ZONAsi Jurnal Sistem Informasi
ZONAsi VOL. 8 NO. 2 Mei 2026
Page: 864 - 877 Akreditasi Sinta 4
ISSN: 2656-7407 (Online)
2656-7393 (Print)

PENERAPAN MODEL TRANSFORMER INDOBERT UNTUK ANALISIS SENTIMEN PADA ULASAN APLIKASI TWITCH

Rizky Gunawan[1], Rizka Dahlia[2], Muhammad Rifqi Firdaus[3], Lady Agustin Fitriana*[4]
[1,2]Informatika,[3,4]Sistem Informasi, Fakultas Teknik & Informatika
Universitas Bina Sarana Informatika
Jl. Abdul Rahman Saleh No.18, Bangka Belitung Laut, Kec. Pontianak Tenggara,
Kota Pontianak, Kalimantan Barat, telp. 0561-583924
e-mail: [1]15225007@bsi.ac.id, [2]rizka.rzl@bsi.ac.id, [3]muhammad.mku@bsi.ac.id, [4]lady.lag@bsi.ac.id

Abstrak

Pemahaman terhadap persepsi publik terhadap aplikasi Twitch menjadi penting karena ulasan pengguna mencerminkan pengalaman penggunaan, kualitas layanan, serta berbagai permasalahan yang dihadapi pengguna. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Twitch berbahasa Indonesia dengan menerapkan model Transformer IndoBERT (indobert-base-p1) melalui proses fine-tuning. Dataset diperoleh dari Google Play Store menggunakan google-play-scraper, kemudian diproses melalui cleaning, case folding, normalisasi, tokenisasi, stopword removal, dan stemming. Pelabelan data dilakukan menggunakan pendekatan lexicon-based sebagai ground truth dengan tiga kelas sentimen, yaitu positif, netral, dan negatif. Selanjutnya data digunakan untuk melatih model IndoBERT dengan konfigurasi learning rate 2e-5, batch size 8, dan lima epoch. Evaluasi dilakukan menggunakan accuracy, precision, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model terbaik diperoleh pada epoch ke-4 dengan akurasi sebesar 89% dan weighted average F1-score sebesar 89%. Analisis confusion matrix memperlihatkan bahwa model memiliki performa terbaik pada kelas negatif dan netral, sedangkan kelas positif masih mengalami sejumlah kesalahan klasifikasi ke kelas netral. Temuan ini menunjukkan bahwa IndoBERT mampu memahami konteks bahasa informal pada ulasan aplikasi berbahasa Indonesia dan layak digunakan sebagai pendekatan analisis sentimen untuk mendukung evaluasi kualitas layanan Twitch.

Kata kunci*:* *Analisis Sentimen, IndoBERT, Transformer, Ulasan Aplikasi, Twitch*

Abstract

*Understanding public perception of the Twitch application is important because user reviews reflect user experiences, service quality, and various issues encountered during application usage. This study aims to analyze the sentiment of Indonesian-language Twitch application reviews by implementing the IndoBERT Transformer model (indobert-base-p1) through a fine-tuning process. The dataset was collected from the Google Play Store using the **google-play-scraper** library and subsequently preprocessed through data cleaning, case folding, text normalization, tokenization, stopword removal, and stemming. The review data were labeled into three sentiment categories, namely positive, neutral, and negative, using a lexicon-based approach as the ground truth. The labeled dataset was then used to fine-tune the IndoBERT model with a learning rate of 2e-5, a batch size of 8, and five training epochs. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. Experimental results showed that the best-performing model was obtained at the fourth epoch, achieving an overall accuracy of **89%** and a weighted average **F1-score of 89%**. Furthermore, confusion matrix analysis indicated that the model performed best in classifying negative and neutral sentiments, while the positive class still experienced several misclassifications into the neutral category. These findings demonstrate that the fine-tuned IndoBERT model is capable of understanding informal Indonesian language commonly found in application reviews and is suitable for sentiment analysis tasks to support the evaluation of Twitch service quality.*

Keywords: *Sentiment Analysis, IndoBERT, Transformer, Application Reviews, Twitch*

1. PENDAHULUAN

Beberapa kajian terdahulu melaporkan penerapan IndoBERT pada berbagai korpus berbahasa Indonesia dan menemukan bahwa model ini umumnya memberikan peningkatan performa dibandingkan pendekatan konvensional atau model multilingual ketika diaplikasikan pada tugas klasifikasi sentimen dan aspek-sentimen [1]. Studi benchmark awal yang memperkenalkan IndoBERT menunjukkan bahwa model pra-latih khusus Bahasa Indonesia tersebut mampu menyediakan representasi kontekstual yang lebih sesuai untuk tugas NLP berbahasa lokal, sehingga menjadi dasar bagi banyak eksperimen lanjutan [2]. Penelitian lain yang menguji *fine-tuning* BERT pada ulasan aplikasi seluler juga memperlihatkan bahwa strategi pra-latih dan penyesuaian domain mampu meningkatkan akurasi klasifikasi pada dataset ulasan aplikasi Indonesia, sehingga mendukung perlunya adaptasi model terhadap karakteristik data target [3]. Penelitian sebelumnya yang dilakukan oleh Sianipar et al. menyatakan bahwa pendekatan analisis sentimen berbasis pembelajaran mesin efektif dalam mengelompokkan opini pengguna terhadap aplikasi digital seperti DANA dan LinkAja [4]. Temuan mereka menunjukkan bahwa model klasifikasi mampu mengenali pola linguistik khas dalam ulasan berbahasa Indonesia, yang kemudian menjadi dasar untuk memahami persepsi publik terhadap kualitas layanan aplikasi. Dengan demikian, hasil penelitian tersebut memperkuat urgensi penerapan model transformer seperti IndoBERT dalam konteks analisis sentimen ulasan aplikasi Twitch yang memiliki karakteristik bahasa serupa.

Analisis sentimen merupakan bidang dalam pengolahan bahasa alami (NLP) yang berfokus pada identifikasi dan klasifikasi ekspresi opini, emosi, atau sikap pengguna dari teks tertulis misalnya ulasan, komentar, atau masukan pengguna ke dalam kategori seperti positif, negatif, atau netral. Sebagai contoh, penelitian pada ulasan hotel Indonesia menunjukkan bahwa teknik deep-learning mampu melakukan analisis multilevel terhadap sentimen pengguna dengan akurasi yang kompetitif. Lebih lanjut, dalam konteks aplikasi digital di Indonesia, telah ditemukan bahwa penerapan model berbasis pra-pelatihan bahasa Indonesia seperti IndoBERT memberikan peningkatan signifikan dalam representasi konteks dan hasil klasifikasi dibanding metode konvensional [5]. Dengan meningkatnya volume ulasan aplikasi berbasis pengguna di platform seperti Google Play atau App Store, maka analisis sentimen terhadap ulasan aplikasi menjadi semakin relevan demi memahami persepsi pengguna secara otomatis[6].

Dalam banyak penelitian, metode machine learning telah diterapkan secara luas untuk tugas klasifikasi sentimen karena kemampuannya dalam mempelajari pola dari data teks yang besar dan kompleks. Sebagai contoh, studi yang menganalisis ulasan aplikasi "MyTelkomsel" menggunakan algoritma seperti Support Vector Machine (SVM) dan Random Forest menunjukkan bahwa pendekatan ini mencapai akurasi hingga 83 % ketika label dibuat oleh pakar bahasa Indonesia, yang mengindikasikan bahwa pemilihan algoritma dan teknik pelabelan memegang peranan penting. Selain itu, studi lain dalam domain bahasa Indonesia membandingkan tiga algoritma Naïve Bayes, SVM, dan Random Forest untuk tugas "aspect category detection" dari ulasan konsumen dan menemukan bahwa Random Forest memberikan skor F1 terbaik (antara 84 % hingga 89 %) ketika digunakan dengan fitur TF-IDF [7]. Namun demikian, meskipun metode machine learning tradisional tersebut cukup unggul dalam sejumlah aplikasi, mereka masih memiliki keterbatasan dalam menangkap dependensi kontekstual jangka panjang dan dinamika bahasa yang kompleks, sehingga memunculkan minat terhadap model pra-latih berbasis transformer yang dapat mengatasi kelemahan tersebut [8].

Natural Language Processing (NLP) dianggap sebagai bidang yang berfokus pada bagaimana sistem komputasi dapat mengolah, memahami, dan menghasilkan bahasa manusia secara otomatis, dan beberapa tinjauan menunjukkan bahwa munculnya teknik pembelajaran mendalam (deep learning) telah memberikan lonjakan besar dalam performa tugas-NLP seperti klasifikasi teks, analisis sentimen, dan pemahaman bahasa [9]. Sementara itu, dalam lingkungan bahasa Indonesia yang memiliki ragam morfologis dan konteks budaya yang khas, penelitian menunjukkan bahwa tantangan utama meliputi ketersediaan data yang terbatas dan representasi linguistik yang belum optimal [10]. Oleh sebab itu, studi-studi terbaru mengusulkan penggunaan model pra-latih khusus bahasa Indonesia atau adaptasi multi-bahasa untuk menguatkan representasi kontekstual, yang pada akhirnya meningkatkan akurasi dan generalisasi dari sistem analisis teks dalam Bahasa Indonesia.

Arsitektur *Transformer* dipandang sebagai tonggak penting dalam perkembangan NLP modern karena kemampuannya dalam mempelajari hubungan kontekstual antar-token secara paralel

menggunakan mekanisme *self-attention* [5]. Peneliti menyebut bahwa pendekatan ini mengatasi keterbatasan arsitektur berbasis RNN dan LSTM yang sulit memproses dependensi panjang dan lambat dalam komputasi sekuensial [11]. Peneliti menyebut bahwa pendekatan ini mengatasi keterbatasan arsitektur berbasis RNN dan LSTM yang sulit memproses dependensi panjang dan lambat dalam komputasi sekuensial [12]. Selain itu, penelitian lain menegaskan bahwa *Transformer* tidak hanya mempercepat pelatihan tetapi juga meningkatkan kualitas representasi semantik karena setiap token mampu mengakses konteks global dalam satu langkah komputasi, sehingga model ini menjadi dasar bagi berbagai varian model pra-latih seperti BERT, GPT, dan RoBERTa yang kemudian diadaptasi untuk berbagai bahasa termasuk bahasa Indonesia.

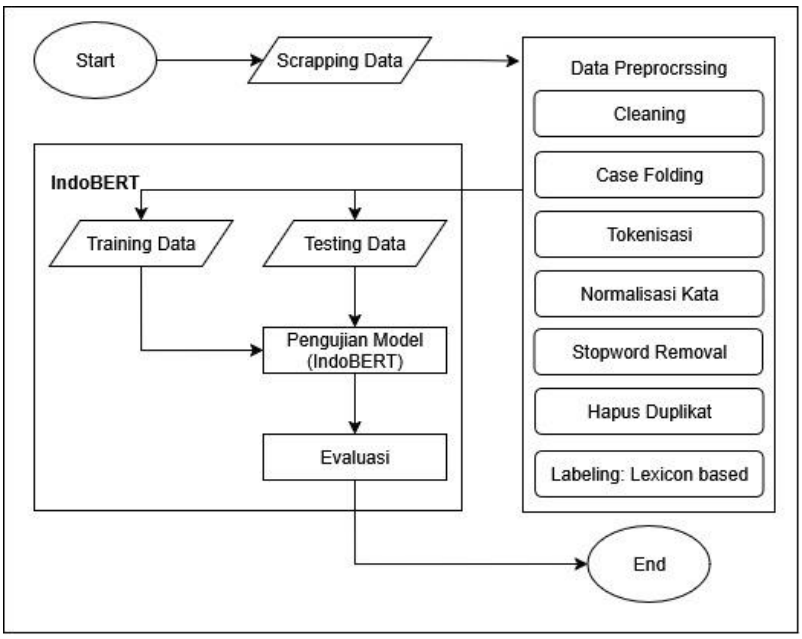
Model BERT diperkenalkan sebagai suatu arsitektur pembelajaran bahasa yang memungkinkan representasi kata secara dua arah dalam teks, dengan cara mempelajari konteks dari kiri dan kanan secara simultan dalam tahap pra-pelatihan [13]. Di Indonesia, beberapa studi menunjukkan bahwa penerapan model BERT atau model yang diadaptasi untuk Bahasa Indonesia seperti IndoBERT telah berhasil meningkatkan hasil tugas klasifikasi teks dibanding model tradisional, karena IndoBERT dilatih khusus dengan korpus Bahasa Indonesia yang mencerminkan ragam morfologi dan penggunaan sehari-hari [14]. Sebagai contoh, penelitian menunjukkan bahwa IndoBERT outperform varian multibahasa pada tugas klasifikasi umpan balik mahasiswa berbahasa Indonesia, sehingga menunjukkan bahwa penyelarasan model dengan bahasa lokal dapat memberi keuntungan signifikan [15].

Model IndoBERT dikembangkan sebagai hasil adaptasi dari arsitektur BERT yang disesuaikan dengan karakteristik linguistik Bahasa Indonesia, dengan tujuan untuk meningkatkan akurasi pemrosesan bahasa alami pada teks berbahasa lokal [1]. Beberapa penelitian di Indonesia menunjukkan bahwa IndoBERT berhasil dalam tugas analisis sentimen: misalnya, penelitian pada ulasan aplikasi menunjukkan bahwa fine-tuning IndoBERT dengan optimasi learning rate dan epoch meningkatkan performanya secara signifikan [16]. Selain itu, studi lain menunjukkan bahwa IndoBERT dengan metode "confident learning" untuk memperbaiki kelabelan data berhasil menggeser akurasi naik dari 85,15 % menjadi 86,03 % dalam tugas sentimen berbahasa Indonesia [17].

Proses *fine-tuning* dianggap sebagai tahapan penting dalam pemanfaatan model pra-latih seperti IndoBERT karena berperan dalam menyesuaikan representasi bahasa dengan domain tugas yang spesifik. Beberapa kajian menunjukkan bahwa penyesuaian bobot model pra-latih terhadap data baru dapat meningkatkan keselarasan model dengan karakteristik bahasa pengguna, termasuk variasi kosakata dan gaya penulisan pada ulasan aplikasi [3]. Penelitian lain menekankan bahwa *fine-tuning* secara domain-spesifik dapat meningkatkan akurasi klasifikasi sentimen karena representasi bahasa yang dihasilkan menjadi lebih kontekstual terhadap topik atau industri tertentu [4]. Selain itu, tahap *fine-tuning* juga memerlukan perhatian terhadap hiperparameter seperti *learning rate*, *batch size*, dan jumlah *epoch* karena ketidakseimbangan konfigurasi dapat menyebabkan *overfitting* atau kehilangan generalisasi model. Walaupun efektivitas *fine-tuning* telah dibuktikan pada berbagai domain seperti ulasan produk atau berita daring, penelitian mengenai penerapannya pada platform hiburan yang menggunakan bahasa informal seperti Twitch masih relatif terbatas. Kondisi ini menunjukkan perlunya evaluasi lebih lanjut mengenai performa IndoBERT ketika disesuaikan dengan karakteristik bahasa pengguna aplikasi hiburan digital.

2. METODE PENELITIAN

Penelitian ini dijelaskan menggunakan pendekatan kuantitatif yang berfokus pada eksperimen berbasis machine learning dengan model transformer IndoBERT. Tahapan yang dilakukan meliputi proses pengumpulan data, praprosesan, pelabelan, pembagian data, pelatihan model, evaluasi, dan analisis hasil yang saling berkaitan secara sistematis sebagaimana digambarkan pada Gambar 1.



Gambar. 1 Tahapan Penelitian

2.1. Subjek dan Sumber Data

Subjek penelitian dijelaskan berupa ulasan pengguna aplikasi Twitch yang tersedia secara publik melalui Google Play Store. Proses pengumpulan data dilakukan dengan memanfaatkan pustaka Python *google-play-scraper* menggunakan parameter bahasa Indonesia (lang='id') dan wilayah Indonesia (country='id'). Data yang diperoleh meliputi teks ulasan, tanggal publikasi, serta nilai *rating* bintang dari pengguna. Pengambilan data dilakukan dalam rentang waktu Januari 2022 hingga September 2025. Seluruh data bersifat anonim dan disimpan dalam format CSV guna mempermudah tahap pengolahan selanjutnya.

Sumber data tersebut dipilih karena ulasan pada Google Play Store dianggap dapat merepresentasikan opini autentik pengguna terhadap kualitas dan kinerja aplikasi, sehingga dinilai sesuai untuk menganalisis persepsi publik menggunakan model berbasis bahasa alami Indonesia.

2.2. Praprosesan Data

Tahap praprosesan dijelaskan bertujuan untuk meningkatkan kualitas teks sebelum digunakan dalam pelatihan model. Proses diawali dengan *data cleaning* yang mencakup penghapusan tanda baca, emoji, URL, angka, serta konversi seluruh huruf menjadi huruf kecil (*case folding*). Langkah berikutnya mencakup *text normalization*, yaitu proses pengubahan kata tidak baku menjadi bentuk baku dengan memanfaatkan kamus normalisasi, misalnya kata “gk” diubah menjadi “tidak” dan “bgt” menjadi “banget”.

Selanjutnya dilakukan tokenisasi menggunakan *tokenizer* IndoBERT berbasis WordPiece yang memecah teks menjadi token *subword* sehingga model dapat mengenali variasi morfologi bahasa Indonesia secara lebih baik. Setelah itu dilakukan *stopword removal* untuk menghapus kata umum yang tidak memiliki makna penting, dengan pengecualian terhadap kata negasi seperti “tidak” dan “bukan” agar konteks sentimen tetap terjaga. Langkah berikutnya adalah *stemming*, yaitu proses untuk mengubah kata-kata yang telah bersih ke bentuk kata dasarnya (misalnya 'mendapatkan' menjadi 'dapat') untuk mengurangi variasi kata. Terakhir, data duplikat serta teks dengan panjang kurang dari tiga kata dihapus untuk mencegah gangguan pada proses pembelajaran model.

2.3. Pelabelan Data

Proses pelabelan dijelaskan menggunakan pendekatan *lexicon-based* dengan memanfaatkan *lexicon* sentimen Bahasa Indonesia dari repositori InSet yang dikembangkan oleh Fajri pada tahun 2021. *Lexicon* InSet berisi daftar kata dalam Bahasa Indonesia yang disertai dengan skor sentimen positif dan negatif. Setiap ulasan teks diurai menjadi token menggunakan metode *tokenisasi* WordPiece, kemudian

sistem menghitung skor total dari seluruh kata yang muncul dalam ulasan berdasarkan nilai sentimen pada InSet. Ulasan diberi label positif apabila jumlah nilai kata positif lebih besar daripada negatif, dan sebaliknya, ulasan diberi label negatif apabila nilai negatif lebih dominan. Apabila jumlah keduanya seimbang atau tidak ditemukan kata yang tercantum dalam *lexicon*, ulasan tersebut diberi label netral.
Hasil akhir dari proses pelabelan digunakan sebagai *ground truth* dalam tahap *fine-tuning* model IndoBERT. Pada tahap ini, seluruh dataset berlabel dimasukkan ke dalam proses pelatihan tanpa pembagian data secara eksplisit, melainkan dengan menerapkan *internal validation* melalui parameter *validation split*. Penyesuaian parameter seperti *learning rate*, *batch size*, dan jumlah *epoch* dilakukan untuk memperoleh performa model yang optimal. Evaluasi terhadap hasil *fine-tuning* dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* berdasarkan hasil validasi model terhadap data yang sama.

2.4. Pelatihan Model IndoBERT

Model yang digunakan dijelaskan merupakan IndoBERT varian *indobert-base-p1* dari repositori HuggingFace, yaitu model *transformer pretrained* yang telah dilatih pada korpus besar berbahasa Indonesia. Model tersebut bersifat dua arah (*bidirectional*) sehingga mampu memahami konteks kalimat secara menyeluruh dari kedua arah pembacaan.
Proses *fine-tuning* dilakukan dengan menggunakan parameter sebagaimana ditunjukkan pada Tabel 1. Pelatihan model diterapkan menggunakan *optimizer* AdamW serta fungsi *loss* CrossEntropyLoss yang telah diimplementasikan secara internal oleh pustaka *Transformers*. Parameter pelatihan ditetapkan dengan *learning rate* sebesar 2e-5, *batch size* 8, jumlah *epoch* 5, dan *weight decay* 0.01. Evaluasi kinerja dilakukan pada setiap *epoch* untuk menentukan model terbaik berdasarkan nilai akurasi tertinggi.

2.5. Evaluasi dan Validasi Model

Model diuji menggunakan data uji yang belum pernah digunakan dalam proses pelatihan. Evaluasi dilakukan menggunakan empat metrik utama, yaitu accuracy, precision, recall, dan F1-score. Metrik tersebut dihitung berdasarkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) yang diperoleh dari confusion matrix. Rumus masing-masing metrik ditunjukkan sebagai berikut :

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1_Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Perhitungan dilakukan menggunakan pustaka scikit-learn dengan fungsi classification_report dan confusion_matrix. Nilai F1-score digunakan sebagai metrik utama karena mampu menyeimbangkan precision dan recall pada dataset yang distribusi kelasnya tidak seimbang.

2.6. Inferensi Model

Tahap inferensi dilakukan setelah model IndoBERT selesai melalui proses *fine-tuning* dan evaluasi. Pada tahap ini, model yang telah terlatih digunakan untuk melakukan prediksi sentimen terhadap teks ulasan baru yang belum pernah dilihat sebelumnya. Proses inferensi bertujuan untuk menguji kemampuan generalisasi model dalam mengidentifikasi polaritas sentimen berdasarkan pola linguistik yang telah dipelajari selama pelatihan.

Proses inferensi diawali dengan pemberian input berupa teks ulasan yang telah melalui tahapan praprosesan sesuai prosedur sebelumnya. Teks tersebut kemudian diubah menjadi representasi numerik melalui proses *tokenisasi* menggunakan WordPiece tokenizer, sehingga setiap kata dikonversi menjadi vektor sesuai dengan kosakata (*vocabulary*) model IndoBERT. Hasil tokenisasi selanjutnya diteruskan ke lapisan-lapisan encoder Transformer untuk menghasilkan representasi kontekstual dari keseluruhan kalimat.

Proses inferensi diawali dengan pemberian input berupa teks ulasan yang telah melalui tahapan praprosesan sesuai prosedur sebelumnya. Teks tersebut kemudian diubah menjadi representasi numerik melalui proses *tokenisasi* menggunakan WordPiece tokenizer, sehingga setiap kata dikonversi menjadi vektor sesuai dengan kosakata (*vocabulary*) model IndoBERT. Hasil tokenisasi selanjutnya diteruskan ke lapisan-lapisan encoder Transformer untuk menghasilkan representasi kontekstual dari keseluruhan kalimat.

Tahapan inferensi ini dijalankan dalam mode *evaluation* untuk memastikan tidak terjadi pembaruan parameter selama proses prediksi. Selain itu, hasil inferensi dapat divisualisasikan dalam bentuk distribusi label sentimen untuk menunjukkan kecenderungan opini pengguna terhadap aplikasi Twitch. Dengan demikian, tahap ini berfungsi sebagai validasi akhir terhadap performa model dalam skenario penerapan nyata (*real-world inference*).

3. HASIL DAN PEMBAHASAN

3.1. Hasil Penelitian

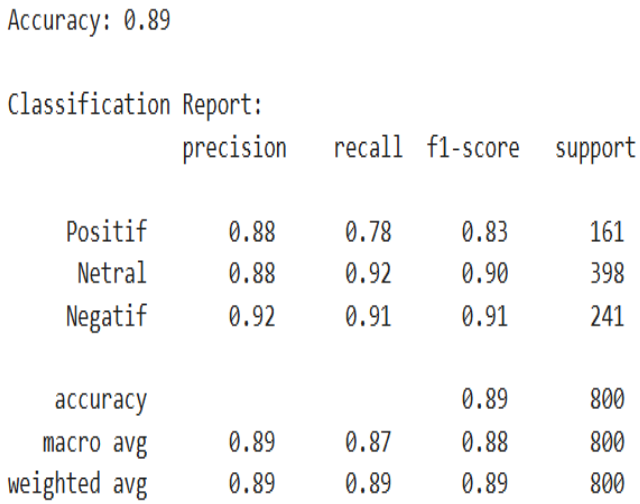
Tabel 1. Hasil Penelitian Menggunakan Model IndoBERT.

Epoch	Training Loss	Validation Loss	Accuracy
1	No Log	0.461330	0.843750
2	0.580000	0.423289	0.868750
3	0.294100	0.521620	0.882500
4	0.162200	0.590849	0.890000
5	0.050600	0.644623	0.881250

Proses pelatihan model IndoBERT menghasilkan performa yang meningkat seiring bertambahnya jumlah *epoch* hingga titik tertentu. Berdasarkan hasil yang ditunjukkan pada Tabel 1, nilai *training loss* mengalami penurunan signifikan dari 0.5800 pada *epoch* ke-2 menjadi 0.0506 pada *epoch* ke-5. Hal ini mengindikasikan bahwa model secara bertahap mampu mempelajari pola pada data latih dengan baik.

Nilai *validation loss* sempat menurun hingga mencapai titik terendah sebesar 0.4233 pada *epoch* ke-2, namun kemudian meningkat pada *epoch* selanjutnya. Peningkatan ini menunjukkan adanya indikasi awal terjadinya *overfitting*, di mana model mulai menyesuaikan diri terlalu spesifik terhadap data latih sehingga performa terhadap data validasi sedikit menurun.

Sementara itu, nilai akurasi meningkat secara konsisten dari 0.8437 pada *epoch* ke-1 hingga mencapai puncak sebesar 0.8900 pada *epoch* ke-4 sebelum sedikit menurun menjadi 0.8813 pada *epoch* ke-5. Berdasarkan tren tersebut, model terbaik diperoleh pada *epoch* ke-4 karena menunjukkan keseimbangan optimal antara *training loss* yang rendah dan akurasi validasi tertinggi.



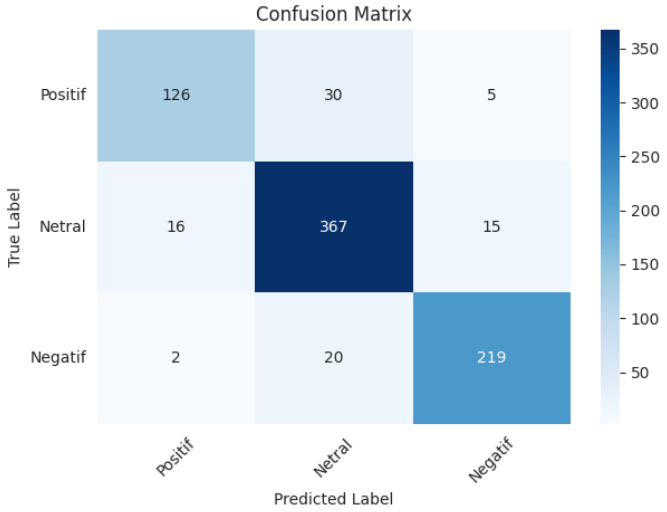
```
Accuracy: 0.89

Classification Report:
              precision    recall  f1-score   support

     Positif       0.88      0.78      0.83       161
      Netral       0.88      0.92      0.90       398
     Negatif       0.92      0.91      0.91       241

    accuracy                           0.89       800
   macro avg       0.89      0.87      0.88       800
weighted avg       0.89      0.89      0.89       800
```

Gambar 2. Classification Report.



Gambar 3. Confusion Matrix.

Untuk memvalidasi hasil pada *Classification Report* pada Gambar 2 dan menganalisis performa model secara mendalam, dilakukan perhitungan manual metrik *precision*, *recall*, dan *F1-score* untuk setiap kelas. Perhitungan ini didasarkan pada nilai *True Positive* (TP), *False Positive* (FP), dan *False Negative* (FN) yang diekstrak dari *Confusion Matrix* pada Gambar 3.

1. Analisis kelas Positif
   a. True Positive (TP): 126 (Data 'Positif' yang diprediksi benar 'Positif')
   b. False Positive (FP): 16 + 2 = 18 (Data 'Netral' dan 'Negatif' yang keliru diprediksi 'Positif')
   c. False Negative (FN): 30 + 5 = 35 (Data 'Positif' yang keliru diprediksi 'Netral' atau 'Negatif')
   d. Support**:** 126 + 35 = 161
   e. $Precision = \frac{126}{(126+18)} = \frac{126}{144} = 0.875 \approx 0.88$
   f. $Recall = \frac{126}{(126+35)} = \frac{126}{161} = 0.7826 \approx 0.78$
   g. $F1 - Score = 2\ x\ \frac{(0.875\ x\ 0.7826}{(0.875+0.7826)} = 0.826 \approx 0.83$
2. Analisa Kelas Netral
   a. True Positive (TP): 367 (Data 'Netral' yang diprediksi benar 'Netral')
   b. False Positive (FP): 30 + 20 = 50 (Data 'Positif' dan 'Negatif' yang keliru diprediksi 'Netral')
   c. False Negative (FN): 16 + 15 = 31 (Data 'Netral' yang keliru diprediksi 'Positif' atau 'Negatif')
   d. Support: 367 + 31 = 398
   e. $Precision = \frac{367}{(367+50)} = \frac{367}{417} = 0.879 \approx 0.88$
   f. $Recall = \frac{367}{(367+31)} = \frac{367}{398} = 0.922 \approx 0.92$
   g. $F1 - Score = 2\ x\ \frac{(0.879\ x\ 0.922)}{(0.879+0.922)} = 0.899 \approx 0.90$
3. Analisis Kelas "Negatif"
   a. True Positive (TP): 219 (Data 'Negatif' yang diprediksi benar 'Negatif')
   b. False Positive (FP): 5 + 15 = 20 (Data 'Positif' dan 'Netral' yang keliru diprediksi 'Negatif')
   c. False Negative (FN): 2 + 20 = 22 (Data 'Negatif' yang keliru diprediksi 'Positif' atau 'Netral')
   d. Support: 219 + 22 = 241
   e. $Precision = \frac{219}{(219+20)} = \frac{219}{239} = 0.916 \approx 0.92$
   f. $Recall = \frac{219}{(219+22)} = \frac{219}{241} = 0.908 \approx 0.91$

   $F1 - Score = 2\ x\ \frac{(0.916\ x\ 0.908)}{(0.916+0.908)} = 0.912 \approx 0.91$
4. Perhitungan Akurasi Keseluruhan (Overall Accuracy)
   a. Total Prediksi Benar (Diagonal): 126 (Positif) + 367 (Netral) + 219 (Negatif) = 712
   b. Total Seluruh Data (Support): 161 + 398 + 241 = 800

$$Akurasi = \frac{Total\ Prediksi\ Benar}{Total\ Seluruh\ Data}$$

$$Akurasi = \frac{712}{800} = 0.89$$

3.2. Pembahasan

3.2.1. Subjek dan Sumber Data

Tabel 2. Hasil Scraping Aplikasi Twitch di Google Play Store.

Review ID	Username	Rating	Review Text	Date
15506208-64fd-4b3e-83bb-c0f4c6600176	Pengguna Google	1	Justice for emiru	2025-10-19 23:29:23
21d26ea9-119c-4310-b282-c63edca06e16	Pengguna Google	2	Yeah, its become lacking	2025-10-19 19:17:37
Eac032f3-c22d-4093-a934-520f710f21ce	Pengguna Google	1	Verifikasi nomor kode nya tidak muncul	2025-10-19 17:03:02
F5c2f46e-5cda-48ad-aa2a-c0f2255c4c2e	Pengguna Google	3	App glitch makes touchscreen untouchable and l…	2025-10-17 15:42:38
Bdf67863-fbb3-41af-b038-2e93f83228db	Pengguna Google	5	Hey bung kenapa saya tidak pernah mendapatkan…	2025-10-16 23:54:17

Berdasarkan Tabel 2, proses *scraping* berhasil memperoleh data ulasan pengguna aplikasi Twitch yang terdiri atas beberapa atribut penting, seperti *review ID, username, rating, review text,* dan *tanggal ulasan*. Keberadaan atribut-atribut tersebut menunjukkan bahwa dataset tidak hanya memuat informasi mengenai isi ulasan, tetapi juga menyediakan informasi pendukung yang dapat digunakan untuk mengidentifikasi karakteristik data.

Dari seluruh atribut yang diperoleh, penelitian ini memanfaatkan *review text* sebagai sumber utama analisis sentimen, sedangkan *rating* digunakan sebagai informasi pendukung untuk memahami kecenderungan opini pengguna. Data yang berasal langsung dari Google Play Store mencerminkan pengalaman aktual pengguna setelah menggunakan aplikasi sehingga memiliki tingkat representatif yang tinggi terhadap kualitas layanan Twitch.

Karakteristik data yang bersifat *user-generated content* juga membuat penelitian ini lebih relevan dibandingkan data yang diperoleh melalui survei, karena opini yang diberikan pengguna muncul secara spontan tanpa intervensi peneliti. Oleh karena itu, dataset hasil *scraping* dinilai telah memenuhi kebutuhan penelitian sebagai dasar proses pelabelan, praprosesan, hingga pelatihan model IndoBERT.

3.2.2. Praprosesan Data

Tabel 3. Hasil proses *cleaning* pada data ulasan pengguna aplikasi Twitch.

Review Text	cleaning
Justice for emiru	Justice for emiru
Yeah, its become lacking	Yeah its become lacking
Verifikasi nomor kode nya tidak muncul muncul	Verifikasi nomor kode nya tidak muncul muncul
App glitch makes touchscreen untouchable and l…	App glitch makes touchscreen untouchable and l…
Hey bung kenapa saya tidak pernah mendapatkan …	Hey bung kenapa saya tidak pernah mendapatkan …

Tabel 3 menunjukkan bahwa proses *cleaning* berhasil menghilangkan karakter yang tidak

relevan, seperti URL, emoji, simbol, dan tanda baca. Penghapusan elemen-elemen tersebut menghasilkan teks yang lebih bersih sehingga mengurangi *noise* yang dapat memengaruhi proses pembelajaran model. Dengan data yang lebih terstruktur, model IndoBERT dapat lebih mudah mempelajari pola bahasa yang berkaitan dengan sentimen pengguna.

1. Proses *case folding* untuk menyeragamkan bentuk huruf pada teks ulasan.

Tabel 4. Hasil Proses C*ase Folding* Pada Data Ulasan Pengguna Aplikasi Twitch.

cleaning	case_folding
Justice for emiru	justice for emiru
Yeah its become lacking	yeah its become lacking
Verifikasi nomor kode nya tidak muncul muncul	verifikasi nomor kode nya tidak muncul muncul
App glitch makes touchscreen untouchable and l…	app glitch makes touchscreen untouchable and l…
Hey bung kenapa saya tidak pernah mendapatkan …	hey bung kenapa saya tidak pernah mendapatkan …

Berdasarkan Tabel 4, seluruh huruf pada data ulasan berhasil dikonversi menjadi huruf kecil (*lowercase*). Proses ini bertujuan untuk menyeragamkan bentuk penulisan sehingga kata dengan makna yang sama tidak diperlakukan sebagai token yang berbeda. Konsistensi tersebut membantu proses tokenisasi dan meningkatkan kualitas representasi teks pada tahap pelatihan model.

2. Proses normalisasi untuk menyeragamkan bentuk penulisan kata tidak baku.

Tabel 5. Hasil Proses *Normalisasi* Pada Data Ulasan Pengguna Aplikasi Twitch.

case_forlding	normalisasi
justice for emiru	justice for emiru
yeah its become lacking	yeah its become lacking
verifikasi nomor kode nya tidak muncul muncul	verifikasi nomor kode ya tidak muncul muncul
app glitch makes touchscreen untouchable and l…	app glitch makes touchscreen untouchable and l…
hey bung kenapa saya tidak pernah mendapatkan …	hey bung kenapa saya tidak pernah mendapatkan …

Tabel 5 memperlihatkan bahwa kata-kata tidak baku berhasil diubah menjadi bentuk baku, misalnya "gk" menjadi "tidak". Tahapan ini penting karena ulasan pengguna aplikasi umumnya menggunakan bahasa informal. Dengan normalisasi, model memperoleh representasi kata yang lebih konsisten sehingga mampu memahami konteks kalimat dengan lebih baik.

3. Proses tokenisasi untuk memecah kalimat menjadi unit kata.

Tabel 6. Hasil Proses Tokenisasi Pada Data Ulasan Pengguna Aplikasi Twitch.

normalisasi	tokenize
justice for emiru	[justice, for, emiru]
yeah its become lacking	[yeah, its, become, lacking]
verifikasi nomor kode ya tidak muncul muncul	[verifikasi, nomor, kode, ya, tidak, muncul, muncul]
app glitch makes touchscreen untouchable and l…	[app, glitch, makes, touchscreen, untouchable,…]
hey bung kenapa saya tidak pernah mendapatkan …	[hey, bung, kenapa, saya, tidak, pernah, mendapatkan …]

Berdasarkan Tabel 6, proses tokenisasi berhasil memecah setiap kalimat menjadi kumpulan

token menggunakan tokenizer WordPiece milik IndoBERT. Berbeda dengan tokenisasi konvensional yang hanya memisahkan kata berdasarkan spasi, WordPiece memungkinkan model mengenali kata yang tidak terdapat pada kosakata utama melalui pembentukan subword. Pendekatan ini membantu IndoBERT mempertahankan informasi semantik pada kata-kata baru maupun kata tidak baku yang sering muncul pada ulasan pengguna aplikasi. Dengan demikian, proses tokenisasi berperan penting dalam menghasilkan representasi teks yang lebih kontekstual sehingga meningkatkan kemampuan model dalam memahami makna setiap ulasan.

4. Proses stopword removal untuk menghilangkan kata yang tidak bermakna signifikan

Tabel 7. Hasil Proses *Stopword Removal* Pada Data Ulasan Pengguna Aplikasi Twitch.

tokenize	stopword removal
[justice, for, emiru]	[justice, for, emiru]
[yeah, its, become, lacking]	[yeah, its, become, lacking]
[verifikasi, nomor, kode, ya, tidak, muncul, muncul]	[verifikasi, nomor, kode, ya, muncul, muncul]
[app, glitch, makes, touchscreen, untouchable,…]	[app, glitch, makes, touchscreen, untouchable,…]
[hey, bung, kenapa, saya, tidak, pernah, mendapatkan …]	[hey, kode, verifikasi, via, sms, telepon…]

Tabel 7 menunjukkan bahwa proses stopword removal berhasil menghilangkan kata-kata yang memiliki frekuensi kemunculan tinggi tetapi tidak memberikan kontribusi yang signifikan terhadap penentuan sentimen. Penghapusan stopword menghasilkan teks yang lebih ringkas sehingga model dapat lebih berfokus pada kata-kata yang mengandung informasi penting mengenai opini pengguna. Namun demikian, kata negasi seperti *tidak* dan *bukan* tetap dipertahankan karena memiliki pengaruh besar terhadap perubahan polaritas sentimen. Oleh karena itu, tahapan ini membantu meningkatkan kualitas fitur yang dipelajari oleh model IndoBERT selama proses pelatihan.

5. Proses stemming untuk mengembalikan kata ke bentuk dasarnya

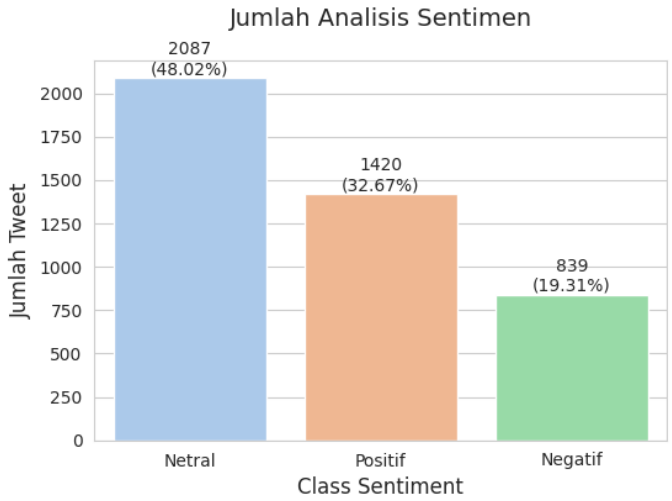
Tabel 8. Hasil Proses *Stemming* Pada Data Pada Data Ulasan Pengguna Aplikasi Twitch.

stopword removal	stemming_data
[justice, for, emiru]	justice for emiru
[yeah, its, become, lacking]	yeah its become lacking
[verifikasi, nomor, kode, ya, muncul, muncul]	verifikasi nomor kode ya muncul muncul
[app, glitch, makes, touchscreen, untouchable,…]	app glitch makes touchscreen untouchabe and l…
[hey, kode, verifikasi, via, sms, telepon…]	hey kode verifikasi via sms telepon apakah hub…

Berdasarkan Tabel 8, proses stemming berhasil mengubah berbagai kata berimbuhan menjadi bentuk kata dasar sehingga variasi bentuk kata dapat diminimalkan. Sebagai contoh, beberapa kata yang memiliki makna sama tetapi berbeda bentuk morfologinya direduksi menjadi satu bentuk dasar. Penyederhanaan tersebut membantu mengurangi kompleksitas kosakata dan meningkatkan konsistensi data. Dengan jumlah variasi kata yang lebih sedikit, model IndoBERT dapat lebih mudah mengenali pola hubungan antar kata sehingga proses klasifikasi sentimen menjadi lebih efektif.

3.2.3. Pelabelan Data

Distribusi hasil pelabelan menunjukkan bahwa kelas netral mendominasi dataset, diikuti kelas positif dan negatif. Kondisi tersebut mengindikasikan bahwa sebagian besar pengguna memberikan ulasan yang bersifat informatif dibandingkan ekspresif. Ketidakseimbangan distribusi kelas berpotensi memengaruhi proses pembelajaran model, terutama pada kelas dengan jumlah data yang lebih sedikit. Meskipun demikian, pendekatan lexicon-based tetap mampu menghasilkan ground truth yang konsisten sebagai dasar proses fine-tuning IndoBERT.



Gambar 4. Distribusi Sentimen Dataset Hasil Pelabelan.

Berdasarkan hasil proses pelabelan data, diperoleh tiga kategori sentimen. Sentimen 'Netral' ditemukan mendominasi dataset dengan jumlah 2.087 ulasan, yang setara dengan 48,02% dari total data. Kategori 'Positif' menempati urutan kedua terbesar dengan perolehan 1.420 ulasan (32,67%). Sementara itu, kategori 'Negatif' memiliki jumlah paling sedikit di antara ketiganya, yaitu sebanyak 839 ulasan (19,31%).

3.2.4. Pelatihan Model IndoBERT

Konfigurasi hyperparameter yang digunakan menunjukkan proses pelatihan yang stabil. Learning rate sebesar 2e-5 memungkinkan penyesuaian bobot model berlangsung secara bertahap sehingga risiko perubahan parameter yang terlalu besar dapat dihindari. Peningkatan akurasi hingga epoch keempat menunjukkan bahwa model berhasil mempelajari pola sentimen pada data ulasan. Meningkatnya validation loss setelah epoch keempat mengindikasikan gejala awal overfitting sehingga pemilihan model terbaik pada epoch keempat dinilai tepat.

Tabel 9. Konfigurasi Hyperparameter untuk *Fine-Tuning*

Parameter	Value
Base Model	indobenchmark/indobert-base-p1
Learning Rate	2e-5
Number of Epochs	5
Training Batch Size	8
Evaluation Batch Size	16
Weight Decay	0.01
Evaluation Strategy	epoch
Save Strategy	epoch
Best Model Metric	accuracy
Load Best Model at End	True

Berdasarkan Tabel 9, konfigurasi hyperparameter dipilih untuk memperoleh keseimbangan antara kualitas pembelajaran model dan efisiensi komputasi. Learning rate sebesar 2×10^{-5} dipilih agar proses pembaruan bobot berlangsung stabil, sedangkan penggunaan batch size 8 disesuaikan dengan kapasitas memori GPU yang digunakan. Pemilihan lima epoch dianggap memadai karena berdasarkan hasil pelatihan model telah mencapai akurasi tertinggi pada epoch keempat sebelum mulai menunjukkan indikasi overfitting.

3.2.5. Evaluasi dan Validasi Model

Hasil evaluasi menunjukkan bahwa model IndoBERT mampu mengklasifikasikan sentimen dengan baik dengan akurasi sebesar 89% dan weighted average F1-score sebesar 89%. Berdasarkan confusion matrix, sebagian besar data berhasil diklasifikasikan dengan benar. Kelas negatif memperoleh performa terbaik karena penggunaan kosakata bernada negatif cenderung lebih eksplisit sehingga lebih mudah dikenali model. Sebaliknya, beberapa data positif masih diprediksi sebagai netral karena memiliki karakteristik bahasa yang relatif mirip. Secara keseluruhan, hasil tersebut menunjukkan bahwa IndoBERT mampu membangun representasi kontekstual yang baik pada ulasan aplikasi Twitch.

3.2.6. Inferensi Model

Visualisasi word cloud sebelum dan sesudah praprosesan menunjukkan bahwa proses praproses berhasil mengurangi kata-kata yang tidak relevan dan menonjolkan kata-kata yang memiliki hubungan langsung dengan konteks aplikasi maupun sentimen pengguna. Dengan kualitas data yang lebih baik, model IndoBERT mampu menghasilkan representasi kontekstual yang lebih akurat sehingga proses inferensi terhadap ulasan baru dapat dilakukan secara lebih andal pada skenario penggunaan nyata.



Gambar 5. Perbandingan word cloud sebelum dan sesudah praprosesan.

Visualisasi "Before Preprocessing" (sebelah kiri) menunjukkan data yang masih mentah (*raw*). Tampilan ini didominasi oleh *stopwords* atau kata-kata umum yang sering muncul namun tidak memiliki makna sentimen yang kuat (misalnya: 'yang', 'di', 'tapi', 'lagi', 'dan') serta kata-kata tidak baku atau slang (misalnya: 'gw', 'yg', 'gak', 'aja', 'banget'). Kata-kata ini, meskipun frekuensinya tinggi, tidak memberikan kontribusi signifikan terhadap penentuan polaritas sentimen dan justru dapat menjadi "noise" yang mengganggu proses analisis model.

Sebaliknya, *word cloud* "After Preprocessing" (sebelah kanan) menunjukkan perubahan yang drastis. Setelah melalui tahapan *case folding*, *stopword removal*, dan normalisasi slang, data menjadi jauh lebih bersih. Hasilnya, kata-kata kunci yang paling relevan dengan konteks aplikasi (seperti

'twitch', 'stream', 'app', 'tonton', 'live') dan kata-kata yang secara langsung membawa sentimen (seperti 'good', 'bagus', 'mantap', 'error', 'bad') menjadi jauh lebih dominan dan terlihat jelas.

Proses "pembersihan" ini sangat krusial untuk tahap inferensi. Dengan memberikan *input* yang bersih dan fokus pada fitur-fitur yang bermakna, model IndoBERT yang telah dilatih dapat melakukan tokenisasi dan ekstraksi fitur kontekstual secara lebih efektif. Hal ini memastikan bahwa prediksi sentimen yang dihasilkan pada teks baru benar-benar didasarkan pada sinyal linguistik yang relevan, sehingga meningkatkan akurasi dan keandalan model dalam skenario penerapan nyata.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan model *transformer* IndoBERT melalui proses *fine-tuning* sangat efektif untuk tugas analisis sentimen pada ulasan aplikasi Twitch berbahasa Indonesia. Model indobert-base-p1 yang telah dilatih selama 5 *epoch* berhasil mencapai performa yang kuat pada data uji, dengan perolehan akurasi keseluruhan sebesar 89% dan weighted average F1-score sebesar 89%. Analisis log pelatihan menunjukkan bahwa model terbaik diperoleh pada *epoch* ke-4, yang berhasil mengidentifikasi pola sentimen secara optimal sebelum terjadinya *overfitting*.

Meskipun model menunjukkan performa sangat baik pada kelas 'Negatif' (F1-score 0.91) dan 'Netral' (F1-score 0.90), tantangan teridentifikasi pada kelas 'Positif' (F1-score 0.83) yang sebagian keliru diklasifikasikan sebagai 'Netral'. Temuan ini membuktikan bahwa IndoBERT mampu menangani bahasa informal yang umum ditemukan dalam ulasan aplikasi dan dapat menjadi alat yang andal untuk mengevaluasi persepsi pengguna.

Untuk penelitian selanjutnya, beberapa pengembangan dapat dilakukan. Pertama, dapat dieksplorasi penggunaan teknik *data augmentation* atau metode *resampling* (seperti SMOTE) untuk menangani ketidakseimbangan data antar kelas, yang berpotensi meningkatkan performa *recall* pada kelas minoritas. Kedua, penelitian di masa depan dapat membandingkan performa model ini dengan varian IndoBERT yang lebih besar (misalnya, indobert-large) atau arsitektur *transformer* lain yang juga spesifik untuk bahasa Indonesia. Terakhir, disarankan untuk menggunakan *dataset* yang dilabeli secara manual oleh anotator manusia sebagai *ground truth* untuk menguji apakah hal tersebut dapat lebih meningkatkan akurasi dan keandalan model dibandingkan pelabelan berbasis *lexicon*.

DAFTAR PUSTAKA

[1] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," *COLING 2020 - 28th Int. Conf. Comput. Linguist. Proc. Conf.*, pp. 757–770, 2020, doi: 10.18653/v1/2020.coling-main.66.

[2] K. S. Nugroho, A. Y. Sukmadewa, H. Wuswilahaken DW, F. A. Bachtiar, and N. Yudistira, "BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews," in *6th International Conference on Sustainable Information Engineering and Technology 2021*, New York, NY, USA: ACM, Sep. 2021, pp. 258–264. doi: 10.1145/3479645.3479679.

[3] S. Apriliani, A. Erfina, and C. Warman, "Fine-Tuned IndoBERT for Aspect-Based Sentiment Analysis of Indonesian Five-Star Hotel Reviews," vol. 14, pp. 437–445, 2025.

[4] A. N. Sianipar, Yuhelmi, and M. Devega, "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI DANA DAN LINKAJA MENGGUNAKAN METODE NAÏVE BAYES," *Zo. J. Sist. Inf.*, vol. 6, no. 2, pp. 510–522, Jun. 2024, doi: 10.31849/zn.v6i2.19907.

[5] R. Kusumaningrum, I. Z. Nisa, R. Jayanto, R. P. Nawangsari, and A. Wibowo, "Deep learning-based application for multilevel sentiment analysis of Indonesian hotel reviews," *Heliyon*, vol. 9, no. 6, p. e17147, 2023, doi: 10.1016/j.heliyon.2023.e17147.

[6] 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, pp. 113–117, 2023, doi: 10.14569/IJACSA.2023.0140813.

[7] A. Algiffary and T. Sutabri, "Indonesian Journal of Computer Science," *Indones. J. Comput. Sci.*, vol. 12, no. 2, pp. 284–301, 2023, [Online]. Available: http://ijcs.stmikindonesia.ac.id/ijcs/index.php/ijcs/article/view/3135

[8] S. A. P. Perdana, T. Bharata Aji, and R. Ferdiana, "Aspect Category Classification dengan Pendekatan Machine Learning Menggunakan Dataset Bahasa Indonesia (Aspect Category

Classification with Machine Learning Approach Using Indonesian Language Dataset)," *J. Nas. Tek. Elektro dan Teknol. Inf.* |, vol. 10, no. 3, pp. 229–235, 2021.

[9] X. Shu, Y. Yang, and B. Wu, "Adaptive segmentation model for liver CT images based on neural network and level set method," *Neurocomputing*, vol. 453, pp. 438–452, Sep. 2021, doi: 10.1016/j.neucom.2021.01.081.

[10] A. N. Ma'aly, D. Pramesti, A. D. Fathurahman, and H. Fakhrurroja, "Exploring Sentiment Analysis for the Indonesian Presidential Election Through Online Reviews Using Multi-Label Classification with a Deep Learning Algorithm," *Inf.*, vol. 15, no. 11, pp. 1–33, 2024, doi: 10.3390/info15110705.

[11] Y. Dou, Y. Yun, Q. Gao, and X. Zhang, "Self-representation and matrix factorization based multi-view clustering," *Neurocomputing*, vol. 459, pp. 395–407, Oct. 2021, doi: 10.1016/j.neucom.2021.06.092.

[12] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning--based Text Classification," *ACM Comput. Surv.*, vol. 54, no. 3, pp. 1–40, Apr. 2022, doi: 10.1145/3439726.

[13] M. B. M. Amin *et al.*, "Deteksi Spam Berbahasa Indonesia Berbasis Teks Menggunakan Model Bert," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 6, pp. 1291–1302, 2024, doi: 10.25126/jtiik.2024118121.

[14] S. Development, G. Using, and T. Learning, "JURNAL RESTI Automatic Classification of Multilanguage Scientific Papers to the," vol. 9, no. 3, pp. 604–612, 2025.

[15] C. J. L. Tobing, IGN Lanang Wijayakusuma, and Luh Putu Ida Harini, "Perbandingan Kinerja IndoBERT dan MBERT Untuk Deteksi Berita Hoaks Politik dalam Bahasa Indonesia," *JST (Jurnal Sains dan Teknol.*, vol. 14, no. 1, pp. 114–123, 2025, doi: 10.23887/jstundiksha.v14i1.92126.

[16] M. N. Zaidan, Y. Sibaroni, and S. S. Prasetyowati, "LEARNING RATE AND EPOCH OPTIMIZATION IN THE FINE-TUNING PROCESS FOR INDOBERT'S PERFORMANCE ON SENTIMENT ANALYSIS OF MYTELKOMSEL APP REVIEWS," *J. Tek. Inform.*, vol. 5, no. 5, pp. 1443–1450, Oct. 2024, doi: 10.52436/1.jutif.2024.5.5.2396.

[17] D. Al Akhdaan, Taufik Edy Sutanto, and Muhaza Liebenlito, "Confident Learning pada IndoBERT: Peningkatan Kinerja Klasifikasi Sentimen," *Indones. J. Comput. Sci.*, vol. 13, no. 5, pp. 8350–8359, 2024, doi: 10.33022/ijcs.v13i5.4401.

***ZONAsi**: Jurnal Sistem Informasi*
Is licensed under a Creative Commons Attribution International (CC BY-SA 4.0)