



Analisis Sentimen Kepuasan Masyarakat terhadap Pelayanan Aplikasi Mobile JKN Menggunakan Algoritma Naïve Bayes

Muhammad Diaz R. Rambe^{1*}, Martiano²

^{1*,2}Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia
Email: ^{1*}diazrrambe148@gmail.com, ²martiano@umsu.ac.id

Abstrak

Program Jaminan Kesehatan Nasional (JKN) yang diselenggarakan oleh BPJS Kesehatan sedang memasuki masa transisi kebijakan Kelas Rawat Inap Standar (KRIS) berdasarkan Peraturan Presiden Nomor 59 Tahun 2024. Perubahan kebijakan tersebut memunculkan beragam tanggapan masyarakat yang tersebar pada ulasan aplikasi Mobile JKN di Google Play Store, namun tanggapan tersebut belum diolah menjadi informasi yang dapat dimanfaatkan sebagai bahan evaluasi layanan. Penelitian ini bertujuan mengklasifikasikan sentimen ulasan pengguna aplikasi Mobile JKN menggunakan algoritma Naïve Bayes serta mengimplementasikan hasil klasifikasi ke dalam sistem informasi berbasis web. Data dikumpulkan melalui teknik web scraping terhadap 4.000 ulasan berbahasa Indonesia pada periode 15 Juli sampai dengan 15 Agustus 2026 dan menghasilkan 2.675 ulasan hasil seleksi. Pelabelan dilakukan menggunakan pendekatan rating pengguna dan kamus sentimen InSet dengan verifikasi manual pada ulasan yang kedua pendekatannya berbeda, sehingga diperoleh 751 ulasan berlabel positif dan negatif. Data diproses melalui cleaning, case folding, tokenizing, stopword removal, dan stemming, kemudian dibobotkan menggunakan TF-IDF. Hasil pengujian menunjukkan model Naïve Bayes memperoleh accuracy 90,07%, precision 90,57%, recall 90,07%, dan F1-score 89,99%, sedangkan 10-fold cross validation memperoleh rata-rata akurasi 90,68% dengan simpangan baku 2,60%. Sebaran sentimen menunjukkan 403 ulasan (53,7%) bersentimen negatif dan 348 ulasan (46,3%) bersentimen positif, dengan keluhan pengguna terpusat pada proses autentikasi, khususnya pengiriman kode OTP dan verifikasi akun. Sistem informasi berbasis web yang dibangun berhasil melewati seluruh 12 skenario pengujian Black Box dan dapat digunakan secara berulang untuk menganalisis ulasan pengguna tanpa harus membacanya satu per satu.

Kata Kunci: analisis sentimen, Naïve Bayes, Mobile JKN, Kelas Rawat Inap Standar, TF-IDF, sistem informasi berbasis web.

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah mengubah cara masyarakat menyampaikan pendapat, kritik, maupun apresiasi terhadap suatu layanan. Apabila sebelumnya penyampaian penilaian dilakukan secara langsung melalui kotak saran atau survei kepuasan, saat ini masyarakat lebih banyak memanfaatkan kanal digital, salah satunya kolom ulasan pada toko aplikasi digital, karena bersifat terbuka, dapat diakses oleh siapa pun, serta tersimpan dalam jangka waktu yang panjang sehingga dapat dimanfaatkan sebagai bahan evaluasi kualitas pelayanan.

Salah satu bentuk pelayanan publik yang menyentuh paling banyak masyarakat Indonesia adalah program Jaminan Kesehatan Nasional (JKN) yang diselenggarakan oleh Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan. Hingga akhir tahun 2025, jumlah peserta program JKN tercatat mencapai 282,7 juta jiwa atau setara dengan 98,62 persen penduduk Indonesia, dengan pemanfaatan layanan mencapai 725,3 juta kali sepanjang tahun tersebut, atau sekitar 1,9 juta layanan setiap harinya (BPJS Kesehatan, 2026). Besarnya cakupan tersebut menjadikan kualitas pelayanan BPJS Kesehatan sebagai isu yang berdampak langsung terhadap hampir seluruh lapisan masyarakat. Seiring dengan transformasi digital pelayanan publik, BPJS Kesehatan menghadirkan aplikasi Mobile JKN sebagai kanal layanan utama berbasis telepon pintar, sehingga kualitas aplikasi ini turut menentukan persepsi masyarakat terhadap kualitas program JKN secara keseluruhan.

Pada saat yang bersamaan, program JKN sedang memasuki masa perubahan kebijakan yang signifikan. Berdasarkan Peraturan Presiden Nomor 59 Tahun 2024 (Republik Indonesia, 2024), pemerintah menghapus sistem kelas perawatan 1, 2, dan 3 dan menggantinya dengan Kelas Rawat Inap Standar (KRIS)

yang mulai diberlakukan pada Agustus 2026. Perubahan ini menimbulkan berbagai pertanyaan, kebingungan, maupun keluhan dari masyarakat yang sebagian besar disampaikan melalui kolom ulasan aplikasi Mobile JKN pada Google Play Store. Permasalahannya, ulasan tersebut berbentuk teks tidak terstruktur dengan jumlah yang sangat besar sehingga pembacaan dan penilaian secara manual menjadi tidak efektif, membutuhkan waktu lama, dan menyulitkan penyelenggara layanan dalam memetakan keluhan yang paling dominan secara objektif. Oleh karena itu, diperlukan suatu metode yang mampu mengolah opini pengguna secara otomatis sehingga menghasilkan informasi yang mudah dipahami dan dapat digunakan sebagai dasar pengambilan keputusan.

Beberapa penelitian terdahulu menunjukkan bahwa algoritma Naïve Bayes mampu memberikan performa yang baik dalam analisis sentimen pada berbagai jenis layanan. Rahmawati dkk. (2024) pada ulasan pengguna aplikasi m-BCA memperoleh akurasi sebesar 82%; Hidayat dkk. (2024) pada layanan Sistem Informasi Akademik Mahasiswa memperoleh akurasi 80,06%; Dzulkarnain dkk. (2024) pada penilaian masyarakat terhadap pelayanan rumah sakit di Kota Malang memperoleh akurasi 90%; sementara Handayanto dkk. (2021) membandingkan Naïve Bayes dengan Support Vector Machine pada ulasan Google Review dan melaporkan Naïve Bayes tetap kompetitif untuk data ulasan berskala kecil hingga menengah. Hartini dkk. (2024) juga menunjukkan performa yang baik pada ulasan pelanggan restoran. Di luar ranah analisis sentimen, penerapan algoritma pembelajaran mesin berbasis teks pengguna juga terus berkembang pada berbagai konteks data, misalnya Sari dan Hutagalung (2026) yang menerapkan algoritma Fuzzy Support Vector Machine untuk mengklasifikasikan preferensi konsumen berdasarkan kebutuhan pengguna pada platform e-commerce. Temuan-temuan tersebut memperkuat bukti bahwa metode klasifikasi berbasis teks, termasuk Naïve Bayes, dapat digunakan secara efektif untuk mengevaluasi persepsi dan kebutuhan pengguna pada layanan digital maupun layanan publik.

Penelitian mengenai analisis sentimen terhadap aplikasi Mobile JKN sebenarnya telah dilakukan oleh beberapa peneliti sebelumnya. Namun, seluruh penelitian tersebut menggunakan data ulasan yang dikumpulkan sebelum kebijakan Kelas Rawat Inap Standar diberlakukan, sehingga belum menggambarkan persepsi masyarakat pada masa transisi kebijakan yang sedang berlangsung. Selain itu, sebagian besar penelitian terdahulu berhenti pada penyajian nilai akurasi model dan belum mengimplementasikan hasil klasifikasi ke dalam sebuah sistem informasi yang dapat digunakan secara berulang oleh penyelenggara layanan. Kedua hal tersebut menjadi kebaruan ilmiah yang diisi melalui penelitian ini, yaitu analisis sentimen pada periode transisi kebijakan KRIS yang diimplementasikan sekaligus ke dalam sistem informasi berbasis web yang dapat dipakai ulang.

Berdasarkan uraian tersebut, penelitian ini bertujuan (1) menerapkan algoritma Naïve Bayes untuk mengklasifikasikan sentimen ulasan pengguna aplikasi Mobile JKN pada Google Play Store ke dalam kategori positif dan negatif, (2) mengukur performa algoritma tersebut berdasarkan nilai accuracy, precision, recall, dan F1-score, serta (3) mengimplementasikan hasil klasifikasi ke dalam sistem informasi berbasis web sebagai sarana evaluasi berkelanjutan terhadap kualitas pelayanan aplikasi Mobile JKN.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen untuk mengetahui kecenderungan opini pengguna terhadap pelayanan aplikasi Mobile JKN berdasarkan data ulasan yang diperoleh dari Google Play Store. Objek penelitian adalah aplikasi Mobile JKN (app.bpjs.mobile) dengan data berupa ulasan berbahasa Indonesia pada periode 15 Juli sampai dengan 15 Agustus 2026, yaitu periode yang mencakup masa pemberlakuan kebijakan Kelas Rawat Inap Standar. Sumber data hanya berasal dari Google Play Store dan tidak mencakup Apple App Store, media sosial, maupun kanal pengaduan resmi BPJS Kesehatan.

Pengumpulan data dilakukan dengan teknik web scraping menggunakan bahasa pemrograman Python dan pustaka google-play-scraper, yang menghasilkan 4.000 ulasan mentah. Data tersebut diseleksi dengan menghapus ulasan duplikat, ulasan yang terdiri atas kurang dari tiga kata, dan ulasan yang tidak berbahasa Indonesia, sehingga diperoleh dataset hasil seleksi sebanyak 2.675 ulasan. Mengingat sebaran rating yang terpolarisasi pada nilai ekstrem, pengambilan sampel untuk pelabelan dilakukan menggunakan metode sampel berimbang (seed = 42) sebanyak 922 ulasan, terdiri atas 400 ulasan dari kelompok rating 1–2 bintang, 122 ulasan dari kelompok rating 3 bintang, dan 400 ulasan dari kelompok rating 4–5 bintang.

Pelabelan sentimen dilakukan menggunakan dua pendekatan yang saling melengkapi, yaitu pendekatan berbasis rating pengguna dan pendekatan lexicon based menggunakan kamus sentimen InSet untuk bahasa Indonesia (Koto & Rahmaningtyas, 2017). Ulasan yang kedua pendekatan menghasilkan label sama dilabeli secara otomatis, sedangkan ulasan yang menghasilkan label berbeda diverifikasi secara manual oleh peneliti. Tingkat kesepakatan kedua pendekatan diukur menggunakan Cohen's Kappa. Ulasan yang teridentifikasi netral atau berupa pertanyaan informatif dikeluarkan dari dataset setelah tahap pelabelan, sehingga dihasilkan dataset akhir sebanyak 751 ulasan berlabel positif dan negatif yang digunakan sebagai ground truth.

Data selanjutnya diproses melalui lima tahap text preprocessing, yaitu cleaning (menghapus tanda baca, angka, emoji, dan URL), case folding (mengubah huruf menjadi huruf kecil), tokenizing (memecah kalimat menjadi token), stopword removal, dan stemming menggunakan pustaka Sastrawi untuk bahasa Indonesia (Librian, 2016). Kata pengingkar seperti “tidak” secara sengaja dipertahankan pada tahap stopword removal karena kata tersebut menentukan polaritas kalimat. Data hasil preprocessing kemudian dibobotkan menggunakan TF-IDF dengan kombinasi unigram dan bigram, menghasilkan 4.004 fitur kata. Proses pembobotan hanya dilatih pada data latih untuk mencegah kebocoran data (data leakage) ke data uji.

Dataset dibagi menjadi data latih dan data uji dengan perbandingan 80:20 menggunakan teknik stratified split agar proporsi kelas tetap seimbang, menghasilkan 600 data latih dan 151 data uji. Model klasifikasi dibangun menggunakan algoritma Multinomial Naïve Bayes dengan nilai alpha sebesar 1,0 (Laplace smoothing) melalui pustaka scikit-learn (Pedregosa dkk., 2011). Evaluasi performa model dilakukan menggunakan Confusion Matrix dengan menghitung nilai accuracy, precision, recall, dan F1-score, baik secara weighted average maupun macro average. Untuk memastikan kestabilan model terhadap pembagian data, dilakukan pula pengujian 10-fold stratified cross validation.

Hasil klasifikasi sentimen selanjutnya diimplementasikan ke dalam sistem informasi berbasis web yang dirancang menggunakan Unified Modeling Language, meliputi Use Case Diagram, Activity Diagram, Sequence Diagram, dan Class Diagram (Rosa & Shalahuddin, 2019), dengan satu aktor pengguna (Admin). Sistem dibangun menggunakan bahasa pemrograman Python dengan kerangka kerja Flask dan basis data SQLite, memuat halaman login, dashboard, dataset, analisis sentimen, dan hasil analisis. Pengujian fungsional sistem dilakukan menggunakan metode Black Box Testing terhadap 12 skenario yang mencakup fungsi login, pengelolaan dataset, proses analisis sentimen, penyajian hasil, serta logout.

HASIL DAN PEMBAHASAN

Hasil Seleksi dan Pelabelan Dataset

Pengumpulan data menghasilkan 4.000 ulasan mentah. Setelah melalui proses seleksi sebagaimana diuraikan pada bagian Metode, diperoleh dataset hasil seleksi sebanyak 2.675 ulasan. Rincian proses seleksi disajikan pada Tabel 1.

Tabel 1. Hasil Seleksi Dataset

Tahap Seleksi	Jumlah Dibuang	Sisa Data
Data mentah hasil scraping	–	4.000
Ulasan duplikat	874	3.126
Ulasan kurang dari tiga kata	443	2.683
Ulasan bukan berbahasa Indonesia	8	2.675
Dataset hasil seleksi	–	2.675

Dari 922 ulasan yang diambil melalui sampel berimbang, sebanyak 622 ulasan (67,5%) memperoleh label yang sama dari pendekatan rating dan pendekatan lexicon InSet, sedangkan 300 ulasan (32,5%) memperoleh label berbeda sehingga diverifikasi secara manual. Tingkat kesepakatan kedua pendekatan menurut Cohen's Kappa adalah 0,438 (kategori sedang), yang menegaskan pentingnya verifikasi manual pada ulasan yang tidak disepakati. Hasil pelabelan akhir disajikan pada Tabel 2.

Tabel 2. Hasil Pelabelan Sentimen

Label	Jumlah	Persentase
Negatif	403	44,1%
Positif	350	38,3%
Netral	161	17,6%
Tidak dilabeli	8	0,9%
Total	922	100%

Ulasan berlabel netral umumnya berupa pertanyaan atau pernyataan informatif yang tidak menyatakan kepuasan maupun ketidakpuasan, misalnya pertanyaan mengenai cara mengubah fasilitas kesehatan. Sesuai ruang lingkup penelitian yang berfokus pada polaritas kepuasan, ulasan netral dan ulasan yang tidak dilabeli dikeluarkan dari dataset, sehingga dataset akhir yang digunakan pada tahap klasifikasi berjumlah 751 ulasan, terdiri atas 403 ulasan negatif dan 348 ulasan positif.

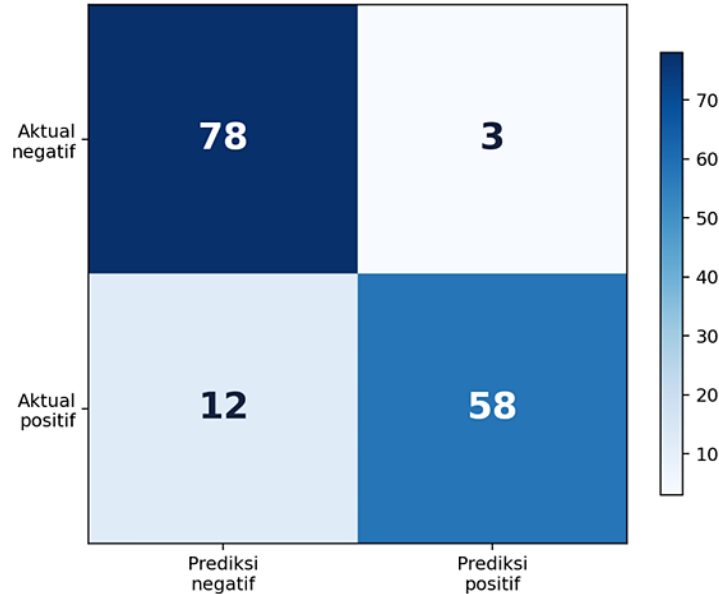
Hasil Klasifikasi dan Evaluasi Model

Model Multinomial Naïve Bayes dilatih menggunakan 600 data latih dan diuji terhadap 151 data uji. Hasil evaluasi menggunakan Confusion Matrix disajikan pada Tabel 3.

Tabel 3. Confusion Matrix Hasil Klasifikasi

Aktual / Prediksi	Prediksi Negatif	Prediksi Positif
Aktual Negatif	78	3
Aktual Positif	12	58

Confusion Matrix Hasil Klasifikasi Naive Bayes



Gambar 1. Visualisasi Confusion Matrix Hasil Klasifikasi

Berdasarkan Tabel 3, sebanyak 78 ulasan negatif dan 58 ulasan positif berhasil diprediksi dengan benar, sedangkan 3 ulasan negatif salah diprediksi sebagai positif dan 12 ulasan positif salah diprediksi sebagai negatif, sehingga total prediksi benar adalah 136 dari 151 data uji. Nilai precision, recall, dan F1-score per kelas disajikan pada Tabel 4, sedangkan nilai evaluasi keseluruhan model disajikan pada Tabel 5.

Tabel 4. Perhitungan TP, TN, FP, dan FN per Kelas

Kelas	TP	TN	FP	FN	Precision	Recall	F1-Score
Negatif	78	58	12	3	86,67%	96,30%	91,23%
Positif	58	78	3	12	95,08%	82,86%	88,55%

Tabel 5. Nilai Evaluasi Model

Metrik	Weighted Average	Macro Average
Accuracy	90,07%	90,07%
Precision	90,57%	90,87%
Recall	90,07%	89,58%
F1-Score	89,99%	89,89%

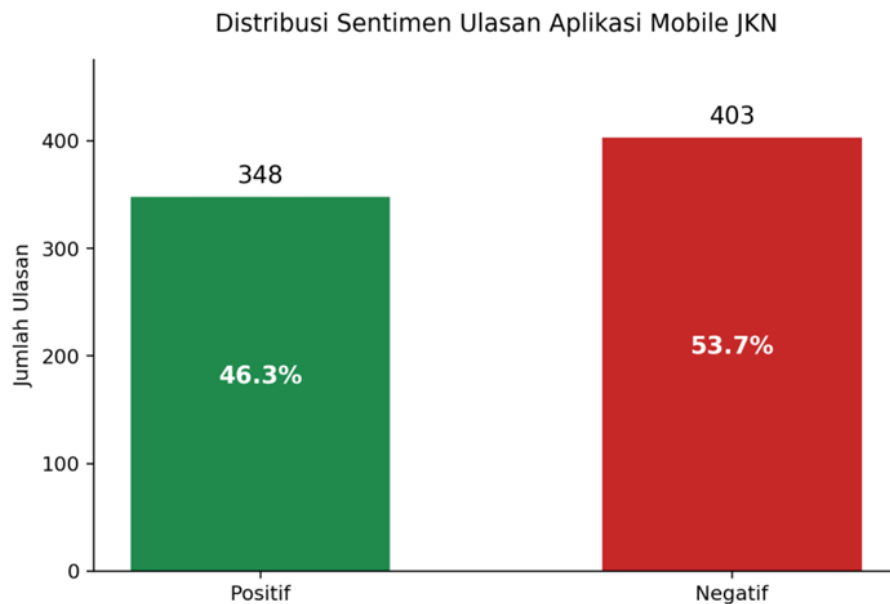
Model memperoleh accuracy 90,07%, precision 90,57%, recall 90,07%, dan F1-score 89,99%. Nilai macro average yang hampir sama dengan weighted average menunjukkan bahwa performa model relatif merata pada kedua kelas, tidak hanya baik pada kelas mayoritas. Untuk memastikan hasil tersebut tidak dipengaruhi oleh pembagian data tertentu, dilakukan pengujian 10-fold cross validation yang hasilnya disajikan pada Tabel 6.

Tabel 6. Hasil 10-Fold Cross Validation

Fold	Akurasi	Fold	Akurasi
1	92,11%	6	88,00%
2	92,00%	7	90,67%
3	86,67%	8	94,67%
4	86,67%	9	93,33%
5	92,00%	10	90,67%

Hasil cross validation memperoleh rata-rata akurasi 90,68% dengan simpangan baku 2,60%. Nilai simpangan baku yang kecil menunjukkan bahwa performa model stabil dan tidak bergantung pada pembagian data tertentu, serta konsisten dengan hasil pengujian pada data uji sehingga memperkuat

keandalan model yang dibangun. Distribusi hasil klasifikasi sentimen terhadap keseluruhan dataset (751 ulasan) menunjukkan 403 ulasan (53,7%) tergolong sentimen negatif dan 348 ulasan (46,3%) tergolong sentimen positif, sebagaimana diilustrasikan pada Gambar 2. Hasil ini menunjukkan bahwa mayoritas opini pengguna terhadap pelayanan aplikasi Mobile JKN pada periode penelitian bersentimen negatif, meskipun selisihnya dengan sentimen positif tidak terlalu besar.



Gambar 2. Distribusi Sentimen Hasil Klasifikasi

Kata Paling Berpengaruh

Analisis terhadap kata paling berpengaruh pada setiap kelas dilakukan untuk memastikan bahwa model benar-benar mempelajari kata yang relevan, bukan sekadar menebak. Hasilnya disajikan pada Tabel 7 di bawah ini.

Tabel 7. Kata Paling Berpengaruh pada Setiap Kelas

Peringkat	Kelas Negatif	Kelas Positif
1	tidak	mudah
2	aplikasi	bantu
3	sulit	bagus
4	daftar	cepat
5	masuk	layan
6	kode	jkn
7	otp	mudah cepat
8	verifikasi	terima kasih
9	login	sehat
10	kode otp	terimakasih

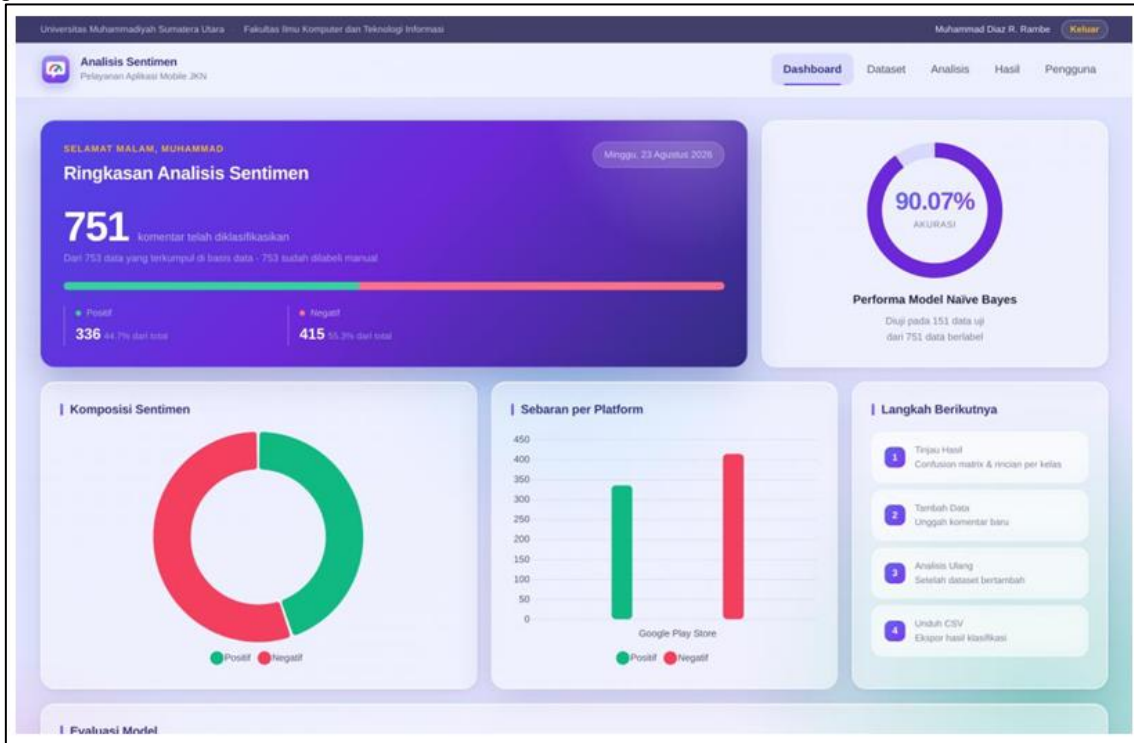
Kata yang dominan pada kelas negatif berkaitan dengan kendala teknis proses masuk dan verifikasi, yaitu sulit, daftar, masuk, kode, otp, verifikasi, dan login. Sebaliknya, kata dominan pada kelas positif berkaitan dengan kemudahan dan manfaat layanan, yaitu mudah, bantu, bagus, cepat, dan layan. Temuan ini menunjukkan bahwa keluhan pengguna tidak tersebar merata pada seluruh fitur, melainkan terpusat pada proses autentikasi, khususnya pengiriman kode OTP dan proses verifikasi akun, sehingga perbaikan pada bagian tersebut berpotensi menurunkan proporsi sentimen negatif secara signifikan.

Analisis Kesalahan Klasifikasi

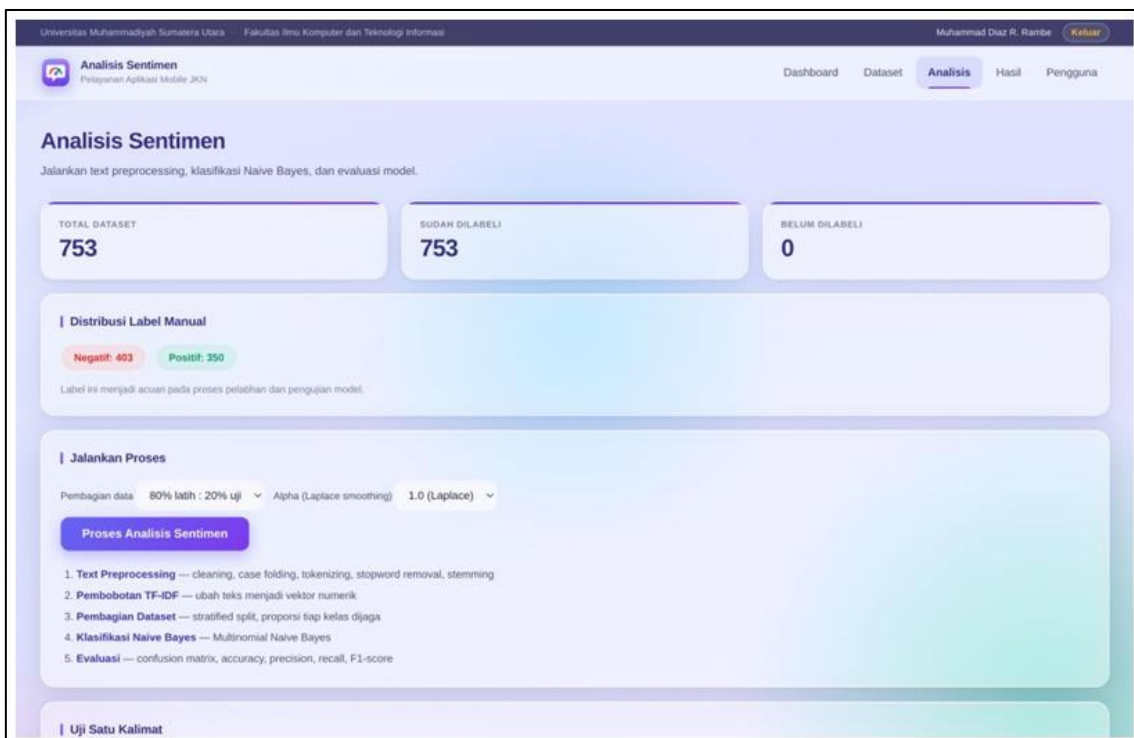
Dari 151 data uji, terdapat 15 ulasan yang salah diklasifikasikan, terdiri atas 12 ulasan positif yang diprediksi sebagai negatif dan 3 ulasan negatif yang diprediksi sebagai positif. Pemeriksaan terhadap kesalahan tersebut menunjukkan dua pola utama. Pola pertama adalah ulasan yang mengandung kata bermuatan negatif meskipun kalimat secara keseluruhan menyampaikan hal positif atau sekadar pertanyaan, misalnya ulasan yang menanyakan mengapa aplikasi tidak dapat dibuka, sehingga model mengenali kata “tidak” dan “dibuka” sebagai penanda negatif dan mengabaikan konteks kalimat secara keseluruhan. Pola kedua adalah keterbatasan asumsi independensi antar-kata pada algoritma Naïve Bayes, sehingga kalimat yang mengandung campuran pujian dan keluhan cenderung diklasifikasikan berdasarkan kata yang paling dominan bobotnya.

Implementasi dan Pengujian Sistem

Hasil analisis sentimen diimplementasikan ke dalam sistem informasi berbasis web menggunakan Flask dan basis data SQLite, dengan satu aktor pengguna (Admin) yang dapat mengelola dataset, menjalankan proses analisis sentimen, dan melihat hasil analisis berupa tabel klasifikasi, grafik distribusi sentimen, Confusion Matrix, dan nilai evaluasi model. Kata sandi pengguna disimpan dalam bentuk hash sehingga tidak dapat dibaca langsung dari basis data. Tampilan beberapa halaman utama sistem ditunjukkan pada Gambar 3 dan Gambar 4.



Gambar 3. Tampilan Halaman Dashboard Sistem Analisis Sentimen



Gambar 4. Tampilan Halaman Hasil Analisis Sentimen pada Sistem

Nilai akurasi yang ditampilkan sistem, yaitu 90,07%, konsisten dengan hasil perhitungan evaluasi model. Jumlah sentimen yang ditampilkan sistem terhadap seluruh 751 ulasan, yaitu 336 ulasan positif dan

415 ulasan negatif, berbeda tipis dari sebaran hasil pelabelan (348 positif dan 403 negatif) karena sebagian kecil ulasan diprediksi berbeda dari labelnya, selisih yang sejalan dengan nilai accuracy model sebesar 90,07%. Pengujian fungsional sistem menggunakan metode Black Box Testing terhadap 12 skenario menunjukkan bahwa seluruh fungsi, meliputi login, pengelolaan dataset, proses analisis sentimen, penyajian hasil analisis, pengelolaan kata sandi, dan logout, berjalan sesuai dengan hasil yang diharapkan sebagaimana dirangkum pada Tabel 8.

Tabel 8. Hasil Pengujian Black Box

No	Fungsi	Skenario Pengujian	Hasil
1	Login	Username dan password benar	Sesuai
2	Login	Password salah	Sesuai
3	Login	Akses halaman tanpa login	Sesuai
4	Dataset	Unggah berkas CSV sesuai format	Sesuai
5	Dataset	Unggah berkas selain CSV	Sesuai
6	Dataset	Menampilkan daftar ulasan	Sesuai
7	Analisis	Menjalankan proses analisis sentimen	Sesuai
8	Hasil	Menampilkan nilai evaluasi model	Sesuai
9	Hasil	Menampilkan Confusion Matrix	Sesuai
10	Hasil	Mengunduh hasil klasifikasi	Sesuai
11	Pengguna	Mengubah kata sandi	Sesuai
12	Logout	Menekan tombol keluar	Sesuai

Pembahasan

Nilai akurasi yang diperoleh pada penelitian ini, yaitu 90,07%, berada pada kisaran yang sebanding dengan penelitian terdahulu mengenai penerapan Naïve Bayes pada analisis sentimen layanan digital. Rahmawati dkk. (2024) pada aplikasi m-BCA memperoleh akurasi 82%, Hidayat dkk. (2024) pada layanan Sistem Informasi Akademik memperoleh 80,06%, sedangkan Dzulkarnain dkk. (2024) pada pelayanan rumah sakit memperoleh 90%. Perbandingan ini sejalan dengan temuan Handayanto dkk. (2021) dan Hartini dkk. (2024) yang juga melaporkan Naïve Bayes sebagai metode yang kompetitif untuk klasifikasi ulasan berskala menengah. Perbedaan utama penelitian ini terletak pada objek dan periode data: seluruh penelitian terdahulu mengenai sentimen aplikasi Mobile JKN menggunakan data yang dikumpulkan sebelum kebijakan Kelas Rawat Inap Standar diberlakukan, sedangkan penelitian ini secara khusus mengambil data pada masa transisi kebijakan tersebut. Selain itu, berbeda dari sebagian besar penelitian terdahulu yang berhenti pada penyajian nilai akurasi, penelitian ini mengimplementasikan hasil klasifikasi ke dalam sistem informasi yang dapat digunakan secara berulang, sejalan dengan kecenderungan penerapan algoritma klasifikasi berbasis teks pada beragam konteks data pengguna seperti yang ditunjukkan oleh Sari dan Hutagalung (2026) pada data preferensi konsumen e-commerce.

Berdasarkan rumusan masalah penelitian, dapat disimpulkan bahwa algoritma Naïve Bayes berhasil diterapkan untuk mengklasifikasikan 751 ulasan pengguna aplikasi Mobile JKN ke dalam kategori positif dan negatif melalui tahapan text preprocessing, pembobotan TF-IDF, dan klasifikasi Multinomial Naïve Bayes, dengan performa accuracy 90,07%, precision 90,57%, recall 90,07%, dan F1-score 89,99%, yang didukung oleh hasil cross validation yang stabil. Hasil klasifikasi tersebut berhasil diimplementasikan ke dalam sistem informasi berbasis web yang seluruh fungsinya berjalan sesuai pengujian Black Box, sehingga dapat digunakan sebagai sarana evaluasi berkelanjutan terhadap kualitas pelayanan aplikasi Mobile JKN, khususnya pada aspek autentikasi dan verifikasi akun yang menjadi sumber keluhan dominan.

KESIMPULAN

Algoritma Naïve Bayes berhasil diterapkan untuk mengklasifikasikan sentimen ulasan pengguna aplikasi Mobile JKN pada masa transisi kebijakan Kelas Rawat Inap Standar, dengan 751 dari 2.675 ulasan hasil seleksi digunakan pada tahap analisis setelah melalui proses pelabelan, menghasilkan 403 ulasan (53,7%) bersentimen negatif dan 348 ulasan (46,3%) bersentimen positif. Model yang dibangun memperoleh accuracy 90,07%, precision 90,57%, recall 90,07%, dan F1-score 89,99%, dengan rata-rata akurasi 10-fold cross validation 90,68% dan simpangan baku 2,60%, yang menunjukkan performa model yang stabil dan tidak bergantung pada pembagian data tertentu. Hasil klasifikasi tersebut berhasil diimplementasikan ke dalam sistem informasi berbasis web yang seluruh 12 skenario pengujian Black Box-nya memberikan hasil sesuai, sehingga sistem dapat digunakan secara berulang untuk menganalisis ulasan pengguna tanpa harus membacanya satu per satu. Analisis kata paling berpengaruh menunjukkan bahwa keluhan pengguna terpusat pada proses autentikasi, khususnya pengiriman kode OTP dan verifikasi akun, sedangkan apresiasi pengguna berkaitan dengan kemudahan dan kecepatan layanan.

Bagi BPJS Kesehatan, perbaikan pada proses masuk, pengiriman kode OTP, dan verifikasi akun disarankan menjadi prioritas karena berpotensi menurunkan proporsi sentimen negatif secara signifikan.

Bagi peneliti selanjutnya, disarankan menambah jumlah data berlabel dan memperluas periode pengambilan data agar dapat membandingkan sentimen sebelum dan sesudah pemberlakuan kebijakan Kelas Rawat Inap Standar secara lebih menyeluruh, mencoba algoritma klasifikasi lain seperti Support Vector Machine atau pendekatan berbasis deep learning untuk pembandingan, serta mengembangkan analisis berbasis aspek agar dapat mengetahui aspek layanan tertentu, seperti autentikasi, antrean, atau pembayaran iuran, yang paling banyak dikeluhkan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara atas dukungan fasilitas dan bimbingan akademik selama penelitian ini dilaksanakan, serta kepada seluruh pihak yang telah membantu proses pengumpulan dan pengolahan data pada penelitian ini.

DAFTAR PUSTAKA

- BPJS Kesehatan. (2026). *Capaian Program Jaminan Kesehatan Nasional Tahun 2025: Peserta JKN Mencapai 282,7 Juta Jiwa*. Jakarta: Badan Penyelenggara Jaminan Sosial Kesehatan.
- Dzulkarnain, T., Ratnawati, D. E., & Rahayudi, B. (2024). Penggunaan metode Naïve Bayes Classifier pada analisis sentimen penilaian masyarakat terhadap pelayanan rumah sakit di Malang. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(5), 993–1000. <https://doi.org/10.25126/jtiik.2024117979>
- Handayanto, R. T., Herlawati, H., Atika, P. D., Khasanah, F. N., Yusuf, A. Y. P., & Septia, D. Y. (2021). Analisis sentimen pada situs Google Review dengan Naïve Bayes dan Support Vector Machine. *Jurnal Komtika (Komputasi dan Informatika)*, 5(2), 153–163. <https://doi.org/10.31603/komtika.v5i2.6280>
- Hartini, P., Suarna, N., & Prihartono, W. (2024). Naive Bayes Classifier untuk analisis sentimen ulasan pelanggan pada Domo Coffee and Resto. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 6(1), 271–280. <https://doi.org/10.36499/jinrpl.v6i1.10315>
- Hidayat, T., Cahyana, R., & Julianto, I. T. (2024). Analisis sentimen layanan sistem informasi akademik mahasiswa menggunakan algoritma Naive Bayes. *Jurnal Algoritma*, 21(1), 119–130. <https://doi.org/10.33364/algoritma/v.21-1.1514>
- Koto, F., & Rahmaningtyas, G. Y. (2017). InSet lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs. In *Proceedings of the 2017 International Conference on Asian Language Processing (IALP)* (pp. 391–394). IEEE. <https://doi.org/10.1109/IALP.2017.8300625>
- Librian, A. (2016). *Sastrawi: High quality stemmer library for Indonesian language*. Diakses dari <https://github.com/sastrawi/sastrawi>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Rahmawati, I., Fitriani, T. R., & Yusuf, A. Y. P. (2024). Analisis sentimen menggunakan algoritma Naïve Bayes pada aplikasi m-BCA berdasarkan ulasan pengguna di Google Play Store. *Jurnal Riset Informatika dan Teknologi Informasi*, 1(2), 38–42. <https://doi.org/10.58776/jriti.v1i2.116>
- Republik Indonesia. (2024). *Peraturan Presiden Republik Indonesia Nomor 59 Tahun 2024 tentang Perubahan Ketiga atas Peraturan Presiden Nomor 82 Tahun 2018 tentang Jaminan Kesehatan*. Jakarta: Sekretariat Negara.
- Rosa, A. S., & Shalahuddin, M. (2019). *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek* (Edisi revisi). Bandung: Informatika Bandung.
- Sari, M., & Hutagalung, F. S. (2026). Analisis preferensi konsumen dalam pembelian laptop berdasarkan kebutuhan pengguna di Tokopedia menggunakan algoritma Fuzzy SVM (Support Vector Machine). *Indonesian Journal of Multidisciplinary Scientific Studies (IJOMSS)*, 4(4), 157–160. <https://doi.org/10.33151/ijomss.v4i4.917>