



Analisis Sentimen Komentar *Trailer* Youtube Film *Pelangi di Mars* Menggunakan IndoBERT

Siska Amelia*¹, Migel Orvin Febryan², Muhammad Rizky Pribadi³

^{1,2,3}Universitas Multi Data Palembang, Palembang, Indonesia

E-mail : siskaamelia_2327250062@mhs.mdp.ac.id *

*Penulis Korespondensi

Received 15 May 2026; Revised 25 May 2026; Accepted 15 Juni 2026

Abstrak - Pertumbuhan industri perfilman Indonesia meningkatkan jumlah opini penonton pada platform digital, khususnya YouTube. Kolom komentar pada trailer film dapat dimanfaatkan untuk mengetahui respons dan persepsi penonton terhadap sebuah film. Namun, volume komentar yang besar serta penggunaan bahasa tidak baku pada media sosial menyebabkan analisis manual menjadi kurang efektif. Penelitian ini bertujuan untuk menganalisis sentimen komentar pengguna pada trailer film *Pelangi di Mars* di YouTube menggunakan model *pre-trained* IndoBERT. Data penelitian diperoleh melalui proses *web scraping* komentar YouTube pada periode 24 November 2025 hingga 5 Mei 2026 dan menghasilkan 2656 komentar. Setelah proses seleksi data, diperoleh 1898 komentar yang digunakan dalam penelitian. Tahap *preprocessing* meliputi *cleaning text*, *case folding*, *tokenizing*, normalisasi kata tidak baku, *stopword removal*, dan *stemming*. Proses klasifikasi sentimen dilakukan menggunakan model *Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis* pada Hugging Face Transformers tanpa proses *fine-tuning*. Hasil klasifikasi menunjukkan terdapat 839 komentar positif, 546 komentar negatif, dan 513 komentar netral. Evaluasi model menghasilkan nilai *accuracy* sebesar 72,23%, *precision* sebesar 80,86%, *recall* sebesar 72,23%, dan *F1-score* sebesar 74,11%. Hasil penelitian menunjukkan bahwa mayoritas pengguna memberikan respons positif terhadap trailer film *Pelangi di Mars*, serta model IndoBERT memiliki kemampuan yang cukup baik dalam melakukan klasifikasi sentimen komentar berbahasa Indonesia pada media sosial.

Kata Kunci: Analisis Sentimen, IndoBERT, YouTube, Trailer Film, Natural Language Processing.

Abstract - The growth of the Indonesian film industry has increased audience opinions on digital platforms, particularly YouTube. Comment sections on movie trailers can be utilized to understand audience responses and perceptions toward a film. However, the large volume of comments and the use of informal language on social media make manual analysis inefficient. This study aims to analyze the sentiment of user comments on the *Pelangi di Mars* movie trailer on YouTube using a *pre-trained* IndoBERT model. The research data were collected through a YouTube comment *web scraping* process from November 24, 2025 to May 5, 2026, resulting in 2656 comments. After data selection, 1898 comments were used in this study. The preprocessing stage included *cleaning text*, *case folding*, *tokenizing*, slangword normalization, *stopword removal*, and *stemming*. Sentiment classification was performed using the *Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis* model from Hugging Face Transformers without a *fine-tuning* process. The classification results showed 839 positive comments, 546 negative comments, and 513 neutral comments. Model evaluation produced an accuracy of 72,23%, precision of 80,86%, recall of 72,23%, and F1-score of 74,11%. The results indicate that most users gave positive responses toward the *Pelangi di Mars* movie trailer, and the IndoBERT model demonstrated fairly good performance in classifying Indonesian-language social media comments.

Keywords: Sentiment Analysis, IndoBERT, YouTube, Movie Trailer, Natural Language Processing.



1. PENDAHULUAN

Industri perfilman Indonesia mengalami pertumbuhan yang signifikan dalam beberapa tahun terakhir, ditandai dengan meningkatnya jumlah produksi film lokal yang mendapatkan perhatian luas dari masyarakat. Pertumbuhan industri perfilman Indonesia yang pesat turut meningkatkan jumlah ulasan penonton di berbagai platform digital (Lestari & Pasaribu, 2026). Salah satu film yang menarik perhatian publik adalah *Pelangi di Mars*, yang mendapatkan respons beragam dari penonton melalui berbagai kanal digital, termasuk YouTube.

YouTube merupakan salah satu platform berbagi video terbesar yang telah berkembang menjadi ruang publik tempat masyarakat menyampaikan opini secara terbuka, termasuk melalui kolom komentar pada trailer film. Besarnya volume komentar dan karakteristik bahasa yang tidak terstruktur menyebabkan analisis opini secara manual menjadi kurang efektif (Apriyanto et al., 2026). Kondisi ini menjadi permasalahan nyata dalam upaya memahami persepsi awal penonton film *Pelangi di Mars* secara otomatis dan akurat.

Analisis sentimen merupakan salah satu cabang *Natural Language Processing* (NLP) yang digunakan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini subjektif dalam teks ke dalam kategori positif, negatif, atau netral (Achmad et al., 2023). Analisis sentimen memiliki beberapa tingkatan, yaitu tingkat dokumen, tingkat kalimat, dan tingkat aspek, yang digunakan untuk mengetahui polaritas opini pada suatu data teks (Walasary, 2022). Dalam penelitian berbasis media sosial, teknik NLP terbukti mampu meningkatkan efisiensi analisis data dengan akurasi yang lebih baik dibandingkan metode konvensional (Sihombing et al., 2025).

Beberapa penelitian sebelumnya menunjukkan bahwa metode *machine learning* tradisional seperti Naïve Bayes dan Support Vector Machine (SVM) masih memiliki keterbatasan dalam memahami konteks bahasa, khususnya pada komentar media sosial yang bersifat tidak formal (Ghosh & Sur, 2025). Oleh karena itu, pendekatan *deep learning* berbasis Transformer seperti BERT (*Bidirectional Encoder Representations from Transformers*) mulai banyak digunakan karena mampu memahami konteks dua arah dalam kalimat (Min et al., 2024). BERT merupakan model bahasa pra-latih berbasis Transformer yang dapat di-*fine tuning* untuk berbagai tugas NLP, termasuk analisis sentimen, klasifikasi teks, dan pemahaman bahasa alami (Min et al., 2024).

Dalam konteks Bahasa Indonesia, IndoBERT dikembangkan sebagai adaptasi BERT yang dilatih menggunakan korpus teks Bahasa Indonesia dalam skala besar (Purnomo & Sutopo, 2024). IndoBERT terbukti efektif dalam berbagai tugas NLP seperti klasifikasi emosi dan analisis sentimen (Ahmadian et al., 2024; Purnomo & Sutopo, 2024). Penelitian sebelumnya menunjukkan bahwa IndoBERT mampu mengungguli model multibahasa maupun metode klasifikasi tradisional dengan tingkat akurasi yang tinggi pada dataset berbahasa Indonesia (Koto et al., 2021). Bahkan, hasil penelitian menunjukkan IndoBERT mencapai akurasi sebesar 95%, lebih tinggi dibandingkan BERT sebesar 92%, SVM sebesar 88%, dan Random Forest sebesar 85% (Setiawan et al., 2025). Temuan tersebut menunjukkan bahwa IndoBERT memiliki kemampuan yang baik dalam memahami konteks linguistik Bahasa Indonesia.

Dalam proses pengumpulan data penelitian, teknik *web scraping* digunakan untuk mengambil komentar dari platform YouTube secara otomatis. *Web scraping* merupakan teknik ekstraksi data dari halaman web menggunakan skrip pemrograman untuk memperoleh informasi dalam jumlah besar secara efisien (Yuda, 2025). Pada penelitian analisis sentimen media sosial, teknik ini banyak dimanfaatkan untuk memperoleh data komentar YouTube yang kemudian diproses dalam tahapan *text mining* dan klasifikasi sentimen (Modami et al., 2025).

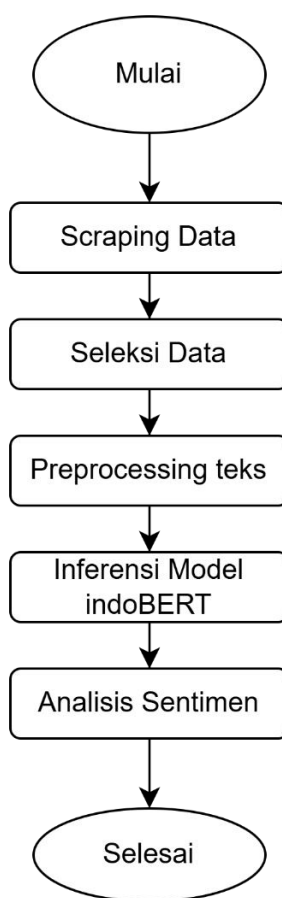
Berdasarkan penelitian terdahulu tersebut, penelitian ini memanfaatkan model pre-trained IndoBERT sebagai pendekatan aplikatif untuk melakukan analisis sentimen komentar trailer film *Pelangi di Mars* di platform YouTube. Penelitian ini tidak berfokus pada



pengembangan arsitektur model maupun proses fine-tuning, melainkan pada implementasi model Transformer berbahasa Indonesia dalam mengklasifikasikan sentimen komentar pengguna media sosial. Dengan demikian, penelitian ini diharapkan dapat memberikan gambaran mengenai kemampuan model IndoBERT dalam memahami opini pengguna pada domain komentar YouTube berbahasa Indonesia.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan *Natural Language Processing* (NLP) untuk melakukan analisis sentimen terhadap komentar pengguna pada trailer film *Pelangi di Mars* di platform YouTube menggunakan model IndoBERT. Tahapan penelitian dilakukan secara bertahap mulai dari pengumpulan data hingga analisis hasil klasifikasi sentimen. Alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Data penelitian diperoleh dari komentar pengguna pada trailer film *Pelangi di Mars* di platform YouTube. Proses pengumpulan data dilakukan menggunakan teknik web scraping dengan bantuan library Python untuk mengambil komentar secara otomatis dari video trailer yang diteliti. Pengambilan data dilakukan pada rentang waktu 24 November 2025 hingga 5 Mei 2026 untuk memperoleh komentar yang relevan dan merepresentasikan respons penonton terhadap trailer film tersebut. Data yang diperoleh berupa teks komentar pengguna yang kemudian disimpan dalam bentuk dataset untuk proses analisis lebih lanjut. Hasil proses



pengumpulan data menghasilkan sebanyak 2656 komentar yang digunakan sebagai data awal penelitian. Hasil pengumpulan data dapat dilihat pada Tabel 1.

Tabel 1. Hasil Scraping

<i>author</i>	<i>comment</i>
@obsценewitch7150	Aku pikirnya malah kyk iklan komersil, taunya film. Ayo dong harus sukses ini film
@taehyungwibuu	Terimakasih sudah bikin jiwa 27th ini bermimpi kembali
@adimasda	Gokil unreal engine 5 bisa di pake buat bikin film besar, biasanya di pake buat game, wah gak sabar buat liat full filmnya Wajib Kita Tonton Sih
@kuspınastıkuspınasti3512	teaser Pelangi di Mars memberi impresi awal yang kuat: visualnya memukau, konsepnya ambisius, dan ceritanya potensial menjadi tontonan yang menarik bagi penonton Indonesia serta menandakan langkah baru dalam kualitas sinema lokal. ✨
@sufinatun_nazwa	Ini animasi atau live action full syutingnya soalnya mirip asli weh kalo animasi sih semoga aja bisa memperbaiki reputasi animasi Indonesia gegara animasi nyari celaka di hutan berkedok nyari bendera (YTТА)
@isaerlando	Filmnya jelek ternyata gasebagus di trailer awkwkwk banyak AI gajelas sama sound campur aduk

2.2. Seleksi Data

Setelah proses pengumpulan data selesai, dilakukan tahap seleksi data untuk memastikan kualitas dataset yang digunakan dalam penelitian. Seleksi data dilakukan dengan menghapus komentar duplikat, komentar kosong, komentar spam, komentar yang hanya berisi emoji atau simbol, serta komentar yang tidak relevan dengan isi trailer film. Proses seleksi data ini bertujuan untuk meningkatkan kualitas dataset sehingga hasil klasifikasi sentimen menjadi lebih akurat dan relevan. Setelah tahap seleksi data dilakukan, jumlah dataset yang digunakan dalam penelitian menjadi sebanyak 1898 komentar. Dataset tersebut kemudian digunakan pada tahap preprocessing teks sebelum diproses menggunakan model IndoBERT.

2.3. Preprocessing Teks

Tahap preprocessing teks dilakukan untuk membersihkan dan mempersiapkan data komentar sebelum diproses menggunakan model IndoBERT. Pada penelitian berbasis Transformer seperti IndoBERT, preprocessing dilakukan secara lebih ringan karena model mampu memahami konteks kalimat secara kontekstual (Chai, 2023). Tahap preprocessing untuk klasifikasi sentimen meliputi *cleaning text*, *case folding*, *tokenizing*, dan normalisasi kata tidak baku (*slangword normalization*). Proses *cleaning text* dilakukan dengan menghapus *mention*, *hashtag*, URL, angka, tanda baca, emoji, dan karakter khusus menggunakan *regular expression*. Selanjutnya dilakukan *case folding* untuk mengubah seluruh huruf menjadi *lowercase*, *tokenizing* untuk memisahkan kalimat menjadi token kata, serta normalisasi kata tidak baku seperti “gk” menjadi “tidak” dan “yg” menjadi “yang”.

Selain itu, preprocessing tambahan berupa *stopword removal* dan *stemming* diterapkan pada visualisasi *word cloud*. *Stopword removal* digunakan untuk menghapus kata-kata umum



yang kurang memiliki makna penting, sedangkan proses *stemming* menggunakan library Sastrawi dilakukan untuk mengubah kata berimbuhan menjadi bentuk dasar agar kata-kata pada *word cloud* lebih representatif terhadap topik utama komentar pengguna. Contoh hasil preprocessing dapat dilihat pada Tabel 2.

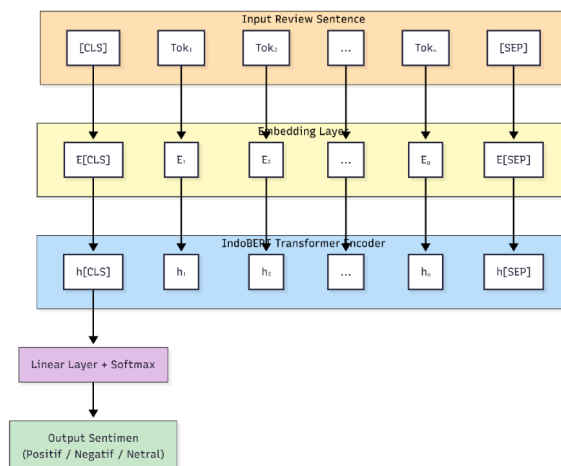
Tabel 2. Contoh Hasil preprocessing

comment	cleaned_text
PLISS GUYS NONTON, BIAR KALO INI MELEDAK NANTINYA BANYAK FILM MAKER YG BIKIN LAGI DI MASA DEPAN ♥♥♥ VISUAL yang bagus dibarengi Alur cerita yang Tak JELAS. Dialog yang berantakan. Robot nya pun berisik amat! Semoga anak bangsa, menjadi seperti karakter protagonis di film ini. yeeey..... IQ 30 Min, itu roket2 yg ninggalin Mars mau mudik ke Bumi kan ya ? Kira2 itu trayeknya cuma sampe Purwokerto apa bablas sampe Wonosobo ya ? btw ijin share Min Sebagai penggemar sci-fi movie pasti akan aku tonton filmnya, asal tahun depan aku masih idup please kali ini dikatain fomo juga aku nerima,dukung terus sampai viral,sampai bisa jadi box office gais. Film indonesia udah sampai sejauh ini,bangga banget bisa menyaksikan kebangkitan film indonesia berkualitas ♥♥ FILM TAK HANYA MELULU TENTANG Visual... tapi yang lebih Penting adalah Alur cerita yang menyenangkan dan Gampang dicerna... apalagi film yg ditujukan buat Anak anak... Moga Kedepan Film Nasional lebih baik & semakin baik dengan Harapan mampu bersaing di pentas Internasional.	pliss guys nonton biar ledak film maker bikin visual bagus bareng alur cerita dialog beranta robot berisik moga anak bangsa karakter protagonis film yeeey iq min roket ninggalin mars mudik bumi trayek sampe purwokerto bablas sampe wonosobo btw ijin share min gemar scifi movie tonton film idup please kali dikatain fomo nerimadukung viralsampai box office gais film indonesia inibangga banget saksi bangkit film indonesia kualitas film melulu visual alur cerita senang gampang cerna film anak anak moga depan film nasional harap saing pentas internasional nasional lebih baik semakin baik

2.4. Inferensi Model IndoBERT

Proses klasifikasi sentimen dilakukan menggunakan *pre-trained model* IndoBERT pada library Hugging Face Transformers tanpa proses *fine-tuning*. Model yang digunakan adalah *Aardiiiy/indobertweet-base-Indonesian-sentiment-analysis* untuk mengklasifikasikan komentar ke dalam kategori positif, negatif, dan netral (*Aardiiiy/Indobertweet-Base-Indonesian-Sentiment-Analysis · Hugging Face*, 2026).

Arsitektur IndoBERT pada penelitian ini terdiri dari tokenizer, *embedding layer*, *transformer encoder*, dan *classification layer*. Teks komentar yang telah melalui preprocessing diubah menjadi token numerik menggunakan tokenizer, kemudian diproses oleh *transformer encoder* untuk memahami hubungan kontekstual antar kata secara dua arah (*bidirectional*) (Devlin et al., 2019). Hasil representasi konteks selanjutnya diteruskan ke *classification layer* untuk menghasilkan prediksi sentimen. Arsitektur model IndoBERT dapat dilihat pada Gambar 2.



Gambar 2. Arsitektur model IndoBERT

Proses inferensi dilakukan secara otomatis terhadap setiap komentar yang telah dipreprocessing. IndoBERT dipilih karena mampu memahami konteks bahasa Indonesia dengan baik sehingga efektif digunakan pada analisis sentimen media sosial (Koto et al., 2020).

2.5. Analisis Sentimen

Hasil klasifikasi sentimen dari model IndoBERT dianalisis berdasarkan kategori sentimen positif, negatif, dan netral untuk mengetahui kecenderungan opini penonton terhadap trailer film *Pelangi di Mars* di platform YouTube. Proses analisis dilakukan dengan menghitung jumlah komentar pada setiap kategori sentimen yang dihasilkan oleh model.

Selanjutnya, hasil klasifikasi digunakan sebagai dasar untuk penyajian visualisasi data dan interpretasi sentimen. Analisis dilakukan untuk melihat distribusi sentimen serta memahami kecenderungan opini pengguna berdasarkan konteks komentar berbahasa Indonesia yang diproses menggunakan model IndoBERT. Contoh Hasil Klasifikasi Sentimen dapat dilihat pada Tabel 3.

Tabel 3. Contoh Hasil Klasifikasi Sentimen

comment	sentiment
PLISS GUYS NONTON, BIAR KALO INI MELEDAK NANTINYA BANYAK FILM MAKER YG BIKIN LAGI DI MASA DEPAN ❤️❤️❤️	Positive
VISUAL yang bagus dibarengi Alur cerita yang Tak JELAS. Dialog yang berantakan. Robot nya pun berisik amat!	Negative
Semoga anak bangsa, menjadi seperti karakter protagonis di film ini. yeeey..... IQ 30	Neutral
Min, itu roket2 yg ninggalin Mars mau mudik ke Bumi kan ya ? Kira2 itu trayeknya cuma sampe Purwokerto apa bablas sampe Wonosobo ya ? btw ijin share Min	Neutral
Sebagai penggemar sci-fi movie pasti akan aku tonton filmnya, asal tahun depan aku masih idup please kali ini dikatain fomo juga aku terima,dukung terus sampai viral,sampai bisa jadi box office gais. Film indonesia udah sampai sejauh ini,bangga banget bisa menyaksikan kebangkitan film indonesia berkualitas ❤️❤️	Positive
FILM TAK HANYA MELULU TENTANG Visual... tapi yang lebih Penting adalah Alur cerita yang menyenangkan dan Gampang dicerna... apalagi film yg ditujukan buat Anak anak... Moga Kedepan Film Nasional lebih baik & semakin baik dengan Harapan mampu bersaing di pentas Internasional	Positive



2.6. Evaluasi Model

Untuk mengukur performa model IndoBERT pada dataset penelitian ini, dilakukan proses validasi menggunakan sampel data berlabel manual sebagai ground truth. Sebanyak 200 komentar dipilih secara acak dari total 1898 komentar yang telah melalui tahap klasifikasi. Sampel tersebut kemudian dilabeli secara manual oleh anotator manusia ke dalam kategori positif, negatif, atau netral berdasarkan konteks komentar secara keseluruhan. Hasil pelabelan manual digunakan sebagai acuan (ground truth) untuk mengevaluasi prediksi model dengan menghitung nilai Confusion Matrix, Accuracy, Precision, Recall, dan F1-Score pada masing-masing kelas sentimen. Metrik-metrik tersebut dihitung menggunakan Persamaan 1, 2, 3, dan 4.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

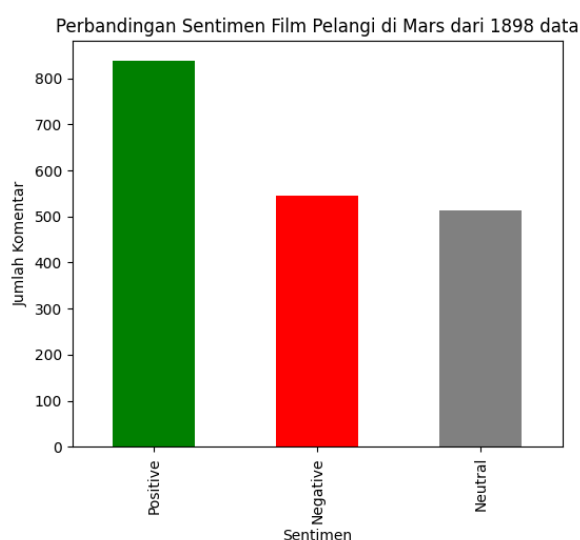
$$F1 - Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \quad (4)$$

di mana TP (True Positive), TN (True Negative), FP (False Positive), dan FN (False Negative) dihitung berdasarkan perbandingan antara hasil prediksi model dengan label manual.

3. HASIL DAN PEMBAHASAN

Hasil penelitian diperoleh dari proses analisis sentimen terhadap komentar pengguna pada trailer film *Pelangi di Mars* di platform YouTube menggunakan model IndoBERT. Data awal yang berhasil dikumpulkan melalui proses *web scraping* sebanyak 2656 komentar. Selanjutnya dilakukan tahap seleksi data sehingga diperoleh sebanyak 1898 komentar yang digunakan dalam penelitian.

Berdasarkan hasil klasifikasi sentimen yang dilakukan oleh model IndoBERT, diperoleh sebanyak 839 komentar dengan sentimen positif, 546 komentar dengan sentimen negatif, dan 513 komentar dengan sentimen netral. Hasil tersebut menunjukkan bahwa sebagian besar komentar penonton terhadap trailer film *Pelangi di Mars* memiliki kecenderungan sentimen positif. Distribusi hasil klasifikasi sentimen dapat dilihat pada Gambar 3.



Gambar 3. Distribusi hasil klasifikasi



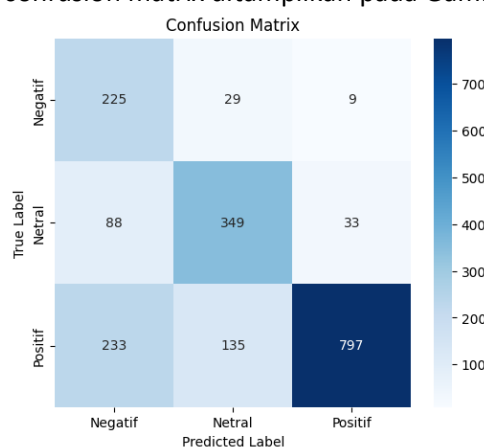
Selain itu, kata-kata seperti “animasi”, “visual”, “robot”, “trailer”, dan “sci fi” menunjukkan bahwa aspek visual, teknologi animasi, serta genre fiksi ilmiah menjadi topik utama yang banyak dibahas oleh pengguna pada kolom komentar YouTube. Visualisasi word cloud tersebut memperlihatkan bahwa perhatian penonton tidak hanya berfokus pada alur cerita film, tetapi juga pada kualitas visual dan konsep futuristik yang ditampilkan dalam trailer film *Pelangi di Mars*.

Untuk mengukur performa model IndoBERT dalam melakukan klasifikasi sentimen komentar trailer film *Pelangi di Mars*, dilakukan evaluasi menggunakan metrik accuracy, precision, recall, dan F1-score. Hasil evaluasi performa model dapat dilihat pada Tabel 4.

Tabel 4. Hasil Evaluasi Performa Model IndoBERT

Metrik	Value
Accuracy	72,23%
Precision	80,86%
Recall	72,23%
F1-Score	74,11%

Hasil evaluasi pada Tabel 4 menunjukkan bahwa model IndoBERT memiliki kemampuan yang cukup baik dalam mengklasifikasikan sentimen komentar berbahasa Indonesia pada platform YouTube. Untuk melihat distribusi hasil prediksi dan kesalahan klasifikasi pada masing-masing kategori sentimen, confusion matrix ditampilkan pada Gambar 6.



Gambar 6. Visualisasi *Confusion Matrix*

Berdasarkan confusion matrix pada Gambar 6, sebagian besar komentar positif berhasil diklasifikasikan dengan cukup baik, meskipun masih terdapat beberapa kesalahan klasifikasi pada kategori negatif dan netral akibat penggunaan bahasa informal, konteks ambigu, serta gaya bahasa sarkasme pada komentar media sosial.

4. KESIMPULAN

Berdasarkan hasil penelitian, model pre-trained IndoBERT dapat digunakan sebagai pendekatan aplikatif dalam analisis sentimen komentar YouTube berbahasa Indonesia. Hasil klasifikasi dan evaluasi menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengidentifikasi sentimen positif, negatif, dan netral pada komentar trailer film *Pelangi di Mars*. Hasil analisis juga menunjukkan bahwa mayoritas pengguna memberikan respons positif terhadap trailer film tersebut.

Meskipun penelitian ini belum melakukan proses fine-tuning maupun pengembangan arsitektur model, IndoBERT tetap mampu memberikan gambaran distribusi sentimen pengguna pada media sosial. Namun, masih terdapat beberapa kesalahan klasifikasi yang dipengaruhi oleh



penggunaan bahasa informal, konteks ambigu, dan gaya bahasa sarkasme pada komentar pengguna. Oleh karena itu, pengembangan penelitian selanjutnya dapat dilakukan melalui fine-tuning menggunakan dataset lokal serta penambahan deteksi sarkasme agar performa klasifikasi sentimen menjadi lebih optimal pada domain komentar media sosial berbahasa Indonesia.

DAFTAR PUSTAKA

- Aardiiiy/indobertweet-base-Indonesian-sentiment-analysis- Hugging Face. (2026). <https://huggingface.co/Aardiiiy/indobertweet-base-Indonesian-sentiment-analysis>
- Achmad, A. A., Iin, K., & Iska, Y. (2023). Analisis Klasifikasi Sentimen Berbasis Topik pada Ulasan Layanan Dana dan Sakuku dengan Convolutional Neural Network. *INFORMASI (Jurnal Informatika Dan Sistem Informasi)*, 15(2), 225–236. <https://doi.org/10.37424/INFORMASI.V15I2.267>
- Ahmadian, H., Abidin, T. F., Riza, H., & Muchtar, K. (2024). Hybrid Models for Emotion Classification and Sentiment Analysis in Indonesian Language. *Applied Computational Intelligence and Soft Computing*, 2024(1), 2826773. <https://doi.org/10.1155/2024/2826773>
- Apriyanto, H. T., Chamid, A. A., Meymaharani, R., Informatika, T., Kudus, U. M., Utara, J. L., Kudus, K., Youtube, K., & Jumbo, F. (2026). Perbandingan Model Decision Tree, Random Forest, Dan Svm Pada Analisis Sentimen Berbasis Aspek Komentar Film Jumbo. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 10(1), 302–309.
- Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, 29(3), 509–553. <https://doi.org/10.1017/S1351324922000213>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 1, 4171–4186.
- Ghosh, K. K., & Sur, C. (2025). *Learning to Extract Cross-Domain Aspects and Understanding Sentiments Using Large Language Models*. <https://arxiv.org/pdf/2501.08974>
- Koto, F., Lau, J. H., & Baldwin, T. (2021). IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization. *EMNLP 2021 - 2021 Conference on Empirical Methods in Natural Language Processing, Proceedings*, 10660–10668. <https://doi.org/10.18653/v1/2021.EMNLP-MAIN.833>
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*, 757–770. <https://doi.org/10.18653/v1/2020.coling-main.66>
- Lestari, S., & Pasaribu, W. S. C. (2026). Analisis Sentimen Review Film Indonesia Berdasarkan Rentang Usia dengan Simple Moving Average untuk Optimasi Sistem Rekomendasi. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, 10(1), 345–350. <https://doi.org/10.35870/jtik.v10i1.5219>
- Min, B., Ross, H., Sulem, E., Veyseh, A. P. Ben, Nguyen, T. H., Sainz, O., Agirre, E., Heintz, I., & Roth, D. (2024). Recent Advances in Natural Language Processing via Large Pre-trained Language Models: A Survey. *ACM Computing Surveys*, 56(2), 30. <https://doi.org/10.1145/3605943;CSUBTYPE:STRING:JOURNAL;SUBPAGE:STRING:BASIS>
- Modami, N., Eleazar, E., Manopo, R., Enditama, D. R., & Ayunda, A. T. (2025). Analisis Sentimen Komentar YouTube terhadap Rumor Peluncuran iPhone 17 Menggunakan Web Scraping dan Studi Komparatif Algoritma Klasifikasi. 9(2), 403–411.



- Purnomo, T. D., & Sutopo, J. (2024). Comparison of Pre-Trained Bert-Based Transformer Models For Regional Language Text Sentiment Analysis in Indonesia. *International Journal Science and Technology*, 3(3), 11–21. <https://doi.org/10.56127/IJST.V3I3.1739>
- Setiawan, V. D., Iswavigra, D. U., & Anggiratih, E. (2025). Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court's Decision Regarding the Minimum Age of Vice Presidential Candidates. *Scientific Journal of Informatics*, 12(3), 397–406. <https://doi.org/10.15294/SJI.V12I3.26320>
- Sihombing, K. E., Ineke Pakereng, M. A., Kristen Satya Wacana, U., & Artikel, R. (2025). Analisis Sentimen Komentar Youtube Terhadap Wawancara Presiden Prabowo Menggunakan Machine Learning Dan Orange Data Mining. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 4(4), 327–336. <https://doi.org/10.55123/STORAGE.V4I4.6478>
- Walasary, T. (2022). Survey Paper tentang Analisis Sentimen. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 2(1), 201–206. <https://doi.org/10.24002/konstelasi.v2i1.5378>
- Yuda, M. A. D. (2025). Implementasi Web Scraping Untuk Ekstraksi Data Penjual dan Produk Panel Surya Di E-Commerce: Indonesia. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 5(1), 155–165. <https://doi.org/10.24002/KONSTELASI.V5I1.11751>