



Evaluasi Kinerja Algoritma Naïve Bayes, SVM, dan IndoBERT pada Analisis Sentimen Ulasan Pengguna Gojek Berbasis Text Mining

I Wayan Aries Agetia^{1,*}, Ni Luh Eka Armoni¹, I Putu Ari Utama Irawan²

¹ Program Studi Manajemen Bisnis Pariwisata, Politeknik Negeri Bali, Badung, Indonesia

² Program Studi Pengelolaan Perhotelan, Politeknik Negeri Bali, Badung, Indonesia

Email: ^{1,*}ariesagetia@pnb.com, ²luhekaarmoni@pnb.ac.id, ³putuariutama@pnb.ac.id

Email Penulis Korespondensi: ariesagetia@pnb.ac.id

Abstrak—Tujuan dari penelitian ini adalah untuk mengevaluasi serta melakukan komparasi performa algoritma Naïve Bayes, Support Vector Machine (SVM), dan IndoBERT dalam tugas klasifikasi sentimen terhadap ulasan pengguna aplikasi Gojek. Data dikumpulkan melalui teknik web scraping dari Google Play Store, kemudian dikategorikan ke dalam tiga jenis sentimen, yaitu sentimen negatif, sentimen netral, dan sentimen positif. Pendekatan yang digunakan adalah kuantitatif dengan desain deskriptif-komparatif. Prosedur penelitian mencakup tahap pengumpulan data, pra-pemrosesan teks, pembagian dataset menjadi data latih dan data uji, pembangunan model, serta untuk mengevaluasi digunakan metrik accuracy, precision, recall, f1-score, dan confusion matrix. Hasil analisis memperlihatkan bahwa algoritma IndoBERT menghasilkan kinerja paling unggul dengan tingkat akurasi sebesar 92,46%, melampaui Naïve Bayes yang memperoleh 87,94% dan SVM sebesar 86,43%. Selain itu, IndoBERT memperlihatkan konsistensi yang lebih baik pada nilai precision, recall, dan f1-score di setiap kategori sentimen. Sebaliknya, Naïve Bayes masih menunjukkan kecenderungan kesalahan klasifikasi pada kelas tertentu, sedangkan SVM relatif lebih stabil meskipun belum mencapai performa optimal. Temuan ini mengindikasikan pendekatan berbasis transformer lebih efektif dalam menangkap kompleksitas konteks bahasa Indonesia. Kontribusi penelitian ini terletak pada penyajian analisis komparatif terhadap metode klasik dan berbasis transformer dalam konteks klasifikasi sentimen berbahasa Indonesia, serta memberikan bukti empiris mengenai keunggulan pendekatan transformer dalam menangkap kompleksitas konteks linguistik pada ulasan pengguna aplikasi digital.

Kata Kunci: Analisis Sentimen; Naïve Bayes; Support Vector Machine; IndoBERT; Text Mining

Abstract—This study aims to evaluate and compare the performance of Naïve Bayes, Support Vector Machine (SVM), and IndoBERT algorithms in the task of sentiment classification of user reviews on the Gojek application. Data were collected through web scraping from the Google Play Store and subsequently labeled into three sentiment categories: negative, neutral, and positive. A quantitative approach with a descriptive-comparative design was employed. The research procedure consisted of data collection, text preprocessing, dataset splitting into training and testing sets, model development, and evaluation using accuracy, precision, recall, F1-score, and confusion matrix metrics. The results indicate that the IndoBERT algorithm achieved the best performance, with an accuracy of 92.46%, outperforming Naïve Bayes (87.94%) and SVM (86.43%). Furthermore, IndoBERT demonstrated greater consistency in precision, recall, and F1-score across all sentiment categories. In contrast, Naïve Bayes exhibited a tendency to misclassify certain classes, while SVM showed relatively stable performance, although it did not reach optimal results. These findings suggest that transformer-based approaches are more effective in capturing the contextual complexity of the Indonesian language. This study contributes by providing a comparative analysis of classical and transformer-based methods in Indonesian sentiment classification and offers empirical evidence of the superiority of transformer-based approaches in capturing linguistic contextual nuances in user reviews of digital applications.

Keywords: Sentiment Analysis; Naïve Bayes; Support Vector Machine; IndoBERT; Text Mining

1. PENDAHULUAN

Teknologi informasi dan komunikasi yang sangat cepat berkembang dalam beberapa tahun terakhir telah mendorong terjadinya digitalisasi di berbagai bidang, termasuk dalam sektor transportasi dan layanan yang mengandalkan aplikasi. Di Indonesia, keberadaan platform digital seperti Gojek telah menjadi komponen yang vital dalam kehidupan masyarakat masa kini, karena mampu menawarkan layanan yang cepat, mudah, dan efisien. Peningkatan jumlah pengguna aplikasi ini disertai dengan tingginya frekuensi interaksi penggunaannya, khususnya dalam bentuk tanggapan di platform seperti Google Play Store. Tanggapan tersebut mencerminkan pengalaman nyata pengguna sehingga memiliki pentingnya strategis dalam mengevaluasi mutu layanan yang diberikan.

Ulasan pengguna termasuk dalam kategori *user-generated content* yang memuat opini, persepsi, dan emosi terhadap suatu layanan. Namun, karakteristik data ulasan yang tidak terstruktur, beragam secara linguistik, serta berjumlah besar menjadikannya sulit untuk dianalisis secara manual. Oleh sebab itu, diperlukan pendekatan komputasional yang mampu mengolah serta mengekstraksi informasi secara efektif dari data tersebut. Suatu pendekatan yang lazim digunakan adalah analisis sentimen, yang mempunyai fungsi guna mengklasifikasikan opini pengguna ke dalam beberapa kategori, dalam penelitian ini digunakan kategori positif, negatif, dan netral.

Data dalam penelitian ini dikumpulkan melalui teknik *web scraping* terhadap ulasan pengguna aplikasi Gojek pada Google Play Store. Teknik ini memungkinkan pengambilan data secara sistematis langsung dari sumber asli dalam jumlah besar, sehingga mencerminkan kondisi aktual pengalaman pengguna. Penggunaan data hasil scraping dari Google Play Store telah banyak dimanfaatkan dalam penelitian analisis sentimen karena dinilai mampu menyediakan data yang relevan dan mutakhir (Nugroho et al., 2021; Putra et al., 2025). Selain itu, metode ini juga memungkinkan pengumpulan informasi tambahan seperti rating dan waktu ulasan yang dapat memperkaya hasil analisis (Putra et al., 2025).



Dalam beberapa tahun terakhir, kajian mengenai analisis sentimen di Indonesia menunjukkan perkembangan yang signifikan. Setiawan (2025) mengemukakan bahwa penerapan analisis sentimen semakin meluas pada berbagai domain, seperti ulasan aplikasi, media sosial, dan opini publik. Hal ini menegaskan peran penting analisis sentimen dalam mendukung pengambilan keputusan berbasis data, khususnya dalam memahami tingkat kepuasan dan kebutuhan pengguna. Meskipun demikian, analisis teks berbahasa Indonesia menghadapi sejumlah tantangan, antara lain penggunaan bahasa yang tidak baku, adanya singkatan dan istilah slang, serta struktur kalimat yang tidak teratur dan ambigu. Penelitian Irmawan (2024) menunjukkan bahwa karakteristik kompleks pada teks ulasan aplikasi di Indonesia menuntut penggunaan model yang mampu memahami konteks secara mendalam.

Beragam cara telah digunakan dalam penelitian sentimen, mulai dari teknik pembelajaran mesin konvensional hingga metode yang didasarkan pada pembelajaran mendalam. Algoritma seperti Support Vector Machine (SVM) serta Naïve Bayes masih sering dipilih dikarenakan keunggulannya dalam hal kesederhanaan, efisiensi pengolahan data, dan hasil yang cukup memuaskan pada dataset berukuran kecil hingga sedang. Raschka et al. (2022) menyebutkan bahwa algoritma tradisional masih memiliki relevansi dalam tugas-tugas klasifikasi, terutama saat digabungkan dengan metode ekstraksi fitur seperti TF-IDF. Meskipun demikian, metode ini memiliki batasan dalam memahami makna semantik dan konteks bahasa, sehingga kinerjanya cenderung menurun ketika dihadapkan pada data yang lebih kompleks.

Seiring dengan kemajuan teknologi, model berbasis transformer seperti BERT dan turunannya mulai banyak diadopsi dalam analisis sentimen. BERT yang Devlin et al. (2019) perkenalkan menjadi tonggak penting dalam pemrosesan bahasa alami karena kemampuannya memahami konteks secara dua arah (*bidirectional*) melalui proses pre-training. Model ini kemudian dikembangkan ke berbagai bahasa, termasuk bahasa Indonesia melalui IndoBERT yang diperkenalkan oleh Koto et al. (2021), yang dilatih menggunakan korpus bahasa Indonesia sehingga lebih adaptif terhadap karakteristik linguistik lokal.

Sejumlah penelitian menunjukkan bahwa pendekatan berbasis *deep learning* memiliki keunggulan dibandingkan metode tradisional dalam klasifikasi teks. Abdullah dan Ahmet (2023) serta Sahoo et al. (2023) mengungkapkan bahwa model *deep learning* mampu menangani kompleksitas data teks tidak terstruktur secara lebih efektif dibandingkan pendekatan berbasis frekuensi kata. Selain itu, Wankhade et al. (2022) menunjukkan bahwa model berbasis transformer secara konsisten memberikan performa yang lebih baik dibandingkan algoritma klasik seperti Naïve Bayes dan SVM dalam berbagai tugas klasifikasi teks, termasuk analisis sentimen.

Dalam konteks bahasa Indonesia, penggunaan IndoBERT juga menunjukkan hasil yang unggul dalam berbagai tugas. Geni et al (2023) menemukan bahwa IndoBERT mampu meningkatkan akurasi klasifikasi sentimen dibandingkan metode konvensional. Temuan ini diperkuat oleh Setiawan dan Iswavigra (2025) yang menyatakan bahwa IndoBERT efektif dalam mengklasifikasikan sentimen teks berbahasa Indonesia. Dilain sisi, penelitian oleh Wijaya dan Sasmita (2024) memperlihatkan bahwa model berbasis BERT mampu melampaui kinerja metode tradisional dalam analisis sentimen pada ulasan aplikasi. Salsabila dan Priyatna (2025) juga menegaskan bahwa kombinasi Naïve Bayes, SVM, dan BERT sering digunakan untuk memberikan perbandingan performa yang komprehensif dalam analisis sentimen berdasarkan data ulasan aplikasi di Google Play Store.

Meskipun demikian, penelitian yang secara langsung membandingkan metode *machine learning* tradisional seperti Naïve Bayes dan SVM dengan model berbasis *transformer* dalam satu kerangka eksperimen yang sama masih terbatas, khususnya pada data ulasan aplikasi Gojek. Selain itu, pemanfaatan data hasil *scraping* dari Google Play Store sebagai sumber utama penelitian dalam konteks layanan transportasi digital di Indonesia juga belum banyak dikaji secara spesifik.

Berdasarkan uraian tersebut, terdapat kesenjangan penelitian dalam hal evaluasi komparatif antara metode tradisional dan pendekatan berbasis *deep learning* dalam analisis sentimen teks berbahasa Indonesia yang bersumber dari ulasan aplikasi. Dengan demikian, penelitian ini mempunyai tujuan guna melakukan perbandingan efektivitas dari tiga algoritma klasifikasi, yaitu SVM, Naïve Bayes, dan IndoBERT, untuk mengategorikan sentimen ulasan pengguna Gojek ke dalam kelas negatif, netral, dan positif. Evaluasi kinerja dilaksanakan dengan perhitungan metrik *accuracy*, *precision*, *f1-score*, dan *recall*, serta *confusion matrix*. Kontribusi penelitian ini terletak pada penyajian analisis komparatif terhadap metode klasik dan berbasis *transformer* dalam konteks klasifikasi sentimen berbahasa Indonesia, serta memberikan bukti empiris mengenai keunggulan pendekatan *transformer* dalam menangkap kompleksitas konteks linguistik pada ulasan pengguna aplikasi digital.

2. METODOLOGI PENELITIAN

2.1 Kerangka Dasar Penelitian

Penelitian ini menerapkan pendekatan kuantitatif dengan desain deskriptif-komparatif yang bertujuan untuk mengevaluasi serta membandingkan kinerja berbagai algoritma klasifikasi analisis sentimen pada ulasan pengguna aplikasi Gojek. Pendekatan kuantitatif dipilih karena penelitian berorientasi pada pengukuran performa model secara numerik melalui indikator seperti *accuracy*, *precision*, *recall*, dan *f1-score*. Sementara itu, desain komparatif digunakan untuk mengidentifikasi perbedaan kinerja antara metode *machine learning* konvensional dan pendekatan berbasis *deep learning* dalam pengolahan teks berbahasa Indonesia. Penggunaan metrik evaluasi tersebut telah menjadi standar dalam sistem klasifikasi karena mampu memberikan gambaran yang komprehensif terhadap performa model, baik secara keseluruhan maupun pada masing-masing kelas (Powers, 2020).



Data yang dimanfaatkan pada penelitian ini berupa data sekunder yang dikumpulkan melalui teknik *web scraping* terhadap ulasan pengguna aplikasi Gojek pada platform layanan Google yaitu Play Store. Dengan teknik *web scrapping* ini, dimungkinkan proses pengambilan data secara otomatis, terstruktur, serta langsung dari sumber asli, sehingga mampu merepresentasikan pengalaman pengguna secara aktual. Penggunaan *web scraping* dalam pengumpulan data ulasan aplikasi telah banyak dimanfaatkan dalam penelitian analisis sentimen karena kemampuannya menghasilkan data berjumlah besar yang relevan dan mutakhir (Nugroho et al., 2021; Putra et al., 2025). Data yang diperoleh selanjutnya melalui tahap pelabelan sentimen ke dalam 3 kategori, yaitu negatif, netral, dan positif, sebelum digunakan dalam proses klasifikasi. Berikutnya, data dibagi kedalam data latih dan data uji dengan rasio 80:20 guna menjamin objektivitas dalam proses pelatihan dan pengujian model.

Penelitian ini menggunakan beberapa variabel yaitu variabel independen dan dependen. Variabel independen mencakup metode klasifikasi yang digunakan, yaitu, SVM, Naïve Bayes, serta IndoBERT. Adapun variabel dependen adalah hasil klasifikasi sentimen yang diukur menggunakan metrik *accuracy*, *precision*, *f1-score*, serta *recall*. Selain itu, *confusion matrix* dimanfaatkan untuk mengevaluasi distribusi hasil prediksi kepada data aktual, sehingga pola kesalahan klasifikasi dapat dianalisis secara lebih mendalam. Alahmadi et al. (2025) menyatakan bahwa kombinasi antara metrik evaluasi kuantitatif dan analisis kesalahan melalui *confusion matrix* merupakan pendekatan yang penting dalam mengevaluasi performa model klasifikasi teks secara menyeluruh.

Hipotesis penelitian ini menyatakan bahwa terdapat perbedaan kinerja di antara metode yang digunakan, di mana model berbasis *deep learning* seperti IndoBERT diasumsikan memiliki performa yang lebih unggul dibandingkan metode tradisional. Hal ini berlandaskan pada hasil kajian yang telah dilakukan sebelumnya yang memperlihatkan bahwa model berbasis BERT mampu memahami konteks bahasa secara lebih mendalam dibandingkan pendekatan berbasis frekuensi kata (V. D. Setiawan et al., 2025). Selain itu, Ramdhan dan Purbolaksono (2024) menunjukkan bahwa IndoBERT menghasilkan tingkat akurasi yang lebih tinggi dalam klasifikasi sentimen bahasa Indonesia dibandingkan Naïve Bayes. Penelitian lain juga mengungkapkan bahwa model berbasis *transformer* mampu meningkatkan kinerja klasifikasi pada data ulasan aplikasi karena kemampuannya dalam menangkap relasi semantik antar kata secara kontekstual (Wijaya et al., 2024). Hal ini diperkuat oleh Nandwani dan Verma (2021) yang menjelaskan bahwa proses *fine-tuning* pada model BERT memungkinkan model beradaptasi secara optimal terhadap tugas klasifikasi teks tertentu.

Kerangka konseptual dalam penelitian ini didasarkan pada perbedaan mendasar antara pendekatan tradisional dan modern dalam pengolahan teks. Metode Naïve Bayes dan SVM memanfaatkan representasi fitur berbasis frekuensi kata seperti TF-IDF yang mengonversi teks ke dalam bentuk numerik sederhana. Sebaliknya, IndoBERT menggunakan arsitektur *transformer* yang mampu memahami keterkaitan konteks tiap kata dalam kalimat secara dua arah (*bidirectional*). Raschka et al. (2022) menyatakan bahwa algoritma *machine learning* klasik masih relevan dalam tugas klasifikasi teks, namun mempunyai batasan dalam menangkap konteks semantik yang kompleks. Di sisi lain, Kowsari et al. (2019) menegaskan bahwa pendekatan berbasis *deep learning* memiliki keunggulan dalam menangani kompleksitas data teks karena kemampuannya dalam mempelajari representasi fitur secara otomatis. Oleh karena itu, penelitian ini secara empiris menguji perbandingan kinerja ketiga metode tersebut pada data ulasan aplikasi Gojek yang didapat dari platform Google yaitu Play Store, yang memiliki karakteristik bahasa tidak baku dan tingkat kompleksitas yang tinggi.

2.2 Tahapan Penelitian

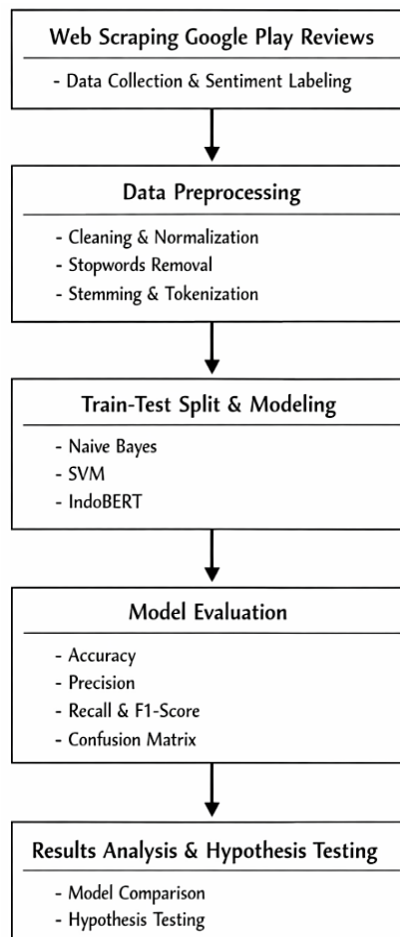
Penelitian ini dilakukan dengan beberapa langkah atau tahapan yang terorganisir untuk memastikan bahwa analisis dilakukan secara teratur dan hasil yang didapat memiliki kemampuan replikasi yang tinggi. Proses penelitian dimulai dengan pengumpulan data menggunakan metode *web scraping*, kemudian dilanjutkan dengan fase pra-pemrosesan, pemodelan, hingga evaluasi hasil dan pengujian hipotesis.

Tahap awal berupa pengumpulan data, tahapan ini melalui penerapan teknik *web scraping* terhadap ulasan pemakai aplikasi Gojek pada platform Play Store. Pendekatan ini memungkinkan akuisisi data secara otomatis dalam skala besar langsung dari sumber asli, sehingga data yang diperoleh merepresentasikan pengalaman pengguna secara aktual. Proses *scraping* dilakukan menggunakan pustaka Python seperti *google-play-scraper* atau metode sejenis yang mendukung pengambilan data ulasan beserta atribut tambahan, seperti rating dan waktu unggahan. Pemanfaatan teknik ini telah jamak diterapkan dalam penelitian analisis sentimen dikarenakan mampu menyediakan data yang relevan, dinamis, dan kontekstual (Nugroho et al., 2021; Putra et al., 2025). Setelah proses pengumpulan selesai, dilakukan pelabelan sentimen ke dalam 3 kategori, yaitu negatif, netral, dan positif, baik secara manual maupun semi-otomatis untuk menjaga kualitas data sebelum tahap klasifikasi.

Tahap berikutnya adalah prapemrosesan teks yang mempunyai tujuan guna memperbaiki tingkat kualitas data sebelum digunakan dalam pemodelan. Proses ini mencakup pembersihan teks (*cleaning*) dari elemen yang tidak sesuai seperti URL, angka, serta simbol, konversi huruf kedalam bentuk huruf kecil (*case folding*), penghapusan *stopwords*, tokenisasi, serta proses *stemming* menggunakan pustaka Sastrawi. Selain itu, dilakukan normalisasi kata untuk mengatasi variasi bahasa tidak baku, termasuk singkatan dan istilah slang yang umum ditemukan dalam ulasan pengguna. Tahap ini memiliki peran krusial karena kualitas data yang baik akan berpengaruh langsung terhadap kinerja model klasifikasi. Purnomo dan Sutopo (2024) menunjukkan bahwa pra-pemrosesan yang optimal dapat meningkatkan akurasi model secara signifikan dalam analisis sentimen.

Selanjutnya, data dibagi menjadi data latih dan data uji menggunakan metode *train-test split* dengan rasio 80:20. Data latih digunakan dalam proses pembangunan model, sedangkan data uji dimanfaatkan untuk mengevaluasi performa model terhadap data yang belum pernah dipelajari sebelumnya. Proses pelatihan dilakukan menggunakan tiga algoritma, yaitu Naïve Bayes, Support Vector Machine (SVM), dan IndoBERT. Pada metode Naïve Bayes dan SVM, digunakan teknik ekstraksi fitur TF-IDF untuk mengonversi teks ke dalam representasi numerik. Sementara itu, IndoBERT memanfaatkan tokenisasi berbasis *subword* yang memungkinkan pemahaman konteks bahasa secara lebih mendalam. Implementasi model dilakukan menggunakan bahasa pemrograman Python dengan dukungan pustaka seperti Scikit-learn untuk metode *machine learning* dan *transformers* untuk IndoBERT.

Langkah berikutnya melibatkan penilaian model melalui metrik *accuracy*, *precision*, *recall*, serta *f1-score*. Akurasi berfungsi guna mengevaluasi sejauh mana keakuratan model secara keseluruhan, sedangkan presisi dan recall bertujuan untuk mengukur kinerja model di setiap kategori sentimen. F1-score dipergunakan sebagai metrik yang mengharmonisasi antara presisi dan recall. Di samping itu, *confusion matrix* dipakai untuk menganalisa distribusi hasil klasifikasi, sehingga dapat teridentifikasi jumlah prediksi yang tepat serta kesalahan di setiap kategori. Pemakaian metrik ini telah menjadi norma dalam penelitian analisis sentimen karena dapat memberikan gambaran tentang performa model secara menyeluruh (Irmawan et al., 2024). Langkah terakhir adalah menganalisis hasil dan menguji hipotesis. Analisis dilakukan dengan membandingkan nilai metrik evaluasi dari masing-masing model untuk menentukan metode dengan kinerja terbaik dalam mengklasifikasikan sentimen ulasan pengguna Gojek. Pengujian hipotesis dilaksanakan dengan pendekatan deskriptif-komparatif merujuk pada perbandingan nilai metrik antar model, menggunakan tingkat signifikansi 0,05 sebagai pondasi pengambilan keputusan. Keseluruhan proses analisis dilaksanakan melalui perangkat lunak Python dalam lingkungan Jupyter Notebook atau Google Colab untuk memastikan transparansi, efisiensi, dan reproduktibilitas penelitian.



Gambar 1. Tahapan Penelitian

3. HASIL DAN PEMBAHASAN

Menganalisis perbandingan dari kinerja tiga metode klasifikasi sentimen, yaitu Naïve Bayes, Support Vector Machine (SVM), dan IndoBERT untuk mengkategorikan teks berbahasa Indonesia yang diperoleh dari platform Play Store merupakan tujuan penelitian ini. Penilaian dilakukan dengan menggunakan beberapa metrik performa yang lazim dipakai dalam evaluasi klasifikasi, yakni akurasi, presisi, f1-score, recall, serta *confusion matrix*. Total dataset yang

dimanfaatkan dalam penelitian ini berjumlah 2000 data. Sementara itu, jumlah data uji yang digunakan dalam studi ini adalah 20% atau sebanyak 400 data, yang dikategorikan ke dalam 3 kelas sentimen, yaitu negatif yang disimbolkan dengan angka 0, netral yang disimbolkan dengan angka 1, dan positif yang disimbolkan dengan angka 2. Adapun data latih yang diambil dalam penelitian ini adalah 80% yang setara dengan 1600 data. Distribusi data uji: Negatif 43,55% (871), Positif 51,9% (1038), Netral 4,55% (91).

3.1 Hasil

Dataset yang digunakan dalam penelitian ini merepresentasikan karakteristik bahasa pengguna aplikasi di Indonesia yang umumnya tidak baku, banyak mengandung singkatan, serta memanfaatkan ragam bahasa informal. Sebelum memasuki tahap pemodelan, terlebih dahulu data diproses melalui pra-pemrosesan teks yang meliputi *cleaning*, *case folding*, tokenisasi, penghapusan *stopword*, serta *stemming* dengan bantuan pustaka Sastrawi. Tahapan ini dilakukan agar kualitas data dapat meningkat sekaligus meminimalkan *noise* yang berpotensi memengaruhi kinerja model dalam proses klasifikasi.

Dalam penelitian ini, data dibagi ke data latih dan data uji berdasarkan proporsi tertentu, di mana data uji berjumlah 400 entri yang telah dilabeli ke dalam tiga kategori sentimen. Distribusi data mencakup tiga kelas utama dengan komposisi yang relatif seimbang, sehingga memungkinkan proses evaluasi model dilakukan secara lebih objektif. Dataset tersebut mencerminkan beragam opini pengguna terhadap layanan Gojek, mulai dari keluhan (negatif), pernyataan informatif (netral), hingga bentuk apresiasi terhadap layanan (positif).

Studi ini juga melakukan perbandingan kinerja tiga algoritma klasifikasi berdasarkan perbedaan pendekatan dalam representasi fitur. Pada metode Naïve Bayes dan SVM, digunakan teknik TF-IDF untuk mengonversi teks menjadi representasi numerik. Sementara itu, IndoBERT mengadopsi pendekatan tokenisasi berbasis subword yang memungkinkan pemahaman konteks bahasa secara lebih komprehensif. Perbedaan pendekatan tersebut menjadi landasan dalam menilai kemampuan masing-masing metode dalam mengelola kompleksitas teks berbahasa Indonesia.

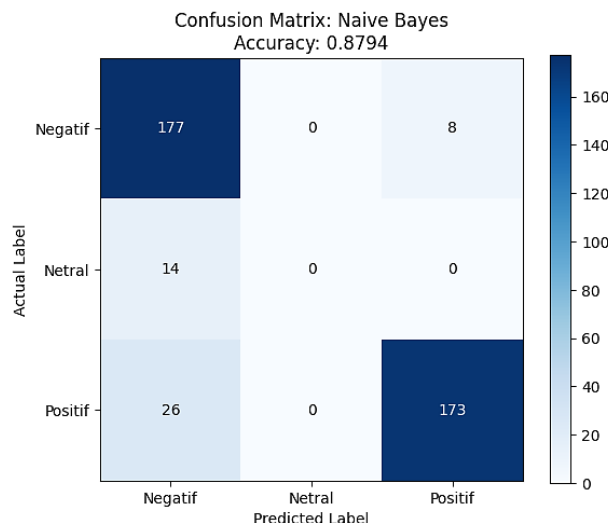
Penilaian kinerja model dilakukan dengan memanfaatkan *confusion matrix* serta metrik akurasi, presisi, recall, dan f1-score. *Confusion matrix* digunakan untuk memberikan gambaran detail terkait distribusi hasil klasifikasi, mencakup jumlah prediksi yang tepat serta kesalahan di setiap kategori. Evaluasi bukan sekadar terfokus pada akurasi keseluruhan, melainkan juga pada keterampilan model dalam mengenali setiap kelas sentimen dengan akurat. Perhatian lebih diberikan pada kesalahan *false positive* dan *false negative*, karena dapat memengaruhi pemahaman terhadap persepsi pengguna. Oleh karena itu, penilaian dilakukan secara menyeluruh untuk menemukan kelebihan dan kelemahan setiap metode dalam konteks pengolahan sentimen terhadap ulasan aplikasi Gojek.

3.1.1 Naïve Bayes

Merujuk hasil dari pengujian pada data uji, model ini mendapat tingkat akurasi sebesar 87,94%. Nilai tersebut menunjukkan kemampuan model dalam mengklasifikasikan sebagian data secara tepat, meskipun masih terdapat sejumlah kesalahan klasifikasi. Secara lebih rinci, hasil evaluasi kinerja model Naïve Bayes diperlihatkan pada Tabel 1.

Tabel 1. Hasil Evaluasi Model Naïve Bayes

Kelas	Precision	Recall	F1-Score	Support
0 (Negatif)	0.82	0.96	0.88	185
1 (Netral)	0.00	0.00	0.00	14
2 (Positif)	0.96	0.87	0.91	199



Gambar 2. Confusion Matrix Model Naïve Bayes

Berdasarkan analisis yang dilakukan melalui *confusion matrix* seperti yang terlihat pada Gambar 2, model Naïve Bayes mencapai akurasi sebesar 0,8794 atau 87,94%. Nilai ini menunjukkan bahwa model memiliki kemampuan yang cukup mumpuni untuk mengkategorikan sentimen dari penilaian pengguna Gojek. Meski demikian, analisis tambahan terhadap distribusi prediksi memperlihatkan adanya ketidakseimbangan performa di antara setiap kelas sentimen.

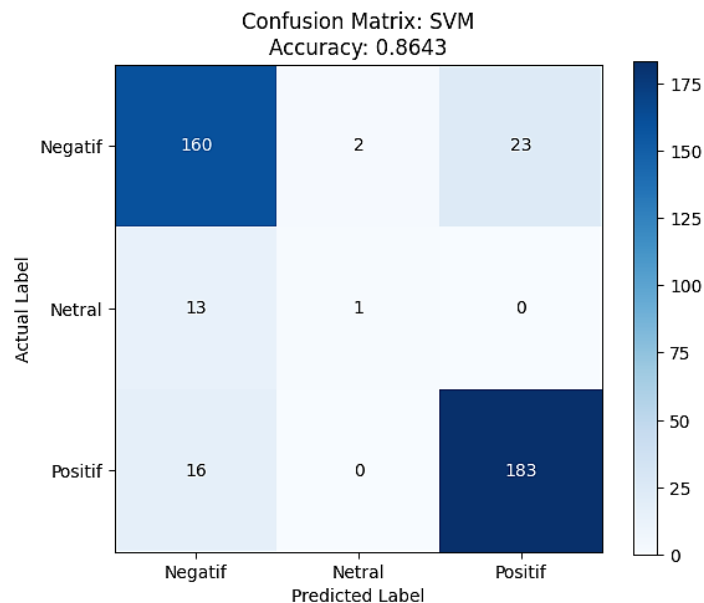
Pada kelas negatif, model mengindikasikan hasil yang sangat optimal. Dari total 185 data aktual negatif, sebanyak 177 data berhasil diidentifikasi dengan tepat sebagai negatif, sementara 8 data lainnya keliru diprediksi sebagai positif. Tidak ada kesalahan klasifikasi yang mengarah pada kelas netral. Hal tersebut memperlihatkan bahwa model mempunyai tingkat recall yang sangat tinggi dalam menangkap sentimen negatif. Sementara itu, kinerja model pada kelas netral tergolong sangat rendah. Dari 14 data aktual netral, tidak ada satupun data yang berhasil diklasifikasikan dengan benar, seluruhnya diprediksi sebagai negatif. Keadaan ini memperlihatkan bahwa pada model Naïve Bayes terjadi kesulitan dalam mengenali karakteristik sentimen netral, yang mungkin disebabkan oleh kesamaan fitur antara teks netral dan negatif atau kurangnya representativitas distribusi fitur dalam kelas tersebut. Pada kelas positif, model menunjukkan performa yang relatif baik. Dari 199 data aktual positif, sebanyak 173 data dapat diidentifikasi dengan benar sebagai positif, sedangkan 26 data lainnya salah diklasifikasikan sebagai negatif. Tidak ditemukan prediksi yang mengarah ke kelas netral. Ini menandakan bahwa model cukup mampu mengenali sentimen positif, meskipun masih terdapat kecenderungan kesalahan ke kelas negatif. Secara keseluruhan, hasil *confusion matrix* memperlihatkan bahwa model Naïve Bayes memiliki kecenderungan kuat dalam mengklasifikasikan data ke kelas negatif dan positif, tetapi kurang efektif dalam mengidentifikasi kelas netral dengan akurat. Kondisi ini bisa jadi disebabkan oleh keterbatasan model dalam menangkap konteks semantik serta kemungkinan ketidakseimbangan distribusi data atau kurangnya fitur diskriminatif pada kelas netral. Oleh karena itu, meski nilai akurasi yang dicapai relatif tinggi, performa model tidak merata di seluruh kelas. Dengan demikian, evaluasi kinerja tidak bisa hanya mengandalkan akurasi, tetapi juga perlu mempertimbangkan metrik lain seperti presisi, f1-score, serta recall untuk mendapatkan gambaran yang lebih menyeluruh.

3.1.2 Support Vector Machine (SVM)

Dari hasil uji coba yang telah dilaksanakan, model SVM memperlihatkan peningkatan performa dibandingkan Naïve Bayes, dengan persentase akurasi mencapai 86,43%. Detail hasil evaluasi untuk model SVM ditunjukkan dalam tabel berikut.

Tabel 2. Hasil Evaluasi Model SVM

Kelas	Precision	Recall	F1-Score	Support
0 (Negatif)	0.85	0.86	0.86	185
1 (Netral)	0.33	0.07	0.12	14
2 (Positif)	0.89	0.92	0.90	199



Gambar 3. Confusion Matrix Model SVM

Hasil evaluasi yang diperoleh dari *confusion matrix*, seperti yang ditunjukkan pada Gambar 3, model SVM berhasil mencapai tingkat akurasi sebesar 0,8643 atau 86,43%. Angka ini mencerminkan bahwa model berkinerja cukup baik dalam mengklasifikasikan sentimen dari ulasan pengguna aplikasi Gojek. Namun, analisis lebih mendalam terhadap sebaran hasil klasifikasi menunjukkan adanya perbedaan kinerja di setiap kelas sentimen. Pada kelas negatif, performa model SVM terlihat cukup tinggi. Dari 185 data aktual negatif, 160 data berhasil diklasifikasikan dengan tepat

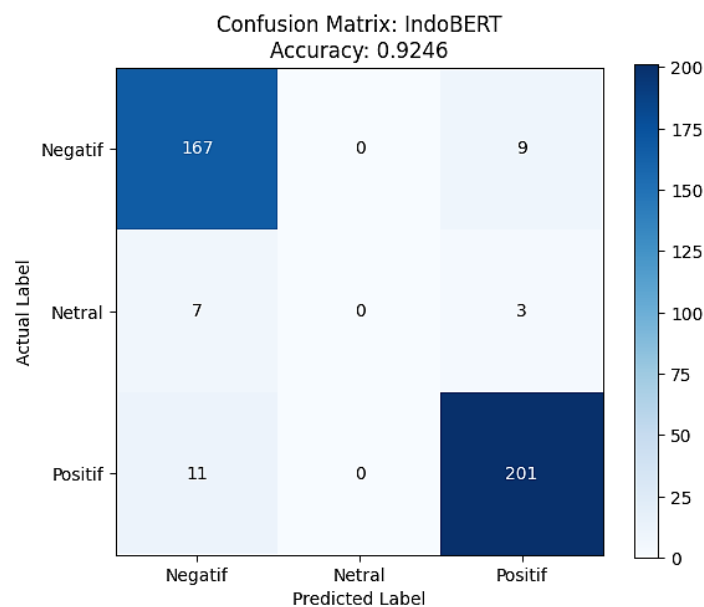
sebagai negatif, sementara 23 data salah terklasifikasi sebagai positif dan 2 data sebagai netral. Hal ini menandakan bahwa model cukup efektif untuk mendeteksi sentimen negatif, meskipun masih terdapat kecenderungan kesalahan yang mengarah ke kelas positif, menunjukkan adanya kesamaan dalam karakteristik fitur antara kedua kelas tersebut. Sebaliknya, pada kelas netral, kinerja model tergolong sangat rendah. Dari total 14 data aktual netral, hanya 1 data yang berhasil diidentifikasi dengan benar, sedangkan 13 data lainnya keliru terklasifikasi sebagai negatif. Tidak ada kesalahan prediksi ke dalam kelas positif. Penemuan ini memperlihatkan bahwa model SVM menghadapi tantangan saat membedakan sentimen netral dari negatif, yang mungkin disebabkan oleh keterbatasan representasi fitur berbasis TF-IDF dalam menangkap nuansa yang ambigu pada teks netral. Pada kelas positif, performa model cukup baik. Dari 199 data aktual positif, sebanyak 183 data berhasil diidentifikasi dengan akurat sebagai positif, sementara 16 data lainnya salah terklasifikasi sebagai negatif. Tidak terdapat kesalahan klasifikasi ke dalam kelas netral. Ini menunjukkan bahwa model mampu mengidentifikasi sentimen positif dengan baik, walaupun masih terdapat sedikit kesalahan yang cenderung menuju kelas negatif. Hasil *confusion matrix* secara keseluruhan memperlihatkan bahwa model SVM memiliki kemampuan baik untuk mengkategorikan sentimen negatif dan positif, tetapi memiliki keterbatasan signifikan dalam mengenali sentimen netral. Model ini cenderung mengelompokkan data ke dalam dua kelas utama, yaitu negatif dan positif, sehingga representasi kelas netral menjadi kurang maksimal. Situasi ini mungkin dipengaruhi oleh kurangnya fitur diskriminatif serta keterbatasan metode dalam memahami konteks semantik teks yang kompleks. Selain itu, meskipun nilai akurasi yang diperoleh cukup tinggi, evaluasi kinerja model perlu mempertimbangkan metrik lain seperti precision, recall, dan f1-score dalam setiap kelas untuk mendapatkan gambaran yang lebih menyeluruh dalam analisis sentimen.

3.1.2 IndoBERT

Berdasarkan hasil yang diperoleh dari pengujian, model ini menunjukkan performa terbaik dengan tingkat ketepatan mencapai 92,46%, yang secara signifikan melebihi kinerja dua model sebelumnya. Detail dari evaluasi model IndoBERT diperlihatkan dalam Tabel 3 berikut.

Tabel 3. Hasil Evaluasi Model IndoBERT

Kelas	Precision	Recall	F1-Score	Support
0 (Negatif)	0.90	0.95	0.93	176
1 (Netral)	0.00	0.00	0.00	10
2 (Positif)	0.94	0.95	0.95	212



Gambar 4. Confusion Matrix Model IndoBERT

Hasil analisis menggunakan *confusion matrix* yang diperlihatkan pada gambar diatas, model IndoBERT berhasil mencapai tingkat akurasi sebesar 0,9246 atau 92,46%, mengindikasikan bahwa model ini berfungsi dengan sangat baik dalam mengklasifikasikan sentimen ulasan dari pengguna aplikasi Gojek. Angka akurasi tersebut merupakan yang tertinggi dibandingkan dengan model lainnya, sehingga menegaskan keunggulan pendekatan berbasis *transformer* dalam memahami konteks bahasa dengan lebih mendalam. Pada kelas negatif, IndoBERT menunjukkan kinerja yang sangat optimal. Dari total 176 data nyata negatif, sebanyak 167 data berhasil diprediksi dengan akurat sebagai negatif, sementara 9 data lainnya terklasifikasi keliru sebagai positif. Tidak ada kesalahan klasifikasi ke dalam kelas netral. Hal ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mendeteksi pola sentimen negatif, dengan tingkat *recall* yang tinggi untuk kelas tersebut. Pada kelas netral, performa model masih menunjukkan kelemahan yang

cukup signifikan. Dari 10 data nyata netral, tidak ada satu pun data yang berhasil diklasifikasikan secara tepat, di mana 7 data diprediksi sebagai negatif dan 3 data sebagai positif. Temuan ini menunjukkan bahwa walaupun IndoBERT tampil unggul secara keseluruhan, model masih kesulitan untuk mengidentifikasi sentimen netral yang umumnya bersifat ambigu dan berada di antara sentimen negatif dan positif. Pada kelas positif, IndoBERT memperlihatkan performa yang sangat tinggi. Dari total 212 data nyata positif, sebanyak 201 data berhasil diklasifikasikan dengan tepat sebagai positif, dan 11 data lainnya salah diprediksi sebagai negatif. Tidak ada kesalahan klasifikasi ke kelas netral. Ini menandakan bahwa model sangat efektif dalam mengenali sentimen positif dengan tingkat akurasi yang tinggi. Secara keseluruhan, hasil *confusion matrix* memperlihatkan bahwa IndoBERT memiliki kemampuan yang sangat baik untuk mengklasifikasikan sentimen negatif serta positif, melalui dominasi prediksi yang benar pada kedua kelas tersebut. Tetapi model masih menghadapi tantangan untuk mengidentifikasi kelas netral yang cenderung sulit terdeteksi dengan akurat. Kondisi ini kemungkinan dipengaruhi oleh distribusi data yang tidak seimbang atau karakteristik linguistik dari teks netral yang tidak memiliki ciri pembeda yang kuat dibandingkan dengan teks positif dan negatif. Meskipun demikian, secara umum IndoBERT tetap menjadi model dengan performa terbaik dalam penelitian ini karena bisa menghasilkan klasifikasi yang lebih akurat dan konsisten dibandingkan metode lainnya.

3.2 Analisis Komparatif Model Naïve Bayes, SVM, dan IndoBERT

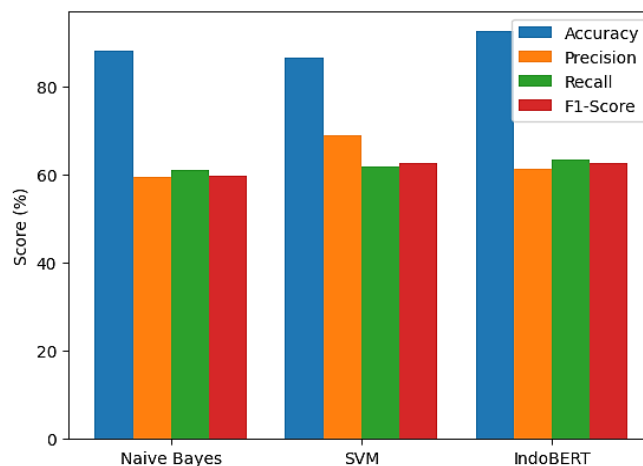
Merujuk dari hasil evaluasi yang telah dilaksanakan, ketiga model klasifikasi, yakni Naïve Bayes, SVM, dan IndoBERT, memperlihatkan variasi kinerja yang cukup signifikan untuk mengklasifikasikan sentimen ulasan dari pengguna aplikasi Gojek. Perbandingan ini dilakukan dengan merujuk pada metrik evaluasi diantaranya yaitu akurasi, presisi, f1-score, recall serta analisis *confusion matrix* guna mengidentifikasi pola kesalahan klasifikasi di setiap model.

Secara umum, IndoBERT menghasilkan performa terbaik dengan tingkat akurasi mencapai 92,46%, diikuti oleh Naïve Bayes pada 87,94% dan SVM pada 86,43%. Penemuan ini menunjukkan bahwa model berbasis *deep learning* dengan arsitektur *transformer* memiliki keunggulan dalam memahami konteks bahasa dibandingkan metode *machine learning* tradisional. Ini sejalan dengan konsep bahwa model berbasis BERT mampu menangkap hubungan semantik antar kata dalam konteks. Dari aspek distribusi hasil klasifikasi, Naïve Bayes menunjukkan kecenderungan kuat dalam mengelompokkan data ke dalam kelas dominan, yaitu negatif dan positif. Walaupun memiliki tingkat akurasi yang relatif tinggi, model ini tidak dapat mengidentifikasi kelas netral dengan tepat, yang menunjukkan keterbatasannya dalam menangkap nuansa bahasa yang ambigu.

Model SVM mempertunjukkan kinerja yang lebih stabil dibandingkan Naïve Bayes, khususnya pada kelas negatif dan positif. Namun demikian, model ini juga mengalami tantangan serupa dalam mengklasifikasikan kelas netral, di mana sebagian besar data netral cenderung diprediksi sebagai negatif. Hal ini menunjukkan bahwa representasi fitur berbasis TF-IDF belum cukup mampu membedakan antara sentimen netral dan negatif secara jelas. Meskipun demikian, SVM memiliki keunggulan dalam menjaga keseimbangan performa antar kelas dibandingkan Naïve Bayes.

Di sisi lain, IndoBERT menunjukkan keunggulan yang lebih signifikan untuk pengklasifikasian sentimen negatif dan positif, dengan jumlah prediksi yang benar dominan pada kedua kelas tersebut. Ini menunjukkan bahwa IndoBERT mampu memahami konteks kalimat secara lebih mendalam, termasuk hubungan antar kata dalam teks. Meskipun demikian, model ini masih menunjukkan keterbatasan dalam mengidentifikasi kelas netral, di mana tidak ada prediksi yang tepat untuk kategori tersebut. Walaupun demikian, tingkat kesalahan klasifikasi di IndoBERT relatif lebih rendah dibandingkan dengan dua model lainnya.

Secara keseluruhan, ketiga model menunjukkan pola kesalahan yang serupa, yaitu kesulitan dalam membedakan sentimen netral dari sentimen negatif. Ini mengindikasikan bahwa tantangan utama dalam analisis sentimen pada dataset ini bukan hanya terletak pada pemilihan model, tetapi juga pada karakteristik data itu sendiri, terutama pada kelas netral yang cenderung ambigu dan kurang memiliki indikator linguistik yang kuat.



Gambar 5. Grafik Analisis Komparatif Model



4. KESIMPULAN

Penelitian ini menilai serta melakukan perbandingan terhadap performa dari tiga metode klasifikasi, yakni Naïve Bayes, SVM, dan IndoBERT, dalam mengkategorikan sentimen dari ulasan pengguna aplikasi Gojek yang diambil melalui proses pengambilan data dari Google Play Store ke dalam 3 kelompok, yaitu negatif, netral, dan positif. Hasil pengujian memperlihatkan bahwa IndoBERT menawarkan performa terbaik dengan tingkat akurasi mencapai 92,46%, diikuti oleh Naïve Bayes di 87,94% dan SVM di 86,43%. Penemuan ini menguatkan hipotesis bahwa model berbasis pembelajaran mendalam (*deep learning*), terutama IndoBERT, menunjukkan performa yang lebih superior dibandingkan metode pembelajaran mesin tradisional dalam analisis sentimen teks berbahasa Indonesia. Selain itu, analisis menunjukkan bahwa Naïve Bayes dan SVM terbatas dalam memahami konteks semantik, terutama dalam membedakan sentimen netral yang cenderung ambigu. Walaupun IndoBERT menunjukkan performa yang lebih baik secara keseluruhan, model ini tetap belum optimal dalam klasifikasi kelas netral. Ini mengindikasikan bahwa tantangan dalam analisis sentimen tidak hanya bersinggungan dengan pemilihan model, tetapi juga dipengaruhi oleh karakteristik data, terutama distribusi kelas dan kompleksitas bahasa dalam ulasan pengguna. Secara praktis, studi ini merekomendasikan penerapan IndoBERT sebagai metode yang efisien untuk menganalisis sentimen dari ulasan pengguna pada aplikasi layanan digital seperti Gojek. Hasil analisis sentimen dapat dimanfaatkan oleh pengembang aplikasi dan stakeholder untuk memahami persepsi pengguna secara lebih tepat, sehingga dapat digunakan sebagai dasar untuk meningkatkan kualitas layanan, mengembangkan fitur, dan menyusun strategi yang lebih terarah. Di samping itu, pendekatan yang berbasis pembelajaran mesin (*machine learning*) dan pembelajaran mendalam (*deep learning*) dalam analisis sentimen juga memiliki potensi untuk diterapkan di berbagai bidang lain yang melibatkan data teks yang tidak terstruktur. Disamping itu, studi ini memiliki beberapa batasan. Pertama, distribusi data yang masih belum seimbang, khususnya pada kelas netral, memengaruhi kemampuan model dalam melakukan klasifikasi secara maksimal. Kedua, proses pelabelan sentimen berisiko mengandung subjektivitas, terutama dalam membedakan antara sentimen netral dan negatif. Ketiga, studi ini hanya meneliti tiga metode klasifikasi sehingga tidak mencakup semua pendekatan yang ada dalam analisis sentimen. Oleh karenanya, dianjurkan bagi penelitian selanjutnya untuk menggunakan teknik penyeimbangan data yaitu *undersampling* atau *oversampling* guna meningkatkan kinerja pada kelas minoritas. Selain itu, penting untuk mengeksplorasi model-model pembelajaran mendalam lain seperti LSTM, BiLSTM, atau varian terbaru dari transformer, serta melakukan fine-tuning yang lebih terfokus pada domain data. Diharapkan dengan penggunaan dataset yang lebih besar atau bahkan beragam, kemampuan generalisasi model akan meningkat, sehingga penelitian selanjutnya dapat menghadirkan temuan yang lebih mendalam dan menyeluruh dalam analisis sentimen teks berbahasa Indonesia.

REFERENCES

- Abdullah, T., & Ahmet, A. (2023). Deep Learning in Sentiment Analysis: Recent Architectures. *ACM Computing Surveys*, 55(8). <https://doi.org/10.1145/3548772>
- Alahmadi, K., Alharbi, S., Chen, J., & Wang, X. (2025). Generalizing sentiment analysis: a review of progress, challenges, and emerging directions. *Social Network Analysis and Mining*, 15(1), 1–28. <https://doi.org/10.1007/s13278-025-01461-8>
- 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.
- Geni, L., Yulianti, E., & Sensuse, D. I. (2023). Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using Bert Language Models. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, 9(3), 746–757. <https://doi.org/10.26555/jiteki.v9i3.26490>
- Irmawan, O. A., Budi, I., Santoso, A. B., & Putra, P. K. (2024). Improving Sentiment Analysis and Topic Extraction in Indonesian Travel App Reviews Through BERT Fine-Tuning. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 13(2), 359–370. <https://doi.org/10.23887/janapati.v13i2.77028>
- 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>
- Kowsari, K., Meimandi, K. J., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text classification algorithms: A survey. *Information (Switzerland)*, 10(4), 1–68. <https://doi.org/10.3390/info10040150>
- Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), 1–19. <https://doi.org/10.1007/s13278-021-00776-6>
- Nugroho, K. S., Sukmadewa, A. Y., Wuswilahaken Dw, H., Bachtiar, F. A., & Yudistira, N. (2021). BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews. *ACM International Conference Proceeding Series, April*, 258–264. <https://doi.org/10.1145/3479645.3479679>
- Powers, D. M. W. (2020). *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation*. 37–63. <http://arxiv.org/abs/2010.16061>
- Putra, V. H. C., Kanugrahan, G., Tho, C., Siswaja, H. D., Prasetyo, R. T., & Ramdhani, Y. (2025). Real-Time Sentiment Analysis of the Salaman App: A Comparative Study of SVM, LSTM, and BERT Models with IoT Integration. *2025 Tenth International Conference on Informatics and Computing (ICIC)*, 1–6.



- <https://doi.org/10.1109/icic68054.2025.11309525>
- Ramdhan, H. M., Dwifebri Purbolaksono, M., & Bunyamin, B. (2024). Sentiment Analysis of Beauty Product Reviews Using the IndoBERT Method and Naive Bayes Classification. *2024 12th International Conference on Information and Communication Technology, ICoICT 2024*, 397–404. <https://doi.org/10.1109/ICoICT61617.2024.10698198>
- Raschka, S., Yuxi, L., & Mirjalili, V. (2022). *Machine Learning with PyTorch and Scikit-Learn* (1st ed.). Packt Publishing.
- 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>
- Sahoo, C., Wankhade, M., & Singh, B. K. (2023). Sentiment analysis using deep learning techniques: a comprehensive review. *International Journal of Multimedia Information Retrieval*, 12(2), 41. <https://doi.org/10.1007/s13735-023-00308-2>
- Setiawan, B. (2025). A Review of Sentiment Analysis Applications in Indonesia Between 2023-2024. *Journal of Information Engineering and Educational Technology*, 8(2), 71–83. <https://doi.org/10.26740/jieet.v8n2.p71-83>
- 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>
- Taufiq Dwi Purnomo, & Joko Sutopo. (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>
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. In *Artificial Intelligence Review* (Vol. 55, Issue 7). Springer Netherlands. <https://doi.org/10.1007/s10462-022-10144-1>
- Wijaya, D. R., Sasmita, G. M. A., & Vihikan, W. O. (2024). Sentiment Analysis of Indonesian Citizens on Electric Vehicle Using FastText and BERT Method. *Journal of Information Systems and Informatics*, 6(3), 1360–1372. <https://doi.org/10.51519/journalisi.v6i3.784>