



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 3 (2025) pp: 4051-4062

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Sentimen Aplikasi Deepseek Menggunakan Metode Naive Bayes dan Support Vector Machine

Muhammad Faqih Baihaqi¹, Lena Magdalena², Rifki Fahrudin³

^{1,2,3}Prodi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Catur Insan Cendikia

¹muhammadfaqihb15@gmail.com, ²lena.magdalena@cic.ac.id, ³rifki.fahrudin@cic.ac.id

Abstrak

Penelitian ini membandingkan metode Naive Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen ulasan aplikasi DeepSeek Assistant AI, yang berhasil menembus 10 besar aplikasi gratis di Google Play Store di 18 negara dalam 16 hari. Analisis dilakukan pada 2400 ulasan dari Januari hingga Juli 2025 dengan proporsi 68% ulasan positif dan 32% negatif. Evaluasi model menggunakan confusion matrix menunjukkan bahwa SVM unggul dalam akurasi (90,1% vs 75,7%), F1-Score (92,7% vs 84,6%), dan presisi (93% vs 74%), sementara Naive Bayes unggul pada recall (99% vs 93%). Perbandingan juga mencakup kinerja model, Area Under Curve, rasio dataset, dan kecepatan klasifikasi. Hasil penelitian menunjukkan bahwa SVM lebih unggul dalam tiga dari empat aspek yang dianalisis, sementara Naive Bayes hanya unggul pada kecepatan klasifikasi. Penelitian ini juga menyediakan visualisasi data dalam sebuah website yang menampilkan hasil analisis sentimen dan perbandingan performa kedua metode. Temuan ini memberikan wawasan penting dalam pemilihan metode yang paling efektif untuk analisis sentimen aplikasi berbasis AI, serta dapat digunakan sebagai referensi bagi pengembang aplikasi yang ingin mengoptimalkan pemahaman terhadap ulasan pengguna.

Kata kunci: Analisis Sentimen, Aplikasi, Deepseek, Naive Bayes, Support Vector Machine.

1. Latar Belakang

Penggunaan kecerdasan buatan (AI) dalam kehidupan sehari-hari semakin meluas, salah satunya adalah chatbot, yang dirancang untuk meniru percakapan manusia melalui pemahaman input pengguna dan memberikan respons otomatis. Berdasarkan Chatbot Statistics 2024, 74% pengguna internet lebih memilih berinteraksi dengan chatbot untuk mencari jawaban [1]. Hal ini menunjukkan bahwa chatbot kini menjadi bagian integral dalam pengalaman digital, memfasilitasi interaksi dari layanan pelanggan hingga pencarian informasi. Seiring perkembangan teknologi chatbot, muncul platform AI generatif baru, seperti Meta AI, ChatGPT, dan Gemini, yang memiliki jutaan pengguna aktif [2]. Teknologi ini tidak hanya meningkatkan interaksi manusia-mesin, tetapi juga membuka peluang baru di berbagai industri, termasuk teknologi, e-commerce, dan pendidikan.

Salah satu inovasi terbaru adalah DeepSeek, chatbot generatif yang diluncurkan oleh startup Tiongkok, yang menggabungkan teknologi generatif dengan pemahaman percakapan manusia canggih melalui algoritma Natural Language Processing (NLP) [3]. DeepSeek menarik perhatian karena dampaknya terhadap industri teknologi, diperkirakan menyebabkan kerugian besar bagi Nvidia, yang berperan dalam pengembangan chip untuk AI generatif [4]. Selain itu, DeepSeek juga meluncurkan model bahasa besar, DeepSeek R1, yang diklaim setara dengan GPT-4 dengan biaya pelatihan yang lebih efisien [5]. Meski begitu, aplikasi ini menghadapi tantangan terkait keamanan data pengguna, menyebabkan beberapa negara seperti AS dan Australia memblokir aplikasi ini [6]. DeepSeek berhasil meraih popularitas, menempati posisi 10 aplikasi gratis teratas di Google Play Store pada 27 Januari 2025.

Namun, ulasan pengguna di Play Store terbatas pada rating bintang, tanpa visualisasi data lebih informatif. Penelitian ini bertujuan menganalisis sentimen ulasan aplikasi DeepSeek menggunakan dua metode machine learning, Naive Bayes dan SVM, untuk mengklasifikasikan ulasan sebagai positif atau negatif dan memberikan gambaran lebih luas. Data yang diperoleh akan diproses menggunakan Google Colab dan divisualisasikan dalam grafik yang dapat diakses melalui website. Penelitian ini juga membandingkan kinerja kedua algoritma berdasarkan empat aspek: Kinerja Model, Area Under Curve (AUC), Akurasi Berdasarkan Rasio Dataset, dan Kecepatan Klasifikasi. Perbandingan ini akan lebih komprehensif dibandingkan dengan penelitian sebelumnya, seperti analisis sentimen terhadap boy band BTS (Noviana & Rasal, 2023) dan layanan paylater (Safira & Hasan,

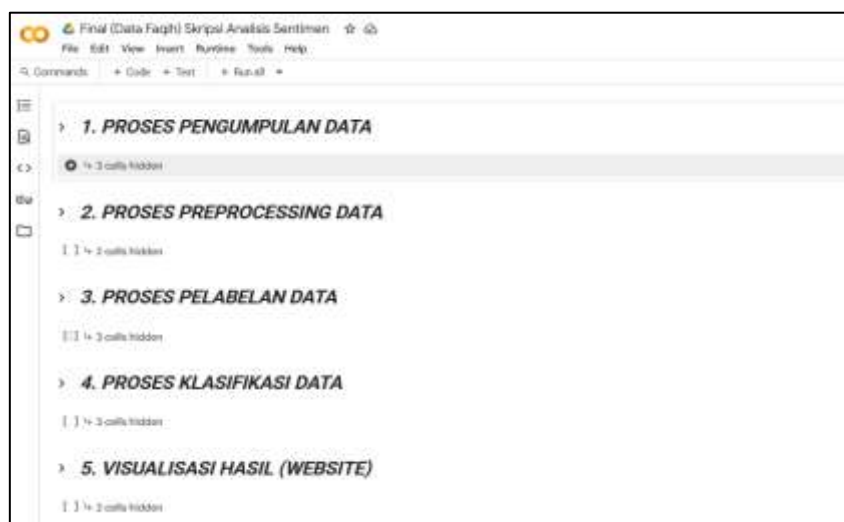
2023) [7][8], serta ulasan aplikasi AI-Qur'an digital (Aisah et al., 2023) [9]. Penelitian ini akan membedakan dengan fokus pada analisis ulasan DeepSeek, mengumpulkan 2.400 ulasan dari Januari-Juli 2025. Dengan pendekatan komprehensif dan visualisasi data, penelitian ini diharapkan memberikan wawasan lebih jelas tentang persepsi pengguna terhadap aplikasi DeepSeek dan dampak teknologi chatbot generatif pada industri teknologi secara keseluruhan.

2. Metode Penelitian



Gambar 1. Alur Pengolahan Data

Pada Gambar 1, digambarkan secara rinci proses Alur Pengolahan Data yang digunakan dalam penelitian analisis sentimen ulasan aplikasi DeepSeek Assistant AI dengan penerapan metode Naïve Bayes dan Support Vector Machine (SVM). Proses dimulai dengan pengumpulan data ulasan pengguna yang diperoleh dari platform seperti Google Play Store, yang kemudian memasuki tahap Pre-Processing Data, di mana data dibersihkan dan dipersiapkan untuk memastikan kualitas yang optimal sebelum analisis lebih lanjut. Setelah itu, pada tahap pelabelan data, setiap ulasan diberikan label sentimen yang dapat berupa positif atau negatif, yang menjadi dasar untuk melatih model klasifikasi. Proses klasifikasi data kemudian dilakukan dengan mengaplikasikan kedua model machine learning, Naïve Bayes dan SVM, untuk menentukan sentimen yang terkandung dalam setiap ulasan. Hasil dari klasifikasi ini selanjutnya disajikan dalam bentuk visualisasi yang dapat diakses melalui website, memungkinkan peneliti dan pengguna untuk mengeksplorasi dan memahami hasil analisis secara lebih mendalam. Dengan demikian, alur ini tidak hanya mencakup pemrosesan data tetapi juga memperkenalkan cara-cara efektif untuk menampilkan hasil yang dapat mempermudah pengambilan keputusan berbasis data.



Gambar 2. Alur Pengolahan Data pada Google Colab

Gambar 2 menggambarkan lima tahapan analisis data yang dilakukan peneliti menggunakan Google Colab, yaitu pengumpulan data, preprocessing data, pelabelan data, klasifikasi data, dan visualisasi hasil dalam bentuk website.

1. Pengumpulan Data

Tahap pertama adalah mengumpulkan data. Dalam konteks ini, data yang dikumpulkan berupa ulasan pengguna aplikasi DeepSeek Assistant AI dari platform Google Play Store sebanyak 2.400 ulasan dari rentan waktu januari hingga juli 2025. Data ini akan menjadi dasar untuk analisis lebih lanjut.

2. Pre-Processing Data

Preprocessing merupakan tahap penting dalam text mining yang berfungsi untuk membersihkan dan mempersiapkan data teks sebelum dianalisis lebih lanjut [15].

3. Pelabelan Data

Lexicon-based adalah metode yang digunakan untuk analisis sentimen dengan memanfaatkan kamus kata yang berisi kata-kata yang sudah diberi label sentimen, seperti positif atau negatif. Setiap kata dalam ulasan dibandingkan dengan kamus tersebut untuk menentukan nilai sentimennya [16].

Setelah data dibersihkan, setiap ulasan akan diberi label sentimen (positif atau negatif). Proses pelabelan dilakukan secara otomatis menggunakan metode Lexicon-Based. Sebuah kamus sentimen Bahasa Indonesia digunakan untuk menghitung skor sentimen dari setiap ulasan. Berdasarkan skor yang diperoleh, ulasan tersebut kemudian dikategorikan sebagai 'positif' atau 'negatif'.

4. Klasifikasi Data

Pada tahap ini, data yang telah dilabeli dibagi menjadi dua bagian: data latih (80%) dan data uji (20%) lalu dilakukan pembobotan TF-IDF. Data latih digunakan untuk melatih dua model klasifikasi, yaitu Naive Bayes dan Support Vector Machine (SVM). Setelah model dilatih, keduanya akan digunakan untuk mengklasifikasikan data uji dan diuji performanya dalam memprediksi sentimen.

a. Naïve Bayes

Naïve Bayes adalah metode klasifikasi yang didasarkan pada teori probabilitas Bayes, yang digunakan untuk menentukan kategori suatu data berdasarkan probabilitas yang dihitung dari fitur-fiturnya [17]. Metode ini dikenal dengan istilah "naïve" karena mengasumsikan bahwa setiap fitur dalam dataset bersifat independen atau tidak saling bergantung, meskipun dalam praktiknya mungkin terdapat korelasi antar fitur [18].

$$P(C_i|X) = \frac{P(X|C)P(C)}{P(X)} \tag{1}$$

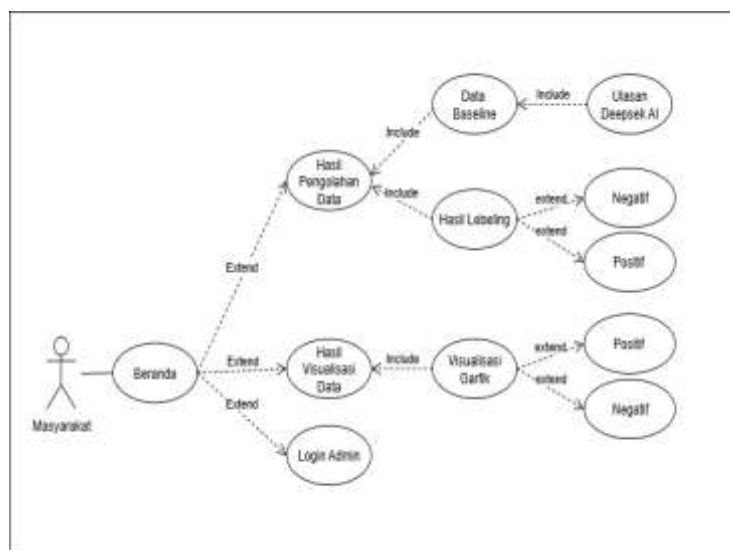
b. Support Vektor Machine

SVM adalah metode algoritma klasifikasi yang digunakan untuk klasifikasi, baik pada data yang bersifat linear maupun non-linear. SVM bekerja dengan memetakan data ke dalam ruang berdimensi tinggi, kemudian mencari hyperplane terbaik yang dapat membagi data ke dalam dua kelas secara optimal[19].

$$f(x) = w \cdot x + b \tag{2}$$

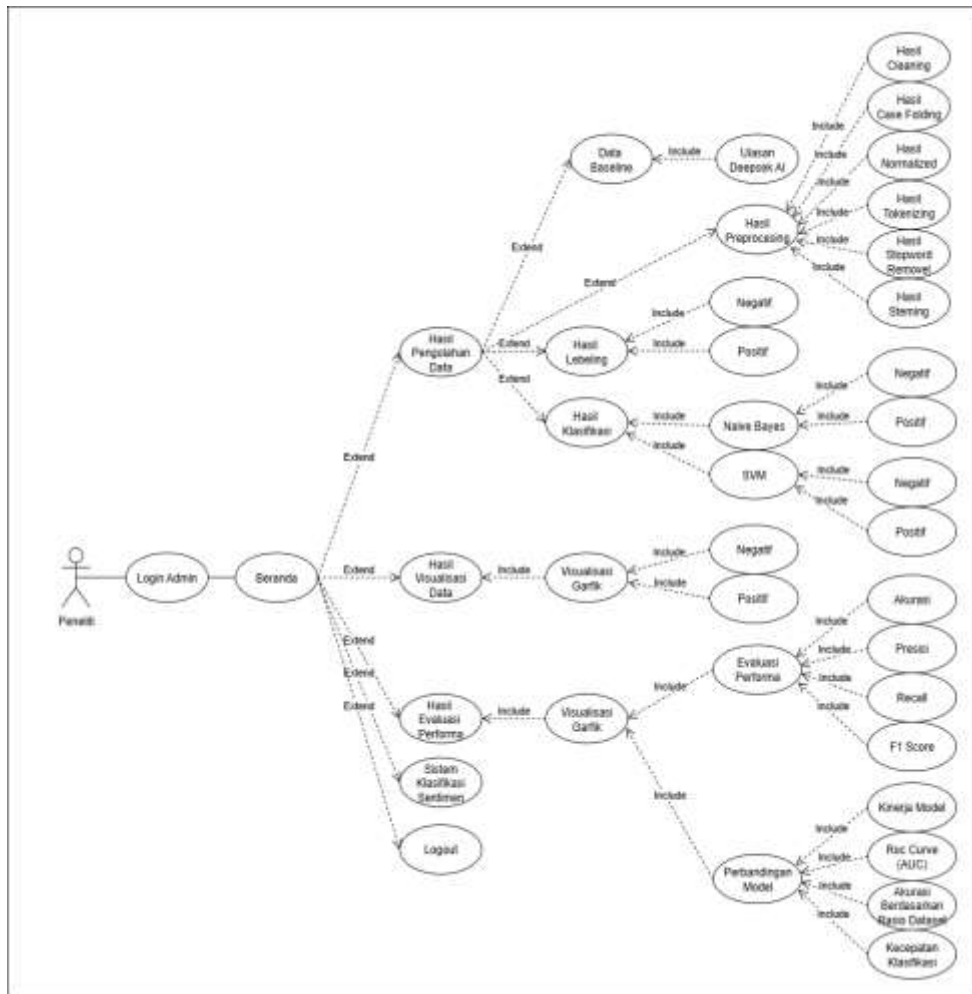
5. Hasil (Visualisasi Website)

Tahap terakhir adalah menyajikan seluruh hasil penelitian dalam bentuk dasbor interaktif. Aplikasi web ini dibangun menggunakan Gradio, yang mengintegrasikan semua data, hasil analisis, visualisasi grafik, dan model klasifikasi yang telah dibuat di Google Colab. Tujuan dari aplikasi ini adalah untuk mempermudah pemahaman hasil analisis.



Gambar 3. Use Case Bagian Masyarakat

Gambar 3 menggambarkan use case untuk masyarakat dalam sistem yang diusulkan, yang memberikan akses langsung kepada pengguna melalui halaman beranda untuk menjelajahi berbagai fitur yang tersedia. Pengguna dapat memulai dengan melihat hasil pengolahan data yang telah diproses, di mana data baseline yang telah diberi label sentimen positif dan negatif tersedia untuk diakses. Setelah itu, pengguna dapat mengeksplorasi hasil visualisasi data yang dihasilkan dari proses analisis, yang disajikan dalam bentuk grafik yang jelas dan mudah dipahami. Grafik ini memberikan gambaran tentang distribusi sentimen dalam ulasan aplikasi DeepSeek, dengan membedakan antara sentimen positif dan negatif, memungkinkan pengguna untuk memperoleh wawasan yang lebih mendalam mengenai persepsi masyarakat terhadap aplikasi tersebut. Visualisasi ini bertujuan untuk memperjelas hasil analisis dan memberikan informasi yang bermanfaat bagi pengguna untuk membuat keputusan berbasis data.



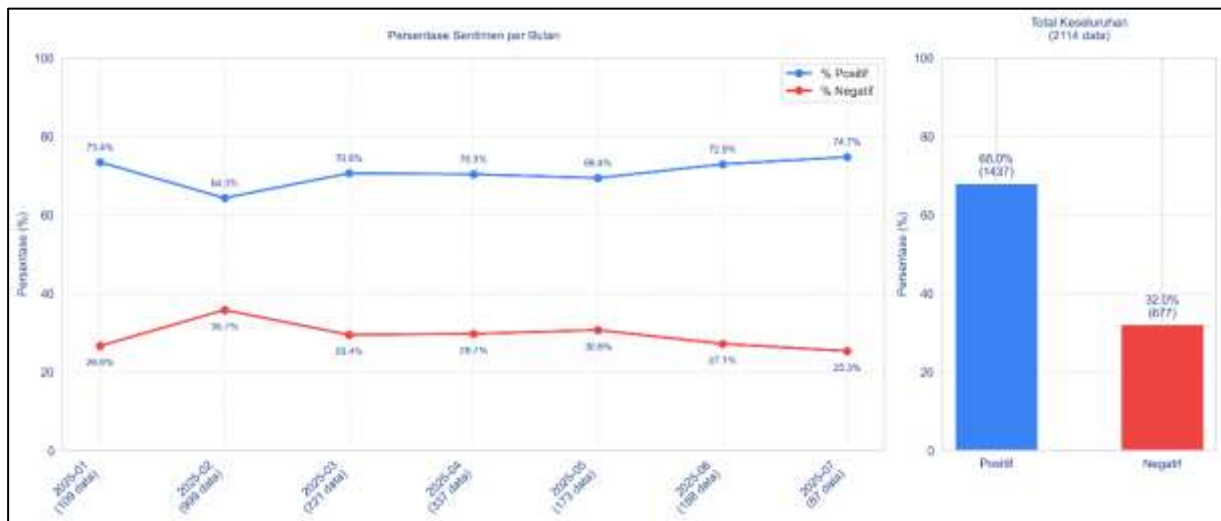
Gambar 4. Use Case Bagian Peneliti

Gambar 4 menggambarkan alur penggunaan website yang dimulai dengan proses login oleh Peneliti, yang mengarah ke halaman beranda yang menyediakan berbagai menu akses. Setelah berhasil login, Peneliti dapat dengan mudah mengakses berbagai fitur, seperti hasil pengolahan data, data baseline, serta hasil dari proses preprocessing, pelabelan, dan klasifikasi menggunakan algoritma Naive Bayes atau SVM. Website ini juga dilengkapi dengan fitur untuk mengevaluasi kinerja model menggunakan Confusion Matrix, yang memberikan gambaran yang jelas tentang akurasi klasifikasi dan kesalahan model. Selain itu, Peneliti dapat memanfaatkan visualisasi data berupa grafik yang menyederhanakan pemahaman hasil analisis, serta menggunakan fitur perbandingan model untuk melihat perbedaan performa antara Naive Bayes dan SVM. Sebagai tambahan, tersedia menu sistem klasifikasi sentimen yang memungkinkan Peneliti untuk menguji model dengan data baru secara langsung, memberikan kemudahan dalam melakukan analisis dan evaluasi model secara lebih mendalam dan menyeluruh. Dengan fitur-fitur tersebut, website ini menawarkan platform yang komprehensif untuk mendalami dan mengoptimalkan proses klasifikasi sentimen dalam penelitian.

3. Hasil dan Diskusi

3.1 Visualisasi Data Sentimen

Bagian ini menyajikan grafik visualisasi hasil klasifikasi sentimen positif dan negatif dari 2114 ulasan yang dikumpulkan antara Januari dan Juli 2025. Analisis ini bertujuan untuk memberikan wawasan mengenai bagaimana persepsi pengguna terhadap aplikasi DeepSeek berkembang seiring waktu dan bagaimana sentimen ini mempengaruhi pengalaman pengguna. Grafik yang disajikan menggambarkan perubahan sentimen per bulan, yang memungkinkan peneliti untuk mengidentifikasi tren atau pola tertentu dalam ulasan sepanjang periode tersebut. Selain itu, word cloud digunakan untuk menampilkan kata-kata yang paling sering muncul dalam ulasan, dengan ukuran kata yang lebih besar menunjukkan frekuensi kemunculannya. Kata-kata yang muncul paling sering, baik dalam ulasan positif maupun negatif, memberikan gambaran tentang topik atau fitur yang paling dibicarakan oleh pengguna, yang dapat mencakup aspek-aspek penting dari aplikasi seperti fungsionalitas, kegunaan, atau masalah teknis. Gabungan antara grafik perubahan sentimen dan word cloud ini memberikan pemahaman yang lebih mendalam mengenai sentimen dan opini pengguna terhadap aplikasi DeepSeek, serta memungkinkan peneliti untuk lebih mudah mengidentifikasi faktor-faktor yang mempengaruhi pengalaman pengguna.



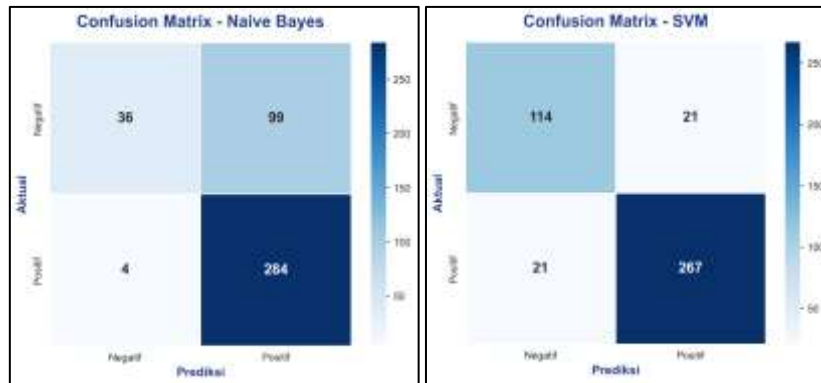
Gambar 5. Hasil Grafik Batang Positif dan Negatif

Gambar 5 adalah grafik visualisasi data klasifikasi yang menunjukkan persentase sentimen positif dan negatif pada data yang diklasifikasikan per bulan dari Januari 2025 hingga Juli 2025, dengan total data 2114 ulasan. Grafik ini terdiri dari dua bagian utama yang memberikan gambaran menyeluruh mengenai sentimen pengguna terhadap aplikasi DeepSeek sepanjang periode waktu tersebut.

- Grafik Garis: Grafik ini menampilkan persentase sentimen positif dan negatif setiap bulan. Garis biru menggambarkan persentase sentimen positif, sementara garis merah menunjukkan persentase sentimen negatif. Pada bulan Januari 2025, yang merupakan bulan peluncuran aplikasi DeepSeek, sentimen positif mencapai 73,4%, yang mencerminkan tanggapan yang sangat baik dari pengguna. Namun, pada bulan Februari 2025, terjadi penurunan persentase sentimen positif menjadi 64%. Penurunan ini mungkin disebabkan oleh beberapa masalah atau kekurangan dalam aplikasi yang dirasakan oleh pengguna. Meskipun demikian, mulai bulan kedua hingga bulan ketujuh (Juli 2025), persentase sentimen positif kembali meningkat, mencapai puncaknya, yang menunjukkan bahwa aplikasi ini telah diperbaiki dan pengguna merasa lebih puas seiring waktu.
- Grafik Batang: Grafik batang menunjukkan persentase total sentimen positif dan negatif dari seluruh data yang terkumpul. Persentase sentimen positif tercatat sebesar 68,0% (1437 ulasan), sementara sentimen negatif sebesar 32,0% (677 ulasan). Meskipun ada penurunan di bulan kedua, secara keseluruhan aplikasi DeepSeek memperoleh respon positif yang lebih dominan dari pengguna, menandakan bahwa meskipun ada tantangan di awal peluncuran, aplikasi ini berhasil meraih opini yang lebih baik seiring waktu.

b) Korelasi Negatif: Grafik ini juga menampilkan korelasi antara kata-kata yang sering muncul dalam ulasan negatif. Sebagai contoh, kata "mantap" memiliki korelasi dengan kata "server" (0.53) dan "kurang" (0.29), yang menunjukkan bahwa meskipun "mantap" umumnya muncul dalam ulasan positif, kata ini juga digunakan dalam ulasan negatif, terutama yang berkaitan dengan masalah teknis seperti kestabilan server. Selain itu, kata "please" memiliki korelasi tinggi dengan kata "mantap" (0.53) dan "server" (0.62), menunjukkan bahwa beberapa pengguna berharap agar aplikasi ini dapat memberikan performa yang lebih baik, terutama dalam hal server.

Grafik ini memberikan gambaran yang jelas mengenai bagaimana kata-kata tertentu berhubungan satu sama lain dalam konteks ulasan positif dan negatif, membantu untuk lebih memahami persepsi pengguna terhadap berbagai aspek aplikasi DeepSeek, baik dalam hal keunggulan maupun masalah yang perlu diperbaiki.

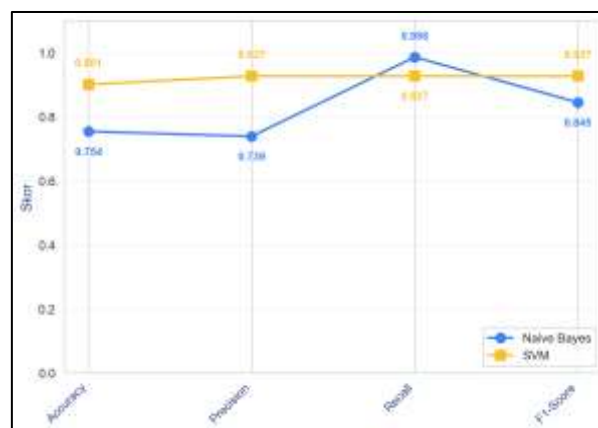


Gambar 8. Visualisasi Hasil Evaluasi Confusion Matrix

Gambar 8 menunjukkan Confusion Matrix yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi dan label aktual [20]. Confusion Matrix ini terbagi menjadi empat kategori utama: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Pada Confusion Matrix Naive Bayes, terdapat 36 (TN), 99 (FP), 4 (FN), dan 284 (TP), yang menunjukkan bahwa model ini menghasilkan 284 prediksi benar untuk sentimen positif, 36 prediksi benar untuk sentimen negatif, tetapi juga menghasilkan 99 prediksi negatif yang salah untuk data positif dan 4 prediksi positif yang salah untuk data negatif. Sementara itu, pada Confusion Matrix SVM, terdapat 114 (TN), 21 (FP), 21 (FN), dan 267 (TP). Ini menunjukkan bahwa model SVM memiliki 21 prediksi positif yang benar dan 267 prediksi negatif yang salah, dengan lebih banyak kesalahan pada kategori negatif dibandingkan Naive Bayes.

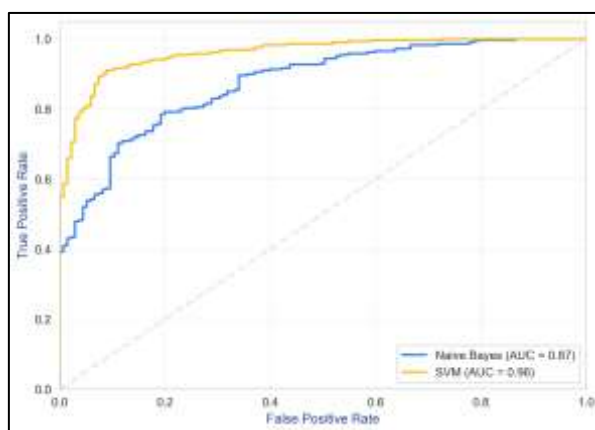
3.2 Hasil Perbandingan Metode

Berikut hasil Perbandingan machine learning Naive Bayes dan Support Vector Machine. berdasarkan empat aspek, yaitu Kinerja Model, Area Under Curve, Akurasi Berdasarkan Rasio Dataset, dan Kecepatan Klasifikasi.



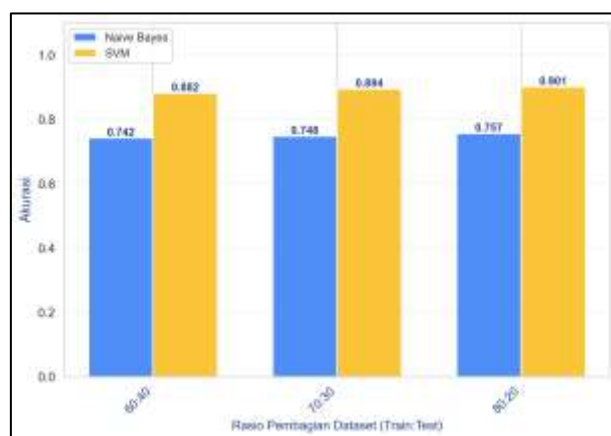
Gambar 9. Grafik Kinerja Model

Gambar 9 adalah grafik yang membandingkan kinerja model Naïve Bayes dan SVM berdasarkan empat metrik evaluasi utama: Akurasi, Precision, Recall, dan F1-Score. Grafik ini menggunakan dua warna untuk mempermudah perbandingan, dengan biru mewakili Naïve Bayes dan kuning mewakili SVM. Berdasarkan Akurasi, SVM unggul dengan nilai 0.901, sementara Naïve Bayes mencatatkan nilai 0.754, menunjukkan bahwa SVM lebih efektif dalam memberikan prediksi yang tepat secara keseluruhan. Dalam hal Precision, yang mengukur ketepatan model dalam mengklasifikasikan kelas positif, SVM kembali lebih unggul dengan nilai 0.927, dibandingkan dengan Naïve Bayes yang hanya mencapai 0.739. Namun, Recall, yang mengukur kemampuan model dalam menangkap semua prediksi positif yang sebenarnya, menunjukkan keunggulan Naïve Bayes dengan nilai 0.986, lebih tinggi daripada SVM yang hanya mencapai 0.927. Untuk F1-Score, yang menggabungkan Precision dan Recall dalam satu metrik, kedua model memiliki nilai yang hampir sama, yakni 0.927 untuk SVM dan sedikit lebih rendah pada Naïve Bayes dengan 0.845. Grafik ini memberikan gambaran yang jelas tentang kekuatan dan kelemahan masing-masing model dalam keempat metrik tersebut, serta membantu menentukan model yang paling sesuai berdasarkan kebutuhan analisis tertentu.



Gambar 10. Grafik Area Under Curve (AUC).

Gambar 10 adalah grafik AUC yang membandingkan performa Naïve Bayes dan SVM berdasarkan ROC Curve. Grafik ini menampilkan dua kurva, dengan Naïve Bayes (biru) yang memiliki nilai AUC = 0.87, menunjukkan kemampuan yang baik dalam membedakan kelas positif dan negatif, meskipun masih kalah dibandingkan dengan SVM. SVM (kuning), di sisi lain, memiliki nilai AUC = 0.96, yang menunjukkan performa yang lebih unggul dalam membedakan antara kedua kelas. AUC (Area Under the Curve) mengukur seberapa baik model dapat memisahkan kelas positif dan negatif, dan semakin mendekati nilai 1, semakin baik kinerja model tersebut. Dalam grafik ini, SVM menunjukkan kurva yang lebih dekat ke sudut kiri atas, yang mengindikasikan bahwa model ini memiliki kemampuan yang lebih baik dalam memprediksi kelas yang benar, dengan sedikit kesalahan pada kedua kelas. Dengan demikian, meskipun Naïve Bayes menunjukkan hasil yang cukup baik, SVM jelas lebih unggul dalam hal kemampuan discriminasi antara kelas.

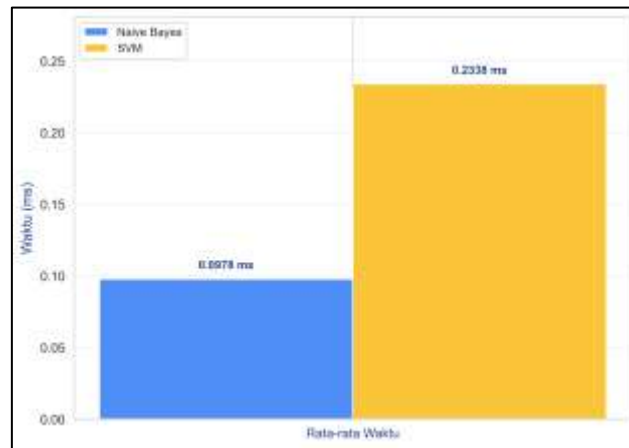


Gambar 11. Grafik Akurasi Berdasarkan Rasio Dataset

DOI: <https://doi.org/10.31004/riggs.v4i3.2511>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Gambar 11 memperlihatkan grafik yang membandingkan akurasi model Naïve Bayes dan SVM berdasarkan berbagai rasio pembagian dataset antara Train dan Test, yaitu 60:40, 70:30, dan 80:20. Pada rasio 60:40, Naïve Bayes memperoleh akurasi 0.742, sedangkan SVM mencatatkan akurasi lebih tinggi sebesar 0.882, menunjukkan keunggulan SVM dalam mengklasifikasikan data. Pada rasio 70:30, meskipun akurasi Naïve Bayes sedikit meningkat menjadi 0.748, SVM tetap unggul dengan akurasi 0.894. Begitu pula pada rasio 80:20, Naïve Bayes memperoleh akurasi 0.757, sementara SVM mencapai akurasi tertinggi sebesar 0.901. Hasil ini secara konsisten menunjukkan bahwa SVM menghasilkan akurasi yang lebih tinggi dibandingkan Naïve Bayes pada setiap rasio pembagian dataset, yang menandakan bahwa SVM lebih efektif dan stabil dalam mengklasifikasikan data, terutama ketika proporsi data latih semakin besar. Grafik ini juga mengindikasikan bahwa meskipun Naïve Bayes menunjukkan kinerja yang baik, SVM lebih unggul dalam menangani kompleksitas dataset yang lebih besar dan lebih variatif.



Gambar 12. Grafik kecepatan klasifikasi

Gambar 12 adalah grafik yang membandingkan kecepatan klasifikasi antara Naïve Bayes dan SVM berdasarkan rata-rata waktu yang dibutuhkan untuk proses klasifikasi. Hasil menunjukkan bahwa Naïve Bayes memerlukan waktu 0.0978 ms, yang jauh lebih cepat dibandingkan dengan SVM yang membutuhkan waktu 0.2338 ms. Grafik ini jelas mengindikasikan bahwa Naïve Bayes lebih unggul dalam hal kecepatan proses klasifikasi, menjadikannya pilihan yang lebih efisien untuk aplikasi yang memerlukan waktu respons cepat, seperti sistem yang beroperasi secara real-time atau dalam lingkungan dengan data besar yang memerlukan klasifikasi instan. Meskipun SVM menunjukkan kinerja yang lebih baik dalam hal akurasi dan AUC, Naïve Bayes tetap lebih unggul dalam situasi yang mengutamakan efisiensi waktu, membuatnya sangat cocok untuk aplikasi yang mengutamakan kecepatan tanpa mengorbankan terlalu banyak akurasi.

3.3 Perbandingan 4 Aspek Perfroma Model

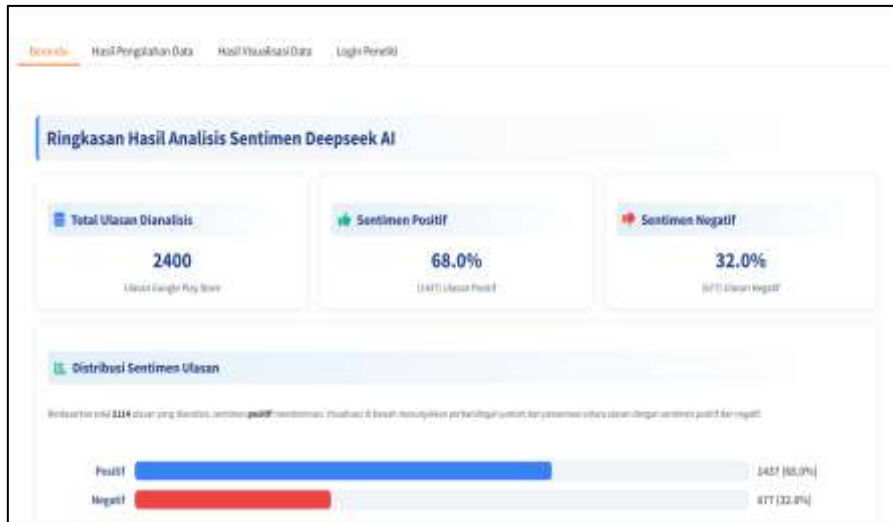
Tabel 1 Hasil Tabel Perbandingan 4 Aspek Perfroma Model

	Naïve Bayes	Support Vektor Machine
Perbandingan Kinerja Model		✓
Perbandingan AUC (ROC Curve)		✓
Perbandingan Akurasi Berdasarkan Rasio Dataset		✓
Perbandingan Kecepatan Klasifikasi	✓	

Tabel 1 menunjukkan perbandingan empat aspek performa model, dengan tanda centang (✓) yang menunjukkan aspek yang diuji pada masing-masing model. Dalam Perbandingan Kinerja Model, SVM lebih unggul dibandingkan Naïve Bayes, sehingga SVM mendapatkan tanda centang pada kolom ini. Untuk Perbandingan AUC (ROC Curve), SVM juga menunjukkan keunggulan dengan nilai AUC = 0.96, lebih tinggi dibandingkan dengan Naïve Bayes yang hanya mencapai AUC = 0.87. Dalam Perbandingan Akurasi Berdasarkan Rasio Dataset, SVM menunjukkan akurasi yang lebih baik pada setiap rasio pembagian dataset dibandingkan Naïve Bayes, menandakan kemampuannya yang lebih baik dalam mengklasifikasikan data dalam berbagai pembagian. Namun, pada

Perbandingan Kecepatan Klasifikasi, Naïve Bayes lebih unggul karena memerlukan waktu yang lebih sedikit untuk melakukan klasifikasi dibandingkan SVM. Secara keseluruhan, SVM unggul dalam kinerja model, AUC, dan akurasi, sementara Naïve Bayes lebih efisien dalam hal kecepatan klasifikasi, menjadikannya lebih cocok untuk aplikasi yang membutuhkan proses klasifikasi yang cepat.

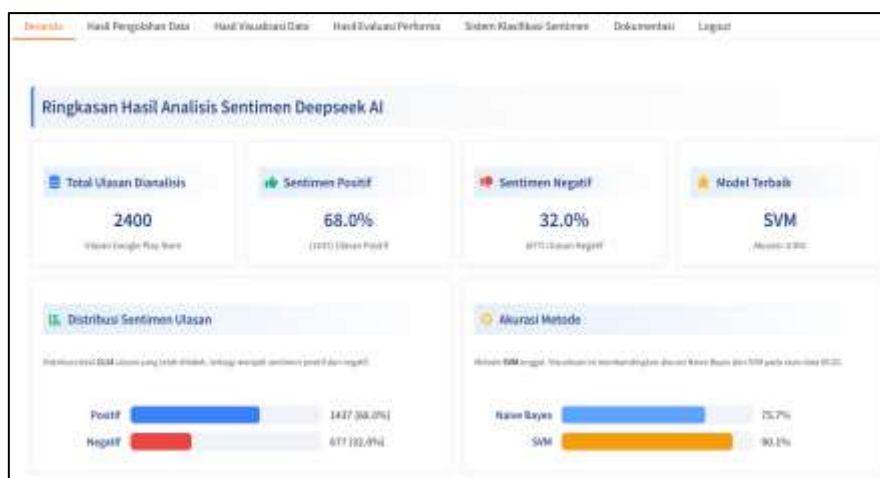
3.4 Visualisasi website bagian Masyarakat



Gambar 13. Visualisasi website bagian Masyarakat

Gambar 13 adalah halaman beranda pada website analisis sentimen aplikasi DeepSeek AI dirancang dengan antarmuka yang ramah pengguna, dilengkapi dengan beberapa menu utama yang memudahkan akses ke berbagai informasi dan hasil analisis secara lebih mendalam. Menu pertama adalah Hasil Pengolahan Data, yang menyediakan informasi terkait proses pengolahan data ulasan pengguna sebelum dilakukan analisis sentimen, termasuk tahap preprocessing dan pelabelan data. Menu kedua, Hasil Visualisasi Data, menyajikan berbagai grafik dan visualisasi yang menggambarkan perbandingan sentimen positif dan negatif, serta metrik evaluasi performa model klasifikasi yang digunakan, seperti akurasi, precision, recall, dan AUC, yang memberikan gambaran menyeluruh tentang efektivitas model. Terakhir, terdapat Menu Login Peneliti, yang memungkinkan peneliti untuk mengakses fitur lanjutan, seperti evaluasi model, analisis lebih lanjut, serta pengujian model dengan dataset baru. Dengan menu-menu ini, halaman beranda menyediakan akses mudah dan efisien bagi pengguna untuk mengeksplorasi dan memahami hasil analisis sentimen secara komprehensif.

3.5 Visualisasi website bagian Peneliti



Gambar 14. Visualisasi website bagian Peneliti

Gambar 14 adalah halaman beranda pada website analisis sentimen aplikasi DeepSeek AI untuk peneliti dirancang dengan berbagai informasi dan menu yang memudahkan peneliti dalam mengakses dan memahami hasil analisis sentimen secara komprehensif. Di bagian atas halaman, terdapat Ringkasan Hasil Analisis Sentimen, yang menampilkan informasi penting seperti total ulasan yang dianalisis, persentase sentimen positif (68%) dan negatif (32%), serta model terbaik yang digunakan (SVM dengan akurasi 90,1%). Di bawahnya, terdapat Distribusi Sentimen Ulasan, yang menyajikan grafik batang untuk memberikan perbandingan visual yang jelas antara ulasan positif dan negatif. Menu Akurasi Metode memungkinkan peneliti untuk melihat perbandingan performa model Naïve Bayes dan SVM berdasarkan akurasi pada rasio data 80:20, memberikan wawasan tentang efektivitas masing-masing model. Website ini juga dilengkapi dengan menu-menu tambahan seperti Hasil Pengolahan Data, yang memberikan informasi mengenai proses pengolahan data ulasan pengguna sebelum dilakukan analisis sentimen, dan Hasil Visualisasi Data, yang menampilkan berbagai grafik dan metrik evaluasi untuk mengevaluasi kinerja model. Hasil Evaluasi Performa menyediakan tampilan rinci tentang kinerja model berdasarkan berbagai metrik, sementara Sistem Klasifikasi Sentimen memungkinkan peneliti untuk menguji model pada data baru. Selain itu, terdapat Dokumentasi, yang menyajikan penjelasan teknis dan panduan penggunaan sistem, serta Menu Logout yang memungkinkan peneliti untuk keluar dari sistem dengan mudah. Dengan fitur-fitur ini, halaman beranda menyediakan navigasi yang intuitif dan akses yang mudah untuk mendalami hasil analisis sentimen dan mengevaluasi performa model secara menyeluruh.

4. Kesimpulan

Analisis sentimen terhadap aplikasi DeepSeek Assistant AI menunjukkan bahwa 68% ulasan bersentimen positif dan 32% bersentimen negatif dari total 2114 data. Model Naive Bayes mengklasifikasikan 90,5% sentimen positif dan 9,5% negatif, sementara SVM mengklasifikasikan 67,8% positif dan 32,2% negatif. Kata-kata seperti "mantap", "ai", dan "server" muncul pada sentimen positif, mencerminkan kepuasan, sedangkan kata "kurang", "sibuk", dan "server" muncul pada sentimen negatif, menunjukkan ketidakpuasan. Dalam perbandingan performa, SVM unggul dengan skor lebih tinggi pada hampir semua aspek evaluasi, termasuk akurasi (SVM: 0.927 vs Naive Bayes: 0.754), precision (SVM: 0.927 vs Naive Bayes: 0.739), dan F1-Score (SVM: 0.927 vs Naive Bayes: 0.854), meskipun Naive Bayes unggul pada recall (Naive Bayes: 0.986 vs SVM: 0.845) dan lebih efisien dalam waktu klasifikasi (Naive Bayes: 0.2486 ms vs SVM: 0.3758 ms). Pada perbandingan ROC Curve (AUC), SVM juga unggul dengan skor AUC 0.96, lebih tinggi dibandingkan Naive Bayes yang memperoleh AUC 0.87. Berdasarkan Akurasi Berdasarkan Rasio Dataset, pada rasio 60:40, SVM mencapai 0.882 sementara Naive Bayes 0.742; pada rasio 70:30, SVM mencapai 0.894 dan Naive Bayes 0.748; dan pada rasio 80:20, SVM mencapai 0.901 dan Naive Bayes 0.757. Website yang dibangun menyediakan visualisasi hasil analisis untuk dua aktor utama: masyarakat yang mendapatkan informasi lebih luas dan komprehensif melalui grafik sentimen, dan peneliti yang dapat mengakses alat analisis lebih mendalam untuk penelitian lanjutan.

Referensi

- [1] Zuzanna Bocian, "Key Chatbot Statistics You Should Follow in 2024." Diakses: 3 Agustus 2025. [Daring]. Tersedia pada: <https://www.chatbot.com/blog/chatbot-statistics/>
- [2] Anthony Cardillo, "40+ Chatbot Statistics (2025)." Diakses: 3 Agustus 2025. [Daring]. Tersedia pada: <https://explodingtopics.com/blog/chatbot-statistics>
- [3] Prasatya, "DeepSeek AI Buat China." Diakses: 3 Agustus 2025. [Daring]. Tersedia pada: <https://www.codepolitan.com/blog/apa-itu-deepseek-ai-buatan-china-yang-siap-gantiin-chatgpt/>
- [4] CNN Indonesia, "7 Fakta DeepSeek yang Guncang AS dan Bikin Rugi Ribuan Triliun Nvidia."
- [5] Anggie Irfansyah, "Efisiensi Biaya Pelatihan Deepseek AI." Diakses: 3 Agustus 2025. [Daring]. Tersedia pada: <https://eduparx.id/blog/insight/munculnya-deepseek-inovasi-ai-tiongkok-yang-mengguncang-dunia-teknologi/>
- [6] Triya andriyani, "DeepSeek Ramai Diblokir Banyak Negara, Pakar UGM Nilai Setiap Negara Punya Kedaulatan Digital." Diakses: 3 Agustus 2025. [Daring]. Tersedia pada: <https://ugm.ac.id/id/berita/deepseek-ramai-diblokir-banyak-negara-pakar-ugm-nilai-setiap-negara-punya-kedaulatan-digital/>
- [7] R. Noviana dan I. Rasal B A Jurusan, "PENERAPAN ALGORITMA NAIVE BAYES DAN SVM UNTUK ANALISIS SENTIMEN BOY BAND BTS PADA MEDIA SOSIAL TWITTER," *JTS*, vol. 2, no. 2, hlm. 59, Jun 2023.
- [8] A. Safira, A. S. Masyarakat...v, dan F. N. Hasan, "ANALISIS SENTIMEN MASYARAKAT TERHADAP PAYLATER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER," *Jurnal Sistem Informasi*, vol. 5, no. 1, Jan 2023.
- [9] I. S. Aisah, B. Irawan, dan T. Suprapti, "ALGORITMA SUPPORT VECTOR MACHINE (SVM) UNTUK ANALISIS SENTIMEN ULASAN APLIKASI AL QUR'AN DIGITAL," Cirebon, Des 2023.
- [10] M. G. Undap, V. P. Rantung, dan P. T. D. Rompas, "Analisis Sentimen Situs Pembajak Artikel Penelitian Menggunakan Metode Lexicon-Based," Des 2021.
- [11] B. Ramadhani dan R. R. Suryono, "Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 2, hlm. 10, Apr 2024, doi: 10.30865/mib.v8i2.7458.
- [12] A. Pebdika, R. Herdiana, dan D. Solihudin, "KLASIFIKASI MENGGUNAKAN METODE NAIVE BAYES UNTUK MENENTUKAN CALON PENERIMA PIP," Cirebon, Feb 2023.
- [13] Tinaliah dan Triana Elizabeth, "Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode Support Vector Machine," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 4, hlm. 6, Des 2022, [Daring]. Tersedia pada: <http://jurnal.mdp.ac.id>

DOI: <https://doi.org/10.31004/riggs.v4i3.2511>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

- [14] Rifina Dwi Pebriant, "ANALISIS SENTIMEN MASYARAKAT PLATFORM X TERHADAP KORUPSI PT. PERTAMINA (PERSERO) MENGGUNAKAN METODE SVM," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 2, Apr 2025, doi: 10.23960/jitet.v13i2.6399.
- [15] N. A. Fadhlurrohman, A. Primajaya, dan A. Nugraha, "ANALISIS SENTIMEN TERHADAP SKEMA STUDENT LOAN UNTUK BIAYA PERGURUAN TINGGI PADA TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES," Karawang, Apr 2025.
- [16] Irma Surya Kumala Idris, Yasin Aril Mustofa, dan Irvan Abraham Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)," *J Inf Sci*, vol. 36, no. 6, hlm. 4–848, Jan 2023, doi: 10.1177/0165551510388123.
- [17] T. Ridwansyah, "Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan Naïve Bayes Classifier," *Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 5, hlm. 2, Apr 2022, [Daring]. Tersedia pada: <https://djournals.com/klik>
- [18] Kardandi Alfarizi Siregar, Salsabila Nasution, dan Putri Nabawy, "Analisis Sentimen Netizen Indonesia Terhadap Kampanye Penggunaan Kecerdasan Buatan Oleh Pemerintah Menggunakan Algoritma Naive Bayes," *Jurnal Ilmu Komputer Dan Informatika*, hlm. 9, Apr 2025.
- [19] Brilian Prilindaputra, Dinda Rima Rachcita Putri, dan Nurissaidah Ulinnuha, "Implementasi Support Vector Machine untuk Analisis Sentimen Aplikasi Deepseek," Surabaya, 2025.
- [20] D. Normawati dan S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," Yogyakarta, Sep 2021.