

ANALISIS SENTIMEN PADA KOMENTAR APLIKASI SEHAT INDONESIAKU MENGGUNAKAN METODE MACHINE LEARNING DAN KLASIFIKASI POSITIF-NEGATIF

Nabila Yudhitya Larasati

Program Studi Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur

e-mail: 23083010086@student.upnjatim.ac.id

Marthalia Kusumarima

Program Studi Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur

Karina Auralia

Program Studi Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur

Shindi Shella May Wara

Program Studi Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur

Aviolla Terza Damaliana

Program Studi Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur

Abstrak

Perkembangan aplikasi kesehatan digital seperti Aplikasi Sehat Indonesiaku (ASIK) memerlukan pemahaman terhadap persepsi pengguna guna meningkatkan kualitas layanan. Penelitian ini bertujuan untuk mengembangkan sistem analisis sentimen terhadap komentar pengguna ASIK di Google Play Store. Sebanyak 1.053 komentar dikumpulkan melalui teknik *web scraping*, lalu diproses menggunakan tahapan normalisasi, *case folding*, *stemming*, penghapusan *stopwords*, dan pembobotan TF-IDF. Visualisasi word cloud digunakan untuk mengidentifikasi kata-kata dominan dalam komentar positif dan negatif. Proses klasifikasi sentimen dilakukan dengan tiga algoritma machine learning, yaitu Random Forest, Gradient Boosting, dan Logistic Regression. Hasil analisis menunjukkan bahwa komentar positif banyak memuat kata seperti "mudah" dan "bermanfaat", sementara komentar negatif didominasi oleh kata "perbaiki" dan "lama". Model Gradient Boosting menunjukkan performa terbaik dengan *accuracy* sebesar 87,83%, *precision* 87,47%, dan *F1-Score* 87,31%. Temuan ini menggambarkan persepsi publik terhadap ASIK serta menyoroti pentingnya perbaikan teknis guna meningkatkan kepuasan pengguna terhadap layanan digital pemerintah.

Kata Kunci: Analisis Sentimen, *Text Mining*, Aplikasi Sehat Indonesiaku

Abstract

The development of digital health applications such as the Sehat Indonesiaku (ASIK) application requires an understanding of user perceptions in order to improve service quality. This study aims to develop a sentiment analysis system for ASIK user comments on the Google Play Store. A total of 1,053 comments were collected using web scraping techniques, then processed using normalization, case folding, stemming, stopword removal, and TF-IDF weighting. Word cloud visualization was used to identify dominant words in positive and negative comments. Sentiment classification was performed using three machine learning algorithms: Random Forest, Gradient Boosting, and Logistic Regression. The analysis results showed that positive comments frequently contained words like "easy" and "useful," while negative comments were dominated by words like "improve" and "slow." The Gradient Boosting model demonstrated the best performance with an accuracy of 87.83%, precision of 87.47%, and F1-Score of 87.31%. These findings reflect public perception of ASIK and highlight the importance of technical improvements to enhance user satisfaction with government digital services.

Keywords: Sentiment Analysis, *Text Mining*, Sehat Indonesiaku Application

PENDAHULUAN (GUNAKAN STYLE SECTION)

Perkembangan aplikasi kesehatan digital telah menjadi fenomena penting di tingkat global maupun nasional. Di Indonesia, Aplikasi Sehat Indonesiaku (ASIK) yang dikembangkan Kementerian Kesehatan RI berperan sebagai platform terpadu untuk memfasilitasi akses layanan kesehatan, pencatatan dan pemantauan pelayanan kesehatan masyarakat. Namun, efektivitas aplikasi ini sangat bergantung pada penerimaan pengguna, yang dapat diukur melalui analisis sentimen terhadap ulasan dan komentar di platform aplikasi.

Analisis sentimen merupakan cara yang efektif untuk memahami bagaimana tanggapan dan pandangan pengguna terhadap suatu produk atau layanan (Kusuma & Ali, 2024). Analisis ini biasa dikenal sebagai *opinion mining*, yang merupakan proses untuk mengenali opini seseorang yang disampaikan dalam bentuk teks, kemudian menentukan apakah opini tersebut bersifat positif atau negatif (Siregar et al., 2024). Semakin banyaknya komentar pengguna yang tersedia secara online, teknik pengumpulan data otomatis seperti *web scraping* menjadi sangat penting untuk mengumpulkan data secara efisien dan terstruktur.

Beberapa penelitian sebelumnya telah melakukan analisis sentimen dengan berbagai pendekatan *machine learning*. Penelitian Terdahulu menunjukkan keberagaman kinerja model tergantung konteks dan bahasa. Penelitian oleh (Rahmawati & Fitriani, 2023) pada analisis sentimen maskapai Lion Air menemukan bahwa Logistic Regression unggul dengan *accuracy* 82% (*precision* 0.82, *recall* 0.78), mengalahkan Naïve Bayes (47%) dan Random Forest (39%). Temuan ini kontras dengan penelitian (Muntiari et al., 2023) pada aplikasi pendidikan (DITA), di mana Gradient Boosting mencapai *accuracy* 97.5%, serta studi Larasati et al. (2022) pada aplikasi DANA yang menunjukkan keefektifan Random Forest (*accuracy* 84%). Perbedaan hasil ini mengindikasikan bahwa tidak ada model universal yang cocok untuk semua domain, sehingga perlu uji komparatif spesifik untuk aplikasi kesehatan digital.

Namun demikian, sebagian besar studi yang ada masih terfokus pada aplikasi komersial atau layanan swasta, sedangkan kajian terhadap aplikasi layanan publik, khususnya milik pemerintah masih

tergolong terbatas. Padahal, dalam era digital yang semakin berkembang, keberadaan aplikasi layanan kesehatan berbasis teknologi menjadi salah satu pilar penting dalam meningkatkan akses dan kualitas pelayanan publik.

Oleh karena itu, penelitian ini bertujuan untuk mengembangkan sistem analisis sentimen yang efektif pada komentar pengguna Aplikasi Sehat Indonesiaku dengan memanfaatkan teknik *web scraping* untuk pengumpulan data dan metode klasifikasi berbasis *machine learning* untuk mengkategorikan sentimen positif dan negatif. Dengan memanfaatkan ulasan daring yang merefleksikan opini publik secara alami, studi ini diharapkan dapat memberikan wawasan berharga bagi pengembang aplikasi dalam meningkatkan kualitas layanan berbasis digital pemerintah.

KAJIAN TEORI (GUNAKAN STYLE SECTION)

DATA MINING

Data mining merupakan proses sistematis untuk memperoleh informasi atau pengetahuan yang bernali dari kumpulan data yang besar dan kompleks. Tujuan utamanya adalah untuk menemukan pola, keterkaitan, atau informasi tersembunyi yang tidak langsung terlihat, sehingga dapat memberikan pemahaman yang lebih mendalam dan bernali dalam pengambilan keputusan (Rahayu et al., 2024).

TEXT MINING

Text mining merupakan proses yang melibatkan analisis mendalam terhadap kumpulan dokumen, di mana pengguna secara aktif berinteraksi dengan data tersebut menggunakan berbagai alat analisis untuk memperoleh pengetahuan dari waktu ke waktu (Findawati & Rosid, 2020).

ANALISIS SENTIMEN

Analisis sentimen merupakan salah satu bidang dalam penelitian text mining yang berfokus pada proses klasifikasi dokumen teks. Melalui analisis ini, opini, emosi, maupun evaluasi tertulis seseorang mengenai suatu topik tertentu dapat diekstraksi dengan memanfaatkan teknik pemrosesan bahasa alami (Natural Language Processing) (Idris et al., 2023).

WEB SCRAPING

Web scraping merupakan teknik yang digunakan untuk mengekstraksi atau mengambil data dari suatu situs web secara otomatis (Dwicahyo & Ratnasari, 2023). Teknik ini memungkinkan pengambilan informasi yang tersedia di halaman web untuk kemudian diolah atau dianalisis lebih lanjut.

RANDOM FOREST

Algoritma ini merupakan pengembangan dari metode pohon keputusan (Decision Tree), di mana sistem membangun sejumlah pohon keputusan secara paralel dan menggabungkan hasil prediksi dari masing-masing pohon untuk memperoleh hasil akhir yang lebih akurat. Random Forest terbukti efektif dalam meminimalkan risiko *overfitting* serta meningkatkan kinerja model, karena pendekatannya yang mengkombinasikan berbagai prediksi secara kolektif dari banyak pohon keputusan (Bintoro et al., 2024).

GRADIENT BOOSTING

Algoritma ini membangun model secara berurutan dengan fokus pada pengurangan kesalahan residu dari model sebelumnya. Setiap model yang dibentuk dilatih untuk memperbaiki kelemahan atau kesalahan yang masih tersisa dari model sebelumnya. Hasil akhir diperoleh dengan menggabungkan prediksi dari seluruh model yang telah dibentuk, sehingga menghasilkan kinerja yang lebih optimal dan akurat (Bintoro et al., 2024).

LOGISTIC REGRESSION

Logistic regression adalah suatu metode untuk membangun model prediksi yang mirip dengan regresi linear. Perbedaan utama terletak pada jenis variabel dependen yang diprediksi. Dalam logistic regression, variabel dependen tersebut bersifat dikotomi, yaitu data nominal yang terdiri dari dua kategori, seperti Ya dan Tidak, Baik dan Buruk, atau Tinggi dan Rendah (Fahmuddin et al., 2023).

ACCURACY

Accuracy merupakan persentase jumlah prediksi yang benar dibandingkan dengan total seluruh prediksi yang dilakukan oleh model (Hartono, n.d.). Metrik ini sering digunakan sebagai indikator awal untuk menilai kinerja model.

PRECISION

Presisi adalah rasio antara jumlah prediksi positif yang benar (*true positive*) dengan seluruh prediksi yang diklasifikasikan sebagai positif (*true positive + false positive*) (Hartono, n.d.). Metrik ini menggambarkan seberapa akurat model dalam mengidentifikasi data yang benar-benar relevan sebagai positif, sehingga semakin tinggi nilai presisi, semakin sedikit kesalahan dalam memprediksi kelas positif.

F1-SCORE

F1-Score merupakan rata-rata harmonis dari nilai presisi dan *recall*, yang digunakan untuk memberikan keseimbangan antara keduanya (Hartono, n.d.). Metrik ini sangat berguna ketika terdapat ketidakseimbangan data, karena *F1-Score* mempertimbangkan baik kemampuan model dalam mengidentifikasi data relevan (presisi) maupun kemampuannya dalam menemukan seluruh data relevan (*recall*). Nilai *F1-Score* yang tinggi menunjukkan bahwa model memiliki kinerja yang baik secara keseluruhan.

METODE

JENIS PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksploratif, dengan tujuan untuk menganalisis komentar pengguna terhadap Aplikasi Sehat Indonesiaku yang dapat ditemukan di Google Play Store. Penelitian dilakukan tanpa mengubah variabel, melainkan dengan mengumpulkan data apa adanya dari sumber online, kemudian mengolah dan mengklasifikasikan data tersebut sesuai dengan pola komentar pada Aplikasi Sehat Indonesia.

SUMBER DATA

Data dalam penelitian ini berupa seluruh komentar pengguna terhadap Aplikasi Sehat Indonesiaku yang diambil dari Google Play Store. Sebanyak 1.053 komentar berhasil dikumpulkan melalui proses *web scraping*. Namun, tidak semua komentar digunakan dalam analisis. Pemilihan sampel dilakukan secara sengaja (*purposive*), yaitu hanya komentar yang mengandung teks bermakna

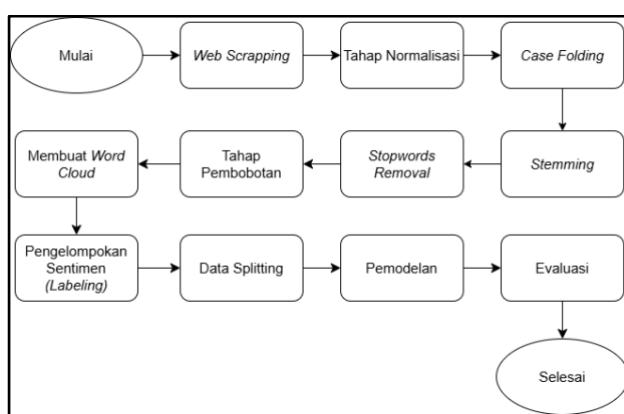
dan dapat dianalisis (tidak kosong atau hanya berisi simbol) yang dijadikan data akhir untuk proses analisis sentimen.

TEKNIK PENGUMPULAN DATA

Data dikumpulkan dengan menggunakan teknik *web scraping*, yaitu metode pengambilan data secara otomatis dari situs web. Dalam penelitian ini, *scraping* dilakukan melalui Python dengan memanfaatkan pustaka *google-play-scrapers* untuk mengambil komentar pengguna dari Aplikasi Sehat Indonesiaku di Google Play Store.

TEKNIK ANALISIS DATA

Proses analisis data dalam penelitian ini dilakukan melalui beberapa tahapan yang saling berurutan dan sistematis. Setiap tahapan dirancang untuk memastikan data yang diambil melalui proses *web scraping* dapat diolah secara tepat hingga menghasilkan klasifikasi sentimen yang akurat. Rangkaian proses tersebut dapat dilihat pada flowchart berikut.



Gambar 1. Flowchart Teknik Analisis Data

Setelah melihat flowchart pada Gambar 1, setiap tahapan dalam proses analisis data dapat dijelaskan sebagai berikut:

WEB SCRAPING

Tahap ini merupakan proses pengambilan data komentar secara otomatis dari Google Play Store menggunakan pustaka *google-play-scrapers*. Data yang dikumpulkan mencakup berbagai kolom seperti *reviewId*, *userName*, *userImage*, *content*, *score*, *thumbsUpCount*, *reviewCreatedVersion*, *at*, *replyContent*, *repliedAt*, dan *appVersion*. Namun, dalam penelitian ini hanya kolom *content* (isi komentar) dan *score* (rating) yang digunakan sebagai variabel dalam analisis sentimen.

TAHAP NORMALISASI

Tahapan normalisasi merupakan tahapan yang bertujuan untuk menyamakan format teks agar lebih konsisten dan mudah dianalisis. Proses ini meliputi beberapa langkah, seperti menghilangkan tanda baca, angka, dan karakter non-huruf lainnya yang tidak memiliki makna semantik dalam konteks analisis sentimen. Karakter non-huruf ini sering muncul dalam komentar pengguna, baik di situs web maupun media sosial, dan keberadaannya perlu dihapus karena tidak memberikan informasi yang relevan dan justru dapat mengganggu proses pemodelan (Idris et al., 2023).

CASE FOLDING

Case folding merupakan tahapan merubah semua huruf dalam komentar menjadi sama (Idris et al., 2023). Pada penelitian ini, semua huruf diubah menjadi huruf kecil untuk menghindari perbedaan makna akibat variasi kapitalisasi. Misalnya, kata "Sehat" dan "sehat" akan diperlakukan sama setelah *case folding*.

STEMMING

Stemming merupakan suatu proses untuk menemukan kata dasar dari sebuah kata. Dengan menghilangkan semua imbuhan (*affixes*) baik yang terdiri dari awalan (*prefixes*), sisipan (*infixes*), akhiran (*suffixes*) dan *confixes* (kombinasi dari awalan dan akhiran) pada kata turunan (Siswandi & Surojudin, 2020). Tujuannya adalah membantu dalam meminimalkan keragaman kata yang sebenarnya mempunyai makna yang sama (Aziz et al., 2024). Hasil dari tahap ini adalah teks yang telah dinormalisasi, di mana kata yang tidak baku diganti dengan yang baku, yang kemudian disusun dalam data frame untuk perbandingan antara teks sebelum dan sesudah normalisasi.

STOPWORDS REMOVAL

Stopword adalah tahapan dimana kata-kata yang tidak mengandung makna penting dan tidak memberikan informasi relevan terhadap topik atau masalah yang sedang dianalisis, biasanya meliputi kata depan, kata hubung, dan kata-kata umum lainnya yang sering muncul dalam teks (Idris et al., 2023). Teks akan dipisah menjadi kata-kata, lalu kata yang termasuk *stopwords* akan dihilangkan. Proses ini bertujuan untuk menyaring kata-kata penting agar analisis dan pemodelan data berjalan lebih efektif.

TAHAP PEMBOBOTAN (TF-IDF)

TF-IDF merupakan metode pembobotan dalam analisis teks yang mengukur pentingnya suatu kata dalam sebuah dokumen berdasarkan frekuensi kemunculannya (TF) dan kelangkaannya di seluruh koleksi dokumen (IDF) (Karan et al., 2025). Tujuan utama dari metode TF-IDF adalah untuk mengevaluasi seberapa relevan atau penting suatu kata dalam sebuah dokumen dalam konteks koleksi dokumen yang lebih besar (Septiani & Isabela, 2022).

MEMBUAT WORD CLOUD

Pada tahap ini, dilakukan visualisasi kata-kata yang paling sering muncul dalam komentar menggunakan *word cloud*. Data TF-IDF digabungkan dengan skor sentimen untuk memisahkan kata-kata pada komentar positif dan negatif, kemudian dibuat *word cloud* masing-masing. Selain itu, dibuat juga *word cloud* untuk seluruh komentar berdasarkan TF-IDF. Visualisasi ini memudahkan identifikasi kata penting dalam data.

PENGELOMPOKAN SENTIMEN (LABELING)

Pada tahap ini, data komentar diklasifikasikan menjadi dua kelompok sentimen, yaitu positif dan negatif, berdasarkan skor yang diberikan. Komentar dengan skor ≥ 4 dikategorikan sebagai sentimen positif, sedangkan skor ≤ 3 dikategorikan sebagai sentimen negatif. Pengelompokan ini memudahkan analisis lebih lanjut terhadap pola dan karakteristik tiap kelompok sentimen.

DATA SPLITTING

Komentar yang telah diberi label dibagi menjadi dua bagian, yaitu data latih dan data uji, untuk memungkinkan model dilatih dan diuji menggunakan data yang berbeda serta mengevaluasi performanya secara objektif. Dalam penelitian ini, digunakan metode *Stratified K-Fold Cross Validation*, yaitu variasi dari *K-Fold Cross Validation* yang data mendistribusikan data secara acak ke dalam k-fold, namun tidak merubah distribusi di seluruh fold (Widodo et al., 2022). Pendekatan ini sangat berguna terutama ketika dataset memiliki ketidakseimbangan kelas, karena menjaga representasi proporsional dari setiap kelas di setiap fold sehingga evaluasi model menjadi lebih adil dan akurat.

PEMODELAN

Tahap ini bertujuan untuk mengembangkan model klasifikasi sentimen berdasarkan data komentar yang sudah melalui proses prapemrosesan dan pelabelan. Penelitian ini menggunakan tiga algoritma *machine learning*, yaitu Random Forest, Gradient Boosting, dan Logistic Regression. Ketiga model tersebut dipilih karena memiliki pendekatan yang beragam, yaitu model ensemble yang didasarkan pada pohon keputusan (Random Forest dan Gradient Boosting) serta model linier sederhana (Logistic Regression).

Random Forest diterapkan karena kemampuannya dalam mengurangi *overfitting* serta menangani banyak fitur. Gradient Boosting dipilih karena tingkat akurasinya yang lumayan tinggi dan kemampuannya untuk memahami pola-pola kompleks secara bertahap. Logistic Regression digunakan karena kemampuannya untuk bekerja dengan baik pada dataset dengan fitur yang sangat banyak. Pemilihan model ini memberikan kesempatan untuk mengevaluasi secara menyeluruh

kinerja klasifikasi, dengan mempertimbangkan akurasi, presisi, serta keseimbangan deteksi di antara berbagai kelas sentimen.

EVALUASI

Model yang telah dilatih akan dievaluasi menggunakan metrik *accuracy*, *precision*, dan *F1-Score*. Akurasi digunakan untuk menghitung tingkat ketepatan klasifikasi suatu dokumen ketika data pada setiap kategori seimbang (Wara et al., 2024). Namun, pada kondisi data yang tidak seimbang, akurasi saja tidak cukup karena model dapat memberikan hasil tinggi dengan lebih mengutamakan kelas mayoritas. Precision mengukur kemampuan model dalam mengidentifikasi komentar positif secara tepat, sedangkan *F1-Score* merupakan gabungan antara *precision* dan *recall* memberikan ukuran yang lebih seimbang (Damayanti et al., 2025)

Dengan menggunakan ketiga metrik ini secara bersamaan, evaluasi model klasifikasi sentimen dapat dilakukan secara komprehensif, memastikan bahwa model tidak hanya akurat secara keseluruhan tetapi juga efektif dalam mengidentifikasi sentimen positif dan negatif dengan tepat.

HASIL DAN PEMBAHASAN

Pada bagian ini, akan disajikan hasil yang diperoleh dari penelitian serta pembahasan terkait dengan analisis yang dilakukan. Berikut ini adalah data yang diperoleh sebelum dilakukan proses *preprocessing*.

Tabel 1. Data Sebelum *Preprocessing*

Content
mohon maaf ini data imunisasi nya tolong kalo udah diinput bulan ini bupan depan nya kok bisa hilang dan harus masukan dari awal lagi padahal datanya sama, dan skrg apk nya malah ga bisa login tolong ya diperbaiki lagi 🙏🙏
bulan kemaren aplikasi asik lancar dipakai! bulan ini knp kok dibuka lgsg layarnya hitam semua...jadi gak bisa masukin data imunisasi... SDH coba restart dan uninstall dan download lagi masih tetap tidak bisa.. mohon dibantu ya.. trimakasih
tolong dong diperbaiki lagi, pelaksanaan posyandu untuk data yg sudah di masukkan oleh salah satu tim kalo bisa bisa di cek oleh anggota lain dari satu tim tsb,

agar tidak ada data ganda. karna ada beberapa data ganda, dan saya bisa melakukan pencatatan lagi data gg sudah dilakukan pencatatan okeh rekan tim saya.. kemudian untuk hasil pencatatan seharusnya setelah pencatatan itu bisa di cek lagi, tapi ini malah tidak ditemukan

...
✍
👍
👎

Setelah melalui tahapan *preprocessing* yang meliputi normalisasi, *case folding*, *stemming*, dan *stopwords removal*, data awal yang terdiri dari 1.053 kolom berhasil diproses dan dikurangi menjadi 723 kolom. Proses ini bertujuan untuk membersihkan dan mempersiapkan data agar lebih siap digunakan dalam analisis selanjutnya. Berikut adalah data yang telah diproses setelah beberapa tahap tersebut.

Tabel 2. Data Sesudah *Preprocessing*

Content
masukan login perbaiki
perbaiki salah
lancar lama lama perbaiki lancar
...
bagus
bagus
bagus

Setelah proses pembersihan kata selesai, langkah selanjutnya adalah melakukan pembobotan dengan metode TF-IDF. Berikut ini adalah hasil pembobotan yang diperoleh.

Tabel 3. TF-IDF

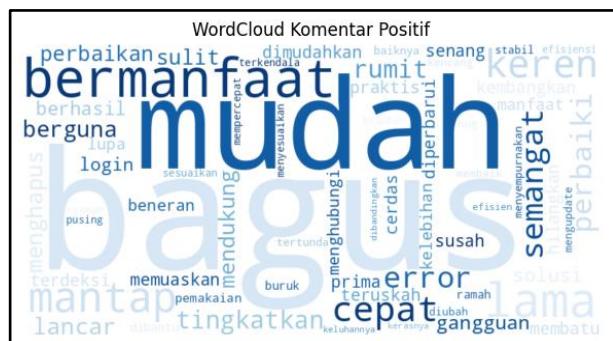
Kata	Jumlah_TFIDF

lama	93,2869
bagus	75,2426
mudah	61,7038
...	...
keluhannya	0,3780
kerasnya	0,3780
dibandingkan	0,3780

Berdasarkan Tabel 3 dapat dilihat bahwa kata "lama", "bagus", dan "mudah" memiliki nilai TF-IDF yang paling tinggi, sehingga dianggap paling relevan dalam dokumen. Sementara itu, kata-kata seperti "keluhannya", "kerasnya", dan "dibandingkan" memiliki nilai TF-IDF yang rendah, yang menunjukkan bahwa kata-kata tersebut kurang penting dalam konteks keseluruhan teks. Hasil ini menunjukkan bahwa metode TF-IDF dapat membantu mengenali kata-kata yang paling berpengaruh dalam analisis teks.

Sebagai lanjutan dari analisis TF-IDF, visualisasi word cloud disusun untuk menggambarkan kata-kata yang paling dominan dalam komentar. Visualisasi ini dibagi berdasarkan sentimen positif, negatif, dan keseluruhan data, sehingga dapat mempermudah dalam mengidentifikasi pola kemunculan kata berdasarkan jenis sentimennya.

WORD CLOUD KOMENTAR POSITIF

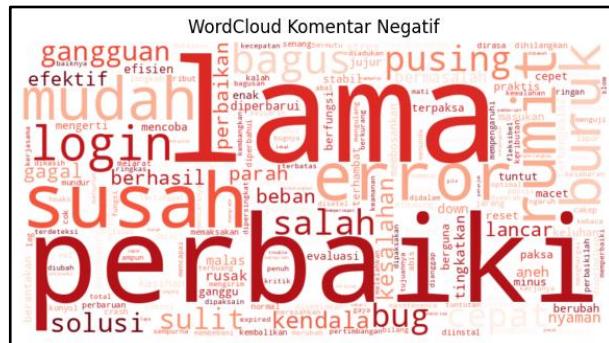


Gambar 2. Word Cloud Komentar Positif

Word cloud komentar positif menunjukkan bahwa kata-kata seperti "mudah", "bagus", "bermanfaat", "cepat", dan "berguna" muncul dengan frekuensi tinggi dalam ulasan positif pengguna. Ini mengindikasikan bahwa aplikasi Sehat Indonesiaku dinilai memudahkan pengguna dalam mengakses

layanan kesehatan digital. Pengguna juga merasa aplikasi ini berguna dan memberi pengalaman yang menyenangkan, terutama dari sisi kemudahan penggunaan, kecepatan layanan, dan dukungan fitur yang membantu kebutuhan mereka.

WORD CLOUD KOMENTAR NEGATIF



Gambar 3. Word Cloud Komentar Negatif

Word cloud komentar negatif menampilkan kata-kata dominan seperti “perbaiki”, “lama”, “susah”, “login”, dan “error”, yang mencerminkan berbagai keluhan dari pengguna. Sebagian besar komentar negatif berkaitan dengan masalah teknis, seperti kesulitan saat masuk ke dalam aplikasi (login), performa aplikasi yang lambat, hingga adanya gangguan atau bug. Kata “perbaiki” yang paling menonjol menunjukkan adanya harapan dari pengguna agar aplikasi ini segera diperbaiki dan ditingkatkan kualitas layanannya.

WORD CLOUD KESELURUHAN KOMENTAR



Gambar 4 Word Cloud Keseluruhan Komentar

Word cloud seluruh komentar menampilkan kombinasi kata-kata dari seluruh ulasan pengguna, baik yang bersentimen positif maupun negatif. Penggunaan kata “bagus” dan “mudah” menunjukkan bahwa sebagian pengguna merasa terbantu dengan fitur dan aksesibilitas aplikasi. Namun, kehadiran kata seperti “lama”, “susah”, “perbaiki”, dan “error” menyoroti adanya keluhan

terhadap performa teknis aplikasi, khususnya terkait kecepatan, gangguan teknis, dan pengalaman login. *Word cloud* ini mencerminkan kompleksitas persepsi publik, di mana pengalaman positif dan negatif saling berdampingan sebagai cerminan nyata dari tanggapan pengguna terhadap layanan digital pemerintah.

Berdasarkan penelusuran komentar, banyak keluhan negatif muncul setelah pengguna melakukan pembaruan aplikasi. Kata-kata seperti "lama", "susah", dan "error" sering kali dikaitkan dengan versi terbaru aplikasi. Hal ini menunjukkan bahwa pembaruan justru menimbulkan pengalaman pengguna yang kurang memuaskan, yang berdampak langsung pada persepsi publik terhadap kualitas layanan digital ini.

Setelah melakukan analisis menggunakan *word cloud*, pada penelitian ini juga dilakukan pemodelan dengan tiga teknik *machine learning*. Berikut adalah hasil dari ketiga pemodelan tersebut:

Tabel 4. Hasil Evaluasi Pemodelan

Model	Accuracy	Precision	F1-Score
Random Forest	0,8769	0,8730	0,8718
Gradient Boosting	0,8783	0,8747	0,8731
Logistic Regression	0,8727	0,8716	0,8636

Berdasarkan Tabel 4, Gradient Boosting menunjukkan kinerja terbaik dalam mengklasifikasikan sentimen komentar dari aplikasi Sehat Indonesiaku. Hal ini dapat dilihat dari nilai accuracy sebesar 0,8783, precision 0,8747, dan F1-Score 0,8731. Nilai ini menunjukkan bahwa model ini tidak hanya tepat, tetapi juga seimbang dalam mengenali komentar positif dan negatif, sehingga dapat diandalkan untuk kegiatan klasifikasi pada data teks singkat seperti komentar pengguna.

Model Random Forest menempati urutan kedua dengan accuracy 0,8769, precision 0,8730, dan F1-Score 0,8718. Meskipun sedikit tertinggal dibandingkan dengan Gradient Boosting, perbedaannya tidak terlalu signifikan, dan model ini tetap menunjukkan performa yang sangat baik serta dapat dipertimbangkan sebagai pilihan model yang kuat.

Sementara itu, Logistic Regression menunjukkan hasil terendah di antara ketiga metode, dengan accuracy 0,8727, precision 0,8716,

dan F1-Score sebesar 0,8636. Meskipun kinerjanya masih berada dalam batas yang dapat diterima, model ini kurang efektif dalam mengelola kompleksitas data komentar dibandingkan dengan dua model lainnya.

Secara keseluruhan, dari ketiga model yang telah diuji, Gradient Boosting dipilih sebagai model yang paling unggul karena memberikan nilai metrik evaluasi tertinggi serta paling seimbang dalam mengklasifikasikan data sentimen dengan akurat dan konsisten.

PENUTUP

SIMPULAN

Penelitian ini bertujuan untuk mengembangkan sistem analisis sentimen pada komentar pengguna Aplikasi Sehat Indonesiaku dengan memanfaatkan web scraping untuk pengumpulan data. Visualisasi word cloud menunjukkan bahwa komentar positif didominasi oleh kata-kata seperti "mudah", "bagus", dan "bermanfaat", yang menggambarkan tingkat kepuasan pengguna terhadap kemudahan serta manfaat yang diberikan oleh aplikasi tersebut. Sebaliknya, komentar negatif lebih banyak mengandung kata-kata seperti "perbaiki", "lama", dan "error", yang mencerminkan keluhan terkait masalah teknis, khususnya setelah pembaruan aplikasi.

Selanjutnya, untuk klasifikasi sentimen, diterapkan tiga metode machine learning, yaitu Random Forest, Gradient Boosting, dan Logistic Regression. Berdasarkan hasil evaluasi, model Gradient Boosting menunjukkan performa terbaik dengan tingkat akurasi sebesar 87,83%, menjadikannya model yang paling tepat dalam mengklasifikasikan komentar positif dan negatif. Model ini juga menghasilkan nilai precision dan F1-Score tertinggi, yang menjadikannya pilihan yang lebih konsisten dan akurat dalam proses klasifikasi sentimen.

Temuan penelitian ini memberikan pemahaman yang lebih mendalam mengenai persepsi pengguna serta memberikan rekomendasi perbaikan untuk layanan aplikasi kesehatan digital yang dikembangkan oleh pemerintah.

SARAN

Berdasarkan hasil penelitian, disarankan agar pengembang aplikasi Sehat Indonesiku memberikan perhatian khusus pada perbaikan teknis yang menjadi sumber utama keluhan pengguna, terutama yang muncul pasca pembaruan aplikasi. Pengembang juga perlu meningkatkan komunikasi dengan pengguna, misalnya melalui penyediaan kanal pelaporan masalah dan respon cepat terhadap keluhan umum seperti kesulitan login dan performa lambat. Selanjutnya, pengembangan model klasifikasi sentimen dapat diperluas dengan menguji algoritma lain yang lebih canggih seperti deep learning untuk meningkatkan akurasi dan kemampuan menangani konteks bahasa Indonesia yang kompleks. Untuk penelitian mendatang, disarankan menggabungkan analisis waktu (*temporal analysis*) untuk melacak perubahan sentimen seiring waktu dan melihat secara lebih spesifik dampak dari versi tertentu aplikasi terhadap persepsi pengguna

DAFTAR PUSTAKA

- Aziz, R., Fahrudin, T. M., & Saputra, W. S. J. (2024, April). Analisis Sentimen Kepuasan Pengguna OYO DiPlaystore Dengan Multinomial Naive Bayes dan Chi-square. *JURNAL FASILKOM*, 14(1). <https://pdfs.semanticscholar.org/1c75/f1525e143ab7ba41ff5ba82c05960ec4ea0c.pdf>
- Bintoro, P., Ratnasari, Wihardjo, E., Putri, I. P., & Asari, A. (2024). Pengantar Machine Learning. PT MAFY MEDIA LITERASI INDONESIA. <https://repository.um.ac.id/5619/1/fulltexts.pdf>
- Damayanti, E., Rahatingsih, N., Dana, R. D., & Rohmat, C. L. (2025). Analisis Sentimen Pengguna Aplikasi Mobile Jaminan Kesehatan Nasional di Playstore Menggunakan Metode Naive Bayes. *Jurnal INTEK (Informatika dan Teknologi Informasi)*, 8(1). <https://jurnal.umpwr.ac.id/intek/article/view/5770>
- Dwicahyo, K., & Ratnasari, C. I. (2023, 11 14). Perbandingan Metode Web Scraping Dalam Pengambilan Data: Kajian Literatur. 4(2). <https://journal.uii.ac.id/AUTOMATA/article/view/28635>
- Fahmuddin, M., Aidid, M. K., & Taslim, M. J. (2023, 12). IMPLEMENTASI ANALISIS REGRESI LOGISTIK DENGAN METODE MACHINE LEARNING UNTUK MENKLASIFIKASI BERITA DI INDONESIA,. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 5(3). <https://jurnalvariansi.unm.ac.id/index.php/variansi/article/view/116>
- Findawati, Y., & Rosid, M. A. (2020). BUKU AJAR TEXT MINING. UMSIDA PRESS. <https://press.umsida.ac.id/index.php/umsidapress/article/view/978-623-6833-19-3/770>
- Hartono. (n.d.). Modul Digital Machine Learning. https://lmssspada.kemdiktisaintek.go.id/pluginfolder.php/795078/mod_resource/content/2/Modul%20Digital%20-Machine%20Learning%20-%20Hartono-compressed.pdf
- Idris, I. S. K., & Mustofa, Y. A. (2022). Typo Checking Menggunakan Algoritma Rabin-Karp. *Jambura Journal of Electrical and Electronics Engineering*, 4(1). <https://ejurnal.ung.ac.id/index.php/jjeee/article/view/12150>
- Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023, 1). Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM). *Jambura Journal of Electrical and Electronics Engineering*, 5(1). <https://ejurnal.ung.ac.id/index.php/jjeee/article/view/16830/5678>
- Karan, Alidin, M., Fadilah, R., & Taslim. (2025). Penerapan K-Nearest Neighbors untuk Klasifikasi Kasus Hukum di Pengadilan Federal Australia. *Jurnal Computer Science and Information Technology (CoSciTech)*, 6(1).
- Kusuma, I. N., & Ali, I. (2024, 4). ANALISIS SENTIMEN PADA PENGGUNA APLIKASI DANA MENGGUNAKAN ALGORITMA NAIVE BAYES. *Jurnal Mahasiswa Teknik Informatika*, 8(2). <https://ejournal.itn.ac.id/index.php/jati/article/view/9041>

- Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(9). <https://jptiik.ub.ac.id/index.php/jptiik/article/view/11562>
- Muntiari, N. R., Hanif, K. H., & Nisa, I. C. (2023). Perbandingan Algoritma Regresi Logistik, Support Vector Machine, dan Gradient Boosting Pada Analisis Sentimen Data Komentar Siswa. *JURNAL ILMU KOMPUTER DAN TEKNOLOGI*, 4(2). https://ejournal.uhb.ac.id/index.php/IKO_MTI/article/view/1286
- Rahayu, P. W., Sudipa, I. G. I., Suryani, Surachman, A., Ridwan, A., Darmawiguna, I. G. M., Sutoyo, M. N., Slamet, I., Harlina, S., & Maysanjaya, I. M. D. (2024). Buku Ajar Data Mining. PT. Sonpedia Publishing Indonesia. https://www.researchgate.net/publication/377415198_BUKU_AJAR_DATA_MINING
- Rahmawati, I., & Fitriani, T. R. (2023). Analisis Sentimen Menggunakan Algoritma Logistic Regression Pada Penerbangan Lion Air berdasarkan Ulasan Pengguna Platform Online. *Jurnal Riset Informatika dan Teknologi Informasi (JRITI)*, 1(1). <https://ejurnal.jejarungppm.org/index.php/jriti/article/view/60>
- Septiani, D., & Isabela, I. (2022). ANALISIS TERM FREQUENCY INVERSE DOCUMENT FREQUENCY (TF-IDF) DALAM TEMU KEMBALI INFORMASI PADA DOKUMEN TEKS. *SINTESIA: Jurnal Sistem dan Teknologi Informasi Indonesia*, 1(2).
- Siregar, M. Y., Wiranata, A. D., & Saputra, R. A. (2024, 4). Analisis Sentimen Pada Ulasan Pengguna Aplikasi Streaming Vedio Menggunakan Metode Naïve Bayes. *Kajian Ilmiah Informatika dan Komputer*, 4(5). <https://djournals.com/klik/article/view/1787>
- Siswandi, A., & Surojudin, N. (2020). ANALISIS DAN PERBANDINGAN STEMMING ALGORITMA PORTER DENGAN ALGORITMA AHMAD YUSOFF SEMBOK DALAM DOKUMEN TEKS BAHASA INDONESIA. Seminar Nasional Teknologi Informasi dan Komunikasi STI&K (SeNTIK), 4(1). <https://ejurnal.jakstik.ac.id/files/journals/2/articles/sentik2020/324/submission/proof/324-13-1121-1-10-20201101.pdf>
- Wara, S. S. M., Adziima, A. F., Nasrudin, M., & Pratama, A. R. (2024). Predictive Analysis of Government Application Comment on Playstore with Clustered Support Vector Machine. 2024 IEEE 10th Information Technology International Seminar (ITIS). <https://doi.org/10.1109/ITIS64716.2024.10845453>
- Widodo, S., Brawijaya, H., & Samudi. (2022). Stratified K-fold cross validation optimisation on machine learning for prediction. *Sinkron : Jurnal dan Penelitian Teknik Informatika*, 6(4). <https://doi.org/10.33395/sinkron.v7i4.11792>