

Penerapan Algoritma K-Nearest Neighbors (KNN) Dalam Analisis Sentimen Ulasan Pengguna Aplikasi JMO (Jamsostek Mobile)

Ghifari Fatihah Rabbani¹, Kresna Ramanda², Sulaeman Hadi Sukmana³,
Qudsiah Nur Azizah⁴, Erma Delima Sikumbang⁵

Universitas Bina Sarana Informatika¹²³⁴⁵

ghifarifr037@gmail.com¹, kresna.kra@bsi.ac.id², sulaeman.sdu@bsi.ac.id³,

qudsiah.qna@bsi.ac.id⁴, erma@bsi.ac.id⁵

Abstrak - JMO (Jamsostek Mobile) merupakan aplikasi resmi yang diluncurkan oleh BPJS Ketenagakerjaan, dirancang untuk mendukung para tenaga kerja dalam memudahkan akses perlindungan sosial. Namun, terdapat beberapa kendala yang dialami oleh pengguna aplikasi, seperti kesalahan yang terjadi dalam aplikasi, kesulitan saat melakukan login, dan saldo JHT yang tidak tampil. Untuk mendapatkan gambaran umum mengenai perasaan pengguna aplikasi JMO, diperlukan suatu evaluasi untuk mengamati tanggapan para pengguna. Oleh karena itu, studi ini akan mengeksplorasi tingkat ketepatan algoritma KNN dalam menganalisis perasaan dari ulasan pengguna aplikasi JMO. Penelitian ini bertujuan untuk melihat dan menganalisis tanggapan pengguna, mengklasifikasi sentimen umum, menerapkan analisis sentimen, dan menguji akurasi algoritma KNN. Penelitian ini bersifat kuantitatif, menggunakan metode analisis data *numerik* dengan pendekatan *text mining*. Hasil dari penelitian dengan mengambil 10,000 ulasan oleh proses *scrapping*, dan melalui *Preprocessing* data mendapatkan 7,185 data yang menunjukkan 51.38% sentimen positif dari keseluruhan data, algoritma KNN mendapatkan hasil evaluasi keakuratan tertinggi 76.2%, *precision* 76.2%, *recall* 78.0%, dan *F1-score* 77.1% dari nilai $K = 21$ pada distribusi data 90%-10%, kinerja model KNN mendapatkan nilai *AUC* 0.7617 menunjukkan hasil yang *Fair* atau cukup baik. Temuan ini mengartikan model klasifikasi algoritma KNN mampu memberikan hasil klasifikasi yang seimbang pada setiap kelas, sehingga evaluasi lebih adil dan bebas dari bias.

Kata Kunci : Analisis Sentimen, JMO, K-Nearest Neighbors, *Text Mining*

Abstract - JMO (Jamsostek Mobile) is an official application launched by BPJS Ketenagakerjaan, designed to support workers in facilitating access to social protection. However, several obstacles are experienced by application users, such as errors in the application, difficulties logging in, and JHT balances not being displayed. To obtain a general overview of JMO application user sentiment, an evaluation is needed to capture user concerns. Therefore, this study will explore the precision of the KNN algorithm in analyzing sentiment from JMO application user reviews. The objectives of this research are to identify and analyze user feedback, classify overall sentiments, implement sentiment analysis, and test the accuracy of the KNN algorithm. This study adopts a quantitative approach, applying numerical data analysis through text mining techniques. From a dataset of 10,000 reviews collected via web scraping and refined Preprocessing, 7,185 reviews were obtained, revealing that 51.38% expressed positive sentiment. The KNN algorithm achieved its highest accuracy 76.2%, precision 76.2%, recall 78.0%, and F1-score 77.1% at $K = 21$ under 90%-10% data split. Furthermore, the model's AUC score of 0.7617 indicated fair and reasonably good performance. These findings suggest that the KNN classification model is capable of providing balanced classification results across classes, leading to a fairer and less biased evaluation.

Keywords: Sentiment Analysis, JMO, K-Nearest Neighbors, *Text Mining*

I. PENDAHULUAN

JMO (Jamsostek Mobile) adalah aplikasi resmi dari instansi Badan Penyelenggara Jaminan Sosial (BPJS) Ketenagakerjaan, aplikasi ini bertujuan untuk membantu para pekerja dengan mempermudah pemberian perlindungan sosial bagi para tenaga kerja di Indonesia. BPJS Ketenagakerjaan berperan dalam memberikan manfaat bagi para pekerja, yaitu dengan menyelenggarakan program-program seperti, yang meliputi program jaminan kecelakaan kerja, jaminan kematian, jaminan hari tua, serta jaminan pensiun untuk pekerja beserta keluarga mereka (Sutrisno, 2020).

BPJS Ketenagakerjaan menghadirkan aplikasi JMO sebagai inovasi digital terkini yang

merupakan kelanjutan dari aplikasi *BPJSTKU* yang telah diperkenalkan sejak tahun 2016. Aplikasi ini menyediakan layanan seperti pembaruan data, pengajuan serta pemantauan klaim JHT, simulasi saldo JHT dan JP, pengaduan, hingga pelaporan (Samudera & Pertiwi, 2022). Berfungsi sebagai aplikasi digital bagi peserta dengan tujuan untuk mempermudah akses informasi, klaim, dan pelayanan jaminan sosial tanpa harus mendatangi kantor cabang.

Tetapi sebagaimana bagusnya aplikasi JMO ini, masih ada beberapa masalah yang dihadapi oleh pengguna aplikasi seperti *error* nya aplikasi, gagal untuk login, dan tidak munculnya saldo JHT. Masalah-masalah ini tentu saja dapat

membuat pengguna frustrasi sehingga mereka menyampaikan keluhannya, salah satu cara para pengguna melakukannya adalah dengan meninggalkan sebuah ulasan pada halaman *Google Play* aplikasi JMO.

Dengan hanya melihat secara sekilas *review* aplikasi JMO pada halaman *Google Play* masih belum tentu semua ulasan yang ada di sana adalah sebuah keluhan pada aplikasi, pastinya banyak juga ulasan yang memuji aplikasi JMO ini. Ulasan yang tercantum pada halaman *Google Play Store* ini masih harus diurutkan secara konsisten untuk mendapatkan hasil yang akurat pada sentimen pengguna terhadap kinerja aplikasi JMO.

Untuk memperoleh hasil sentimen umum pengguna aplikasi JMO, diperlukan sebuah analisis terhadap sentimen suara pengguna pada halaman *Google Play* untuk melihat reaksi para pengguna aplikasi bagaimana pengalaman dan opini mereka terhadap aplikasi, maupun itu positif atau negatif. Teknik ini banyak dimanfaatkan dalam pemantauan media sosial guna memperoleh gambaran umum mengenai kecenderungan opini, apakah bernuansa positif atau negatif (Dewi & Sulastri, 2022). Data mengenai sentimen pengguna dapat dimanfaatkan untuk merancang aplikasi yang lebih optimal, mengatur *alokasi* sumber daya secara efisien, serta menyusun strategi pengembangan aplikasi sesuai dengan kebutuhan dan harapan pengguna (Kevin et al., 2024).

Dalam melakukan analisis sentimen pengguna, dibutuhkan sebuah metode klasifikasi untuk melihat bagaimana akurat data ulasan yang didapatkan, salah satu algoritma yang biasa dipakai untuk proses klasifikasi data ulasan adalah Algoritma K-Nearest Neighbors (KNN), yaitu merupakan sebuah metode yang dapat digunakan pada prediksi maupun pengklasifikasi data, dengan penerapannya disesuaikan pada karakteristik serta jenis data pada himpunan data yang dianalisis (Khairi et al., 2021). Metode K-Nearest Neighbor (KNN) bekerja berdasarkan prinsip pengukuran jarak, yaitu dengan menentukan kedekatan suatu data uji terhadap sejumlah K tetangga terdekat yang terdapat dalam himpunan data latih. Dalam tahap klasifikasi, nilai K ditentukan sebagai sebuah konstanta oleh pengguna. Klasifikasi kemudian ditentukan berdasarkan mayoritas kelas dari tetangga terdekat tersebut (Akbar, 2024).

Penelitian yang dilakukan oleh (Astuti & Nuris, 2022) pada analisis sentimen *review* aplikasi Peduli Lindungi, berkesimpulan bahwa metode algoritma KNN dengan nilai $K=20$ bisa mendapatkan nilai akurasi 81,7% dengan nilai *Area Under the Curve* sebesar 0,861 yang termasuk dalam klasifikasi baik (*Good*). Dan penelitian lainnya yang menggunakan algoritma KNN oleh (Pratmanto & Imaniawan, 2023) pada analisis sentimen terhadap aplikasi Canva

menunjukkan algoritma KNN menghasilkan performa yang lebih bagus daripada algoritma *Naive Bayes* pada proses klasifikasi data, dengan capaian akurasi mencapai 83,92%, sementara *Naive Bayes* memperoleh akurasi 77,41%. Kedua penelitian ini menunjukkan bahwa algoritma KNN dapat digunakan untuk sebuah analisis sentimen sebuah aplikasi. Selain itu penelitian terdahulu pada aplikasi JMO yang dilakukan oleh (Fitriyana et al., 2022) pada data ulasan tahun 2018-2022 menunjukkan analisis sentimen dengan metode SVM pada aplikasi JMO menunjukkan akurasi tinggi sebesar 0,96, penelitian oleh (Rizaldi et al., 2023) menggunakan algoritma *Naive Bayes* dengan data tahun 2018-2023 dengan hasil 95% pada *accuracy*, dan penelitian oleh (Azmi et al., 2023) pada aplikasi JMO mendapatkan hasil bahwa metode *Random Forest* mencapai akurasi 93% dengan ulasan positif lebih dominan. Penelitian sebelumnya menunjukkan penggunaan algoritma seperti SVM, *Naive Bayes*, dan *Random Forest* telah diterapkan pada analisis sentimen di aplikasi JMO.

Berdasarkan kajian tersebut Sasaran dari penelitian ini adalah untuk mengamati dan mengevaluasi reaksi pengguna terhadap aplikasi JMO, mengelompokkan sentimen umum dari ulasan pengguna ke dalam dua kategori, yaitu sentimen positif/baik dan sentimen negatif/buruk. Selain itu, penelitian ini akan menerapkan analisis sentimen untuk menghasilkan klasifikasi sentimen dari ulasan pengguna aplikasi JMO dan juga menguji seberapa akurat algoritma KNN dalam analisis sentimen tersebut.

II. METODOLOGI PENELITIAN

Penelitian ini menerapkan teknik analisis data angka dengan pendekatan text mining, yang merupakan suatu langkah untuk mengeksplorasi dan menganalisis data teks dalam volume besar yang tidak teratur, dengan menggunakan perangkat lunak untuk mengidentifikasi pola, konsep, kata penting, tema, serta atribut lain yang terdapat di dalamnya (Fathonah & Herliana, 2021). Metode ini dipilih karena kumpulan data yang akan digunakan merupakan tulisan ulasan konsumen di halaman *Google Play*, yang kemudian akan diproses dan diperiksa dengan menggunakan teknik analisis sentimen untuk mendapatkan pengelompokan opini.

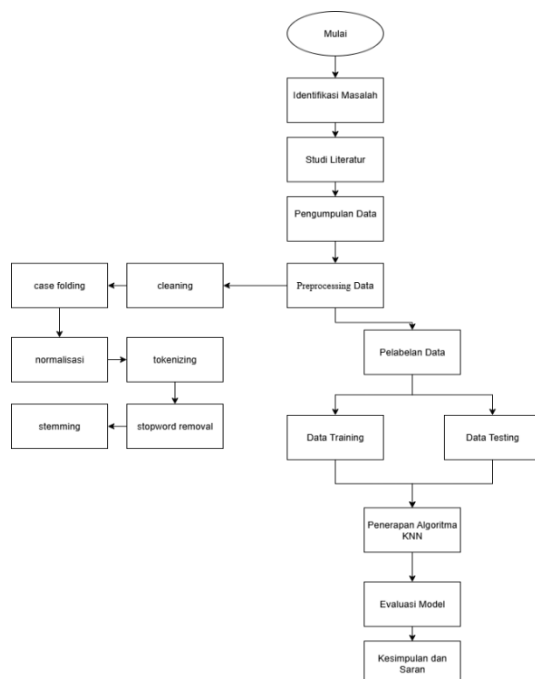
Proses pengambilan data akan dilakukan melalui cara web *scraping* menggunakan perpustakaan API Python '*google-play-scrapers*', Pada proses pengambilan data ini data diambil dari bagian *review* *Google Play Store* menggunakan perpustakaan API dengan bahasa program Python yang dijalankan pada *tools* *Google Colaboratory* untuk mengekstraksi *review* pengguna (Anshari et al., 2023). Data penelitian diperoleh dari *review* pengguna

aplikasi JMO yang terdapat di Google *Play Store*. Ulasan ini berisi opini, komentar, keluhan, maupun pujian yang diberikan pengguna terhadap kinerja aplikasi.

Langkah penelitian selanjutnya adalah dengan melakukan beberapa proses sebelum menggunakan algoritma klasifikasi yaitu melalui *Preprocessing* data seperti: (1) pembersihan data, misalnya menghapus angka, tanda baca, serta karakter lainnya; (2) *case folding* dengan mengganti seluruh huruf menjadi huruf *lowercase*; (3) *normalisasi* kata dengan mengganti kata tidak baku menjadi bentuk baku; (4) *tokenizing* yaitu proses pemecahan kalimat memisahkan kata menjadi sendiri; (5) penghapusan *stopword*, yakni kata-kata yang tidak memiliki makna penting seperti 'dan', 'yang', atau 'di'; serta (6) *stemming* untuk mengembalikan kata ke bentuk paling dasar.

Data ulasan yang sudah diproses kemudian dianalisis guna menentukan apakah ulasan termasuk ulasan sentimen baik (positif) atau buruk (negatif), lalu menggunakan metode klasifikasi, dan algoritma yang akan digunakan pada penelitian ini adalah algoritma K-Nearest Neighbors (KNN) untuk mendapatkan tingkat akurasi. Algoritma KNN dikenal sebagai algoritma yang sederhana namun memiliki tingkat efektivitas yang baik, dengan dasar pemikiran bahwa objek data yang memiliki kesamaan atribut cenderung dikategorikan pada kelas yang sama (Kusuma & Cahyono, 2023).

Metode yang digunakan bersifat kuantitatif dengan menggunakan strategi *machine learning*, dan tahapan secara keseluruhan ditunjukkan pada Gambar 1.



Sumber:(Rabbani et al., 2025)

Gambar 1. Tahapan Penelitian

Dari Gambar 1 tahapan penelitian dijelaskan sebagai berikut:

1. Identifikasi permasalahan, mengidentifikasi permasalahan yaitu banyaknya keluhan maupun *apresiasi* pengguna aplikasi JMO pada ulasan Google *Play Store*, tetapi belum ada kajian analisis sentimen menggunakan algoritma KNN untuk mengetahui kecenderungan opini pengguna.
2. Studi Literatur, pada tahap ini dilakukan penelusuran referensi berupa jurnal, buku, artikel ilmiah, dan penelitian terdahulu yang relevan. Tujuannya adalah untuk memperkuat dasar teori serta menemukan kesenjangan penelitian.
3. Pengumpulan Data, dataset ulasan JMO dikumpulkan melalui Google *Play Scraper*, dengan *dataset* yang berisi teks *review* beserta *rating* pengguna dengan mengambil jumlah 10,000 ulasan yang dikumpulkan menggunakan Python di Google *Colab*. Proses ini dilakukan dengan mengambil data ulasan pengguna aplikasi JMO dari Google *Play Store* menggunakan *tools web scraping*. Proses pengumpulan data dilakukan melalui teknik *web scraping* pada Google *Play Store* dengan memanfaatkan bahasa pemrograman Python yang dijalankan pada platform Google *Colab* serta menggunakan *library* Google *Play Scraper*. Hasil dari proses tersebut berupa kumpulan ulasan pengguna, yang kemudian diolah dalam bentuk *DataFrame* dan diekspor ke dalam format file CSV untuk selanjutnya dianalisis secara lebih mendalam (Mumtaz et al., 2025).
4. *Preprocessing* Data, data ulasan yang telah dikumpulkan kemudian dibersihkan dan diolah melalui beberapa tahap, seperti *cleaning*, *case folding*, *normalisasi*, *tokenizing*, *stopword removal*, dan *stemming*. *Preprocessing* atau Pra-pemrosesan data merupakan tahapan penting yang bertujuan untuk mengurangi gangguan (*noise*) dan memperbaiki ketidak sesuaian pada data, sehingga data yang dihasilkan lebih berkualitas dan siap digunakan pada tahap analisis selanjutnya (Adhitya et al., 2023). Tahapan ini bertujuan untuk membersihkan, menormalkan, serta mempersiapkan data agar dapat diolah secara optimal oleh algoritma analisis maupun model *machine learning*. Pada *dataset* ini, dilakukan proses pembersihan, perbaikan, dan transformasi tanpa mengubah keaslian maupun validitas isi yang terkandung di dalamnya (Putra et al., 2024).
5. Pelabelan Data dan Penerapan Algoritma KNN, Selanjutnya dataset akan kemudian dilakukan pelabelan berdasarkan *rating* ulasan untuk melihat kategori sentimen ulasan positif, dan negatif, kemudian diklasifikasikan dengan algoritma KNN untuk melihat tingkat akurasi. Algoritma K-Nearest

Neighbor (KNN) digolongkan sebagai metode *supervised learning*, sebab proses klasifikasinya dilakukan dengan menentukan label data uji berdasarkan mayoritas kelas dari sejumlah K tetangga terdekat pada data latih (Cholil et al., 2021). Data kemudian dibagi menjadi dua bagian, yakni data *training* dan data *testing*. Pada penelitian ini, algoritma KNN diuji dengan variasi distribusi data *training* dan data *testing* mulai dari 70% banding 30%, 80% banding 20%, dan 90% banding 10%, dan juga variasi nilai K dari 1 hingga 21 untuk membandingkan tingkat akurasi. Hasil pengujian menunjukkan bahwa nilai K yang optimal diperoleh pada penggunaan KNN dengan akurasi tertinggi (Astuti & Nuris, 2022).

- Evaluasi Model, setelah itu model dievaluasi menggunakan *confusion matrix* dan *metrik* seperti akurasi, *precision*, *recall*, dan *f1-score* untuk menilai performa klasifikasi. *Confusion matrix* digunakan untuk mengevaluasi hasil klasifikasi dengan menampilkan jumlah prediksi benar dan salah. Dari hasil *confusion matrix* maka akan dapat perhitungan metrik evaluasi seperti akurasi, *precision*, *recall*, *f1-score*, serta tingkat *error* model (Adhitya et al., 2023). *Confusion Matrix* digunakan untuk mengevaluasi keakuratan klasifikasi melalui perhitungan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sesuai dengan persamaan (1),(2),(3),(4) yang tercantum pada Formula dibawah (Prasetyo et al., 2023).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

Keterangan:

- TP = prediksi benar untuk data positif
- TN = prediksi benar untuk data negatif
- FP = data negatif salah diprediksi positif
- FN = data positif salah diprediksi negatif.

Pada tahap akhir penelitian dilakukan proses evaluasi, yang bertujuan untuk mengukur kinerja model melalui perhitungan *Confusion Matrix*, *Receiver Operating Characteristic* (ROC), serta *Area Under Curve* (AUC) (Pramudita et al., 2024). Dengan pengukuran akurasi hasil klasifikasi dapat dilakukan dengan menggunakan nilai AUC (*Area Under Curve*), dengan kategori penilaian (Astuti & Nuris, 2022) sebagai berikut:

- excellent* = 0,90–1,00 sangat baik
- good* = 0,80–0,90 baik
- fair* = 0,70–0,80 cukup
- poor* = 0,60–0,70 kurang

e. *failure* = 0,50–0,60 gagal

- Kesimpulan dan Saran, hal terakhir pada penelitian ini adalah membuat kesimpulan berdasarkan hasil penelitian dan memberikan rekomendasi yang dapat berguna bagi pengembangan aplikasi JMO maupun penelitian berikutnya.

III. HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Data yang digunakan pada penelitian ini adalah ulasan-ulasan pengguna pada halaman Google *Play Store* yang telah diambil dengan bahasa pemrograman Python menggunakan API *google-play-scraper* pada *Google Colab*, dengan mengambil data ulasan *Review ID*, *Username*, *Rating*, *Review text*, *Date* dan menghasilkan sebuah file data dengan format CSV, dapat dilihat pada Gambar 2.

Review ID	Username	Rating	Review text	Date
6d844b4-2642-42ab-89ab-155816a487ae	Abdul Rofiq	3	list nand	2025-08-31 01:16:05
16ea92432-1783-46a1-8323-4c1c16694a82	Isyori Muliem	5	good	2025-08-31 01:12:16
1419898f-596a-4069-915d-4175697199ac	Andi Saprol	2	jmo mental tk bisa dibuka	2025-08-31 03:59:01
9856d6e-7a2c-431b-ae3e-3c3a69ab88dc	Okta Omaruddin	1	Kanapa akun JMO saya gabisa dibuka woi saya ud...	2025-08-31 03:54:29
608be637-576f-4826-9674-77edf8234418	Carot Nitro	4	baik	2025-08-31 03:50:55
cc781304-fcd1-4d72-9ed3-5216b403399b	Friska Sarah	5	bagus	2025-08-31 03:47:55
56c2ba29-c09f-4a11-8486-49670aa05c39	Eko Susanto	5	Alasan mudah	2025-08-31 03:29:39
8f259027-5a2b-4b52-ae07-dcaaf9625a8c	Muhammad Arifanyah	2	setelah di update malah gabisa dibuka	2025-08-31 03:04:04
562d1015-b56a-431d-b0ed-c204b8765d44	Amar Ikhsanudin	5	top	2025-08-31 03:03:49
c679336d-4293-4486-9a34-ea8bb1154882	Pipang Putra	4	ok	2025-08-31 03:03:44
c30440aa-2351-4571-9ee8-e1d8-7272bd4d	Heru Siswanto	5	baik kenapa sekarang jmo gak bisa di buka	2025-08-31 02:50:36

Sumber: (Rabbani et al., 2025)

Gambar 2. Data Scraping Ulasan

Langkah selanjutnya adalah melakukan *Preprocessing* Data untuk mempersiapkan data sebelum melakukan modeling algoritma KNN, dikarenakan data yang baru didapat masih belum bersih dan masih banyak langkah-langkah yang harus dilakukan untuk memproses data.

2. Preprocessing Data

Pada langkah *Preprocessing* Data ada beberapa hal yang dilakukan kepada data yang sudah didapat agar dapat diolah secara optimal oleh algoritma KNN, yaitu:

- Cleaning* pembersihan data dengan menghapus tanda baca, angka, dan karakter khusus yang terdapat pada teks ulasan,

Tabel 1. *Cleaning* Data

Data sebelum diolah	Data sesudah diolah
Tidak bisa buka aplikasi JMO. Dari layar pembukaan langsung keluar. Sdh coba install ulang.	Tidak bisa buka aplikasi JMO Dari layar pembukaan langsung keluar Sdh coba install ulang

Sumber: (Rabbani et al., 2025)

- Case folding* mengubah semua huruf menjadi huruf kecil,

Tabel 2. *Case folding* Data

Data sebelum diolah	Data sesudah diolah
Tidak bisa buka aplikasi JMO Dari	tidak bisa buka aplikasi jmo dari layar

ulasan negatif banyak berisi keluhan pengguna terkait aplikasi.



Sumber: (Rabbani et al., 2025)

Gambar 5. WordCloud Sentimen Positif

Sedangkan pada wordcloud sentimen positif di Gambar 5, kata-kata yang dominan adalah “mudah”, “bantu”, “cepat”, “bagus”, “layanan”, “proses”, dan “terima kasih”, Kata-kata ini menunjukkan bahwa pengguna yang memberikan ulasan positif merasa aplikasi membantu, mudah digunakan, prosesnya cepat, serta memberikan layanan yang bermanfaat.

4. Penerapan Algoritma KNN

Proses selanjutnya adalah mendistribusikan data menjadi data *training* dan data *testing*, dengan distribusi 70%-30%, 80%-20%, dan 90%-10%, untuk bereksperimen mencari distribusi mana yang akan memberikan tingkat akurasi yang paling tinggi berdasarkan nilai K dari 1 sampai 21, yang akan mendapatkan hasil paling tinggi tingkat akurasinya.

Tabel 7. Eksperimen dengan Distribusi data *Training* & data *Testing* pada Perubahan Nilai K

Distribusi 70%-30%		Distribusi 80%-20%		Distribusi 90%-10%	
Nilai K	Akurasi	Nilai K	Akurasi	Nilai K	Akurasi
K = 1		K = 1		K = 1	
K = 3		K = 3		K = 3	
K = 5		K = 5		K = 5	
K = 7		K = 7		K = 7	
K = 11		K = 11		K = 11	
K = 13	65.4%	K = 13	66.7%	K = 13	68.2%
K = 15	68.5%	K = 15	68.8%	K = 15	68.7%
K = 17	70.5%	K = 17	70.9%	K = 17	70.5%
K = 19	71.9%	K = 19	72.5%	K = 19	72.0%
K = 21	73.9%	K = 21	73.2%	K = 21	72.5%
			73.1%		73.3%
			73.7%		74.0%
			73.8%		75.9%
			73.9%		75.8%
			73.6%		76.2%

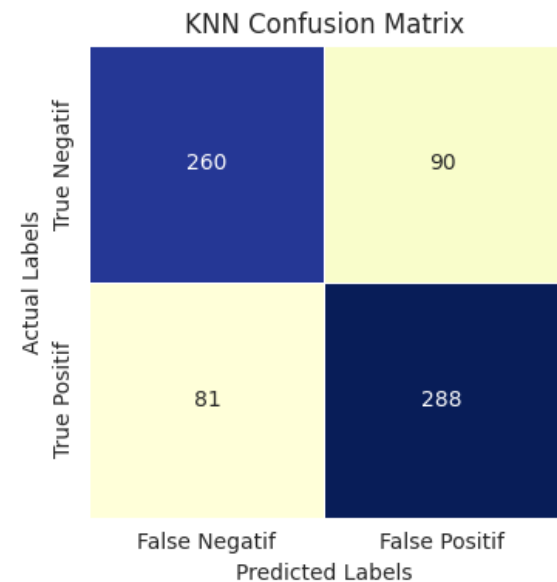
Sumber: (Rabbani et al., 2025)

Berdasarkan dari Tabel 7 dengan bereksperimen pada jumlah distribusi data dan nilai K, telah mendapatkan hasil akurasi tertinggi pada algoritma KNN penelitian ini adalah pada distribusi data *Training* 90% - data *Testing* 10% dengan nilai K = 21 yang mendapatkan akurasi

76.2%.

5. Evaluasi Data

Dari hasil penerapan algoritma di atas yang telah mendapatkan hasil akurasi 76.2% dari nilai K = 21 pada distribusi data 90%-10%, maka *Confusion Matrix* dari data tersebut dapat dilihat di Gambar 6.



Sumber: (Rabbani et al., 2025)

Gambar 6. Confusion Matrix

Dari confusion matrix ini bisa mendapatkan perhitungan nilai *Accuracy*, *Precision*, *Recall*, *F1-Score*, Kurva ROC, dan nilai AUC pada di bawah ini:

a. Perhitungan akurasi dengan rumus:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$Accuracy = \frac{288+260}{288+260+90+81}$$

$$Accuracy = \frac{548}{719}$$

$$Accuracy = 0.762 = 76.2\%$$

b. Perhitungan *Precision* dengan rumus:

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Precision = \frac{288}{288+90}$$

$$Precision = \frac{288}{378}$$

$$Precision = 0.7619 = 76.2\%$$

c. Perhitungan *Recall* dengan rumus:

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$Recall = \frac{288}{288+81}$$

$$Recall = \frac{288}{369}$$

$$Recall = 0.7805 = 78.0\%$$

d. Perhitungan *F1-Score* dengan rumus:

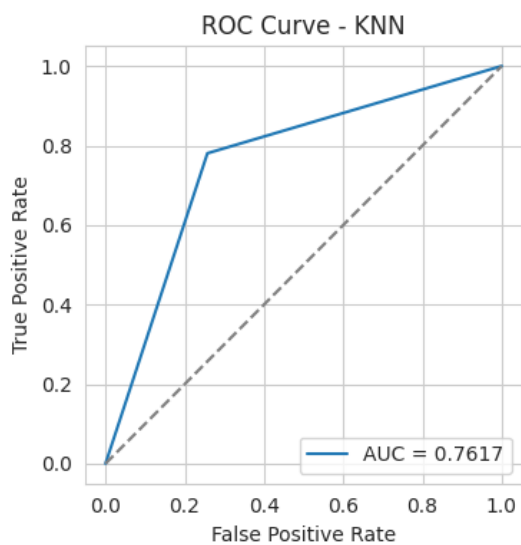
$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

$$F1 = 2 \times \frac{0.7619 \times 0.7805}{0.7619 + 0.7805}$$

$$F1 = 2 \times \frac{0.5946}{1.5424}$$

$$F1 = 0.771 = 77.1\%$$

e. Kurva ROC, dan nilai AUC:



Sumber: (Rabbani et al., 2025)

Gambar 7. Kurva ROC dan nilai AUC

Dengan perhitungan yang telah dilakukan bisa dilihat dari penerapan algoritma KNN telah mendapatkan hasil evaluasi keakuratan klasifikasi data dengan akurasi 76.2%, *precision* 76.2%, *recall* 78.0%, dan *F1-score* 77.1%. Dan untuk mengukur kinerja model KNN telah terdapat Kurva ROC, dan nilai AUC 0.7617 menunjukkan hasil Fair atau hasil yang cukup.

IV. KESIMPULAN

Penelitian ini dilakukan untuk melihat bagaimana reaksi pengguna terhadap aplikasi JMO, dan menganalisis seperti apa sentimen umum pengguna aplikasi JMO (Jamsostek Mobile) menurut ulasan-ulasan *review* yang ada pada halaman Google *Play Store*, menggunakan algoritma klasifikasi data K-Nearest Neighbor (KNN). dengan melakukan penelitian analisis sentimen ini, berikutlah kesimpulan yang dapat diambil dari penelitian ini:

1. Dari hasil penelitian yang telah dilakukan, dengan langkah mengambil data yang berjumlah 10,000 data ulasan oleh proses *scrapping*, dan selanjutnya mengolah data melalui *Preprocessing* data untuk mendapatkan hasil 7,185 data yang telah dibersihkan, dari data tersebut menunjukkan bahwa dari keseluruhan data 51.38% sentimen yang didapat adalah sentimen positif, dengan kata-kata seperti "mudah", "bantu", "cepat", "bagus", "layanan", "proses", dan "terima kasih" menunjukkan bahwa sentimen positif sedikit lebih dominan dibandingkan sentimen negatif dengan selisih data pada 2.76%.
 2. Meskipun sudah banyak pengguna yang merasa terbantu, masih banyak pula pengguna yang merasakan kendala pada aplikasi dan meninggalkan keluhan mereka pada halaman Google *Play Store*, hal ini dapat dilihat dari hasil pelabelan data yang menunjukkan bahwa masih terdapat 48.62% sentimen negatif.
 3. Dengan penerapan algoritma KNN pada analisis sentimen ini, dimana nilai akurasi didapatkan dari eksperimentasi pembagian data *training* dengan data *testing*, dan pencarian nilai K, menunjukkan pada distribusi data 90% *training* - 10% *testing* menggunakan nilai K = 21 sudah bisa mendapatkan keakuratan klasifikasi data dengan nilai akurasi tertinggi pada 76.2%.
 4. Hasil evaluasi yang didapatkan pada penggunaan algoritma KNN ini menunjukkan nilai *precision* 76.2%, nilai *recall* 78.0%, nilai *F1-score* 77.1%, dan juga kinerja model KNN mendapatkan nilai *Area Under Curve* (AUC) = 0.7617, hal ini menunjukkan bahwa hasil algoritma KNN mendapatkan nilai atau hasil yang *Fair* atau cukup baik, mengartikan model klasifikasi algoritma KNN mampu memberikan hasil klasifikasi yang seimbang pada setiap kelas, sehingga evaluasi lebih adil dan bebas dari bias.
- Mengacu pada hasil penelitian ini, beberapa rekomendasi dapat dirumuskan sebagai berikut:
1. Penelitian selanjutnya disarankan menggunakan jumlah dan variasi data ulasan yang lebih besar agar hasil analisis sentimen lebih representatif.
 2. Perlu dilakukan perbandingan dengan metode dan algoritma klasifikasi lain untuk memperoleh gambaran kinerja algoritma KNN yang lebih menyeluruh.
 3. Bagi pengembang aplikasi JMO (Jamsostek Mobile), hasil penelitian ini dapat dimanfaatkan sebagai masukan untuk meningkatkan kualitas layanan, kinerja aplikasi, dan pengalaman pengguna berdasarkan sentimen negatif.
 4. Bagi BPJS Ketenagakerjaan, analisis sentimen ulasan pengguna dapat dijadikan bahan evaluasi berkala dalam

pengembangan layanan digital agar lebih sesuai dengan kebutuhan pengguna.

V. REFERENSI

- Adhitya, R. R., Witanti, W., & Yuniarti, R. (2023). Perbandingan Metode Cart Dan Naïve Bayes Untuk Klasifikasi Customer Churn. *INFOTECH Journal*, 9(2), 307–318. <https://doi.org/10.31949/infotech.v9i2.5641>
- Akbar, F. M. N. (2024). Metode KNN (K-Nearest Neighbor) untuk Menentukan Kualitas Air. *Jurnal Tekno Kompak*, 18(1), 28–40. <https://doi.org/10.33365/jtk.v18i1.3241>
- Anshari, R. AL, Alam, S., & Hafid T, M. (2023). Komparasi Payment Digital Untuk Analisis Sentimen Berdasarkan Ulasan Di Google Playstore Menggunakan Metode Support Vector Machine. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 2(3), 118–128. <https://doi.org/10.55123/storage.v2i3.2337>
- Astuti, P., & Nuris, N. (2022). Penerapan Algoritma KNN Pada Analisis Sentimen Review Aplikasi Peduli Lindungi. *Computer Science (CO-SCIENCE)*, 2(2), 137–142. <https://doi.org/10.31294/coscience.v2i2.1258>
- Azmi, T. A. U., Hakim, L., Novitasari, D. C. R., & Utami, W. D. U. (2023). Application Random Forest Method for Sentiment Analysis in Jamsostek Mobile Review. *Telematika*, 20(1), 117. <https://doi.org/10.31315/telematika.v20i1.8868>
- Cholil, S. R., Handayani, T., Prathivi, R., & Ardianita, T. (2021). Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 6(2), 118–127. <https://doi.org/10.31294/ijcit.v6i2.10438>
- Dewi, A. K., & Sulastri. (2022). Analisis Sentimen Ekspedisi Sicepat Dari Ulasan Google Play Menggunakan Algoritma Naïve Bayes. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 796–805. <https://doi.org/10.35957/jatisi.v9i2.1802>
- Fathonah, F., & Herliana, A. (2021). Penerapan Text Mining Analisis Sentimen Mengenai Vaksin Covid - 19 Menggunakan Metode Naïve Bayes. *Jurnal Sains Dan Informatika*, 7(2), 155–164. <https://doi.org/10.34128/jsi.v7i2.331>
- Fitriyana, V., Hakim, L., Novitasari, D. C. R., & Asyhar, A. H. (2022). Analisis Sentimen Ulasan Aplikasi Jamsostek Mobile Menggunakan Metode Support Vector Machine. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(4), 3436–3442. <https://doi.org/10.24002/jbi.v14i01.6909>
- Kevin, K., Enjeli, M., & Wijaya, A. (2024). Analisis Sentimen Penggunaan Aplikasi Kinemaster Menggunakan Metode Naive Bayes. *Jurnal Ilmiah Computer Science*, 2(2), 89–98. <https://doi.org/10.58602/jics.v2i2.24>
- Khairi, A., Ghozali, A. F., & Hidayah, A. D. N. (2021). Implementasi K-Nearest Neighbor (KNN) untuk Mengklasifikasi Masyarakat Pra-Sejahtera Desa Sapikerep Kecamatan Sukapura. *TRILOGI: Jurnal Ilmu Teknologi, Kesehatan, Dan Humaniora*, 2(3), 319–323. <https://doi.org/10.33650/trilogi.v2i3.2878>
- Kusuma, I. H., & Cahyono, N. (2023). Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Informatika: Jurnal Pengembangan IT*, 8(3), 302–307. <https://doi.org/10.30591/jpit.v8i3.5734>
- Mumtaz, J. A., Komariansyah, K. K., Holik, W., Gumelar, M. G., Pratama, R., & Lestari, H. R. (2025). Analisis Sentimen Ulasan Aplikasi HeyJapan di Google Play Store Menggunakan Algoritma NLP pembelajaran bahasa Jepang. Analisis ini dilakukan dengan memanfaatkan metode NLP memberikan gambaran komprehensif mengenai persepsi pengguna terhadap kedua aplikasi. 3. <https://doi.org/10.61132/pragmatik.v3i3.1801>
- Pramudita, R., Muis, S., Safitri, N., & Shafirawati, F. (2024). Optimasi Algoritma Machine Learning Menggunakan Teknik Bagging Pada Klasifikasi Diagnosis Kanker Payudara. *Tematik*, 11(1), 128–134. <https://doi.org/10.38204/tematik.v11i1.1928>
- Prasetyo, S. D., Hilabi, S. S., & Nurapriani, F. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN. *Jurnal KomtekInfo*, 10, 1–7. <https://doi.org/10.35134/komtekinfo.v10i1.330>
- Pratmanto, D., & Imaniawan, F. F. D. (2023). Analisis Sentimen Terhadap Aplikasi Canva Menggunakan Algoritma Naive Bayes Dan K-Nearest Neighbors. *Computer Science (CO-SCIENCE)*, 3(2), 110–117. <https://doi.org/10.31294/coscience.v3i2.1917>
- Putra, F., Tahiyat, H. F., Ihsan, R. M., Rahmaddeni, & Efrizoni, L. (2024). Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia. *Institut*

- Riset Dan Publikasi Indonesia (IRPI)*
MALCOM: Indonesian Journal of Machine Learning and Computer Science, 4(January), 273–281.
<https://doi.org/10.57152/malcom.v4i1.1085>
- Rabbani, G. F., Ramanda, K., Sukmana, S. H., Azizah, Q. N., & Sikumbang, E. D. (2025). *Laporan Akhir Penelitian: Penerapan Algoritma K-Nearest Neighbors (KNN) Dalam Analisis Sentimen Ulasan Pengguna Aplikasi JMO (Jamsostek Mobile)*.
- Rizaldi, S. A., Alam, S., & Kurniawan, I. (2023). Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 2(3), 109–117.
<https://doi.org/10.55123/storage.v2i3.2334>
- Samudera, M. F. D., & Pertiwi, V. I. (2022). Inovasi Pelayanan Publik Melalui Jamsostek Mobile (JMO) (Studi Kasus di BPJS Ketenagakerjaan Cabang Rungkut Kota Surabaya). *JPAP: Jurnal Penelitian Administrasi Publik*, 8(1), 152–172.
<https://doi.org/10.30996/jpap.v8i1.6617>
- Sutrisno, H. (2020). Pengaruh Bpjs Ketenagakerjaan Dalam Meningkatkan Kesejahteraan Tenaga Kerja. *PREPOTIF: Jurnal Kesehatan Masyarakat*, 4(1), 78–84.
<https://doi.org/10.31004/prepotif.v4i1.670>