



Klasifikasi Sentimen Ulasan Aplikasi Claude Menggunakan Naive Bayes dan Support Vector Machine

Muhammad Zidna Mubarak^{1*}, Saiful Rizqi², Ramadhan Muchamad Rizal³, Andhita Bagus Noor Syalim⁴, Nur Iksan⁵

^{1,2,3,4,5} Teknik Elektro, Fakultas Teknik, Universitas Negeri Semarang, Indonesia

Info Artikel

Riwayat Artikel:Diterima : **12 Mei 2025**Direvisi : **28 Juni 2025**Diterbitkan : **30 Juni 2025****Kata Kunci:**

Analisis Sentimen,

Naive Bayes,

Support Vector Machine

ABSTRAK

Analisis sentimen adalah teknik penting yang digunakan untuk memproses data tekstual untuk mengetahui pendapat pengguna mengenai suatu produk. Penelitian ini menggunakan teknik *machine learning*, khususnya Multinomial Naive Bayes dan Support Vector Machine (SVM) Linear, untuk mengkategorikan sentimen ulasan untuk aplikasi Claude yang tersedia di Google Play Store. Data ulasan diperoleh melalui *web scraping* dan kemudian dibersihkan dan diproses untuk ekstraksi fitur menggunakan metode TF-IDF. Evaluasi model dilakukan dengan menggunakan matriks seperti akurasi, presisi, *recall*, dan *f1-score*. Temuan menunjukkan bahwa kedua model menunjukkan kemampuan klasifikasi yang kuat, dengan SVM Linear mencapai akurasi tertinggi sebesar 99,8%, presisi 99,9008%, *recall* 99,7446%, dan *f1-score* 99,8221%. Visualisasi melalui WordCloud dan analisis frekuensi kata membantu dalam mengidentifikasi istilah-istilah kunci yang terkait dengan setiap kategori sentimen. Hasil ini menggambarkan bahwa teknik *machine learning* efektif dalam menangkap sentimen pengguna dan menjadi dasar untuk menciptakan aplikasi berbasis AI yang lebih berorientasi pada pengguna.

Keywords:

Sentiment Analysis,

Naive Bayes,

Support Vector Machine

ABSTRACT

Sentiment analysis is an essential technique used to process textual data in order to ascertain user opinions regarding a product. This research employs machine learning techniques, specifically Multinomial Naive Bayes and Support Vector Machine (SVM) Linear, to categorize the sentiment of reviews for the Claude application available on the Google Play Store. Review data were obtained via web scraping and subsequently cleaned and processed for feature extraction using the TF-IDF method. Model evaluation was performed utilizing metrics such as accuracy, precision, recall, and f1-score. The findings indicate that both models demonstrated strong classification capabilities, with SVM Linear achieving the highest accuracy of 99.8%, precision of 99.9008%, recall of 99.7446%, and f1-score of 99.8221%. Visualization through WordCloud and an analysis of word frequency aided in identifying the key terms associated with each sentiment category. These results illustrate that machine learning techniques are effective in capturing user sentiments and lay the groundwork for creating more user-oriented AI-driven applications.

Penulis Korespondensi:

Muhammad Zidna Mubarak,
Program Studi Teknik Elektro,
Universitas Negeri Semarang

Email: zidnamubarak8@students.unnes.ac.id

This is an open access article under the [CC BY-SA](#) license



1. PENDAHULUAN

Kecerdasan buatan (*Artificial Intelligence* atau AI) telah berkembang pesat dalam beberapa dekade terakhir dan diaplikasikan ke berbagai bidang, termasuk dalam pengembangan asisten digital berbasis AI seperti Claude dari Anthropic. Claude memiliki setidaknya dua keunggulan utama yang membuatnya sangat menarik untuk dimanfaatkan dalam kegiatan pembelajaran, yaitu kemampuannya untuk menerima unggahan file serta kapasitas pemrosesan teks yang mencapai 75.000 kata [1]. Claude menawarkan kebijakan privasi yang lebih terbuka dan memberi pengguna kendali penuh atas data mereka, termasuk opsi untuk menghapus data pribadi kapan saja serta tidak menyimpan percakapan kecuali diminta, sehingga menjadi pilihan ideal bagi pengguna yang mengutamakan privasi dan kontrol data [2]. Aplikasi ini memiliki ribuan ulasan pengguna, menjadikannya salah satu produk berbasis bahasa alami yang tersedia di Google Play Store. Ulasan tersebut mengandung informasi penting yang mencerminkan persepsi dan kepuasan pengguna, yang apabila dianalisis dengan baik, dapat menjadi dasar untuk evaluasi performa dan pengembangan lebih lanjut terhadap aplikasi.

Analisis sentimen berfungsi sebagai alat yang ampuh untuk menganalisis data ulasan untuk menentukan apakah opini itu positif, negatif, atau netral. Aplikasi spesifik dari *text mining* adalah analisis sentimen, yang berusaha mengungkap opini yang diungkapkan dalam teks [3]. *Text mining* memiliki beberapa cabang salah satunya adalah analisis sentimen, yang berusaha menilai sudut pandang individu atau kolektif mengenai subjek tertentu, mengidentifikasinya sebagai positif atau negatif [4]. Teknik ini semakin relevan seiring meningkatnya volume data teks yang dihasilkan secara daring. Pendekatan *lexicon-based* memiliki keunggulan dalam kemudahan implementasi, namun cenderung terbatas dalam menangkap konteks semantik dan ekspresi bahasa yang kompleks [5][6]. Oleh karena itu, metode *machine learning* dengan algoritma *Support Vector Machine (SVM)* serta *Naive Bayes* menjadi alternatif yang lebih efisien dan fleksibel [7][8][9]. *Naive Bayes* merupakan teknik pengelompokan probabilistik elementer yang berguna untuk pemrosesan bahasa alami, terutama ketika bekerja dengan kumpulan teks yang kompleks [10]. Sebaliknya, *SVM* dikenal dapat menangani data dengan dimensi tinggi dan kasus yang tidak seimbang dengan baik [11]. Kedua metode ini telah terbukti mampu memiliki akurasi dalam klasifikasi opini pengguna terhadap berbagai layanan digital [12].

Penelitian ini menggunakan teknik *supervised learning* seperti *SVM* maupun *Naive Bayes* guna melakukan pengkajian sentimen terhadap umpan balik pengguna mengenai aplikasi Claude. Data dikumpulkan melalui Google Play Store dengan cara *web scraping* dan setelah itu diproses dan dicadangkan dalam format .CSV. Tahap *preprocessing* melibatkan pembersihan teks, melakukan tokenisasi, dan mengekstraksi fitur menggunakan *TF-IDF*. Penelitian ini juga mengusulkan kerangka kerja klasifikasi sentimen yang didasarkan pada pembelajaran mesin yang disesuaikan dengan konteks lokal, menilai kemampuan model dengan presisi, *recall*, akurasi serta matriks *F1-score*, dan menawarkan visualisasi sentimen pengguna yang dapat digunakan untuk peningkatan aplikasi berbasis AI.

2. METODE PENELITIAN

2.1. Analisis Sentimen

Analisis sentimen ialah metode dalam pemrosesan bahasa alamiah (*NLP*) yang berguna untuk mengenali serta mengkategorikan argumen yang ditemukan dalam dokumen. Metode tersebut terbukti bermanfaat untuk mengevaluasi sikap pengguna terhadap produk digital, termasuk aplikasi AI seperti Claude. Pengumpulan bahan berupa ulasan pengguna menggunakan teknik *web scraping* dari website Google Play Store untuk menganalisis tren opini dan kinerja aplikasi dari sudut pandang pengguna. *Web scraping* merupakan pendekatan mekanis guna mengekstraksi data dari sebuah *website*, yang mencakup pengambilan tautan gambar dan pemilihan setiap gambar yang ada, dibandingkan dengan melakukan tugas-tugas ini secara manual [13].

2.2. Algoritma *Naive Bayes* dan *SVM*

Pendekatan *machine learning* digunakan sebab fleksibel serta dapat disesuaikan dengan berbagai bahasa yang ditemukan dalam teks. Penelitian ini memanfaatkan dua metode *supervised learning* yaitu *Support Vector Machine (SVM)* dan *Naive Bayes* dimana teknik ini banyak digunakan dan efisien. *Naive Bayes* merupakan teknik klasifikasi yang bergantung pada anggapan independensi di antara fitur-fitur, yang diturunkan dari teorema Bayes. Meskipun memiliki data latih yang relatif kecil dan proses komputasi yang efisien, algoritma ini dikenal

mampu memberikan hasil klasifikasi yang cukup baik [10]. Sementara itu, algoritma *SVM* bertujuan untuk mengidentifikasi *hyperplane* optimal yang membedakan data yang termasuk dalam dua kelas yang berbeda. Teknik ini efektif untuk set data berdimensi tinggi seperti teks, karena dapat mengurangi *overfitting* dan mencapai *margin* klasifikasi sebesar mungkin. *SVM* juga menunjukkan performa yang konsisten dalam kasus klasifikasi biner dan sangat efektif pada dataset dengan distribusi tidak seimbang [11].

2.3. Google Colab

Google Colab merupakan alat analisis berbasis *cloud* yang disukai para peneliti karena kesederhanaan dan keefektifannya. Platform ini menyediakan lingkungan Python terintegrasi dengan Google Drive serta mendukung pustaka open-source untuk *machine learning* dan *NLP*, tanpa memerlukan instalasi atau konfigurasi sistem secara lokal. Google Colab digunakan sebagai platform pemrosesan karena menyediakan lingkungan komputasi yang kuat dan akses gratis ke GPU, yang mempercepat proses pelatihan model [14]. Dalam konteks penelitian ini, seluruh proses analisis mulai dari *preprocessing*, pelatihan model, hingga evaluasi dilakukan pada platform Google Colab.

2.4. Tahapan Penelitian

a. Pengumpulan Data

Informasi yang dikumpulkan selama tahap pengumpulan diperoleh dengan menggunakan metode web scraping yang diterapkan pada Google Play Store. Data yang dikumpulkan terdiri dari ulasan dan peringkat pengguna untuk aplikasi Claude. Sebanyak 5000 ulasan berhasil dikumpulkan dalam rentang waktu 12 November 2024 hingga 24 April 2025. Data ini kemudian disimpan dalam format file .CSV agar mudah diakses dan dianalisis menggunakan pustaka Python.

b. Preprocessing Data

Pada langkah ini, teks dibersihkan dari karakter, simbol, dan angka yang tidak penting. Selanjutnya, proses normalisasi teks dilakukan, termasuk tokenisasi, mengubah seluruh huruf menjadi *lowercase*, dan penghapusan *stopwords* dengan menggunakan pustaka *nlTK*. *Stopwords* merupakan istilah yang biasa keluar dalam bahasa tertulis yang tidak memegang arti penting [15]. Setelah data dibersihkan, teknik *TF-IDF* (*Term Frequency-Inverse Document Frequency*) dimanfaatkan guna mengekstrak fitur. Frekuensi suatu istilah dalam data ditentukan dengan menggunakan model frekuensi kata dan frekuensi dokumen terbalik (*IDF*), di mana *TF* mengukur berapa kali sebuah istilah muncul [16].

Untuk memberikan gambaran yang lebih jelas mengenai proses pembersihan data, berikut ditampilkan Gambar 1, cuplikan data sebelum (kolom komentar) dan sesudah *preprocessing* (kolom *cleaned*).

	komentar	cleaned	label	sentimen
0	sangat manjur	sangat manjur	2	positif
1	mantap.. semuanya mudah	mantap semuanya mudah	2	positif
2	Ai nya bagus ,cuman kirim file yg mau di jawab...	ai nya bagus cuman kirim file yg mau jawab lan...	2	positif
3	membantu sekali	membantu sekali	2	positif
4	cukup memuaskan	cukup memuaskan	2	positif
...	...	...	...	...
4995	belum. bisa mengerjakan soal pilihan ganda	mengerjakan soal pilihan ganda	0	negatif
4996	Keren	keren	2	positif
4997	sudah bagus. Berharap ke depannya bisa di kemb...	bagus berharap depannya kembangkan lebih baik	2	positif
4998	cepat merespon dan membantu dalam memahami	cepat merespon membantu memahami	2	positif
4999	Apk sudah bagus, cuma ada limit nya perhari ja...	apk bagus cuma limit nya perhari jadi digunain...	2	positif

5000 rows × 4 columns

Gambar 1. Cuplikan Data Sebelum dan Sesudah *Preprocessing*

c. Pelatihan Model *Machine Learning*

Machine learning merupakan bagian dari kecerdasan buatan yang menjadikan sistem guna memahami dan meningkatkan kemampuannya sendiri tanpa memerlukan pemrograman yang komprehensif untuk setiap tugas [17]. Pada tahap ini, dataset yang telah diproses kemudian dibagi menjadi dua bagian menggunakan metode *train_test_split* dari pustaka *scikit-learn*, yaitu 80% data digunakan untuk pelatihan (*training*) dan 20% untuk pengujian (*testing*).

Pemilihan rasio 80:20 antara data pelatihan dan data pengujian merupakan praktik umum dalam penelitian *machine learning*. Rasio ini dipilih untuk memastikan bahwa model memperoleh data yang cukup untuk belajar dan mengenali pola-pola yang ada dalam data, namun tetap menyisakan proporsi data yang memadai untuk menguji performa model secara objektif. Dengan 80% data untuk pelatihan, model dapat membangun generalisasi yang baik, sementara 20% data pengujian cukup representatif untuk mengevaluasi kinerja model terhadap data yang belum pernah dilihat sebelumnya. Selain itu, rasio ini juga membantu mengurangi risiko *overfitting*, yaitu kondisi di mana model terlalu menyesuaikan diri dengan data pelatihan sehingga gagal mengeneralisasi pada data baru.

Selanjutnya, model kemudian dievaluasi memanfaatkan dua metode, Multinomial *Naive Bayes* dan *Support Vector Machine (SVM)*. Masing-masing model dibangun dan dioptimalkan untuk mempelajari pola sentimen berdasarkan hasil ekstraksi fitur *TF-IDF*.

d. Pengujian Model

Pengujian model dilakukan guna mengukur sejauh mana metode dapat mengelompokkan sentimen, memanfaatkan matriks evaluasi misalnya presisi, *recall*, akurasi serta matriks *f1-score*. Hasilnya diilustrasikan melalui grafik frekuensi kata, *confusion matrix*, dan *WordCloud*. *Confusion matrix* merupakan bagan untuk menampilkan total dokumen uji yang dikategorikan benar dan salah, serta meringkas hasil prediksi metode pengelompokan [18]. *Confusion matrix* dimanfaatkan guna menilai distribusi prediksi model terhadap kelas sentimen. Berdasarkan *confusion matrix*, *precision*, *recall*, *accuracy* serta matriks *f1-score* dihitung menggunakan rumus:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

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

Selain itu, *WordCloud* dimanfaatkan guna mengilustrasikan sejauh mana kata sering muncul dalam ulasan, dengan ukuran kata yang mencerminkan frekuensi kemunculannya. *WordCloud* memudahkan identifikasi kata-kata kunci yang paling sering digunakan dalam kategori sentimen yang berbeda (positif, netral, dan negatif) [19]. Analisis lebih rinci dilakukan dengan grafik batang yang menampilkan 25 kata teratas dalam setiap kategori sentimen, untuk membantu mengidentifikasi pola kata yang dominan.

3. HASIL DAN ANALISIS

3.1. Pengumpulan Data

Dalam tahap awal ini, informasi dihimpun dengan menganalisis ulasan aplikasi Claude di Google Playstore, terutama berfokus pada ulasan dalam bahasa Indonesia. Data diperoleh melalui metode *web scraping* menggunakan tools berbasis Python, yang memungkinkan ekstraksi otomatis dari halaman aplikasi.

Sebanyak 5000 ulasan berhasil dikumpulkan dalam rentang waktu 12 November 2024 hingga 24 April 2025. Setiap ulasan mencakup teks, rating, dan tanggal, namun analisis difokuskan pada isi teks untuk keperluan klasifikasi sentimen. Data disimpan dalam format CSV agar mudah diproses lebih lanjut. Gambar 2 memperlihatkan distribusi data berdasarkan kategori sentimen.

	username	tanggal	rating	komentar
0	Andika Ramadan	2025-04-24 15:19:17	5	sangat manjur
1	Atack ransomware	2025-04-22 11:59:26	5	mantap.. semuanya mudah
2	Shulton arkaan wibowo	2025-04-22 01:56:07	5	Ai nya bagus ,cuman kirim file yg mau di jawab...
3	2021SA009_Andriyanto	2025-04-19 23:32:55	5	membantu sekali
4	Hsnl Khtmh	2025-04-19 17:49:50	5	cukup memuaskan
...	...	...	...	...
4995	Piza Imoet	2024-11-13 11:24:41	1	belum. bisa mengerjakan soal pilihan ganda
4996	Afif Fatchur Rozzy	2024-11-13 10:39:55	5	Keren
4997	Aldi R.	2024-11-13 00:35:00	5	sudah bagus. Berharap ke depannya bisa di kemb...
4998	Lukman Nulhakim	2024-11-12 12:39:45	5	cepat merespon dan membantu dalam memahami
4999	Andika M.S.	2024-11-12 12:14:06	5	Apk sudah bagus, cuma ada limit nya perhari ja...

5000 rows × 4 columns

Gambar 2. Hasil Pengumpulan Data Sentimen Aplikasi Calude AI

3.2. Preprocessing Data

Sebelum proses *preprocessing*, teks menjalani proses pembersihan untuk menghilangkan elemen-elemen asing seperti emoji, karakter khusus, angka, dan spasi yang tidak perlu setelah data dikumpulkan. Untuk mencegah duplikasi dari kapitalisasi yang berbeda-beda, seluruh teks diubah menjadi huruf kecil, dan kata henti dihilangkan dengan menggunakan pustaka Sastrawi untuk berkonsentrasi pada istilah-istilah kunci. Tokenisasi dilakukan secara implisit saat teks diubah menjadi representasi fitur vektor. Penelitian ini tidak menerapkan *stemming* atau lemmatisasi secara eksplisit. Selanjutnya, ulasan diberi label sentimen secara otomatis berdasarkan rating, 4–5 (positif/label 2), 3 (netral/label 1), dan 1–2 (negatif/label 0). Contoh hasil *preprocessing* dan pelabelan ditunjukkan pada Gambar 3.

	cleaned	label	sentimen
0	sangat manjur	2	positif
1	mantap semuanya mudah	2	positif
2	ai nya bagus cuman kirim file yg mau jawab lan...	2	positif
3	membantu sekali	2	positif
4	cukup memuaskan	2	positif
...	...	...	...
4995	mengerjakan soal pilihan ganda	0	negatif
4996	keren	2	positif
4997	bagus berharap depannya kembangkan lebih baik	2	positif
4998	cepat merespon membantu memahami	2	positif
4999	apk bagus cuma limit nya perhari jadi digunain...	2	positif

5000 rows × 3 columns

Gambar 3. Hasil *Preprocessing* Data

3.3. TF-IDF

Setelah tahap *preprocessing*, dokumen teks diubah menjadi format numerik yang memungkinkan model pembelajaran mesin berfungsi secara efektif. Metode ekstraksi fitur yang digunakan adalah *Term Frequency-Inverse Document Frequency (TF-IDF)*. *TF-IDF* merupakan metode komputasi yang dimanfaatkan guna mengukur level kepentingan suatu kata atau istilah dalam data tertentu dalam kaitannya pada keseluruhan koleksi dokumen (korpus) [20]. *TF* mengukur frekuensi sebuah kata dalam dokumen tertentu, sedangkan *IDF* menyampaikan nilai lebih terhadap kata yang kurang umum di semua dokumen, menyoroti pentingnya istilah yang lebih bermakna.

Setiap ulasan kemudian direpresentasikan sebagai vektor angka berdasarkan skor *TF-IDF*. Hasilnya disusun dalam bentuk matriks dengan dimensi besar, di mana setiap baris mewakili satu ulasan dan setiap kolom adalah kata unik. Gambar 4 memperlihatkan cuplikan hasil representasi *TF-IDF* pada sebagian data.

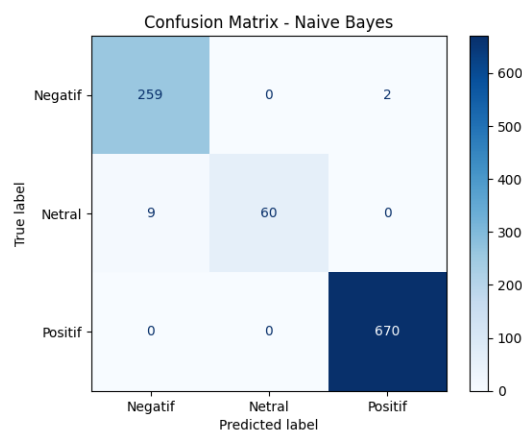
	aamiin	abis	agus	ai	aja	akal	akhirnya	akibatnya	akses	aktivitas
0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.113037	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...
4995	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0
4996	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0
4997	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0
4998	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0
4999	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0

5000 rows × 10 columns

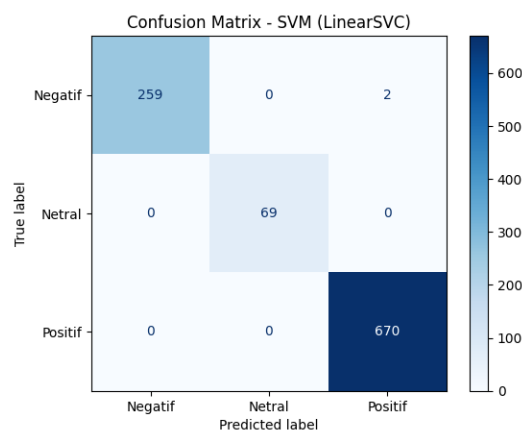
Gambar 4. Hasil Representasi *TF-IDF*

3.4. Evaluasi Model Menggunakan Confusion Matrix

Performa kedua model tersebut diukur dengan menggunakan *matrix confusion* untuk melakukan evaluasi. Gambar 5 menunjukkan *confusion matrix* dari metode Multinomial Naive Bayes, sedangkan Gambar 6 menunjukkan *confusion matrix* dari model SVM Linear.



Gambar 5. Confusion Matrix Metode Multinomial Naive Bayes



Gambar 6. Confusion Matrix Model SVM Linear

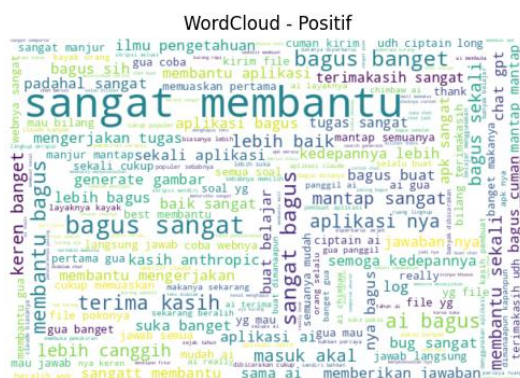
Confusion matrix menggambarkan bagaimana prediksi model didistribusikan di berbagai kategori sentimen. Model Multinomial *Naive Bayes* menunjukkan akurasi tinggi meskipun ada sedikit kesalahan pada sentimen netral, sedangkan *SVM Linear* mengklasifikasikan seluruh data dengan sempurna.

3.5. Analisis Frekuensi Kata dan Perbandingan Metode Klasifikasi

Setelah klasifikasi sentimen selesai, langkah selanjutnya adalah memeriksa hasil distribusi kata-kata pada masing-masing kategori sentimen dan membandingkan kinerja dua metode klasifikasi, *Support Vector Machine (SVM)* serta *Multinomial Naive Bayes*.

a. Visualisasi WordCloud

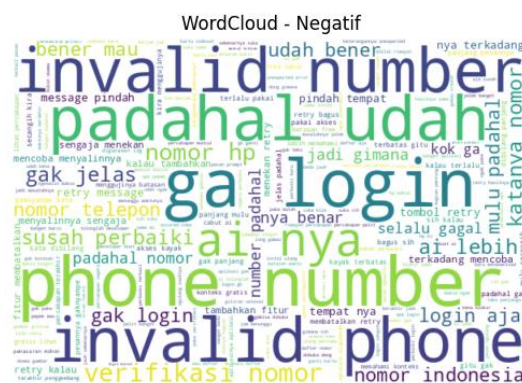
Pembuatan visualisasi *WordCloud* digunakan untuk mendapatkan pemahaman yang lebih baik mengenai istilah yang kerap keluar pada tiap jenis sentimen. Ukuran kata dalam gambar menunjukkan frekuensi kata tersebut. Visualisasi *WordCloud* untuk masing-masing kategori ditunjukkan dalam Gambar 7 sampai Gambar 9.



Gambar 7. WordCloud Sentimen Positif



Gambar 8. *WordCloud* Sentimen Netral

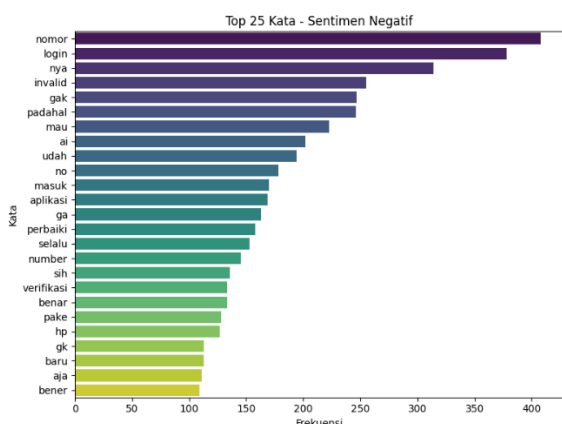


Gambar 9. WordCloud Sentimen Negatif

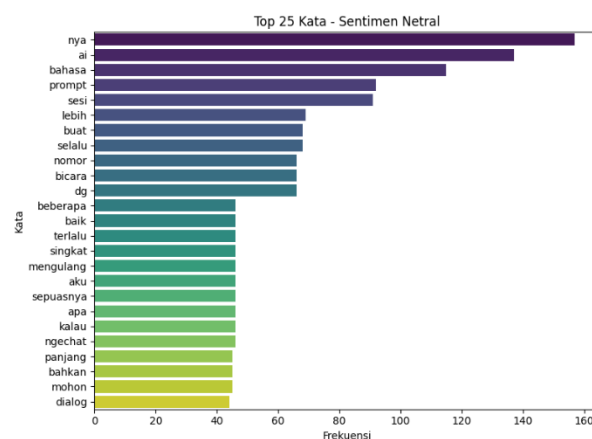
Dari visualisasi *WordCloud*, terlihat jelas bahwa kata-kata tertentu sering muncul di setiap kategori sentimen. Hal ini mengindikasikan istilah-istilah penting yang umumnya terkait dengan perasaan atau pandangan pengguna tentang aplikasi.

b. Analisis Frekuensi Kata

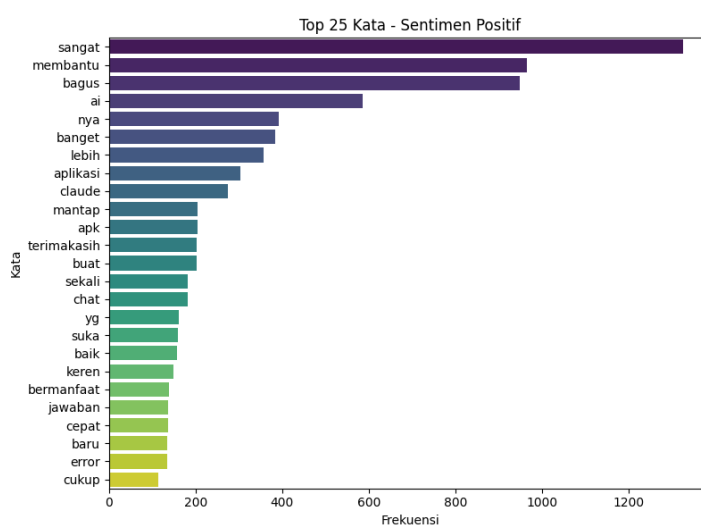
Di samping *WordCloud*, analisis tambahan yang dilakukan adalah menentukan frekuensi kata-kata dalam setiap sentimen. Diagram batang menggambarkan 25 kata yang paling sering muncul untuk setiap kategori, yang menyoroti keunggulannya. Visualisasi frekuensi kata ditunjukkan pada Gambar 10 hingga Gambar 12.



Gambar 10. Frekuensi Kata Sentimen Negatif



Gambar 11. Frekuensi Kata Sentimen Netral



Gambar 12. Frekuensi Kata Sentimen Positif

Ada perbedaan kata-kata yang dominan dalam kategori sentimen, seperti yang ditunjukkan oleh analisis ini. Misalnya, kata-kata tertentu lebih sering ditemukan dalam ulasan negatif dibandingkan dengan ulasan positif atau netral.

c. Perbandingan Metode Klasifikasi

Tabel 1. Perbandingan Performansi Model Multinomial *Naive Bayes* dan *SVM Linear*

Metode	Accuracy	Precision	Recall	F1-Score
Multinomial <i>Naive Bayes</i>	98,9%	98,7814%	95,3967%	96,9316%
<i>SVM Linear</i>	99,8%	99,9008%	99,7446%	99,8221%

Hasil yang ditunjukkan dalam Tabel 1, menunjukkan bahwa kedua teknik klasifikasi berkinerja sangat baik dalam menganalisis data ulasan aplikasi Claude. Berdasarkan hasil yang diperoleh, model *SVM Linear* mengungguli model Multinomial *Naive Bayes* dalam klasifikasi sentimen. Model *SVM Linear* mencapai *accuracy* 99.8%, *precision* 99.9008%, *recall* 99.7446%, dan *f1-score* 99.8221%. Nilai-nilai ini menunjukkan bahwa *SVM Linear* mampu mengklasifikasikan ulasan dengan tingkat kesalahan yang sangat rendah dan koherensi perkiraan yang tinggi di seluruh kategori sentimen. Sementara itu, model Multinomial *Naive Bayes* juga menunjukkan performa yang cukup baik dengan *accuracy* sebesar 98,9%, *precision* sebesar 98,7814%, *recall* sebesar 95,3967%, dan *f1-score* sebesar 96,9316%. Walaupun akurasinya sedikit lebih rendah dibandingkan *SVM*, model ini tetap memberikan hasil prediksi yang kuat dan efisien.

Secara umum, nilai *accuracy* menunjukkan persentase prediksi yang benar dari seluruh data yang diuji, sedangkan *precision* menggambarkan seberapa banyak prediksi positif yang benar-benar relevan atau tepat sasaran. *Recall* mengukur seberapa banyak data positif yang berhasil diidentifikasi dengan benar oleh model, dan *f1-score* merupakan rata-rata harmonis antara *precision* dan *recall* yang memberikan gambaran keseimbangan antara keduanya, terutama penting ketika distribusi data tidak seimbang.

Secara keseluruhan, kedua metode mampu mengklasifikasikan sentimen ulasan dengan tingkat akurasi tinggi. Namun, *SVM Linear* lebih unggul dalam hal kestabilan dan konsistensi performa dibandingkan Multinomial *Naive Bayes*. Hal tersebut memperlihatkan bahwasanya *SVM* lebih unggul untuk mengidentifikasi model yang signifikan dalam data ulasan aplikasi, sehingga menghasilkan prediksi yang lebih tepat. Keunggulan *SVM Linear* ini dapat dijelaskan karena algoritma *SVM* sangat efektif dalam menangani data berdimensi tinggi seperti teks, mampu menemukan *hyperplane* optimal yang memisahkan kelas dengan margin terbesar, serta lebih tahan terhadap *overfitting* pada dataset yang besar dan kompleks, sehingga prediksi yang dihasilkan menjadi lebih stabil dan konsisten dibandingkan dengan metode probabilistik seperti *Naive Bayes*.

4. KESIMPULAN

Analisis sentimen dilakukan terhadap ulasan aplikasi Claude dengan memanfaatkan algoritma *Support Vector Machine (SVM)* dan Multinomial *Naive Bayes*. Sebelum melatih model, *preprocessing* dilakukan melalui pembersihan teks dan ekstraksi fitur melalui metode *TF-IDF* untuk meningkatkan kualitas data. Kedua model tersebut dapat mengelompokkan sentimen dalam tingkat akurasi yang tinggi. Model *SVM Linear* mencatatkan *accuracy* 99,8%, sedangkan Multinomial *Naive Bayes* mencatatkan *accuracy* 98,9% setelah melakukan evaluasi menggunakan *matrix confusion*. Selain *accuracy* yang menunjukkan persentase prediksi yang benar dari seluruh data, *precision* pada model *SVM Linear* mencapai 99,9008% dan pada Multinomial *Naive Bayes* sebesar 98,7814%. Nilai *precision* ini menunjukkan proporsi prediksi sentimen yang benar-benar tepat dari seluruh prediksi sentimen yang dihasilkan model. *Recall* pada *SVM Linear* sebesar 99,7446% dan pada Multinomial *Naive Bayes* sebesar 95,3967%, yang berarti model mampu menemukan hampir seluruh data sentimen yang benar dalam dataset. Sementara itu, *f1-score* pada *SVM Linear* sebesar 99,8221% dan pada Multinomial *Naive Bayes* sebesar 96,9316%, yang merupakan rata-rata harmonis antara *precision* dan *recall*, sehingga memberikan gambaran seberapa seimbang model dalam menghasilkan prediksi yang tepat dan lengkap, terutama ketika distribusi data tidak seimbang.

Untuk kategori sentimen tertentu, analisis visual yang digunakan melalui *WordCloud* dan grafik frekuensi kata meningkatkan pemahaman kita tentang kata kunci dominan. Secara keseluruhan, kedua metode aktual dalam pengelompokan sentimen ulasan pengguna. *SVM Linear* menunjukkan keunggulan dalam hal *accuracy*, konsistensi prediksi, dan stabilitas performa. Keunggulan *SVM Linear* ini dapat dijelaskan karena algoritma *SVM* sangat efektif dalam menangani data berdimensi tinggi seperti teks, mampu menemukan *hyperplane* optimal yang memisahkan kelas dengan margin terbesar, serta lebih tahan terhadap *overfitting* pada dataset yang besar dan kompleks. Dengan demikian, prediksi yang dihasilkan menjadi lebih stabil dan konsisten dibandingkan dengan metode probabilistik seperti *Naive Bayes*. Hasil ini dapat menjadi dasar bagi pengembang untuk menilai dan meningkatkan layanan aplikasi berbasis AI.

REFERENSI

- [1] U. Prajogo, B. Bunyamin, S. Munfaqiroh, L. Lindananty, L. Andiani, and S. Sunarto, "Pelatihan Pemanfaatan Artificial Intelligence (AI) dalam Pembelajaran pada Guru Pondok Pesantren El Jasmeen," *Community Reinforcement and Development Journal*, vol. 3, no. 1, pp. 11–16, 2024, doi: 10.35584/reinforcementanddevelopmentjournal.v3i1.166.
- [2] T. Tinaliah and T. Elizabeth, "Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode Support Vector Machine," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 9, no. 4, pp. 3436–3442, 2022, doi: 10.35957/jatisi.v9i4.3586.
- [3] A. Suganda, "Memilih Ai Yang Tepat Untuk Guru: Perbandingan Fitur Gemini, Chatgpt, Dan Claude Ai," *Jurnal Inovasi Teknologi dan Edukasi Teknik*, vol. 3, no. 11, p. 2, 2023, doi: 10.17977/um068.v3.i11.2023.2.

- [4] D. Angraina and A. Putri, "Analisis Sentimen Pengguna Aplikasi Google Meet Menggunakan Algoritma Support Vector Machine," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 3, no. 3, pp. 472–478, 2022, doi: 10.37859/coscitech.v3i3.4260.
- [5] Fitri Wulandari, Elin Haerani, Muhammad Fikry, and Elvia Budianita, "Analisis sentimen larangan penggunaan obat sirup menggunakan algoritma naive bayes classifier," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 4, no. 1, pp. 88–96, 2023, doi: 10.37859/coscitech.v4i1.4781.
- [6] F. Amaliah and I. K. Dwi Nuryana, "Perbandingan Akurasi Metode Lexicon Based Dan Naive Bayes Classifier Pada Analisis Sentimen Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter," *Journal of Informatics and Computer Science (JINACS)*, vol. 3, no. 03, pp. 384–393, 2022, doi: 10.26740/jinacs.v3n03.p384-393.
- [7] Alfandi Safira and F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," *ZONasi: Jurnal Sistem Informasi*, vol. 5, no. 1, pp. 59–70, 2023, doi: 10.31849/zn.v5i1.12856.
- [8] U. Kusnia, F. Kurniawan, and S. Artikel, "Analisis Sentimen Review Aplikasi Media Berita Online Pada Google Play menggunakan Metode Algoritma Support Vector Machines (SVM) Dan Naive Bayes INFO ARTIKEL ABSTRAK," *Jurnal Keilmuan dan Aplikasi Teknik Informatika*, vol. 14(1), no. 36, pp. 24–25, 2022.
- [9] F. M. Malik *et al.*, "Komparasi Sentiment Analysis Pada Review Aplikasi Tokopedia Dan Shopee Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine Dengan Metode Tf- Idf," vol. 11, no. 6, pp. 5872–5878, 2024.
- [10] I. Loelianto, Moh. S. S. Thayf, and H. Angriani, "Implementasi Teori Naive Bayes Dalam Klasifikasi Calon Mahasiswa Baru Stmik Kharisma Makassar," *SINTECH (Science and Information Technology) Journal*, vol. 3, no. 2, pp. 110–117, 2020, doi: 10.31598/sintechjournal.v3i2.651.
- [11] Y. Ayuningtyas and I. Made Suartana, "Klasifikasi Penyakit Stroke Menggunakan Support Vector Machine (SVM) dan Particle Swarm Optimization (PSO)," *JINACS: (Journal of Informatics and Computer Science)*, vol. 04, no. 04, pp. 452–457, 2023.
- [12] M. R. Saputra, "Analisis sentimen twitter terhadap konflik di papua menggunakan perbandingan naive bayes dan svm," vol. 10, no. 2, pp. 1197–1208, 2025.
- [13] L. C. O. Id, "MENGGUNAKAN MULTIPLE LINEAR REGRESSION (STUDI KASUS: KABUPATEN WEBSITE," vol. 13, no. 2, 2025.
- [14] M. Ulfi, T. Savira, and P. Ayu, "Implementasi Deep Learning dengan Convolutional Neural Network untuk Tingkat Akurasi Citra Image Hama Sawi Hijau Menggunakan Google Colab," vol. 10, no. 2, pp. 116–125, 2024.
- [15] F. F. Irfani, "Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine," *JBMI (Jurnal Bisnis, Manajemen, dan Informatika)*, vol. 16, no. 3, pp. 258–266, 2020, doi: 10.26487/jbmi.v16i3.8607.
- [16] J. Algoritme *et al.*, "Analisis Sentimen Komentar Pada Saluran Youtube Beauty Vlogger Berbahasa Indonesia Menggunakan Metode Support Vector Machine," vol. 5, no. 2, pp. 164–175, 2025, doi: 10.35957/algoritme.v5i2.9692.
- [17] A. C. T. Angel, V. H. Pranatawijaya, and W. Widiatry, "Analisis Sentimen dan Emosi dari Ulasan Google Maps Untuk Layanan Rumah Sakit di Palangka Raya Menggunakan Machine Learning," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 4, no. 1, pp. 35–49, Jun. 2024, doi: 10.24002/konstelasi.v4i1.8924.
- [18] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.
- [19] S. A. S. Mola, D. L. B. Baun, I. O. Nunes, and M. M. A. R. Sani, "ANALISIS SENTIMEN APLIKASI HALO BCA DI GOOGLE PLAY STORE MENGGUNAKAN METODE NAIVE BAYES, SUPPORT VECTOR MACHINE DAN RANDOM FOREST," *HOAQ (High Education of Organization Archive Quality): Jurnal Teknologi Informasi*, vol. 15, no. 2, pp. 69–79, Dec. 2024, doi: 10.52972/hoaq.vol15no2.p69-79.
- [20] M. N. Saadah, R. W. Atmagi, D. S. Rahayu, and A. Z. Arifin, "SISTEM TEMU KEMBALI DOKUMEN TEKS DENGAN PEMBOBOTAN TF-IDF DAN LCS," *JUTI: Jurnal Ilmiah Teknologi Informasi*, pp. 19–22, Jan. 2013, doi: 10.12962/j24068535.v11i1.a16.