

Analisis Sentimen Pengguna terhadap Aplikasi SatuSehat pada Platform Google Play Store Menggunakan Model IndoBERT

Mukhtarul Hadi^{*1}, Budi Prasetyo², Regina Ayumi Ulayyaa³

^{1,2,3}Jurusan Ilmu Komputer, FMIPA UNNES, Semarang

e-mail: *mukhtarulhadi@students.unnes.ac.id, bprasetyo@mail.unnes.ac.id,

reginaayumi@students.unnes.ac.id

Abstrak

Penelitian ini menganalisis sentimen 1.048 ulasan pengguna aplikasi SatuSehat di Google Play Store menggunakan model IndoBERT (mdhugol/indonesia-bert-sentiment-classification) yang di-pra-fine-tune untuk konteks kesehatan digital. Setelah tahap preprocessing (lowercasing, penghapusan karakter non-alfabet, tokenisasi, dan penghilangan stopword—kecuali kata negasi seperti “tidak”, “bukan”, “gagal”, “error”), model mengklasifikasikan 87,3% ulasan sebagai negatif, 8,4% netral, dan 4,3% positif. Analisis n-gram mengungkap tiga frasa negatif teratas—“tolong diperbaiki” (58 kali), “scan barcode” (49 kali), dan “kode OTP” (26 kali)—yang menyoroti permasalahan teknis mendesak. Di sisi lain, frasa positif meliputi “terima kasih” (4 kali), “bagus cek” (2 kali), dan “tekanan darah” (2 kali), menandakan apresiasi pengguna terhadap fitur pemantauan kesehatan tertentu. Confidence score rata-rata untuk ulasan negatif mencapai 0,96, sedangkan ulasan netral dan positif masing-masing rata-ratanya 0,82 dan 0,87, menunjukkan model paling yakin dalam memprediksi keluhan teknis. Berdasarkan hasil ini, direkomendasikan agar pengembang SatuSehat segera memprioritaskan perbaikan fitur login/OTP dan pemindaian barcode, serta mengembangkan lebih lanjut fitur monitoring tekanan darah untuk meningkatkan kepuasan pengguna.

Kata kunci—Analisis Sentimen, SatuSehat, IndoBERT, Tenaga Kesehatan, mHealth

Abstract

This study analyzes the sentiment of 1,048 user reviews of the SatuSehat app on the Google Play Store using the IndoBERT (mdhugol/indonesia-bert-sentiment-classification) model pre-fine-tuned for the digital health context. After preprocessing (lowercasing, non-alphabetic character removal, tokenization, and stopword removal-except for negation words such as “no”, “not”, “fail”, “error”), the model classified 87.3% of reviews as negative, 8.4% as neutral, and 4.3% as positive. The n-gram analysis revealed the top three negative phrases—“please fix” (58 times), “scan barcode” (49 times), and “OTP code” (26 times)—which highlighted pressing technical issues. On the other hand, positive phrases include “thank you” (4 times), “good check” (2 times), and “blood pressure” (2 times), signaling users’ appreciation of certain health monitoring features. The average confidence score for negative reviews reached 0.96, while neutral and positive reviews averaged 0.82 and 0.87 respectively, indicating the model was most confident in predicting technical complaints. Based on these results, it is recommended that SatuSehat developers immediately prioritize improvements to the login/OTP and barcode scanning features, and further develop the blood pressure monitoring feature to increase user satisfaction.

Keywords—Sentiment Analysis, SatuSehat, IndoBERT, Health workers, mHealth

1. PENDAHULUAN

Analisis sentimen telah menjadi komponen krusial dalam kajian text mining untuk mengevaluasi pendapat publik terhadap aplikasi kesehatan digital, karena ulasan pengguna sering kali mencerminkan pengalaman dan kepuasan yang dapat memengaruhi adopsi aplikasi tersebut [1], [2]. Sejak pandemi COVID-19, penggunaan *mobile health* (mHealth) melonjak tajam, dan data ulasan di media sosial atau *app store* menyediakan informasi *real-time* mengenai kritik dan puji yang dialamatkan ke aplikasi kesehatan [3]. Dengan memanfaatkan teknik NLP, peneliti dapat mengelompokkan ulasan menjadi beberapa sentimen—positif, netral, maupun negatif—sehingga penyedia layanan mHealth dapat mengetahui area mana yang membutuhkan perbaikan atau peningkatan fitur [2], [3]. Selain itu, analisis sentimen memungkinkan pengambilan keputusan berbasis data untuk mengoptimalkan fungsi aplikasi, meningkatkan retensi pengguna, dan merancang strategi komunikasi yang lebih efektif dengan pengguna [1], [3].

Aplikasi SatuSehat resmi menggantikan PeduliLindungi pada 1 Maret 2023 sebagai *platform* terintegrasi oleh Kementerian Kesehatan RI untuk membantu masyarakat memantau status vaksinasi dan riwayat kesehatan secara digital [4]. Berbeda dengan *contact tracing* murni, SatuSehat mengombinasikan fitur rekam medis elektronik dan integrasi data vaksinasi, sehingga cakupan fungsinya lebih luas dibandingkan pendahulunya [4]. Meskipun telah banyak penelitian yang menganalisis ulasan PeduliLindungi menggunakan metode *lexicon-based* maupun *machine learning* klasik seperti Naïve Bayes dan SVM dengan berbagai tingkat akurasi (75%–91%), studi mendalam terkait persepsi pengguna SatuSehat masih sangat terbatas [5], [6]. Selain itu, implementasi semacam itu sebagian besar berfokus pada data Twitter atau media sosial lain, bukan ulasan resmi di Google Play Store yang cenderung mencakup keluhan teknis dan pengalaman penggunaan aplikasi langsung [5], [7]. Oleh karena itu, analisis sentimen pada ulasan SatuSehat di Google Play Store menjadi langkah penting untuk memahami isu-isu utama pengguna seperti kesulitan *login*, *bug* pembaruan, atau sinkronisasi data kesehatan.

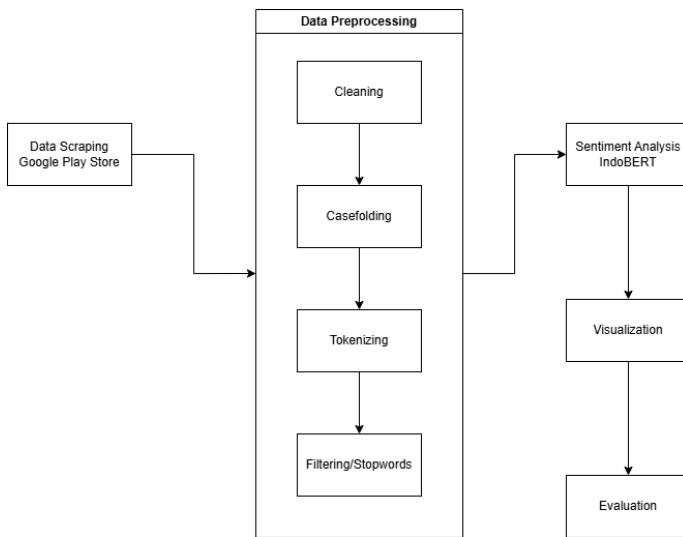
Sejak diperkenalkannya IndoBERT—varian BERT khusus bahasa Indonesia—oleh Koto & Rahimi, banyak penelitian yang menunjukkan keunggulan model ini dalam tugas klasifikasi sentimen teks berbahasa Indonesia dibandingkan model berbasis multilingual. Misalnya, Singgalen (2025) melaporkan bahwa *fine-tuning* IndoBERT pada *dataset* ulasan hotel menghasilkan akurasi 92,52% dan F1-score 96,09% untuk kelas positif, menunjukkan performa unggul meski menghadapi ketidakseimbangan kelas [8]. Perwira et al. (2025) menerapkan IndoBERT khusus *domain* untuk analisis sentimen aspek pada ulasan pariwisata, memperoleh akurasi klasifikasi sebesar 84% pada deteksi aspek dan kelas sentimen utama [9]. Selain itu, Asri et al. (2023) memadukan IndoBERT dengan *lexicon* Indonesia untuk 1.000 ulasan PLN dan mendapatkan akurasi hingga 81% pada dataset yang relatif kecil, membuktikan bahwa model ini dapat diadaptasi untuk berbagai konteks aplikasi layanan publik [10]. Penelitian lain seperti Saputra et al. (2023) juga menunjukkan bahwa IndoBERT efektif dalam prediksi emosi bahasa Indonesia dengan akurasi 73%, meski fokusnya pada kategori emosi, yang menunjukkan fleksibilitas IndoBERT di ranah analisis bahasa Indonesia [11].

Meskipun beberapa penelitian telah memanfaatkan IndoBERT untuk ulasan aplikasi kesehatan atau layanan publik, hingga saat ini belum ada studi yang secara khusus menerapkan IndoBERT untuk mengkaji ulasan SatuSehat di Google Play Store [12]. Penelitian terdahulu yang berfokus pada PeduliLindungi cenderung menggunakan metode *machine learning* klasik atau *lexicon-based* pada data Twitter, tanpa mengeksplorasi konteks tersendiri dari Google Play Store yang mencakup komentar teknis terkait fungsionalitas aplikasi [13], [14]. Selain itu, studi aspek-pariwisata dan hotel menggunakan IndoBERT belum menangani karakteristik ulasan pengguna SatuSehat yang unik, seperti istilah medis, kode vaksinasi, dan masalah sinkronisasi data kesehatan [8], [9]. Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan melakukan analisis sentimen terhadap 1.048 ulasan SatuSehat menggunakan model IndoBERT yang telah terlatih khusus, sehingga diharapkan memberikan wawasan lebih mendalam mengenai persepsi serta keluhan pengguna yang spesifik pada aplikasi ini.

Penelitian ini akan mengumpulkan data 1.048 ulasan SatuSehat dari Google Play Store sebagai *input*, kemudian menjalankan proses *preprocessing* yang mencakup *lowercasing*, penghapusan karakter non-alfabet, dan penghilangan *stopword* dengan menggunakan pustaka NLTK, sesuai praktik yang umum diterapkan dalam penelitian Sentiment Analysis berbahasa Indonesia [12], [15], [16]. Setelah pembersihan data, setiap ulasan akan di-tokenisasi dan disiapkan dalam format yang sesuai untuk *fine-tuning* IndoBERT (model mdhugol/indonesia-bert-sentiment-classification) menggunakan pustaka HuggingFace Transformers; pelatihan dan validasi dilakukan dengan pembagian data *train-test* sebesar 80:20, di mana metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score* akan dihitung untuk menilai kualitas klasifikasi [8]. Hasil klasifikasi kemudian akan divisualisasikan melalui diagram batang distribusi sentimen, *wordcloud* untuk kata dominan di masing-masing kelas sentimen, serta grafik sebaran *confidence score* menggunakan pustaka Matplotlib dan Seaborn untuk memberikan ilustrasi menyeluruh tentang pola sentimen pengguna [8]. Dengan demikian, pembaca akan mendapatkan gambaran komprehensif mengenai bagaimana model IndoBERT memahami nuansa ulasan SatuSehat dan aspek-elemen penting yang sering dibahas pengguna, sehingga hasil penelitian dapat menjadi rujukan bagi pengembang aplikasi untuk perbaikan fitur serta strategi komunikasi yang lebih tepat sasaran dalam meningkatkan *user experience*.

2. METODE PENELITIAN

Penelitian ini dilakukan dalam empat tahapan utama: (1) pengumpulan data ulasan pengguna, (2) *preprocessing* teks, (3) klasifikasi sentimen menggunakan IndoBERT, dan (4) visualisasi serta evaluasi hasil. Secara garis besar, alur kerja diawali dengan mengambil data mentah dari Google Play Store, membersihkan dan mempersiapkannya agar sesuai untuk analisis, kemudian mengklasifikasikan setiap ulasan menggunakan model IndoBERT (mdhugol/indonesia-bert-sentiment-classification) yang sudah di-pra-fine-tune untuk konteks aplikasi kesehatan digital. Setelah klasifikasi, hasil dikaji melalui teknik distribusi persentase, visualisasi *wordcloud*, analisis n-gram, dan sebaran nilai *confidence score*. Desain keseluruhan metode dapat dilihat pada Gambar 1.



Gambar 1 Flowchart Sentimen Analisis

2. 1 Pengumpulan Data

Proses pengumpulan menggunakan Google Play Scraper untuk mengekstrak seluruh ulasan pengguna SatuSehat di Google Play Store. Dari awal scraping dihasilkan sekitar 1.200 baris data mentah. Selanjutnya, dilakukan pembersihan awal dengan menghapus entri duplikat dan baris yang kosong, sehingga tersisa 1.048 ulasan. Peneliti kemudian memilih hanya dua kolom penting—"content" (isi ulasan) dan "rating" (skor bintang)—untuk menyederhanakan analisis berikutnya. Dataset akhir disimpan dalam format CSV dengan struktur sederhana, memudahkan penerapan tahapan *preprocessing*.

2. 2 Preprocessing

Tahap *preprocessing* dilakukan untuk membersihkan data teks ulasan dari elemen-elemen yang tidak relevan dan menyiapkannya agar optimal digunakan dalam pemodelan. Proses dimulai dengan mengubah seluruh teks ulasan menjadi huruf kecil (*lowercasing*), diikuti dengan penghapusan karakter-karakter yang tidak diperlukan seperti URL, angka, tanda baca, serta simbol atau karakter khusus lainnya. Selanjutnya, dilakukan tokenisasi untuk memisahkan setiap kata dalam kalimat. Setelah itu, *stopword* dihapus menggunakan daftar *stopword* dari pustaka NLTK yang telah dikombinasikan dengan daftar kata umum tambahan yang sering muncul namun tidak membawa makna penting, seperti kata "app", "satusehat", "kesehatan", dan sebagainya. Perlu dicatat bahwa beberapa kata penting seperti negasi ("tidak", "bukan", "jangan") sengaja dipertahankan karena berperan penting dalam menentukan makna kalimat dan sentimennya. Setelah semua proses ini dilakukan, teks hasil pembersihan disimpan dalam kolom baru untuk dianalisis lebih lanjut oleh model IndoBERT.

Tabel 1 Perbandingan Sebelum dan Sesudah Preprocessing

Sebelum Preprocessing dan Sesudah Preprocessing	
Sebelum	aplikasi jelek, pas zaman nya covid pernah daftar, tp skrng mau masuk lagi malah data nya ilang, cuma email dan nomor hp yang masih terverifikasi, giliran udah masuk cek profil malah kosong, data nya, mau di perbaharui gak bisa
Sesudah	jelek pas zaman covid daftar skrng masuk data ilang email nomor terverifikasi giliran udah masuk cek profil kosong data perbaharui

Sebagai ilustrasi, Tabel 1 memperlihatkan bagaimana proses *preprocessing* mengubah ulasan mentah yang panjang, berisi tanda baca, dan kata-kata pengisi menjadi teks yang lebih ringkas dan fokus pada kata kunci bermakna. Sebelum *preprocessing*, ulasan memuat frasa seperti “aplikasi jelek, pas zaman nya covid pernah daftar, tp skrng mau masuk lagi malah data nya ilang, cuma email dan nomor hp yang masih terverifikasi, giliran udah masuk cek profil malah kosong, data nya, mau di perbaharui gak bisa”—di mana terdapat banyak kata hubung, ungkapan waktu, dan tanda baca yang hanya menambah *noise*. Setelah *preprocessing*, teks berubah menjadi “jelek pas zaman covid daftar skrng masuk data ilang email nomor terverifikasi giliran udah masuk cek profil kosong data perbaharui”, di mana semua elemen non-alfabet dan *stopword* yang tidak esensial telah dihapus (tetapi mempertahankan kata-kata negasi seperti “gak”/“tidak” untuk menjaga konteks negatif). Dengan menghilangkan kata-kata seperti “aplikasi”, “nya”, “yang”, dan tanda baca, proses ini memastikan bahwa model IndoBERT hanya menerima input berupa kata-kata kunci yang benar-benar mewakili sentimen pengguna—misalnya “jelek”, “data ilang”, “cek profil kosong”—sehingga meningkatkan akurasi klasifikasi dan meminimalkan gangguan dari *noise* teks.

2. 3 Analisis Sentimen dengan IndoBERT

Analisis sentimen dilakukan menggunakan model IndoBERT yang sudah dilatih sebelumnya, yaitu mdhugol/indonesia-bert-sentiment-classification dari HuggingFace Transformers. Model ini mampu mengklasifikasikan sentimen ke dalam tiga kategori: positif, netral, dan negatif. Proses klasifikasi dilakukan dengan pendekatan *pipeline* dari Transformers. Setiap ulasan dianalisis satu per satu dengan pemrosesan batch untuk efisiensi. *Threshold confidence* disesuaikan (misalnya 0.6) agar hasil klasifikasi tetap diperoleh meskipun skor kepercayaan rendah. Label sentimen dan confidence score hasil klasifikasi disimpan dalam *dataframe* untuk keperluan visualisasi dan evaluasi.

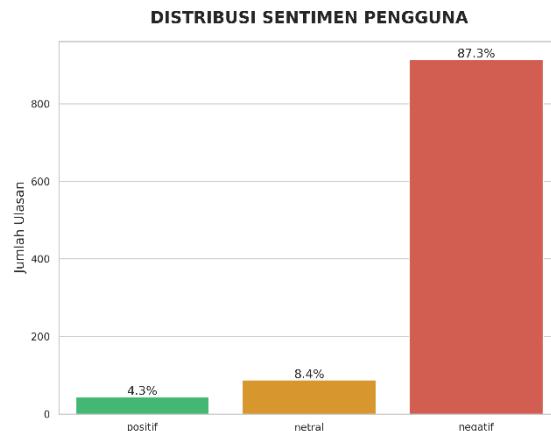
2. 4 Visualisasi dan Analisis Lanjutan

Setelah proses klasifikasi selesai, hasilnya dianalisis dan divisualisasikan untuk memberikan gambaran yang lebih jelas mengenai pola sentimen yang muncul. Visualisasi utama berupa diagram batang digunakan untuk menunjukkan distribusi sentimen berdasarkan jumlah ulasan yang tergolong ke dalam kategori positif, netral, dan negatif. Selain itu, visualisasi berbentuk wordcloud digunakan untuk menggambarkan kata-kata yang paling sering muncul dalam masing-masing kelompok sentimen, yang dapat memberikan indikasi topik dominan atau isu yang sering dibahas oleh pengguna. Wordcloud ini dibuat terpisah untuk ulasan positif dan negatif. Di sisi lain, untuk menilai kepercayaan model terhadap prediksi yang dihasilkan, sebaran *confidence score* dari setiap prediksi ditampilkan dalam bentuk grafik KDE (*Kernel Density Estimation*). Analisis ini juga dilengkapi dengan penerapan metode n-gram seperti *unigram* dan *bigram* untuk menelusuri frasa atau kombinasi kata yang sering muncul, yang dapat mengungkap pola keluhan atau puji yang lebih spesifik dalam ulasan pengguna.

3. HASIL DAN PEMBAHASAN

3. 1 Distribusi Sentimen

Hasil klasifikasi sentimen pada 1.048 ulasan pengguna aplikasi SatuSehat menggunakan model IndoBERT menunjukkan bahwa 87,3% ulasan bersifat negatif (915 ulasan), 8,4% ulasan netral (88 ulasan), dan 4,3% ulasan positif (45 ulasan) seperti yang terlihat pada Gambar 2 dan Tabel 2.



Gambar 2 Grafik Distribusi Sentimen Pengguna

Tabel 2 Distribusi Data Sentimen

Sentimen	Data
Positif	45
Netral	88
Negatif	915
Total	1048

Proporsi ini merefleksikan dominasi kritik teknis dalam pengalaman pengguna, di mana mayoritas merasa fitur inti aplikasi—seperti proses *login*, verifikasi OTP, dan pemindai *barcode*—belum berjalan sesuai harapan. Sebagian kecil ulasan netral (8,4%) biasanya berisi pernyataan deskriptif tanpa penekanan emosional kuat, misalnya “aplikasi telah diperbarui, tetapi masih butuh penyempurnaan”. Sementara itu, ulasan positif (4,3%) menandakan bahwa hanya segelintir pengguna yang merasa puas, seringkali berkaitan dengan fungsi monitoring data kesehatan yang berjalan lancar, seperti tampilan tekanan darah atau notifikasi vaksinasi yang tepat waktu.

3. 2 Wordcloud Kata Dominan

Wordcloud digunakan untuk mengidentifikasi kata-kata yang paling sering muncul dalam setiap kategori sentimen. Visualisasi ini memberikan gambaran visual yang cepat dan intuitif mengenai topik-topik yang dominan di mata pengguna. Untuk ulasan dengan sentimen negatif (Gambar 3), *wordcloud* memperlihatkan berbagai kata yang sering dikaitkan dengan keluhan, seperti kata-kata yang menggambarkan masalah teknis, performa aplikasi, atau fitur yang tidak berfungsi.



Gambar 3 Wordcloud Kata Negatif



Gambar 4 Wordcloud Kata Positif

Di sisi lain, *wordcloud* dari ulasan positif (Gambar 4) cenderung menampilkan kata-kata apresiatif atau bentuk ungkapan terima kasih dari pengguna. *Wordcloud* ini tidak hanya berfungsi sebagai ilustrasi frekuensi kata, tetapi juga sebagai alat eksplorasi topik yang memungkinkan peneliti dan pengembang aplikasi untuk menangkap pola-pola utama dalam persepsi pengguna tanpa harus membaca seluruh data secara manual.

3. 3 Sebaran Confidence Score

Setelah ulasan diklasifikasikan oleh model IndoBERT (mdhugol/indonesia-bert-sentiment-classification), setiap prediksi dilengkapi dengan nilai *confidence score* yang mencerminkan keyakinan model dalam menentukan kategori sentimen. Untuk ulasan negatif, nilai *confidence score* berkisar antara 0.412 hingga 0.9995, dengan rata-rata sebesar 0.9632 dan median 0.9966, menandakan sebagian besar prediksi negatif memiliki keyakinan yang sangat tinggi (banyak skor > 0.98). Sebaliknya, ulasan netral menunjukkan rentang *confidence score* yang lebih lebar, yakni 0.456 hingga 0.9989, dengan rata-rata 0.8163 dan median 0.8568, yang menggambarkan adanya beberapa prediksi netral dengan tingkat ketidakpastian cukup tinggi (skor di bawah 0.70) sebelum *threshold*. Untuk ulasan positif, *confidence score* berada di antara 0.566 dan 0.9968, dengan rata-rata 0.8694 dan median 0.9174, menunjukkan bahwa meski jumlahnya lebih sedikit, sebagian besar prediksi positif juga memiliki keyakinan moderat hingga tinggi.



Gambar 5 Grafik Sebaran Confidence Score

Secara keseluruhan, grafik KDE pada Gambar 5 memperlihatkan puncak distribusi untuk kelas negatif di dekat skor 1.0, sedangkan kelas netral dan positif menampilkan distribusi yang lebih melebar di rentang 0.50–0.90. Temuan ini menandakan bahwa model cenderung paling yakin ketika memprediksi ulasan negatif, sedangkan untuk prediksi netral dan positif terdapat lebih banyak ketidakpastian yang perlu diwaspadai, terutama ketika *confidence score* berada di bawah threshold 0.60.

3. 4 Analisis Frasa Umum (*n*-gram)

Analisis *n*-gram dilakukan untuk mengidentifikasi frasa-frasa yang sering muncul dalam ulasan pengguna, dengan fokus pada ulasan negatif dan positif agar dapat memahami isu-isu utama serta aspek yang diapresiasi. Dari keseluruhan 1.048 ulasan, ditemukan bahwa beberapa frasa negatif paling dominan adalah “tolong diperbaiki” (58 kemunculan), “*scan barcode*” (49 kemunculan), dan “kode otp” (26 kemunculan). Frasa “tolong diperbaiki” mencerminkan permintaan eksplisit pengguna agar fitur-fitur tertentu diperbaiki, terutama terkait keandalan fungsi aplikasi. Selanjutnya, “*scan barcode*” menunjukkan masalah teknis dalam proses pemindaian, yang berdampak pada kegagalan akses data vaksinasi di fasilitas kesehatan. Frasa “kode otp” mengindikasikan tantangan dalam proses autentikasi—banyak pengguna mengalami kesulitan ketika menerima atau memasukkan kode verifikasi yang diperlukan untuk login. Kombinasi ketiga frasa ini menggarisbawahi area perbaikan mendesak yang harus menjadi prioritas pengembang.

Tabel 3 Frasa Umum (*bigram*) Sentimen Positif dan Negatif

No.	Negatif	Positif
1.	tolong diperbaiki (58)	terima kasih (4)
2.	scan barcode (49)	bagus cek (2)
3.	kode otp (26)	tekanan darah (2)

Di sisi lain, ulasan dengan sentimen positif menunjukkan frasa-frasa yang relatif lebih singkat dan terfokus pada apresiasi fungsionalitas tertentu. Frasa “terima kasih” muncul sebanyak 4 kali, menandakan bentuk pengakuan dan kepuasan pengguna atas layanan yang berjalan sesuai harapan. Frasa “bagus cek” (2 kemunculan) mengindikasikan pengguna menghargai kemudahan dalam memeriksa data seperti status vaksin atau rekam medis secara cepat. Kemudian, frasa “tekanan darah” (2 kemunculan) menggambarkan pengguna yang memanfaatkan fitur pemantauan data kesehatan, seperti tekanan darah, dan merasa antarmuka atau tampilan hasilnya memadai. Frasa-frasa positif ini memperlihatkan bahwa walaupun persentase ulasan positif relatif kecil, terdapat aspek tertentu dari aplikasi yang sudah memenuhi kebutuhan pengguna, khususnya terkait penyajian data kesehatan.

4. KESIMPULAN

Penelitian ini membuktikan bahwa model IndoBERT pra-fine-tune sangat efektif untuk mengklasifikasikan sentimen ulasan berbahasa Indonesia pada aplikasi SatuSehat. Dari 1.048 ulasan yang diproses, 87,3% teridentifikasi sebagai negatif, 8,4% sebagai netral, dan 4,3% sebagai positif. Nilai *confidence score* yang tinggi pada kategori negatif (rata-rata $\approx 0,96$) menunjukkan model sangat yakin dalam mengenali pola keluhan teknis. Analisis *unigram/bigram* mengungkap frasa-frasa kritis: “tolong diperbaiki”, “scan barcode”, dan “kode OTP” menjadi tiga keluhan utama, sedangkan frasa “terima kasih”, “bagus cek”, dan “tekanan darah” menyoroti fitur-fitur monitoring kesehatan yang sudah diapresiasi pengguna. Hasil ini menegaskan bahwa perbaikan segera perlu difokuskan pada proses *login/OTP* dan pemindaian *barcode*, agar poin-poin kegagalan teknis dapat diminimalkan. Sementara itu, fitur visualisasi tekanan darah dapat dikembangkan lebih lanjut sebagai nilai tambah yang positif.

Sebagai tindak lanjut, disarankan menerapkan *aspect-based sentiment analysis* dan melakukan *fine-tuning* IndoBERT di domain kesehatan spesifik, sehingga model dapat menangkap istilah medis dan nuansa bahasa pengguna dengan lebih presisi. Temuan ini tidak hanya memvalidasi keunggulan IndoBERT dalam konteks aplikasi kesehatan digital berbahasa Indonesia, tetapi juga memberikan rekomendasi berbasis data bagi tim pengembang SatuSehat untuk meningkatkan kualitas dan kepuasan pengguna.

DAFTAR PUSTAKA

- [1] S. Binkheder *et al.*, “Public Perceptions around mHealth Applications during COVID-19 Pandemic: A Network and Sentiment Analysis of Tweets in Saudi Arabia,” *Int. J. Environ. Res. Public. Health*, vol. 18, no. 24, p. 13388, Dec. 2021, doi: 10.3390/ijerph182413388.
- [2] C. L. Petersen *et al.*, “Using Natural Language Processing and Sentiment Analysis to Augment Traditional User-Centered Design: Development and Usability Study,” *JMIR MHealth UHealth*, vol. 8, no. 8, p. e16862, Aug. 2020, doi: 10.2196/16862.
- [3] J. Meyer and S. Okuboyejo, “User Reviews of Depression App Features: Sentiment Analysis,” *JMIR Form. Res.*, vol. 5, no. 12, p. e17062, Dec. 2021, doi: 10.2196/17062.
- [4] R. Wulandari and A. N. Hidayanto, “Measuring contact tracing service quality using sentiment analysis: a case study of PeduliLindungi Indonesia,” *Qual. Quant.*, vol. 58, no. 2, pp. 1409–1424, Apr. 2024, doi: 10.1007/s11135-023-01695-8.
- [5] I. Yunanto and S. Yulianto, “TWITTER SENTIMENT ANALYSIS PEDULILINDUNGI APPLICATION USING NAÏVE BAYES AND SUPPORT VECTOR MACHINE,” *J. Tek. Inform. Jutif*, vol. 3, no. 4, pp. 807–814, Aug. 2022, doi: 10.20884/1.jutif.2022.3.4.292.
- [6] R. Rahmatullah, J. N. Elquthb, F. N. Al-Qurani, and A. U. Khasanah, “Sentiment Analysis of PeduliLindungi User Using Naïve Bayes Classifier Algorithm and Support Vector Machine,” *J. Ind. Eng. Halal Ind.*, vol. 5, no. 1, pp. 36–42, Aug. 2024, doi: 10.14421/jiehis.4672.

- [7] R. Kurniawan and R. Rachmawati, “Indonesian Twitter User Sentiment towards PeduliLindungi App in Strengthening Smart Living during COVID-19,” *IKAT Indones. J. Southeast Asian Stud.*, vol. 6, no. 2, p. 150, Mar. 2024, doi: 10.22146/ikat.v6i2.89824.
 - [8] Y. A. Singgalen, “Performance Analysis of IndoBERT for Sentiment Classification in Indonesian Hotel Review Data,” *J. Inf. Syst. Res. JOSH*, vol. 6, no. 2, Art. no. 2, Jan. 2025, doi: 10.47065/josh.v6i2.6505.
 - [9] R. I. Perwira, V. A. Permadi, D. I. Purnamasari, and R. P. Agusdin, “Domain-Specific Fine-Tuning of IndoBERT for Aspect-Based Sentiment Analysis in Indonesian Travel User-Generated Content,” *J. Inf. Syst. Eng. Bus. Intell.*, vol. 11, no. 1, pp. 30–40, Mar. 2025, doi: 10.20473/jisebi.11.1.30-40.
 - [10] Y. Asri, D. Kuswardani, W. N. Suliyanti, Y. O. Manullang, and A. R. Ansyari, “Sentiment analysis based on Indonesian language lexicon and IndoBERT on user reviews PLN mobile application,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 38, no. 1, p. 677, Apr. 2025, doi: 10.11591/ijeeecs.v38.i1.pp677-688.
 - [11] A. C. Saputra, A. S. Saragih, and D. Ronaldo, “PREDIKSI EMOSI DALAM TEKS BAHASA INDONESIA MENGGUNAKAN MODEL INDOBERT,” *J. Teknol. Inf. J. Keilmuan Dan Apl. Bid. Tek. Inform.*, vol. 19, no. 1, Art. no. 1, Jan. 2025, doi: 10.47111/jti.v19i1.17617.
 - [12] H. Imaduddin, F. Y. A’la, and Y. S. Nugroho, “Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, 2023, doi: 10.14569/IJACSA.2023.0140813.
 - [13] M. Hanif Razka, T. Theresiawati, and N. Chamidah, “Analisis Sentimen terhadap Aplikasi PeduliLindungi pada Jejaring Sosial Twitter Menggunakan Algoritma Naïve Bayes dan Seleksi Fitur Particle Swarm Optimization,” *Inform. J. Ilmu Komput.*, vol. 19, no. 1, pp. 68–80, May 2023, doi: 10.52958/iftk.v19i1.4688.
 - [14] R. I. Syah, H. Hoiriyyah, and M. Walid, “ANALISIS SENTIMEN PENGGUNA MEDIA SOSIAL TERHADAP APLIKASI M-HEALTH PEDULI LINDUNG DENGAN METODE LEXICON BASED DAN NAÏVE BAYES,” *Indones. J. Bus. Intell. IJUBI*, vol. 6, no. 1, Jun. 2023, doi: 10.21927/ijubi.v6i1.3275.
 - [15] K. S. Nugroho, A. Y. Sukmadewa, H. W. DW, F. A. Bachtiar, and N. Yudistira, “BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews,” 2021, doi: 10.48550/ARXIV.2107.06802.
 - [16] D. A. Akhdaan, Taufik Edy Sutanto, and Muhamza Liebenlito, “Confident Learning pada IndoBERT: Peningkatan Kinerja Klasifikasi Sentimen,” *Indones. J. Comput. Sci.*, vol. 13, no. 5, Oct. 2024, doi: 10.33022/ijcs.v13i5.4401.
-