

Perbandingan Algoritma K-NN dan SVM dalam Analisis Sentimen Ulasan Pengguna Aplikasi JMO

Comparison of K-NN and SVM Algorithms in JMO Application for Sentiment Analysis

Sevia Anggreini Simanjuntak¹, Martaleli Bettiza², Fortia Magfira*³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik dan Teknologi Kemaritiman
Universitas Maritim Raja Ali Haji

Jl. Politeknik, Senggarang, Tanjungpinang Provinsi Kepulauan Riau 29115

e-mail:

12101020083@student.umrah.ac.id, mbettiza@umrah.ac.id, [*3fortiamagfira@umrah.ac.id](mailto:fortiamagfira@umrah.ac.id)

Abstrak

Perkembangan teknologi digital mendorong lembaga layanan publik untuk mengadopsi aplikasi berbasis mobile guna meningkatkan kualitas dan efisiensi pelayanan kepada masyarakat. Salah satu implementasi layanan tersebut adalah aplikasi Jamsostek Mobile (JMO) yang dikembangkan oleh BPJS Ketenagakerjaan sebagai sarana pelayanan digital bagi peserta jaminan sosial tenaga kerja. Namun, tingginya jumlah ulasan negatif di Google PlayStore mengindikasikan adanya permasalahan terkait kinerja aplikasi, khususnya pada aspek kemudahan penggunaan, stabilitas sistem, dan proses layanan seperti login serta klaim manfaat. Oleh karena itu, penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap kinerja dan kualitas layanan aplikasi JMO berdasarkan ulasan pengguna, serta membandingkan performa algoritma Support Vector Machine (SVM) dan K-Nearest Neighbor (K-NN) dalam klasifikasi sentimen. Data yang digunakan berupa 10.000 ulasan pengguna aplikasi JMO yang dikumpulkan dari Google PlayStore. Data tersebut diproses melalui tahapan preprocessing meliputi pembersihan teks, normalisasi, tokenisasi, penghapusan stopword, dan stemming. Representasi fitur dilakukan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Dataset dibagi dengan rasio 80% sebagai data latih dan 20% sebagai data uji. Kinerja model dievaluasi menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma SVM memberikan performa yang lebih baik dengan akurasi sebesar 90,00% dibandingkan K-NN sebesar 78,25%. Dengan demikian, algoritma SVM dinilai lebih optimal untuk mengklasifikasikan sentimen pengguna terhadap aplikasi JMO.

Kata kunci— Analisis Sentimen, K-NN, SVM, Aplikasi JMO, Machine Learning

Abstract

The rapid development of digital technology has encouraged public service institutions to adopt mobile-based applications to improve service quality and efficiency. One such implementation is the Jamsostek Mobile (JMO) application developed by BPJS Ketenagakerjaan to facilitate digital services for social security participants. However, the high number of negative reviews on Google Play Store indicates persistent issues related to application performance, particularly in terms of usability, system stability, and service processes such as login and benefit claims. Therefore, this study aims to analyze user sentiment toward the performance and service quality of the JMO application based on user reviews, as well as to compare the classification

performance of the Support Vector Machine (SVM) and K-Nearest Neighbor (K-NN) algorithms. The dataset consists of 10,000 user reviews collected from Google Play Store. The data were processed through several preprocessing stages, including text cleaning, normalization, tokenization, stopword removal, and stemming. Feature representation was conducted using the Term Frequency–Inverse Document Frequency (TF-IDF) method. The dataset was divided into 80% training data and 20% testing data. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. The experimental results show that the SVM algorithm outperforms K-NN, achieving an accuracy of 90.00%, while K-NN achieved 78.25%. These findings indicate that SVM is more effective for sentiment classification of user reviews related to the JMO application.

Keywords—Sentiment Analysis, KNN, SVM, JMO Application, Machine Learning

1. PENDAHULUAN

Perkembangan teknologi digital yang semakin pesat mendorong berbagai sektor untuk beradaptasi melalui penerapan layanan berbasis aplikasi guna meningkatkan efisiensi dan kemudahan akses bagi penggunanya. Salah satu layanan digital yang diterapkan dalam bidang ketenagakerjaan di Indonesia adalah Jamsostek *Mobile* (JMO) yang dikembangkan oleh BPJS Ketenagakerjaan sebagai sarana untuk memberikan kemudahan bagi peserta dalam mengakses informasi kepesertaan, pengajuan klaim manfaat, serta layanan jaminan sosial tenaga kerja secara daring [1]. Aplikasi JMO hadir sebagai bentuk transformasi layanan publik menuju sistem digital yang adaptif terhadap kebutuhan masyarakat modern. Berdasarkan laporan dari Media Indonesia [2], hingga Desember 2024 aplikasi JMO telah diunduh lebih dari 25,3 juta kali di *Google PlayStore*, menunjukkan tingginya tingkat adopsi pengguna di seluruh wilayah Indonesia. Namun, seiring dengan meningkatnya penggunaan, muncul pula berbagai permasalahan seperti kesulitan dalam proses pendaftaran, kegagalan saat klaim, dan gangguan teknis yang menimbulkan keluhan dari pengguna. Kondisi tersebut mencerminkan bahwa meskipun aplikasi ini telah mempermudah akses layanan, masih terdapat aspek fungsional yang perlu ditingkatkan.

Beragamnya ulasan pengguna di *Google PlayStore* menjadi sumber informasi penting untuk mengevaluasi kualitas aplikasi. Penelitian sebelumnya yang menggunakan algoritma *Naïve Bayes* terhadap 5.000 ulasan pengguna JMO menunjukkan mayoritas ulasan bersentimen negatif dengan kata dominan seperti *login*, *update*, dan *gagal* [3]. Hasil tersebut menunjukkan perlunya penelitian lanjutan dengan data terbaru dan penerapan algoritma yang memiliki kemampuan klasifikasi lebih baik guna memperoleh hasil analisis yang lebih akurat. Dalam bidang *Artificial Intelligence* (AI), khususnya *Machine Learning* (ML), analisis sentimen merupakan salah satu penerapan penting dari *Natural Language Processing* (NLP) yang bertujuan untuk memahami dan menginterpretasikan opini atau emosi dalam teks secara otomatis [4]. Berbagai algoritma ML telah digunakan untuk analisis sentimen, seperti *Support Vector Machine* (SVM), *K-Nearest Neighbor* (K-NN), *Naïve Bayes*, *Decision Tree*, dan *Logistic Regression* [5]. Setiap algoritma memiliki kelebihan dan kekurangan masing-masing dalam hal kompleksitas, efisiensi, serta akurasi hasil klasifikasi.

Algoritma SVM dikenal memiliki performa tinggi dalam mengklasifikasikan data berdimensi besar serta mampu menemukan hyperplane optimal yang memisahkan data ke dalam kategori tertentu [6]. Sementara itu, K-NN memiliki keunggulan dalam kesederhanaan implementasi dengan prinsip pengelompokan berdasarkan kedekatan jarak antar data [7]. Beberapa penelitian menunjukkan bahwa SVM unggul dalam hal akurasi klasifikasi teks, namun K-NN tetap kompetitif pada dataset dengan pola distribusi tertentu [8]. Penelitian sebelumnya [9] menunjukkan bahwa algoritma SVM mencapai akurasi tertinggi sebesar 83,33% dalam klasifikasi opini publik mengenai vaksinasi COVID-19 dibandingkan *Naïve Bayes* dan *Decision Tree*. Sementara itu, penelitian [6] menunjukkan bahwa K-NN memberikan akurasi tertinggi sebesar 81% dalam prediksi pembatalan reservasi hotel dibandingkan *Logistic Regression*. Hasil tersebut

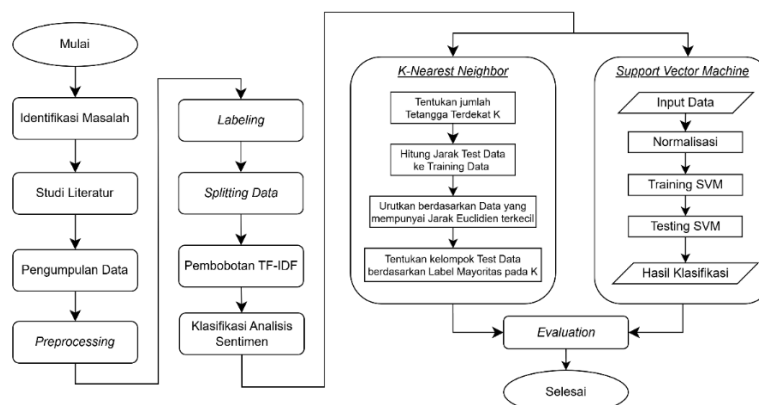
Perbandingan Algoritma K-NN dan SVM dalam Analisis Sentimen Ulasan Pengguna Aplikasi JMO

memperlihatkan bahwa performa kedua algoritma dapat berbeda tergantung pada karakteristik dan kompleksitas data yang digunakan.

Penelitian ini bertujuan untuk menganalisis dan membandingkan performa algoritma *K-Nearest Neighbor* (K-NN) dan *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen ulasan pengguna terhadap kualitas layanan aplikasi Jamsostek Mobile (JMO) di *Google PlayStore*, yang mencerminkan persepsi pengguna terkait kemudahan penggunaan, keandalan sistem, dan pengalaman layanan BPJS Ketenagakerjaan secara digital. Dataset yang digunakan terdiri dari 10.000 ulasan pengguna yang dikumpulkan melalui *web scraping* pada periode Januari 2023 hingga Desember 2024, kemudian diproses melalui tahapan *text preprocessing* meliputi pembersihan teks, normalisasi, tokenisasi, penghapusan *stopword*, dan *stemming*. Evaluasi performa kedua algoritma dilakukan menggunakan metrik akurasi, presisi, *recall*, *F1-score*, serta waktu komputasi, dengan perbandingan terhadap label aktual berdasarkan kamus InSet *Lexicon*. Gap penelitian ini terletak pada masih terbatasnya studi yang membandingkan secara langsung algoritma K-NN dan SVM pada dataset ulasan JMO yang besar dan terbaru. Oleh karena itu, penelitian ini bertujuan untuk menentukan algoritma terbaik dalam klasifikasi sentimen ulasan pengguna aplikasi JMO serta memberikan gambaran objektif mengenai kualitas layanan aplikasi berdasarkan opini pengguna.

2. METODE PENELITIAN

2.1 Alur Penelitian



Gambar 1. Alur Penelitian

Alur penelitian ini diawali dengan identifikasi masalah untuk menentukan fokus penelitian terkait analisis sentimen ulasan pengguna aplikasi JMO. Selanjutnya dilakukan studi literatur guna memperkuat dasar teori dan menentukan metode yang digunakan. Tahap berikutnya adalah pengumpulan data ulasan dari *Google PlayStore* menggunakan *google-play-scraper*, kemudian dilakukan *preprocessing* mencakup pembersihan teks, normalisasi, tokenisasi, penghapusan *stopword*, dan *stemming*. Setelah itu dilakukan labeling menggunakan kamus InSet *Lexicon* serta pembagian data menjadi 80% data latih dan 20% data uji [14]. Proses dilanjutkan dengan pembobotan TF-IDF untuk mengubah teks menjadi representasi numerik. Data yang telah diproses kemudian diklasifikasikan menggunakan algoritma *K-Nearest Neighbor* (K-NN) dan *Support Vector Machine* (SVM). Tahap terakhir adalah evaluasi model berdasarkan metrik akurasi, presisi, *recall*, dan *F1-score* untuk menilai performa klasifikasi sentimen.

2.2 Metode K-Nearest Neighbor

K-Nearest Neighbor (K-NN) merupakan algoritma *supervised learning* yang digunakan untuk klasifikasi dengan prinsip mencari sejumlah tetangga terdekat (*nearest neighbors*) dari data uji berdasarkan jarak tertentu, seperti *Euclidean Distance*. Nilai kelas ditentukan berdasarkan

mayoritas kelas dari K tetangga terdekat. Pemilihan nilai K berpengaruh terhadap akurasi model; nilai K kecil berpotensi menyebabkan *overfitting*, sedangkan nilai besar dapat menyebabkan *underfitting*. K-NN mudah diimplementasikan namun memiliki kelemahan dalam efisiensi komputasi karena setiap prediksi memerlukan perhitungan jarak ke seluruh data latih [7]. Untuk menghitung Jarak *Euclidean* dalam K-NN dinyatakan pada persamaan (1).

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 \dots + (x_n - y_n)^2} \quad (1)$$

2.3 Metode Support Vector Machine

Support Vector Machine (SVM) merupakan algoritma *supervised learning* yang digunakan untuk tugas klasifikasi dengan cara mencari hyperplane optimal yang memisahkan data ke dalam kelas berbeda dengan margin maksimum [14]. Semakin besar margin antara kelas, semakin baik kemampuan generalisasi model terhadap data baru. SVM dikenal unggul dalam menangani data berdimensi tinggi dan efektif digunakan untuk analisis teks seperti klasifikasi sentimen. Menurut penelitian [10], metode *Sequential Training SVM* digunakan untuk menemukan hyperplane pemisah dengan akurasi terbaik melalui proses pelatihan bertahap. Dalam penelitian ini, metode tersebut diterapkan untuk membangun model klasifikasi sentimen ulasan aplikasi JMO karena dinilai lebih efisien dibandingkan pendekatan konvensional. Tahapan implementasinya meliputi:

1. Inisialisasi Parameter, seperti *Lagrange Multiplier* (α_i), konstanta kernel (c), parameter regulasi (λ), *gamma* (γ), *complexity* (C), dan jumlah iterasi maksimum.
2. Perhitungan *Matriks Hessian* (D) menggunakan kernel *polynomial* derajat dua menggunakan rumus pada persamaan (2), yang merepresentasikan hubungan antar data dalam ruang fitur. Rumus perhitungan *Matriks Hessian* menggunakan persamaan (3).

$$K(x_i, x_j) = (x_i \cdot x_j + c)^d \quad (2)$$

$$D_{ij} = y_i \cdot y_j \cdot (K(x_i, x_j)) + \lambda^2 \quad (3)$$

3. Perhitungan *Error* (E_i), variabel tunggal (δa_i), dan pembaruan nilai *Lagrange Multiplier* (a_i) untuk meminimalkan kesalahan klasifikasi. Maka, dilakukan perhitungan menggunakan persamaan (4)-(6).

$$E_i = \sum_{j=1}^N a_i \cdot D_{ij} \quad (4)$$

$$\delta a_i = \min\{\max[\gamma(1 - E_i); -a_i]; C - a_i\} \quad (5)$$

$$a_i = a_i + \delta a_i \quad (6)$$

4. Iterasi Pelatihan dilakukan hingga parameter optimal diperoleh.
5. Penentuan *Support Vector* (SV) berdasarkan nilai $a_i > 0$.
6. Perhitungan Bias (b) untuk menentukan posisi *hyperplane* pemisah. Nilai bias dihitung menggunakan rumus pada persamaan (7).

$$b = -\frac{1}{2} \left(\sum_{i=0}^N a_i y_i K(x_i, x^+) + \sum_{i=0}^N a_i y_i K(x_i, x^-) \right) \quad (7)$$

7. Klasifikasi Sentimen dilakukan dengan fungsi keputusan SVM, di mana hasil $h(x) > 0$ dikategorikan positif, $h(x) < 0$ negatif, dan $h(x) = 0$ netral. Perhitungan hasil klasifikasi menggunakan persamaan (8).

$$h(x) = \left(\sum_{i=0}^N a_i y_i K(x, x_i) + b \right) \quad (8)$$

2.4 Google Collaboratory

Dalam penelitian ini, tools yang digunakan meliputi bahasa pemrograman *Python* sebagai bahasa utama pengolahan data, *Google Collab* sebagai lingkungan pengembangan untuk proses *preprocessing*, pelatihan, dan pengujian model, serta library *Scikit-learn* untuk implementasi algoritma *K-Nearest Neighbor* (K-NN) dan *Support Vector Machine* (SVM). Selain itu, *Pandas* dan *NumPy* digunakan untuk pengolahan data, *Sastrawi* untuk proses stemming bahasa Indonesia, *Matplotlib* dan *WordCloud* untuk visualisasi hasil analisis, serta *google-play-scraper* untuk pengambilan data ulasan dari *Google PlayStore*.

2.5 Tahapan Pengumpulan Data Set

Dataset yang digunakan dalam penelitian ini berasal dari ulasan pengguna aplikasi *Jamsostek Mobile* (JMO) di *Google PlayStore*. Data dikumpulkan menggunakan library *google-play-scraper* pada bahasa pemrograman *Python*, sehingga proses pengambilan data dapat dilakukan secara otomatis, cepat, dan terstruktur tanpa *parsing* manual. Data yang diperoleh berupa teks ulasan pengguna yang mencerminkan pengalaman, keluhan, maupun kepuasan terhadap layanan aplikasi JMO. Total dataset yang digunakan berjumlah 10.000 ulasan, yang dikumpulkan pada periode Januari 2023 hingga Desember 2024, dan dijadikan sebagai dasar analisis sentimen dalam penelitian ini.

2.6 Tahapan Preprocessing Data

Pada tahap *preprocessing* data, ulasan yang telah dikumpulkan diproses agar bersih dan siap digunakan dalam klasifikasi. Proses ini mencakup beberapa langkah utama, yaitu:

1. *Cleaning Data*, untuk menghapus elemen yang tidak relevan.
2. *Case Folding*, mengubah seluruh huruf menjadi huruf kecil.
3. *Normalisasi*, mengganti kata tidak baku atau slang menjadi bentuk baku.
4. *Tokenization*, memecah kalimat menjadi token kata.
5. *Stopword Removal*, menghapus kata umum yang tidak memiliki makna penting.
6. *Stemming*, mengembalikan setiap kata ke bentuk dasarnya agar analisis lebih efisien.

2.7 Tahapan Labeling dan Splitting Data

Setelah *preprocessing*, dilakukan proses labeling sentimen menggunakan kamus *Indonesian Sentiment* (InSet) *Lexicon*. Setiap ulasan diberi skor polaritas berdasarkan penjumlahan nilai sentimen dari kata-kata penyusunnya, kemudian diklasifikasikan ke dalam tiga kelas sentimen, yaitu positif (skor > 0), negatif (skor < 0), dan netral (skor $= 0$). Selanjutnya, dataset dibagi menjadi data latih dan data uji dengan rasio 80:20, di mana 80% digunakan sebagai data latih dan 20% sebagai data uji. Rasio ini dipilih karena umum digunakan dalam penelitian *machine learning* [14] untuk memberikan porsi data yang cukup besar pada proses pembelajaran model, sekaligus mempertahankan data uji yang representatif untuk evaluasi performa secara objektif. Sebelum pembagian, data diacak terlebih dahulu untuk menjaga proporsi kelas sentimen.

2.8 Tahapan Pembobotan TF-IDF

Pada tahap ini, peneliti menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF) sebagai teknik pembobotan untuk mengubah teks ulasan menjadi

representasi numerik yang dapat diproses oleh model klasifikasi. Pembobotan ini bertujuan untuk menilai tingkat kepentingan setiap kata dalam keseluruhan dokumen [12]. Proses pembobotan dilakukan terlebih dahulu pada data latih, sedangkan untuk data uji, pembobotan diterapkan secara otomatis saat tahap klasifikasi berlangsung. Perhitungan TF-IDF menggunakan persamaan (9)-(11).

$$TF(t, d) = \frac{f_{t,d}}{N_d} \quad (9)$$

$$IDF(t) = \log \log \left(\frac{N}{df_t} \right) \quad (10)$$

$$TF-IDF(t, d) = TF(t, d) \times IDF(t) \quad (11)$$

2.9 Evaluasi Confusion Matrix

Confusion Matrix digunakan untuk mengevaluasi performa model klasifikasi dengan membandingkan hasil prediksi terhadap label sebenarnya dalam bentuk tabel [13]. Dalam penelitian ini, digunakan *Confusion Matrix* 3×3 karena klasifikasi dilakukan terhadap tiga kelas sentimen, yaitu positif, netral, dan negatif. Metode ini memungkinkan analisis mendalam terhadap tingkat keberhasilan model dalam mengklasifikasikan setiap kelas (*True Positive*) maupun kesalahan klasifikasi (*False Positive* dan *False Negative*). Perhitungan matrik evaluasi mencakup *Accuracy*, *Precision*, dan *Recall*, masing-masing dihitung menggunakan persamaan (12)-(14).

$$Accuracy = \frac{Xx+Yy+Zz}{Xx+Xy+Xz+Yx+Yy+Yz+Zx+Zy+Zz} \times 100\% \quad (12)$$

$$Precision = \frac{1}{3} \left(\frac{Xx}{Xx+Yx+Zx} + \frac{Yy}{Xy+Yy+Zy} + \frac{Zz}{Xz+Yz+Zz} \right) \times 100\% \quad (13)$$

$$Recall = \frac{1}{3} \left(\frac{Xx}{Xx+Xy+Xz} + \frac{Yy}{Yx+Yy+Yz} + \frac{Zz}{Zx+Zy+Zz} \right) \times 100\% \quad (14)$$

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data Set

Pengumpulan dataset dilakukan untuk menganalisis sentimen ulasan pengguna terhadap kualitas layanan aplikasi Jamsostek *Mobile* (JMO) yang mencerminkan persepsi pengguna mengenai kemudahan penggunaan aplikasi, keandalan sistem, serta pengalaman layanan BPJS Ketenagakerjaan secara digital. Data berupa teks ulasan pengguna dikumpulkan dari *platform Google PlayStore* menggunakan library *google-play-scraper*. Sebanyak 10.000 ulasan pengguna aplikasi JMO berhasil dikumpulkan dan digunakan sebagai data penelitian. Ulasan tersebut selanjutnya diproses melalui tahapan preprocessing untuk membersihkan dan menyiapkan data agar layak digunakan dalam proses pelatihan dan pengujian model klasifikasi sentimen. Data ulasan hasil proses *scraping* disajikan pada Gambar 2.

Perbandingan Algoritma K-NN dan SVM dalam Analisis Sentimen Ulasan Pengguna Aplikasi JMO

Context
 Sebenarnya ini aplikasi buat apa sih.... Tiap log in selalu salah password padahal password nya ya itu2 aja ga jauh2 dr email & email pun udh dari jaman purbu Aplikasi sudah bagus cuma Sedikit saran dan masukan kekurangannya saat logging Aplikasi masih pakai Email seharusnya ada penambahan Fitur logging pakai sidi Tidak bisa melakukan update pengkinian data, muncul notifikasi tidak bisa pengkinian data karena ada proses klaim. Tidak bisa di buka problem nya dan tidak ti Kemarin udah bisa login, tapi pas diperbarui aplikasinya malah ga bisa masuk dan disuruh login lagi, password dan email sama ga ada yang berubah, dicoba berk Sudah Gonta ganti akun email akhirnya bisa. Ya tapi masak harus sribet itu? Imi aplikasi aku bari download yang tentunya sudah paling update dong, email juga Sudah coba berkali kali untuk ajukan klaim sampai biometrik 3x kemudian pergi ke kantor terdekat buat buka blokir...antre lama ...sudah dibuka sama stafnya. Mau cek saldo jht aja susah bgt, layar blank tiba2 lag layar hitam ga jalan sama sekali, bgiu juga pengkinian data, sering tiba2 apk terhenti..... mksdnya apa cob aplikasi menyulitkan pengguna inilah bukannya mempermudah, udah foto beberapa kali pas berhasil keluar aplikasi data tidak tersimpan, padahal langkah 2 te hayyyy... JMO aplikasi super keren buat pencairan saldo JHT, proses gampang saldo di bawah 10 juta hitungan jam saldo langsung di transfer ke rekening kalian Gimana rakyat mo percaya ma semua pelayanan pemerintah,bisanya cuma ambil untung dan nyusahin saja,kl ga bisa kerja memding keluar,ups kerjanyaan cu Aplikasi yg bagus dan praktis, semoga semakin berkembang dan semakin bertambah tingkat keamanan aplikasinya, supaya tidak mudah dibobol oleh cracker. T Setiap mau masuk selalu gagal katanya email anda tidak terdeteksi jelas jelas email saya tidak bermasalah aktif tapi aplikasi bodok ini yang bermasalah tolong: sementara saya kasih bintang 2 aplikasinya kurang bagus, gara-gara di update longuot sendiri dan tidak bisa login lagi, padahal login akun nya sama seperti akur Mohon maaf ini saya mau tanya ya, saya mau buka jmo tulisannya harus update tapi pas buka di Google play gak ada update cuma ada buka dan uninstal aplik Katanya JMO untuk mem permudah klaim,Nyatanya bukan gampang malah sulit, Waktu itu saya klaim di bulan Oktober, 2 minggu gak masuk duitnya, malah ha Foto biometrik nya ribet salah 3 kali harus ke kantor ga reset otomatis, sama pemeliharaan sistem jadi lama proses pencairan nya ... Kalo bisa app nya seperti j Maap bintang nya aku ubah dari 5 ke 1. Semua aduan ku gak ada yg nanggapi. luran ku bulan Oktober dan November gak masuk ke saldo. Tapi autodebet suks Data makin ambarudul, tidak sesuai dengan data yang pernah ada di aplikasi BPJSTKU... TOLONG PERBAIKAN SEGERA... NO. KPJ kenapa dihapus sepihak dan n Kenapa ini klaw mau masuk apk nya malah di suruh perbarui, pas sudah di perbarui malah tetap gak bisa masih di suruh perbarui padahal sudaah. Tolong dong Daftarnya gampang. Pas buka lgi minta di perbarui. Udh di perbarui malah kendala lgi. Ngerjain bgt ni aplikasi. mana ke kantor cabang mesti ngambil antrian p Wahhhh gila siii lama banget proses pencarian nya, padahal kan itu uang kita sendiri yaaa. Ini sudah hampir 3 minggu belum cair-juga semenjak pengajuan p Pertama instal dan buka aman2 aja sih, tapi kok tiba2 ada update perbaharui aplikasi, tapi pas di buka di PlayStore gk bisa update. Tolong di perbaiki lagi

Gambar 2. Data Ulasan Hasil Scrapping

3.2 Preprocessing Data

Berdasarkan data ulasan yang telah di *scraping* pada Gambar 2, kemudian dilakukan proses *preprocessing* data ulasan untuk mempersiapkan teks agar siap digunakan dalam tahap klasifikasi. Hasil dari proses *preprocessing* tersebut ditunjukkan pada Gambar 3 yang memperlihatkan data ulasan yang telah dibersihkan dan dinormalisasi.

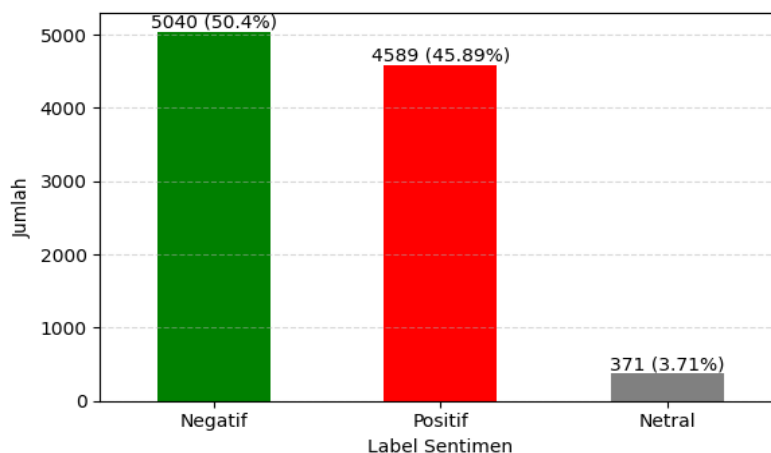
cleaning_text	case_folding	normalized_text	tokenized_text	stopword_removal	stemming_text	result_preprocessing
Sebenarnya ini aplikasi buat apaan sih Tiap lo...	sebenarnya ini aplikasi buat apaan sih tiap lo...	sebenarnya ini aplikasi buat apaan sih tiap lo...	['sebenarnya', 'ini', 'aplikasi', 'buat', 'apa...]	['aplikasi', 'sih', 'log', 'salah', 'password'...]	['aplikasi', 'sih', 'log', 'salah', 'password'...]	aplikasi sih log salah password password nya y...
Aplikasi sudah bagus cuma Sedikit saran dan ma...	aplikasi sudah bagus cuma sedikit saran dan ma...	aplikasi sudah bagus cuma sedikit saran dan ma...	['aplikasi', 'sudah', 'bagus', 'cuma', 'sediki...]	['aplikasi', 'bagus', 'saran', 'masukan', 'kek...]	['aplikasi', 'bagus', 'saran', 'masuk', 'kuran...]	aplikasi bagus saran masuk kurang loging aplik...
Tidak bisa melakukan update pengkinian data mu...	tidak bisa melakukan update pengkinian data mu...	tidak bisa melakukan update pengkinian data mu...	['tidak', 'bisa', 'melakukan', 'update', 'peng...]	['update', 'pengkinian', 'data', 'muncul', 'no...]	['update', 'kini', 'data', 'muncul', 'notifika...]	update kini data muncul notifikasi kini data a...
Kemarin udah bisa login tapi pas diperbarui ap...	kemarin udah bisa login tapi pas diperbarui ap...	kemarin sudah bisa login tapi pas diperbarui a...	['kemarin', 'sudah', 'bisa', 'login', 'tapi', ...]	['kemarin', 'login', 'pas', 'diperbarui', 'apl...]	['kemarin', 'login', 'pas', 'baru', 'aplikasi'...]	kemarin login pas baru aplikasi masuk suruh lo...
Sudah Gonta ganti akun email akhirnya bisa Ya ...	sudah gonta ganti akun email akhirnya bisa ya ...	sudah gonta ganti akun email akhirnya bisa ya ...	['sudah', 'gonta', 'ganti', 'akun', 'email', 'mas...]	['gonta', 'ganti', 'akun', 'email', 'ya', 'mas...]	['gonta', 'ganti', 'akun', 'email', 'ya', 'mas...]	gonta ganti akun email ya masak sribet imi apl...

Gambar 4. Proses Preprocessing Data

3.3 Labeling Data

Setelah proses *preprocessing* selesai dilakukan, tahap selanjutnya adalah pelabelan data (*labeling*) menggunakan kamus *Indonesian Sentiment (InSet) Lexicon*. Proses ini dilakukan secara otomatis dengan menghitung skor polaritas dari setiap ulasan. Berdasarkan nilai total polaritas, ulasan diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif.

Distribusi jumlah ulasan pada masing-masing kategori sentimen ditunjukkan pada Gambar 5 yang menampilkan frekuensi label sentimen hasil pelabelan.



Gambar 5. Frekuensi Label Sentimen

3.4 Pembobotan TF-IDF

Term Frequency–Inverse Document Frequency (TF-IDF) merupakan metode pembobotan yang digunakan untuk mengukur tingkat kepentingan suatu kata dalam kumpulan ulasan pengguna aplikasi JMO. Metode ini menghitung frekuensi kemunculan kata dalam satu dokumen (TF) dan membandingkannya dengan frekuensi kemunculan kata tersebut di seluruh dokumen (IDF). TF-IDF dinilai efisien, sederhana, dan mampu memberikan hasil pembobotan yang akurat dalam analisis teks.

Misalkan kata “aplikasi” muncul sebanyak 2 kali dalam satu dokumen yang terdiri dari 7 kata, maka:

$$TF(t, d) = \frac{f_{t,d}}{N_d} \rightarrow TF(\text{aplikasi}, d) = \frac{2}{7} \rightarrow 0,286$$

Dengan jumlah total dokumen $N=8$ dan kata “aplikasi” muncul pada $d_{ft}=7$ dokumen, maka:

$$IDF(t) = \log \log \left(\frac{N}{d_{ft}} \right) \rightarrow IDF(\text{aplikasi}) = \log \log \left(\frac{8}{7} \right) \rightarrow 0,058$$

$$TF-IDF(t, d) = TF(t, d) \times IDF(t)$$

$$TF-IDF(\text{aplikasi}, d) = 0,142 \times 0,057$$

$$TF-IDF(\text{aplikasi}, d) = 0,008$$

3.5 Pemodelan K-Nearest Neighbor dan Support Vector Machine

Pada tahap ini dilakukan pemodelan menggunakan algoritma *K-Nearest Neighbor* (K-NN) dan *Support Vector Machine* (SVM) untuk klasifikasi sentimen ulasan pengguna aplikasi JMO. Proses dimulai dengan data hasil pembobotan TF-IDF yang telah melalui tahap *preprocessing* dan pelabelan. Algoritma K-NN mengklasifikasikan data berdasarkan jarak

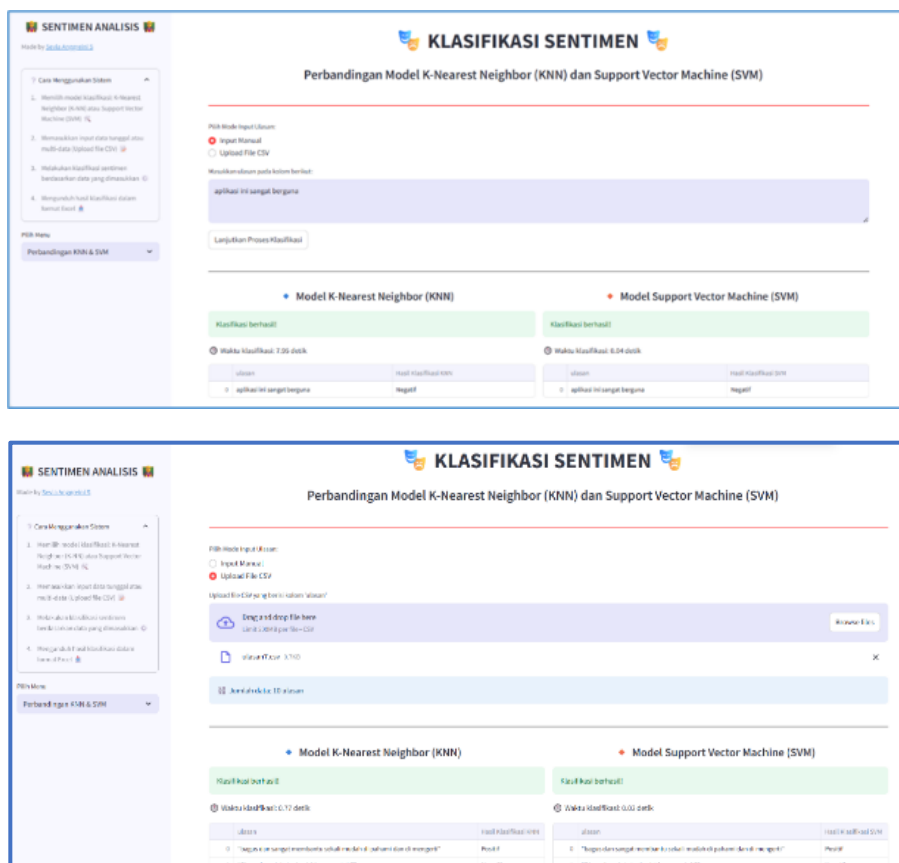
Perbandingan Algoritma K-NN dan SVM dalam Analisis Sentimen Ulasan Pengguna Aplikasi JMO

terdekat (*Euclidean Distance*) dengan menentukan kelas mayoritas dari K tetangga terdekat. Sementara itu, SVM bekerja dengan mencari *hyperplane* optimal yang memisahkan kelas sentimen positif, netral, dan negatif secara maksimal. Kedua model ini kemudian dibandingkan berdasarkan hasil akurasi, presisi, recall, F1-score, serta efisiensi waktu pemrosesan.

3.6 Pengujian Model

Setelah model *K-Nearest Neighbor* (K-NN) dan *Support Vector Machine* (SVM) dibangun dan dievaluasi, langkah selanjutnya adalah melakukan analisis sentimen terhadap data uji menggunakan model yang telah dilatih. Pada tahap ini, kedua algoritma akan memprediksi kelas sentimen dari setiap ulasan pengguna yang belum memiliki label, apakah termasuk positif, netral, atau negatif, berdasarkan pola yang telah dipelajari pada data latih. Hasil prediksi dari masing-masing model kemudian dibandingkan untuk mengetahui algoritma yang memberikan akurasi dan performa terbaik dalam mengklasifikasikan sentimen ulasan aplikasi JMO.

Hasil dari analisis sentimen menunjukkan bahwa algoritma K-NN dan SVM mampu mengklasifikasikan ulasan yang belum memiliki label sentimen pada data uji. Visualisasi hasil pelabelan sentimen dari masing-masing model ditampilkan melalui antarmuka sistem klasifikasi sentimen yang dikembangkan, sebagaimana ditunjukkan pada Gambar 6, yang memperlihatkan hasil prediksi label sentimen untuk setiap ulasan pengguna berdasarkan model K-NN dan SVM.

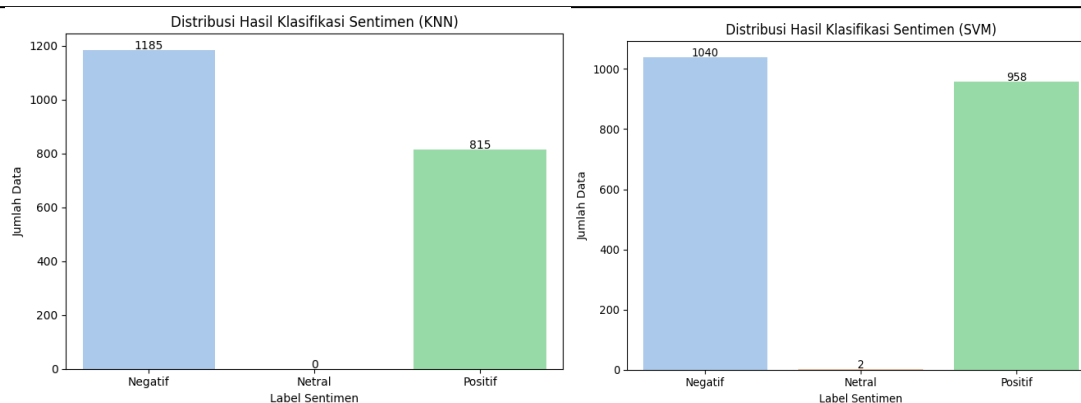


Gambar 6. Antarmuka Sistem Labeling Model K-NN dan SVM

3.7 Hasil Akurasi

Berdasarkan hasil pengujian, algoritma *Support Vector Machine* (SVM) menunjukkan performa yang lebih unggul dibandingkan *K-Nearest Neighbor* (K-NN) dalam klasifikasi sentimen ulasan pengguna aplikasi JMO. Model SVM mencapai akurasi sebesar 90,00%, melampaui akurasi tertinggi K-NN yang hanya 78,25%. Akurasi optimal pada SVM diperoleh

Perbandingan Algoritma K-NN dan SVM dalam Analisis Sentimen Ulasan Pengguna Aplikasi JMO



Gambar 8. Hasil Prediksi Data Uji K-NN dan SVM

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, hasil analisis sentimen terhadap ulasan pengguna aplikasi Jamsostek Mobile (JMO) menunjukkan bahwa algoritma *Support Vector Machine* (SVM) memiliki performa yang lebih baik dibandingkan *K-Nearest Neighbor* (K-NN). Model SVM mencapai tingkat akurasi sebesar 90,00%, sedangkan K-NN memperoleh akurasi sebesar 78,25%. Hasil klasifikasi menggunakan model SVM menunjukkan bahwa ulasan pengguna didominasi oleh sentimen positif dan negatif, dengan masing-masing sebanyak 958 ulasan positif dan 1.040 ulasan negatif, serta hanya 2 ulasan yang tergolong netral. Temuan ini mengindikasikan bahwa pengguna secara aktif menyampaikan pengalaman mereka dalam menggunakan aplikasi JMO, baik dalam bentuk apresiasi maupun masukan. Ulasan positif mencerminkan manfaat aplikasi dalam mendukung layanan digital BPJS Ketenagakerjaan, sementara ulasan negatif dapat dijadikan sebagai bahan evaluasi untuk pengembangan dan peningkatan kualitas aplikasi di masa mendatang. Dengan demikian, hasil penelitian ini diharapkan dapat memberikan kontribusi sebagai bahan pertimbangan dalam pengambilan keputusan terkait pengembangan layanan aplikasi JMO secara berkelanjutan.

5. SARAN

Saran untuk penelitian selanjutnya dapat difokuskan pada beberapa aspek berikut. Pertama, disarankan untuk meningkatkan jumlah dan keragaman data, khususnya pada kelas netral, agar model dapat belajar dengan lebih seimbang dan menghasilkan klasifikasi yang lebih akurat untuk ketiga kategori sentimen. Kedua, penelitian selanjutnya dapat mengeksplorasi metode klasifikasi lain seperti Random Forest, Naïve Bayes, atau LSTM, guna membandingkan performanya dengan K-NN dan SVM.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ibu Martaleli Bettiza dan Ibu Fortia Magfira yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] M. F. D. Samudera and V. I. Pertiwi, "Inovasi Pelayanan Publik Melalui Jamsostek Mobile (JMO) (Studi Kasus di BPJS Ketenagakerjaan Cabang Rungkut Kota Surabaya)," JPAP: Jurnal Penelitian Administrasi Publik, vol. 8, no. 1, pp. 152–172, 2022.

- [2] G. Buana, “BPJS Ketenagakerjaan Raih Penghargaan WOW Brand 2025, Aplikasi JMO Makin Diminati!,” *Media Indonesia*, Mar. 16, 2025. [Online]. Available: <https://mediaindonesia.com/humaniora/747563/bpjs-ketenagakerjaan-raih-penghargaan-wow-brand-2025-aplikasi-jmo-makin-diminati>.
- [3] S. A. R. Rizaldi, S. Alam, and I. Kurniawan, “Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes,” *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 2, no. 3, pp. 109–117, 2023.
- [4] M. Harahap, B. P. Sihombing, O. A. Laia, B. T. Saragih, and K. Dharma, “Analisis Sentimen Review Penjualan Produk UMKM pada Kabupaten Nias dengan Komparasi Algoritma Klasifikasi Machine Learning,” *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, vol. 5, no. 2, pp. 147–154, 2021.
- [5] Saiful, “3 Jenis ML: Supervised, Unsupervised, Reinforcement Learning,” *VPS Labs*, 2024. [Online]. Available: <https://vpslabs.net/supervised-unsupervised-reinforcement-learning/>. [Accessed: Feb. 13, 2025].
- [6] J. E. B. Sinulingga and H. C. K. Sitorus, “Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF,” *Jurnal Manajemen Informatika (JAMIKA)*, vol. 14, no. 1, pp. 42–53, 2024.
- [7] K. Mustaqim, F. A. Amaresti, and I. N. Dewi, “Analisis Sentimen Ulasan Aplikasi PosPay untuk Meningkatkan Kepuasan Pengguna dengan Metode K-Nearest Neighbor (KNN),” *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 1, pp. 11–20, 2024.
- [8] V. Vinne, D. U. Sinurat, and Y. Prasetianti, “Penerapan Algoritma K-Nearest Neighbors (KNN) dalam Menganalisis Sentimen Ulasan Aplikasi Seabank pada Google Play Store,” *Journal of Information Systems Management and Digital Business*, vol. 2, no. 2, pp. 103–113, 2025.
- [9] C. Z. V. Junus, T. Tarno, and P. Kartikasari, “Klasifikasi Menggunakan Metode Support Vector Machine dan Random Forest untuk Deteksi Awal Risiko Diabetes Melitus,” *Jurnal Gaussian*, vol. 11, no. 3, pp. 386–396, 2023.
- [10] V. K. S. Que, A. Iriani, and H. D. Purnomo, “Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization,” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 9, no. 2, 2020.
- [11] D. Musfiroh, U. Khaira, P. E. P. Utomo, and T. Suratno, “Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 24–33, 2021.
- [12] S. Syafrizal, M. Afdal, and R. Novita, “Analisis Sentimen Ulasan Aplikasi PLN Mobile Menggunakan Algoritma Naive Bayes Classifier dan K-Nearest Neighbor,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 10–19, 2024.

- [13] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier dan Confusion Matrix pada Analisis Sentimen Berbasis Teks pada Twitter," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 5, no. 2, pp. 697–711, 2021.
- [14] N. S. Marga, A. R. Isnain, and D. Alita, "Sentimen analisis tentang kebijakan pemerintah terhadap kasus corona menggunakan metode Naive Bayes," *Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA)*, vol. 2, no. 4, pp. 453–463, 2021.