



Analisis Sentimen Multi-Kelas Ulasan Aplikasi SatuSehat Menggunakan Indobert Berbasis Transformer

Lilik Kuswati*, Bakhrudin All Habsy

Universitas Negeri Surabaya, Indonesia

Email: 25011355030@mhs.unesa.ac.id*, bakhrudinhabsy@unesa.ac.id

Abstract

Keywords:

Sentiment Analysis, IndoBERT, SatuSehat Application, Transformer

Digital transformation in the healthcare sector has driven the development of the SatuSehat application as an integrated health service platform in Indonesia. However, the increasing number of users is accompanied by diverse user reviews that reflect user experiences and satisfaction, requiring systematic analysis to understand these perceptions. This study aims to analyze user sentiment in SatuSehat application reviews using a transformer-based deep learning model, namely IndoBERT. The research method employed is Knowledge Discovery from Data (KDD), which includes data collection through web scraping from the Google Play Store, preprocessing, text transformation, classification, and model evaluation. The dataset consists of 3,000 reviews categorized into three sentiment classes: negative, neutral, and positive. The results indicate that the IndoBERT model achieves an accuracy of 86.56%, with precision of 86.01%, recall of 86.56%, and an F1-score of 84.33%. The findings reveal that negative sentiment dominates, primarily related to technical issues such as system errors and login difficulties. In conclusion, IndoBERT is effective in understanding the context of the Indonesian language in informal reviews and can serve as a foundation for developing automated sentiment monitoring systems to improve digital healthcare service quality.

Kata Kunci:

Analisis Sentimen, IndoBERT, Aplikasi SatuSehat, Transformer

Abstrak

Transformasi digital di sektor kesehatan mendorong pengembangan aplikasi SatuSehat sebagai platform layanan kesehatan terintegrasi di Indonesia. Namun, meningkatnya jumlah pengguna juga diiringi dengan beragam ulasan yang mencerminkan pengalaman dan kepuasan pengguna, sehingga diperlukan analisis yang sistematis untuk memahami persepsi tersebut. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi SatuSehat menggunakan model deep learning berbasis transformer, yaitu IndoBERT. Metode yang digunakan adalah Knowledge Discovery from Data (KDD) yang meliputi tahap pengumpulan data melalui web scraping dari Google Play Store, preprocessing, transformasi teks, klasifikasi, dan evaluasi model. Dataset yang digunakan terdiri dari 3.000 ulasan yang diklasifikasikan ke dalam tiga kategori sentimen: negatif, netral, dan positif. Hasil penelitian menunjukkan bahwa model IndoBERT mampu mencapai akurasi sebesar 86,56%, dengan precision 86,01%, recall 86,56%, dan F1-score 84,33%. Temuan menunjukkan dominasi sentimen negatif yang berkaitan dengan kendala teknis seperti error sistem dan kesulitan login. Kesimpulannya, IndoBERT efektif dalam memahami konteks bahasa Indonesia pada ulasan informal dan dapat digunakan sebagai dasar pengembangan sistem monitoring sentimen otomatis untuk meningkatkan kualitas layanan aplikasi kesehatan digital.

PENDAHULUAN

Transformasi digital pada sektor kesehatan di Indonesia terus mengalami perkembangan yang signifikan (Firdaus et al., 2025; Lukito & Gani, 2024; Santoso et al., 2025). Kementerian Kesehatan Republik Indonesia mengembangkan aplikasi SatuSehat sebagai bagian dari program transformasi digital kesehatan nasional (Ammara & Hafiza, 2025; Kurniawan & Rojabi, 2026; Widjaya et al., 2023). Aplikasi ini merupakan pengembangan dari PeduliLindungi dan saat ini telah menjadi platform layanan kesehatan utama yang menyediakan berbagai fitur seperti booking layanan kesehatan, pelacakan vaksinasi, serta rekam medis elektronik. Dengan jumlah pengguna yang terus meningkat, kualitas layanan aplikasi menjadi aspek penting yang perlu dievaluasi secara berkala.

Google Play Store menyediakan fitur rating dan ulasan yang memungkinkan pengguna memberikan opini terhadap aplikasi yang digunakan (As' ari et al., 2025; Berliana & Mustikasari, 2024; Helmayanti et al., 2023; Husain et al., 2024). Ulasan tersebut mengandung informasi mengenai pengalaman pengguna, baik berupa apresiasi, kritik, maupun keluhan terhadap performa aplikasi. Oleh karena itu, diperlukan metode yang mampu mengolah dan menganalisis opini tersebut secara otomatis dan sistematis. Salah satu pendekatan yang dapat digunakan adalah analisis sentimen berbasis *Natural Language Processing (NLP)* untuk mengklasifikasikan opini ke dalam kategori sentimen tertentu (Rizki et al., 2024)(Ayu et al., n.d.)(Ardianto & Marhoon, 2020).

Beberapa penelitian sebelumnya telah melakukan analisis sentimen terhadap aplikasi SatuSehat menggunakan algoritma berbasis *Recurrent Neural Network (RNN)* seperti *Long Short-Term Memory (LSTM)* dan *Gated Recurrent Unit (GRU)*. Penelitian tersebut menunjukkan tingkat akurasi antara 85% hingga 90% pada klasifikasi dua kelas sentimen. Meskipun demikian, pendekatan tersebut belum memanfaatkan arsitektur transformer yang memiliki kemampuan pemahaman konteks dua arah (*bidirectional*) melalui mekanisme attention. Model *Bidirectional Encoder Representations from Transformers (BERT)* telah terbukti mampu meningkatkan performa analisis sentimen pada berbagai domain, termasuk ulasan aplikasi publik, media sosial, dan dokumen kesehatan (Rizki et al., 2024)(Wahyudi & Sibaroni, 2022).

Model BERT, khususnya *IndoBERT* yang telah dilatih menggunakan korpus Bahasa Indonesia, memiliki kemampuan yang lebih baik dalam memahami struktur dan konteks bahasa Indonesia dibandingkan model *RNN* konvensional (Immanuela Leihitu & Kusumarini, 2025). Dengan kemampuan tersebut, model transformer berpotensi memberikan hasil klasifikasi yang lebih akurat dan stabil pada data ulasan aplikasi kesehatan digital (Sabrina Amanda Salsabila et al., 2025)(Atlas et al., 2025)(Kenton et al., 2019).

Berdasarkan hal tersebut, penelitian ini mengimplementasikan model *IndoBERT* untuk melakukan analisis sentimen multi-kelas (negatif, netral, dan positif) terhadap ulasan aplikasi SatuSehat pada periode 1 Januari hingga 7 November 2025. Hasil penelitian ini akan dianalisis untuk mengetahui distribusi sentimen pengguna serta dibandingkan dengan hasil penelitian terdahulu yang menggunakan *LSTM* dan *GRU* sebagai baseline. Penelitian ini bertujuan untuk mengetahui efektivitas pendekatan berbasis *transformer* dalam domain aplikasi kesehatan digital di Indonesia serta memberikan gambaran persepsi pengguna terhadap kualitas layanan aplikasi SatuSehat.

Sejumlah penelitian terdahulu telah mengkaji analisis sentimen terhadap ulasan aplikasi menggunakan berbagai pendekatan machine learning dan deep learning. Misalnya, penelitian oleh Devlin et al. memperkenalkan model BERT yang mampu memahami konteks bahasa secara bidirectional dan memberikan peningkatan signifikan dalam berbagai tugas NLP. Penelitian lain oleh Wahyudi dan Sibaroni menunjukkan bahwa model berbasis LSTM mampu

mencapai akurasi yang cukup tinggi dalam analisis sentimen, meskipun masih memiliki keterbatasan dalam menangkap konteks jangka panjang. Selain itu, studi oleh Aisy dan Karyono mengungkapkan bahwa penggunaan IndoBERT memberikan hasil yang lebih baik dibandingkan metode konvensional dalam klasifikasi sentimen berbahasa Indonesia.

Meskipun demikian, sebagian besar penelitian sebelumnya masih berfokus pada klasifikasi sentimen dua kelas (positif dan negatif) serta menggunakan dataset yang relatif terbatas. Selain itu, penggunaan model transformer seperti IndoBERT dalam konteks aplikasi kesehatan digital, khususnya SatuSehat, masih belum banyak dieksplorasi secara mendalam. Kesenjangan penelitian ini menunjukkan bahwa diperlukan kajian yang lebih komprehensif dengan pendekatan multi-kelas yang mampu menangkap nuansa sentimen secara lebih detail, termasuk kategori netral yang seringkali memiliki karakteristik ambigu.

Urgensi penelitian ini semakin meningkat seiring dengan pentingnya kualitas layanan aplikasi kesehatan digital dalam mendukung sistem kesehatan nasional. Ulasan pengguna yang didominasi oleh sentimen negatif, terutama terkait kendala teknis, menunjukkan adanya permasalahan yang perlu segera ditangani oleh pengembang aplikasi. Tanpa adanya sistem analisis yang efektif, informasi berharga dari pengguna berpotensi tidak termanfaatkan secara optimal. Oleh karena itu, diperlukan sistem analisis sentimen otomatis yang tidak hanya akurat tetapi juga mampu memberikan insight yang relevan bagi pengambilan keputusan strategis.

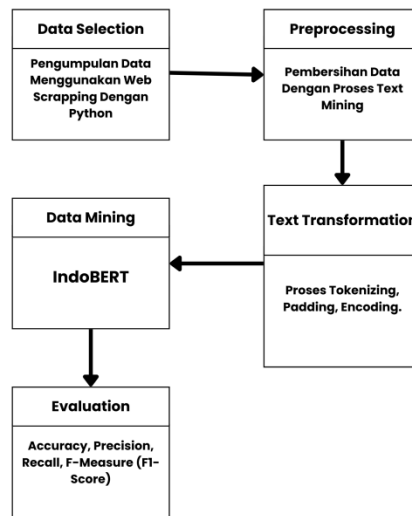
Penelitian ini menawarkan kebaruan dengan mengimplementasikan model IndoBERT untuk klasifikasi sentimen multi-kelas pada ulasan aplikasi SatuSehat, yang sebelumnya lebih banyak dianalisis menggunakan pendekatan RNN atau metode tradisional. Selain itu, penelitian ini juga mengintegrasikan tahapan Knowledge Discovery from Data (KDD) secara sistematis, mulai dari pengumpulan data hingga evaluasi model, sehingga menghasilkan proses analisis yang lebih terstruktur dan komprehensif. Kebaruan lainnya terletak pada penggunaan dataset yang lebih besar dan aktual, sehingga mampu merepresentasikan kondisi nyata persepsi pengguna terhadap aplikasi.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi SatuSehat menggunakan model IndoBERT serta mengevaluasi performa model dalam klasifikasi multi-kelas. Selain itu, penelitian ini juga bertujuan untuk mengidentifikasi distribusi sentimen pengguna sebagai dasar evaluasi kualitas layanan aplikasi. Manfaat penelitian ini diharapkan dapat memberikan kontribusi teoritis dalam pengembangan metode analisis sentimen berbasis transformer untuk bahasa Indonesia, serta kontribusi praktis bagi pengembang aplikasi dalam meningkatkan kualitas layanan berdasarkan masukan pengguna secara berbasis data.

METODE PENELITIAN

Pada penelitian ini dilakukan proses pengambilan data (*web scraping*) pada *Google Play Store* menggunakan bahasa pemrograman *Python* dengan memasukkan link aplikasi SatuSehat. Metode penelitian yang digunakan adalah *Knowledge Discovery from Data (KDD)*.

KDD merupakan metode yang terdiri dari beberapa tahapan sistematis dalam proses pengolahan data hingga menghasilkan informasi yang bermanfaat. Tahapan dalam *KDD* meliputi *Data Selection*, *Preprocessing*, *Transformation*, *Data Mining*, dan *Evaluation*. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1 Alur Penelitian

Data Selection

Data dikumpulkan dari *Google Play Store* berupa ulasan pengguna aplikasi SatuSehat periode 1 Januari hingga 7 November 2025. Proses *scraping* dilakukan menggunakan *library Python* untuk memperoleh data berupa teks ulasan, rating, tanggal, dan informasi pendukung lainnya.

Dataset yang terkumpul sebanyak 3.000 ulasan. Data kemudian diklasifikasikan berdasarkan rating menjadi tiga kelas sentimen, yaitu:

- Negatif (rating 1–2)
- Netral (rating 3)
- Positif (rating 4–5)

Distribusi kelas menunjukkan adanya ketidakseimbangan data (*imbalanced dataset*), di mana jumlah ulasan negatif lebih dominan dibandingkan kelas lainnya.

Preprocessing

Tahap preprocessing merupakan tahap awal dalam *text mining* yang bertujuan untuk membersihkan dan menyiapkan data sebelum dilakukan proses klasifikasi. Tahap ini diperlukan untuk menghilangkan *noise* serta meningkatkan kualitas data teks.

Pada penelitian ini, *preprocessing* meliputi:

1. *Cleaning*, yaitu menghapus karakter khusus, emoji, URL, dan simbol yang tidak diperlukan.
2. *Case Folding*, yaitu mengubah seluruh huruf menjadi huruf kecil.
3. Normalisasi, yaitu memperbaiki kata tidak baku atau singkatan menjadi bentuk standar Bahasa Indonesia.
4. *Labeling*, yaitu pemberian label sentimen berdasarkan rating yang diberikan pengguna.

Tahap *preprocessing* dilakukan untuk memastikan data siap diproses pada tahap transformasi dan klasifikasi model.

Text Transformation

Text transformation merupakan tahap untuk mengubah data teks menjadi representasi yang dapat diproses oleh model *machine learning* atau *deep learning*.

Pada penelitian ini, transformasi dilakukan menggunakan *tokenizer* bawaan *IndoBERT*. Proses ini meliputi:

- *Tokenizing*, memecah teks menjadi token sesuai vocabulary model.
- *Padding dan Truncation*, menyamakan panjang input sesuai kebutuhan model.

- *Encoding*, mengubah token menjadi representasi numerik (input IDs dan attention mask).

Tahap ini memungkinkan teks ulasan diproses oleh arsitektur transformer secara optimal.

Data Mining

Pada tahap klasifikasi digunakan model *IndoBERT* (*Bidirectional Encoder Representations from Transformers*). Model ini merupakan arsitektur transformer yang telah melalui proses *pre-training* pada korpus Bahasa Indonesia (Aisy & Karyono, 2025).

Proses klasifikasi dibagi menjadi dua tahap, yaitu:

1. *Training*, yaitu proses pelatihan model menggunakan data latih untuk menyesuaikan bobot model terhadap tugas klasifikasi sentimen.
2. *Testing*, yaitu proses pengujian model menggunakan data uji untuk mengukur performa model.

Dataset dibagi menjadi 80% data *training* dan 20% data *testing*. Model *IndoBERT* dilakukan *fine-tuning* dengan menambahkan layer klasifikasi pada bagian akhir untuk menghasilkan prediksi tiga kelas sentimen (negatif, netral, positif) (Anggraini & Wahyuni, 2025).

Model *transformer* dipilih karena memiliki kemampuan memahami konteks kalimat secara dua arah melalui mekanisme *self-attention*, sehingga mampu menangkap hubungan antar kata dengan lebih baik dibandingkan model *RNN* konvensional.

Evaluasi

Tahap evaluasi dilakukan untuk mengetahui performa model klasifikasi yang digunakan dalam penelitian ini.

Tolak ukur yang digunakan meliputi:

1. *Accuracy*
Mengukur tingkat kesesuaian antara hasil prediksi model dengan label aktual.
2. *Precision*
Mengukur tingkat ketepatan model dalam memprediksi kelas tertentu. *Precision* penting ketika nilai *False Positive* tinggi.
3. *Recall*
Mengukur kemampuan model dalam mengidentifikasi seluruh data yang benar-benar termasuk dalam suatu kelas (*True Positive*). *Recall* penting ketika nilai *False Negative* tinggi.
4. *F-Measure (F1-Score)*
Merupakan rata-rata harmonik antara *precision* dan *recall* yang digunakan untuk menilai keseimbangan performa model.

Untuk memberikan gambaran kinerja yang lebih jelas, digunakan ***confusion matrix*** yang menampilkan jumlah *True Positive*, *True Negative*, *False Positive*, dan *False Negative* pada masing-masing kelas sentimen.

HASIL DAN PEMBAHASAN

Hasil dan pembahasan pada penelitian ini merupakan implementasi analisis sentimen ulasan aplikasi SatuSehat menggunakan model *IndoBERT* yang diolah menggunakan *Google Colab* dengan dukungan *GPU Tesla T4*. Contoh data *scraping*, pelabelan, *preprocessing*, proses *training*, serta hasil evaluasi model disajikan dibawah.

Scraping Data

Pada tahap ini dilakukan *scraping* data dengan menyantumkan link aplikasi SatuSehat dari *Google Play Store*. Proses *scraping* menggunakan *library google_play_scraper* pada *Python* dengan jumlah data sebanyak 3.308 ulasan.

Output hasil *crawling* menampilkan dataset mentah yang kemudian disimpan dalam bentuk DataFrame untuk proses selanjutnya.

	review	rating	date
0	sudah akses tspi susah online	1	2025-11-07 00:14:25
1	buruk banget, aku mau daftar padahal cuma scan...	1	2025-11-06 23:48:41
2	setelah update ga bisa login malah error,hadeh...	1	2025-11-06 22:41:00
3	jelek bgt aplikasinya. setiap login masukin ko...	1	2025-11-06 22:20:35
4	aplikasi gak jelas nik ktp saya tidak terdaftar	1	2025-11-06 18:03:34
5	setelah update yang terbaru, malah gak bisa lo...	1	2025-11-06 15:47:03
6	baik	5	2025-11-06 14:14:44
7	sangat baik	5	2025-11-06 14:00:15
8	baru coba2	5	2025-11-06 13:19:56
9	tidak bisa login dengan berbagai cara	1	2025-11-06 11:10:41

Gambar 2 Hasil Crawling Data

Preprocessing

Tahap *preprocessing* dilakukan untuk membersihkan data teks sebelum digunakan dalam pelatihan model *IndoBERT*.

Berikut contoh proses *preprocessing* pada satu sampel ulasan:

Table 1 Data ulasan Bersih

Tahapan	Hasil
Data Mentah	“Sudah lima hari tidak bisa <i>login</i> , aplikasi terus-menerus <i>error</i> !!!”
<i>Case Folding</i>	sudah lima hari tidak bisa <i>login</i> aplikasi terus menerus <i>error</i>
<i>Cleaning</i>	sudah lima hari tidak bisa login aplikasi terus menerus <i>error</i>
<i>Tokenizing (IndoBERT)</i>	sudah, lima, hari, tidak, bisa, <i>login</i> , aplikasi, terus, menerus, <i>error</i>

Berbeda dengan metode tradisional, pada *IndoBERT* tidak dilakukan *stopword removal* dan *stemming* karena arsitektur *transformer* mampu memahami konteks secara *bidirectional* melalui mekanisme *self-attention*.

Tahapan ini menghasilkan input berupa *input_ids* dan *attention_mask* yang siap diproses pada tahap *training*.

Pelabelan Otomatis

Setelah proses *scraping*, dilakukan pelabelan otomatis berdasarkan rating yang diberikan pengguna. Skema *labeling* menggunakan tiga kelas, yaitu:

- Negatif (rating 1–2)
- Netral (rating 3)
- Positif (rating 4–5)

Hasil pelabelan menambahkan kolom baru bernama **Label** yang menunjukkan kategori sentimen dari setiap ulasan. Distribusi hasil pelabelan menunjukkan bahwa kelas Negatif mendominasi dataset sebesar 74,4%.

```

=====
📄 CONTOH ULASAN PER KELAS SENTIMEN
=====

--- NEGATIF SENTIMENT ---
1. gak sesuai update di aplikasi dan di lapangan, tanggal sabtu atau tanggalan
2. 5 hari gak bisa login. aplikasi te
3. ini apk nya ribet bgt, lupa nomor telepon tapi ttp ngirim kode verif ke no 1

--- NETRAL SENTIMENT ---
1. tidak bisa konek ke google fit. selalu force close.
2. sangat sulit mendaftar di akun satusehat!
3. apk nya error terus

--- POSITIF SENTIMENT ---
1. maaf tanya aplikasi ini sama jkn bedanya apa ya?
2. alhamdulillah bgs
3. sangat bermanfaat

```

Gambar 3 Ulasan Per Kelas Sentimen

Training Model IndoBERT

Tahap klasifikasi dilakukan dengan proses fine-tuning model *IndoBERT* menggunakan data *training* sebanyak 80% dari total dataset.

Konfigurasi training:

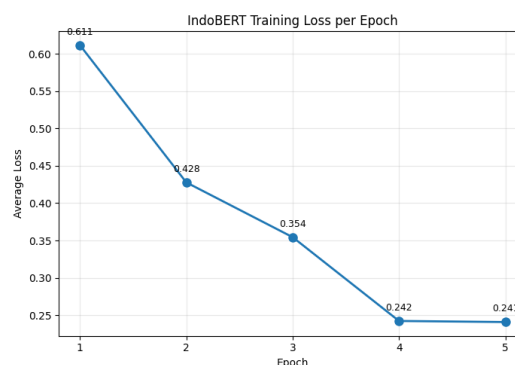
- *Epoch: 5*
- *Batch size: 16*
- *Learning rate: 2e-5*
- *Optimizer: AdamW*

Training dilakukan menggunakan *GPU Tesla T4* dengan waktu eksekusi ± 4 menit 40 detik. Perkembangan nilai *training loss* dapat dilihat pada tabel berikut.

Tabel 2 Training Loss

Epoch	Training Loss
1	0,6114
2	0,4278
3	0,3545
4	0,2424
5	0,2409

Penurunan *loss* yang konsisten menunjukkan bahwa model mencapai konvergensi dengan baik tanpa indikasi *overfitting*.



Gambar 4 Training Loss

Hasil Accuracy dan Evaluasi Model

Tahap evaluasi dilakukan menggunakan data *testing* sebanyak 662 ulasan (20% dari total dataset).

Hasil evaluasi menunjukkan:

- *Accuracy*: 86,56%
- *Precision (Weighted)*: 86,01%
- *Recall (Weighted)*: 86,56%
- *F1-score (Weighted)*: 84,33%

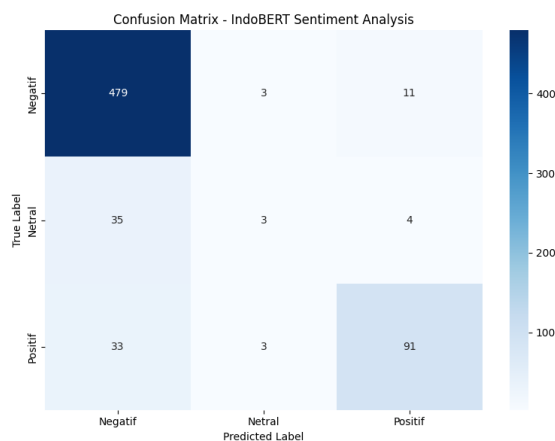
Hasil ini menunjukkan bahwa *IndoBERT* mampu mengklasifikasikan sentimen dengan performa yang baik pada klasifikasi tiga kelas.

Confusion Matrix

Model menunjukkan performa sangat baik dalam membedakan kelas Negatif dan Positif. Namun, kelas Netral masih sulit dikenali akibat distribusi data yang tidak seimbang.

1. **Negatif vs Positif**: Model sangat baik membedakan kedua kelas ekstrem ini, hanya 11 dari 493 negatif diprediksi sebagai positif (2,2% error), dan 32 dari 127 positif diprediksi negatif (25,2% error).
2. **Tantangan Kelas Netral**: Dari 42 ulasan netral aktual:
 - a. 35 diprediksi sebagai Negatif (83,3%) — *overlap* karakteristik kritik ringan.
 - b. 3 diprediksi benar sebagai Netral (7,1%) — sangat rendah.
 - c. 4 diprediksi sebagai Positif (9,5%).

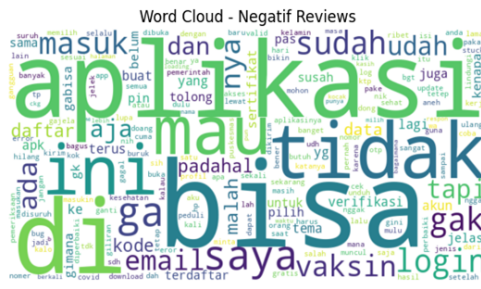
Pattern Error Dominan: Model memiliki bias kuat terhadap kelas mayoritas (Negatif), mencerminkan *class imbalance* 74,4% vs 6,3% pada training data.



Gambar 5 Confusion Matrix

Visualisasi Word Cloud Sentimen

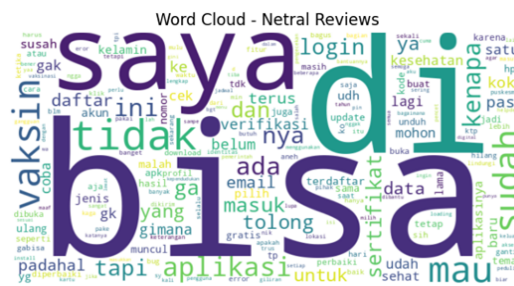
Untuk memperoleh gambaran umum mengenai karakteristik teks pada setiap kategori sentimen, dilakukan visualisasi menggunakan metode *word cloud*. Visualisasi ini menampilkan frekuensi kemunculan kata dalam bentuk ukuran yang proporsional, di mana semakin besar ukuran kata maka semakin sering kata tersebut muncul dalam ulasan.



1. Sentimen Negatif

Berdasarkan Gambar Word Cloud sentimen negatif, kata-kata yang paling dominan antara lain “tidak”, “apk”, “bisa”, “login”, “daftar”, dan “error”. Dominasi kata-kata tersebut menunjukkan bahwa sebagian besar pengguna mengalami kendala teknis dalam penggunaan aplikasi, terutama terkait proses autentikasi dan akses sistem.

Selain itu, kemunculan kata seperti “*susah*”, “*gagal*”, dan “*kode*” mengindikasikan adanya permasalahan pada fitur verifikasi dan stabilitas layanan. Hal ini menunjukkan bahwa aspek performa sistem dan pengalaman pengguna masih menjadi isu utama dalam aplikasi SatuSehat.



2. Sentimen Netral

Pada sentimen netral, kata-kata yang sering muncul meliputi “*saya*”, “*tidak*”, “*login*”, “*aplikasi*”, dan “*data*”. Ulasan pada kategori ini cenderung bersifat informatif, berupa pertanyaan, laporan singkat, atau deskripsi pengalaman tanpa ekspresi emosi yang kuat.

Kemunculan kata seperti “*verifikasi*”, “*daftar*”, dan “*email*” menunjukkan bahwa pengguna masih berinteraksi dengan fitur utama aplikasi, namun tidak secara eksplisit menunjukkan kepuasan atau ketidakpuasan. Hal ini menjelaskan mengapa kelas netral sulit diklasifikasikan oleh model, karena memiliki karakteristik yang ambigu dan berada di antara sentimen positif dan negatif.



3. Sentimen Positif

Pada sentimen positif, kata-kata dominan antara lain “*aplikasi*”, “*baik*”, “*bagus*”, “*membantu*”, “*mudah*”, dan “*kesehatan*”. Hal ini menunjukkan bahwa pengguna yang memberikan ulasan positif cenderung menyoroti manfaat aplikasi dalam mendukung layanan kesehatan digital. Kata seperti “*vaksin*”, “*data*”, dan “*pelayanan*” mengindikasikan bahwa fitur utama aplikasi dinilai bermanfaat oleh pengguna. Selain itu, kemunculan kata “*mudah*” dan “*cepat*” menunjukkan bahwa aspek *usability* menjadi faktor penting dalam kepuasan pengguna.

Secara keseluruhan, hasil visualisasi word cloud menunjukkan adanya perbedaan pola bahasa yang signifikan antar kelas sentimen. Sentimen negatif didominasi oleh kata-kata yang berkaitan dengan masalah teknis, sementara sentimen positif lebih menekankan manfaat dan kemudahan penggunaan aplikasi.

Temuan ini sejalan dengan hasil klasifikasi model *IndoBERT*, di mana kelas negatif memiliki performa tertinggi karena pola bahasanya lebih konsisten dan mudah dikenali. Sebaliknya, kelas netral memiliki karakteristik yang ambigu, sehingga lebih sulit diklasifikasikan oleh model.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan model transformer berbasis *IndoBERT* terbukti efektif dalam melakukan klasifikasi sentimen multi-kelas terhadap ulasan pengguna aplikasi SatuSehat. Model mampu mencapai performa yang baik dengan nilai *accuracy* sebesar 86,56% serta didukung oleh nilai *precision*, *recall*, dan *F1-score* yang relatif tinggi. Hasil analisis menunjukkan bahwa mayoritas ulasan pengguna didominasi oleh sentimen negatif yang berkaitan dengan kendala teknis seperti error sistem, kesulitan login, dan stabilitas layanan. Selain itu, model menunjukkan kemampuan yang sangat baik dalam membedakan sentimen ekstrem (negatif dan positif), meskipun masih mengalami kesulitan dalam mengklasifikasikan sentimen netral akibat ketidakseimbangan distribusi data. Temuan ini mengindikasikan bahwa pendekatan berbasis transformer memiliki keunggulan dalam memahami konteks bahasa Indonesia, khususnya pada data teks yang bersifat informal, serta dapat menjadi solusi dalam pengembangan sistem analisis sentimen otomatis untuk evaluasi layanan digital kesehatan.

Untuk penelitian selanjutnya, disarankan untuk mengatasi permasalahan ketidakseimbangan data dengan menerapkan teknik *balancing* seperti *oversampling*, *undersampling*, atau penggunaan metode *Synthetic Minority Over-sampling Technique (SMOTE)* agar performa klasifikasi pada kelas netral dapat ditingkatkan. Selain itu, penelitian lanjutan dapat mengembangkan model dengan pendekatan *ensemble learning* atau membandingkan performa *IndoBERT* dengan varian model transformer lainnya seperti *RoBERTa* atau *XLNet* yang telah diadaptasi untuk Bahasa Indonesia. Penggunaan data yang lebih besar dan beragam dari berbagai platform digital juga direkomendasikan untuk meningkatkan generalisasi model. Lebih lanjut, integrasi analisis sentimen dengan analisis aspek (*aspect-based sentiment analysis*) dapat memberikan insight yang lebih mendalam terkait faktor-faktor spesifik yang mempengaruhi kepuasan pengguna. Dengan demikian, hasil penelitian di masa depan diharapkan tidak hanya meningkatkan akurasi model, tetapi juga memberikan kontribusi praktis dalam pengambilan keputusan strategis untuk pengembangan layanan aplikasi kesehatan digital di Indonesia.

REFERENSI

- Aisy, A. R., & Karyono, G. (2025). *Analisis Sentimen Terhadap Ulasan Google Play Store Aplikasi Lazada, Shopee, dan Tokopedia Menggunakan Algoritma IndoBERT*. 7(3), 2127–2135. <https://doi.org/10.47065/bits.v7i3.8745>
- Ammara, D., & Hafiza, G. (2025). Evaluasi Layanan Publik Berbasis Digital: Studi Efektivitas Implementasi Cek Kesehatan Gratis (CKG) pada Aplikasi Satusehat di UPT. Puskesmas Medan Johor. *Recht Studiosum Law Review*, 4(2), 210–220.
- Anggraini, D., & Wahyuni, E. I. (2025). ANALISIS SENTIMEN PENGGUNA TERHADAP APLIKASI GLINTS DI GOOGLE PLAY STORE MENGGUNAKAN METODE NAIVE BAYES Eka Indah Wahyuni 2 Politeknik Belitung. *JURNAL MEDIA AKADEMIK (JMA)*, 3(5), 3031–5220.
- Ardianto, R., & Marhoon, H. M. (2020). Understanding User Sentiment: Analysis of SATUSEHAT Application Reviews on Google Play Store. *Journal of Advanced Health Informatics Research*, 1(2), 83–94. <https://doi.org/10.59247/jahir.v1i2.44>
- As' ari, W. R., Arifin, M., Fithri, D. L., & Setiaji, P. (2025). Analisis Sentimen Ulasan Pengguna Aplikasi Spotify Di Google Play Store Menggunakan Algoritma Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(2), 3600–3607.
- Atlas, L. G., Arockiam, D., Muthusamy, A., Balusamy, B., Selvarajan, S., Al-Shehari, T., & Alsadhan, N. A. (2025). A modernized approach to sentiment analysis of product reviews using BiGRU and RNN based LSTM deep learning models. *Scientific Reports*, 15(1), 1–24. <https://doi.org/10.1038/s41598-025-01104-0>
- Ayu, I., Cahya, M., Dharmendra, I. K., & Setiasih, N. W. (n.d.). *ANALISIS SENTIMEN REVIEW APLIKASI SATU SEHAT MOBILE MENGGUNAKAN MODEL SAMPLING TOMEK LINKS*. 497–504.
- Berliana, A. S., & Mustikasari, M. (2024). Analisis Sentimen Pada Ulasan Aplikasi Jakartanotebook Di Google Play Menggunakan Metode Recurrent Neural Network (RNN). *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3).
- Firdaus, R., Syeira, K., & Wijaya, N. (2025). Transformasi Digital Sistem Informasi Kesehatan Menuju Layanan Kesehatan Yang Terkoneksi Dan Berpusat Pada Pasien. *Economics and Digital Business Review*, 6(2), 1045–1055.
- Helmayanti, S. A., Hamami, F., & Fa'rifah, R. Y. (2023). Penerapan algoritma Tf-Idf dan Naïve Bayes untuk analisis sentimen berbasis aspek ulasan aplikasi Flip pada Google Play Store. *Jurnal Indonesia: Manajemen Informatika Dan Komunikasi*, 4(3), 1822–1834.
- Husain, N. P., Sukirman, S., & SAJIAH, S. (2024). Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine. *Journal of System and Computer Engineering*, 5(1), 91–102.
- Immanuela Leihitu, E., & Kusumarini, C. D. (2025). *Eksplorasi Sentimen Aplikasi BCA Mobile Pada Google Playstore Menggunakan Metode Bidirectional Encoder Representations from Transformers (BERT)*. 03(03), 153–160.
- Kenton, M. C., Kristina, L., & Devlin, J. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 4171–4186.
- Kurniawan, D., & Rojabi, M. A. (2026). *SatuSehat: Integrasi kesehatan nasional Indonesia*. Afdan Rojabi Publisher.
- Lukito, M., & Gani, A. (2024). Pelayanan kesehatan yang efisien dan terjangkau melalui transformasi kesehatan digital via telemedicine di Indonesia. *Jurnal Kesehatan: Jurnal Ilmiah Multi Sciences*, 14(2), 107–117.

- Rizki, B., Hidayat, N., Sanjaya, R., Studi, P., Informatika, T., Adhirajasa, U., Sanjaya, R., Studi, P., Informasi, S., Adhirajasa, U., & Sanjaya, R. (2024). *Penerapan Text Mining Dengan Algoritma Random Forest Menganalisis Sentimen Ulasan SATUSEHAT Mobile*. 5(2), 209–222.
- Sabrina Amanda Salsabila, Bayu Priyatna, Agustia Hananto, & Tukino. (2025). Komparasi Kinerja Model Naive Bayes, SVM, dan BERT dalam Klasifikasi Sentimen Ulasan Pada Aplikasi YUMMY. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 4(2), 42–47. <https://doi.org/10.55123/storage.v4i2.5120>
- Santoso, F. S., Ramadhani, P. A., Amnamuchlisah, D., & Purba, S. H. (2025). Transformasi Digital Dalam Sektor Kesehatan Kajian Literatur Untuk Mendukung Inovasi dan Efisiensi Layanan Kesehatan. *Cindoku: Jurnal Keperawatan Dan Ilmu Kesehatan*, 2(1), 1–12.
- Wahyudi, D., & Sibaroni, Y. (2022). Deep Learning for Multi-Aspect Sentiment Analysis of TikTok App using the RNN-LSTM Method. *Building of Informatics, Technology and Science (BITS)*, 4(1), 169–177. <https://doi.org/10.47065/bits.v4i1.1665>
- Widjaya, D., Putra, H. D., & Vanny, C. (2023). *Accountability Brief: Tantangan dalam Mewujudkan Transformasi Teknologi Kesehatan*.