

Analisis Sentimen Ulasan Pengguna Aplikasi Famiapps Menggunakan *Naïve Bayes*, SVM, dan Random Forest

Jullius Calvin¹, Simon Prananta Barus²

Universitas Matana

jullius.celvin@gmail.com

Abstrak: Aplikasi FamiApps yang tersedia di *Google Play Store* telah digunakan oleh banyak pengguna, namun belum ada penelitian komprehensif untuk mengevaluasi pola emosional mereka. Penelitian ini bertujuan untuk mengeksplorasi sentimen pengguna dan memberikan wawasan mengenai persepsi mereka terhadap aplikasi. Tiga metode *machine learning* yang dibandingkan adalah *Random Forest*, *Naïve Bayes*, dan *Support Vector Machine (SVM)*. Dataset terdiri dari 1.079 ulasan pengguna yang dikategorikan menjadi sentimen positif, negatif, dan netral. Proses analisis melibatkan preprocessing data seperti pembersihan, tokenisasi, *stemming*, dan pembobotan TF-IDF untuk mengubah teks menjadi format numerik. Evaluasi model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa *SVM* memiliki kinerja terbaik dengan *accuracy* tertinggi sebesar 98,58% pada skenario pemisahan data 80:20 dan 98,24% pada skenario 70:30. Keunggulan *SVM* terletak pada kemampuannya menangani data berdimensi tinggi dan menemukan hiperplane optimal untuk klasifikasi lebih akurat. *Random Forest* dan *Naïve Bayes* cenderung lebih rentan terhadap *noise* dan fitur yang kurang relevan. Dengan demikian, penelitian ini menyimpulkan bahwa *SVM* adalah algoritma paling andal dalam klasifikasi sentimen ulasan pengguna FamiApps. Temuan ini dapat membantu pengembang dalam meningkatkan kualitas aplikasi secara lebih tepat sasaran.

Kata Kunci: Analisis Sentimen, *Machine Learning*, *Naïve Bayes*, *Random Forest*, *Support Vector Machine*.

Abstract: The FamiApps application, available on the *Google Play Store*, has been used by many users, yet no comprehensive research has been conducted to evaluate their emotional patterns. This study aims to explore user sentiment and provide insights into their perception of the application. Three machine learning methods were compared: *Random Forest*, *Naïve Bayes*, and *Support Vector Machine (SVM)*. The dataset consists of 1,079 user reviews categorized into positive, negative, and neutral sentiments. The analysis process involves data preprocessing, including cleaning, tokenization, *stemming*, and TF-IDF weighting to convert text into numerical format. Model evaluation was conducted using *accuracy*, *precision*, *recall*, and *F1-score* metrics. The results show that *SVM* performs the best, achieving the highest *accuracy* of 98.58% in the 80:20 data split scenario and 98.24% in the 70:30 scenario. *SVM*'s advantage lies in its ability to handle high-dimensional data and find the optimal hyperplane for more accurate classification. In contrast, *Random Forest* and *Naïve Bayes* tend to be more susceptible to noise and less relevant features. Thus, this study concludes that *SVM* is the most reliable algorithm for classifying user sentiment in FamiApps reviews. These findings can help developers improve application quality in a more targeted manner.

Keywords: Sentiment Analysis, *Machine Learning*, *Naïve Bayes*, *Random Forest*, *Support Vector Machine*.

1. Pendahuluan

Kemajuan teknologi yang pesat, terutama melalui kehadiran Internet, telah mengubah cara berpikir dan gaya hidup masyarakat, mulai dari cara berkomunikasi, interaksi di media sosial, hingga aktivitas belanja online (Putra dkk., 2021). Di era digital ini, transformasi tersebut menghasilkan big data atau kumpulan data yang besar dan kompleks yang memungkinkan analisis mendalam seperti analisis sentimen. Teknik ini menjadi sangat relevan untuk mengungkap persepsi dan preferensi pengguna, yang menjadi fokus utama penelitian ini (Sulistyawati & Munawir, 2024).

Salah satu platform yang mengumpulkan big data secara signifikan adalah Google Play Store, yang menyediakan beragam aplikasi, game, buku, dan konten digital lainnya. Review dan rating dari pengguna tidak hanya menilai kualitas aplikasi, tetapi juga memberikan umpan balik penting bagi pengembang. Dalam konteks penelitian ini, ulasan yang terdapat di Google Play Store digunakan untuk menganalisis sentimen pengguna, dengan tujuan mengaitkannya langsung pada evaluasi aplikasi FamiApps (Herlinawati dkk., 2020).

FamiApps merupakan aplikasi ritel yang dikembangkan oleh PT. Fajar Mitra Indah sebagai solusi atas antrean panjang di toko FamilyMart dan sebagai dukungan terhadap protokol kesehatan selama pandemi *COVID-19*. Melalui FamiApps, pengguna dapat memesan produk secara digital dan mengambilnya tanpa harus mengantri (Firmansyah dkk., 2023). Pentingnya analisis sentimen terhadap aplikasi ini muncul dari minimnya studi yang secara khusus mengeksplorasi persepsi pengguna, padahal ulasan dari platform seperti Google Play Store menawarkan wawasan berharga untuk perbaikan dan inovasi aplikasi (Tarwoto dkk., 2025).

Berbagai penelitian telah mengeksplorasi analisis sentimen menggunakan pendekatan algoritma pembelajaran mesin, seperti penerapan *Naïve Bayes* dalam mengklasifikasikan tingkat kepopuleran *e-commerce* di Indonesia (Widagdo dkk., 2020). Namun, studi mendalam mengenai analisis sentimen pada aplikasi ritel, khususnya FamiApps, masih sangat terbatas. Penelitian ini berupaya mengisi celah tersebut dengan mengevaluasi dan membandingkan efektivitas beberapa model pembelajaran mesin dalam memproses ulasan pengguna, sehingga memberikan kontribusi yang berbeda dari studi-studi sebelumnya.

Kebaruan dari penelitian ini terletak pada fokus khusus terhadap aplikasi FamiApps, yang hingga saat ini belum banyak dieksplorasi dalam studi analisis sentimen, meskipun memiliki peran penting dalam transformasi digital ritel di Indonesia. Selain itu, penelitian ini menggunakan dataset ulasan yang diambil langsung dari Google Play Store pada periode 2022–2024, sehingga mencerminkan persepsi terkini pengguna. Berbeda dari sebagian besar studi terdahulu yang hanya menggunakan satu metode klasifikasi dalam analisis sentimen, penelitian ini membandingkan kinerja tiga algoritma populer yaitu, *Random Forest*, *Naïve Bayes*, *Support Vector Machine* pada dataset ulasan FamiApps, sehingga memberikan wawasan komparatif yang lebih komprehensif. Hingga saat ini, belum ada studi ilmiah yang menguji ketiga model tersebut secara bersama-sama pada dataset FamiApps; oleh karena itu, penelitian ini diharapkan dapat memberikan sudut pandang baru dalam pemilihan algoritma terbaik untuk analisis sentimen di industri ritel digital Indonesia.

Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja tiga model pembelajaran mesin *Random Forest*, *Naïve Bayes*, dan *Support Vector Machine*

(*SVM*) dalam mengklasifikasikan sentimen pengguna terhadap FamiApps (Safrudin dkk., 2024). Ketiga model ini dipilih karena mewakili pendekatan yang berbeda: *Naïve Bayes* dikenal cepat dan sederhana untuk klasifikasi teks berbahasa Indonesia, *SVM* unggul dalam menangani data berdimensi tinggi dan memiliki kemampuan margin yang baik untuk memisahkan kelas sentimen, serta *Random Forest* mampu menangani fitur yang kompleks dan cenderung lebih tahan terhadap *overfitting* pada *dataset* berukuran sedang. Melalui pendekatan ini, diharapkan dapat diidentifikasi model yang paling efektif dan diperoleh rekomendasi strategis bagi pengembang untuk meningkatkan aplikasi sesuai dengan kebutuhan serta persepsi pengguna di Indonesia.

Hasil penelitian ini diharapkan dapat menyajikan rekomendasi mengenai model pembelajaran mesin yang paling efektif untuk analisis sentimen pengguna aplikasi ritel seperti FamiApps. Selain itu, penelitian ini juga bertujuan untuk membantu pengembang aplikasi dalam memahami pola sentimen pengguna sehingga mereka dapat merancang strategi pengembangan yang lebih sesuai dengan kebutuhan masyarakat Indonesia. Dengan demikian, penelitian ini tidak hanya memiliki kontribusi akademik, tetapi juga memberikan manfaat nyata bagi industri ritel digital di Indonesia.

2. Metode Penelitian

2.1 Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen yang didukung oleh teknik data mining dan algoritma pembelajaran mesin. Secara khusus, penelitian ini mengadopsi metode eksperimental untuk membandingkan kinerja tiga model klasifikasi (*Random Forest*, *Naïve Bayes*, dan *Support Vector Machine*) dalam mengkategorikan sentimen ulasan pengguna (R. N. Rahman dkk., 2025).

2.2 Tempat dan Waktu

Penelitian dilakukan secara daring dengan memanfaatkan data yang dikumpulkan dari Google Play Store. Tempat pengambilan data berasal dari ulasan pengguna aplikasi FamiApps oleh FamilyMart. Data yang dikumpulkan adalah ulasan dan review dari periode 2022-2024 untuk mengakomodasi perubahan dan tren ulasan dalam rentang waktu yang cukup representatif.

2.3 Subjek Penelitian

Subjek penelitian adalah ulasan atau review yang diberikan oleh pengguna aplikasi FamiApps. Dataset yang digunakan terdiri dari 1.079 ulasan yang dikumpulkan secara otomatis dari platform Google Play Store, yang kemudian diklasifikasikan menjadi tiga kategori sentimen: positif, negatif, dan netral.

2.4 Teknik Pengumpulan Data

Data dikumpulkan melalui teknik *crawling* menggunakan Google Play Scraper, sebuah pustaka Python yang dirancang untuk mengambil informasi ulasan secara sistematis dari Google Play Store. Prosedur pengumpulan data meliputi identifikasi aplikasi FamiApps, ekstraksi ulasan pengguna, serta penyimpanan data dalam format yang telah ditentukan untuk kemudian diproses lebih lanjut (Rahmadani dkk., 2024).

2.5 Teknik Analisis Data

Tahapan analisis data diawali dengan proses *preprocessing* menggunakan metode *Natural Language Processing (NLP)* untuk mengubah teks ulasan yang tidak terstruktur menjadi data yang siap dianalisis (Urva dkk., 2023). Langkah-langkah utama meliputi:

1. *Cleaning*: Menghapus elemen-elemen seperti mention, hashtag, karakter khusus, tanda baca, spasi berlebih, dan URL (Ilyas & Chu, 2019).
2. *Case-Folding*: Mengubah teks menjadi huruf kecil.
3. *Tokenizing*: Membagi kalimat menjadi token atau kata (O. H. Rahman dkk., 2021).
4. *Filtering*: Menggunakan stopwords removal untuk menyaring kata-kata tidak relevan (Kartika & Gondohanindijo, 2020).
5. *Stemming*: Untuk mendapatkan bentuk dasar kata (Dinata dkk., 2020).

Setelah *preprocessing*, data diubah menjadi representasi vektor dengan metode *TF-IDF* (Sinulingga & Sitorus, 2024). Vektor tersebut kemudian dianalisis menggunakan tiga model pembelajaran mesin (*Random Forest*, *Naïve Bayes*, dan *SVM*) yang hasil kinerjanya dievaluasi menggunakan *confusion matrix*. Metrik evaluasi yang dipakai adalah *accuracy*, *precision*, *recall*, dan *F1-score*.

2.6 Siklus Penelitian

Siklus penelitian terdiri dari beberapa tahapan iteratif, yakni:

- Pengumpulan Data: Melakukan *crawling* dan ekstraksi data ulasan dari Google Play Store.
- Preprocessing Data: Melakukan *cleaning*, *case folding*, *tokenizing*, *filtering*, dan *stemming*.
- Ekstraksi Fitur: Mengkonversi teks menjadi vektor dengan metode *TF-IDF*.
- Pembangunan Model: Mengimplementasikan dan melatih model *Random Forest*, *Naïve Bayes*, serta *Support Vector Machine*.
- Evaluasi Model: Menghitung metrik evaluasi (*accuracy*, *precision*, *recall*, *F1-score*) dengan menggunakan *confusion matrix*.
- Analisis Hasil: Membandingkan kinerja ketiga model untuk menentukan model yang paling optimal dalam mengklasifikasikan sentimen ulasan pengguna.

3. Hasil dan Pembahasan

a. Pengumpulan Data

Penelitian ini mencakup 1.079 data ulasan pengguna dari Google Play Store, dimana 859 ulasan memiliki sentimen positif, 190 ulasan memiliki sentimen negatif, dan 30 ulasan netral. Data diperoleh melalui teknik scraping dengan menggunakan pustaka Python bernama Google Play Scraper. Informasi yang sudah diperoleh diubah ke dalam format “.csv”.

b. Preprocessing Data

Setelah data berhasil di *scraping* dari *Google Play Store*, tahapan selanjutnya adalah melakukan *preprocessing data* untuk menyeleksi fitur ataupun atribut yang kurang bermakna, akan ada 5 tahap *preprocessing* yaitu, *cleaning*, *case folding*, *tokenizing*, *filtering*, dan *stemming*. Berikut adalah contoh hasil Preprocessing data pada Tabel 2.

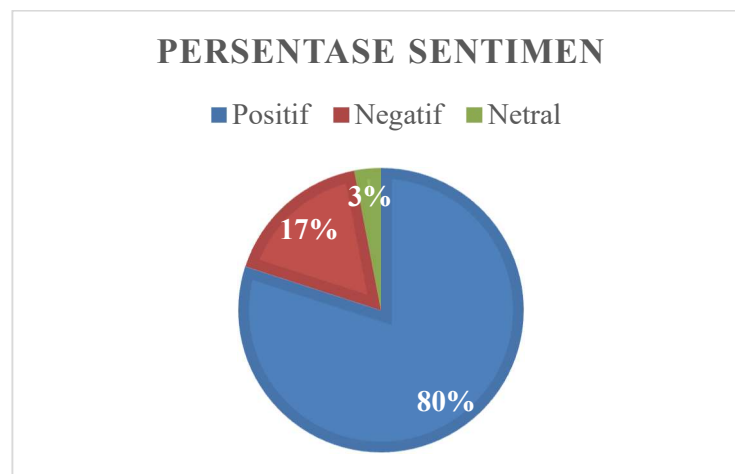
Tabel 1. Proses *Preprocessing Data*

No	Proses	Hasil
1	Data Ulasan	Aplikasi nya sangat membantu mempermudah order, dan juga bisa mendapatkan cashback 👍
2	<i>Cleaning</i>	Aplikasi nya sangat membantu mempermudah order dan juga bisa mendapatkan cashback
3	<i>Case-Folding</i>	aplikasi nya sangat membantu mempermudah order dan juga bisa mendapatkan cashback
4	<i>Tokenizing</i>	"aplikasi", "nya", "sangat", "membantu", "mempermudah", "order", "dan", "juga", "bisa", "mendapatkan", "cashback".
5	<i>Filtering</i>	"aplikasi", "membantu", "mempermudah", "order", "mendapatkan", "cashback".
6	<i>Stemming</i>	"aplikasi", "bantu", "permudah", "order", "dapat", "cashback".
Hasil Akhir		aplikasi bantu permudah order dapat cashback

c. Analisis Sentimen

Setelah dilakukan preprocessing data seperti pada contoh pada Tabel 2 proses selanjutnya adalah mengetahui 3 kategori sentimen yaitu sentimen positif, sentimen negatif, dan sentimen netral dengan menggunakan tiga model *machine learning* yang menggunakan pembobotan *TF-IDF* menghasilkan hasil yang baik.

Berdasarkan pada Gambar 1 menunjukkan bahwa distribusi persentase sentimen pengguna terhadap aplikasi FamiApps berdasarkan analisis sentimen.



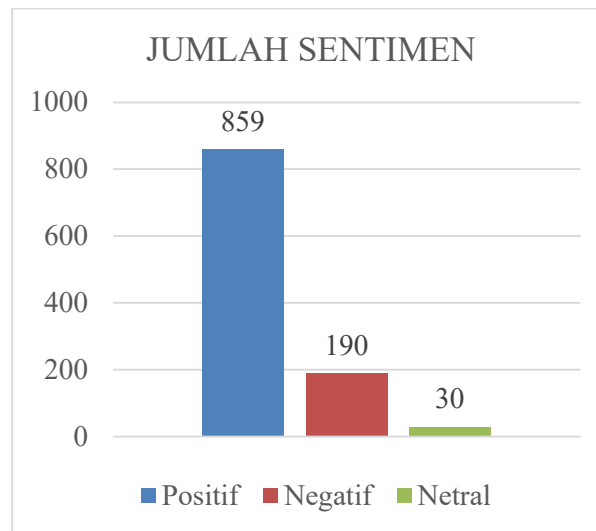
Gambar 1. Persentase Sentimen

Diagram lingkaran ini membagi sentimen ke dalam tiga kategori utama:

1. Sentimen Positif (80%): Mayoritas ulasan dari pengguna bersentimen positif, menunjukkan bahwa sebagian besar pengguna merasa puas dengan aplikasi ini dan memiliki pengalaman yang baik.
2. Sentimen Negatif (17%): Sebagian kecil ulasan mencerminkan ketidakpuasan pengguna, yang menunjukkan adanya masalah atau kekurangan dalam aplikasi yang perlu diperbaiki.
3. Sentimen Netral (3%): Persentase kecil ulasan menunjukkan sentimen netral, di mana pengguna memberikan komentar yang tidak terlalu menunjukkan kepuasan maupun ketidakpuasan.

Dari hasil ini, dapat disimpulkan bahwa aplikasi *FamiApps* secara umum mendapatkan respons yang sangat positif dari penggunanya, meskipun tetap terdapat ruang untuk perbaikan berdasarkan ulasan negatif.

Gambar 2 menampilkan grafik batang yang menunjukkan jumlah ulasan pengguna berdasarkan kategori sentimen terhadap aplikasi *FamiApps*.



Gambar 2. Grafik Jumlah Sentimen

Grafik ini mengilustrasikan jumlah data yang diklasifikasikan sebagai sentimen positif, negatif, dan netral:

1. Sentimen Positif: Ditandai dengan batang berwarna biru, kategori ini memiliki jumlah ulasan tertinggi, mencapai sekitar 900 ulasan. Hal ini menunjukkan bahwa mayoritas pengguna memiliki pengalaman yang sangat memuaskan dengan aplikasi *FamiApps*.
2. Sentimen Negatif: Ditandai dengan batang berwarna merah, kategori ini memiliki jumlah ulasan sekitar 200. Angka ini mengindikasikan adanya sejumlah pengguna yang merasa kurang puas dengan aplikasi.

3. Sentimen Netral: Ditandai dengan batang berwarna hijau, kategori ini memiliki jumlah ulasan yang paling sedikit, sekitar 30. Ulasan ini menunjukkan bahwa pengguna memberikan tanggapan yang netral tanpa menunjukkan perasaan puas atau tidak puas.

Grafik ini memberikan gambaran yang lebih konkret tentang distribusi jumlah ulasan berdasarkan kategori sentimen, mendukung analisis bahwa sentimen positif mendominasi ulasan pengguna aplikasi *FamiApps*. Hasil yang dihasilkan menunjukkan bahwa *Support Vector Machine* memiliki performa yang sedikit lebih baik dibandingkan *Random Forest* dan *Naïve Bayes*

Tabel 2 dan Tabel 3 Menunjukkan bahwa hasil tertinggi dari Skenario 80:20 dan 70:30 dipegang oleh *Support Vector Machine (SVM)* dengan hasil *accuracy* sebesar 98,58%, *precision* sebesar 98,57%, *recall* sebesar 98,58% dan *F1-score* sebesar 98,58%.

Tabel 2. Confusion matrix dengan skenario 80:20

<i>Model</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
<i>RF</i>	97,76%	97,80%	97,76%	97,77%
<i>SVM</i>	98,58%	98,57%	98,58%	98,58%
<i>NB</i>	94,31%	94,65%	94,31%	94,29%

Tabel 3. Confusion matrix dengan skenario 70:30

<i>Model</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
<i>RF</i>	96,48%	96,55%	96,48%	96,48%
<i>SVM</i>	98,24%	98,24%	98,24%	98,23%
<i>NB</i>	94,04%	94,45%	94,04%	94,05%

Aplikasi *FamiApps* yang tersedia di *Google Play Store* telah banyak digunakan oleh pengguna, namun hingga saat ini belum terdapat penelitian mendalam yang menganalisis pola emosional mereka. Studi ini bertujuan untuk memahami sentimen pengguna serta menggali wawasan terkait persepsi mereka terhadap aplikasi. Penelitian ini membandingkan tiga algoritma *machine learning*, yaitu *Random Forest*, *Naïve Bayes*, dan *Support Vector Machine (SVM)*. Dataset yang digunakan mencakup 1.079 ulasan pengguna, yang diklasifikasikan menjadi tiga kategori sentimen: positif, negatif, dan netral. Analisis dilakukan melalui serangkaian tahapan *preprocessing data*, seperti *cleaning*, *tokenizing*, *stemming*, serta pembobotan *TF-IDF* untuk mengonversi teks ke dalam format numerik. Model dievaluasi berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

Hasil penelitian menunjukkan bahwa *SVM* unggul dalam kinerja, dengan akurasi tertinggi sebesar 98,58% pada skenario pemisahan data 80:20 dan 98,24% pada skenario 70:30. Keunggulan ini dikarenakan *SVM* memiliki kemampuan superior dalam mengolah data berdimensi tinggi yang dihasilkan dari ekstraksi fitur *TF-IDF*, dimana algoritma ini dapat menemukan hiperplane optimal yang memaksimalkan margin antar kelas sentimen. Hal ini sejalan dengan temuan Rabbani et. al., (2023) yang menunjukkan

bahwa *SVM* konsisten memberikan performa terbaik pada analisis sentimen dengan fitur *TF-IDF*.

Sementara itu, *Random Forest* menunjukkan performa yang baik dalam menangani data kompleks dan mengurangi risiko *overfitting* melalui mekanisme *ensemble learning*, namun mengalami penurunan akurasi ketika berhadapan dengan fitur yang kurang relevan atau *noise* dalam data teks. Temuan ini konsisten dengan penelitian Wahid dkk. (2024) yang menyatakan bahwa *Random Forest* cenderung mengalami dilusi performa ketika jumlah fitur tidak relevan meningkat dalam analisis sentimen. Adapun *Naïve Bayes*, meskipun unggul dalam kecepatan komputasi dan kesederhanaan implementasi, menunjukkan keterbatasan signifikan dalam menangani asumsi independensi fitur yang jarang terpenuhi dalam data teks natural, sejalan dengan Husin, (2023), yang menunjukkan bahwa asumsi independensi *Naïve Bayes* sering dilanggar dalam konteks analisis sentimen.

Implikasi praktis dari temuan ini menunjukkan bahwa untuk aplikasi analisis sentimen skala besar seperti FamiApps, implementasi *SVM* dapat memberikan insight yang lebih akurat untuk pengambilan keputusan bisnis, seperti identifikasi aspek aplikasi yang perlu diperbaiki berdasarkan sentimen negatif atau pengembangan fitur baru berdasarkan feedback positif pengguna. Selain itu, akurasi tinggi yang dicapai *SVM* (98,58%) memberikan fondasi yang *solid* untuk sistem rekomendasi atau *early warning system* terkait kepuasan pengguna aplikasi.

4. Kesimpulan dan Saran

Penelitian ini menunjukkan bahwa *Support Vector Machine (SVM)* merupakan algoritma paling efektif dalam mengklasifikasikan sentimen ulasan pengguna aplikasi FamiApps di *Google Play Store*. Pada skenario 80:20, *SVM* mencapai *accuracy* 98,58%, *precision* 98,57%, *recall* 98,58%, dan *F1-score* 98,58%, melampaui performa *Random Forest* (*accuracy* 97,76%, *F1-score* 97,77%) serta *Naïve Bayes* (*accuracy* 94,31%, *F1-score* 94,29%). Keunggulan serupa juga terlihat dalam skenario 70:30, di mana *SVM* tetap menunjukkan performa tertinggi dengan *accuracy* 98,24%, *precision* 98,24%, *recall* 98,24%, dan *F1-score* 98,23%, dibandingkan *Random Forest* (*accuracy* 96,48%, *F1-score* 96,48%) dan *Naïve Bayes* (*accuracy* 94,04%, *F1-score* 94,05%). Konsistensi hasil ini mengindikasikan bahwa *SVM* memiliki kemampuan yang lebih handal dalam mengklasifikasikan sentimen pengguna secara akurat dan stabil.

Selain itu, hasil penelitian ini memiliki implikasi praktis dalam pengelolaan ulasan pengguna untuk pengembangan aplikasi. Dengan mengadopsi model *SVM* dalam analisis sentimen, pengembang dapat secara otomatis mendeteksi pola sentimen pengguna dengan lebih presisi, sehingga memungkinkan mereka untuk meningkatkan pengalaman pengguna, merespons masukan secara lebih cepat, dan memahami kebutuhan pelanggan dengan lebih baik. Oleh karena itu, *SVM* dapat dianggap sebagai pendekatan yang paling optimal untuk analisis sentimen yang membutuhkan tingkat ketepatan dan akurasi tinggi.

Penelitian selanjutnya dapat memperluas cakupan dengan mengeksplorasi algoritma berbasis *deep learning*, seperti *LSTM* atau *Transformer-based models (BERT)*, untuk meningkatkan akurasi analisis sentimen. Selain itu, analisis temporal dapat dilakukan untuk memahami tren sentimen pengguna terhadap pembaruan aplikasi dari waktu ke waktu. Penelitian juga dapat memasukkan data ulasan dari platform lain, seperti media sosial atau *App Store*, untuk mendapatkan perspektif yang lebih menyeluruh. Dengan menambahkan

analisis faktor-faktor yang memengaruhi sentimen, seperti fitur tertentu dalam aplikasi, hasil penelitian dapat memberikan rekomendasi yang lebih spesifik bagi pengembang untuk meningkatkan pengalaman pengguna.

Daftar Pustaka

- Dinata, R. K., Safwandi, S., Hasdyna, N., & Mahendra, R. (2020). Kombinasi Algoritma Brute Force dan Stemming pada Sistem Pencarian Mashdar. *CESS (Journal of Computer Engineering, System and Science)*, 5(2), 273. <https://doi.org/10.24114/cess.v5i2.17989>
- Firmansyah, M. M., Sunandar, M. A., & Komara, M. A. (2023). REDESAIN UI/UX FAMI APPS MENGGUNAKAN METODE GOAL DIRECTED DESIGN DAN COGNITIVE WALKTHROUGH. *Jurnal Mahasiswa Teknik Informatika*, 7(5), 3281–3287. <https://doi.org/10.36040/jati.v7i5.7326>
- Herlinawati, N., Yuliani, Y., Faizah, S., Gata, W., & Samudi. (2020). ANALISIS SENTIMEN ZOOM CLOUD MEETINGS DI PLAY STORE MENGGUNAKAN NAÏVE BAYES DAN SUPPORT VECTOR MACHINE. *CESS (Journal of Computer Engineering System and Science)*, 5(2), 293–298. <https://doi.org/10.24114/cess.v5i2.18186>
- Husin, N. (2023). Komparasi Algoritma Random Forest, Naïve Bayes, dan Bert Untuk Multi-Class Classification Pada Artikel Cable News Network (CNN) Nanang Husin. *Jurnal Esensi Infokom*, 7(1), 75–84. <https://pdfs.semanticscholar.org/f048/ea7a08cfca1a0a705eb5fd57efd7ba798fa2.pdf>
- Ilyas, I. F., & Chu, X. (2019). *Data Cleaning*. Association for Computing Machinery. <https://doi.org/10.1145/3310205>
- Kartika, E., & Gondohanindijo, J. (2020). RANCANG BANGUN MODEL SENTIMEN ANALISIS REVIEW PRODUK PADA TOKO ONLINE MENGGUNAKAN NAIVE BAYES. *Seminar Nasional Hasil Peneltiaan (SNHP)*, 201–212.
- Putra, D. H., Asfi, M., & Fahrudin, R. (2021). PERANCANGAN UI/UX MENGGUNAKAN METODE DESIGN THINKING BERBASIS WEB PADA LAPORTEA COMPANY. *Jurnal Ilmiah Teknologi Informasi Terapan*, 8(1), 111–117. <https://doi.org/10.33197/jitter.vol8.iss1.2021.730>
- Rabbani, S., Safitri, D., Rahmadhani, N., Sani, A. A. F., & Anam, M. K. (2023). Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(2), 153–160. <https://doi.org/10.57152/malcom.v3i2.897>
- Rahmadani, R., Rahim, A., & Rudiman. (2024). ANALISIS SENTIMEN ULASAN “OJOL THE GAME” DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES DAN MODEL EKSTRAKSI FITUR TF-IDF UNTUK MENINGKATKAN KUALITAS GAME. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, 12(3), 2928–2936. <https://doi.org/10.23960/jitet.v12i3.4988>
- Rahman, O. H., Abdillah, G., & Komarudin, A. (2021). Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(1), 17–23. <https://doi.org/10.29207/resti.v5i1.2700>

- Rahman, R. N., Rahim, A., & Pranoto, W. J. (2025). Analisis Sentimen Ulasan Game eFootball 2024 Pada Playstore menggunakan Algoritma Naïve Bayes. *Jurnal Ilmiah Informatika (JIF)*, 13(01), 38–44. <https://doi.org/10.33884/jif.v13i01.9913>
- Safrudin, M., Martanto, & Hayati, U. (2024). PERBANDINGAN KINERJA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI SENTIMEN ULASAN GAME GENSHIN IMPACT. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 3182–3188. <https://doi.org/10.36040/jati.v8i3.8415>
- Sinulingga, J. E. B., & Sitorus, H. C. K. (2024). Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF. *Jurnal Manajemen Informatika (JAMIKA)*, 14(1), 42–53. <https://doi.org/10.34010/jamika.v14i1.11946>
- Sulistiyawati, U. S., & Munawir. (2024). Decoding Big Data: Mengubah Data Menjadi Keunggulan Kompetitif dalam Pengambilan Keputusan Bisnis. *Jurnal Manajemen dan Teknologi (JMT)*, 1(2), 58–71. <https://doi.org/10.35870/jmt.vxix.1114>
- Tarwoto, Nugroho, R., Azka, N., & Graha, W. S. R. (2025). Analisis Sentimen Ulasan Aplikasi Mobile JKN di Google PlayStore Menggunakan IndoBERT. *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, 9(2), 495–505. <https://doi.org/10.35870/jtik.v9i2.3340>
- Urva, G., Desyanti, Albanna, I., Sungkar, M. S., Gunawan, I. M. A. O., Adhicandra, I., Ramadhan, S., Rahardian, R. L., Herlawati, Handayanto, R. T., Ariana, A. A. G. B., Tartatik, Atika, P. D., & Junaidi, S. (2023). *PENERAPAN DATA MINING DI BERBAGAI BIDANG : Konsep, Metode, dan Studi Kasus*.
- Wahid, A. M., Turino, Nugroho, K. A., Safitri, T., Darmono, & Utomo Fandy Setyo. (2024). Optimasi Logistic Regression dan Random Forest untuk Deteksi Berita Hoax Berbasis TF-IDF. *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, 4(8), 381–392. <https://doi.org/https://doi.org/10.52436/1.jpti.602>
- Widagdo, A. S., A. B. S. W., & Nasiri, A. (2020). Analisis Tingkat Kepopuleran E-Commerce Di Indonesia Berdasarkan Sentimen Sosial Media Menggunakan Metode Naïve Bayes. *Jurnal INFORMA Politeknik Indonusa Surakarta*, 6(1), 1–5. <https://doi.org/10.46808/informa.v6i1.159>