. U. Albab, Y. Karuniawati P, and M. N. Fawaiq, "Optimization of the Stemming Technique on Text preprocessing President 3 Periods Topic," *J. Transform.*, vol. 20, no. 2, pp. 1–10, 2023, doi: 10.26623/transformatika.v20i2.5374.

[16] S. J. Angelina, A. B. P. Negara, and H. Muhardi, "Analisis Pengaruh Penerapan Stopword Removal pada Performa Klasifikasi Sentimen Tweet Bahasa Indonesia," *J. Apl. dan Ris. Inform.*, vol. 02, no. 1, pp. 165–173, 2023, doi: 10.26418/juara.v2i1.69680.

[17] D. Septiani and I. Isabela, "Analisis Term Frequency Inverse Document Frequency (TF-IDF) dalam Temu Kembali Informasi pada Dokumen Teks," *J. Sist. dan Teknol. Inf. Indones.*, vol. 1, no. 2, pp. 81–88, 2022, [Online]. Available: https://journal.unj.ac.id/unj/index.php/SINTESIA/article/view/39364

[18] Yulistiani and Styawati, "Analisis Sentimen Terhadap Calon Presiden Indonesia 2024 dengan Metode Extreme Gradient Boosting (XGBOOST)," *J. Pengemb. IT*, vol. 9, no. 3, pp. 322–328, 2024, doi: 10.30591/jpit.v9i3.6127.

[19] A. Aurellia, S. Nooriansyah, and Y. Amrozi, "Pemanfaatan UML dalam Perancangan Sistem Informasi Produk Kreatif Daur Ulang Sampah Berbasis Web," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 3, pp. 1739–1749, 2025, doi: 10.23960/jitet.v13i3S1.8073.

[20] H. L. Walingkas and P. O. N. Saian, "Penerapan Framework Flask pada Pembangunan Sistem Informasi Pemasok Barang," *J. Teknol. Inf. dan Komun.*, vol. 7, no. 2, pp. 227–234, 2023, doi: 10.35870/jtik.v7i2.729.

[21] Herfandi, M. Julkarnain, and M. Hanif, "Desain dan Implementasi Restful Web Services untuk Integrasi Data dan Aplikasi," *J. Inform. Teknol. dan Sains*, vol. 4, no. 1, pp. 36–41, 2022, doi: 10.51401/jinteks.v4i1.1529.

[22] V. W. D. Thomas and F. Rumaisa, "Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF," *J. Media Inform. Budidarma*, vol. 6, no. 3, pp. 1767–1774, 2022, doi: 10.30865/mib.v6i3.4218.

DOI: https://doi.org/10.31004/riggs.v5i2.10584
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)