

**ANALISIS SENTIMEN KEPUASAN MASYARAKAT
TERHADAP PELAYANAN APLIKASI MOBILE JKN
MENGUNAKAN ALGORITMA NAÏVE BAYES**

SKRIPSI

DISUSUN OLEH

MUHAMMAD DIAZ R. RAMBE

NPM : 2209010031



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

MEDAN

2026

**ANALISIS SENTIMEN KEPUASAN MASYARAKAT
TERHADAP PELAYANAN APLIKASI MOBILE JKN
MENGUNAKAN ALGORITMA NAÏVE BAYES**

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
(S.Kom) pada Program Studi Sistem Informasi Fakultas Ilmu Komputer dan
Teknologi Informasi Universitas Muhammadiyah Sumatera Utara

DISUSUN OLEH

MUHAMMAD DIAZ R. RAMBE
NPM : 2209010031

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2026**

LEMBAR PENGESAHAN

Judul Skripsi : ANALISIS SENTIMEN KEPUASAN
MASYARAKAT TERHADAP PELAYANAN
APLIKASI MOBILE JKN MENGGUNAKAN
ALGORITMA NAÏVE BAYES
Nama Mahasiswa : MUHAMMAD DIAZ R. RAMBE
NPM : 2209010031
Program Studi : SISTEM INFORMASI

Menyetujui
Komisi Pembimbing



(Martiano, S.Pd., S.Kom., M.Kom.)
NIDN. 0128029302

Ketua Program Studi



(Mahardika Abdi Prawira Tanjung,
M.Kom.)
NIDN. 0117019301

Dekan



(Dr. A. R. Khowarizmi, S.Kom.,
M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

Saya yang bertanda tangan di bawah ini:

Nama : MUHAMMAD DIAZ R. RAMBE

NPM : 2209010031

Program Studi : Sistem Informasi

Fakultas : Ilmu Komputer dan Teknologi Informasi

dengan ini menyatakan dengan sesungguhnya bahwa skripsi berjudul
**“ANALISIS SENTIMEN KEPUASAN MASYARAKAT TERHADAP
PELAYANAN APLIKASI MOBILE JKN MENGGUNAKAN ALGORITMA
NAÏVE BAYES”** beserta seluruh isinya adalah benar-benar karya saya sendiri.
Saya tidak melakukan penjiplakan atau pengutipan dengan cara yang tidak sesuai
dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung risiko atau sanksi yang dijatuhkan
kepada saya apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika
keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian
karya saya ini.

Medan, 13 Oktober 2026
Yang membuat pernyataan,

Materai
Rp10.000

01079AOX188820382
MUHAMMAD DIAZ R. RAMBE
NPM : 2209010031

PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas akademika Universitas Muhammadiyah Sumatera Utara, saya yang bertanda tangan di bawah ini:

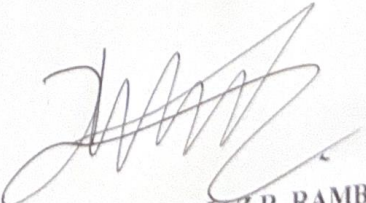
Nama : MUHAMMAD DIAZ R. RAMBE
NPM : 2209010031
Program Studi : Sistem Informasi
Fakultas : Ilmu Komputer dan Teknologi Informasi
Jenis Karya : Skripsi

demikian pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty Free Right) atas karya ilmiah saya yang berjudul **"ANALISIS SENTIMEN KEPUASAN MASYARAKAT TERHADAP PELAYANAN APLIKASI MOBILE JKN MENGGUNAKAN ALGORITMA NAÏVE BAYES"** beserta perangkat yang ada.

Dengan Hak Bebas Royalti Noneksklusif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalihmediakan, mengelola dalam bentuk pangkalan data, merawat, dan memublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Medan, 13 Oktober 2026
Yang menyatakan,


MUHAMMAD DIAZ R. RAMBE
NPM : 2209010031

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Muhammad Diaz R. Rambe
Tempat dan Tanggal Lahir : Binjai, 18 Maret 2003
Alamat Rumah : Jl. Soekarno Hatta Lk. V Binjai Timur
Telepon/Faks/HP : 081269606307
E-mail : diazrrambe148@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD	: SD Ahmad Yani Binjai	TAMAT: 2015
SMP	: SMP Ahmad Yani Binjai	TAMAT: 2018
SMA	: SMA Negeri 2 Binjai	TAMAT: 2021

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji dan syukur penulis panjatkan ke hadirat Allah Subhanahu wa Ta'ala atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **“ANALISIS SENTIMEN KEPUASAN MASYARAKAT TERHADAP PELAYANAN APLIKASI MOBILE JKN MENGGUNAKAN ALGORITMA NAÏVE BAYES”**. Selawat dan salam senantiasa tercurah kepada Nabi Muhammad Shallallahu ‘alaihi wa sallam.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.

Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, dan dukungan berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Akrim, M.Pd., Rektor Universitas Muhammadiyah Sumatera Utara.
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom., Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.
3. Ibu Dr. Firahmi Rizky, S.Kom, M.Kom., selaku Wakil Dekan 1 Program Studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
4. Bapak Muhammad Basri, S.Si, M.Kom., selaku Wakil Dekan 3 Program Studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
5. Bapak Mahardika Abdi Prawira Tanjung, M.Kom., Ketua Program Studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.

6. Bapak Mulkan Azhari, S.Kom., M.Kom., Sekretaris Program Studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.
7. Bapak Martiano, S.Pd., S.Kom., M.Kom., selaku Dosen Pembimbing yang telah meluangkan waktu, tenaga, dan pikiran dalam memberikan bimbingan, arahan, koreksi, serta motivasi selama penyusunan skripsi ini.
8. Seluruh dosen dan staf Program Studi Sistem Informasi yang telah memberikan ilmu dan pelayanan selama masa perkuliahan.
9. Kedua orang tua dan seluruh keluarga yang senantiasa memberikan doa, dukungan, dan semangat kepada penulis.
10. Rekan-rekan mahasiswa Program Studi Sistem Informasi angkatan 2022 yang telah memberikan dukungan selama proses perkuliahan dan penyusunan skripsi.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan bagi pengembangan ilmu pengetahuan.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Medan, 13 Oktober 2026
Penulis,

MUHAMMAD DIAZ R. RAMBE
NPM : 2209010031

ANALISIS SENTIMEN KEPUASAN MASYARAKAT TERHADAP PELAYANAN APLIKASI MOBILE JKN MENGGUNAKAN ALGORITMA NAÏVE BAYES

ABSTRAK

Program Jaminan Kesehatan Nasional (JKN) yang diselenggarakan oleh BPJS Kesehatan 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 digunakan 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 dua pendekatan, yaitu rating pengguna dan kamus sentimen InSet, dengan verifikasi manual pada ulasan yang kedua pendekatannya menghasilkan label berbeda, sehingga diperoleh 751 ulasan berlabel positif dan negatif. Data diproses melalui tahapan cleaning, case folding, tokenizing, stopword removal, dan stemming, kemudian dibobotkan menggunakan TF-IDF. Hasil pengujian menunjukkan model Naïve Bayes memperoleh accuracy sebesar 90,07%, precision 90,57%, recall 90,07%, dan F1-score 89,99%, sedangkan pengujian 10-fold cross validation memperoleh rata-rata akurasi 90,68% dengan simpangan baku 2,60%. Sebaran sentimen menunjukkan 403 ulasan atau 53,7% bersentimen negatif dan 348 ulasan atau 46,3% bersentimen positif. Analisis kata paling berpengaruh menunjukkan 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

**SENTIMENT ANALYSIS OF PUBLIC SATISFACTION WITH THE
SERVICES OF THE MOBILE JKN APPLICATION USING THE NAÏVE
BAYES ALGORITHM**

ABSTRACT

The National Health Insurance (JKN) program organised by BPJS Kesehatan is undergoing a policy transition towards the Standard Inpatient Class (KRIS) based on Presidential Regulation Number 59 of 2024. This policy change has generated a wide range of public responses scattered across user reviews of the Mobile JKN application on the Google Play Store, yet these responses have not been processed into information that can be used to evaluate the service. This research aims to classify the sentiment of Mobile JKN user reviews using the Naïve Bayes algorithm and to implement the classification results into a web-based information system. Data were collected through web scraping of 4,000 Indonesian-language reviews from 15 July to 15 August 2026, resulting in 2,675 selected reviews. Labelling was carried out using two approaches, namely user ratings and the InSet sentiment lexicon, with manual verification of reviews for which the two approaches produced different labels, yielding 751 reviews labelled as positive or negative. The data were processed through cleaning, case folding, tokenizing, stopword removal, and stemming, and then weighted using TF-IDF. The test results show that the Naïve Bayes model achieved an accuracy of 90.07%, precision of 90.57%, recall of 90.07%, and an F1-score of 89.99%, while 10-fold cross validation obtained an average accuracy of 90.68% with a standard deviation of 2.60%. The sentiment distribution shows 403 reviews or 53.7% with negative sentiment and 348 reviews or 46.3% with positive sentiment. Analysis of the most influential words indicates that user complaints are concentrated on the authentication process, particularly OTP code delivery and account verification. The web-based information system passed all 12 Black Box test scenarios and can be reused to analyse user reviews without reading them one by one.

Keywords: *sentiment analysis, Naïve Bayes, Mobile JKN, Standard Inpatient Class, TF-IDF, web-based information system*

DAFTAR ISI

LEMBAR PENGESAHAN	ii
PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI.....	iv
RIWAYAT HIDUP	v
KATA PENGANTAR.....	vi
ABSTRAK	viii
ABSTRACT	ix
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian.....	6
BAB II LANDASAN TEORI	8
2.1. Pelayanan Publik	8
2.2. Kepuasan Masyarakat.....	9
2.3. Program JKN dan Aplikasi Mobile JKN	10
2.4. Analisis Sentimen	11
2.4.1. Pengertian Analisis Sentimen	11
2.4.2. Tujuan Analisis Sentimen	12
2.4.3. Klasifikasi Sentimen.....	13
2.4.4. Tahapan Analisis Sentimen.....	13
2.5. Text Mining	13
2.5.1. Pengertian Text Mining	13
2.5.2. Tahapan Text Mining	14
2.6. Natural Language Processing (NLP).....	15
2.6.1. Pengertian Natural Language Processing (NLP).....	15
2.6.2. Tahapan Natural Language Processing.....	16
2.6.3. Peran Natural Language Processing dalam Analisis Sentimen	18
2.7. Pembobotan Kata TF-IDF	18
2.8. Algoritma Naïve Bayes	19

2.8.1. Pengertian Algoritma Naïve Bayes.....	19
2.8.2. Teorema Bayes.....	20
2.8.3. Cara Kerja Algoritma Naïve Bayes	21
2.8.4. Kelebihan dan Kekurangan Algoritma Naïve Bayes	21
2.8.5. Penerapan Algoritma Naïve Bayes dalam Analisis Sentimen	22
2.8.6. Laplace Smoothing	23
2.9. Google Play Store dan Ulasan Pengguna	23
2.10. Web Scraping.....	25
2.11. Confusion Matrix.....	25
2.12. Unified Modeling Language (UML).....	27
2.12.1. Use Case Diagram	28
2.12.2. Activity Diagram	28
2.12.3. Sequence Diagram	29
2.12.4. Class Diagram.....	30
2.13. Black Box Testing	30
2.14. Penelitian Terdahulu.....	31
2.15. Kerangka Berpikir	33
BAB III METODOLOGI PENELITIAN	36
3.1. Lokasi dan Waktu Penelitian	36
3.2. Pendekatan dan Tahapan Penelitian	37
3.3. Data Penelitian	40
3.3.1. Data Primer	41
3.3.2. Data Sekunder.....	41
3.4. Teknik Pengumpulan Data	42
3.4.1. Pengumpulan Data Ulasan Google Play Store	42
3.4.2. Seleksi Dataset.....	42
3.5. Alur Pengolahan Sentimen	43
3.5.1. Text Preprocessing.....	44
3.5.2. Pelabelan Sentimen.....	46
3.5.3. Pembagian Dataset	46
3.5.4. Klasifikasi Menggunakan Algoritma Naïve Bayes	47
3.5.5. Evaluasi Model	47

3.5.6. Visualisasi Hasil Analisis.....	47
3.6. Perancangan Sistem.....	48
3.6.1. Use Case Diagram	48
3.6.2. Activity Diagram	50
3.6.3. Sequence Diagram.....	51
3.6.4. Class Diagram.....	51
3.7. Entity Relationship Diagram (ERD)	52
3.8. Perancangan Interface	53
3.8.1. Halaman Login	54
3.8.2. Halaman Dashboard	54
3.8.3. Halaman Dataset.....	55
3.8.4. Halaman Analisis Sentimen.....	56
3.8.5. Halaman Hasil Analisis	57
3.9. Pengujian Sistem	58
3.9.1. Pengujian Fungsional	58
3.9.2. Pengujian Hasil Klasifikasi	58
BAB IV HASIL DAN PEMBAHASAN.....	59
4.1. Hasil Pengumpulan Data	59
4.2. Hasil Pengambilan Sampel.....	60
4.3. Hasil Pelabelan Sentimen.....	60
4.4. Hasil Text Preprocessing	62
4.5. Hasil Pembobotan TF-IDF	62
4.6. Hasil Pembagian Dataset.....	63
4.7. Hasil Klasifikasi Algoritma Naïve Bayes.....	63
4.8. Hasil Evaluasi Model	63
4.9. Implementasi Sistem	67
4.9.1. Halaman Login	67
4.9.2. Halaman Dashboard	67
4.9.3. Halaman Dataset.....	68
4.9.4. Halaman Analisis Sentimen.....	69
4.9.5. Halaman Hasil Analisis	70
4.9.6. Potongan Kode Program Utama	72

4.10. Hasil Pengujian Sistem.....	76
4.11. Pembahasan	76
4.11.1. Kata Paling Berpengaruh.....	76
4.11.2. Analisis Kesalahan Klasifikasi.....	77
4.11.3. Perbandingan dengan Penelitian Terdahulu.....	78
4.11.4. Jawaban atas Rumusan Masalah.....	79
BAB V PENUTUP	80
5.1. Kesimpulan.....	80
5.2. Saran.....	81
DAFTAR PUSTAKA.....	82
LAMPIRAN.....	84

DAFTAR TABEL

Tabel 2.1 Confusion Matrix (Handayanto et al., 2021).....	26
Tabel 2.2 Simbol Use Case Diagram	28
Tabel 2.3 Simbol Activity Diagram	28
Tabel 2.4 Simbol Sequence Diagram	29
Tabel 2.5 Simbol Class Diagram.....	30
Tabel 2.6 Penelitian Terdahulu.....	31
Tabel 4.1 Hasil Seleksi Dataset.....	59
Tabel 4.2 Sebaran Rating pada Dataset.....	60
Tabel 4.3 Alokasi Sampel per Kelompok Rating	60
Tabel 4.4 Hasil Pelabelan Sentimen.....	61
Tabel 4.5 Contoh Hasil Text Preprocessing	62
Tabel 4.6 Pembagian Dataset	63
Tabel 4.7 Contoh Hasil Klasifikasi	63
Tabel 4.8 Confusion Matrix	63
Tabel 4.9 Perhitungan TP, TN, FP, dan FN per Kelas	64
Tabel 4.10 Nilai Evaluasi Model.....	64
Tabel 4.11 Hasil 10-Fold Cross Validation	65
Tabel 4.12 Hasil Pengujian Black Box	76
Tabel 4.13 Kata Paling Berpengaruh pada Setiap Kelas.....	76

DAFTAR GAMBAR

Gambar 2.1 Kerangka Berpikir	35
Gambar 3.1 Tahapan Penelitian	39
Gambar 3.2 Alur Pengolahan Sentimen	44
Gambar 3.3 Use Case Diagram	49
Gambar 3.4 Activity Diagram	50
Gambar 3.5 Sequence Diagram	51
Gambar 3.6 Class Diagram	52
Gambar 3.7 Entity Relationship Diagram	53
Gambar 3.8 Rancangan Halaman Login	54
Gambar 3.9 Rancangan Halaman Dashboard	55
Gambar 3.10 Rancangan Halaman Dataset	56
Gambar 3.11 Rancangan Halaman Analisis Sentimen	56
Gambar 3.12 Rancangan Halaman Hasil Analisis	57
Gambar 4.1 Confusion Matrix Hasil Klasifikasi	64
Gambar 4.2 Distribusi Sentimen Hasil Klasifikasi	66
Gambar 4.3 Tampilan Halaman Login	67
Gambar 4.4 Tampilan Halaman Dashboard	68
Gambar 4.5 Tampilan Halaman Dataset	69
Gambar 4.6 Tampilan Halaman Analisis Sentimen	70
Gambar 4.7 Tampilan Halaman Hasil Analisis	71
Gambar 4.8 Tampilan Detail Hasil Klasifikasi pada Sistem	71
Gambar 4.9 Potongan Kode Tahapan Text Preprocessing	73
Gambar 4.10 Potongan Kode Pembobotan TF-IDF dan Pelatihan Model Naïve Bayes	74
Gambar 4.11 Potongan Kode Evaluasi Model	75
Gambar 4.12 Potongan Kode Pengujian 10-Fold Cross Validation	75

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Perkembangan teknologi informasi dan komunikasi telah mengubah cara masyarakat dalam menyampaikan pendapat, kritik, maupun apresiasi terhadap suatu layanan. Jika sebelumnya penyampaian penilaian dilakukan secara langsung melalui kotak saran atau survei kepuasan, saat ini masyarakat lebih banyak memanfaatkan kanal digital untuk menyampaikan pengalaman yang mereka terima. Salah satu kanal yang paling banyak digunakan adalah kolom ulasan pada toko aplikasi digital, karena ulasan tersebut bersifat terbuka, dapat diakses oleh siapa pun, serta tersimpan dalam jangka waktu yang panjang. Informasi yang terkandung di dalamnya dapat dimanfaatkan sebagai bahan evaluasi untuk meningkatkan kualitas pelayanan kepada masyarakat.

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 upaya transformasi digital pelayanan publik, BPJS Kesehatan menghadirkan aplikasi Mobile JKN sebagai kanal layanan utama berbasis telepon

pintar. Melalui aplikasi tersebut, peserta dapat melakukan pendaftaran kepesertaan, mengubah data, mengambil nomor antrean pada fasilitas kesehatan, memeriksa status pembayaran iuran, hingga mengakses layanan skrining riwayat kesehatan tanpa harus datang langsung ke kantor cabang. Dengan demikian, Mobile JKN tidak lagi sekadar aplikasi pendukung, melainkan telah menjadi wajah utama pelayanan BPJS Kesehatan di mata masyarakat, 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, pemerintah menghapus sistem kelas perawatan 1, 2, dan 3 dan menggantinya dengan Kelas Rawat Inap Standar (KRIS) yang mulai diberlakukan pada Agustus 2026. Kebijakan ini merupakan salah satu perubahan terbesar dalam penyelenggaraan program JKN karena berdampak langsung terhadap seluruh peserta. Perubahan tersebut menimbulkan berbagai pertanyaan, kebingungan, maupun keluhan dari masyarakat, dan sebagian besar disampaikan melalui kanal digital, termasuk kolom ulasan aplikasi Mobile JKN pada Google Play Store.

Permasalahannya, ulasan yang disampaikan masyarakat berbentuk teks tidak terstruktur dengan jumlah yang sangat besar. Kondisi tersebut menyebabkan proses pembacaan dan penilaian secara manual menjadi tidak efektif serta membutuhkan waktu yang sangat lama. Akibatnya, penyelenggara layanan kesulitan mengidentifikasi ulasan yang bersifat positif maupun negatif, sehingga keluhan yang paling dominan tidak dapat dipetakan secara objektif. Padahal informasi tersebut sangat dibutuhkan sebagai dasar evaluasi dan penentuan

prioritas perbaikan layanan. Oleh karena itu, diperlukan suatu metode yang mampu mengolah opini pengguna secara otomatis sehingga menghasilkan informasi yang lebih mudah dipahami dan dapat digunakan sebagai dasar pengambilan keputusan.

Beberapa penelitian sebelumnya menunjukkan bahwa algoritma Naïve Bayes mampu memberikan performa yang baik dalam analisis sentimen. Penelitian yang dilakukan oleh Rahmawati et al. (2024) mengenai analisis sentimen ulasan pengguna aplikasi m-BCA memperoleh akurasi sebesar 82%. Penelitian Hidayat et al. (2024) mengenai layanan Sistem Informasi Akademik Mahasiswa juga menunjukkan bahwa algoritma Naïve Bayes mampu menghasilkan akurasi sebesar 80,06%. Penelitian lain yang dilakukan oleh Dzulkarnain et al. (2024) pada penilaian masyarakat terhadap pelayanan rumah sakit di Kota Malang memperoleh nilai akurasi mencapai 90%, sehingga membuktikan bahwa metode Naïve Bayes dapat digunakan sebagai metode yang efektif dalam mengevaluasi kualitas pelayanan publik.

Penelitian mengenai analisis sentimen terhadap aplikasi Mobile JKN sebenarnya telah dilakukan oleh beberapa peneliti sebelumnya (Putra et al., 2025; Amaa et al., 2025; Paliling & Sutoyo, 2026). 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 saat ini. 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 celah penelitian yang berupaya diisi melalui penelitian ini.

Berdasarkan uraian tersebut, penulis tertarik untuk melakukan penelitian dengan judul “Analisis Sentimen Kepuasan Masyarakat terhadap Pelayanan Aplikasi Mobile JKN Menggunakan Algoritma Naïve Bayes”. Penelitian ini diharapkan dapat memberikan gambaran yang objektif mengenai tingkat kepuasan masyarakat terhadap pelayanan digital BPJS Kesehatan serta menghasilkan sistem informasi yang dapat dimanfaatkan sebagai sarana evaluasi berkelanjutan.

1.2. Rumusan Masalah

Berdasarkan penjelasan dari masalah yang telah dipaparkan pada latar belakang, berikut ini merupakan rumusan masalah dalam penelitian ini, yaitu:

1. Bagaimana penerapan algoritma Naïve Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Mobile JKN pada Google Play Store ke dalam kategori positif dan negatif?
2. Bagaimana tingkat performa algoritma Naïve Bayes dalam mengklasifikasikan sentimen tersebut berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score*?
3. Bagaimana hasil klasifikasi sentimen tersebut diimplementasikan ke dalam sistem informasi berbasis web sehingga dapat digunakan sebagai gambaran tingkat kepuasan masyarakat terhadap pelayanan aplikasi Mobile JKN?

1.3. Batasan Masalah

Agar penelitian ini lebih terarah dan tidak menyimpang dari tujuan yang telah ditetapkan, maka batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Objek penelitian adalah aplikasi Mobile JKN BPJS Kesehatan yang tersedia pada Google Play Store dengan nama paket app.bpjs.mobile.
2. Data yang digunakan berupa ulasan pengguna 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.
3. Tahapan text preprocessing yang digunakan meliputi cleaning, case folding, tokenizing, stopword removal, dan stemming, sedangkan pembobotan kata dilakukan menggunakan metode TF-IDF.
4. Algoritma klasifikasi yang digunakan adalah Multinomial Naïve Bayes dan tidak dibandingkan dengan algoritma klasifikasi lainnya.
5. Evaluasi performa model dilakukan menggunakan *Confusion Matrix* dengan parameter *accuracy*, *precision*, *recall*, dan *F1-score*.
6. Sistem yang dibangun berbasis web dengan satu aktor, yaitu Admin, dan diuji menggunakan metode Black Box Testing.
7. Penelitian ini tidak membahas aspek hukum, kebijakan, maupun kinerja kelembagaan BPJS Kesehatan, melainkan terbatas pada opini pengguna terhadap aplikasi Mobile JKN.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan penelitian ini adalah sebagai berikut:

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 tingkat performa algoritma Naïve Bayes dalam mengklasifikasikan sentimen tersebut berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score*.
3. Mengimplementasikan hasil klasifikasi sentimen ke dalam sistem informasi berbasis web sebagai sarana penyajian gambaran tingkat kepuasan masyarakat terhadap pelayanan aplikasi Mobile JKN.

1.5. Manfaat Penelitian

a. Manfaat Teoritis

1. Menambah referensi ilmiah mengenai penerapan analisis sentimen menggunakan algoritma Naïve Bayes pada data ulasan aplikasi layanan publik.
2. Memberikan kontribusi terhadap pengembangan ilmu pengetahuan, khususnya pada bidang Data Mining, *Text Mining*, dan Natural Language Processing (NLP) dalam menganalisis opini pengguna.
3. Menjadi referensi bagi peneliti selanjutnya yang ingin mengembangkan penelitian mengenai analisis sentimen pada layanan digital publik, khususnya pada masa perubahan kebijakan.

b. Manfaat Praktis

1. Bagi Penulis

Penelitian ini dapat meningkatkan pengetahuan dan pengalaman penulis dalam menerapkan metode analisis sentimen menggunakan

algoritma Naïve Bayes serta membangun sistem informasi berbasis web.

2. Bagi BPJS Kesehatan

Hasil penelitian diharapkan dapat menjadi bahan evaluasi dalam mengetahui persepsi dan tingkat kepuasan masyarakat terhadap pelayanan aplikasi Mobile JKN, sehingga dapat digunakan sebagai dasar dalam meningkatkan kualitas layanan digital.

3. Bagi Masyarakat

Penelitian ini diharapkan dapat memberikan informasi mengenai kecenderungan opini pengguna terhadap pelayanan aplikasi Mobile JKN.

BAB II

LANDASAN TEORI

2.1. Pelayanan Publik

Berdasarkan Undang-Undang Nomor 25 Tahun 2009 tentang Pelayanan Publik, pelayanan publik didefinisikan sebagai kegiatan atau rangkaian kegiatan dalam rangka pemenuhan kebutuhan pelayanan sesuai dengan peraturan perundang-undangan bagi setiap warga negara dan penduduk atas barang, jasa, dan/atau pelayanan administratif yang disediakan oleh penyelenggara pelayanan publik. Penyelenggara pelayanan publik meliputi instansi pemerintah, korporasi, lembaga independen yang dibentuk berdasarkan undang-undang, serta badan hukum lain yang dibentuk semata-mata untuk kegiatan pelayanan publik.

Kualitas pelayanan publik umumnya diukur menggunakan lima dimensi kualitas layanan yang dikenal dengan model SERVQUAL (Parasuraman et al., 1988), yaitu tangible (bukti fisik), reliability (kehandalan), responsiveness (daya tanggap), assurance (jaminan), dan empathy (empati). Dalam konteks pelayanan berbasis aplikasi digital, kelima dimensi tersebut tercermin melalui ketersediaan fitur, kestabilan sistem, kecepatan respons aplikasi, keamanan data pengguna, serta kemudahan penggunaan.

Seiring dengan kebijakan transformasi digital pemerintahan, penyelenggaraan pelayanan publik saat ini banyak dipindahkan ke kanal digital. Pergeseran tersebut menuntut adanya mekanisme evaluasi yang mampu menangkap persepsi masyarakat terhadap layanan digital secara cepat

dan berkelanjutan, tidak lagi semata-mata melalui survei konvensional yang bersifat periodik.

2.2. Kepuasan Masyarakat

Kepuasan merupakan perasaan senang atau kecewa yang timbul dari perbandingan antara kinerja suatu produk atau layanan yang dirasakan dengan harapan yang dimiliki oleh penggunanya. Apabila kinerja yang dirasakan berada di bawah harapan, maka pengguna akan merasa tidak puas; sebaliknya apabila kinerja memenuhi atau melampaui harapan, maka pengguna akan merasa puas. Dalam konteks pelayanan publik, kepuasan masyarakat menggambarkan sejauh mana layanan yang diberikan penyelenggara mampu memenuhi harapan masyarakat sebagai penerima layanan.

Pemerintah mengatur pengukuran kepuasan masyarakat melalui Peraturan Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Nomor 14 Tahun 2017 tentang Pedoman Penyusunan Survei Kepuasan Masyarakat Unit Penyelenggara Pelayanan Publik. Peraturan tersebut menetapkan sembilan unsur penilaian, yaitu persyaratan, sistem dan prosedur, waktu penyelesaian, biaya atau tarif, produk spesifikasi jenis pelayanan, kompetensi pelaksana, perilaku pelaksana, penanganan pengaduan, serta sarana dan prasarana. Hasil pengukuran tersebut dinyatakan dalam bentuk Indeks Kepuasan Masyarakat (IKM).

Pengukuran kepuasan masyarakat melalui survei memiliki keterbatasan, yaitu bersifat periodik, memerlukan biaya dan waktu yang tidak sedikit, serta jumlah respondennya terbatas. Sebaliknya, ulasan yang disampaikan masyarakat secara sukarela pada kanal digital tersedia secara terus-menerus

dalam jumlah yang jauh lebih besar dan mencerminkan pengalaman aktual pengguna. Oleh karena itu, analisis sentimen terhadap ulasan pengguna dapat berperan sebagai pendekatan alternatif untuk memperoleh gambaran kepuasan masyarakat secara lebih cepat dan berkelanjutan.

Pada penelitian ini, tingkat kepuasan masyarakat diukur melalui proporsi sentimen positif dan negatif yang dihasilkan dari klasifikasi ulasan pengguna aplikasi Mobile JKN. Sentimen positif diinterpretasikan sebagai indikasi kepuasan, sedangkan sentimen negatif sebagai indikasi ketidakpuasan.

2.3. Program JKN dan Aplikasi Mobile JKN

Program Jaminan Kesehatan Nasional (JKN) merupakan program jaminan sosial di bidang kesehatan yang diselenggarakan oleh Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan. Hingga akhir tahun 2025, jumlah peserta program JKN tercatat mencapai 282,7 juta jiwa atau 98,62 persen dari total penduduk Indonesia, dengan pemanfaatan layanan sebanyak 725,3 juta kali sepanjang tahun tersebut (BPJS Kesehatan, 2026).

Mobile JKN merupakan aplikasi berbasis telepon pintar yang dikembangkan BPJS Kesehatan sebagai kanal pelayanan digital bagi peserta program JKN. Aplikasi ini menyediakan berbagai layanan, antara lain pendaftaran peserta, perubahan data kepesertaan, informasi tagihan dan pembayaran iuran, pengambilan nomor antrean pada fasilitas kesehatan, konsultasi dokter, serta skrining riwayat kesehatan. Kehadiran aplikasi ini bertujuan mempersingkat waktu layanan dan mengurangi kebutuhan peserta untuk datang langsung ke kantor cabang.

Pada tahun 2026, penyelenggaraan program JKN memasuki masa transisi kebijakan sebagai akibat diberlakukannya Kelas Rawat Inap Standar (KRIS) berdasarkan Peraturan Presiden Nomor 59 Tahun 2024. Kebijakan tersebut menghapus sistem kelas perawatan 1, 2, dan 3 dan menggantinya dengan standar pelayanan rawat inap yang sama bagi seluruh peserta. Implementasi kebijakan ini dimulai pada Agustus 2026 secara bertahap dan berdampak langsung terhadap seluruh peserta program JKN.

2.4. Analisis Sentimen

2.4.1. Pengertian Analisis Sentimen

Analisis sentimen merupakan salah satu cabang dari Natural Language Processing (NLP) dan *text mining* yang digunakan untuk mengidentifikasi, mengolah, serta mengklasifikasikan opini atau pendapat yang terdapat pada data berbentuk teks ke dalam kategori sentimen tertentu, seperti positif maupun negatif. Analisis sentimen bertujuan untuk memahami persepsi atau sikap seseorang terhadap suatu objek, layanan, maupun peristiwa berdasarkan informasi yang disampaikan dalam bentuk tulisan. Penerapan analisis sentimen saat ini banyak digunakan untuk mengolah data yang berasal dari media sosial, forum diskusi, maupun platform ulasan daring karena mampu membantu pengambilan keputusan berdasarkan opini pengguna (Hidayat et al., 2024).

Dalam perkembangannya, analisis sentimen telah dimanfaatkan pada berbagai bidang, seperti evaluasi kualitas pelayanan publik, layanan kesehatan, pendidikan, e-commerce, hingga sektor pemerintahan. Melalui analisis sentimen, instansi dapat mengetahui kecenderungan opini pengguna

terhadap suatu layanan secara lebih cepat dibandingkan dengan proses analisis manual. Selain mampu mengolah data dalam jumlah besar, analisis sentimen juga memberikan informasi yang dapat digunakan sebagai bahan evaluasi dan penyusunan strategi peningkatan kualitas layanan (Dzulkarnain et al., 2024).

Pada penelitian ini, analisis sentimen digunakan untuk mengidentifikasi tingkat kepuasan masyarakat terhadap pelayanan aplikasi Mobile JKN berdasarkan komentar yang diperoleh dari Google Play Store. Hasil analisis sentimen akan dikelompokkan ke dalam dua kategori, yaitu sentimen positif dan negatif, sehingga dapat memberikan gambaran mengenai persepsi masyarakat terhadap kualitas pelayanan yang diberikan oleh aplikasi Mobile JKN.

2.4.2. Tujuan Analisis Sentimen

Analisis sentimen bertujuan untuk mengetahui kecenderungan opini pengguna terhadap suatu objek atau layanan berdasarkan data teks yang diperoleh dari berbagai sumber. Selain mengidentifikasi polaritas sentimen, analisis sentimen juga digunakan untuk membantu organisasi memahami kebutuhan pengguna, mengevaluasi kualitas layanan, serta mendukung proses pengambilan keputusan berbasis data. Dalam penelitian ini, analisis sentimen bertujuan untuk mengetahui tingkat kepuasan masyarakat terhadap pelayanan aplikasi Mobile JKN melalui pengelompokan komentar menjadi sentimen positif dan negatif.

2.4.3. Klasifikasi Sentimen

Secara umum, hasil analisis sentimen dibedakan menjadi beberapa kategori berdasarkan polaritasnya. Pada penelitian ini digunakan dua kategori sentimen, yaitu:

1. Sentimen Positif, yaitu opini yang menunjukkan rasa puas, apresiasi, atau penilaian baik terhadap pelayanan.
2. Sentimen Negatif, yaitu opini yang mengandung kritik, keluhan, atau ketidakpuasan terhadap pelayanan yang diterima.

Adapun ulasan yang bersifat informatif atau berupa pertanyaan, yang tidak menyatakan penilaian terhadap layanan, tidak diikutsertakan dalam analisis karena tidak menunjukkan kecenderungan kepuasan maupun ketidakpuasan.

2.4.4. Tahapan Analisis Sentimen

Secara umum proses analisis sentimen terdiri atas beberapa tahapan, yaitu:

1. Pengumpulan data (*data collection*).
2. *Text preprocessing*.
3. Ekstraksi atau representasi fitur.
4. Klasifikasi menggunakan algoritma *machine learning* (Naïve Bayes).
5. Evaluasi hasil menggunakan *Confusion Matrix*.

2.5. Text Mining

2.5.1. Pengertian Text Mining

Text mining merupakan proses penggalian informasi dari sekumpulan data berbentuk teks untuk memperoleh pola, hubungan, maupun informasi yang bermanfaat. Berbeda dengan data mining yang mengolah data

terstruktur, *text mining* berfokus pada pengolahan data tidak terstruktur, seperti komentar, ulasan, dokumen, maupun informasi yang berasal dari media sosial. Melalui teknik *text mining*, data teks yang semula sulit dipahami dapat diolah menjadi informasi yang lebih terstruktur sehingga dapat digunakan sebagai dasar dalam proses analisis dan pengambilan keputusan (Hidayat et al., 2024).

Dalam perkembangannya, *text mining* telah banyak diterapkan pada berbagai bidang, seperti analisis ulasan pelanggan, evaluasi pelayanan publik, analisis media sosial, hingga sistem rekomendasi. Salah satu penerapan *text mining* yang paling banyak digunakan adalah analisis sentimen, yaitu proses mengidentifikasi kecenderungan opini pengguna terhadap suatu objek berdasarkan data teks yang tersedia (Dzulkarnain et al., 2024).

Pada penelitian ini, *text mining* digunakan untuk mengolah ulasan pengguna yang diperoleh dari Google Play Store aplikasi Mobile JKN. Proses tersebut dilakukan untuk menghasilkan data yang siap digunakan dalam proses klasifikasi sentimen menggunakan algoritma Naïve Bayes.

2.5.2. Tahapan Text Mining

Secara umum, proses *text mining* terdiri atas beberapa tahapan, yaitu:

1. Pengumpulan Data (Data Collection)

Tahap ini dilakukan dengan mengumpulkan komentar atau ulasan masyarakat yang berasal dari Google Play Store aplikasi Mobile JKN sebagai sumber data penelitian.

2. *Text Preprocessing*

Tahapan ini bertujuan untuk membersihkan data dari karakter yang tidak diperlukan sehingga menghasilkan data yang lebih terstruktur. Proses *preprocessing* meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*.

3. Representasi Data

Data teks yang telah diproses kemudian diubah menjadi bentuk numerik agar dapat diproses oleh algoritma klasifikasi.

4. Klasifikasi

Data yang telah direpresentasikan selanjutnya diklasifikasikan menggunakan algoritma Naïve Bayes untuk menentukan kategori sentimen positif atau negatif.

5. Evaluasi

Tahap terakhir dilakukan untuk mengukur performa model menggunakan *confusion matrix* dengan parameter *accuracy*, *precision*, *recall*, dan *F1 Score*.

2.6. Natural Language Processing (NLP)

2.6.1. Pengertian Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan salah satu cabang ilmu kecerdasan buatan (Artificial Intelligence) yang mempelajari interaksi antara komputer dan bahasa manusia. NLP memungkinkan komputer untuk memahami, mengolah, menganalisis, serta menghasilkan informasi dari data

yang berbentuk bahasa alami, baik berupa teks maupun suara. Dalam perkembangannya, NLP banyak diterapkan pada berbagai bidang, seperti penerjemahan bahasa, chatbot, sistem rekomendasi, analisis dokumen, serta analisis sentimen (Hidayat et al., 2024).

Data yang berasal dari ulasan pengguna di media sosial maupun platform ulasan umumnya masih berbentuk teks tidak terstruktur. Oleh karena itu, diperlukan proses NLP agar data tersebut dapat diproses oleh komputer menjadi informasi yang lebih terstruktur. Proses tersebut meliputi pembersihan data, pemisahan kata, penghapusan kata yang tidak memiliki makna penting, hingga pengubahan kata menjadi bentuk dasar sebelum dilakukan proses klasifikasi (Rahmawati et al., 2024).

2.6.2. Tahapan Natural Language Processing

Dalam analisis sentimen, NLP berperan sebagai tahapan awal untuk mempersiapkan data teks sebelum dilakukan proses klasifikasi. Penerapan NLP membantu mengurangi data yang tidak relevan, memperbaiki kualitas data, serta meningkatkan akurasi algoritma klasifikasi. Oleh karena itu, proses NLP menjadi bagian penting dalam penelitian analisis sentimen karena berpengaruh terhadap kualitas hasil klasifikasi yang dihasilkan.

Secara umum, proses NLP terdiri atas beberapa tahapan yang saling berkaitan. Tahapan tersebut bertujuan untuk mengubah data teks menjadi bentuk yang lebih mudah dipahami oleh sistem komputer sehingga dapat dianalisis lebih lanjut. Tahapan tersebut meliputi:

1. Lexical Analysis

Tahap awal yang bertujuan untuk mengenali karakter, kata, maupun tanda baca yang terdapat dalam suatu teks. Pada tahap ini dilakukan proses identifikasi terhadap setiap komponen penyusun kalimat

2. *Syntactic Analysis*

Tahap ini dilakukan untuk menganalisis struktur kalimat berdasarkan aturan tata Bahasa sehingga hubungan antar kata dalam suatu kalimat dapat dipahami dengan lebih baik.

3. *Semantic Analysis*

Setelah struktur kalimat diketahui, tahap berikutnya adalah memahami makna dari setiap kata maupun keseluruhan kalimat sehingga informasi yang terkandung di dalam teks dapat diinterpretasikan secara tepat.

4. *Discourse integration*

Tahap ini bertujuan untuk memahami hubungan antar kalimat sehingga makna suatu kalimat dapat dipahami berdasarkan konteks pembicaraan secara keseluruhan.

5. *Pragmatic Analysis*

Tahap terakhir bertujuan memahami maksud atau tujuan sebenarnya dari suatu kalimat dalam mempertimbangkan konteks penggunaan Bahasa dalam komunikasi

2.6.3. Peran Natural Language Processing dalam Analisis Sentimen

Dalam penelitian analisis sentimen, NLP memiliki peran penting sebagai teknik yang digunakan untuk mengolah data teks sebelum dilakukan proses klasifikasi. Komentar masyarakat yang berasal dari Google Play Store umumnya masih mengandung berbagai karakter yang tidak diperlukan, seperti tanda baca, angka, simbol, emoji, maupun kata-kata yang tidak memiliki makna penting. Kondisi tersebut dapat memengaruhi kualitas hasil klasifikasi apabila tidak dilakukan pengolahan terlebih dahulu.

Melalui penerapan NLP, data komentar dibersihkan dan disiapkan melalui tahapan *text preprocessing* sehingga menghasilkan data yang lebih terstruktur. Data yang telah diproses kemudian digunakan sebagai masukan bagi algoritma Naïve Bayes untuk mengidentifikasi sentimen positif maupun negatif. Dengan demikian, NLP berfungsi sebagai teknik pendukung yang membantu meningkatkan kualitas data sehingga proses klasifikasi dapat menghasilkan tingkat akurasi yang lebih baik (Dzulkarnain et al., 2024).

Pada penelitian ini, penerapan NLP difokuskan pada tahapan *text preprocessing*, yaitu *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Kelima tahapan tersebut menjadi dasar dalam mempersiapkan data ulasan pengguna sebelum dilakukan analisis sentimen menggunakan algoritma Naïve Bayes.

2.7. Pembobotan Kata TF-IDF

Term Frequency – Inverse Document Frequency (TF-IDF) merupakan metode pembobotan yang digunakan untuk mengubah data teks menjadi bentuk numerik agar dapat diproses oleh algoritma klasifikasi. Metode ini

memberikan bobot pada setiap kata berdasarkan dua komponen, yaitu frekuensi kemunculan kata dalam suatu dokumen (term frequency) dan kebalikan dari jumlah dokumen yang memuat kata tersebut (inverse document frequency).

Prinsip dasar TF-IDF adalah bahwa kata yang sering muncul pada satu dokumen tetapi jarang muncul pada dokumen lain memiliki daya pembeda yang tinggi sehingga diberi bobot besar. Sebaliknya, kata yang muncul hampir pada seluruh dokumen dianggap kurang informatif sehingga diberi bobot kecil. Secara matematis, nilai TF-IDF dihitung menggunakan persamaan $W(t,d) = tf(t,d) \times \log (N / df(t))$, dengan $W(t,d)$ adalah bobot kata t pada dokumen d , $tf(t,d)$ adalah frekuensi kemunculan kata t pada dokumen d , N adalah jumlah seluruh dokumen, dan $df(t)$ adalah jumlah dokumen yang memuat kata t (Manning et al., 2008).

Pada penelitian ini, TF-IDF digunakan untuk mengubah ulasan pengguna yang telah melalui tahapan *text preprocessing* menjadi representasi numerik sebelum diklasifikasikan menggunakan algoritma Naïve Bayes.

2.8. Algoritma Naïve Bayes

2.8.1. Pengertian Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan salah satu metode klasifikasi yang didasarkan pada Teorema Bayes dengan asumsi bahwa setiap atribut atau fitur bersifat independen terhadap atribut lainnya. Meskipun asumsi tersebut tergolong sederhana, algoritma Naïve Bayes mampu memberikan hasil klasifikasi yang baik pada berbagai permasalahan, khususnya dalam pengolahan data teks (*text classification*) dan analisis sentimen. Metode ini

bekerja dengan menghitung probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan informasi yang diperoleh dari data pelatihan (*training data*) (Hidayat et al., 2024).

Dalam bidang *text mining*, algoritma Naïve Bayes banyak digunakan karena memiliki proses perhitungan yang sederhana, kebutuhan komputasi yang relatif rendah, serta mampu menangani data dalam jumlah besar. Selain itu, algoritma ini tetap mampu menghasilkan tingkat akurasi yang baik meskipun digunakan pada data teks yang memiliki dimensi tinggi. Oleh karena itu, Naïve Bayes menjadi salah satu metode yang paling banyak diterapkan dalam penelitian analisis sentimen terhadap ulasan pengguna maupun komentar pada media sosial (Dzulkarnain et al., 2024).

Pada penelitian ini, algoritma Naïve Bayes digunakan untuk mengklasifikasikan ulasan pengguna mengenai pelayanan aplikasi Mobile JKN ke dalam dua kategori sentimen, yaitu positif dan negatif. Data yang telah melalui proses *text preprocessing* selanjutnya diproses menggunakan algoritma Naïve Bayes untuk mengetahui kecenderungan opini pengguna terhadap kualitas pelayanan yang diberikan.

2.8.2. Teorema Bayes

Algoritma Naïve Bayes didasarkan pada **Teorema Bayes**, yaitu metode untuk menghitung probabilitas suatu kejadian berdasarkan informasi atau bukti yang telah diketahui sebelumnya. Secara matematis, Teorema Bayes dinyatakan sebagai berikut:

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)}$$

Keterangan:

- a. $P(C|X)$ = probabilitas suatu data termasuk ke dalam kelas C berdasarkan data X.
- b. $P(X|C)$ = probabilitas munculnya data X pada kelas C.
- c. $P(C)$ = probabilitas awal (prior probability) dari kelas C.
- d. $P(X)$ = probabilitas munculnya data X.

Berdasarkan persamaan tersebut, algoritma Naïve Bayes akan memilih kelas yang memiliki nilai probabilitas paling besar sebagai hasil klasifikasi.

2.8.3. Cara Kerja Algoritma Naïve Bayes

Secara umum, proses klasifikasi menggunakan algoritma Naïve Bayes dilakukan melalui beberapa tahapan sebagai berikut:

- a. Mengumpulkan data latih (training data) yang telah memiliki label sentimen.
- b. Menghitung probabilitas kemunculan setiap kata pada masing-masing kelas sentimen.
- c. Menghitung probabilitas setiap komentar berdasarkan Teorema Bayes.
- d. Membandingkan nilai probabilitas pada setiap kelas.
- e. Menentukan kelas sentimen dengan nilai probabilitas terbesar sebagai hasil klasifikasi.

2.8.4. Kelebihan dan Kekurangan Algoritma Naïve Bayes

Kelebihan

Algoritma Naïve Bayes memiliki beberapa kelebihan, antara lain:

- a. Memiliki proses perhitungan yang sederhana sehingga mudah diimplementasikan.

- b. Memerlukan waktu komputasi yang relatif cepat.
- c. Mampu mengolah data dalam jumlah besar.
- d. Memiliki performa yang baik pada klasifikasi data teks.
- e. Tetap memberikan hasil yang baik meskipun jumlah data pelatihan terbatas.

Kekurangan

Selain memiliki kelebihan, algoritma Naïve Bayes juga memiliki beberapa keterbatasan, yaitu:

- a. Mengasumsikan setiap fitur bersifat independen, padahal pada data nyata hubungan antar fitur sering kali saling berkaitan.
- b. Sensitif terhadap distribusi data yang tidak seimbang.
- c. Kinerja klasifikasi dapat menurun apabila data pelatihan kurang representatif.

2.8.5. Penerapan Algoritma Naïve Bayes dalam Analisis Sentimen

Algoritma Naïve Bayes telah banyak digunakan dalam penelitian analisis sentimen karena mampu mengklasifikasikan opini pengguna secara efektif berdasarkan probabilitas kemunculan kata. Berbagai penelitian menunjukkan bahwa metode ini memberikan tingkat akurasi yang baik pada analisis ulasan pengguna aplikasi, pelayanan rumah sakit, layanan akademik, maupun pelayanan publik (Rahmawati et al., 2024; Hidayat et al., 2024; Dzulkarnain et al., 2024).

Dalam penelitian ini, algoritma Naïve Bayes diterapkan untuk menganalisis ulasan pengguna mengenai pelayanan aplikasi Mobile JKN

yang diperoleh dari Google Play Store. Hasil klasifikasi berupa sentimen positif dan negatif diharapkan dapat memberikan gambaran mengenai tingkat kepuasan masyarakat terhadap kualitas pelayanan yang diberikan serta menjadi bahan evaluasi bagi aplikasi Mobile JKN dalam meningkatkan mutu pelayanan publik.

2.8.6. Laplace Smoothing

Salah satu keterbatasan algoritma Naïve Bayes adalah munculnya nilai probabilitas nol ketika terdapat kata pada data uji yang tidak pernah muncul pada data latih untuk suatu kelas tertentu. Karena perhitungan Naïve Bayes dilakukan dengan mengalikan seluruh probabilitas kata, satu nilai nol akan menyebabkan hasil perkalian keseluruhan menjadi nol, sehingga kelas tersebut tidak akan pernah terpilih meskipun kata-kata lainnya sangat mendukung.

Untuk mengatasi permasalahan tersebut digunakan teknik *Laplace Smoothing*, yaitu dengan menambahkan nilai konstanta pada setiap frekuensi kemunculan kata sehingga tidak ada probabilitas yang bernilai nol. Perhitungannya dinyatakan sebagai $P(t|c) = (\text{count}(t,c) + 1) / (\text{total_kata}(c) + |V|)$, dengan $|V|$ adalah jumlah kata unik pada keseluruhan kosakata (Manning et al., 2008). Pada penelitian ini, *Laplace Smoothing* diterapkan untuk memastikan seluruh kata pada data uji tetap dapat diperhitungkan dalam proses klasifikasi.

2.9. Google Play Store dan Ulasan Pengguna

Google Play Store merupakan layanan distribusi aplikasi digital resmi untuk perangkat bersistem operasi Android. Selain berfungsi sebagai sarana

pengunduhan aplikasi, Google Play Store menyediakan fitur ulasan yang memungkinkan pengguna memberikan penilaian berupa *rating* berskala satu sampai lima bintang serta komentar dalam bentuk teks berdasarkan pengalaman mereka menggunakan aplikasi tersebut.

Ulasan pengguna pada Google Play Store memiliki beberapa karakteristik yang menjadikannya sumber data yang baik untuk penelitian analisis sentimen. Pertama, ulasan bersifat publik sehingga dapat diakses tanpa memerlukan izin khusus. Kedua, setiap ulasan disertai metadata berupa *rating*, tanggal, dan versi aplikasi, sehingga memudahkan proses seleksi data. Ketiga, jumlah ulasan pada aplikasi layanan publik berskala nasional dapat mencapai ratusan ribu, sehingga tersedia data yang memadai untuk proses pelatihan dan pengujian model klasifikasi.

Bagi penyedia layanan, ulasan pengguna berfungsi sebagai umpan balik langsung mengenai kelebihan dan kekurangan aplikasi. Namun demikian, jumlah ulasan yang sangat besar menyebabkan proses pembacaan secara manual menjadi tidak efektif, sehingga diperlukan teknik analisis sentimen untuk mengelompokkan opini pengguna secara otomatis.

Pada penelitian ini, Google Play Store digunakan sebagai satu-satunya sumber data penelitian. Ulasan pengguna aplikasi Mobile JKN dikumpulkan, kemudian diproses melalui tahapan *text preprocessing* sebelum diklasifikasikan menggunakan algoritma Naïve Bayes ke dalam kategori sentimen positif dan negatif.

2.10. Web Scraping

Web scraping merupakan teknik pengambilan data dari halaman web secara otomatis menggunakan program komputer. Teknik ini digunakan untuk memperoleh data dalam jumlah besar yang apabila dikumpulkan secara manual akan membutuhkan waktu yang sangat lama. Proses *web scraping* umumnya terdiri atas tiga tahapan, yaitu pengiriman permintaan ke server, pengambilan data yang dikembalikan server, serta penyimpanan data ke dalam format terstruktur seperti CSV atau basis data.

Pada penelitian ini, proses *web scraping* dilakukan menggunakan bahasa pemrograman Python dengan bantuan pustaka *google-play-scraper*. Pustaka tersebut memungkinkan pengambilan ulasan pengguna suatu aplikasi beserta metadata pendukungnya, seperti isi ulasan, *rating*, tanggal, dan versi aplikasi. Data hasil pengambilan kemudian disimpan dalam berkas berformat CSV untuk diproses pada tahapan berikutnya.

2.11. Confusion Matrix

Confusion Matrix merupakan salah satu metode evaluasi yang digunakan untuk mengukur kinerja model klasifikasi dengan membandingkan hasil prediksi yang dihasilkan oleh model terhadap kelas sebenarnya. Melalui *Confusion Matrix*, tingkat keberhasilan model dalam mengklasifikasikan data dapat diketahui berdasarkan jumlah prediksi yang benar maupun prediksi yang salah. Pada permasalahan klasifikasi biner, hasil klasifikasi dibagi ke dalam empat kategori, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) (Handayanto et al., 2021).

Keempat komponen tersebut memiliki fungsi yang berbeda dalam mengevaluasi performa model. *True Positive* (TP) menunjukkan jumlah data positif yang berhasil diprediksi sebagai positif. *True Negative* (TN) merupakan jumlah data negatif yang berhasil diprediksi sebagai negatif. *False Positive* (FP) menunjukkan data negatif yang salah diprediksi sebagai positif, sedangkan *False Negative* (FN) merupakan data positif yang salah diprediksi sebagai negatif. Nilai-nilai tersebut menjadi dasar dalam menghitung berbagai metrik evaluasi, seperti akurasi (*accuracy*), presisi (*precision*), *recall*, dan *F1-score* (Berliansyah et al., 2024; Handayanto et al., 2021).

Prediksi	Aktual Positif	Aktual Negatif
Prediksi Positif	True Positive (TP)	False Positive (FP)
Prediksi Negatif	False Negative (FN)	True Negative (TN)

Tabel 2.1 Confusion Matrix (Handayanto et al., 2021)

Berdasarkan nilai yang diperoleh pada *Confusion Matrix*, salah satu metrik evaluasi yang paling sering digunakan adalah akurasi (*accuracy*). Akurasi menunjukkan persentase jumlah data yang berhasil diklasifikasikan dengan benar dibandingkan dengan seluruh data yang diuji. Semakin tinggi nilai akurasi, maka semakin baik kemampuan model dalam melakukan klasifikasi terhadap data yang digunakan (Handayanto et al., 2021).

Rumus akurasi adalah sebagai berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

Keterangan:

- TP (*True Positive*)** = jumlah data positif yang diprediksi benar.
- TN (*True Negative*)** = jumlah data negatif yang diprediksi benar.

- c. **FP (*False Positive*)** = jumlah data negatif yang diprediksi positif.
- d. **FN (*False Negative*)** = jumlah data positif yang diprediksi negatif.

Pada penelitian ini, *Confusion Matrix* digunakan untuk mengevaluasi hasil klasifikasi sentimen menggunakan algoritma Naïve Bayes terhadap data ulasan yang diperoleh dari Google Play Store. Melalui evaluasi tersebut dapat diketahui tingkat akurasi model dalam mengklasifikasikan sentimen masyarakat terhadap pelayanan aplikasi Mobile JKN.

2.12. Unified Modeling Language (UML)

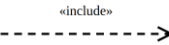
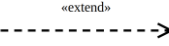
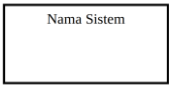
Unified Modeling Language (UML) merupakan bahasa pemodelan standar yang digunakan untuk menggambarkan, merancang, dan mendokumentasikan sistem perangkat lunak berorientasi objek. UML menyediakan sekumpulan notasi grafis yang membantu pengembang memvisualisasikan struktur maupun perilaku sistem sebelum sistem tersebut dibangun (Rosa & Shalahuddin, 2019).

Diagram UML secara umum dikelompokkan menjadi dua kategori, yaitu diagram struktur yang menggambarkan susunan statis sistem, serta diagram perilaku yang menggambarkan aktivitas dan interaksi yang terjadi di dalam sistem. Pada penelitian ini digunakan empat jenis diagram, yaitu Use Case Diagram yang menggambarkan interaksi antara aktor dengan fungsi yang disediakan sistem, Activity Diagram yang menggambarkan urutan aktivitas pada suatu proses, Sequence Diagram yang menggambarkan urutan pesan yang dipertukarkan antarobjek berdasarkan waktu, serta Class Diagram yang menggambarkan struktur kelas, atribut, metode, dan relasi antarkelas.

2.12.1. Use Case Diagram

Use Case Diagram menggambarkan interaksi antara aktor dengan fungsi-fungsi yang disediakan sistem. Diagram ini digunakan untuk menunjukkan siapa saja yang dapat menggunakan sistem serta layanan apa yang tersedia bagi masing-masing pengguna. Simbol yang digunakan pada Use Case Diagram ditunjukkan pada Tabel 2.2.

Tabel 2.2 Simbol Use Case Diagram

No.	Simbol	Nama Simbol	Keterangan
1		Aktor (Actor)	Orang, perangkat, atau sistem lain yang berinteraksi dengan sistem yang dibangun. Aktor berada di luar batas sistem.
2		Use Case	Fungsi atau layanan yang disediakan sistem dan dapat diakses oleh aktor. Dinamai menggunakan kata kerja.
3		Asosiasi (Association)	Hubungan yang menunjukkan bahwa seorang aktor menggunakan suatu use case.
4		Include	Relasi yang menunjukkan bahwa suatu use case selalu memanggil use case lain sebagai bagian wajib dari prosesnya.
5		Extend	Relasi yang menunjukkan bahwa suatu use case dapat diperluas oleh use case lain pada kondisi tertentu.
6		Batas Sistem (System Boundary)	Garis yang memisahkan fungsi-fungsi di dalam sistem dari aktor yang berada di luar sistem.

2.12.2. Activity Diagram

Activity Diagram menggambarkan urutan aktivitas yang terjadi pada suatu proses, termasuk percabangan keputusan dan aktivitas yang berjalan bersamaan. Diagram ini membantu memperjelas alur kerja sistem dari awal hingga akhir. Simbol yang digunakan pada Activity Diagram ditunjukkan pada Tabel 2.3.

Tabel 2.3 Simbol Activity Diagram

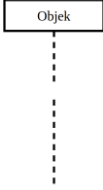
No.	Simbol	Nama Simbol	Keterangan
-----	--------	-------------	------------

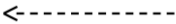
No.	Simbol	Nama Simbol	Keterangan
1		Status Awal (Initial Node)	Titik awal dimulainya rangkaian aktivitas pada suatu proses.
2		Aktivitas (Action)	Aktivitas atau kegiatan yang dilakukan oleh aktor maupun sistem.
3		Percabangan (Decision)	Titik pengambilan keputusan yang menghasilkan lebih dari satu kemungkinan alur berdasarkan suatu kondisi.
4		Fork / Join	Pemisahan satu alur menjadi beberapa alur yang berjalan bersamaan, atau penggabungan kembali beberapa alur menjadi satu.
5		Alur Kendali (Control Flow)	Garis berarah yang menunjukkan urutan perpindahan dari satu aktivitas ke aktivitas berikutnya.
6		Swimlane (Partition)	Pembagian kolom yang menunjukkan pihak mana yang bertanggung jawab atas setiap aktivitas, misalnya pengguna atau sistem.
7		Status Akhir (Final Node)	Titik berakhirnya seluruh rangkaian aktivitas pada suatu proses.

2.12.3. Sequence Diagram

Sequence Diagram menggambarkan urutan pesan yang dipertukarkan antarobjek berdasarkan urutan waktu. Diagram ini memperlihatkan bagaimana suatu fungsi dijalankan mulai dari permintaan pengguna hingga hasil dikembalikan kepada pengguna. Simbol yang digunakan pada Sequence Diagram ditunjukkan pada Tabel 2.4.

Tabel 2.4 Simbol Sequence Diagram

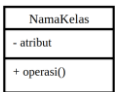
No.	Simbol	Nama Simbol	Keterangan
1		Aktor (Actor)	Pihak yang memulai interaksi dengan sistem pada rangkaian pesan yang digambarkan.
2		Objek (Object)	Objek, kelas, atau antarmuka yang terlibat dalam pertukaran pesan.
3		Garis Hidup (Lifeline)	Garis putus-putus vertikal yang menyatakan keberadaan objek sepanjang waktu berjalannya proses.

No.	Simbol	Nama Simbol	Keterangan
4		Aktivasi (Activation)	Batang persegi pada garis hidup yang menyatakan bahwa objek sedang aktif menjalankan suatu operasi.
5		Pesan (Message)	Panah bergaris penuh yang menyatakan pengiriman pesan atau pemanggilan operasi dari satu objek ke objek lain.
6		Pesan Balasan (Return Message)	Panah bergaris putus-putus yang menyatakan pengembalian hasil kepada objek pengirim.

2.12.4. Class Diagram

Class Diagram menggambarkan struktur kelas yang membentuk sistem beserta atribut, operasi, dan hubungan antarkelas. Diagram ini menjadi dasar dalam perancangan basis data dan struktur program. Simbol yang digunakan pada Class Diagram ditunjukkan pada Tabel 2.5.

Tabel 2.5 Simbol Class Diagram

No.	Simbol	Nama Simbol	Keterangan
1		Kelas (Class)	Kotak bersusun tiga yang memuat nama kelas, atribut, dan operasi yang dimiliki kelas tersebut.
2		Asosiasi (Association)	Hubungan struktural antara dua kelas yang saling berkaitan.
3		Multiplisitas (Multiplicity)	Keterangan jumlah objek yang dapat terlibat dalam suatu hubungan, misalnya satu ke banyak.
4		Generalisasi (Generalization)	Hubungan pewarisan, yaitu suatu kelas merupakan bentuk khusus dari kelas yang lebih umum.
5		Agregasi (Aggregation)	Hubungan bagian dan keseluruhan, dengan bagian yang tetap dapat berdiri sendiri tanpa keseluruhannya.
6		Komposisi (Composition)	Hubungan bagian dan keseluruhan, dengan bagian yang tidak dapat berdiri sendiri apabila keseluruhannya dihapus.

2.13. Black Box Testing

Black Box Testing merupakan metode pengujian perangkat lunak yang berfokus pada pemeriksaan fungsionalitas sistem tanpa memperhatikan struktur kode program di dalamnya. Pengujian dilakukan dengan memberikan

masukan tertentu pada sistem, kemudian membandingkan keluaran yang dihasilkan dengan keluaran yang diharapkan (Rosa & Shalahuddin, 2019).

Metode ini banyak digunakan karena dapat dilakukan tanpa memerlukan pengetahuan mengenai implementasi internal sistem, serta efektif untuk menemukan kesalahan pada fungsi, antarmuka, maupun proses masukan dan keluaran. Pada penelitian ini, Black Box Testing digunakan untuk menguji fungsi *login*, unggah *dataset*, proses analisis sentimen, penyajian hasil analisis, dan *logout* pada sistem yang dibangun.

2.14. Penelitian Terdahulu

Penelitian terdahulu digunakan sebagai acuan dalam penyusunan penelitian ini. Melalui kajian terhadap penelitian sebelumnya, peneliti dapat mengetahui metode yang telah digunakan, hasil yang diperoleh, serta menemukan perbedaan dan kebaruan penelitian yang akan dilakukan. Ringkasan penelitian terdahulu yang relevan dengan penelitian ini disajikan pada Tabel 2.6.

Tabel 2.6 Penelitian Terdahulu

N o.	Peneliti	Judul Penelitian	Metode	Hasil Penelitian	Kelebihan / Kekurangan	Research Gap dengan Penelitian Ini
1	Handayanto et al. (2021)	Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine	Naïve Bayes, Support Vector Machine	Menggunakan 2.143 data Google Review Summarecon Mall Bekasi. Naïve Bayes memperoleh akurasi 80,95%, sedangkan SVM	Kelebihan: Membandingkan dua algoritma sehingga diketahui performanya. Kekurangan: Hanya menggunakan Google Review.	Menggunakan data ulasan Google Play Store pada masa transisi kebijakan KRIS serta mengimplementasikan hasil klasifikasi ke dalam sistem informasi berbasis web.

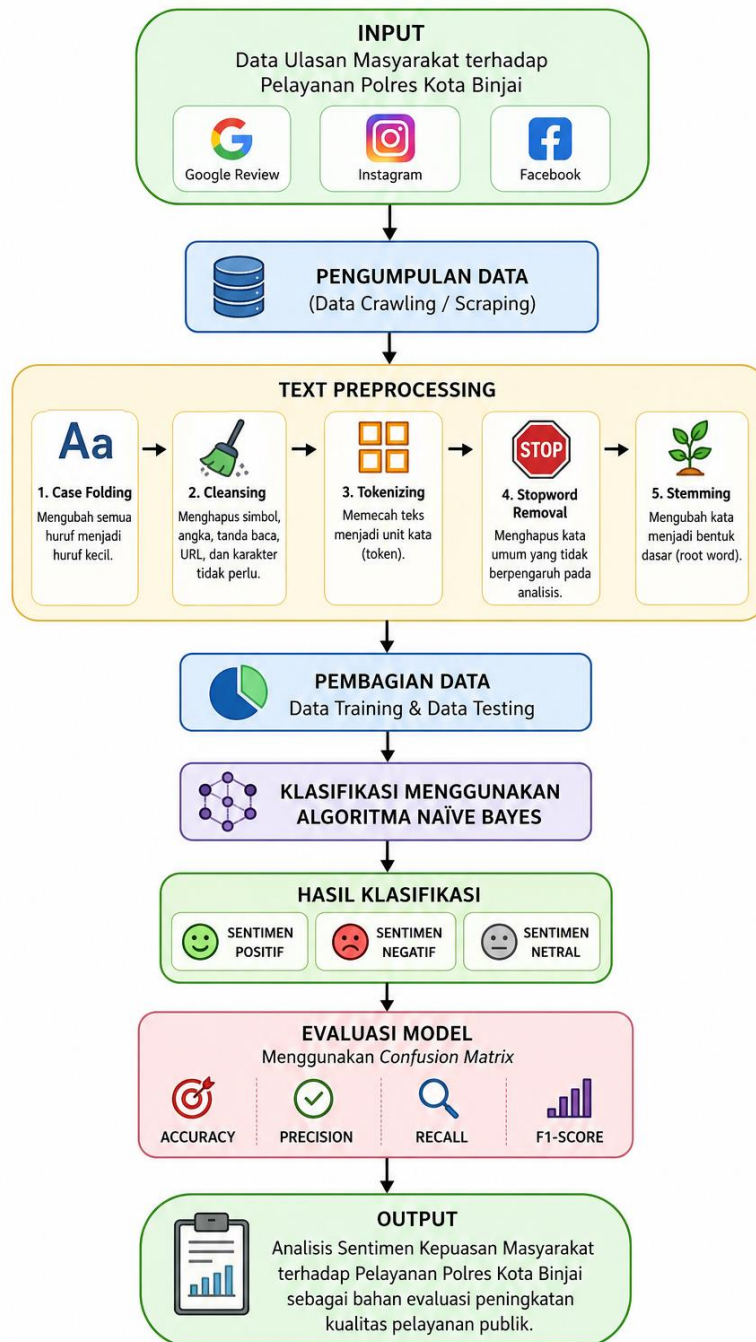
				mencapai 100%.		
2	Rahmawati et al. (2024)	Analisis Sentimen Menggunakan Algoritma Naïve Bayes pada Aplikasi m-BCA Berdasarkan Ulasan Pengguna di Google Play Store	Naïve Bayes	Akurasi 82%, precision 84%, recall 82%, F1-score 81%.	Kelebihan: Efektif pada ulasan aplikasi digital. Kekurangan: Objek terbatas pada m-BCA.	Diterapkan pada layanan publik bidang kesehatan dengan cakupan peserta nasional, serta menghasilkan sistem informasi yang dapat digunakan berulang.
3	Hidayat et al. (2024)	Analisis Sentimen Layanan Sistem Informasi Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes	Naïve Bayes	Accuracy 80,06%, precision 83,11%, recall 75,21%.	Kelebihan: Memberikan rekomendasi peningkatan layanan akademik. Kekurangan: Hanya layanan akademik.	Menggunakan ulasan pengguna aplikasi layanan kesehatan nasional dengan cakupan peserta yang jauh lebih luas serta karakteristik data ulasan yang berbeda.
4	Dzulkarnain et al. (2024)	Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Penilaian Masyarakat terhadap Pelayanan Rumah Sakit di Malang	Naïve Bayes, TF-IDF, Cross Validation	Akurasi 90% pada data uji dan 82,97% pada cross validation.	Kelebihan: Mengelompokkan aspek pelayanan. Kekurangan: Fokus rumah sakit.	Mengkaji pelayanan aplikasi Mobile JKN pada masa transisi kebijakan Kelas Rawat Inap Standar.
5	Berliansyah et al. (2024)	Analisis Sentimen Pada Media Sosial Instagram Terhadap	Naïve Bayes Classifier	1000 komentar Instagram; accuracy 84,1%, precision 85,5%,	Kelebihan: Efektif pada Instagram. Kekurangan: Satu media sosial.	Menggunakan ulasan Google Play Store yang disertai rating sebagai pembanding pelabelan, bukan

		Akun Presiden Joko Widodo Menggunakan Metode Naïve Bayes Classifier		recall 92,4%.		komentar media sosial.
6	Hartini et al. (2024)	Naïve Bayes Classifier untuk Analisis Sentimen Ulasan Pelanggan pada Domo Coffee and Resto	Naïve Bayes Classifier	Akurasi 98,22%, precision 97,58%, recall 100%, F1-score 98,78%.	Kelebihan: Evaluasi berdasarkan aspek makanan, suasana, pelayanan. Kekurangan: Sektor kuliner.	Berfokus pada kepuasan masyarakat terhadap pelayanan publik.
7	Putra et al. (2025)	Pemanfaatan Multinomial Naïve Bayes untuk Analisis Sentimen Ulasan Aplikasi Mobile JKN	Multinomial Naïve Bayes	Akurasi 85,5%; sangat baik mengenali ulasan negatif, kurang baik pada kelas netral dan positif.	Kelebihan: Objek penelitian sama, yaitu aplikasi Mobile JKN. Kekurangan: Kelas netral sulit dikenali.	Menggunakan data pada masa transisi kebijakan KRIS serta mengimplementasikan hasil klasifikasi ke dalam sistem informasi berbasis web.
8	Paliling & Sutoyo (2026)	Analisis Sentimen Ulasan Mobile JKN Menggunakan TF-IDF dan Logistic Regression	TF-IDF, Logistic Regression	5.000 ulasan; accuracy 92,15%, precision 96,42%, recall 89,79%, F1-score 92,99%.	Kelebihan: Memetakan keluhan per kategori layanan, terutama login dan OTP. Kekurangan: Tidak menggunakan Naïve Bayes.	Menggunakan algoritma Naïve Bayes pada data masa transisi KRIS serta membangun sistem informasi berbasis web.

2.15. Kerangka Berpikir

Kerangka berpikir merupakan gambaran alur penelitian yang digunakan sebagai pedoman dalam pelaksanaan penelitian. Kerangka berpikir menunjukkan

hubungan antara permasalahan penelitian, teori yang digunakan, tahapan analisis, hingga hasil yang diharapkan. Dalam penelitian ini, kerangka berpikir disusun berdasarkan proses analisis sentimen menggunakan algoritma Naïve Bayes terhadap data yang diperoleh dari Google Play Store terkait pelayanan aplikasi Mobile JKN.



Gambar 2.1 Kerangka Berpikir

BAB III

METODOLOGI PENELITIAN

3.1. Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan dengan mengambil objek penelitian berupa aplikasi Mobile JKN yang dikembangkan oleh Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan. Oleh karena objek penelitian berupa aplikasi digital, penelitian ini tidak terikat pada lokasi fisik tertentu, melainkan dilaksanakan terhadap data ulasan pengguna yang dipublikasikan secara terbuka pada layanan Google Play Store.

Pemilihan objek penelitian didasarkan pada luasnya cakupan program Jaminan Kesehatan Nasional yang mencapai 98,62 persen penduduk Indonesia, serta pada posisi Mobile JKN sebagai kanal pelayanan digital utama program tersebut. Ulasan yang diberikan pengguna pada Google Play Store memuat berbagai opini, pengalaman, kritik, maupun apresiasi terhadap pelayanan yang diberikan, sehingga dapat dimanfaatkan sebagai sumber data dalam analisis sentimen.

Pengambilan data dilakukan terhadap ulasan pada periode 15 Juli sampai dengan 15 Agustus 2026, yaitu periode yang mencakup masa pemberlakuan kebijakan Kelas Rawat Inap Standar. Penelitian secara keseluruhan dilaksanakan pada periode Februari 2026 sampai dengan Agustus 2026, meliputi identifikasi permasalahan, studi literatur, pengumpulan data, *text preprocessing*, pelabelan data, pembangunan model klasifikasi menggunakan algoritma Naïve Bayes, evaluasi model menggunakan *confusion matrix*, implementasi hasil penelitian ke dalam aplikasi berbasis web, serta penyusunan laporan skripsi.

3.2. Pendekatan dan Tahapan Penelitian

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. Pendekatan kuantitatif dipilih karena penelitian berfokus pada pengolahan data teks menjadi informasi yang dapat diukur dan dievaluasi menggunakan algoritma klasifikasi serta metrik pengujian kinerja model.

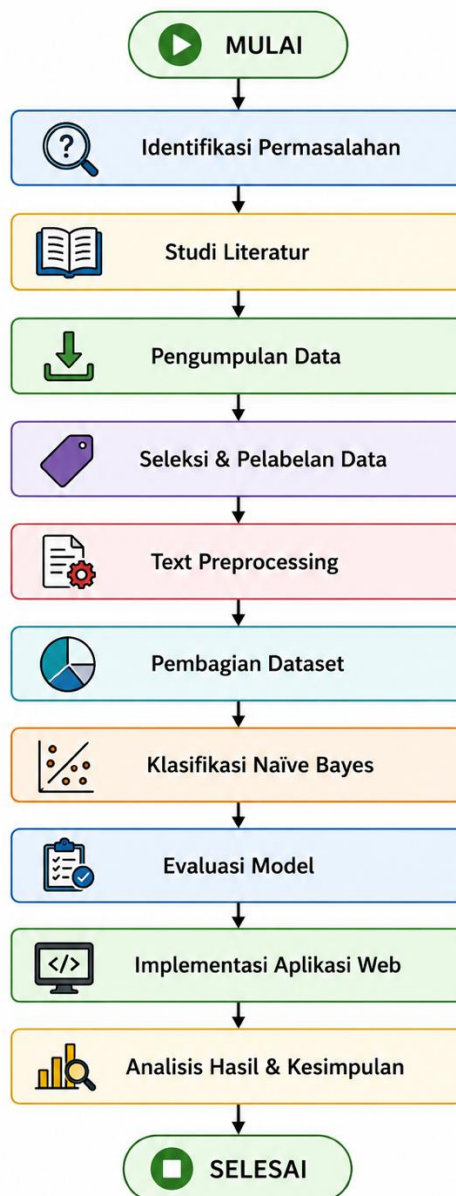
Tahapan penelitian disusun secara sistematis agar proses analisis sentimen dapat dilakukan secara terarah. Penelitian diawali dengan pengumpulan data ulasan pengguna dari Google Play Store yang berkaitan dengan aplikasi Mobile JKN. Data yang telah diperoleh kemudian diseleksi untuk memastikan bahwa hanya ulasan yang relevan dengan objek penelitian yang digunakan pada tahap selanjutnya.

Selanjutnya dilakukan proses *text preprocessing* yang bertujuan untuk membersihkan dan menyeragamkan data teks sebelum dilakukan klasifikasi. Tahapan ini meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Data hasil *preprocessing* kemudian diubah menjadi representasi numerik menggunakan pembobotan TF-IDF sebelum diproses oleh algoritma klasifikasi.

Setelah proses *preprocessing* selesai, setiap data diberikan label sentimen sebagai data acuan (*ground truth*), kemudian dibagi menjadi training data dan testing data. Training data digunakan untuk membangun model klasifikasi menggunakan algoritma Naïve Bayes, sedangkan testing data digunakan untuk menguji kemampuan model dalam mengklasifikasikan

sentimen ke dalam kategori positif dan negatif. Hasil klasifikasi kemudian dievaluasi menggunakan *confusion matrix* dengan menghitung nilai *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil evaluasi tersebut digunakan untuk mengetahui tingkat kinerja algoritma Naïve Bayes dalam melakukan klasifikasi sentimen terhadap data ulasan masyarakat.

Tahap akhir penelitian adalah mengimplementasikan model analisis sentimen ke dalam sebuah aplikasi berbasis web. Aplikasi tersebut dirancang untuk memudahkan proses pengelolaan data ulasan, pelaksanaan analisis sentimen, serta penyajian hasil klasifikasi dalam bentuk tabel maupun visualisasi sehingga hasil penelitian dapat dimanfaatkan sebagai bahan evaluasi terhadap kualitas pelayanan aplikasi Mobile JKN.



Gambar 3.1 Tahapan Penelitian

Berdasarkan Gambar 3.1, penelitian diawali dengan proses identifikasi permasalahan mengenai kepuasan masyarakat terhadap pelayanan aplikasi Mobile JKN berdasarkan ulasan yang dipublikasikan pada Google Play Store. Setelah permasalahan ditetapkan, dilakukan studi literatur sebagai dasar dalam memahami konsep analisis sentimen, Natural Language Processing (NLP), pembobotan TF-

IDF, algoritma Naïve Bayes, serta metode evaluasi menggunakan *confusion matrix*.

Tahap berikutnya adalah pengumpulan data ulasan pengguna dari Google Play Store. Data yang telah diperoleh kemudian melalui proses *text preprocessing* untuk membersihkan dan menyeragamkan data sebelum dilakukan proses analisis. Selanjutnya dilakukan pembobotan TF-IDF, pelabelan data, serta pembagian data menjadi training data dan testing data.

Model klasifikasi kemudian dibangun menggunakan algoritma Naïve Bayes. Setelah proses klasifikasi selesai, dilakukan evaluasi model menggunakan *confusion matrix* dengan menghitung nilai *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil analisis selanjutnya diimplementasikan ke dalam aplikasi berbasis web sehingga pengguna dapat melakukan analisis sentimen secara lebih mudah. Tahap terakhir penelitian adalah melakukan analisis terhadap hasil yang diperoleh serta menyusun kesimpulan berdasarkan tujuan penelitian.

3.3. Data Penelitian

Data merupakan komponen utama dalam penelitian analisis sentimen karena menjadi dasar dalam proses pelatihan dan pengujian model klasifikasi. Pada penelitian ini, data yang digunakan berupa ulasan pengguna mengenai aplikasi Mobile JKN yang diperoleh dari Google Play Store. Data tersebut selanjutnya diolah melalui beberapa tahapan, mulai dari *text preprocessing*, pembobotan TF-IDF, pelabelan data, hingga proses klasifikasi menggunakan algoritma Naïve Bayes.

Berdasarkan sumbernya, data yang digunakan dalam penelitian ini dibedakan menjadi data primer dan data sekunder.

3.3.1. Data Primer

Data primer merupakan data utama yang digunakan dalam penelitian. Data ini diperoleh secara langsung dari ulasan pengguna mengenai aplikasi Mobile JKN yang dipublikasikan pada Google Play Store. Pengambilan data dilakukan dengan memperhatikan keterkaitan ulasan terhadap pelayanan aplikasi sehingga data yang diperoleh relevan dengan tujuan penelitian.

Data yang berhasil dikumpulkan selanjutnya diseleksi untuk menghilangkan ulasan yang tidak berkaitan dengan objek penelitian, ulasan duplikat, ulasan yang terlalu pendek, maupun ulasan yang tidak dapat diproses. Setelah proses seleksi selesai, data digunakan sebagai *dataset* untuk tahapan *text preprocessing*, pelabelan sentimen, proses pelatihan (training), dan pengujian (testing) model klasifikasi.

3.3.2. Data Sekunder

Data sekunder merupakan data pendukung yang digunakan untuk memperkuat landasan teori dan mendukung proses penelitian. Data ini diperoleh dari berbagai sumber, seperti buku, artikel ilmiah, jurnal nasional maupun internasional, prosiding, serta referensi lain yang berkaitan dengan analisis sentimen, Natural Language Processing (NLP), *text mining*, pembobotan TF-IDF, algoritma Naïve Bayes, dan metode evaluasi menggunakan *confusion matrix*. Selain itu, dokumentasi resmi mengenai program Jaminan Kesehatan Nasional dan aplikasi Mobile JKN juga digunakan sebagai informasi pendukung dalam penelitian ini.

Data sekunder dimanfaatkan sebagai acuan dalam penyusunan metodologi penelitian, proses analisis, serta pembahasan hasil penelitian sehingga dapat mendukung keabsahan dan kualitas penelitian yang dilakukan.

3.4. Teknik Pengumpulan Data

Teknik pengumpulan data merupakan tahapan untuk memperoleh data yang akan digunakan dalam proses analisis sentimen. Pada penelitian ini, data dikumpulkan dari satu sumber, yaitu ulasan pengguna aplikasi Mobile JKN pada Google Play Store. Seluruh data yang diperoleh berupa ulasan dalam bentuk teks yang selanjutnya digunakan sebagai *dataset* penelitian.

3.4.1. Pengumpulan Data Ulasan Google Play Store

Data diperoleh dari ulasan pengguna terhadap aplikasi Mobile JKN yang tersedia pada Google Play Store. Pengumpulan data dilakukan dengan teknik *web scraping* menggunakan bahasa pemrograman Python dengan bantuan pustaka `google-play-scraper`. Data yang dikumpulkan meliputi isi ulasan, *rating*, tanggal ulasan, serta versi aplikasi yang diperlukan dalam proses analisis sentimen. Pengambilan data dilakukan terhadap ulasan berbahasa Indonesia pada periode 15 Juli sampai dengan 15 Agustus 2026.

3.4.2. Seleksi Dataset

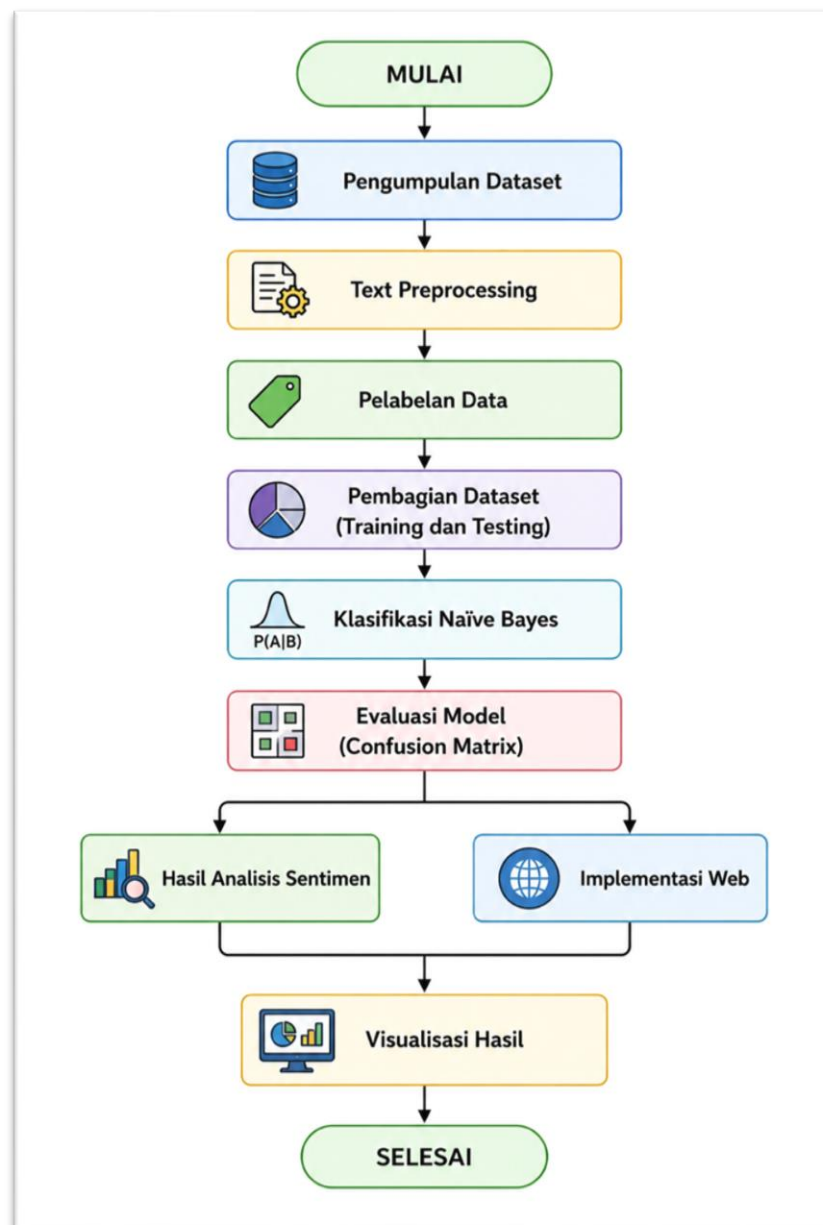
Setelah data berhasil dikumpulkan, dilakukan proses seleksi untuk memastikan bahwa data yang digunakan sesuai dengan tujuan penelitian. Pada tahap ini dilakukan penghapusan terhadap ulasan yang kosong, ulasan duplikat, ulasan yang terdiri atas kurang dari tiga kata, serta ulasan yang tidak menggunakan bahasa Indonesia.

Dari 4.000 ulasan yang berhasil diambil, sebanyak 874 ulasan dibuang karena duplikat, 443 ulasan dibuang karena terdiri atas kurang dari tiga kata, dan 8 ulasan dibuang karena tidak berbahasa Indonesia. Proses seleksi tersebut menghasilkan *dataset* final sebanyak 2.675 ulasan yang selanjutnya digunakan pada tahapan berikutnya.

Proses seleksi *dataset* bertujuan untuk memperoleh kumpulan data yang relevan sehingga dapat digunakan pada tahap pengolahan berikutnya. *Dataset* hasil seleksi selanjutnya akan diproses melalui tahapan *text preprocessing*, pembobotan TF-IDF, pelabelan sentimen, pembagian data, klasifikasi menggunakan algoritma Naïve Bayes, dan evaluasi model.

3.5. Alur Pengolahan Sentimen

Alur pengolahan sentimen merupakan tahapan utama dalam penelitian ini yang digunakan untuk mengolah data ulasan masyarakat menjadi informasi mengenai kecenderungan sentimen terhadap pelayanan aplikasi Mobile JKN. Data yang telah diperoleh dari Google Play Store diproses secara bertahap mulai dari *text preprocessing*, pelabelan sentimen, pembagian *dataset*, proses klasifikasi menggunakan algoritma Naïve Bayes, hingga evaluasi model menggunakan *confusion matrix*. Seluruh tahapan tersebut dilakukan secara sistematis agar model klasifikasi yang dihasilkan mampu memberikan hasil analisis yang akurat dan sesuai dengan tujuan penelitian. Alur pengolahan sentimen pada penelitian ini ditunjukkan pada Gambar 3.2.



Gambar 3.2 Alur Pengolahan Sentimen

3.5.1. Text Preprocessing

Tahap *text preprocessing* merupakan proses awal pengolahan data teks yang bertujuan untuk membersihkan dan menyeragamkan data sebelum dilakukan proses klasifikasi. Proses ini dilakukan agar data yang digunakan memiliki kualitas yang lebih baik sehingga dapat meningkatkan kinerja algoritma Naïve Bayes dalam melakukan klasifikasi sentimen. Pada

penelitian

ini, tahapan *text preprocessing* meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*.

a. *Cleaning*

Tahap *cleaning* dilakukan untuk menghapus karakter yang tidak diperlukan, seperti tanda baca, angka, simbol, emoji, tautan (*URL*), serta karakter khusus lainnya yang tidak memiliki pengaruh terhadap proses analisis sentimen. Proses ini bertujuan agar data teks menjadi lebih bersih dan mudah diproses pada tahapan berikutnya.

b. *Case Folding*

Tahap *case folding* bertujuan untuk mengubah seluruh huruf pada teks menjadi huruf kecil (*lowercase*). Proses ini dilakukan agar tidak terjadi perbedaan penulisan kata akibat penggunaan huruf kapital dan huruf kecil.

c. *Tokenizing*

Tahap *tokenizing* dilakukan dengan memecah kalimat menjadi kumpulan kata (*token*). Setiap kata yang dihasilkan akan diproses secara individual pada tahapan selanjutnya sehingga memudahkan proses klasifikasi.

d. *Stopword Removal*

Tahap *stopword removal* dilakukan untuk menghilangkan kata-kata umum yang tidak memberikan informasi penting terhadap proses analisis sentimen, seperti "dan", "yang", "di", "ke", serta kata-kata lain yang tidak memiliki bobot makna.

e. *Stemming*

Tahap *stemming* bertujuan untuk mengubah kata yang memiliki imbuhan menjadi bentuk kata dasar. Dengan proses ini, variasi kata dapat dikurangi sehingga kata yang memiliki makna sama dapat dikenali sebagai satu bentuk kata.

3.5.2. Pelabelan Sentimen

Setelah melalui proses *text preprocessing*, setiap ulasan diberi label sentimen. Pelabelan dilakukan menggunakan dua pendekatan yang saling melengkapi, yaitu pendekatan berbasis *rating* pengguna dan pendekatan *lexicon based* menggunakan kamus sentimen bahasa Indonesia. Ulasan yang kedua pendekatan tersebut menghasilkan label sama dilabeli secara otomatis, sedangkan ulasan yang kedua pendekatannya berbeda diverifikasi secara manual oleh peneliti. Tingkat kesepakatan kedua pendekatan diukur menggunakan Cohen's Kappa (Cohen, 1960). Ulasan yang teridentifikasi bersifat informatif atau berupa pertanyaan dikeluarkan dari *dataset* setelah tahap pelabelan, sehingga analisis difokuskan pada ulasan yang menyatakan penilaian, yaitu positif dan negatif. Hasil pelabelan ini digunakan sebagai *ground truth* dalam proses pelatihan dan pengujian model klasifikasi.

3.5.3. Pembagian Dataset

Dataset yang telah diberi label kemudian dibagi menjadi **data latih** (*training data*) dan **data uji** (*testing data*). Data latih digunakan untuk membangun model klasifikasi menggunakan algoritma Naïve Bayes, sedangkan data uji digunakan untuk mengukur kemampuan model dalam

mengklasifikasikan sentimen terhadap data yang belum pernah dipelajari sebelumnya.

3.5.4. Klasifikasi Menggunakan Algoritma Naïve Bayes

Pada tahap ini dilakukan proses klasifikasi menggunakan algoritma Naïve Bayes. Model dibangun berdasarkan data latih yang telah memiliki label sentimen. Selanjutnya model digunakan untuk memprediksi kategori sentimen pada data uji ke dalam dua kelas, yaitu positif dan negatif. Hasil prediksi kemudian dibandingkan dengan label sebenarnya untuk mengetahui tingkat ketepatan model.

3.5.5. Evaluasi Model

Tahap evaluasi dilakukan untuk mengetahui kinerja model klasifikasi yang telah dibangun. Evaluasi menggunakan metode *confusion matrix* dengan menghitung nilai *accuracy*, *precision*, *recall*, dan *F1-score*. Nilai-nilai tersebut digunakan sebagai indikator dalam menilai kemampuan algoritma Naïve Bayes dalam mengklasifikasikan sentimen masyarakat terhadap pelayanan aplikasi Mobile JKN.

3.5.6. Visualisasi Hasil Analisis

Setelah proses klasifikasi dan evaluasi model selesai dilakukan, hasil analisis sentimen disajikan dalam bentuk visualisasi agar lebih mudah dipahami oleh pengguna. Visualisasi hasil bertujuan untuk memberikan gambaran mengenai distribusi sentimen masyarakat terhadap pelayanan aplikasi Mobile JKN berdasarkan data yang telah dianalisis menggunakan algoritma Naïve Bayes.

Pada penelitian ini, visualisasi hasil analisis ditampilkan dalam aplikasi berbasis web yang menyajikan informasi berupa tabel hasil klasifikasi, grafik distribusi sentimen, serta nilai evaluasi model yang diperoleh dari *confusion matrix*. Penyajian informasi dalam bentuk visual diharapkan dapat membantu pengguna dalam memahami kecenderungan sentimen masyarakat secara lebih efektif serta mendukung proses evaluasi terhadap kualitas pelayanan aplikasi Mobile JKN.

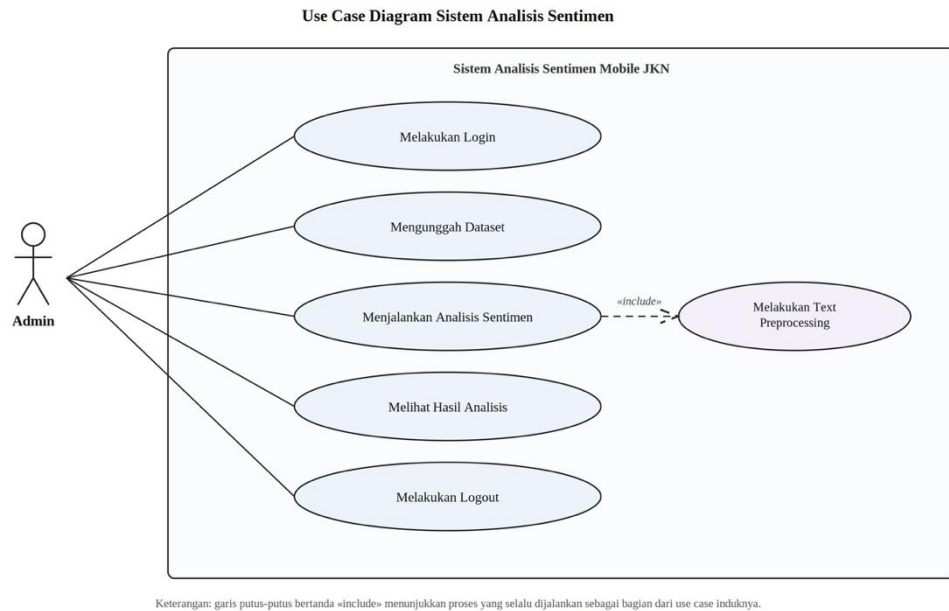
3.6. Perancangan Sistem

Perancangan sistem merupakan tahapan yang dilakukan untuk menggambarkan alur kerja aplikasi yang akan dibangun. Aplikasi berbasis web pada penelitian ini dirancang sebagai media untuk mengelola data ulasan, melakukan proses analisis sentimen menggunakan algoritma Naïve Bayes, serta menampilkan hasil analisis dalam bentuk tabel dan visualisasi. Perancangan sistem dilakukan menggunakan Unified Modeling Language (UML) yang terdiri atas Use Case Diagram, Activity Diagram, Sequence Diagram, dan Class Diagram. Adapun rancangan basis data disajikan secara terpisah pada subbab 3.7 dalam bentuk Entity Relationship Diagram (ERD), karena ERD bukan merupakan bagian dari notasi UML.

3.6.1. Use Case Diagram

Use Case Diagram digunakan untuk menggambarkan interaksi antara pengguna dengan sistem yang dibangun. Pada penelitian ini hanya terdapat satu aktor, yaitu Admin, yang bertugas mengelola seluruh proses analisis sentimen. Admin dapat melakukan *login* ke dalam sistem, mengunggah *dataset*, menjalankan proses analisis sentimen, melihat hasil analisis, dan

keluar dari sistem (*logout*). Rancangan Use Case Diagram ditunjukkan pada Gambar 3.3.



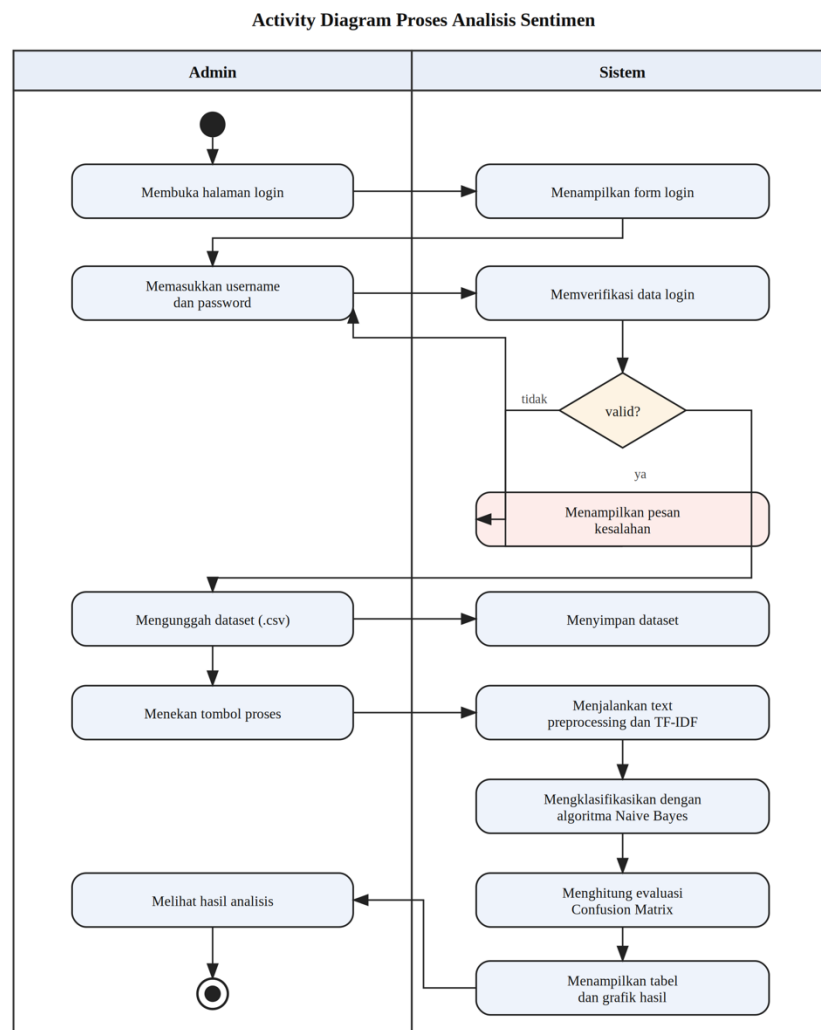
Gambar 3.3 Use Case Diagram

Adapun aktivitas yang dapat dilakukan oleh Admin meliputi:

1. *Login* ke dalam sistem.
2. Mengunggah *dataset*.
3. Menjalankan proses *text preprocessing*.
4. Melakukan analisis sentimen menggunakan algoritma Naïve Bayes.
5. Melihat hasil analisis sentimen.
6. *Logout* dari sistem.

3.6.2. Activity Diagram

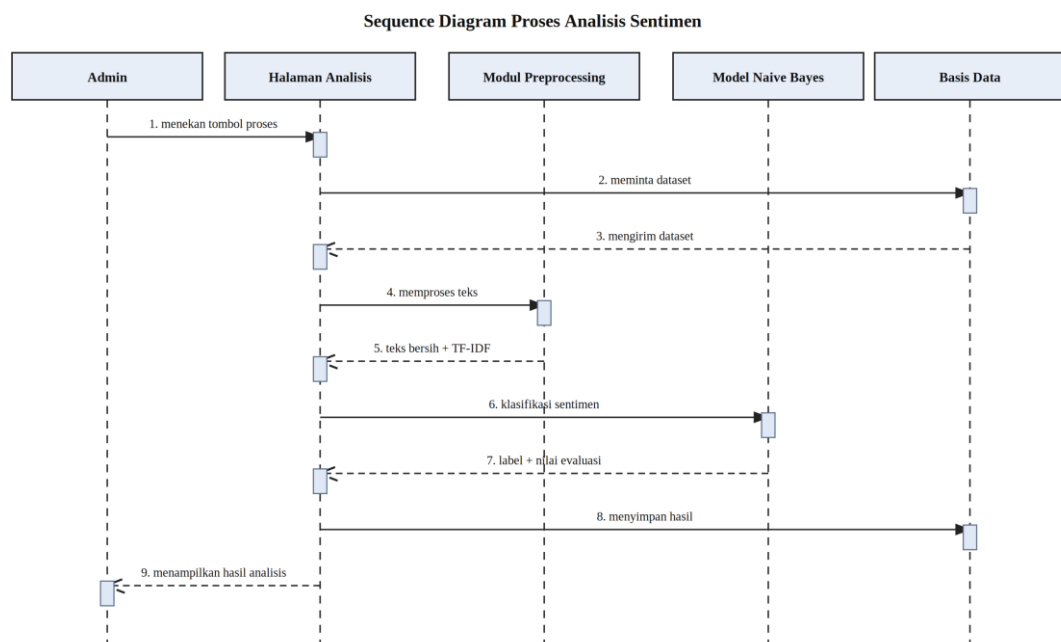
Activity Diagram digunakan untuk menggambarkan alur aktivitas pengguna dalam menggunakan aplikasi analisis sentimen. Diagram ini menunjukkan urutan proses yang dilakukan mulai dari pengguna masuk ke sistem hingga memperoleh hasil analisis sentimen. Rancangan Activity Diagram ditunjukkan pada Gambar 3.4.



Gambar 3.4 Activity Diagram

3.6.3. Sequence Diagram

Sequence Diagram digunakan untuk menggambarkan urutan pesan yang dipertukarkan antarobjek di dalam sistem berdasarkan urutan waktu. Diagram ini menunjukkan alur proses analisis sentimen mulai dari Admin menekan tombol proses, pengambilan *dataset* dari basis data, pemrosesan teks dan pembobotan TF-IDF, klasifikasi menggunakan algoritma Naïve Bayes, hingga penyimpanan dan penyajian hasil analisis. Rancangan Sequence Diagram ditunjukkan pada Gambar 3.5.



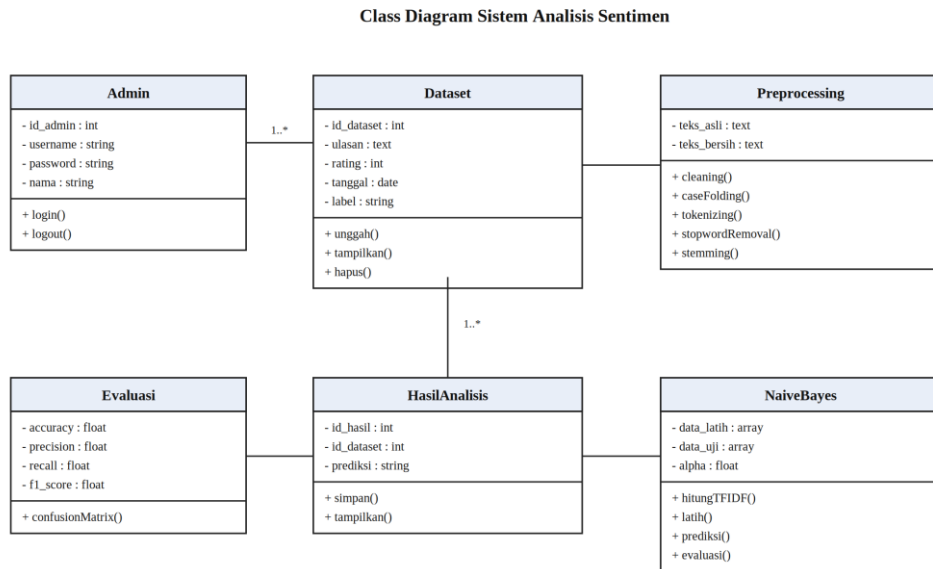
Gambar 3.5 Sequence Diagram

3.6.4. Class Diagram

Class Diagram digunakan untuk menggambarkan struktur kelas yang membentuk sistem beserta atribut, metode, dan relasi antarkelas. Sistem yang dibangun terdiri atas enam kelas utama, yaitu Admin, *Dataset*, *Preprocessing*, NaiveBayes, HasilAnalisis, dan Evaluasi. Tanda minus menunjukkan atribut

bersifat private, sedangkan tanda plus menunjukkan metode bersifat public.

Rancangan Class Diagram ditunjukkan pada Gambar 3.6.

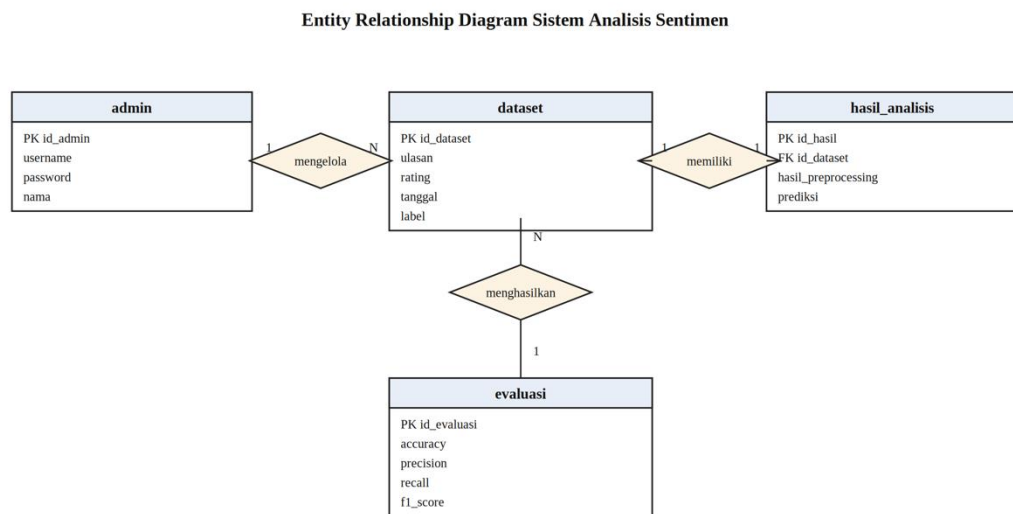


Relasi digambarkan sebagai asosiasi dengan multiplisitas pada ujung garis.

Gambar 3.6 Class Diagram

3.7. Entity Relationship Diagram (ERD)

Entity Relationship Diagram (ERD) digunakan untuk menggambarkan struktur basis data yang digunakan pada aplikasi analisis sentimen. ERD menggunakan notasi tersendiri dan bukan merupakan bagian dari Unified Modeling Language, sehingga disajikan pada subbab terpisah. Basis data dirancang sesuai kebutuhan sistem, yaitu untuk menyimpan data pengguna, *dataset* ulasan, hasil analisis sentimen, dan nilai evaluasi model. Rancangan ERD ditunjukkan pada Gambar 3.7.



PK = Primary Key, FK = Foreign Key. Angka pada garis menunjukkan kardinalitas relasi.

Gambar 3.7 Entity Relationship Diagram

Basis data terdiri dari empat entitas utama, yaitu:

1. **Admin**, digunakan untuk menyimpan data pengguna yang memiliki hak akses ke dalam sistem.
2. **Dataset**, digunakan untuk menyimpan data ulasan pengguna yang diperoleh dari Google Play Store.
3. **Hasil Analisis**, digunakan untuk menyimpan hasil klasifikasi sentimen yang dihasilkan oleh algoritma Naïve Bayes.
4. **Evaluasi**, digunakan untuk menyimpan nilai *accuracy*, *precision*, *recall*, dan *F1-score* yang dihasilkan pada proses evaluasi model.

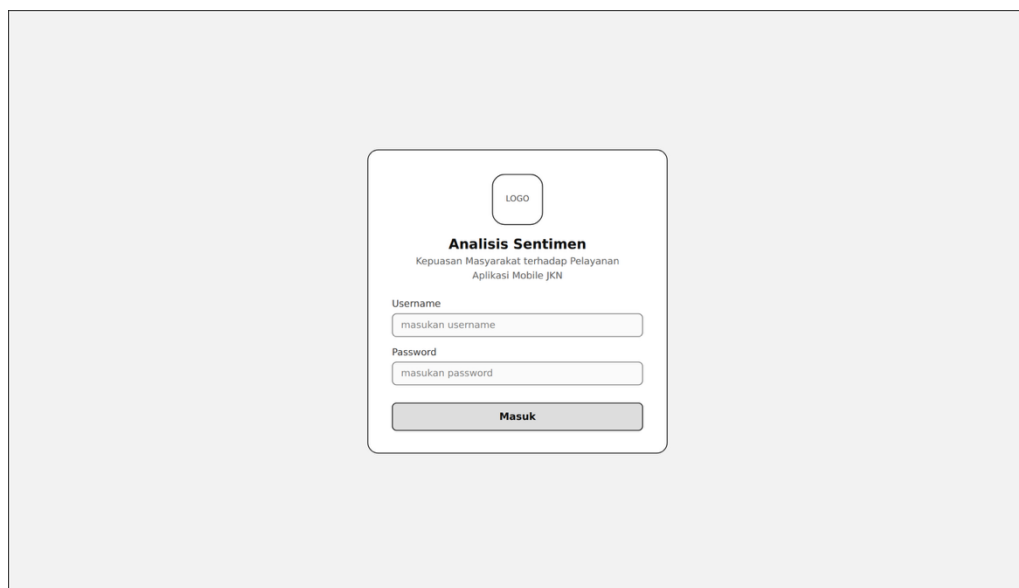
3.8. Perancangan Interface

Perancangan interface bertujuan untuk memberikan gambaran mengenai tampilan aplikasi yang akan dibangun. Antarmuka sistem dirancang sederhana agar memudahkan pengguna dalam melakukan proses analisis sentimen terhadap

ulasan masyarakat mengenai pelayanan aplikasi Mobile JKN. Perancangan interface pada penelitian ini terdiri dari beberapa halaman utama, yaitu halaman *login*, *dashboard*, *dataset*, analisis sentimen, dan hasil analisis.

3.8.1. Halaman Login

Halaman *login* merupakan halaman awal yang digunakan sebagai proses autentikasi pengguna sebelum mengakses sistem. Pada halaman ini pengguna diminta memasukkan **username** dan **password** yang telah terdaftar agar dapat masuk ke dalam aplikasi.



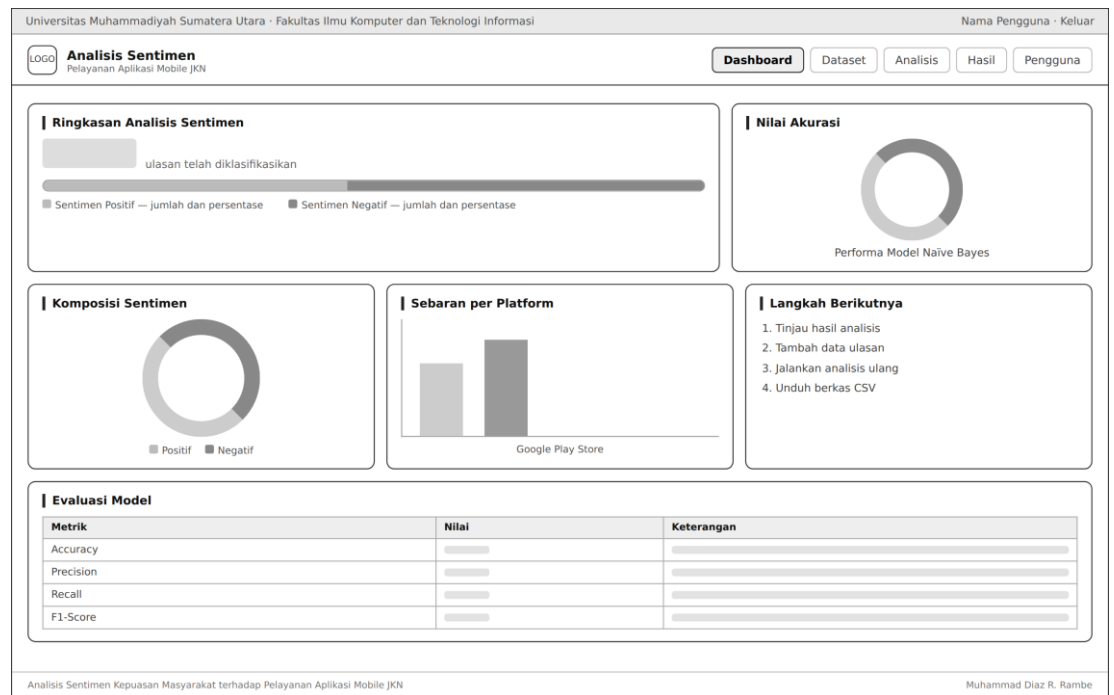
The image shows a login form centered on a light gray background. The form is a white rounded rectangle containing the following elements from top to bottom: a circular logo placeholder labeled 'LOGO'; the title 'Analisis Sentimen' in bold; the subtitle 'Kepuasan Masyarakat terhadap Pelayanan Aplikasi Mobile JKN'; a 'Username' label above a text input field with the placeholder 'masukan username'; a 'Password' label above a text input field with the placeholder 'masukan password'; and a gray button labeled 'Masuk'.

Gambar 3.8 Rancangan Halaman Login

3.8.2. Halaman Dashboard

Halaman *dashboard* merupakan halaman utama setelah pengguna berhasil melakukan *login*. Halaman ini menampilkan informasi ringkas mengenai *dataset* dan hasil analisis sentimen yang telah dilakukan. Informasi yang ditampilkan meliputi:

- a. Jumlah *Dataset*
- b. Sentimen Positif
- c. Sentimen Negatif
- d. Nilai Akurasi



Gambar 3.9 Rancangan Halaman Dashboard

3.8.3. Halaman Dataset

Halaman *dataset* digunakan untuk mengelola data yang akan dianalisis.

Pada halaman ini pengguna dapat mengunggah *dataset* serta melihat data yang berhasil dimasukkan ke dalam sistem. Fitur yang tersedia antara lain:

- a. *Upload Dataset (.csv)*
- b. Menampilkan Daftar Komentar

Universitas Muhammadiyah Sumatera Utara · Fakultas Ilmu Komputer dan Teknologi Informasi
Nama Pengguna · Keluar


Analisis Sentimen
Pelayanan Aplikasi Mobile JKN

Dashboard
Dataset
Analisis
Hasil
Pengguna

Dataset
Kelola data ulasan dari Google Play Store.

Unggah Dataset (.csv)

pilih berkas .csv
Unggah

Kolom wajib: komentar · Kolom opsional: sumber, tanggal, label

Daftar Ulasan

cari ulasan
semua platform
Cari

No	Sumber	Tanggal	Komentar	Label
1				Positif
2				Negatif
3				Negatif
4				Positif
5				Positif

« Sebelumnya · Halaman 1 dari n · Berikutnya »

Analisis Sentimen Kepuasan Masyarakat terhadap Pelayanan Aplikasi Mobile JKN
Muhammad Diaz R. Rambe

Gambar 3.10 Rancangan Halaman Dataset

3.8.4. Halaman Analisis Sentimen

Halaman analisis sentimen digunakan untuk menjalankan proses pengolahan data menggunakan algoritma Naïve Bayes. Pengguna cukup menekan tombol proses sehingga sistem akan melakukan *text preprocessing*, klasifikasi sentimen, dan evaluasi model.

Universitas Muhammadiyah Sumatera Utara · Fakultas Ilmu Komputer dan Teknologi Informasi
Nama Pengguna · Keluar


Analisis Sentimen
Pelayanan Aplikasi Mobile JKN

Dashboard
Dataset
Analisis
Hasil
Pengguna

Analisis Sentimen
Jalankan text preprocessing, klasifikasi Naïve Bayes, dan evaluasi model.

TOTAL DATASET

SUDAH DILABELI

BELUM DILABELI

Sebaran Label

Positif — jumlah
Negatif — jumlah

Jalankan Proses

Pembagian data
80% latih : 20% uji
Alpha (Laplace smoothing)
1.0

Proses Analisis Sentimen

- Text Preprocessing** — cleaning, case folding, tokenizing, stopword removal, stemming
- Pembobotan TF-IDF** — mengubah teks menjadi vektor numerik
- Pembagian Dataset** — data latih dan data uji
- Klasifikasi Naïve Bayes** — Multinomial Naïve Bayes
- Evaluasi** — confusion matrix, accuracy, precision, recall, F1-score

Uji Satu Kalimat

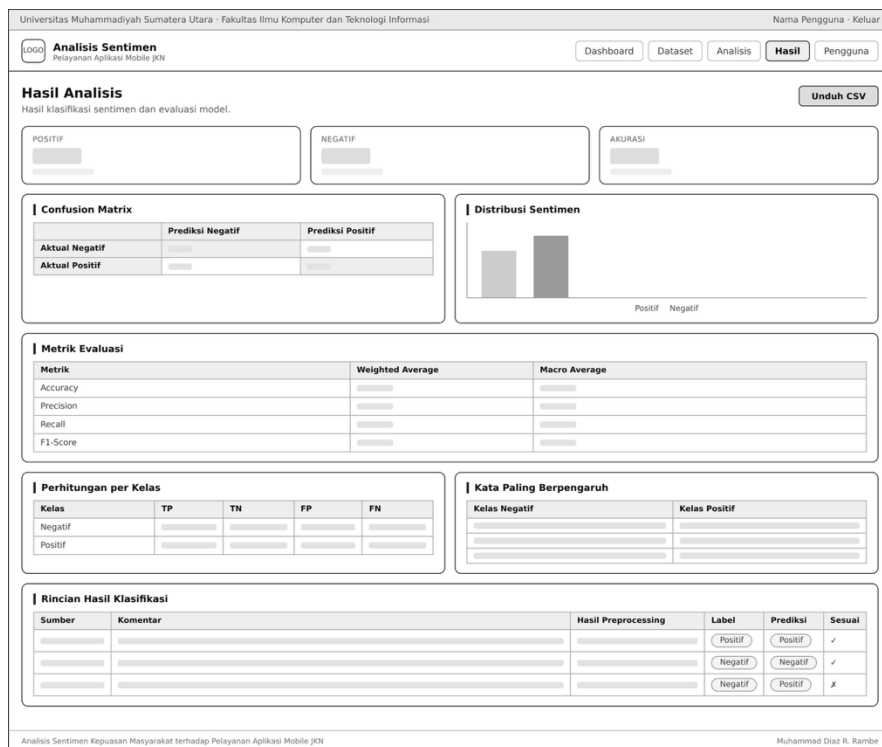
masukan kalimat yang akan diuji
Uji

Analisis Sentimen Kepuasan Masyarakat terhadap Pelayanan Aplikasi Mobile JKN
Muhammad Diaz R. Rambe

Gambar 3.11 Rancangan Halaman Analisis Sentimen

3.8.5. Halaman Hasil Analisis

Halaman hasil analisis merupakan halaman yang menampilkan keluaran akhir proses klasifikasi. Halaman ini dirancang memuat jumlah ulasan pada setiap kategori sentimen, nilai akurasi model, *confusion matrix*, tabel metrik evaluasi, kata paling berpengaruh pada setiap kelas, serta tabel rincian hasil klasifikasi setiap ulasan. Rancangan halaman hasil analisis ditunjukkan pada Gambar 3.12.



Gambar 3.12 Rancangan Halaman Hasil Analisis

3.9. Pengujian Sistem

3.9.1. Pengujian Fungsional

Pengujian fungsional dilakukan menggunakan metode **Black Box Testing**. Pengujian ini bertujuan untuk memastikan bahwa setiap fitur pada aplikasi dapat berjalan sesuai dengan fungsi yang telah dirancang tanpa memperhatikan struktur kode program. Fitur yang diuji meliputi:

- a. *Login*
- b. *Upload Dataset*
- c. Proses Analisis Sentimen
- d. Menampilkan Hasil Analisis
- e. *Logout*

3.9.2. Pengujian Hasil Klasifikasi

Pengujian hasil klasifikasi dilakukan untuk mengetahui tingkat performa algoritma Naïve Bayes dalam mengklasifikasikan sentimen masyarakat terhadap pelayanan aplikasi Mobile JKN. Evaluasi dilakukan menggunakan *Confusion Matrix* dengan menghitung nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Nilai tersebut digunakan sebagai indikator dalam menilai kualitas model klasifikasi yang dibangun.

BAB IV

HASIL DAN PEMBAHASAN

4.1. Hasil Pengumpulan Data

Pengumpulan data dilakukan menggunakan teknik *web scraping* terhadap ulasan pengguna aplikasi Mobile JKN pada Google Play Store. Pengambilan data dilakukan dengan bahasa pemrograman Python menggunakan pustaka *google-play-scraper*. Ulasan diambil berurutan mulai dari ulasan terbaru, dengan periode pengambilan 15 Juli sampai dengan 15 Agustus 2026, yaitu periode yang mencakup masa pemberlakuan kebijakan Kelas Rawat Inap Standar.

Pemeriksaan terhadap halaman aplikasi menunjukkan bahwa aplikasi Mobile JKN memiliki nilai *rating* 4,4 dari 5,0 dengan jumlah ulasan mencapai 1,02 juta dan jumlah pemasangan lebih dari 50 juta. Ketersediaan data tersebut jauh melampaui kebutuhan penelitian ini.

Data mentah yang berhasil dikumpulkan berjumlah 4.000 ulasan. Data tersebut selanjutnya melalui proses seleksi sebagaimana dijelaskan pada BAB III. Hasil seleksi *dataset* disajikan pada Tabel 4.1.

Tabel 4.1 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

Berdasarkan Tabel 4.1, dari 4.000 ulasan mentah diperoleh 2.675 ulasan yang memenuhi kriteria penelitian. Ulasan duplikat dibuang karena tidak memberikan informasi tambahan, ulasan yang terdiri atas kurang dari tiga kata dibuang karena tidak memiliki konteks yang cukup untuk diklasifikasikan, sedangkan ulasan berbahasa asing dibuang karena berada di luar ruang lingkup penelitian.

Sebaran *rating* pada *dataset* hasil seleksi disajikan pada Tabel 4.2.

Tabel 4.2 Sebaran Rating pada Dataset

Rating	Jumlah Ulasan	Persentase
1 bintang	1.348	50,2%
2 bintang	180	6,7%
3 bintang	122	4,5%
4 bintang	108	4,0%
5 bintang	925	34,5%
Total	2.683	100%

Tabel 4.2 menunjukkan bahwa sebaran rating bersifat terpolarisasi, yaitu menumpuk pada nilai ekstrem 1 bintang dan 5 bintang, sedangkan rating tengah sangat sedikit. Karakteristik ini lazim ditemukan pada ulasan aplikasi karena pengguna umumnya menuliskan ulasan ketika merasa sangat puas atau sangat tidak puas.

4.2. Hasil Pengambilan Sampel

Mengingat sebaran kelas yang tidak seimbang, pengambilan sampel untuk pelabelan dilakukan menggunakan metode sampel berimbang, yaitu mengambil jumlah yang setara dari setiap kelompok *rating*. Metode ini dipilih agar setiap kategori sentimen memiliki jumlah data latih yang memadai sehingga model tidak bias terhadap kelas mayoritas. Nilai seed yang digunakan adalah 42 agar pengambilan sampel dapat diulang dengan hasil yang sama.

Tabel 4.3 Alokasi Sampel per Kelompok Rating

Kelompok Rating	Tersedia	Diambil
1-2 bintang (kandidat negatif)	1.527	400
3 bintang (kandidat netral)	122	122
4-5 bintang (kandidat positif)	1.026	400
Total sampel	2.675	922

4.3. Hasil Pelabelan Sentimen

Pelabelan dilakukan menggunakan dua pendekatan otomatis yang saling melengkapi, yaitu pendekatan berbasis *rating* pengguna dan pendekatan lexicon based menggunakan kamus sentimen bahasa Indonesia. Ulasan yang kedua

pendekatannya menghasilkan label sama dilabeli secara otomatis, sedangkan ulasan yang kedua pendekatannya berbeda diverifikasi secara manual oleh peneliti.

Dari 922 ulasan pada sampel, sebanyak 622 ulasan atau 67,5 persen mendapat label yang sama dari kedua pendekatan, sedangkan 300 ulasan atau 32,5 persen menghasilkan label berbeda sehingga diverifikasi manual. Tingkat kesepakatan kedua pendekatan diukur menggunakan Cohen's Kappa dan memperoleh nilai 0,438 yang termasuk kategori sedang.

Nilai kesepakatan yang tidak tinggi tersebut justru menjadi dasar diperlukannya verifikasi manual. Semakin banyak ulasan yang tidak disepakati kedua pendekatan, semakin banyak pula ulasan yang dibaca dan diputuskan langsung oleh peneliti. Label akhir terdiri atas label yang disepakati kedua pendekatan ditambah label hasil keputusan manual peneliti.

Hasil pelabelan terhadap 922 ulasan disajikan pada Tabel 4.4.

Tabel 4.4 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%

Pemeriksaan terhadap ulasan berlabel netral menunjukkan bahwa sebagian besar berupa pertanyaan atau pernyataan informatif yang tidak menyatakan penilaian terhadap layanan, misalnya pertanyaan mengenai cara mengubah fasilitas kesehatan atau cara mendaftarkan anggota keluarga. Kalimat semacam ini tidak menyatakan kepuasan maupun ketidakpuasan.

Sesuai ruang lingkup penelitian yang berfokus pada polaritas kepuasan, ulasan berlabel netral dan ulasan yang tidak dilabeli dikeluarkan dari *dataset*

setelah tahap pelabelan. Perlu ditegaskan bahwa pengeluaran tersebut dilakukan setelah seluruh ulasan dibaca dan dinilai, sehingga merupakan keputusan berbasis hasil pelabelan, bukan asumsi awal. *Dataset* akhir yang digunakan pada tahap analisis berjumlah 751 ulasan, terdiri atas 403 ulasan negatif dan 348 ulasan positif.

4.4. Hasil Text Preprocessing

Seluruh ulasan pada *dataset* akhir diproses melalui lima tahapan *text preprocessing*, yaitu *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Contoh hasil setiap tahapan disajikan pada Tabel 4.5.

Tabel 4.5 Contoh Hasil Text Preprocessing

Tahapan	Hasil
Teks asli	aplikasi tidak bisa di akses di hp Iqoo
Cleaning	aplikasi tidak bisa di akses di hp iqoo
Case folding	aplikasi tidak bisa di akses di hp iqoo
Tokenizing	aplikasi, tidak, bisa, di, akses, di, hp, iqoo
Stopword removal	aplikasi, tidak, akses, iqoo
Stemming	aplikasi, tidak, akses, iqoo

Tabel 4.5 memperlihatkan bahwa kata “tidak” tetap dipertahankan pada tahap *stopword removal*. Hal ini dilakukan secara sengaja karena kata pengingkaran menentukan polaritas kalimat. Apabila kata “tidak” dibuang, frasa “tidak bisa diakses” akan berubah menjadi “bisa diakses” sehingga makna sentimennya terbalik.

4.5. Hasil Pembobotan TF-IDF

Data hasil *preprocessing* selanjutnya diubah menjadi representasi numerik menggunakan pembobotan TF-IDF dengan kombinasi *unigram* dan *bigram*. Proses pembobotan menghasilkan 4.004 fitur kata yang digunakan sebagai masukan bagi algoritma klasifikasi. Pembobotan dilakukan hanya pada data latih,

kemudian diterapkan pada data uji, untuk mencegah terjadinya kebocoran data (*data leakage*).

4.6. Hasil Pembagian Dataset

Dataset dibagi menjadi data latih dan data uji dengan perbandingan 80 berbanding 20 menggunakan teknik stratified split, sehingga proporsi setiap kelas pada data latih dan data uji tetap seimbang. Hasil pembagian *dataset* disajikan pada Tabel 4.6.

Tabel 4.6 Pembagian Dataset

Keterangan	Jumlah	Persentase
Data latih	600	80%
Data uji	151	20%
Total	751	100%

4.7. Hasil Klasifikasi Algoritma Naïve Bayes

Model klasifikasi dibangun menggunakan algoritma Multinomial Naïve Bayes dengan nilai alpha sebesar 1,0 atau *Laplace smoothing*. Model dilatih menggunakan 600 data latih, kemudian diuji terhadap 151 data uji. Contoh hasil klasifikasi disajikan pada Tabel 4.7.

Tabel 4.7 Contoh Hasil Klasifikasi

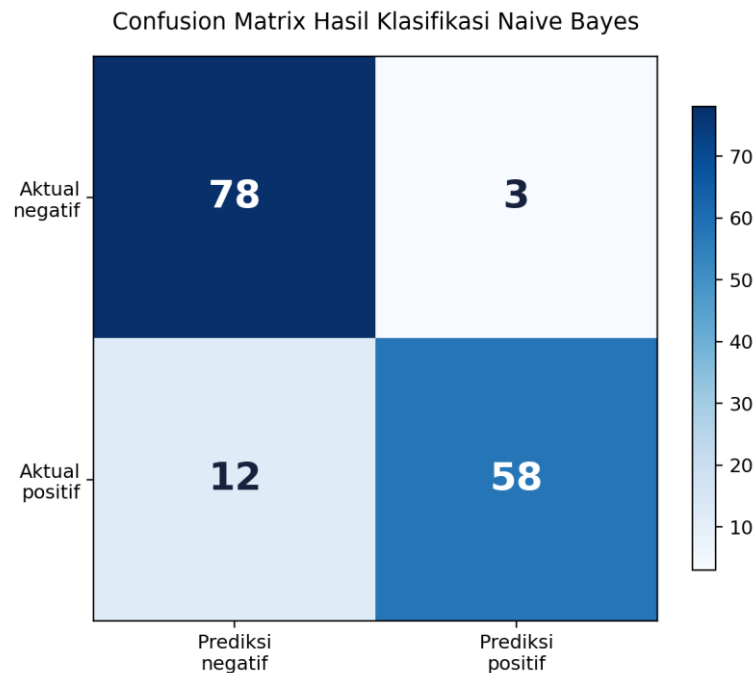
Ulasan	Label Aktual	Prediksi	Status
aplikasi tidak bisa di akses di hp Iqoo	Negatif	Negatif	Benar
sangat bagus dan baik	Positif	Positif	Benar
sangat membantu dalam proses berobat	Positif	Positif	Benar
ini mau login tidak bisa, aplikasi bangke	Negatif	Negatif	Benar
kenapa jkn mobile tidak bisa dibuka ya	Positif	Negatif	Salah

4.8. Hasil Evaluasi Model

Evaluasi model dilakukan menggunakan *Confusion Matrix* terhadap 151 data uji. Hasil *Confusion Matrix* disajikan pada Tabel 4.8 dan Gambar 4.1.

Tabel 4.8 Confusion Matrix

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



Gambar 4.1 Confusion Matrix Hasil Klasifikasi

Berdasarkan Tabel 4.8, sebanyak 78 ulasan negatif berhasil diprediksi dengan benar dan 58 ulasan positif berhasil diprediksi dengan benar. Terdapat 3 ulasan negatif yang salah diprediksi sebagai positif, serta 12 ulasan positif yang salah diprediksi sebagai negatif. Total prediksi benar adalah 136 dari 151 data uji.

Perhitungan nilai *True Positive*, *True Negative*, *False Positive*, dan *False Negative* untuk setiap kelas disajikan pada Tabel 4.9.

Tabel 4.9 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%

Nilai evaluasi model secara keseluruhan disajikan pada Tabel 4.10.

Tabel 4.10 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%

Berdasarkan Tabel 4.10, model memperoleh nilai *accuracy* sebesar 90,07 persen, *precision* 90,57 persen, *recall* 90,07 persen, dan *F1-score* 89,99 persen.

Nilai macro average yang hampir sama dengan weighted average menunjukkan bahwa performa model relatif merata pada kedua kelas, tidak hanya baik pada kelas mayoritas.

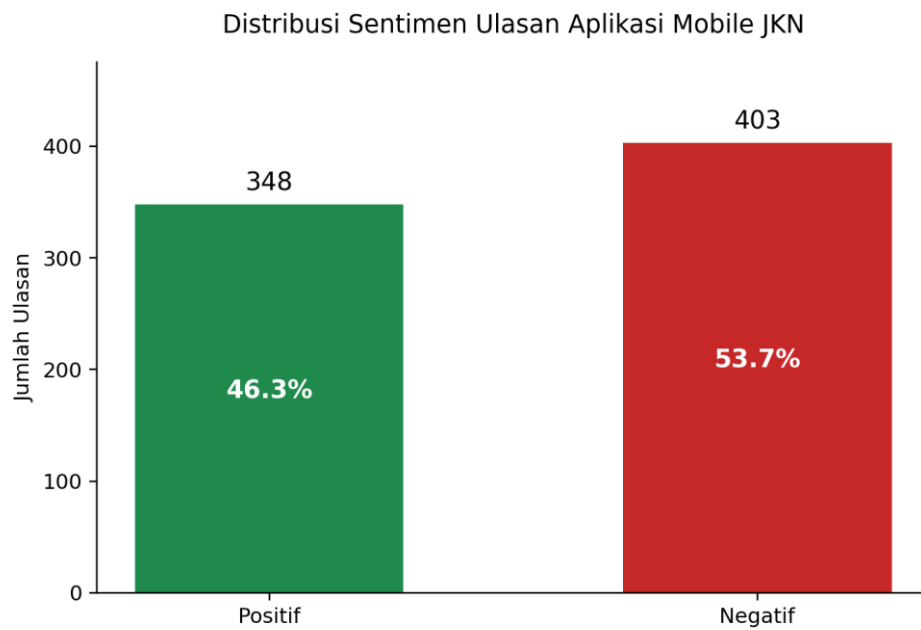
Untuk memastikan bahwa hasil tersebut tidak dipengaruhi oleh pembagian data tertentu, dilakukan pengujian *10-fold cross validation*. Hasilnya disajikan pada Tabel 4.11.

Tabel 4.11 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%
Rata-rata	90,68%	Simpangan baku	2,60%

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

Distribusi hasil klasifikasi sentimen terhadap keseluruhan *dataset* disajikan pada Gambar 4.2.



Gambar 4.2 Distribusi Sentimen Hasil Klasifikasi

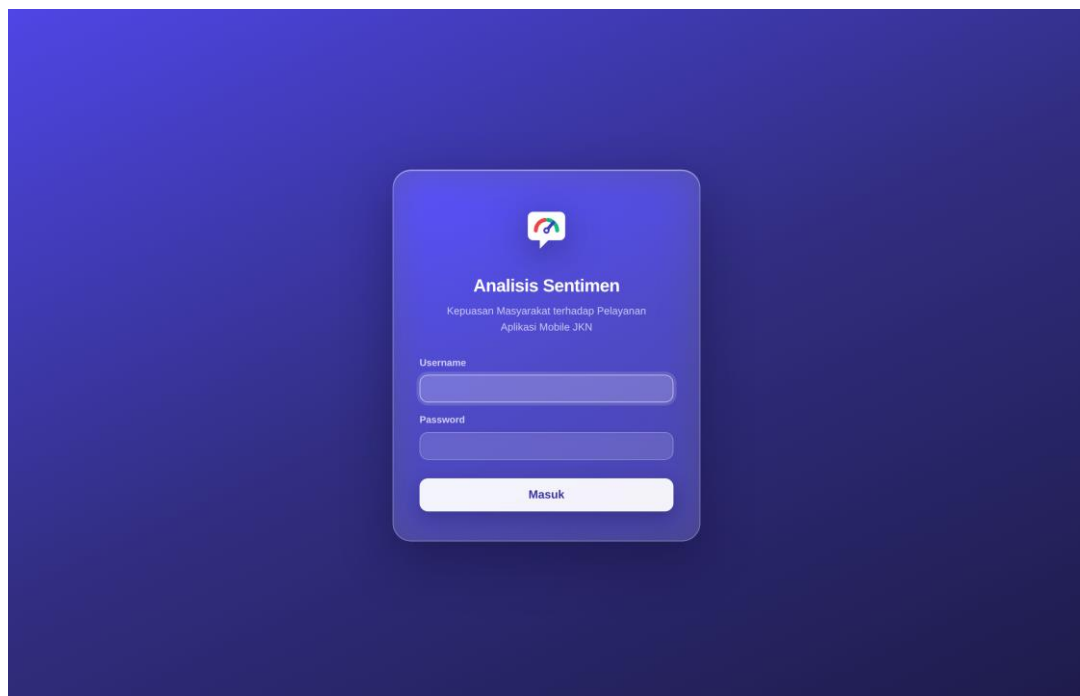
Berdasarkan Gambar 4.2, sebanyak 403 ulasan atau 53,7 persen tergolong sentimen negatif dan 348 ulasan atau 46,3 persen tergolong sentimen positif. Hasil ini menunjukkan bahwa mayoritas opini pengguna terhadap pelayanan aplikasi Mobile JKN pada periode penelitian bersentimen negatif, meskipun selisihnya dengan sentimen positif tidak terlalu besar.

4.9. Implementasi Sistem

Hasil analisis sentimen diimplementasikan ke dalam sistem informasi berbasis web sesuai rancangan pada BAB III. Sistem dibangun menggunakan bahasa pemrograman Python dengan kerangka kerja Flask dan basis data SQLite. Sistem memiliki satu aktor, yaitu Admin, yang dapat mengelola *dataset*, menjalankan proses analisis sentimen, dan melihat hasil analisis.

4.9.1. Halaman Login

Halaman *login* merupakan halaman awal yang digunakan sebagai proses autentikasi pengguna sebelum mengakses sistem. Kata sandi disimpan dalam bentuk hash, bukan teks biasa, sehingga tidak dapat dibaca langsung dari basis data. Tampilan halaman *login* ditunjukkan pada Gambar 4.3.

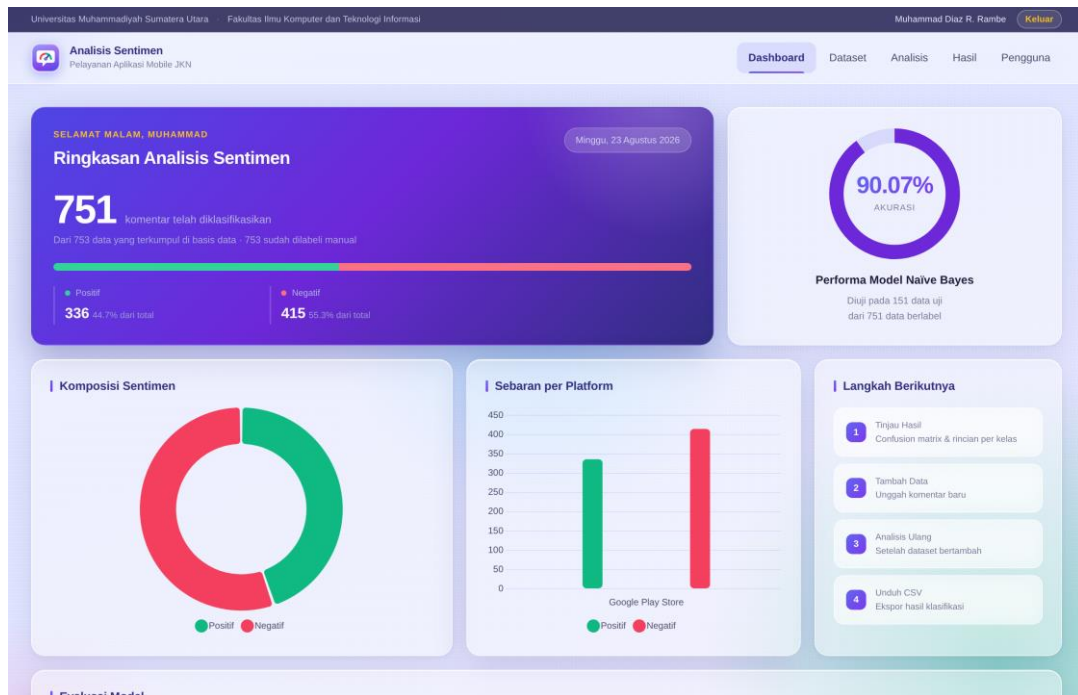


Gambar 4.3 Tampilan Halaman Login

4.9.2. Halaman Dashboard

Halaman *dashboard* merupakan halaman utama setelah pengguna berhasil melakukan *login*. Halaman ini menampilkan ringkasan jumlah *dataset*, distribusi

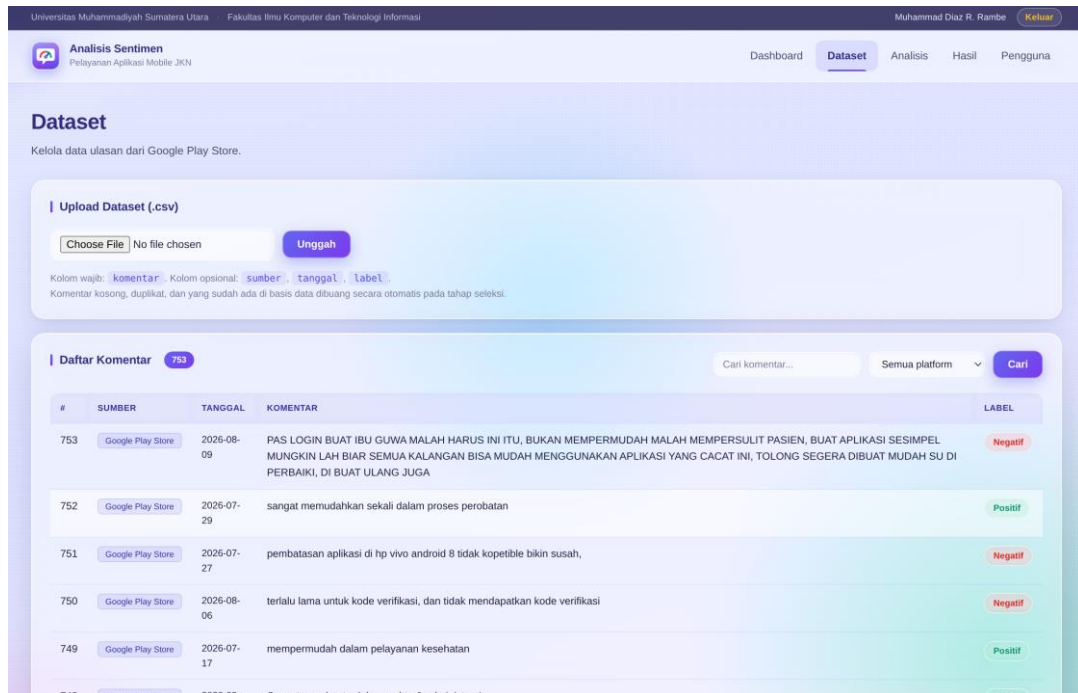
sentimen, dan nilai akurasi model. Tampilan halaman *dashboard* ditunjukkan pada Gambar 4.4.



Gambar 4.4 Tampilan Halaman Dashboard

4.9.3. Halaman Dataset

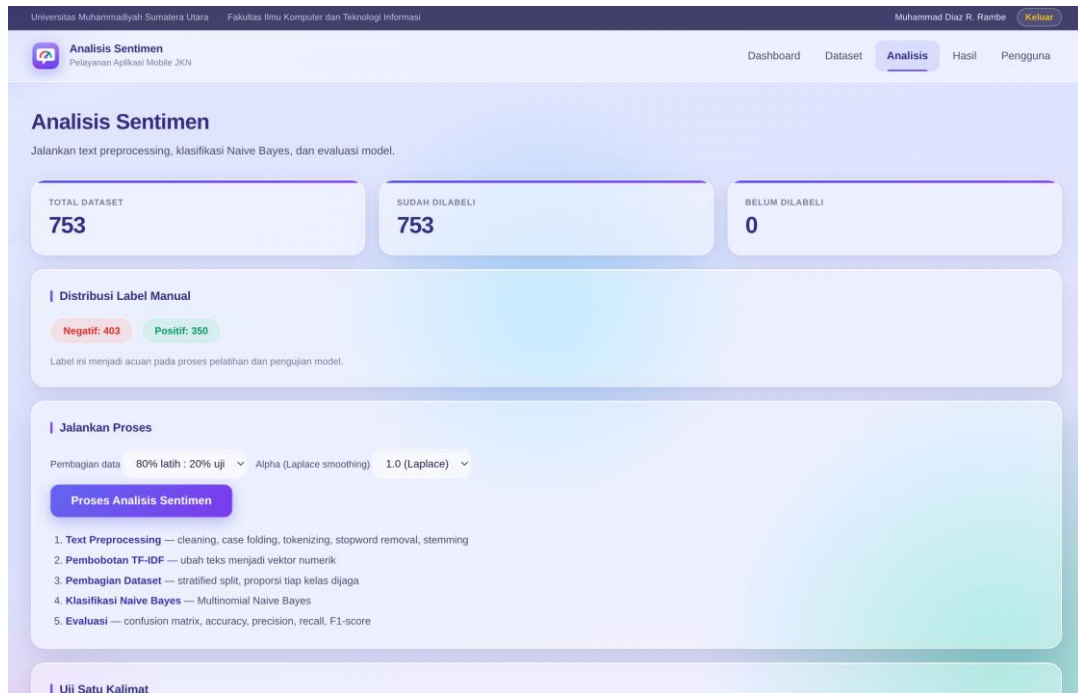
Halaman *dataset* digunakan untuk mengunggah dan menampilkan data ulasan yang akan dianalisis. Sistem menerima berkas berformat CSV dan secara otomatis membuang ulasan kosong maupun duplikat sesuai tahapan seleksi *dataset*. Tampilan halaman *dataset* ditunjukkan pada Gambar 4.5.



Gambar 4.5 Tampilan Halaman Dataset

4.9.4. Halaman Analisis Sentimen

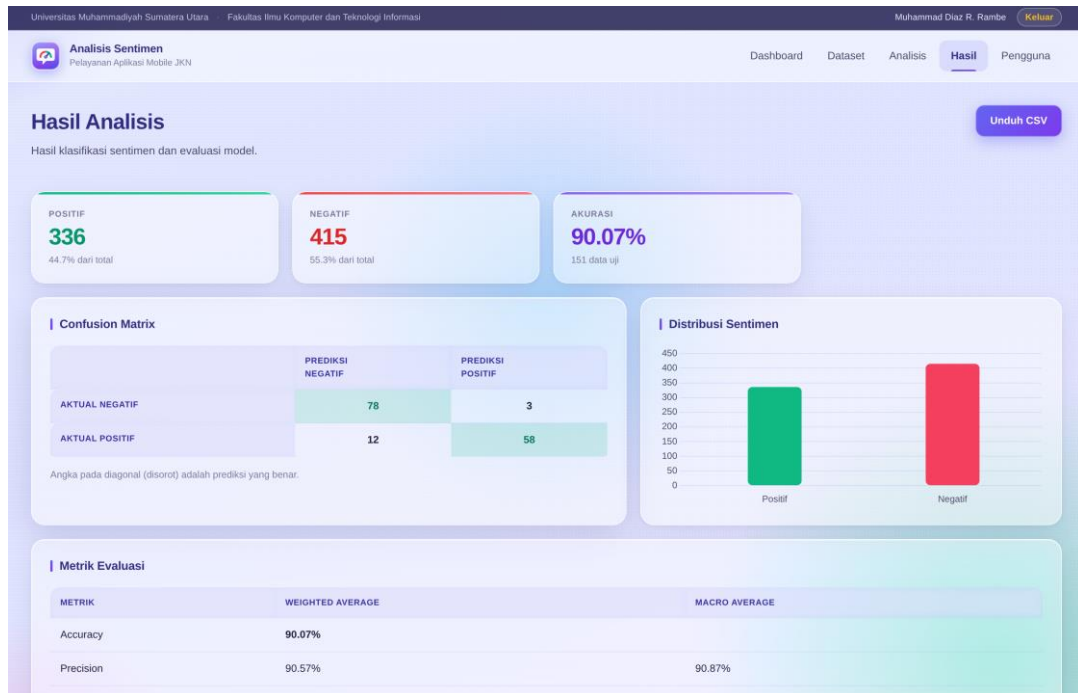
Halaman analisis sentimen digunakan untuk menjalankan proses pengolahan data. Ketika tombol proses ditekan, sistem menjalankan *text preprocessing*, pembobotan TF-IDF, klasifikasi menggunakan algoritma Naïve Bayes, dan evaluasi model secara berurutan. Tampilan halaman analisis sentimen ditunjukkan pada Gambar 4.6.



Gambar 4.6 Tampilan Halaman Analisis Sentimen

4.9.5. Halaman Hasil Analisis

Halaman hasil analisis menampilkan keluaran klasifikasi sentimen beserta evaluasi model, meliputi jumlah ulasan pada setiap kategori, *Confusion Matrix*, tabel metrik evaluasi, perhitungan per kelas, kata paling berpengaruh, serta tabel hasil klasifikasi setiap ulasan. Tampilan halaman hasil analisis ditunjukkan pada Gambar 4.7.



Gambar 4.7 Tampilan Halaman Hasil Analisis

Bagian detail hasil klasifikasi pada halaman yang sama ditunjukkan pada Gambar 4.8. Tabel tersebut menampilkan komentar asli, hasil *preprocessing*, label hasil pelabelan, hasil prediksi model, serta penanda kesesuaian antara keduanya.

Detail Hasil Klasifikasi 751					
SUMBER	KOMENTAR	HASIL PREPROCESSING	LABEL MANUAL	PREDIKSI	SESUAI
Google Play Store	anehhh, tagihan udah di bayarkan/dialihkan oleh instansi, tapi di aplikasi masih ada tunggakan! alasannya sistem. anggarab banyak, tapi sistem BOBROK! ntah lari kemana tuh 🤔🤔🤔	anehh tagih bayar alih instansi aplikasi nggak alas sistem anggarab sistem bobrok ntah lari mana	Negatif	Negatif	✓
Google Play Store	aplikasi tidak bisa di akses di hp Iqoo	aplikasi tidak akses iqoo	Negatif	Negatif	✓
Google Play Store	sangat bagus dan baik	bagus	Positif	Positif	✓
Google Play Store	TOLONG BANGET, PERBAIKI APLIKASINYA!!! BERKALI KALI DAFTAR TAPI OTP GA SAMPE? D NOTIF, INI GIMANASIH? MAU BUAT KEPENTINGAN JADI REPOT BGT GBS BSA!!	tolong banget perbaiki aplikasi kali daftar otp tidak sampe notif gimanasih penting repot banget tidak bsa	Negatif	Negatif	✓
Google Play Store	sangat bagus dan mudah di gunakan SIIP sangat membantu	bagus mudah siip bantu	Positif	Positif	✓
Google Play Store	ini mau login GK bisa,,apk bangke	login tidak apk bangke	Negatif	Negatif	✓
Google Play Store	Sering tidak nyambung, terutama pergantian bulan	tidak nyambung ganti	Negatif	Negatif	✓
Google Play Store	sangat membantu dim proses berobat	bantu proses obat	Positif	Positif	✓
Google Play Store	kenapa hp saya tidak bisa lagi mendownload mobile jkn kesehatan versi lama dengan nik lama dan nik baru saya buat	tidak mendownload mobile jkn sehat versi nik nik	Negatif	Negatif	✓
Google Play Store	ini apk error mulu perasaan, saat mau di pake	apk error mulu asa pake	Negatif	Negatif	✓
Google Play Store	aplikasi ga jelas ngrim kode otp ga bisa bisa	aplikasi tidak ngrim kode otp tidak	Negatif	Negatif	✓
Google Play Store	sejauh ini sangat membantu	bantu	Positif	Positif	✓
Google Play Store	terimakasih banyak atas mobile jkn mempermudah berobat di mana?	terimakasih mobile jkn muat obat mana	Positif	Positif	✓

Gambar 4.8 Tampilan Detail Hasil Klasifikasi pada Sistem

Nilai akurasi yang ditampilkan sistem, yaitu 90,07 persen, konsisten dengan hasil perhitungan yang disajikan pada subbab 4.8. Hal ini menunjukkan bahwa implementasi sistem telah sesuai dengan proses analisis yang dirancang.

Jumlah sentimen yang ditampilkan sistem, yaitu 336 ulasan positif dan 415 ulasan negatif, merupakan hasil prediksi model terhadap seluruh 751 ulasan. Jumlah tersebut berbeda tipis dari sebaran hasil pelabelan pada Gambar 4.2, yaitu 348 ulasan positif dan 403 ulasan negatif, karena sebagian kecil ulasan diprediksi berbeda dari labelnya. Selisih tersebut sejalan dengan nilai *accuracy* model sebesar 90,07 persen.

4.9.6. Potongan Kode Program Utama

Bagian ini menyajikan potongan kode program yang menjalankan tahapan utama penelitian, mulai dari *text preprocessing*, pembobotan TF-IDF, pelatihan model Naïve Bayes, evaluasi, sampai pengujian *cross validation*. Kode ditulis menggunakan bahasa pemrograman Python dengan pustaka Sastrawi untuk *stemming* bahasa Indonesia dan pustaka scikit-learn untuk pembobotan, klasifikasi, serta evaluasi model.

Potongan kode tahapan *text preprocessing* ditunjukkan pada Gambar 4.9, diambil dari berkas *sentimen/preprocessing.py*. Seluruh tahapan dijalankan berurutan melalui fungsi *preprocess*, sesuai rancangan pada BAB III.

```

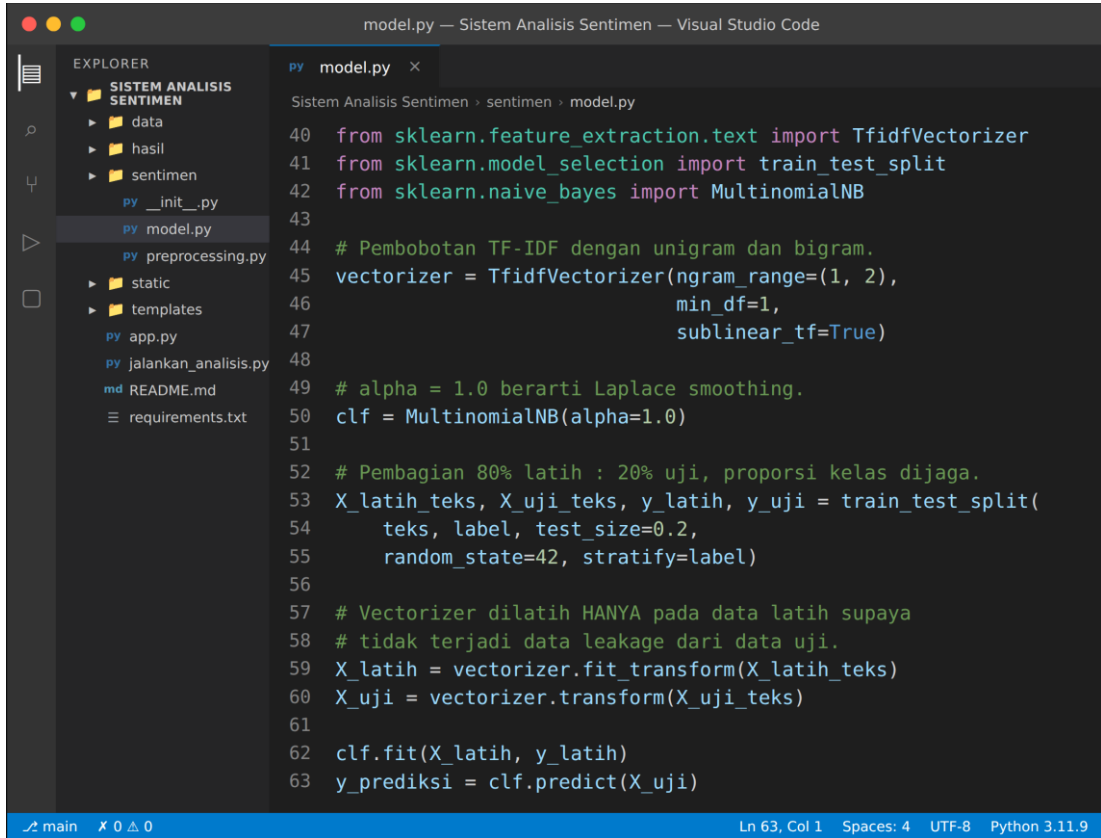
preprocessing.py — Sistem Analisis Sentimen — Visual Studio Code

62 def cleaning(teks):
63     """Hapus URL, mention, tagar, angka, dan tanda baca."""
64     teks = teks.lower() # case folding
65     teks = re.sub(r"http\S+|www.\S+", " ", teks)
66     teks = re.sub(r"@[\w._]+", " ", teks) # mention
67     teks = re.sub(r"#[\w]+", " ", teks) # tagar
68     teks = re.sub(r"\d+", " ", teks) # angka
69     teks = re.sub(r"[^\w\s]", " ", teks) # tanda baca
70     teks = re.sub(r"(\.|\!|\?){2,}", r"\1\1", teks) # huruf berulang
71     return re.sub(r"\s+", " ", teks).strip()
72
73
74 def tokenizing(teks):
75     """Pecah kalimat menjadi daftar kata."""
76     return teks.split()
77
78
79 def stopword_removal(tokens):
80     """Buang kata umum. Kata negasi tetap dipertahankan."""
81     return [t for t in tokens
82             if t not in STOPWORDS and len(t) > 2]
83
84
85 def stemming(tokens):
86     """Ubah kata berimbuhan menjadi kata dasar (Sastrawi)."""
87     return [stemmer.stem(t) for t in tokens]
88
89
90 def preprocess(teks):
91     """Jalankan seluruh tahapan pada satu ulasan."""
92     hasil = cleaning(teks)
93     hasil = tokenizing(hasil)
94     hasil = normalisasi(hasil)
95     hasil = stopword_removal(hasil)
96     hasil = stemming(hasil)
97     return " ".join(hasil)

```

Gambar 4.9 Potongan Kode Tahapan Text Preprocessing

Potongan kode pembobotan TF-IDF dan pelatihan model Naïve Bayes ditunjukkan pada Gambar 4.10. Parameter alpha bernilai 1,0 yang berarti model menggunakan *Laplace smoothing*, sedangkan ngram_range bernilai (1, 2) yang berarti model menggunakan *unigram* dan *bigram* sehingga frasa negasi seperti “tidak ramah” tetap terbaca sebagai satu fitur. Objek vectorizer sengaja hanya dilatih pada data latih agar informasi dari data uji tidak bocor ke proses pelatihan. Kode tersebut berada pada berkas sentimen/model.py.



```

model.py — Sistem Analisis Sentimen — Visual Studio Code

EXPLORER
  SISTEM ANALISIS SENTIMEN
    data
    hasil
    sentimen
    __init__.py
    model.py
    preprocessing.py
    static
    templates
    app.py
    jalankan_analisis.py
    README.md
    requirements.txt

40 from sklearn.feature_extraction.text import TfidfVectorizer
41 from sklearn.model_selection import train_test_split
42 from sklearn.naive_bayes import MultinomialNB
43
44 # Pembobotan TF-IDF dengan unigram dan bigram.
45 vectorizer = TfidfVectorizer(ngram_range=(1, 2),
46                             min_df=1,
47                             sublinear_tf=True)
48
49 # alpha = 1.0 berarti Laplace smoothing.
50 clf = MultinomialNB(alpha=1.0)
51
52 # Pembagian 80% latih : 20% uji, proporsi kelas dijaga.
53 X_latih_teks, X_uji_teks, y_latih, y_uji = train_test_split(
54     teks, label, test_size=0.2,
55     random_state=42, stratify=label)
56
57 # Vectorizer dilatih HANYA pada data latih supaya
58 # tidak terjadi data leakage dari data uji.
59 X_latih = vectorizer.fit_transform(X_latih_teks)
60 X_uji = vectorizer.transform(X_uji_teks)
61
62 clf.fit(X_latih, y_latih)
63 y_prediksi = clf.predict(X_uji)

```

Gambar 4.10 Potongan Kode Pembobotan TF-IDF dan Pelatihan Model Naïve Bayes

Potongan kode evaluasi model ditunjukkan pada Gambar 4.11. Nilai *accuracy*, *precision*, *recall*, dan *F1-score* dihitung menggunakan pustaka *scikit-learn* berdasarkan hasil prediksi pada data uji.

```

model.py — Sistem Analisis Sentimen — Visual Studio Code

EXPLORER
  SISTEM ANALISIS SENTIMEN
    data
    hasil
    sentimen
    py __init__.py
    py model.py
    py preprocessing.py
    static
    templates
    py app.py
    py jalankan_analisis.py
    md README.md
    requirements.txt

112 from sklearn.metrics import (accuracy_score,
113                               precision_score,
114                               recall_score, f1_score,
115                               confusion_matrix)
116
117 kelas = sorted(set(label))      # ["negatif", "positif"]
118 cm = confusion_matrix(y_uji, y_prediksi, labels=kelas)
119
120 accuracy = accuracy_score(y_uji, y_prediksi)
121 precision = precision_score(y_uji, y_prediksi,
122                             average="weighted")
123 recall = recall_score(y_uji, y_prediksi,
124                       average="weighted")
125 f1 = f1_score(y_uji, y_prediksi, average="weighted")
126
127 print(f"Accuracy : {accuracy * 100:.2f}%")
128 print(f"Precision : {precision * 100:.2f}%")
129 print(f"Recall : {recall * 100:.2f}%")
130 print(f"F1-Score : {f1 * 100:.2f}%")

Ln 130, Col 1 Spaces: 4 UTF-8 Python 3.11.9

```

Gambar 4.11 Potongan Kode Evaluasi Model

Potongan kode pengujian *10-fold cross validation* ditunjukkan pada Gambar 4.12. Pembagian lipatan dilakukan menggunakan `StratifiedKFold` sehingga proporsi kelas positif dan negatif tetap terjaga pada setiap lipatan.

```

model.py — Sistem Analisis Sentimen — Visual Studio Code

EXPLORER
  SISTEM ANALISIS SENTIMEN
    data
    hasil
    sentimen
    py __init__.py
    py model.py
    py preprocessing.py
    static
    templates
    py app.py
    py jalankan_analisis.py
    md README.md
    requirements.txt

151 from sklearn.model_selection import (cross_val_score,
152                                     StratifiedKFold)
153
154 # Cross validation memakai seluruh data berlabel.
155 X = vectorizer.fit_transform(teks)
156
157 # StratifiedKFold menjaga proporsi kelas tiap lipatan.
158 skf = StratifiedKFold(n_splits=10, shuffle=True,
159                       random_state=42)
160
161 skor = cross_val_score(MultinomialNB(alpha=1.0), X,
162                        label, cv=skf, scoring="accuracy")
163
164 print(f"Rata-rata akurasi : {skor.mean() * 100:.2f}%")
165 print(f"Simpangan baku : {skor.std() * 100:.2f}%")

Ln 165, Col 1 Spaces: 4 UTF-8 Python 3.11.9

```

Gambar 4.12 Potongan Kode Pengujian 10-Fold Cross Validation

4.10. Hasil Pengujian Sistem

Pengujian sistem dilakukan menggunakan metode Black Box Testing terhadap seluruh fungsi utama sistem. Hasil pengujian disajikan pada Tabel 4.12.

Tabel 4.12 Hasil Pengujian Black Box

No	Fungsi	Skenario	Hasil Diharapkan	Hasil
1	Login	Memasukkan username dan password benar	Sistem menampilkan halaman dashboard	Sesuai
2	Login	Memasukkan password salah	Sistem menampilkan pesan kesalahan	Sesuai
3	Login	Mengakses halaman tanpa login	Sistem mengalihkan ke halaman login	Sesuai
4	Dataset	Mengunggah berkas CSV sesuai format	Data tersimpan ke basis data	Sesuai
5	Dataset	Mengunggah berkas selain CSV	Sistem menolak berkas	Sesuai
6	Dataset	Menampilkan daftar ulasan	Sistem menampilkan tabel data	Sesuai
7	Analisis	Menjalankan proses analisis sentimen	Sistem menampilkan hasil analisis	Sesuai
8	Hasil	Menampilkan nilai evaluasi model	Sistem menampilkan metrik evaluasi	Sesuai
9	Hasil	Menampilkan Confusion Matrix	Sistem menampilkan tabel matriks	Sesuai
10	Hasil	Mengunduh hasil klasifikasi	Sistem mengunduh berkas CSV	Sesuai
11	Pengguna	Mengubah kata sandi	Kata sandi berhasil diperbarui	Sesuai
12	Logout	Menekan tombol keluar	Sesi berakhir, kembali ke login	Sesuai

Berdasarkan Tabel 4.12, seluruh dari 12 skenario pengujian memberikan hasil yang sesuai dengan yang diharapkan. Dengan demikian dapat disimpulkan bahwa seluruh fungsi sistem berjalan sebagaimana mestinya.

4.11. Pembahasan

4.11.1. 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 4.13.

Tabel 4.13 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

Tabel 4.13 menunjukkan bahwa kata yang dominan pada kelas negatif berkaitan dengan kendala teknis proses masuk dan verifikasi, yaitu kata 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 memberikan informasi yang bermanfaat bagi penyelenggara layanan. Keluhan pengguna tidak tersebar merata pada seluruh fitur, melainkan terpusat pada proses autentikasi, khususnya pengiriman kode OTP dan proses verifikasi akun. Perbaikan pada bagian tersebut berpotensi menurunkan proporsi sentimen negatif secara signifikan.

4.11.2. 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 keseluruhan kalimatnya menyampaikan hal positif atau sekadar pertanyaan, misalnya ulasan yang menanyakan mengapa aplikasi tidak dapat

dibuka. Model mengenali kata tidak dan dibuka sebagai penanda negatif sehingga mengabaikan konteks kalimat secara keseluruhan.

Pola kedua adalah keterbatasan algoritma Naïve Bayes yang mengasumsikan setiap kata bersifat independen. Asumsi tersebut menyebabkan model tidak mempertimbangkan hubungan antarkata dalam satu kalimat, sehingga kalimat yang mengandung campuran pujian dan keluhan cenderung diklasifikasikan berdasarkan kata yang paling dominan bobotnya.

4.11.3. Perbandingan dengan Penelitian Terdahulu

Nilai akurasi yang diperoleh pada penelitian ini, yaitu 90,07 persen, berada pada kisaran yang sebanding dengan penelitian terdahulu. Penelitian Rahmawati et al. (2024) pada aplikasi m-BCA memperoleh akurasi 82 persen, penelitian Hidayat et al. (2024) pada layanan Sistem Informasi Akademik memperoleh 80,06 persen, sedangkan penelitian Dzulkarnain et al. (2024) pada pelayanan rumah sakit memperoleh 90 persen. Pada objek yang sama, Putra et al. (2025) memperoleh akurasi 85,5 persen menggunakan Multinomial Naïve Bayes, sedangkan Paliling & Sutoyo (2026) memperoleh akurasi 92,15 persen menggunakan Logistic Regression.

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, sehingga belum menggambarkan persepsi pengguna pada masa transisi kebijakan. Selain itu, sebagian besar penelitian terdahulu berhenti pada penyajian nilai akurasi dan belum mengimplementasikan hasil klasifikasi ke dalam sistem informasi yang dapat digunakan secara berulang. Meskipun demikian, temuan

bahwa keluhan pengguna terpusat pada proses login dan pengiriman kode OTP sejalan dengan penelitian Paliling & Sutoyo (2026) yang juga menemukan sentimen negatif paling dominan pada kategori login dan autentikasi.

4.11.4. Jawaban atas Rumusan Masalah

Rumusan masalah pertama mengenai penerapan algoritma Naïve Bayes telah dijawab pada subbab 4.4 sampai 4.7. 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.

Rumusan masalah kedua mengenai tingkat performa algoritma telah dijawab pada subbab 4.8. Model memperoleh *accuracy* 90,07 persen, *precision* 90,57 persen, *recall* 90,07 persen, dan *F1-score* 89,99 persen, dengan rata-rata *cross validation* 90,68 persen dan simpangan baku 2,60 persen.

Rumusan masalah ketiga mengenai implementasi ke dalam sistem informasi telah dijawab pada subbab 4.9 dan 4.10. Sistem informasi berbasis web berhasil dibangun sesuai rancangan, seluruh fungsinya berjalan sesuai pengujian Black Box, dan nilai akurasi yang ditampilkan sistem konsisten dengan hasil perhitungan.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan, dapat disimpulkan beberapa hal sebagai berikut:

1. Algoritma Naïve Bayes berhasil diterapkan untuk mengklasifikasikan sentimen ulasan pengguna aplikasi Mobile JKN pada Google Play Store. Dari 4.000 ulasan yang dikumpulkan, diperoleh 2.675 ulasan hasil seleksi, dan 751 ulasan digunakan pada tahap analisis setelah melalui proses pelabelan. Hasil klasifikasi menunjukkan 403 ulasan atau 53,7 persen bersentimen negatif dan 348 ulasan atau 46,3 persen bersentimen positif.
2. Model yang dibangun memperoleh nilai *accuracy* sebesar 90,07 persen, *precision* 90,57 persen, *recall* 90,07 persen, dan *F1-score* 89,99 persen. Pengujian *10-fold cross validation* memperoleh rata-rata akurasi 90,68 persen dengan simpangan baku 2,60 persen, yang menunjukkan bahwa performa model stabil dan tidak bergantung pada pembagian data tertentu.
3. Hasil klasifikasi sentimen berhasil diimplementasikan ke dalam sistem informasi berbasis web. Seluruh dari 12 skenario pengujian Black Box memberikan hasil yang sesuai, dan nilai akurasi yang ditampilkan sistem konsisten dengan hasil perhitungan. Sistem tersebut dapat digunakan secara berulang untuk menganalisis ulasan pengguna tanpa harus membaca seluruh ulasan secara manual.

Berdasarkan analisis kata paling berpengaruh, keluhan pengguna terpusat pada proses autentikasi, khususnya pengiriman kode OTP dan proses verifikasi

akun, sedangkan apresiasi pengguna berkaitan dengan kemudahan dan kecepatan layanan.

5.2. Saran

Berdasarkan hasil penelitian, terdapat beberapa saran yang dapat disampaikan:

1. Bagi BPJS Kesehatan, hasil penelitian menunjukkan bahwa keluhan pengguna paling banyak berkaitan dengan proses masuk, pengiriman kode OTP, dan verifikasi akun. Perbaikan pada bagian tersebut disarankan menjadi prioritas karena berpotensi menurunkan proporsi sentimen negatif secara signifikan.
2. 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.
3. Penelitian selanjutnya dapat mencoba algoritma klasifikasi lain seperti Support Vector Machine atau pendekatan berbasis *deep learning*, serta membandingkan hasilnya dengan algoritma Naïve Bayes untuk memperoleh gambaran performa yang lebih lengkap.
4. Klasifikasi sentimen pada penelitian ini terbatas pada dua kategori. Penelitian selanjutnya dapat mengembangkan analisis berbasis aspek agar dapat mengetahui aspek layanan tertentu yang paling banyak dikeluhkan, seperti aspek autentikasi, antrean, atau pembayaran iuran.

DAFTAR PUSTAKA

- Amaa, R. P., Ali, I., Rahaningsih, N., & Prihartono, W. (2025). Kinerja Multinomial Naïve Bayes pada analisis sentimen ulasan pengguna aplikasi Mobile JKN. *J-ENSITEC (Journal of Engineering and Sustainable Technology)*, 12(1), 10338–10345. <https://doi.org/10.31949/j-ensitec.v12i01.16656>
- Berliansyah, D., Anisatur, U., & Alfaruq, H. A. (2024). Analisis sentimen pada media sosial Instagram terhadap akun Presiden Joko Widodo menggunakan metode Naïve Bayes Classifier. *Journal of Digital Innovation & Information Technology*, 1(1), 1–11.
- BPJS Kesehatan. (2026). *Capaian Program Jaminan Kesehatan Nasional tahun 2025: Peserta JKN mencapai 282,7 juta jiwa*. Jakarta: Badan Penyelenggara Jaminan Sosial Kesehatan.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- 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>
- Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi. (2017). *Peraturan Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Nomor 14 Tahun 2017 tentang Pedoman Penyusunan Survei Kepuasan Masyarakat Unit Penyelenggara Pelayanan Publik*. Jakarta: Kementerian PANRB.
- 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>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge: Cambridge University Press.

- Paliling, A., & Sutoyo, M. N. (2026). Analisis sentimen ulasan Mobile JKN menggunakan TF-IDF dan logistic regression. *Buletin Sistem Informasi dan Teknologi Islam (BUSITI)*, 7(1), 47–55. <https://doi.org/10.33096/busiti.v7i1.3319>
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12–40.
- 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.
- Putra, I. E., Asriyanik, A., & Azzahra, F. F. (2025). Pemanfaatan Multinomial Naive Bayes untuk analisis sentimen ulasan aplikasi Mobile JKN. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(2), 3283–3291. <https://doi.org/10.36040/jati.v9i2.13311>
- 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. (2009). *Undang-Undang Republik Indonesia Nomor 25 Tahun 2009 tentang Pelayanan Publik*. Lembaran Negara Republik Indonesia Tahun 2009 Nomor 112. Jakarta: Sekretariat Negara.
- 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.

LAMPIRAN

Lampiran 1. SK Dosen Pembimbing



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi Unggul Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 174/SK/BAK-PT/IAK-P/PT/18/2024
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631093
 Email: info@umsu.ac.id | website@umsu.ac.id | fms@umsu.ac.id | umsumedan | umsumedan | umsumedan | umsumedan

KEPUTUSAN DEKAN
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
 Nomor : 466/KEP/II.3.AU/UMSU-09/F/2026

Tentang :
PENGANGKATAN DOSEN PEMBIMBING SKRIPSI
PRODI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA



Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara, setelah:

Menimbang : 1. Bahwa sehubungan dengan pelaksanaan Tugas Akhir Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara;

2. Bahwa untuk memenuhi maksud tersebut perlu diterbitkan surat keputusan Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.

Mengingat : 1. Undang-undang Republik Indonesia No. 12 Tahun 2012 tentang Pendidikan Tinggi;

2. Peraturan Pemerintah Nomor 4 Tahun 2014 tentang Penyelenggaraan Pendidikan Tinggi dan Pengelolaan Perguruan Tinggi;

3. Pedoman Perguruan Tinggi Muhammadiyah;

4. Statuta Universitas Muhammadiyah Sumatera Utara

5. Keputusan Rektor No. 1001/KEP/II.3-AU/UMSU/D/2025 tanggal 25 Oktober 2025 tentang Dekan Fakultas Ilmu Komputer dan Teknologi Informasi UMSU Masa Jabatan 2025-2029

6. Keputusan Rektor No. 624/KEP/II.3.AU/UMSU/D/2025 tentang Prodi Fakultas Ilmu Komputer dan Teknologi Informasi UMSU Masa Jabatan 2025-2029

MEMUTUSKAN

Menetapkan :
 KESATU : Menetapkan nama-nama dalam lampiran surat keputusan ini sebagai Dosen Pembimbing Prodi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi UMSU.

KEDUA : Keputusan ini berlaku sejak tanggal ditetapkan, dan akan ditinjau serta diperbaiki kembali jika terdapat kekeliruan di kemudian hari.

Ditetapkan di : Medan
 Pada tanggal : 09 Ramadhan 1447 H
 26 Februari 2026 M





Dekan
Dr. M. Khowarizmi, S.Kom., M.Kom
 STN 0127099201

Tembusan:

1. Yth. Wakil Dekan I & III UMSU di Medan,-
2. Ka. Prodi SI dan Sek. Prodi SI,-
3. Pertinggal.







MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi Unggul Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 174/SK/BAAN-PT/AK-Psy/PT/BB/2024
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6623474 - 6631903
<http://www.umsu.ac.id> Email: info@umsu.ac.id [umsu.medan](https://www.facebook.com/umsu.medan) [umsu.medan](https://www.instagram.com/umsu.medan) [umsu.medan](https://www.youtube.com/channel/UC...) [umsu.medan](https://www.tiktok.com/@umsu.medan)

Lampiran Dosen Pembimbing Prodi Sistem Informasi
 Nomor : 466/KEP/II.3.AU/UMSU-09/F/2026
 Tanggal : 09 Ramadhan 1447 H / 26 Februari 2026 M

PENGANGKATAN DOSEN PEMBIMBING SKRIPSI
PRODI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

NO	NAMA	NPM	JUDUL	DOSEN
1	Muhammad Diaz R Rambe	2209010031	Analisis Sentimen Kepuasan Masyarakat terhadap Pelayanan Kepolisian Sektor (Polsek) Menggunakan Natural Language Processing	Asrar Aspia Manurung, S.Pd., M.Pd
2	Alfin Juanda	2109010104	Pengembangan Software Analisis Keranjang Belanja (Market Basket) untuk Rekomendasi Bundling Menu pada PT Modular Kuliner Indonesia menggunakan Apriori	Asrar Aspia Manurung, S.Pd., M.Pd
3	Sarah nur fadhillah	2209010190	Analisis Sentimen dan Pemodelan Topik Keluhan Warga Terkait Banjir Rob Berbasis Data Media Sosial dan Berita Lokal Menggunakan Fine-Tuning IndoBERT dan BERTopic untuk Pemetaan Isu Dominan Periode Waktu (Studi Kasus: Kota Tanjungbalai)	Dr. Irvan, M.Si
4	Siti Aspurina	2209010015	Pembangunan aplikasi deteksi plating pada emas palsu melalui analisis cacat permukaan menggunakan CNN - Based Anomaly detection dengan backbone ResNet	Indah Purnama Sari, S.T., M.Kom
5	Irvan	2409010342P	Penerapan Metode Classical Test Theory untuk Analisis Tingkat Kesukaran dan Daya Beda Soal pada Sistem Bank Soal Berbasis Komputer	Dr. Zainal Azis, M.Si
6	Farhan Mubin	2209010032	Pembangunan Sistem Absensi Cerdas Berbasis QR Code dan SPK untuk Penentuan Tingkat Disiplin Kehadiran Siswa (Studi Kasus: SMAN 1 Hamparan Perak)	Mulkan Azhari, S.Kom., M.Kom.



Dekan

Dr. M. Bawarizmi, S.Kom., M.Kom
 STDN 0127099201



Lampiran 2. Berita Acara Bimbingan



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 89/SK/BAN-PT/Akred/PT/2019
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003
<https://id.umsu.ac.id> info@umsu.ac.id [umsu.ac.id](#) [umsu.ac.id](#) [umsu.ac.id](#) [umsu.ac.id](#)

Berita Acara Pembimbingan Proposal

Nama Mahasiswa : Muhammad Diaz R. Rambe
 NPM : 2209010031
 Program Studi : Sistem Informasi
 Nama Dosen Pembimbing : Martiano, S.Pd., S.Kom., M.Kom.
 Judul Penelitian : Analisis Sentimen Kepuasan Masyarakat terhadap Pelayanan Aplikasi Mobile JKN Menggunakan Algoritma Naïve Bayes

Tanggal Bimbingan	Hasil Evaluasi	Paraf Dosen
03/07	Revisi BAB 1.	
14/07	Revisi BAB 2 dan 3.	
22/07	ACC Sempro	
17/08	Revisi Sempro	
18/08	Revisi BAB 4 dan 5.	
24/8	Exhibition.	
26/8	ACC Sidang	

Diketahui oleh :
Ketua Program Studi



(Mahardika Prawira Tanjung, M.Kom.)

Medan, 26 Januari 2026
Disetujui oleh :
Dosen Pembimbing



(Martiano, S.Pd., S.Kom., M.Kom.)



Lampiran 3. Form Revisi Sidang

MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN FIMPINAN PURAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UMSU Terakreditasi Unggul Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 174/SKOTIAN/PTIAK.Ppg/PT/2024
Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax (061) 6625474 - 6631003
Website: www.umsu.ac.id Email: info@umsu.ac.id
Instagram: @umsu_medan Facebook: UMSU Medan Twitter: @umsu_medan

FORMULIR PERBAIKAN UJIAN SKRIPSI

Pada hari ini, Selasa 1 September 2026 telah dilaksanakan Ujian Skripsi bagi mahasiswa Fakultas Ilmu Komputer Dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.

Sbb:

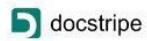
Nama Mahasiswa : Muhammad Diaz R Rambe
 NPM : 2209010031
 Program Studi : Sistem Informasi
 Judul Proposal : Analisis Sentimen Kepuasan Masyarakat Terhadap Pelayanan Aplikasi Mobile Menggunakan Algoritma Naïve Bayes

Materi/Point yang Diperbaiki :

		Paraf
Martiano S.Kom, M.Kom		
Dr.Firahmi Rizky.S.Kom.,M.Kom		
Indah Purnama Sari, S.T., M.Kom	- Perbaiki Numbering pada Kata Pengantar - Perbaiki Penulisan rumusan masalah (kasih numbering) - Batasan masalah sesuai dengan rumusan masalah	

Berita acara ini ditandatangani setelah skripsi diperbaiki sesuai petunjuk/arahan dari Pembimbing dan Penguji/Pembahas.

Lampiran 4. Hasil Turnitin



Plagiarism Report

Submission ID 6acdcc56d09507177640f0fc

Plagiarism Report

Skripsi_Diaz_-_Wisuda_Revisi.pdf

Document Details

Submission ID	File Size
6acdcc56d09507177640f0fc	3.3 MB
File Name	Total Pages
Skripsi_Diaz_-_Wisuda_Revisi.pdf	99 Page
Account Name	Total Words
ITentix Plagiarism	14,665 Words
Submission Date	Total Characters
13 Oct 2026, 13:14 GMT+7	133,205 Characters
Release Date	
13 Oct 2026, 13:20 GMT+7	

Filter Options

Filter settings used in this submission report.

- Exclude Bibliography



Plagiarism Report

Submission ID 6acdcc56d09507177640f0fc

17%

Similarity

215

Matched Sources

Match Group

0%	Internet sources
17%	Publications
0%	Submitted works (Student papers)

Top Sources

The sources with the highest number of matches within the submission.

1	Publication	<1%
repository.uin-suska.ac.id		
2	Publication	<1%
jurnal.polibatam.ac.id		
3	Publication	<1%
jurnal.polibatam.ac.id		
4	Publication	<1%
ejournal.upbatam.ac.id		
5	Publication	<1%
jurnal.polibatam.ac.id		
6	Publication	<1%
jurnal.poltekkespim.ac.id		
7	Publication	<1%
ojs.stmik-banjarbaru.ac.id		
8	Publication	<1%
sistemasi.ftik.unisi.ac.id		
9	Publication	<1%
eprints.ums.ac.id		
10	Publication	<1%
jurnal-perspektif.org		
11	Publication	<1%
publikasi.dinus.ac.id		

Lampiran 5. LOA



Indonesian Journal of Multidisciplinary Scientific Studies

E-ISSN: 2988-7046

Sekretariat: STAI Raudhatul Akmal, Jln. Nusa Indah Gg. Melati No. 66, Batang Kuis, 20372

<https://ojs.staira.ac.id/index.php/IJOMSS/index>

Deli Serdang, 13 Oktober 2026

No : 039/IJOMSS/LoA/X/2026
Lamp : -
Hal : Surat Penerimaan Naskah Jurnal

Kepada Yth.
Para Penulis
di Tempat

Letter of Accepted (LOA)

Terima kasih telah mengirimkan naskah anda ke **Indonesian Journal of Multidisciplinary Scientific Studies (IJOMSS)** dengan data sebagai berikut:

Penulis	Judul
MUHAMMAD DIAZ R. RAMBE	ANALISIS SENTIMEN KEPUASAN MASYARAKAT TERHADAP PELAYANAN APLIKASI MOBILE JKN
MARTIANO	MENGGUNAKAN ALGORITMA NAÏVE BAYES

Berdasarkan hasil review akhir yang telah dilakukan, dengan ini kami menyatakan bahwa naskah tersebut **DITERIMA** untuk dipublikasikan pada *Indonesian Journal of Multidisciplinary Scientific Studies (IJOMSS)*, edisi **Vol. 4 No. 6 (Oktober - November 2026)**.

Demikian surat penerimaan ini kami sampaikan. Atas perhatian dan kerja sama yang baik, kami ucapkan terima kasih.

Hormat Kami,



Dr. Muhammad Azhari, M.Pd
Editor in Chief

INDEX COPERNICUS
INTERNATIONAL

GARUDA
KARYA BUDAYA DIGITAL

Google
Scholar

BASE