



Prediksi Tingkat Kepuasan Pengguna Aplikasi Grab di *Google Play Store* Menggunakan Algoritma *Naive Bayes*

Aditya Monalisa¹, Yudistira Bagus Pratama², Rike Pradila³

^{1,2,3} Universitas Muhammadiyah Bangka Belitung

Email: adityamonalisa04@gmail.com

Article Info

Article history:

Received April 15, 2026

Revised April 23, 2026

Accepted Mei 13, 2026

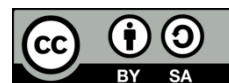
Keywords:

Naive Bayes, Sentiment Analysis, User Satisfaction, Online Transport, Machine Learning

ABSTRACT

The rapid growth of online transportation services in recent years has increased the importance of evaluating user satisfaction as a basis for improving service quality. Grab, one of the leading online transportation applications, has a large number of reviews on the Google Play Store. However, these review data are still unstructured, making them difficult to analyze manually. This study aims to develop an automatic prediction system for user satisfaction levels by classifying reviews into positive, neutral, and negative sentiments, as well as grouping them into categories such as driver, system, and others. The method used in this study is Multinomial Naive Bayes with a multilabel data labeling approach. Data were collected through web scraping techniques covering the period from January 2020 to February 2026. This research applied a quantitative approach with interactive analysis to examine the data, while the model development process followed the CRISP-DM methodology, including business understanding, data understanding, data preparation, modeling, evaluation, and deployment through a web-based system using Streamlit. The results show that the Naive Bayes model was able to classify reviews effectively based on evaluation metrics such as accuracy, precision, recall, and F1-score. The dataset consisted of 234,878 reviews, which were reduced to 183,079 data after preprocessing, providing sufficient variation for model training. The developed system was also able to present analysis results through an informative and user-friendly dashboard. Therefore, this study can help understand user satisfaction levels automatically and serve as a basis for strategic decision-making to improve the quality of online transportation services.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Article Info

Article history:

Received April 15, 2026

Revised April 23, 2026

Accepted Mei 13, 2026

Keywords:

Naive Bayes, Analisis Sentimen, Kepuasan Pengguna, Transportasi Online, Machine Learning

ABSTRAK

Perkembangan transportasi online yang meningkat dalam beberapa terakhir mendorong pentingnya evaluasi tingkat kepuasan pengguna sebagai dasar dalam peningkatan kualitas layanan. Grab, salah satu aplikasi transformasi online yang memiliki banyak ulasan di *Google Play Store*. Namun, data ulasan tersebut masih bersifat tidak terstruktur sehingga sulit untuk dianalisis secara manual. Penelitian ini bertujuan membangun sistem prediksi tingkat kepuasan pengguna secara otomatis dengan mengklasifikasikan ulasan ke dalam sentimen puas, netral, dan tidak puas, serta mengelompokkan ke dalam kategori driver, sistem, dan lainnya. Metode yang digunakan penelitian ini adalah *Multinomial Naive Bayes* dengan pendekatan labeling data menggunakan *Multilabel*. Data diperoleh melalui teknik *web scraping*



dengan periode Januari 2020 hingga Februari 2026. Penelitian ini menggunakan pendekatan kuantitatif dengan analisis interaktif untuk menganalisis data dan proses pengembangan model yang digunakan merupakan metode CRISP-DM yang meliputi pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan deployment dalam sistem berbasis web menggunakan *Streamlit*. Hasil penelitian menunjukkan bahwa model Naive Bayes mampu mengklasifikasikan ulasan dengan baik berdasarkan metrik evaluasi seperti accuracy, precision, recall, dan F1-score. Dataset yang digunakan berjumlah 234.878 ulasan dan setelah *preprocessing* menjadi 183.079 data, sehingga memberikan variasi data yang cukup besar untuk pelatihan model. Sistem yang dikembangkan juga mampu menampilkan hasil analisis dalam bentuk dashboard yang informatif dan mudah dipahami. Dengan demikian, penelitian ini dapat membantu dalam memahami tingkat kepuasan pengguna secara otomatis serta menjadi dasar pengambilan keputusan strategis bagi peningkatan kualitas layanan transportasi online.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Tsabitha Putri Anjani
Universitas Muhammadiyah Bangka Belitung
Email: tsabitha266@gmail.com

PENDAHULUAN

Perkembangan teknologi digital dalam beberapa tahun terakhir telah membawa perubahan signifikan pada berbagai aspek kehidupan, termasuk sektor transportasi. Hal ini ditandai dengan meningkatnya jumlah kendaraan bermotor di Indonesia. Menurut Badan Pusat Statistik, pada periode 2020–2024 jumlah sepeda motor meningkat sebanyak 24.426.974 unit atau sekitar 17,52% (BPS, 2025). Peningkatan tersebut turut mendorong berkembangnya layanan transportasi berbasis aplikasi, terutama di wilayah perkotaan. Pemerintah juga telah menetapkan regulasi terkait operasional transportasi online melalui Peraturan Menteri Perhubungan Nomor PM 32 Tahun 2016 (Heristian et al., 2025). Salah satu perusahaan transportasi online terbesar yang banyak digunakan masyarakat adalah Grab, perusahaan asal Singapura yang didirikan tahun 2012 dan mulai beroperasi di Indonesia sejak Juni 2014 (Raya et al., 2025).

Meningkatnya kebutuhan terhadap layanan transportasi online menjadikan kepuasan pelanggan sebagai aspek penting bagi perusahaan dalam mempertahankan bisnis dan meningkatkan keuntungan. Kepuasan pelanggan dapat berdampak positif terhadap loyalitas pengguna dan citra layanan yang diberikan (Rahman & Sutanto, 2023). Berbagai penelitian sebelumnya telah mengkaji kepuasan pengguna transportasi online menggunakan teknik data mining dan machine learning. Penelitian Irawan et al. (2025) menunjukkan masih terdapat 17,53% pelanggan yang merasa tidak puas, namun penelitian tersebut hanya berfokus pada pengelompokan kepuasan dan belum menghasilkan model prediksi otomatis. Penelitian lain



oleh Rahayu & Cahyono (2022) menggunakan metode Decision Tree pada layanan Gojek, Grab, dan Maxim, tetapi belum menyajikan ukuran akurasi model secara kuantitatif.

Selain itu, penelitian sentiment analysis terhadap ulasan pengguna Grab di Google Play Store menunjukkan bahwa data ulasan aplikasi sangat potensial untuk menilai tingkat kepuasan pengguna secara otomatis. Berbagai studi telah menggunakan algoritma machine learning, khususnya Naive Bayes, untuk mengklasifikasikan sentimen menjadi puas dan tidak puas. Penelitian Ananda et al. (2024) menunjukkan adanya perbedaan tingkat sentimen berdasarkan komentar pengguna serta menegaskan bahwa ulasan di Google Play Store mencerminkan pengalaman nyata pengguna terhadap kualitas layanan. Penggunaan algoritma Naive Bayes juga terbukti efektif dalam mengolah data teks tidak terstruktur dengan evaluasi menggunakan akurasi, presisi, dan recall. Namun, sebagian besar penelitian masih berfokus pada analisis sentimen umum dan belum secara khusus memprediksi tingkat kepuasan pengguna berdasarkan variabel layanan transportasi online secara menyeluruh.

Berdasarkan keterbatasan tersebut, penelitian ini bertujuan mengembangkan model prediksi tingkat kepuasan pengguna layanan Grab menggunakan data ulasan aplikasi di Google Play Store periode Januari 2020 hingga Februari 2026. Dataset bersifat global sehingga mampu memberikan gambaran luas mengenai kualitas layanan. Penelitian ini menggunakan klasifikasi sentimen (Puas, Tidak Puas, Netral) yang dipadukan dengan kategori Driver, Sistem Aplikasi, dan Lainnya agar dapat mengidentifikasi aspek layanan secara lebih mendalam (Rahayu & Cahyono, 2022).

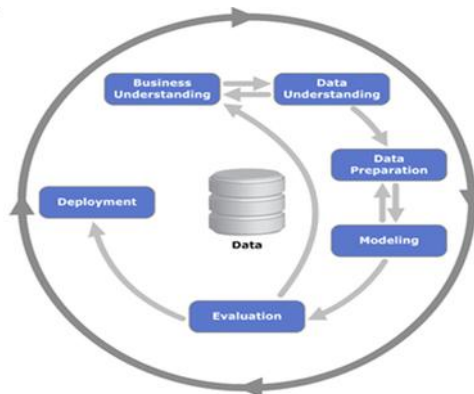
METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode algoritma Naive Bayes. Metode kuantitatif ini bertujuan untuk memberikan penjelasan terhadap fenomena berdasarkan pengukuran data analisis sehingga hasil dari penelitian ini dapat diuji secara ilmiah dan terstruktur. Tahapan analisis data yang dilakukan secara teratur mulai dari pengolahan data, pemodelan, hingga penilaian hasil prediksi dengan menggunakan metrik kuantitatif seperti akurasi, *precision*, *recall*, dan *F1-score*. Metode ini memberikan kesempatan bagi peneliti untuk mendapatkan pemahaman yang jelas mengenai kepuasan pelanggan layanan transportasi online berdasarkan analisis kuantitatif dari kata ulasan pengguna yang menggunakan teknik *machine learning* (Jaha et al., 2020).

Dalam penelitian ini, data diperoleh melalui *platform Google Play Store*, yaitu toko aplikasi resmi untuk perangkat Android yang menawarkan ulasan dan penilaian dari pengguna aplikasi. Ulasan dan penilaian yang diberikan oleh pengguna di *Google Play Store* dapat berfungsi sebagai data empiris untuk melakukan *Sentiment Analysis* supaya memahami respon pengguna terhadap fitur, kualitas layanan, dan pengalaman pengguna secara lebih rinci, karena informasi tersebut berasal dari opini langsung pengguna yang tersedia untuk umum dan dapat diunduh untuk dianalisis (Anissa et al., 2025).

Pada tahap pengembangan, penelitian ini menggunakan metode *Cross Industry Standard Process for Data Mining* (CRISP-DM). CRISP-DM merupakan metode yang digunakan dalam penambangan data (data mining) yang dapat dilaksanakan secara lebih sistematis dan terstruktur (Siswoyo et al., 2025). Proses analisis dan pengolahan data dilakukan

dengan menerapkan metode *Cross Industry Standard Process for Data Mining* (CRISP-DM), dapat dilihat pada gambar berikut (Rossi *et al.*, 2024).



Gambar 1. Alur Tahapan Metode CRISP-DM

Tahapan metode *waterfall* dalam penelitian ini mencakup enam fase berurutan, yaitu pemahaman bisnis (Rossi *et al.*, 2024), pemahaman data, persiapan data, pemodelan (Rossi *et al.*, 2024), evaluasi, penerapan,

HASIL DAN PEMBAHASAN

Data Penelitian

Penelitian ini menggunakan data yang diambil dari salah satu *platform* digital, yakni aplikasi transportasi online Grab di *platform Google Play Store* yang bersifat internasional. Hal ini dikarenakan *Google Play Store* menampilkan komentar dari seluruh pengguna di berbagai negara, sehingga data yang dikumpulkan mencerminkan pandangan pengguna secara global. Dengan demikian, hasil analisis tidak hanya menunjukkan keadaan lokal, tetapi juga meliputi beragam pengalaman pengguna dari berbagai konteks budaya, bahasa, dan lokasi geografis. Data ulasan ini mengambil fitur ulasan dan rating dalam aplikasi, dimana opini dari pengguna ini menjadi salah satu petunjuk penting untuk menilai sejauh mana kepuasan pelanggan terhadap mutu layanan yang disediakan. Kepuasan layanan yang diberikan oleh pengguna merupakan salah satu faktor penting dalam proses suatu bisnis (Wastia *et al.*, 2024).

Secara keseluruhan, total data yang dikumpulkan dari proses *scraping* yaitu sebanyak 234.878 ulasan yang kemudian setelah dilakukan tahap *preprocessing* total keseluruhan datanya menjadi 183.079 ulasan. Hal ini menunjukkan bahwa kumpulan data yang diterapkan dalam penelitian ini tergolong besar, sehingga meningkatkan kemampuan model untuk mengenali pola sentimen pengguna. Selain itu, pemanfaatan data selama periode Januari 2020 hingga Februari 2026 memungkinkan penelitian ini untuk menangkap perubahan kepuasan pengguna dengan lebih mendalam, terutama selama masa pandemi dan setelahnya yang mempengaruhi cara orang menggunakan transportasi online.

Berikut ini tahap perolehan data dari aplikasi Grab yang digunakan sebagai data penelitian:



a. Scraping Data

Scraping data merupakan metode pengambilan data secara otomatis untuk memudahkan proses pengumpulan informasi dalam jumlah besar dari sistem digital (Adila *et al.*, 2022). *Scraping* data yang digunakan untuk mengambil data pada penelitian ini menggunakan bahasa pemrograman *Python* melalui *platform Google Colab* dengan mengimpor *library* yang dibutuhkan.

Kemudian mengambil total data secara keseluruhan, penelitian juga melakukan pemeriksaan distribusi jumlah data per tahun untuk memahami sebaran ulasan dari pengguna pada setiap periode penelitian. Pemeriksaan ini penting untuk mengerti bagaimana jumlah ulasan pengguna aplikasi Grab berubah seiring waktu, serta untuk memastikan bahwa data yang digunakan terwakili dengan baik di setiap tahun. Dari hasil klasifikasi data per tahun, diperoleh jumlah data awal dan setelah proses pembersihan data sebagai berikut:

Table 1. Detail Dataset Penelitian

Periode Tahun	Jumlah Data Awal	Jumlah Data Akhir
Jan 2020 - Des 2020	55.361	48.948
Jan 2021 - Des 2021	51.226	42.533
Jan 2022 - Des 2022	34.742	27.679
Jan 2023 - Des 2023	17.052	14.368
Jan 2024 - Des 2024	24.906	21.398
Jan 2025 - Des 2025	46.700	27.102
Jan 2026 - Feb 2026	4.891	1.051
Total	234.878	183.079

b. Penggabungan Data

Data *scraping* ulasan dari periode Januari 2020- Februari 2026 digabungkan menjadi satu dataset utama menggunakan *Microsoft Excel* agar dapat melihat kemampuannya untuk memeriksa data secara visual guna mencegah kesalahan dan pengulangan. Proses penggabungan dilakukan dengan memastikan bahwa struktur kolom di setiap file adalah identik, selanjutnya digabungkan dalam satu lembar kerja melalui metode salin-tempel atau menggunakan fitur *Power Query* (Dapatkan Data → Dari Folder).

c. Pembersihan Data

Hasil data yang telah digabungkan dalam satu file akan dilakukan tahapan pembersihan data seperti *case folding*, menghapus emoji, simbol, dan angka, *tokenizing* (pemecahan kata), *stopword removal*, dan yang terakhir *stemming*. Selanjutnya, Hasil dari fungsi ini selanjutnya diterapkan pada kolom content dalam dataset dan disimpan dalam kolom baru dengan nama *clean_text*. Dengan melakukan pengolahan awal ini, data menjadi lebih teratur, konsisten, dan relevan, sehingga dapat meningkatkan kualitas fitur yang digunakan dalam proses pelatihan dan evaluasi model *machine learning*.



d. Labeling Data

Hasil dari proses tersebut kemudian disatukan kembali menjadi satu string bersih. Setelah dilakukan proses pembersihan, data kemudian dilakukan proses *labeling* untuk memberikan tanda pada data ulasan sehingga dapat dimanfaatkan dalam *machine learning*. Kemudian, Proses pelabelan data menghasilkan dua label yaitu sentimen dan kategori, sehingga menciptakan struktur data yang multilabel. Penggunaan metode multilabel karena setiap ulasan yang dianalisis dapat memiliki beberapa label sekaligus. Dalam hal ini, suatu kalimat sering kali membawa lebih dari satu informasi, seperti perpaduan antara perasaan dan aspek pelayanan. Klasifikasi multilabel digunakan ketika data tidak bersifat saling eksklusif, sehingga satu contoh dapat tergolong dalam beberapa kategori secara bersamaan (Alfiana *et al.*, 2026). Oleh karena itu, pendekatan multilabel dipilih untuk lebih baik dalam menggambarkan kerumitan data dibandingkan dengan metode yang hanya menggunakan satu label.

Hasil Rancangan Model

a) Pembuatan Model

Dataset yang telah dihasilkan dari proses pembersihan dan proses labeling data kemudian dilakukan proses pemodelan data. Pada proses pemodelan, langkah awal yang dilakukan setelah proses labeling adalah membagi data tes latih dan data uji pada dataset. Data kemudian dibagi menjadi dua kelompok, yakni data latih untuk tahun 2020–2024 dan data uji untuk tahun 2025–2026. Metode ini diterapkan untuk meniru kondisi nyata, di mana model dilatih menggunakan data sejarah dan diuji pada data yang terkini.

b) Evaluasi Model

Evaluasi terhadap model dalam dilakukan untuk menilai seberapa baik model dapat memprediksi data multilabel, yang berarti setiap *instance* dapat memiliki lebih dari satu label. Pengujian dilakukan dengan menggunakan data uji dari tahun 2025 hingga 2026 untuk menilai kemampuan model dalam menggeneralisasi data baru setelah dilatih pada data dari tahun 2020 hingga 2024.

Evaluasi dilakukan dengan mengandalkan metrik seperti *precision*, *recall*, dan *F1-score* untuk setiap kelas label yang sudah dihasilkan oleh MultiLabelBinarizer, sehingga memungkinkan untuk menilai kinerja model di masing-masing kategori. Lalu menghitung subset accuracy, yaitu tingkat akurasi prediksi yang harus benar untuk semua label dalam satu data sekaligus. Selain itu, *hamming loss* juga diterapkan untuk menilai frekuensi kesalahan prediksi untuk setiap label. Penggunaan berbagai metrik ini bertujuan untuk memberikan pemahaman yang lebih menyeluruh tentang kinerja model dalam menangani klasifikasi multilabel yang rumit.

```

==== REPORT ====
              precision    recall  f1-score   support

 Driver          0.88      0.85      0.86      1889
 Lainnya         0.98      0.88      0.93      8886
 Netral          0.96      0.62      0.76      6097
 Puas           0.72      0.96      0.82      3847
 Sistem         0.51      0.97      0.67      1269
 Tidak Puas     0.50      0.96      0.65       945

 micro avg       0.82      0.83      0.82     22053
 macro avg       0.76      0.87      0.78     22053
 weighted avg    0.87      0.83      0.83     22053
 samples avg     0.84      0.83      0.83     22053

 Accuracy (Subset Accuracy): 0.687666452383139
 Micro F1: 0.8233624011356723
 Macro F1: 0.7816858527609148
 Hamming Loss: 0.1199834695564331

```

Gambar 2. Evaluasi Model



Hasil dari evaluasi model menunjukkan kinerja yang cukup memuaskan, namun terdapat ketidaksamaan antar kelas. Secara keseluruhan, kelas Driver, Lainnya, dan Puas menunjukkan nilai precision dan recall yang cukup tinggi, sementara kelas Sistem dan Tidak Puas memiliki recall yang baik tetapi nilai precision yang rendah, yang mengindikasikan adanya banyak false positive. Kelas Netral menunjukkan recall yang lebih rendah, sehingga sering kali tidak teridentifikasi dengan baik. Dari segi metrik keseluruhan, micro F1-score sebesar 0.82 menunjukkan kinerja secara keseluruhan yang baik, sedangkan macro F1-score yang mencapai 0.78 menandakan adanya ketidakseimbangan dalam kinerja antar kelas. Nilai subset accuracy yang mencapai 0.68 menunjukkan bahwa hanya sekitar 68% data yang prediksi seluruh labelnya benar dengan sempurna, sementara hamming loss yang mencapai 0.1199 menunjukkan sekitar 12% kesalahan dalam prediksi label. Secara keseluruhan, model ini sudah menunjukkan hasil yang baik dalam klasifikasi multilabel, tetapi masih perlu pengembangan, terutama dalam hal keseimbangan antar kelas dan pengurangan false positive.

c) Penyimpanan Model

Model *Naïve Bayes* yang telah dievaluasi akan disimpan untuk memungkinkan penggunaannya lagi di kemudian hari. Dalam proses ini, digunakan *library joblib*, yaitu sebuah *library Python* yang dibuat khusus untuk menyimpan atau memuat objek *Python*, termasuk model *machine learning*.

Kode ini digunakan untuk melakukan penyimpanan model *machine learning* dan komponen yang diperlukan agar dapat digunakan kembali tanpa perlu mengulangi proses pelatihan. Pertama, *library joblib* digunakan untuk menyimpan objek model ke dalam sebuah file, sementara `os` berfungsi untuk memverifikasi apakah folder penyimpanan tersedia. Selanjutnya, dibentuk folder baru dengan nama "models" menggunakan `os.makedirs()` jika folder tersebut belum ada sebelumnya. Setelah itu, model *Multinomial Naive Bayes* dengan pendekatan *One-vs-Rest* disimpan dalam file `model_naive_bayes.pkl`, sedangkan objek *TF-IDF Vectorizer* disimpan di file `tfidf_vectorizer.pkl`, yang dibutuhkan untuk mengubah teks baru menjadi format numerik saat melakukan prediksi. Di samping itu, *MultiLabelBinarizer* (mlb) juga disimpan dalam file `mlb.pkl` untuk memastikan proses penguraian label dari hasil prediksi tetap sejalan dengan proses pelatihan. Akhirnya, sistem menampilkan pesan "Model berhasil disimpan!" sebagai tanda bahwa semua komponen penting untuk pipeline prediksi telah berhasil diekspor dan siap digunakan kembali dalam tahap penerapan atau implementasi aplikasi.

Tabel 2. Hasil Pengujian Model

No.	Username	Content	Score	At	Sentimen	Kategori
1.	Cruz Jp	good	50.0	01/09/2022 11:40	Puas	['Netral', 'Lainnya']
2.	Moira Broughton	friendly and patient driver	50.0	01/09/2022 11:33	Puas	['Netral', 'Driver']
3.	Ummu HTHTH	nothing but the bestest of the bestðŸ	50.0	01/09/2022 11:01	Puas	['Netral', 'Lainnya']



4.	Trixie Maximillia	mantap lumayan banyak promonya tingkatin terus promonya thanks	50.0	01/09/2022 11:00	Puas	['Puas', 'Lainnya']
5.	Neneng Widasari	fitur chat aplikasi bermasalah sulit dibuka. selalu diarahkan ke sms atau chat app lain	20.0	01/09/2022 10:49	Puas	['Netral', 'Sistem']
6.	aweal halim	terbaik	50.0	01/09/2022 10:46	Puas	['Puas', 'Lainnya']
7.	nora Penang	saya suka naik grab sebab tak payah nak jalan jauh2 sampai jalan besar.ambik dalam flat saya duduk hantar depan tempat kerja	50.0	01/09/2022 10:26	Puas	['Tidak Puas', 'Lainnya']
8.	Mark angelo Binuya	your app sucks when it comes to food delivery i cant pin specific adress on the map. its not convinient at all. food panda is way better.	10.0	01/09/2022 09:54	Puas	['Netral', 'Lainnya']
9.	KC Yap	it remove my promo code when i click place order button. cancel order button disable within a second. at the end i need to pay without any discount.	10.0	01/09/2022 09:46	Puas	['Netral', 'Driver']
10.	Deva Malika Subhan	bagus	50.0	01/09/2022 09:30	Puas	['Puas', 'Lainnya']

Hasil dari pengujian model menunjukkan bahwa sistem ini dapat dengan efektif mengkategorikan ulasan pengguna ke dalam jenis sentimen dan jenis layanan. Ini menegaskan bahwa model yang dilatih tidak hanya berfungsi dalam teori, tetapi juga dapat diterapkan dalam



sistem yang nyata. Pengujian tersebut juga mengindikasikan bahwa pendekatan yang berbasis pada mesin pembelajaran bisa dimanfaatkan untuk menyederhanakan analisis ratusan ribu ulasan menjadi informasi yang lebih mudah dipahami. Dengan adanya sistem ini, baik pengguna maupun perusahaan dapat dengan cepat dan efisien mendapatkan gambaran tentang kepuasan pelanggan. Meskipun demikian, hasil tes ini juga mengindikasikan bahwa masih ada kemungkinan terjadinya kesalahan dalam klasifikasi, terutama pada ulasan yang bersifat ambigu atau memiliki lebih dari satu konteks (contohnya, ulasan yang memuji pengemudi tetapi mengeluhkan aplikasi). Ini menjadi salah satu keterbatasan dari model yang perlu diperhatikan.

Implementasi Sistem Berbasis Website

Model yang berhasil dibuat dan dinilai untuk klasifikasi sentimen dan kategori, langkah berikutnya dalam penelitian ini adalah menerapkan model tersebut ke dalam suatu sistem yang berbasis website. Tujuan dari sistem ini adalah untuk memudahkan pengguna dalam menganalisis sentimen dari ulasan aplikasi secara otomatis dan langsung. Website ini diciptakan dengan menggunakan *framework Streamlit*, yang memungkinkan integrasi langsung antara model machine learning dan antarmuka pengguna. Dengan memanfaatkan *Streamlit*, proses penyebaran model menjadi lebih mudah dan efisien, sehingga hasil dari penelitian ini tidak hanya bersifat teori, melainkan juga dapat diwujudkan dalam bentuk sistem yang nyata.

a) Fitur Sistem

Berdasarkan implementasi yang telah dilakukan, sistem memiliki beberapa fitur utama sebagai berikut:

1. Pemilihan Aplikasi dan Tahun

Sistem menyediakan pilihan beberapa aplikasi transportasi online seperti Grab, Gojek, Maxim, dan inDrive. Pengguna juga dapat memilih tahun analisis (2020–2026) untuk melihat data ulasan secara spesifik.

2. *Scraping* Data Otomatis

Dalam prosesnya, program mengumpulkan data secara bersamaan dalam kelompok yang berisi 1000 ulasan dengan menggunakan ``continuation_token``, kemudian mengubahnya menjadi DataFrame, serta menyaring dan menyimpan hanya ulasan yang relevan dengan tahun yang ditargetkan ke dalam daftar ``all_year_data``. Proses ini akan secara otomatis berhenti ketika sistem mendeteksi bahwa tanggal ulasan sudah lebih tua dari tahun yang diinginkan atau ketika sudah tidak ada data yang tersedia, lalu diakhiri dengan penggabungan semua data sekaligus menghapus ulasan yang sama pada kolom konten.

3. Prediksi Sentimen dan Kategori

Setelah pengumpulan data selesai, sistem melanjutkan dengan menganalisis informasi menggunakan model pembelajaran mesin yang telah tersedia (*Multinomial Naïve Bayes + TF-IDF*). Model ini berfungsi untuk menghasilkan dua elemen penting, yaitu Sentimen: puas, netral, tidak puas dan Kategori: pengemudi, sistem/aplikasi, lainnya

b) Proses Kerja Sistem

Proses kerja sistem ini dimulai dengan mengklik tombol ``run_btn``, yang kemudian menginstruksikan sistem untuk mengumpulkan data ulasan aplikasi berdasarkan tahun yang dipilih. Setelah data berhasil diambil, langkah selanjutnya adalah membersihkan data dengan menghapus entri yang duplikat berdasarkan isi dari ulasan untuk mencegah adanya data ganda

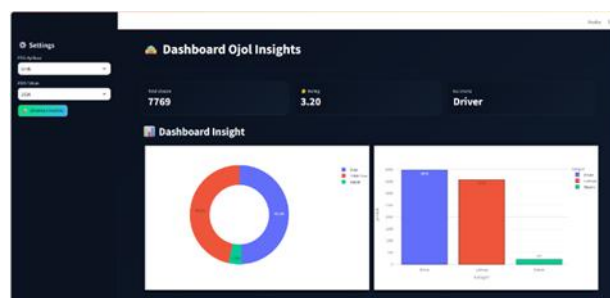


yang dapat mempengaruhi hasil analisis. Setelah itu, jika model, TF-IDF, dan label binarizer tersedia, teks ulasan akan diolah menjadi fitur numerik menggunakan TF-IDF, dan kemudian dilakukan prediksi dengan model *Multinomial Naive Bayes* yang berbasis *One-vs-Rest* untuk menghasilkan label multilabel. Hasil dari prediksi itu kemudian diubah kembali ke label aslinya lalu dipetakan ke dalam kategori sentimen (puas, netral, tidak puas) serta kategori masalah seperti Driver, Sistem, atau kategori lainnya.

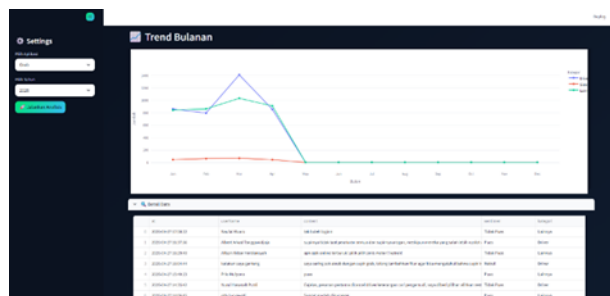
Setelah prediksi dilakukan, sistem akan menampilkan berbagai metrik utama seperti total ulasan, rata-rata rating, dan isu yang paling sering muncul. Data ini kemudian divisualisasikan dalam bentuk grafik lingkaran untuk distribusi sentimen, grafik batang untuk kategori masalah, serta grafik garis untuk menunjukkan tren bulanan selama 12 bulan yang telah distandarisasi agar semua bulan tetap terlihat meski tidak ada data. Selain itu, pengguna juga diberi kemudahan untuk melihat detail data dalam tabel interaktif dan dapat mengekspor data, baik yang utuh maupun yang telah disaring berdasarkan sentimen dan kategori tertentu. Dengan alur ini, sistem dapat memberikan wawasan komprehensif dari data ulasan secara otomatis, mulai dari pengambilan data hingga analisis dan visualisasi yang siap dipakai untuk mengevaluasi kualitas layanan aplikasi transportasi online.

c) Hasil Implementasi Sistem

Hasil penerapan menunjukkan bahwa sistem memiliki kemampuan untuk mengumpulkan data ulasan dalam skala besar secara otomatis, mengidentifikasi sentimen dan kategori dengan cepat, dan menampilkan hasil analisis dalam format visual yang mudah dimengerti. Sistem juga dapat menunjukkan metrik penting seperti Jumlah total ulasan, Rata-rata rating dari pengguna, Kategori isu yang paling sering muncul. Visualisasi yang berbentuk grafik pie dan bar chart memudahkan dalam memahami penyebaran sentimen serta isu utama yang kerap diperbincangkan oleh pengguna.



Gambar 3. Hasil Implementasi Sistem



Gambar 4. Hasil Implementasi Trend dan Data



d) Pengujian Sistem

Pengujian sistem adalah langkah penting dalam proses pengembangan perangkat lunak untuk memastikan bahwa sistem yang dibuat dapat memenuhi harapan pengguna dan berfungsi sesuai dengan spesifikasi yang telah ditetapkan. Salah satu teknik pengujian yang diterapkan dalam penelitian ini adalah metode BlackBox Testing. Metode BlackBox Testing merupakan salah satu metode yang bisa dibilang simpel untuk diterapkan, karena berfokus pada pengujian fungsi sistem tanpa memperhatikan struktur internal atau kode program, melainkan hanya berdasarkan input yang diberikan dan output yang dihasilkan (Huda *et al.*, 2022).

BlackBox Testing yang diterapkan dalam pengujian prediksi kepuasan pengguna aplikasi Grab berbasis website yang dikembangkan menggunakan *Streamlit* dan Algoritma *Naive Bayes*, dilakukan untuk memastikan bahwa seluruh fitur sistem berjalan sesuai dengan kebutuhan pengguna dan mampu menghasilkan output yang benar berdasarkan input yang diberikan. Dalam penelitian ini, fitur yang diuji meliputi:

- Pemilihan aplikasi transportasi online
- Pemilihan tahun data ulasan (2020-2026)
- Proses scraping data otomatis dari *Google Play Store*
- Proses *preprocessing* data
- Proses prediksi sentimen dan kategori
- Tampilan *dashboard* hasil analisis

Pengujian dilakukan dengan memberikan berbagai input pada sistem, kemudian mengamati apakah output yang dihasilkan sudah sesuai dengan yang diharapkan. Pendekatan ini sesuai dengan yang penelitian lain yang menyatakan bahwa *BlackBox Testing* digunakan untuk memastikan fungsi sistem berjalan dengan baik sebelum sistem diimplementasikan (Praniffa *et al.*, 2023). Tabel pengujian BlackBox Testing yang dilakukan pada penelitian ini dapat dilihat pada gambar 4.27 berikut:

Tabel 3. Tabel Hasil Pengujian *BlackBox Testing*

No.	Fitur yang Diuji	Skenario Pengujian	Input	Output yang Diharapkan	Hasil	Status
1.	Pemilihan Aplikasi	User memilih aplikasi	Grab, Gojek, Maxim, Indrive	Sistem menerima aplikasi	Berhasil menampilkan pilihan	Valid
2.	Pemilihan Tahun	User memilih tahun data	2020-2026	Sistem memfilter data sesuai tahun	Data sesuai tahun tampil	Valid
3.	Scraping Data	User klik tombol analisis	ID aplikasi	Sistem mengambil data ulasan otomatis	Data berhasil diambil	Valid
4.	Preprocessing Data	Sistem memproses teks	Data ulasan	Data dibersihkan	Data bersih berhasil diproses	Valid
5.	Prediksi Sentimen	Sistem melakukan klasifikasi	Teks Ulasan	Output sentimen (puas, netral, tidak puas)	Prediksi sesuai model	Valid



6.	Prediksi Kategori	Sistem klasifikasi kategori	Teks Ulasan	Output kategori (driver, sistem aplikasi, lainnya)	Kategori sesuai	Valid
7.	Tampilan Dashboard	Sistem menampilkan hasil	Data hasil analisis	Grafik & Tabel ditampilkan	Dashboard tampil dengan baik	Valid
8.	Error Handling	Input kosong atau salah	Tidak ada input	Sistem memberi notifikasi error	Notifikasi muncul	Valid

KESIMPULAN

Penelitian ini berhasil mengembangkan sistem prediksi untuk menganalisis tingkat kepuasan pengguna aplikasi transportasi online Grab dengan berdasarkan pada data ulasan di *Google Play Store* melalui metode *machine learning*. Latar belakang penelitian ini berfokus pada pentingnya pemahaman kepuasan pelanggan sebagai evaluasi terhadap kualitas layanan seiring dengan meningkatnya penggunaan transportasi online. *Algoritma Naïve Bayes* digunakan sebagai model prediksi karena kemampuannya yang baik dalam klasifikasi teks serta efisiensinya dalam mengelola data ulasan yang tidak terstruktur dalam jumlah besar. Oleh karena itu, penelitian ini menunjukkan bahwa penggunaan algoritma *Naïve Bayes* dalam sistem prediksi berbasis web dapat menghasilkan analisis yang efektif mengenai tingkat kepuasan pelanggan secara otomatis. Sistem yang dikembangkan mampu mengolah data dalam jumlah besar dengan cepat dan menyajikan hasil analisis dalam format yang mudah dimengerti. Dengan cara ini, penelitian ini dapat berfungsi sebagai alat bantu dalam pengambilan keputusan bagi penyedia layanan transportasi online dan juga sebagai dasar untuk penelitian yang lebih mendalam dan komprehensif di masa depan.

Saran

Terdapat beberapa rekomendasi yang dapat diajukan untuk pengembangan lebih lanjut dan perbaikan sistem yang sudah ada. Penelitian selanjutnya disarankan membandingkan algoritma lain seperti SVM, Random Forest, atau *Deep Learning* untuk meningkatkan akurasi model. Selain itu, penggunaan metode representasi fitur seperti Word2Vec, GloVe, atau BERT juga dapat dipertimbangkan. Sumber data sebaiknya diperluas tidak hanya dari *Google Play Store*, tetapi juga dari media sosial dan platform digital lainnya agar hasil analisis lebih representatif. Sistem juga perlu dikembangkan dengan fitur real-time, pembaruan data otomatis, serta visualisasi yang lebih interaktif agar lebih bermanfaat bagi masyarakat maupun perusahaan.

**DAFTAR PUSTAKA**

- Adila, N., Sembiring, F., & Jatmiko, W. (2022). Implementation of Web Scraping for Journal Data Collection on the SINTA Website. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, 7(4), 2478–2485. <https://doi.org/https://doi.org/10.33395/sinkron.v7i4.11576> e-ISSN
- Alfiana, H. N., Doewes, A., & Widoyono, B. (2026). Aspect-Based Sentiment Analysis of Access by KAI Application Reviews Using IndoBERT for Multi-Label Classification Tasks. *Jurnal Teknik Informatika (JUTIF)*, 7(1), 286–306. <https://doi.org/https://doi.org/10.52436/1.jutif.2026.7.2.5402> Aspect-Based
- Ananda, R. F., Syahri, A., Hasan, F. N., Technology, I., & Literature, A. (2024). Sentiment Analysis Of Customer Satisfaction In Gojek And Grab. *Jurnal Teknik Informatika (JUTIF)*, 5(1), 233–241. <https://doi.org/https://doi.org/10.52436/1.jutif.2024.5.1.1680>
- Anissa, C. R., Tania, K. D., Sari, W. K., Informasi, S., & Sriwijaya, U. (2025). Sentiment Analysis on Google Play Store Reviews to Measure User Perception of the Gojek Application Using CNN. *Journal of Applied Informatics and Computing (JAIC)*, 9(6), 3322–3328. <https://jurnal.polibatam.ac.id/index.php/JAIC/article/view/11084/3192>
- BPS. (2025). *Perkembangan Jumlah Kendaraan Bermotor Menurut Jenis (Unit), 2024*. Badan Statistik Indonesia. <https://www.bps.go.id/id/statistics-table/2/NTcjMg==/number-of-motor-vehicle-by-type.html>
- Heristian, S., Napiah, M., & Erawati, W. (2025). Analisis Sentimen Ulasan Pelanggan Menggunakan Algoritma Naive Bayes pada Aplikasi Gojek. *Computer Science (CO-SCIENCE)*, 5(1), 35–41. <https://doi.org/https://doi.org/10.31294/coscience.v5i1.7775>
- Huda, M. N., Burhan, M., Satibi, A., & Pradita, H. A. (2022). Implementasi Black Box Testing pada Aplikasi Sistem Kasir dengan Menggunakan Teknik Equivalence Partitions. *Jurnal Teknologi Sistem Informasi Dan Aplikasi*, 5(2), 120–127. <https://doi.org/10.32493/jtsi.v5i2.17645>
- Irawan, R. A., Marpaung, F. A., Saputra, I. I., & Widarti, D. W. (2025). Analysis of Online Transportation Customer Satisfaction Using C4 . 5 Algorithm. *JITEEHA: Journal of Information Technology Application in Education, Economy, Health and Agriculture*, 02(01), 2–7. <https://journal.iteeacademy.org/index.php/jiteeha/article/view/38>
- Jaha, A., Deepthi, N. S., Suryakanth, G., & Eswari, G. S. K. (2020). Text Sentiment Analysis using Naïve Bayes Classifier. *International Journal of Computer Trends and Technology*, 68(4), 261–265. <https://doi.org/https://doi.org/10.14445/22312803/IJCTT-V68I4P141>



- Praniffa, A. C., Syahri, A., Sandes, F., Fariha, U., & Giansyah, Q. A. (2023). *Jurnal Testing dan Implementasi Sistem Informasi PARKIR BERBASIS WEB BLACK BOX AND WHITE BOX TESTING OF WEB-BASED PARKING*. 1(1), 1–16.
- Rahayu, P., & Cahyono, R. P. (2022). *PENGELOMPOKAN DATA KEPUASAN KONSUMEN PADA APLIKASI TRANSPORTASI ONLINE MENGGUNAKAN ALGORITMA DECISION TREE*. 2(12), 1–10.
- Rahman, R., & Sutanto, F. A. (2023). Data Mining Untuk Memprediksi Tingkat Kepuasan Konsumen Gojek Menggunakan Algoritma Naive Bayes. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 18, 8–18. <https://doi.org/DOI : https://doi.org/10.35969/interkom.v18i1.280>
- Raya, K. K., Setia, P., & Firayanti, Y. (2025). Jurnal Sains Ekonomi dan Edukasi Pengaruh Motivasi Kerja , Disiplin Kerja , Pelatihan Kerja dan Komitmen Organisasi Terhadap Kinerja Driver Grab Bike di. *Jurnal Sains Ekonomi Dan Edukasi*, 2(8), 1990–2012. <https://doi.org/https://doi.org/10.62335>
- Rossi, A., Maranto, K., Damayanti, L., & Ramadika, I. R. (2024). Perbandingan Algoritma C4 . 5 dan Naïve Bayes dalam Prediksi Loyalitas Pelanggan. *Journal Binary Digital - Technology*, 7(2). <https://doi.org/10.32877/bt.v7i2.1825>
- Siswoyo, J., Fajarizka, D., Wira, S., Daud, I., Rumaropen, P., Martha, H., Ayorbaba, A. K., Prasetyo, A., Teknik, F., Informatika, T., & Papua, U. (2025). KLASIFIKASI TINGKAT KEPUASAN PELANGGAN FACEBOOK MARKETPLACE. *JOISIE (Journal Of Information Systems And Informatics Engineering)*, 9(1), 237–245. <https://doi.org/https://doi.org/10.35145/joisie.v9i1.5048>
- Sofwatillah, Risnita, Jailani, M. S., & Saksitha, D. A. (2024). TEHNIK ANALISIS DATA KUANTITATIF DAN KUALITATIF DALAM PENELITIAN ILMIAH. *Genta Mulia*, 15(2), 79–91.
- Wastia, A., Hadiwibowo, T., Nabilla, F. P., Yunizar, A., & Yusuf, P. (2024). Analisis Tingkat Kepuasan Pengguna Shopee Berdasarkan Rating Dan Ulasan Google Play Store Menggunakan Naïve Bayes. *Jurnal Riset Informatika Dan Teknologi Informasi (JRITI)*, 1(2), 43–47. <https://doi.org/https://doi.org/10.58776/jriti.v1i2.122>