

Analisis Sentimen Publik Terhadap Transportasi Umum di Jabodetabek Menggunakan Algoritma SVM Berbasis Web

Dian Gustina¹, Yudi Irawan Chandra^{*2}

¹Universitas Persada Indonesia UPI YAI

Jl. Pangeran Diponegoro No.74, Kenari, Senen, Jakarta Pusat 10430

²Sekolah Tinggi Manajemen Informatika dan Komputer Jakarta STI&K

Jl. BRI No.17 Radio Dalam, Kebayoran Baru, Jakarta Selatan 12140

¹dgus4006@gmail.com, ^{2*}yirawanc@gmail.com

*) Corresponding Author

Intisari— Pelayanan transportasi umum di wilayah Jabodetabek masih menjadi perhatian utama masyarakat, terutama dalam hal kenyamanan, ketepatan waktu, dan keamanan. Di era digital saat ini, media sosial dan platform digital lainnya menjadi sarana yang sering digunakan masyarakat untuk menyampaikan pendapat dan keluhan. Data opini yang tersebar di dunia maya ini dapat dimanfaatkan untuk mengetahui tingkat kepuasan publik terhadap layanan transportasi umum. Namun, banyaknya data yang tersedia membutuhkan metode analisis yang efektif dan akurat. Salah satu pendekatan yang dapat digunakan adalah algoritma Support Vector Machine (SVM), yang dikenal unggul dalam klasifikasi teks, termasuk analisis sentimen. Penelitian ini bertujuan untuk membangun sistem berbasis web yang mampu melakukan analisis sentimen secara otomatis terhadap opini publik mengenai layanan transportasi umum di Jabodetabek dengan memanfaatkan algoritma SVM. Proses penelitian mencakup pengumpulan data dari media sosial, pembersihan data, pelabelan sentimen, serta pelatihan dan pengujian model SVM. Sistem ini dikembangkan agar mudah diakses oleh pihak terkait untuk keperluan evaluasi dan pengambilan kebijakan. Hasil penelitian menunjukkan bahwa algoritma SVM mampu mengklasifikasikan sentimen positif, negatif, dan netral dengan tingkat akurasi yang cukup tinggi, mencapai 87% pada data uji. Sistem berbasis web yang dibangun juga terbukti responsif dan user-friendly. Temuan ini memberikan kontribusi positif dalam penyediaan alat bantu analisis opini publik secara real-time, yang dapat dimanfaatkan oleh penyedia layanan maupun pemerintah daerah dalam meningkatkan kualitas transportasi umum di Jabodetabek.

Kata kunci— Analisis Sentimen, Support Vector Machine, Transportasi Umum, Jabodetabek, Aplikasi Web.

Abstract— Public transportation services in the Greater Jakarta area remain a major concern for the public, particularly in terms of comfort, punctuality, and safety. In today's digital age, social media and other digital platforms have become the most commonly used means for the public to express their opinions and complaints. The opinions scattered across the digital landscape can be utilized to assess public satisfaction with public transportation services. However, the vast amount of available data requires effective and accurate analysis methods. One approach that can be employed is the Support Vector Machine (SVM) algorithm, which is known for its excellence in text classification, including sentiment analysis. This study aims to develop a web-based system capable of automatically analyzing public sentiment toward public transportation services in the Jabodetabek region using the SVM algorithm. The research process includes data collection from social media, data cleaning, sentiment labeling, as well as training and testing the SVM model. The system is designed to be easily accessible by relevant parties for evaluation and policy-making purposes. The results of the study show that the SVM algorithm is capable of classifying positive, negative, and neutral sentiments with a fairly high level of accuracy, reaching 87% on the test data. The web-based system that was built also proved to be responsive and user-friendly. These findings contribute positively to the provision of real-time public opinion analysis tools, which can be utilized by service providers and local governments to improve the quality of public transportation in Jabodetabek.

Keywords— Sentiment Analysis, Support Vector Machine, Public Transportation, Greater Jakarta Area, Web Application.

1. PENDAHULUAN

Transportasi umum merupakan salah satu elemen penting dalam mendukung mobilitas masyarakat, khususnya di kawasan perkotaan yang padat seperti Jakarta, Bogor, Depok, Tangerang, dan Bekasi (Jabodetabek) [1], [2]. Keberadaan moda transportasi seperti KRL, TransJakarta, MRT, LRT, serta angkutan umum lainnya menjadi andalan masyarakat dalam menjalani aktivitas harian. Namun, seiring meningkatnya jumlah pengguna dan dinamika kota besar, berbagai permasalahan mulai bermunculan. Keluhan masyarakat terkait ketepatan waktu, kebersihan, kenyamanan, keamanan, dan pelayanan dari

petugas kerap disuarakan, terutama melalui media sosial dan platform digital lainnya [3], [4].

Di era digital saat ini, media sosial telah menjadi ruang ekspresi publik yang sangat aktif. Masyarakat lebih terbuka dalam menyampaikan opini, baik berupa kritik maupun apresiasi, terhadap layanan publik, termasuk transportasi umum. Data-data opini ini apabila dikelola dengan baik dapat menjadi sumber informasi yang berharga bagi pengambil kebijakan dan pengelola layanan transportasi. Sayangnya, volume data yang besar serta bentuknya yang tidak terstruktur membuat proses analisis menjadi tidak sederhana. Oleh karena itu, diperlukan

sebuah sistem yang mampu menganalisis data opini publik secara otomatis, cepat, dan akurat [5].

Salah satu pendekatan yang efektif dalam menangani analisis opini dalam bentuk teks adalah dengan metode machine learning, khususnya algoritma Support Vector Machine (SVM). Algoritma ini dikenal memiliki performa yang baik dalam klasifikasi teks dan telah banyak digunakan dalam berbagai penelitian analisis sentimen. SVM bekerja dengan mencari hyperplane terbaik untuk memisahkan data ke dalam beberapa kelas, seperti positif, negatif, dan netral. Dengan kombinasi teknik pra-proses teks dan pelabelan yang tepat, SVM dapat menghasilkan model klasifikasi sentimen yang akurat [6], [7].

Namun, penelitian ini tidak hanya berfokus pada pembangunan model analisis sentimen, tetapi juga pada implementasi sistem berbasis web yang memungkinkan pengguna, baik dari pihak penyedia transportasi maupun masyarakat umum, untuk mengakses hasil analisis secara mudah dan interaktif. Sistem ini diharapkan dapat membantu dalam pengambilan keputusan berbasis data, evaluasi layanan, serta meningkatkan partisipasi publik dalam perbaikan transportasi umum.

Berdasarkan uraian tersebut, rumusan masalah dalam penelitian ini adalah bagaimana membangun sistem berbasis web yang dapat melakukan analisis sentimen publik terhadap transportasi umum di Jabodetabek dengan menggunakan algoritma SVM. Penelitian ini juga ingin menjawab sejauh mana tingkat akurasi model yang dibangun dalam mengklasifikasikan opini publik ke dalam kategori sentimen tertentu.

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut: Data yang digunakan berasal dari media sosial atau platform digital tertentu yang membahas transportasi umum di wilayah Jabodetabek; Jenis sentimen yang dianalisis dibatasi pada tiga kategori, yaitu positif, negatif, dan netral; Metode klasifikasi yang digunakan adalah algoritma Support Vector Machine; dan Sistem dikembangkan dalam bentuk aplikasi berbasis web untuk keperluan visualisasi hasil analisis.

Tujuan dari penelitian ini adalah untuk membangun sistem analisis sentimen publik terhadap pelayanan transportasi umum di Jabodetabek yang berbasis web dengan menggunakan algoritma SVM, serta mengevaluasi tingkat akurasi dari model klasifikasi yang dihasilkan.

II. LATAR BELAKANG

Analisis sentimen merupakan metode komputasi yang berfokus pada pendapat, sentimen, emosi, dan penilaian dari berbagai individu terhadap entitas seperti layanan, produk, organisasi, masalah, peristiwa, dan topik lainnya [8], [9]. Terdapat dua pendekatan dalam analisis sentimen, yaitu pelabelan otomatis yang menggunakan kamus kata-kata positif dan negatif untuk mencocokkan dengan kalimat yang akan dianalisis, dan pelabelan kalimat secara manual. Analisis sentimen bertujuan untuk menemukan informasi berharga dari data yang tidak terstruktur. Pendekatan ini melibatkan Penerapan

pemrosesan bahasa alami, evaluasi teks, studi linguistik komputasional, dan teknologi biometrik guna sistematisasi pengenalan, pengekstrakan, pengukuran, serta penelitian terhadap aspek afektif dan informasi subjektif.

Sentimen yang mengandung ekspresi positif, dukungan, apresiasi, pujian, atau respon yang baik terhadap suatu konten, video, tulisan, atau karya lainnya. Sentimen positif ditulis oleh individu yang memberikan respons yang mendukung dan menghargai isi dari apa yang mereka tonton, baca, atau dihadapi.

Sentimen positif dapat muncul di berbagai platform, termasuk YouTube, blog, media sosial, situs web, dan forum diskusi. Tujuan dari Sentimen positif adalah untuk memberikan umpan balik yang positif, mendorong dan memotivasi pencipta konten atau penulis, serta menciptakan lingkungan interaksi yang ramah dan mendukung.

Sentimen negatif adalah jenis komentar yang mengandung kritik, ketidakpuasan, atau ekspresi negatif terhadap suatu konten, video, tulisan, alat atau karya lainnya. Sentimen negatif ditulis oleh individu yang mengekspresikan ketidaksetujuan, ketidakpuasan, atau rasa tidak puas terhadap isi dari apa yang mereka tonton, baca, atau hadapi. Sentimen negatif dapat muncul di berbagai platform, seperti YouTube, blog, media sosial, situs web, dan forum diskusi. Tujuan dari Sentimen negatif dapat bervariasi, mulai dari memberikan kritik konstruktif untuk membantu meningkatkan kualitas Transportasi Umum, hingga menyampaikan rasa tidak puas atau ketidaksetujuan terhadap suatu yang ada di masyarakat. Penting untuk memahami bahwa tidak semua Sentimen negatif bersifat merendahkan atau bersifat negatif dengan tujuan yang buruk. Beberapa komentar negatif dapat membantu meningkatkan kualitas konten dan memberikan sudut pandang alternatif.

Klasifikasi komentar merupakan proses untuk mengelompokkan komentar dari pengguna atau penumpang ke dalam kategori tertentu, seperti positif atau negatif, berdasarkan isi dan ciri-ciri komentar tersebut. Tujuan utama dari proses ini adalah untuk memahami maksud atau sentimen yang terkandung dalam komentar yang diberikan. Dalam penelitian ini, proses klasifikasi bertujuan untuk membedakan antara komentar yang bernada positif dan komentar yang bersifat negatif [10], [11].

Secara umum, tahapan dalam klasifikasi komentar meliputi:

1. **Penyusunan Data Latih:** Langkah awal adalah menyiapkan kumpulan data komentar yang telah diberi label sebagai positif atau negatif. Data ini digunakan untuk melatih model agar dapat mengenali pola tertentu yang membedakan dua jenis sentimen tersebut.
2. **Ekstraksi Fitur:** Komentar kemudian dikonversi menjadi data numerik agar bisa diproses oleh model. Proses ini melibatkan pengambilan fitur penting seperti kata-kata yang sering muncul atau struktur kalimat yang umum pada masing-masing sentimen.
3. **Pelatihan Model:** Algoritma Support Vector Machine (SVM) diterapkan untuk membangun model klasifikasi berdasarkan fitur yang telah diekstraksi. SVM bekerja

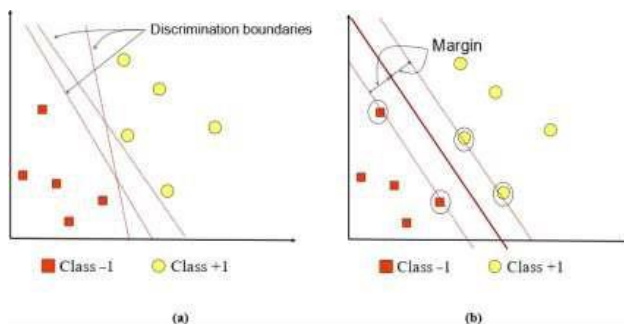
dengan mencari garis pemisah terbaik antara komentar positif dan negatif.

4. Evaluasi Kinerja: Setelah model selesai dilatih, pengujian dilakukan menggunakan data uji untuk menilai tingkat keakuratan model. Ukuran performa yang digunakan antara lain akurasi, presisi, recall, dan F1-score.
5. Prediksi Komentar Baru: Model yang sudah dilatih dapat digunakan untuk menganalisis komentar baru yang belum pernah dilihat sebelumnya dan menentukan kategorinya, apakah termasuk komentar positif atau negatif.

Proses klasifikasi ini memberikan manfaat penting dalam memahami opini masyarakat terhadap layanan transportasi umum. Informasi yang diperoleh dapat menjadi dasar pertimbangan bagi pihak terkait, seperti Dinas Perhubungan atau pengelola layanan, dalam meningkatkan mutu pelayanan serta memperbaiki komunikasi dan interaksi dengan masyarakat di wilayah Jabodetabek.

Pada tahun 1992, Vapnik bersama dengan rekannya Bernhard Boser dan Isabelle Guyon memperkenalkan algoritma Support Vector Machine (SVM) yang inovatif. SVM merupakan sebuah metode yang beroperasi dengan memanfaatkan konsep pemetaan nonlinier untuk mengubah ruang data pelatihan awal menjadi dimensi yang lebih tinggi [12]. Dalam dimensi yang baru ini, algoritma SVM mencari hyperplane yang mampu memisahkan data dari berbagai kelas secara linier. Melalui pemetaan nonlinier yang cermat ke dimensi yang lebih tinggi ini, data dari dua kelas yang berbeda dapat selalu terpisahkan dengan jelas oleh hyperplane yang dihasilkan. Proses ini dilakukan dengan memanfaatkan konsep support vector dan margin yang ada dalam SVM [13].

Metode SVM merupakan teknik dalam mencari hyperplane terbaik karena memiliki fungsi sebagai pemisah antara dua buah class yang berbeda pada sebuah input space dengan menggunakan suatu fungsi dot product [14].

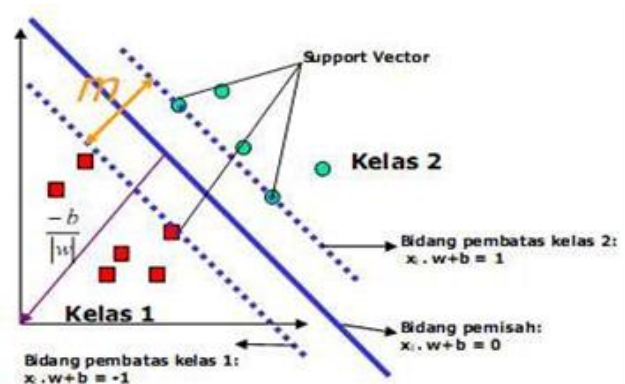


Gambar 1. Menentukan Hyperplane

Pada gambar 1 terlihat bahwa beberapa pola merupakan anggota dari dua kelas, yaitu +1 dan -1. Pola yang terikat pada kelas -1 dilambangkan dengan bentuk persegi berwarna merah, sedangkan pola yang terikat pada kelas +1 disimbolkan dengan lingkaran berwarna kuning. Menentukan garis hyperplane yang memisahkan dua buah

kelas merupakan salah satu teknik dalam permasalahan klasifikasi. Gambar di atas menunjukkan bahwa terdapat berbagai alternatif cara untuk menentukan garis pemisah dengan discrimination boundaries [15], [16].

Menentukan hyperplane pemisah terbaik antara dua kelas dengan mengukur nilai dari margin sebuah hyperplane dan mencari titik optimal hyperplane tersebut [17]. Margin adalah jarak antara garis hyperplane dan pola terdekat dari setiap kelas. Pola terdekat adalah support vector. Pada gambar di atas, garis solid yang terletak tepat di tengah kedua kelas tersebut menunjukkan hyperplane terbaik, adapun titik merah dan kuning pada lingkaran hitam menunjukkan support vector. Usaha dalam mencari lokasi hyperplane ini merupakan inti dari metode SVM. Adapun Gambar 2 menunjukkan proses dalam menentukan hyperplane pada linearly separable data.



Gambar 2. Menentukan Hyperplane pada linearly separable data

Jarak antara garis hyperplane dengan garis pemisah (H_2) dapat didapatkan menggunakan persamaan berikut :

$$d^- = \|x\| = \sqrt{x'x} = \frac{(-b-1)^2}{w'w} = \frac{-b-1}{\|w\|} \quad (1)$$

Nilai margin terbesar diperoleh dengan memaksimalkan jarak antara garis pemisah (H_1) serta (H_2) yang disebut supporting hyperplane, dengan menggunakan persamaan berikut :

$$|d^+ - d^-| = \left| \frac{(1-b)}{\|w\|} - \frac{(-b-1)}{\|w\|} \right| = \frac{2}{\|w\|} \quad (2)$$

Langkah terakhir dalam metode SVM adalah untuk menentukan class testing berdasarkan dari nilai fungsi keputusan :

$$f(x) = wx + b$$

$$f(x) = \sum_{i=1}^{Ns} \alpha_i x'_i y_i x + b \quad (3)$$

Dengan,

Ns = Jumlah support vector

x_i = support vector

Selanjutnya, menghitung nilai bias atau b dengan persamaan berikut :

$$b = \frac{1}{Ns} \sum_{i=1}^{Ns} \alpha_i x'_i y_i x \quad (4)$$

Persamaan di atas hanya dapat diaplikasikan pada kasus data yang dapat dipisahkan secara linier.

Teknik data mining atau machine learning banyak dikembangkan dengan asumsi linieritas. Sehingga algoritma yang dihasilkan lebih untuk kasus-kasus yang linier. Umumnya kasus-kasus yang sering terjadi bukanlah kasus yang linier. Untuk mengatasi sifat yang tidak linier tersebut dapat menggunakan metode kernel. Dengan metode kernel suatu data x input space di mapping ke feature space F dengan dimensi yang lebih tinggi. Adapun fungsi kernel yang biasanya digunakan dalam SVM yaitu:

Kernel linear : $x^T x$

Kernel polynomial: $(x^T x_i + 1)^p$

Kernel RBF: $\exp(-\frac{1}{2\sigma^2} \|x - x_i\|^2)$

Dataset adalah kumpulan data yang berisi informasi yang terkumpul dan disusun dalam format tertentu. Dataset merupakan bagian penting dalam pengolahan data dan analisis data karena berisi kumpulan nilai, atribut, atau pengamatan yang berkaitan dengan suatu topik atau fenomena tertentu. Dataset dapat berupa data numerik, teks, gambar, atau kombinasi dari berbagai jenis data lainnya, tergantung pada sifat data yang ingin dikumpulkan dan dianalisis. Dataset dapat dikumpulkan dari berbagai sumber, seperti survei, pengukuran, rekaman, atau data yang dikumpulkan dari berbagai platform online.

III. HASIL DAN PEMBAHASAN

Saat membangun sebuah sistem sentimen, diperlukan analisis atas perancangan sistem baru. Analisis dilakukan untuk mengetahui kebutuhan setiap individu pengguna sistem. memenuhi beberapa persyaratan yakni dapat diakses dengan mudah, memiliki interface yang memudahkan pengguna, serta dapat memproses dan menampilkan data sebagai informasi dengan baik.

Analisis Masalah

Setelah proses pengumpulan data dilakukan, penulis mengidentifikasi sejumlah permasalahan utama yang menjadi dasar dalam perancangan aplikasi untuk analisis dan klasifikasi komentar. Permasalahan pertama

adalah **ketidakseimbangan data latih** (imbalanced data), di mana jumlah komentar positif dan negatif tidak seimbang sehingga model cenderung memprediksi kelas mayoritas dan mengabaikan kelas minoritas. Untuk mengatasi hal ini, dapat diterapkan teknik seperti *undersampling*, *oversampling*, atau menggunakan algoritma SVM yang telah disesuaikan dengan data tidak seimbang. Permasalahan berikutnya adalah **kurangnya data latih yang representatif**, yang membuat model kesulitan dalam mengenali variasi bahasa dan konteks pada komentar, khususnya yang berasal dari kuisioner atau angket. Solusinya adalah dengan memperbanyak dan memperkaya data latih yang digunakan. Selain itu, **tahapan preprocessing yang tidak tepat** juga dapat memengaruhi akurasi model. Proses seperti pembersihan data, penghapusan stopwords, atau stemming harus dilakukan dengan mempertimbangkan konteks bahasa Indonesia agar tidak menghilangkan informasi penting dalam komentar.

Masalah lainnya yang juga perlu diperhatikan adalah **penentuan parameter yang tidak optimal** dalam algoritma SVM. Parameter seperti C dan jenis *kernel* harus disesuaikan melalui proses eksperimental agar tidak menimbulkan overfitting atau underfitting. Selain itu, komentar yang mengandung **bahasa kasar atau tidak senonoh** juga menjadi tantangan tersendiri karena dapat memengaruhi hasil klasifikasi. Oleh karena itu, dibutuhkan filter khusus atau pendekatan NLP lanjutan untuk mengatasi hal ini. Permasalahan lainnya adalah **evaluasi model yang kurang memadai**, di mana penggunaan metrik yang tidak tepat dapat menghasilkan penilaian performa yang menyesatkan. Oleh karena itu, penting untuk menggunakan ukuran evaluasi yang relevan seperti akurasi, presisi, recall, dan F1-score. Terakhir, **pemilihan fitur yang tidak optimal** juga dapat melemahkan kemampuan model dalam membedakan komentar berdasarkan sentimen. Fitur yang kurang relevan atau tidak informatif akan mengurangi akurasi prediksi, sehingga perlu dilakukan seleksi fitur yang cermat sesuai dengan karakteristik data komentar.

Analisis Pemecahan Masalah

Dalam proses pengembangan sistem analisis sentimen, terdapat beberapa langkah penting yang harus diperhatikan untuk meningkatkan akurasi dan kinerja model. Pertama, pengumpulan data training yang representatif sangat krusial. Data yang digunakan harus mencerminkan keragaman komentar yang mungkin muncul dalam kuisioner transportasi umum, baik dari sisi isi maupun gaya bahasa pengguna. Pengumpulan data dapat dilakukan melalui teknik *web scraping* dari media sosial atau melalui survei langsung kepada pengguna. Setelah data terkumpul, tahap berikutnya adalah *preprocessing* data, yang mencakup pembersihan teks seperti penghilangan tanda baca, konversi ke huruf kecil, penghapusan karakter khusus, tokenisasi, dan penghilangan stopwords dalam bahasa Indonesia agar fitur yang diekstraksi menjadi lebih bermakna. Selain itu, penting untuk menangani data yang tidak seimbang antara komentar positif dan negatif, misalnya dengan teknik *oversampling*, *undersampling*, atau

metode sintetis seperti *SMOTE* untuk meningkatkan representasi kelas minoritas secara proporsional.

Selanjutnya, proses ekstraksi fitur yang relevan memiliki peran besar dalam menentukan kualitas hasil klasifikasi. Fitur seperti kata kunci, panjang komentar, frekuensi kata, dan indikasi sentimen dapat membantu model membedakan sentimen positif dan negatif secara lebih akurat. Selain itu, pemilihan parameter yang optimal dalam algoritma SVM, seperti nilai *C* dan jenis *kernel*, perlu dilakukan secara hati-hati melalui metode *grid search* atau *randomized search* untuk mendapatkan hasil terbaik. Evaluasi model juga menjadi tahap yang tidak kalah penting. Penggunaan metrik evaluasi seperti akurasi, presisi, recall, F1-score, dan AUC-ROC akan memberikan gambaran yang lebih komprehensif mengenai performa model, terutama bila dilakukan pada data uji yang belum pernah digunakan saat pelatihan. Terakhir, setelah model diterapkan dalam lingkungan nyata, pemantauan performa secara berkala menjadi hal yang esensial. Hal ini dikarenakan distribusi data dapat berubah seiring waktu, yang dapat memengaruhi akurasi model. Dengan memperhatikan seluruh aspek tersebut, sistem analisis sentimen berbasis algoritma Support Vector Machine diharapkan dapat berfungsi lebih efektif dan menghasilkan klasifikasi yang lebih akurat terhadap opini publik mengenai layanan transportasi umum di Jabodetabek.

Analisis Kebutuhan

Analisis kebutuhan merupakan tahap awal yang krusial dalam pengembangan aplikasi analisis dan klasifikasi komentar terhadap transportasi umum dengan metode Support Vector Machine (SVM). Proses ini mencakup pemahaman terhadap tujuan utama dari sistem, kondisi lingkungan implementasi, serta ketersediaan sumber daya yang mendukung pengembangan aplikasi. Setelah kebutuhan diidentifikasi secara menyeluruh, langkah selanjutnya adalah merancang spesifikasi teknis yang dibutuhkan agar sistem yang dikembangkan dapat berjalan secara optimal dan sesuai dengan sasaran yang telah ditentukan. Melalui analisis kebutuhan, sistem atau proyek yang dibangun diharapkan dapat memberikan solusi yang tepat, fungsional, dan mudah digunakan oleh target pengguna.

Hasil dari analisis kebutuhan ini memberikan penulis pemahaman yang lebih jelas mengenai elemen-elemen yang harus dipenuhi dalam merancang aplikasi web yang akan dikembangkan. Dalam proses ini, kebutuhan dibagi menjadi dua kategori utama, yaitu kebutuhan fungsional dan kebutuhan non-fungsional. Kebutuhan fungsional mencakup fitur-fitur utama yang akan tersedia dalam aplikasi, seperti klasifikasi komentar, visualisasi hasil, dan pengelolaan data. Sementara itu, kebutuhan non-fungsional berkaitan dengan aspek teknis pendukung seperti perangkat keras, perangkat lunak, keamanan sistem, serta performa aplikasi. Dengan mengidentifikasi kedua aspek ini secara rinci, pengembangan aplikasi dapat lebih terarah dan sesuai dengan ekspektasi pengguna maupun tujuan sistem.

Analisis Kebutuhan Fungsional

Kebutuhan fungsional merupakan aspek penting yang harus dipenuhi oleh suatu sistem, produk, atau proyek agar mampu menjalankan tugas sesuai dengan tujuan yang telah ditetapkan. Kebutuhan ini umumnya diidentifikasi melalui proses analisis kebutuhan dan ditentukan oleh pengguna akhir atau admin yang akan berinteraksi langsung dengan sistem. Fungsionalitas yang dimaksud merujuk pada fitur-fitur utama yang harus tersedia agar sistem dapat beroperasi sebagaimana mestinya. Sebagai contoh, kebutuhan fungsional pada perangkat komputer mencakup kemampuan untuk menyimpan data, menjalankan aplikasi, dan terhubung ke jaringan internet. Dengan kata lain, kebutuhan fungsional memastikan bahwa sistem atau aplikasi memiliki fungsi inti yang benar-benar dibutuhkan oleh pengguna.

Dalam konteks pengembangan aplikasi analisis sentimen publik terhadap transportasi umum di Jabodetabek menggunakan algoritma Support Vector Machine (SVM), kebutuhan fungsional merujuk pada fitur-fitur yang harus tersedia dalam aplikasi agar dapat menjalankan proses klasifikasi dan visualisasi data secara optimal. Kebutuhan ini mencakup berbagai komponen antarmuka dan fungsi sistem, seperti tampilan awal (guest), halaman login, tampilan data latih, data pengujian, hasil pengujian, grafik visualisasi, dan fitur logout. Setiap elemen tersebut dirancang untuk mendukung pengalaman pengguna serta kelancaran proses analisis sentimen. Selain itu, kebutuhan fungsional bersifat dinamis dan dapat dikembangkan lebih lanjut seiring perkembangan teknologi serta meningkatnya kebutuhan pengguna. Oleh karena itu, pemahaman yang tepat terhadap kebutuhan fungsional akan membantu memastikan bahwa aplikasi tidak hanya berfungsi secara teknis, tetapi juga relevan dan berguna bagi penggunaannya.

Analisis Kebutuhan Non Fungsional

Kebutuhan non-fungsional adalah jenis kebutuhan yang tidak secara langsung berkaitan dengan fungsi utama dari suatu sistem atau aplikasi, namun memiliki peran penting dalam menentukan kualitas, kinerja, dan pengalaman pengguna. Berbeda dengan kebutuhan fungsional yang berfokus pada apa yang harus dilakukan sistem, kebutuhan non-fungsional menjelaskan bagaimana sistem tersebut bekerja dalam kondisi nyata. Kebutuhan ini biasanya diidentifikasi melalui proses analisis kebutuhan dan ditentukan oleh pengguna atau pemangku kepentingan yang memahami standar kualitas yang diinginkan. Contoh kebutuhan non-fungsional dalam aplikasi berbasis web meliputi kecepatan akses (loading), tingkat keamanan data pengguna, kompatibilitas dengan berbagai perangkat, efisiensi ukuran aplikasi, serta kepatuhan terhadap standar aksesibilitas.

Pemenuhan kebutuhan non-fungsional sangat penting agar aplikasi dapat beroperasi secara optimal dan memberikan pengalaman yang baik kepada pengguna. Kebutuhan ini juga bersifat dinamis dan dapat dikembangkan seiring dengan kemajuan teknologi serta perubahan kebutuhan pengguna. Selain aspek performa dan

keamanan, kebutuhan non-fungsional juga mencakup spesifikasi teknis yang berkaitan dengan perangkat keras dan perangkat lunak yang digunakan dalam proses pengembangan aplikasi. Dalam hal ini, pengembang harus mempertimbangkan komponen-komponen pendukung yang diperlukan untuk menunjang keberhasilan implementasi sistem, seperti spesifikasi minimum server, jenis browser yang kompatibel, sistem operasi yang didukung, serta framework dan bahasa pemrograman yang digunakan. Dengan memperhatikan kebutuhan non-fungsional secara menyeluruh, aplikasi yang dibangun tidak hanya berfungsi dengan baik, tetapi juga dapat diandalkan, aman, dan mudah diakses oleh berbagai kalangan pengguna.

1. Perangkat Keras (Hardware)

Perangkat Keras yang digunakan penulis dalam pembuatan aplikasi Analisis Sentimen Publik Terhadap Transportasi dengan metode Support Vector Machine yaitu : Personal Computer (PC) dengan spesifikasi Laptop ASUS AMD Ryzen 3, Ram 8 GB dan Harddisk SSD 512 GB

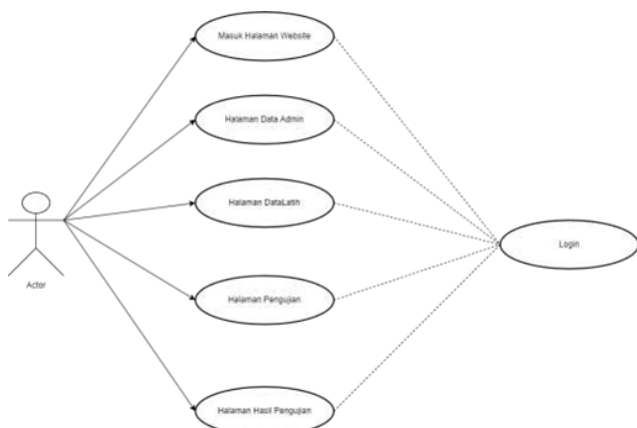
2. Perangkat Lunak (Software)

Perangkat lunak yang dibutuhkan sebagai berikut:

- Sistem Operasi Windows 11 64-bit
- Visual Studio Code
- MySQL
- PhpMyAdmin
- LARAVEL
- Node.js

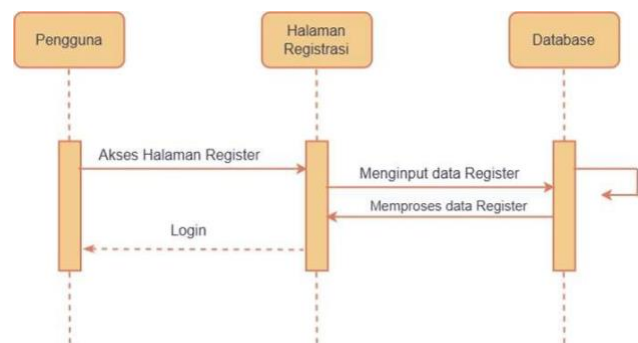
Rancangan Basis Data dengan UML

Gambar 3 menyajikan rancangan *use case* dari aplikasi berbasis web untuk klasifikasi analisis sentimen. Berdasarkan ilustrasi tersebut, dapat disimpulkan bahwa terdapat dua aktor utama yang terlibat dalam proses klasifikasi sentimen terhadap layanan transportasi umum. Kedua aktor tersebut memiliki peran dan interaksi yang berbeda terhadap sistem, yang menggambarkan bagaimana aplikasi ini digunakan untuk menjalankan fungsi-fungsi utama dalam analisis sentimen.



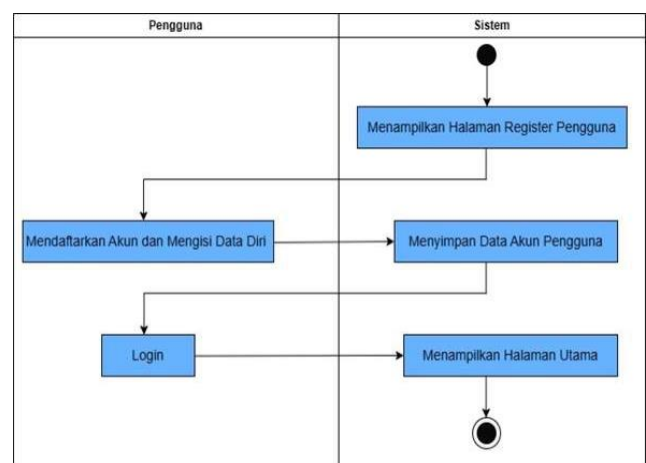
Gambar 3. Use Case Diagram

Sequence diagram atau diagram urutan merupakan representasi visual yang digunakan untuk menggambarkan urutan langkah-langkah atau peristiwa dalam suatu proses secara kronologis. Diagram ini berfungsi untuk menunjukkan alur komunikasi dan interaksi antar objek atau entitas dalam sistem, sehingga memudahkan pemahaman terhadap mekanisme kerja sistem secara dinamis. Tujuan utama dari sequence diagram adalah untuk menjelaskan bagaimana objek saling berinteraksi dalam situasi atau skenario tertentu yang telah didefinisikan dalam *use case*. Salah satu contohnya adalah sequence diagram pada proses registrasi, yang ditampilkan pada Gambar 4.



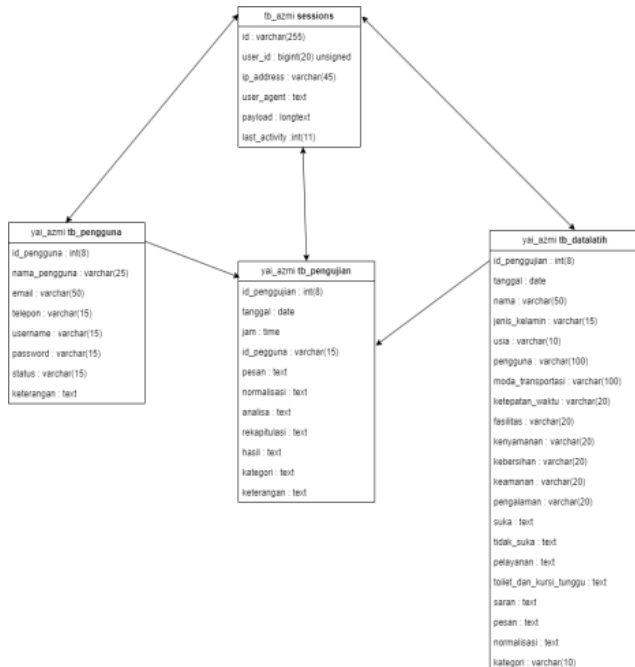
Gambar 4. Sequence Diagram Registrasi

Activity Diagram merupakan representasi visual yang menggambarkan urutan alur kerja, mencakup serangkaian aktivitas, tindakan, serta kemungkinan adanya percabangan atau pengulangan dalam proses. Dalam *Unified Modeling Language* (UML), diagram ini digunakan untuk memodelkan proses sistem, baik berupa aktivitas komputasi maupun alur kerja dalam organisasi. Selain menjelaskan tahapan proses secara terstruktur, activity diagram juga memberikan gambaran umum mengenai alur kontrol yang terjadi dalam sistem. Sebagai contoh, *Activity Diagram* untuk proses pengguna baru pada aplikasi website ditampilkan pada Gambar 5, yang memperlihatkan langkah-langkah interaksi dari awal hingga akhir dalam proses pendaftaran atau registrasi pengguna.



Gambar 5. Activity Diagram Registrasi

Berikut ini disajikan Class Diagram dari aplikasi website, sebagaimana ditampilkan pada Gambar 6. Diagram ini menggambarkan struktur kelas, atribut, metode, serta hubungan antar kelas yang membentuk dasar arsitektur sistem secara menyeluruh.



Gambar 6. Class Diagram

Rancangan Desain

Rancangan ini merupakan tahap awal dalam pengembangan aplikasi berbasis web, yang menampilkan antarmuka pengguna untuk proses login. Pada tampilan tersebut, pengguna diberikan opsi untuk masuk ke dalam sistem dengan memasukkan *username* dan *password* yang telah terdaftar sebelumnya. Seperti ditunjukkan pada Gambar 7, terdapat tombol "Masuk" yang apabila diklik akan mengarahkan pengguna menuju halaman utama (*home*) aplikasi.



Gambar 7. Tampilan Login Aplikasi

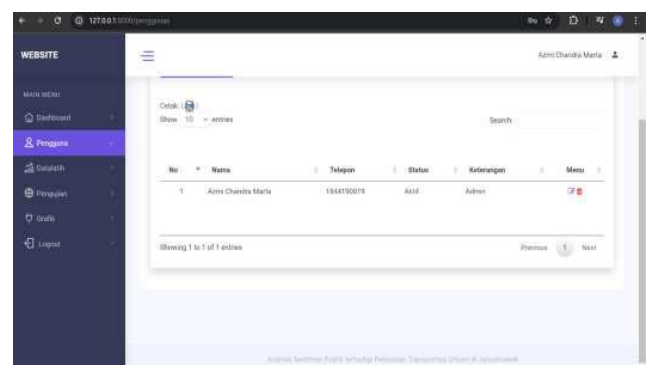
Gambar 8 adalah tampilan menu *Register* yang menyediakan sejumlah kolom input untuk mengisi informasi data pengguna. Pengguna diwajibkan mengisi data seperti *username*, *password*, *email*, nomor telepon, serta jenis kelamin sebagai identitas dasar yang akan disimpan dalam sistem dan digunakan sebagai referensi

dalam penggunaan aplikasi. Setelah seluruh data diisi dengan lengkap, pengguna dapat menekan tombol "Simpan" untuk melanjutkan, yang kemudian akan secara otomatis diarahkan ke halaman menu *Login*.



Gambar 8 Tampilan Register

Gambar 9 adalah tampilan Menu Admin/Pengguna yang memuat daftar data pengguna. Apabila seorang pengguna ingin mengubah informasi seperti nama atau nomor telepon, maka pengguna dapat mengklik bagian *Pengguna* untuk melakukan pembaruan data tersebut. Selain itu, pada menu ini juga tersedia opsi untuk menghapus data pengguna yang sudah tidak diperlukan.



Gambar 9 Tampilan Dashboard Admin

Pengumpulan Data

Proses pengumpulan data dilakukan melalui penyebaran kuesioner menggunakan platform Google Form. Metode ini dipilih karena mampu menjangkau berbagai responden secara luas dan efisien. Data yang diperoleh berasal langsung dari hasil isian responden pada Google Form, kemudian diekspor dalam format Excel untuk keperluan analisis lebih lanjut. Melalui proses *scraping*, diperoleh sebanyak 88 entri data yang siap digunakan dalam tahap pengolahan dan pelatihan model. Gambar 10 memperlihatkan contoh dataset yang diambil dari hasil kuesioner tersebut melalui Google Form.

Gambar 10 Dataset melalui Google Form

Dalam proses perancangan sistem, tahap awal yang dilakukan adalah pengumpulan dataset yang akan digunakan sebagai dasar dalam membangun model analisis sentimen. Dataset diperoleh secara independen dengan mengambil sampel data dari hasil kuesioner yang telah disebarluaskan sebelumnya. Setelah data terkumpul, langkah berikutnya adalah proses *labeling*, yaitu pemberian penanda atau klasifikasi sentimen (seperti positif atau negatif) pada setiap entri data. Tahapan ini penting agar data dapat dikenali dan diproses secara optimal oleh model yang akan dikembangkan.

Filtering

Filtering merupakan bagian dari tahap pra-proses data dalam penggunaan algoritma Support Vector Machine (SVM) untuk mengelompokkan komentar berdasarkan sentimen, seperti positif atau negatif. SVM adalah salah satu metode *machine learning* yang umum digunakan untuk klasifikasi. Sebelum data dianalisis, perlu dilakukan beberapa tahap pra-proses, seperti *case folding* (mengubah huruf menjadi kecil), *tokenizing* (memecah teks menjadi kata-kata), dan *filtering* (menghapus kata-kata yang tidak penting atau *stop words*).

Setelah proses filtering selesai, langkah berikutnya adalah membentuk fitur dari data. Komentar yang telah diproses diubah ke dalam bentuk angka agar bisa dikenali oleh SVM. Salah satu metode yang digunakan adalah TF-IDF (Term Frequency-Inverse Document Frequency), yang memberikan nilai bobot pada setiap kata berdasarkan frekuensinya. Setelah itu, data diberi label sesuai dengan jenis sentimen, seperti positif atau negatif. Pelabelan ini bisa dilakukan secara manual oleh admin atau secara otomatis menggunakan kamus sentimen. Proses ini penting agar model SVM dapat mengenali pola dan melakukan klasifikasi dengan hasil yang baik.

Case Folding

Case Folding merupakan tahapan mengganti seluruh case dalam sebuah dokumen menjadi bentuk standar (huruf kecil). Sedangkan karakter lainnya dianggap sebagai delimiter atau pembatas. Hasil data setelah melalui proses case folding.

Stopword

Stopword adalah kata-kata umum dalam suatu bahasa yang sering muncul tetapi tidak memiliki makna penting dalam analisis teks. Dalam Bahasa Indonesia, contoh stopwords antara lain "dan", "yang", "di", "ke", dan "dari". Pada proses analisis teks, biasanya dilakukan penghapusan stopwords sebelum atau sesudah tahap *stemming*—yaitu proses mengubah kata ke bentuk dasarnya, misalnya "berlari" menjadi "lari". Penghapusan stopwords bertujuan untuk menghilangkan kata-kata yang tidak memberikan informasi penting, karena kata-kata tersebut tidak berkontribusi banyak terhadap pemahaman isi atau makna teks.

Stemming

Stemming merupakan tahapan setiap kata akan diubah dari kata berimbuhan menjadi kata dasar. Stemming yang digunakan adalah algoritma Nazief dan Adriani yang terdapat pada library sastrawi.

Unifikasi

Unifikasi adalah menggambarkan bagaimana proses pemrosesan data, representasi fitur, dan penggunaan model SVM menyatukan data komentar menjadi sebuah model klasifikasi yang dapat mengidentifikasi komentar sebagai positif atau negatif dengan akurasi yang tinggi. Dengan melakukan unifikasi ini, aplikasi analisis klasifikasi komentar dapat memberikan informasi berharga tentang sentimen dan tanggapan pengguna terhadap Transportasi Umum di Jabodetabek.

Hasil pengujian terhadap aplikasi analisis sentimen publik pada moda transportasi umum di wilayah Jabodetabek, yang menggunakan metode Support Vector Machine (SVM), menunjukkan performa yang sangat baik dalam mengklasifikasikan komentar menjadi sentimen positif dan negatif. Algoritma SVM mampu menghasilkan tingkat akurasi mencapai 100% dalam proses klasifikasi yang dilakukan, sebagaimana ditampilkan pada Gambar 11. Meskipun hasil grafik menunjukkan pembagian yang tegas antara komentar positif dan negatif, penting untuk tetap mempertimbangkan ukuran dan keseimbangan dataset dalam interpretasi akurasi tersebut. Secara keseluruhan, temuan ini mengindikasikan bahwa SVM merupakan metode yang efektif dan andal dalam menganalisis sentimen publik terhadap layanan transportasi umum.

Gambar 11 Hasil dari Pengujian data analisis

Berdasarkan hasil dan temuan penelitian, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan lebih lanjut. Pertama, perlu dilakukan penambahan volume dan variasi sumber data karena keragaman data sangat berpengaruh terhadap performa model; oleh karena itu, disarankan untuk memperluas cakupan data dari berbagai platform media sosial dan forum diskusi guna meningkatkan generalisasi sistem. Kedua, klasifikasi sentimen dapat dikembangkan lebih lanjut ke dalam kategori yang lebih beragam, tidak hanya terbatas pada sentimen positif dan negatif, tetapi juga mencakup kategori netral atau kategori spesifik seperti pujian terhadap fasilitas atau kritik terhadap ketepatan waktu. Ketiga, meskipun algoritma Support Vector Machine (SVM) telah memberikan hasil yang cukup baik, eksplorasi terhadap metode pembelajaran mesin lainnya seperti Random Forest, Naive Bayes, atau model berbasis deep learning seperti LSTM dan BERT dapat meningkatkan akurasi sistem. Keempat, hasil analisis sentimen sebaiknya diintegrasikan dengan sistem pendukung keputusan agar penyedia layanan dapat memberikan respon yang lebih cepat dan tepat terhadap isu yang berkembang. Dengan peningkatan dan pengembangan lanjutan, sistem analisis sentimen ini berpotensi menjadi alat strategis dalam mendukung evaluasi dan perbaikan berkelanjutan layanan transportasi umum, khususnya di wilayah Jabodetabek.

- [1] N. Brotodewo, "Penilaian Indikator Transportasi Berkelanjutan Pada Kawasan Metropolitan Di Indonesia," *J. Perenc. Wil. Dan Kota*, vol. 21, no. 3, 2010.
- [2] Y. W. R. Putra, F. N. Styaningsih, and W. H. Herviana, "Analisis Perkembangan Transportasi Online di Indonesia di Era 4.0 Dengan Metode Penelitian Deskriptif," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 4, no. 1, Art. no. 1, Jan. 2022, doi: 10.47233/jteksis.v4i1.389.
- [3] E. N. Julianto, "Hubungan Antara Kecepatan, Volume Dan Kepadatan Lalu Lintas Ruas Jalan Siliwangi Semarang," *J. Tek. Sipil Dan Perenc.*, vol. 12, no. 2, Art. no. 2, 2010, doi: 10.15294/jtsp.v12i2.1348.
- [4] B. Saputra and D. Savitri, "Analisis Hubungan antara Volume, Kecepatan dan Kepadatan Lalu-Lintas Berdasarkan Model Greenshield, Greenberg dan Underwood," *J. Manaj. Aset Infrastruktur Fasilitas*.

- vol. 5, no. 1, Art. no. 1, Jan. 2021, doi: 10.12962/j26151847.v5i1.8742.
- [5] I. K. Sukesha and E. Papyrakis, "Hubungan Antara Pertumbuhan Ekonomi dan Infrastruktur Transportasi di Indonesia," *J. Ekon. Dan Pembang. Indones.*, vol. 23, no. 2, Jan. 2023, doi: 10.21002/jepi.2023.10.
- [6] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, Feb. 2021, doi: 10.34148/teknika.v10i1.311.
- [7] D. S. Utami and A. Erfina, "Analisis Sentimen Pinjaman Online Di Twitter Menggunakan Algoritma Support Vector Machine (Svm)," *Pros. Semin. Nas. Sist. Inf. Dan Manaj. Inform. Univ. Nusa Putra*, vol. 1, pp. 299–305, Sep. 2021.
- [8] F. Fitroh and F. Hudaya, "Systematic Literature Review: Analisis Sentimen Berbasis Deep Learning," *J. Nas. Teknol. Dan Sist. Inf.*, vol. 9, no. 2, Art. no. 2, Aug. 2023, doi: 10.25077/TEKNOSI.v9i2.2023.132-140.
- [9] T. M. P. Aulia, N. Arifin, and R. Mayasari, "Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinisasi Covid-19," *SINTECH Sci. Inf. Technol. J.*, vol. 4, no. 2, Art. no. 2, Oct. 2021, doi: 10.31598/sintechjournal.v4i2.762.
- [10] T. A. Saputra and M. Devega, "Analisis Sentimen Cyberbullying Pada Komentar Instagram Menggunakan Metode Klasifikasi Support Vector Machine (SVM) | ZONasi: Jurnal Sistem Informasi," Dec. 2024, Accessed: Jul. 16, 2025. [Online]. Available: <https://journal.unilak.ac.id/index.php/zn/article/view/23122>
- [11] A. N. Husna, "Klasifikasi Komentar Bullying Pada Kolom Komentar Streamer Youtube," undergraduate, Universitas Muhammadiyah Malang, 2023. Accessed: Jul. 16, 2025. [Online]. Available: <http://www.jurnal.iaii.or.id/index.php/JOSEIT/article/view/5016>
- [12] W. S. Dharmawan, "Komparasi Algoritma Klasifikasi Svm-Pso Dan C4.5-Pso Dalam Prediksi Penyakit Jantung," *N F O R M T K A*, vol. 13, no. 2, Art. no. 2, Jan. 2022, doi: 10.36723/juri.v13i2.301.
- [13] A. Muhammadin and I. A. Sobari, "Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma Svm Dan Nbc," *Reputasi J. Rekayasa Perangkat Lunak*, vol. 2, no. 2, Art. no. 2, Dec. 2021, doi: 10.31294/reputasi.v2i2.785.
- [14] C. M. Fiduccia, E. R. Scheinerman, A. Trenk, and J. S. Zito, "Dot product representations of graphs," *Discrete Math.*, vol. 181, no. 1, pp. 113–138, Feb. 1998, doi: 10.1016/S0012-365X(97)00049-6.
- [15] E. R. Scheinerman and K. Tucker, "Modeling graphs using dot product representations," *Comput. Stat.*, vol. 25, no. 1, pp. 1–16, Mar. 2010, doi: 10.1007/s00180-009-0158-8.
- [16] A. Athreya *et al.*, "Statistical Inference on Random Dot Product Graphs: a Survey," *J. Mach. Learn. Res.*, vol. 18, no. 226, pp. 1–92, 2018.
- [17] Y. Arora and S. K. Gupta, "Intuitionistic fuzzy twin proximal SVM with fuzzy hyperplane and its application in EEG signal classification," *Appl. Soft Comput.*, vol. 163, p. 111816, Sep. 2024, doi: 10.1016/j.asoc.2024.111816.