

ANALISIS SENTIMEN PENGGUNA MYPERTAMINA PADA GOOGLE PLAYSTORE MENGGUNAKAN NAÏVE BAYES CLASSIFIER

Ekie Nurrohman¹, Rudi Kurniawan², Bani Nurhakim³, Ratna Dewi Lestyorini⁴

Program Studi Teknik Informatika¹²
Program Studi Manajemen Informatika³
Program Studi Sistem Informasi⁴

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
ekienurrohman03@gmail.com

(*) Corresponding Author : ekienurrohman03@gmail.com
Published : 30 Juni 2026

Abstract—The development of digital services in the energy sector has driven the use of the MyPertamina app as a platform for fuel transactions and energy subsidy management by the public. As the number of users increases, the reviews posted on the Google Play Store continue to grow, reflecting a variety of perceptions, experiences, and levels of user satisfaction with the app. However, the sheer volume of available reviews makes manual analysis inefficient and difficult to conduct systematically. This situation creates a need for an automated method capable of processing and classifying user opinions quickly and accurately. Therefore, this study aims to analyze user sentiment toward the MyPertamina app by classifying reviews into positive, negative, and neutral categories using a machine learning approach. The method used in this study is the Multinomial Naïve Bayes Classifier algorithm with feature representation using Term Frequency-Inverse Document Frequency (TF-IDF). The dataset consists of 9,104 user reviews obtained from the Google Play Store and has undergone text preprocessing stages, including data cleaning, tokenization, stopword removal, and stemming. To address the issue of data distribution imbalance across sentiment classes, this study employs a class prior approach to reduce prediction bias toward the majority class. Model evaluation was conducted using accuracy, precision, recall, and F1-score metrics to measure classification performance. The results show that the Multinomial Naïve Bayes model is capable of classifying user sentiment with an accuracy rate of 80.56%. The model performed well on the positive and negative classes, but its performance on the neutral class remained relatively low due to the imbalance in data distribution and the ambiguous nature of the review language. Overall, the applied method provides insight into user perceptions of the MyPertamina app and can serve as a basis for developers to evaluate and improve the app's service quality.

Keywords: Sentiment Analysis, Naïve Bayes, MyPertamina, TF-IDF, Class Prior.

Abstrak—Perkembangan layanan digital pada sektor energi mendorong penggunaan aplikasi MyPertamina sebagai sarana transaksi bahan bakar dan pengelolaan subsidi energi oleh masyarakat. Seiring meningkatnya jumlah pengguna, ulasan yang diberikan pada *Google Play Store* juga terus bertambah dan mencerminkan berbagai persepsi, pengalaman, serta tingkat kepuasan pengguna terhadap aplikasi tersebut. Namun, banyaknya ulasan yang tersedia membuat proses analisis secara manual menjadi tidak efisien dan sulit dilakukan secara sistematis. Kondisi ini menimbulkan kebutuhan akan metode otomatis yang mampu mengolah dan mengklasifikasikan opini pengguna secara cepat dan akurat. Oleh karena itu, penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi MyPertamina dengan mengklasifikasikan ulasan ke dalam kategori positif, negatif, dan netral menggunakan pendekatan machine learning. Metode yang digunakan dalam penelitian ini adalah algoritma Multinomial Naïve Bayes Classifier dengan representasi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Dataset yang digunakan terdiri dari 9.104 ulasan pengguna yang diperoleh dari Google Play Store dan telah melalui tahapan preprocessing teks yang meliputi *data cleaning*, *tokenization*, *stopword removal*, dan *stemming*. Untuk mengatasi permasalahan ketidakseimbangan distribusi data antar kelas sentimen, penelitian ini menerapkan pendekatan class prior guna mengurangi bias prediksi terhadap kelas mayoritas. Proses evaluasi model dilakukan dengan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur kinerja klasifikasi. Hasil penelitian menunjukkan bahwa model Multinomial Naïve Bayes mampu mengklasifikasikan sentimen pengguna dengan tingkat akurasi sebesar 80,56%. Model menunjukkan

performa yang baik pada kelas positif dan negatif, namun kinerja pada kelas netral masih relatif rendah akibat ketidakseimbangan distribusi data dan karakteristik bahasa ulasan yang ambigu. Secara keseluruhan, metode yang diterapkan mampu memberikan gambaran mengenai persepsi pengguna terhadap aplikasi MyPertamina serta dapat menjadi dasar evaluasi bagi pengembang dalam meningkatkan kualitas layanan aplikasi.

Kata Kunci: Analisis Sentimen, *Naïve Bayes*, MyPertamina, TF-IDF, *Class Prior*.

INTRODUCTION

Perkembangan teknologi informasi pada era digital saat ini telah membawa perubahan yang signifikan dalam berbagai aspek kehidupan masyarakat, khususnya dalam pemanfaatan layanan berbasis aplikasi. Transformasi digital mendorong instansi pemerintah maupun sektor swasta untuk terus berinovasi dalam menyediakan layanan yang lebih efisien, cepat, dan mudah diakses. Hal ini didukung oleh meningkatnya penggunaan perangkat mobile dan akses internet yang semakin luas, sehingga aplikasi digital menjadi salah satu media utama dalam interaksi antara penyedia layanan dan pengguna. Dalam era digital saat ini, pertumbuhan data teks dari berbagai platform seperti media sosial dan aplikasi mobile meningkat secara signifikan, sehingga diperlukan metode otomatis untuk menganalisis opini pengguna secara efisien. Analisis sentimen menjadi solusi yang banyak digunakan karena mampu mengelompokkan opini pengguna secara sistematis dan membantu dalam pengambilan keputusan berbasis data (Tf-idf, 2026).

Salah satu bentuk implementasi transformasi digital tersebut dapat dilihat pada sektor energi melalui hadirnya aplikasi MyPertamina yang dikembangkan oleh PT Pertamina (Persero). Aplikasi ini berfungsi sebagai sarana transaksi non-tunai serta pengelolaan layanan bahan bakar minyak (BBM) yang lebih modern dan terintegrasi. Dengan jumlah pengguna yang terus meningkat, aplikasi MyPertamina memiliki peran penting dalam menunjang aktivitas masyarakat sehari-hari, khususnya dalam proses pembelian BBM di SPBU secara lebih praktis dan efisien.

Namun, seiring dengan meningkatnya jumlah pengguna, berbagai permasalahan mulai muncul yang tercermin melalui ulasan pengguna pada platform Google Play Store. Ulasan tersebut menunjukkan beragam pengalaman pengguna, mulai dari kepuasan terhadap kemudahan layanan hingga keluhan terkait kendala teknis seperti kesulitan login, error sistem, proses transaksi yang lambat, serta tampilan antarmuka yang kurang optimal. Ulasan ini

menjadi sumber informasi yang sangat penting karena mencerminkan persepsi pengguna secara langsung terhadap kualitas layanan aplikasi. Ulasan pengguna menjadi sumber informasi yang penting karena mencerminkan pengalaman dan persepsi pengguna secara langsung terhadap kualitas layanan aplikasi. Hal ini sejalan dengan penelitian yang menyatakan bahwa ulasan pengguna pada platform digital dapat digunakan sebagai indikator dalam mengevaluasi kualitas layanan serta mengidentifikasi permasalahan yang dialami pengguna secara lebih objektif (Chairuddin et al., 2021)

Di sisi lain, jumlah ulasan pengguna yang terus bertambah dalam jumlah besar menyebabkan proses analisis secara manual menjadi tidak efektif dan memerlukan waktu yang lama. Selain itu, karakteristik data ulasan yang tidak terstruktur, penggunaan bahasa tidak baku, singkatan, serta kesalahan penulisan menambah kompleksitas dalam proses analisis. Kondisi ini menunjukkan perlunya suatu pendekatan yang mampu mengolah data teks secara otomatis, cepat, dan akurat untuk menghasilkan informasi yang lebih terstruktur dan dapat digunakan sebagai dasar pengambilan keputusan.

Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah analisis sentimen, yaitu teknik dalam text mining yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini pengguna ke dalam kategori tertentu seperti positif, negatif, dan netral. Dengan menggunakan analisis sentimen, data ulasan yang semula tidak terstruktur dapat diubah menjadi informasi yang lebih terorganisir, sehingga memudahkan dalam memahami kecenderungan persepsi pengguna terhadap suatu aplikasi.

Analisis sentimen sangat penting untuk mengevaluasi persepsi publik terhadap aplikasi pembayaran digital karena ekstraksi skala besar dan real-time dari ulasan pengguna di toko aplikasi untuk menginformasikan prioritas fitur, mendeteksi masalah kegunaan atau kepercayaan, dan mengarahkan respons kebijakan atau produk (Barik et al., 2023). Kemajuan terbaru kerangka kerja deep learning hibrida dan LSTM yang dioptimalkan, peningkatan penimbangan istilah dan pemilihan fitur, serta pengujian benchmark yang kokoh terhadap model dasar klasik seperti

Naïve Bayes telah secara signifikan meningkatkan akurasi dan wawasan tingkat aspek untuk korpus ulasan e-commerce dan perbankan seluler (Sianipar et al., 2024). Ulasan sistematis dan bibliometrik lebih lanjut mengonfirmasi kematangan bidang ini sambil menyoroti tantangan praktis (ketidakseimbangan kelas, sarkasme, transfer domain) yang langsung relevan dengan ulasan pengguna Google Play Store untuk MyPertamina. Bersama-sama, studi-studi terbuka dan direview oleh rekan sejawat ini menunjukkan bahwa analisis sentimen menyediakan metrik berbasis bukti tentang kepuasan pengguna dan sinyal peringatan dini yang esensial untuk memantau adopsi, kepercayaan, dan perbaikan layanan dalam aplikasi pembayaran digital.

Meskipun penelitian terkait analisis sentimen telah banyak dilakukan pada berbagai domain seperti e-commerce dan media sosial, penelitian yang secara khusus mengkaji ulasan pengguna aplikasi MyPertamina masih relatif terbatas. Selain itu, berdasarkan kajian penelitian terdahulu, sebagian besar penelitian masih belum secara optimal menangani permasalahan ketidakseimbangan data (class imbalance), yang merupakan kondisi di mana distribusi jumlah data antar kelas tidak merata. Hal ini dapat menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas dan kurang akurat dalam mengenali kelas minoritas.

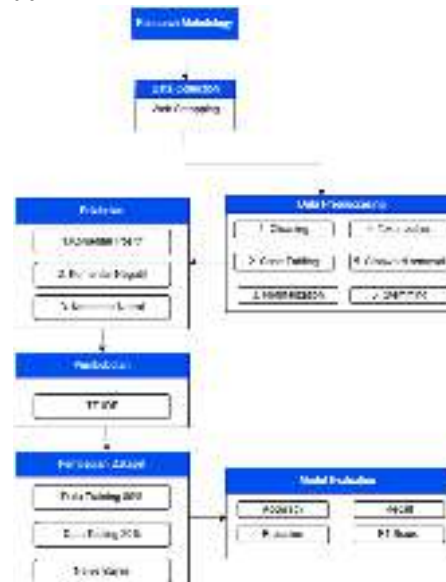
Lebih lanjut, beberapa penelitian sebelumnya yang mengkaji topik serupa pada aplikasi MyPertamina cenderung hanya berfokus pada perbandingan metode klasifikasi tanpa memperhatikan optimasi distribusi data atau pendekatan khusus dalam meningkatkan performa model. Oleh karena itu, diperlukan suatu pendekatan yang tidak hanya mampu melakukan klasifikasi sentimen, tetapi juga dapat mengatasi permasalahan ketidakseimbangan data agar hasil yang diperoleh lebih akurat dan representatif.

Berdasarkan permasalahan dan kesenjangan tersebut, peneliti tertarik untuk melakukan penelitian analisis sentimen terhadap ulasan pengguna aplikasi MyPertamina dengan menggunakan algoritma *Naïve Bayes Classifier*. Metode ini dipilih karena memiliki keunggulan dalam klasifikasi teks yang sederhana, efisien, dan memiliki performa yang cukup baik. Selain itu, penelitian ini juga menerapkan pendekatan *class prior* untuk mengatasi ketidakseimbangan data, sehingga diharapkan dapat meningkatkan kemampuan

model dalam mengenali seluruh kategori sentimen secara lebih optimal.

MATERIALS AND METHODS

Desain penelitian ini disusun untuk menjelaskan secara eksplisit hasil yang ingin dicapai dalam menjawab pertanyaan penelitian yang telah dirumuskan sebelumnya. Setiap tujuan menggambarkan langkah sistematis yang dilakukan peneliti dalam memperoleh solusi terhadap permasalahan yang telah diidentifikasi. Dengan demikian, tujuan penelitian ini berperan sebagai arah dan sasaran akhir dari seluruh rangkaian kegiatan penelitian yang dilakukan. Desain penelitian ini dapat dilihat pada gambar 1 desain penelitian



Gambar 1 Desain Penelitian

Gambar tersebut menunjukkan alur metodologi penelitian yang digunakan dalam proses pembangunan model analisis sentimen pada ulasan aplikasi MyPertamina. Tahapan penelitian dimulai dari proses pengumpulan data menggunakan metode web scraping untuk memperoleh komentar pengguna dari *Google Play Store*. Setelah data terkumpul, dilakukan proses pelabelan sentimen ke dalam tiga kategori, yaitu positif, negatif, dan netral sebagai dasar klasifikasi.

Berikutnya adalah data preprocessing, yang meliputi cleaning, case folding, normalisasi, tokenisasi, stopwords removal, dan stemming. Langkah ini bertujuan untuk membersihkan serta menyederhanakan teks agar lebih siap diolah pada tahap berikutnya. Setelah preprocessing selesai, data diubah menjadi representasi numerik melalui metode TF-IDF pada tahap feature extraction.

Selanjutnya, data dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data

uji. Pembagian ini digunakan untuk melatih dan menguji model klasifikasi menggunakan algoritma Multinomial Naïve Bayes. Tahap akhir dari metodologi ini adalah evaluasi model menggunakan metrik akurasi, precision, recall, dan f1-score untuk mengukur seberapa baik model mampu mengklasifikasikan sentimen berdasarkan hasil prediksi.

RESULTS AND DISCUSSION

Penerapan algoritma Naïve Bayes Classifier dalam penelitian ini dilakukan dengan membagi dataset menjadi data latih dan data uji menggunakan perbandingan 80:20. Sebanyak 80% data digunakan sebagai data latih untuk membangun model klasifikasi, di mana algoritma mempelajari pola distribusi kata terhadap masing-masing kelas sentimen berdasarkan hasil ekstraksi fitur. Proses pelatihan dilakukan dengan menghitung probabilitas prior dan likelihood dari setiap fitur terhadap kelas sentimen.

Setelah tahap pembobotan fitur atau feature extraction dengan metode TF-IDF selesai dilakukan, langkah berikutnya adalah melatih model klasifikasi sentimen menggunakan algoritma Multinomial Naïve Bayes. Algoritma ini dipilih karena memiliki kemampuan yang baik dalam mengolah data teks yang telah diubah ke dalam bentuk numerik, serta mampu bekerja secara efisien pada jumlah fitur yang besar. Sebelum proses pelatihan dimulai, dataset dibagi menjadi dua bagian, yaitu data latih (training data) dan data uji (testing data) dengan menggunakan fungsi `train_test_split` dari pustaka `scikit-learn`. Pembagian Data dibagi menjadi dua bagian dengan proporsi 80% sebagai data latih dan 20% sebagai data uji.

Tujuan dari pembagian ini adalah agar model dapat mempelajari pola sentimen secara optimal dari sebagian besar data, sekaligus diuji performanya pada data yang tidak pernah dilihat sebelumnya. Dengan cara ini, kemampuan generalisasi model terhadap ulasan baru dapat dievaluasi secara lebih akurat. Dapat dilihat pada Tabel 1 Menunjukkan Hasil Pembagian Dataset.

Tabel 1 Hasil Pembagian Dataset

Jenis Data	Jumlah	Persentasi(%)
Data Latih	7283	80%
Data Uji	1821	20%
Total	9104	100%

Pembagian dataset dilakukan dengan proporsi 80% untuk data latih dan 20% untuk data uji. Dari total data setelah *preprocessing* 9.104 ulasan yang tersedia, sebanyak 7.283 data digunakan sebagai data latih atau *training* untuk membangun dan menyesuaikan model klasifikasi, sedangkan 1.821 data sisanya dialokasikan sebagai data uji guna mengukur kinerja model secara objektif. Pembagian ini memastikan bahwa model memperoleh jumlah data yang memadai untuk proses pembelajaran sekaligus memiliki porsi data independen yang cukup untuk proses evaluasi, sehingga hasil pengujian dapat mencerminkan kemampuan model dalam melakukan generalisasi pada data yang belum pernah dilihat sebelumnya.

Tabel 2 Hasil Pelatihan Model Naïve Bayes

No	Parameter	Nilai
1	Jumlah Data Latih	7283
2	Jumlah Data Uji	1821
3	Jumlah Fitur TF-IDF	6246
4	Alpha (Laplace Smoothing)	1.0
5	Class Prior	[0.33, 0.34, 0.33]
6	Fit Prior	True
bf	Force Alpha	True

Berdasarkan hasil pengujian model, jumlah data yang digunakan pada tahap pelatihan dan pengujian model Naïve Bayes dari total data yang ada mencapai 7.283 sampel data latih, sementara data uji terdiri dari 1.821 sampel. Proses ekstraksi fitur menggunakan metode TF-IDF menghasilkan 6.246 fitur yang menjadi representasi teks dalam bentuk numerik. Model Naive Bayes dilatih dengan nilai alpha sebesar 1.0 sebagai penerapan Laplace smoothing untuk mencegah probabilitas nol pada kata yang jarang muncul. Selain itu, model menerapkan *class prior* dengan distribusi [0.33, 0.34, 0.33] yang memberikan bobot awal yang seimbang pada ketiga kelas sentimen. Pengaturan Fit Prior diaktifkan (True), sehingga model menggunakan prior dalam perhitungan probabilitas, dan *Force Alpha* juga disetel ke True, memastikan *smoothing* diterapkan secara konsisten. Secara keseluruhan, parameter-parameter ini menjadi dasar konfigurasi akhir model yang digunakan untuk menghasilkan performa evaluasi pada penelitian ini.

Penggunaan Class Prior

Berdasarkan hasil eksperimen sebelumnya, algoritma *Naïve Bayes* tidak mampu mengenali

kelas netral secara konsisten karena jumlah datanya jauh lebih sedikit dibandingkan kelas positif dan negatif. Ketimpangan ini menyebabkan model membentuk probabilitas yang bias terhadap kelas mayoritas, sehingga kelas netral sering tidak muncul dalam prediksi. Untuk memperbaiki kondisi tersebut, diterapkan *class prior* sebagai penyeimbang probabilitas awal setiap kelas sebelum proses perhitungan likelihood dilakukan. Penetapan *class prior* membuat model memberikan bobot tambahan pada kelas dengan jumlah data lebih rendah, sehingga *Naive Bayes* tetap mempertimbangkan keberadaan kelas netral meskipun representasinya terbatas. Dampaknya terlihat pada tahap evaluasi, di mana model menjadi lebih stabil, kemampuan klasifikasi meningkat, dan seluruh kelas dapat terdeteksi dengan lebih baik.

Kelebihan penggunaan *class prior* adalah kemampuannya menyeimbangkan probabilitas awal pada setiap kelas sehingga model tidak bias terhadap kelas yang jumlah datanya lebih banyak. Dengan adanya *class prior*, *Naive Bayes* dapat mempertimbangkan kelas minoritas seperti kelas netral dalam penelitian ini secara lebih proporsional, sehingga hasil prediksi menjadi lebih stabil, akurat, dan mampu membaca seluruh kelas secara lebih konsisten.

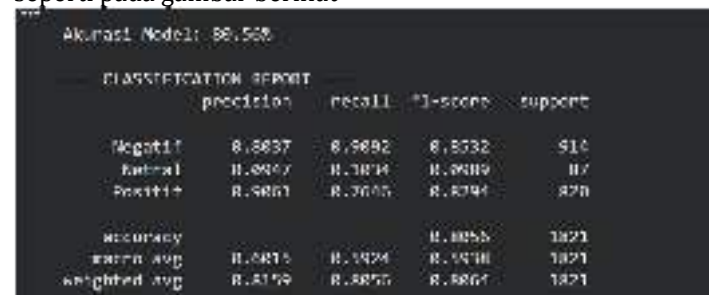
Kelemahan dari penggunaan *class prior* adalah model menjadi sangat bergantung pada nilai probabilitas awal yang ditentukan. Jika *class prior* tidak ditetapkan secara tepat atau tidak mencerminkan karakteristik data sebenarnya, maka model dapat menghasilkan prediksi yang bias dan kurang akurat. Selain itu, penentuan prior yang terlalu besar pada kelas tertentu dapat membuat model mengabaikan pola alami dalam data dan justru menurunkan kinerja klasifikasi secara keseluruhan.

Pada tahap evaluasi model, kinerja algoritma *Naive Bayes Classifier* diukur menggunakan beberapa metrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Evaluasi dilakukan dengan membandingkan hasil prediksi model terhadap data uji dengan label sebenarnya. *Accuracy* digunakan untuk mengetahui tingkat ketepatan model secara keseluruhan dalam mengklasifikasikan sentimen, sedangkan *precision* mengukur tingkat ketepatan prediksi pada masing-masing kelas sentimen. *Recall* digunakan untuk mengukur kemampuan model dalam mengidentifikasi seluruh data yang relevan pada setiap kelas, dan *F1-score* merupakan nilai harmonisasi antara *precision* dan *recall* yang

memberikan gambaran keseimbangan kinerja model.

2. Classification Report

Gambar ini menyajikan hasil evaluasi kinerja model *Naive Bayes* dalam mengklasifikasikan sentimen ulasan pengguna. Laporan tersebut memperlihatkan nilai *accuracy*, *precision*, *recall*, dan *f1-score* untuk setiap kelas sentimen, sehingga memberikan gambaran menyeluruh mengenai kemampuan model dalam membedakan ulasan negatif, netral, dan positif, seperti pada gambar berikut

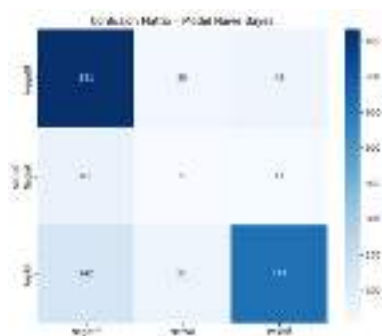


Akurasi Model: 88.56%				
	precision	recall	f1-score	support
Negatif	0.8037	0.9892	0.8932	910
Netral	0.0947	0.1034	0.0989	117
Positif	0.9061	0.7665	0.8394	828
accuracy			0.8856	1855
macro avg	0.4815	0.4424	0.4418	1855
weighted avg	0.8159	0.8875	0.8867	1855

Gambar 3 Classification Report

Berdasarkan hasil evaluasi model, akurasi yang diperoleh adalah sebesar 80,56%, yang menunjukkan bahwa model mampu melakukan prediksi dengan tingkat ketepatan yang cukup baik. Pada kelas sentimen negatif dan positif, model menunjukkan performa yang tinggi dengan nilai *precision* masing-masing 0,8037 dan 0,9061, serta *f1-score* sebesar 0,8532 dan 0,8294. Hal ini menunjukkan bahwa model mampu mengenali kedua kelas tersebut dengan baik. Namun, performa model pada kelas netral masih rendah, ditunjukkan oleh nilai *precision* 0,0947, *recall* 0,1034, dan *f1-score* 0,0989. Kondisi ini mengindikasikan bahwa model masih kesulitan membedakan kelas netral dari kelas lain, yang disebabkan oleh ketidakseimbangan jumlah data. Secara keseluruhan, nilai *macro average f1-score* sebesar 0,5938 menunjukkan adanya variasi performa antar kelas, sedangkan nilai *weighted average* 0,8064 menggambarkan bahwa model bekerja lebih baik pada kelas mayoritas. Dengan demikian, model sudah cukup baik dalam klasifikasi dua kelas dominan, tetapi memerlukan peningkatan dalam mendeteksi kelas netral.

Gambar berikut menyajikan *confusion matrix* yang digunakan untuk menggambarkan kinerja model dalam melakukan klasifikasi sentimen. Melalui matriks ini, dapat dilihat perbandingan antara label aktual dengan hasil prediksi model, sehingga memudahkan analisis terkait tingkat akurasi, pola kesalahan prediksi, serta kemampuan model dalam membedakan setiap kelas sentimen yang diuji. Seperti pada gambar berikut



Gambar 4 Visualisasi Confusion Matrix Naïve Bayes

Berdasarkan *confusion matrix model Naïve Bayes* menunjukkan kemampuan yang baik dalam mengklasifikasikan sentimen negatif dan positif, namun masih kurang optimal dalam mendeteksi sentimen netral. Pada kelas negatif, sebanyak 831 data berhasil diklasifikasikan dengan benar, sementara sebagian kecil salah diprediksi sebagai netral (35 data) dan positif (48 data). Pada kelas positif, model juga menunjukkan performa baik dengan 627 prediksi benar, meskipun terdapat kesalahan klasifikasi ke kelas negatif (142 data) dan kelas netral (51 data). Sebaliknya, kelas netral menunjukkan hasil yang paling rendah, dengan hanya 9 data yang terklasifikasi dengan benar. Sebagian besar data netral justru diprediksi sebagai negatif (61 data) atau positif (17 data), yang menunjukkan kesulitan model dalam membedakan kelas ini. Secara keseluruhan, hasil ini menggambarkan bahwa model lebih mampu mengenali kelas mayoritas (negatif dan positif), sementara performa pada kelas netral masih perlu ditingkatkan.

CONCLUSION

Penelitian ini berhasil mengolah dan menganalisis ribuan ulasan pengguna aplikasi MyPertamina yang sebelumnya sulit dianalisis secara manual. Melalui proses otomatisasi text preprocessing dan analisis sentimen, persepsi pengguna dapat dipetakan menjadi tiga kategori sentimen, yaitu positif, negatif, dan netral.

Penelitian ini telah berhasil membangun sistem otomatis klasifikasi sentimen menggunakan algoritma Multinomial Naïve Bayes dengan representasi TF-IDF. Sistem ini mampu melakukan proses klasifikasi secara lebih cepat dan efisien dibandingkan analisis manual.

Evaluasi performa model menunjukkan bahwa algoritma Multinomial Naïve Bayes memiliki akurasi sebesar 80,56%, sehingga dapat dikatakan cukup baik dalam

mengklasifikasikan sentimen ulasan pengguna aplikasi MyPertamina, terutama pada kelas positif dan negatif, meskipun performa pada kelas netral masih perlu ditingkatkan

REFERENCE

- Alenazi, F. S., & Hindi, K. El. (2023). *applied sciences Complement-Class Harmonized Naïve Bayes Classifier*. 1–18.
- Ammar, H., & Wiranata, A. D. (2025). *Perbandingan Tingkat Akurasi Algoritma Naïve Bayes, Decision Tree dan Random Forest dalam Analisis Sentimen Pengguna Aplikasi Samsat Digital Nasional*. 10(1), 59–73.
- Analisis, U., & Berbasis, S. (2023). *Jurnal Indonesia : Manajemen Informatika dan Komunikasi PENERAPAN ALGORITMA TF-IDF DAN NAÏVE BAYES Jurnal Indonesia : Manajemen Informatika dan Komunikasi*. 4(3), 1822–1834.
- Apriyani, M. E., Nur, A. F., & Astuti, E. S. (2024). *Performance Comparison of Naïve Bayes and SVM Algorithms in Sentiment Analysis on JKN Application Data*. 04(02), 180–188.
- Araujo, A. F., Gôlo, M. P. S., & Marcacini, R. M. (2022). Opinion mining for app reviews: an analysis of textual representation and predictive models. *Automated Software Engineering*, 29(1). <https://doi.org/10.1007/s10515-021-00301-1>
- Ataei, P., Regula, S., Staegemann, D., & Malgaonkar, S. (2024). Filtering Useful App Reviews Using Naïve Bayes—Which Naïve Bayes? *AI (Switzerland)*, 5(4), 2237–2259. <https://doi.org/10.3390/ai5040110>
- Chairuddin, I., Putra, T. N., Tinggi, S., Pekanbaru, T., Pakar, S., & Chaining, F. (2021). *Perancangan Sistem Pakar Jenis Kepribadian Menggunakan Metode Forward Chaining Berbasis Web*. 8(3), 1200–1213.
- Dąbrowski, J., Letier, E., Perini, A., & Susi, A. (2022). Analysing app reviews for software engineering: a systematic literature review. *Empirical Software Engineering*, 27(2). <https://doi.org/10.1007/s10664-021-10065-7>
- Dana, A. (2023). *Hybrid Machine Learning for Sentiment Analysis of Dana Application Reviews*. 7(2), 51–58. <https://doi.org/10.21070/jicte.v7i2.1651>
- Gôlo, M. P. S., Araújo, A. F., Rossi, R. G., & Marcacini, R. M. (2022). Detecting relevant app reviews for software evolution and maintenance through multimodal one-class learning. *Information and Software Technology*, 151. <https://doi.org/10.1016/j.infsof.2022.10699>

- 8
- Hadi, M. A., & Fard, F. H. (2023). Evaluating pre-trained models for user feedback analysis in software engineering: a study on classification of app-reviews. *Empirical Software Engineering*, 28(4). <https://doi.org/10.1007/s10664-023-10314-x>
- Hartati, T., Sohadi, R. T., Tohidi, E., & Wahyudin, E. (2024). Penerapan Algoritma Naive Bayes pada Analisis Sentimen Ulasan Aplikasi Whoosh – Kereta Cepat Di Google Play Store. 6(1), 244–249.
- Heti Aprilianti, Khothibul Umam, & Maya Rini Handayani. (2025). Comparative Study of SVM, KNN, and Naïve Bayes for Sentiment Analysis of Religious Application Reviews. *Journal of Applied Informatics and Computing*, 9(3), 920–927. <https://doi.org/10.30871/jaic.v9i3.9482>
- Khairunnas, R., Pagua, J. A., Fitriya, G., & Ruldeviyani, Y. (2025). User sentiment dynamics in social media : a comparative analysis of X and Threads. 14(1), 447–456. <https://doi.org/10.11591/ijai.v14.i1.pp447-456>
- Khairunnisa, S., & Faraby, S. Al. (2021). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi. 5(April), 406–414. <https://doi.org/10.30865/mib.v5i2.2835>
- Khasanah, A. U., & Febriyanti, A. (2022). Sentiment Analysis of JNE User Perception using Naïve Bayes Classifier Algorithm. 15(1), 124–130.
- Krasadakis, P., Sakkopoulos, E., & Verykios, V. S. (2024). A Survey on Challenges and Advances in Natural Language Processing with a Focus on Legal Informatics and Low-Resource Languages. *Electronics (Switzerland)*, 13(3). <https://doi.org/10.3390/electronics13030648>
- Lestari, V. B., & Hutagalung, C. A. (n.d.). Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM , Logistic Regression , and Naive Bayes. 021, 36–44.
- Li, Q., Peng, H., Li, J., Xia, C., Yang, R., Sun, L., & Yu, P. S. (2022). Twenty years of machine-learning-based text classification: A systematic review. *Algorithms*, 16(5). <https://doi.org/10.3390/algorithms16050236>
- Li, Q., Peng, H., Li, J., Xia, C., Yang, R., Sun, L., Yu, P. S., & He, L. (2021). Learning. *ACM Transactions on Intelligent Systems and Technology*, 37(4).
- Luo, Y., Liu, Y., Liu, J., Tang, S., Harkin, J., & Cao, Y. (2021). Counteracting dynamical degradation of a class of digital chaotic systems via Unscented Kalman Filter and perturbation. *Information Sciences*, 556, 49–66. <https://doi.org/10.1016/j.ins.2020.12.065>
- Madjid, M. F., Ratnawati, D. E., & Rahayudi, B. (2023). Sentiment Analysis on App Reviews Using Support Vector Machine and Naïve Bayes Classification. *Sinkron*, 8(1), 556–562. <https://doi.org/10.33395/sinkron.v8i1.12161>
- Maharani Pakpahan, M., Halmi Dar, M., & Nirmala Sari Hasibuan, M. (2025). Performance Evaluation of Machine Learning Algorithms in Aspect-Based Sentiment Analysis on E-Commerce User Reviews. *International Journal of Science, Technology & Management*, 6(4), 958–965. <https://doi.org/10.46729/ijstm.v6i4.1255>
- Nainggolan, N., Sarumaha, B., Lumbanbatu, J. R., & Aisyah, S. (2024). Analisis Sentimen Pengguna Dompot Digital Menggunakan Algoritma Multivariat Bernoulli (Studi Kasus: Ovo Dan Gopay). *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, 7(1), 345. <https://doi.org/10.37600/tekinkom.v7i1.1223>
- Palomino, M. A. (2022). applied sciences Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis.
- Rainio, O. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, 1–14. <https://doi.org/10.1038/s41598-024-56706-x>
- Sandu, A., Cotfas, L. A., Stănescu, A., & Delcea, C. (2024). A Bibliometric Analysis of Text Mining: Exploring the Use of Natural Language Processing in Social Media Research. *Applied Sciences (Switzerland)*, 14(8). <https://doi.org/10.3390/app14083144>
- Saraswati, D. Y., Handayani, M. R., Umam, K., & Mustofa, M. I. (2025). Sentiment Classification of MyPertamina Reviews Using Naïve Bayes and Logistic Regression. 9(4).
- Setiawan, B. (2024). A Review of Sentiment Analysis Applications in Indonesia Between 2023-2024. 08, 71–83.
- Suryana, A., Purnamasari, A. I., & Ali, I. (2024). MENGOPTIMALKAN KEPUASAN PENGGUNA : ANALISIS SENTIMEN REVIEW APLIKASI GRAB DI INDONESIA. 8(3), 3396–3404.
- Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research. *Applied Sciences*

(Switzerland), 13(7).
<https://doi.org/10.3390/app13074550>