 ISSN:2302-8149
e-ISSN:2540-9719

Perbandingan Kinerja Algoritma *Naïve Bayes, Random Forest, dan Support Vector Machine* dalam Analisis Sentimen Aplikasi *Weverse*

Performance Comparison of Naïve Bayes, Random Forest, and Support Vector Machine Algorithms in Sentiment Analysis of the Weverse Application

[1]Anisa Fitriyani*, [2]Ichsan Ibrahim
[1,2]Program Studi Teknik Informatika, STMIK IM
[1,2]Jl. Belitung No.7, Merdeka, Kec. Sumur Bandung, Kota Bandung, Jawa Barat 40113
*e-mail: *fitriyanianisa18@gmail.com*

(***received*:** 28 November 2025, ***revised*:** 9 January 2026, ***accepted*:** 10 January 2026)

Abstrak

Perkembangan teknologi digital telah meningkatkan interaksi antara artis dan penggemar melalui berbagai platform komunitas, salah satunya Weverse. Ulasan pengguna pada platform ini mencerminkan persepsi dan tingkat kepuasan yang penting bagi pengembang dalam meningkatkan kualitas layanan. Namun, banyaknya ulasan yang tidak terstruktur menimbulkan kesulitan dalam mengidentifikasi pola opini secara sistematis. Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja algoritma Naïve Bayes, Random Forest, dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen positif dan negatif pada 5.000 ulasan pengguna Weverse yang diperoleh dari Google Play Store melalui teknik *web scraping*. Data diproses melalui tahapan *preprocessing*, diterjemahkan ke Bahasa Inggris, dan diberi label secara otomatis menggunakan VADER, sedangkan pembobotan fitur dilakukan dengan metode TF-IDF. Hasil pengujian menunjukkan bahwa Random Forest memiliki performa terbaik dengan akurasi 83%, diikuti SVM sebesar 80,2% dan Naïve Bayes sebesar 74,2%. Temuan ini menegaskan bahwa Random Forest lebih efektif dalam menangani data teks berdimensi tinggi secara stabil dan konsisten. Penelitian ini diharapkan dapat menjadi referensi dalam pengembangan model analisis sentimen berbahasa Indonesia pada platform komunitas digital di masa mendatang.

Kata kunci: *naïve bayes, random forest,* analisis sentimen, *support vector machine, weverse*

Abstract

The rapid development of digital technology has enhanced interactions between artists and fans through various community platforms, one of which is Weverse. User reviews on this platform reflect perceptions and satisfaction levels that are valuable for developers in improving service quality. However, the large volume of unstructured reviews creates challenges in systematically identifying opinion patterns. This study aims to analyze and compare the performance of Naïve Bayes, Random Forest, and Support Vector Machine (SVM) algorithms in classifying positive and negative sentiments from 5,000 Weverse user reviews collected from the Google Play Store through web scraping techniques. The data were processed through preprocessing stages, translated into English, and automatically labeled using VADER, while feature weighting was performed using the TF-IDF method. The experimental results indicate that Random Forest achieved the best performance with an accuracy of 83%, followed by SVM with 80.2% and Naïve Bayes with 74.2%. These findings confirm that Random Forest is more effective in handling high-dimensional text data in a stable and consistent manner. This study is expected to serve as a reference for the development of Indonesian-language sentiment analysis models on digital community platforms in future research.

Keywords: *naïve bayes, random forest, sentiment analysis, support vector machine, weverse*

 ISSN:2302-8149
e-ISSN:2540-9719

1 Pendahuluan

Seiring dengan pesatnya perkembangan teknologi digital, berbagai platform interaksi antara artis dan penggemar mulai bermunculan. Salah satu yang populer adalah Weverse, aplikasi asal Korea Selatan yang dibangun oleh Weverse Company di bawah naungan HYBE Corporation. Fokus aplikasi ini adalah komunikasi dan distribusi materi digital antara artis dan penggemar melalui fitur interaktif seperti *posting*, *live streaming*, dan *direct message*. Selain menyediakan konten eksklusif, Weverse juga memungkinkan pengguna untuk berpartisipasi dalam komunitas global, memberikan komentar, serta membeli *merchandise* resmi artis favorit mereka. Dengan fitur-fitur tersebut, Weverse telah menjadi inovasi digital penting dalam industri hiburan Korea Selatan karena mampu menciptakan ruang komunikasi yang aman, inklusif, dan interaktif antara artis dan penggemar di seluruh dunia [1].

Ulasan pengguna, termasuk rekomendasi dan penilaian, berperan penting dalam meningkatkan performa aplikasi sekaligus mendukung pengembangannya secara berkelanjutan. Namun, beberapa ulasan tidak merepresentasikan kondisi aplikasi dengan tepat, karena dapat dipengaruhi oleh bug, masalah teknis, maupun fitur yang belum tersedia. Situasi tersebut dapat menurunkan ketertarikan pengguna dan berdampak pada reputasi aplikasi. Oleh karena itu, analisis ulasan harus dilakukan untuk mengidentifikasi perasaan positif dan negatif guna membantu pengembang dalam melakukan evaluasi dan perbaikan [2]. Analisis sentimen menjadi penting untuk memahami bagaimana pengguna melihat aplikasi, mengingat data ulasan di Google Play Store bersifat tidak terstruktur dan berjumlah sangat besar. Ulasan tersebut perlu diolah agar dapat menghasilkan informasi yang bermanfaat untuk meningkatkan kualitas aplikasi serta pengalaman pengguna [3]. Secara umum, analisis sentimen merupakan metode komputasional yang memproses teks digital guna mengidentifikasi apakah suatu kalimat atau dokumen memuat opini positif maupun negatif. Metode ini membantu mengukur tingkat kepuasan pengguna sekaligus mendukung pengembang dalam menyesuaikan layanan terhadap perubahan kebutuhan pasar [4].

Beragam cara untuk mengklasifikasikan data telah dibuat untuk meningkatkan keakuratan dalam menganalisis sentimen, antara lain Naïve Bayes, Support Vector Machine (SVM), dan Random Forest. Metode Naïve Bayes adalah teknik yang menggunakan peluang dan terkenal cepat serta efisien dalam tugas klasifikasi teks [5]. SVM berfungsi dengan menemukan *hyperplane* yang paling efektif untuk memilah data ke dalam kelas [6]. Adapun Random Forest adalah metode pembelajaran yang menggabungkan berbagai pohon keputusan untuk menggabungkan hasil dan meningkatkan performa serta stabilitas model klasifikasi [7].

Meskipun telah banyak penelitian yang membahas tentang analisis sentimen dari ulasan aplikasi di platform Google Play Store, sampai sekarang belum ada yang secara khusus meneliti ulasan pengguna aplikasi Weverse dengan membandingkan kinerja tiga algoritma, yaitu Naïve Bayes, Random Forest, dan SVM. Dengan demikian, penelitian ini diharapkan dapat menutup kesenjangan tersebut melalui analisis komparatif yang mendalam terhadap kinerja ketiga model dalam proses klasifikasi sentimen.

Dengan demikian, tujuan penelitian ini adalah untuk menganalisis dan membandingkan kinerja algoritma Naïve Bayes, Random Forest, dan SVM dalam mengklasifikasikan sentimen dari ulasan pengguna aplikasi Weverse dengan menggunakan cara pemberian bobot yang dikenal sebagai *Term Frequency–Inverse Document Frequency* (TF-IDF). Dari perspektif teori, penelitian ini diharapkan dapat meningkatkan pemahaman di bidang *text mining* dan analisis sentimen. Dari perspektif praktik, hasil penelitian ini dapat memfasilitasi pengembang aplikasi memahami persepsi pengguna yang lebih objektif dan meningkatkan mutu layanan pada aplikasi Weverse maupun platform sejenis.

2 Tinjauan Literatur

Analisis sentimen merupakan pendekatan berbasis teks yang digunakan untuk mengekstraksi opini publik terhadap suatu layanan atau aplikasi digital melalui tahapan *preprocessing*, pembobotan teks seperti *Term Frequency–Inverse Document Frequency* (TF-IDF), dan klasifikasi menggunakan algoritma *machine learning* seperti Naïve Bayes, Random Forest, dan Support Vector Machine (SVM) [8]. Pendekatan ini banyak dimanfaatkan untuk menganalisis persepsi pengguna terhadap suatu aplikasi, terutama melalui ulasan yang diperoleh dari platform seperti Google Play Store [9]. Ulasan tersebut biasanya bersifat tidak terstruktur, sehingga pemilihan algoritma dan teknik *preprocessing* menjadi aspek penting untuk memperoleh performa klasifikasi yang optimal [10].

 ISSN:2302-8149
e-ISSN:2540-9719

Sejumlah penelitian terdahulu telah menguji performa ketiga algoritma pada berbagai domain aplikasi. Tarigan *et al*. (2025) misalnya, menemukan bahwa SVM mencapai akurasi tertinggi sebesar 82% pada ulasan aplikasi Sirekap, diikuti Random Forest dengan 81% dan Naïve Bayes dengan 71%. Penelitian tersebut menunjukkan bahwa SVM cenderung unggul dalam mengenali pola opini pengguna berdasarkan fitur teks yang kompleks [11]. Kurniawan *et al*. (2025) menguatkan kecenderungan tersebut melalui studi terhadap aplikasi Ajaib, yang menunjukkan bahwa SVM masih menghasilkan akurasi tertinggi sebesar 91%. Di sisi lain, Random Forest memberikan performa yang lebih seimbang, dengan *recall* mencapai 95% dan *F1-score* sebesar 91%. Penemuan ini mengindikasikan bahwa Random Forest memiliki stabilitas yang baik dalam mendeteksi sentimen positif, sementara SVM tetap unggul dalam hal ketepatan klasifikasi secara keseluruhan [12].

Penelitian Irawan *et al*. (2024) pada aplikasi CapCut memperkuat temuan sebelumnya. SVM kembali mencatat akurasi tertinggi sebesar 86%, diikuti Random Forest sebesar 83% dan Naïve Bayes sebesar 70%. Menariknya, nilai *precision* yang tinggi pada kelas netral menunjukkan kemampuan SVM dalam memahami konteks kalimat yang lebih kompleks pada teks berbahasa Indonesia [13]. Konsistensi dominasi SVM juga tampak pada penelitian Mustofa dan Idris (2024) yang mengkaji ulasan pada dua aplikasi besar, Shopee dan Zoom. Pada aplikasi Zoom, Random Forest mencapai akurasi tertinggi sebesar 94,15% dan SVM sedikit di bawahnya dengan nilai di atas 93%. Namun pada aplikasi Shopee, SVM unggul dengan akurasi sebesar 80,69%. Hal ini menunjukkan bahwa SVM dan Random Forest dapat beradaptasi dengan baik terhadap karakteristik data yang berbeda, terutama pada dataset besar dengan variasi opini pengguna yang tinggi [14].

Temuan berbeda diperoleh oleh Amelia *et al*. (2024) yang meneliti ulasan aplikasi MyPertamina. Dalam studi tersebut, Random Forest mencatat akurasi tertinggi sebesar 99,77%, diikuti oleh SVM sebesar 99,31% dan Naïve Bayes sebesar 90,24%. Ketiga model tetap menunjukkan performa tinggi pada *precision*, *recall*, dan *F1-score*, menegaskan bahwa pembobotan TF-IDF efektif dalam meningkatkan akurasi klasifikasi teks ulasan [15]. Sebaliknya, penelitian Rizki *et al*. (2025) terhadap aplikasi MyICON+ kembali menunjukkan bahwa SVM memiliki keunggulan dalam hal akurasi 97%, meskipun Random Forest dan Naïve Bayes tetap memberikan hasil kompetitif [16]. Variasi temuan ini mengindikasikan bahwa tidak ada algoritma yang secara universal paling unggul; performa sangat ditentukan oleh karakteristik dataset yang digunakan, seperti panjang teks, jumlah data, tingkat repetisi kata, hingga pola bahasa pengguna.

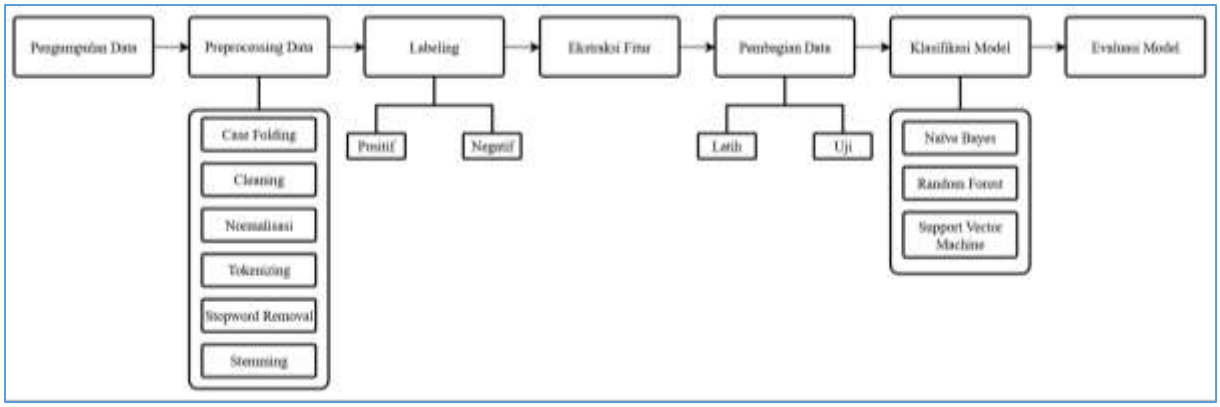
Berdasarkan keseluruhan literatur tersebut, dapat disimpulkan bahwa SVM cenderung menjadi algoritma yang paling konsisten dan unggul, terutama pada dataset dengan konteks yang kompleks dan struktur bahasa yang beragam. Sementara itu, Random Forest sering menunjukkan stabilitas pada dataset dengan distribusi kelas yang relatif seimbang dan jumlah data besar. Naïve Bayes umumnya menghasilkan performa paling rendah, namun tetap efektif pada teks sederhana dengan pola kosakata yang homogen. Meskipun demikian, terdapat kontradiksi antar penelitian yang menegaskan bahwa efektivitas suatu algoritma tidak dapat dilepaskan dari karakteristik domain aplikasi, struktur bahasa pengguna, serta teknik *preprocessing* yang diterapkan.

Dari sudut pandang kesenjangan penelitian, studi-studi terdahulu masih berfokus pada aplikasi umum seperti *e-commerce*, keuangan, dan layanan publik. Belum ditemukan penelitian yang secara khusus menganalisis sentimen pengguna aplikasi Weverse, yaitu platform komunitas hiburan yang memiliki karakteristik bahasa unik—lebih informal, sarat ekspresi emosional, dan dipengaruhi oleh budaya fandom. Dengan demikian, penelitian ini berupaya mengisi kekosongan tersebut dengan mengevaluasi performa Naïve Bayes, Random Forest, dan SVM terhadap ulasan Weverse menggunakan pembobotan TF-IDF. Pendekatan ini diharapkan dapat memberikan kontribusi empiris dalam menentukan algoritma yang paling efektif untuk teks ulasan dalam konteks aplikasi komunitas penggemar, sekaligus memperkaya literatur analisis sentimen berbahasa Indonesia pada domain hiburan digital yang selama ini belum banyak diteliti.

3 Metode Penelitian

Metode kuantitatif komparatif digunakan dalam penelitian ini untuk menilai dan membandingkan tingkat akurasi tiga algoritma klasifikasi teks, yaitu Naïve Bayes, Random Forest, dan SVM, untuk mengidentifikasi sentimen ulasan pengguna aplikasi Weverse di Google Play Store. Gambar 1 menunjukkan langkah-langkah yang diambil dalam penelitian ini.

 ISSN:2302-8149
e-ISSN:2540-9719



Gambar 1 Alur proses penelitian

3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini dikumpulkan melalui proses *web scraping* menggunakan *library google-play-scraper* pada bahasa pemrograman Python [17]. *Web scraping* merupakan teknik pengambilan data terstruktur dari situs web secara otomatis melalui bantuan perangkat lunak atau kode pemrograman tertentu [18]. Teknik ini digunakan untuk memperoleh data ulasan pengguna aplikasi Weverse yang tersedia di platform Google Play Store. Hasil *scraping* kemudian disimpan dalam format *Comma Separated Values* (*.csv*) untuk memudahkan pengolahan pada tahap *preprocessing* dan analisis selanjutnya. Penggunaan *google-play-scraper* memungkinkan peneliti memperoleh data secara terstruktur, efisien, serta mengurangi potensi kesalahan manusia dalam pengambilan data secara manual. Pendekatan ini dinilai efektif untuk kebutuhan analisis sentimen berbasis *machine learning* [19].

3.2 *Preprocessing* Data

Dalam tahap *preprocessing*, dataset yang telah dikumpulkan akan melalui proses pembersihan guna memastikan konsistensi dan kualitas data sebelum dilakukan analisis lebih lanjut. Tujuan utamanya yaitu menghilangkan data yang tidak valid, duplikat, atau hilang, mempersiapkan data agar lebih mudah diolah dan lebih akurat dalam proses analisis [20]. Beberapa langkah yang dilakukan pada tahap *preprocessing* meliputi:

1. *Case Folding*
   Case folding adalah proses mengubah seluruh karakter dalam dokumen menjadi huruf kecil (*lowercase*) guna menyeragamkan bentuk teks dan mencegah perbedaan antara huruf besar dan kecil [21].
2. *Cleaning*
   Cleaning merupakan proses penyederhanaan data teks melalui penghapusan elemen yang tidak relevan atau tidak bermakna. Langkah ini dilaksanakan untuk memperlancar proses pengolahan data teks dan memperoleh data dengan kualitas yang lebih baik [22].
3. Normalisasi
   Normalisasi adalah proses untuk menyesuaikan kata tidak baku menjadi bentuk baku, termasuk mengganti singkatan atau kata *slang* agar sesuai dengan kaidah bahasa yang benar [23]. Pada penelitian ini, normalisasi dilakukan menggunakan kamus *slang* bahasa Indonesia yang berisi pasangan kata tidak baku dan padanan bakunya. Setiap kata pada teks ulasan dibandingkan dengan entri dalam kamus tersebut dan digantikan dengan bentuk baku apabila ditemukan kecocokan. Tahap ini bertujuan untuk mengurangi variasi bahasa informal yang umum digunakan dalam ulasan fandom Weverse sehingga meningkatkan konsistensi representasi teks sebelum pembobotan TF-IDF dan proses klasifikasi sentimen.
4. *Tokenizing*
   Tokenizing adalah proses memecah teks menjadi bagian-bagian kecil yang disebut *token*, yang dapat berupa kata, frasa, atau simbol tertentu dalam suatu kalimat [17].
5. *Stopword Removal*
   Stopword removal adalah proses eliminasi kata-kata umum yang tidak memiliki makna signifikan dalam analisis, seperti "dan", "ke", “di”, "yang", dan sebagainya [7].

 ISSN:2302-8149
e-ISSN:2540-9719

6. *Stemming*
   Langkah akhir dalam proses *preprocessing* adalah *stemming*, yaitu proses mengonversi kata ke bentuk dasarnya agar lebih sederhana dan seragam. Dalam penelitian ini, *library* Sastrawi digunakan untuk proses *stemming* [24].

3.3 *Labeling*

Pada tahap *labeling*, dataset diberi label sentimen untuk setiap data yang telah dikumpulkan, di mana label tersebut berperan penting dalam mengukur tingkat akurasi pada tahap analisis berikutnya. Proses pelabelan dilakukan dengan menggunakan modul VADER (*Valence Aware Dictionary and sEntiment Reasoner*) *lexicon*, yang berisi kumpulan kata dengan skor sentimen positif dan negatif. Setiap kata dalam ulasan diberi bobot berdasarkan nilai sentimen yang tercatat dalam kamus tersebut, dengan rentang skor antara -1 hingga 1. Setelah setiap kata memperoleh bobotnya, seluruh bobot dijumlahkan untuk menentukan sentimen keseluruhan dari ulasan. Jika total skor positif lebih besar daripada skor negatif, maka ulasan dikategorikan sebagai positif, sedangkan jika sebaliknya, ulasan diberi label negatif. Penelitian ini membatasi kategori sentimen menjadi dua, yaitu positif dan negatif, guna menyederhanakan proses analisis serta meningkatkan konsistensi hasil klasifikasi. Pendekatan dua kelas ini juga umum digunakan pada penelitian analisis sentimen berbasis *lexicon* seperti VADER, karena menghasilkan performa yang lebih stabil dan mudah diinterpretasikan. Kategori netral diabaikan karena jumlahnya relatif sedikit dan sering kali bersifat ambigu secara konteks. Penulis menyadari bahwa proses terjemahan berpotensi memengaruhi interpretasi sentimen yang dihasilkan oleh VADER. Untuk meminimalkan potensi bias, dilakukan validasi manual secara kualitatif terhadap 200 ulasan yang dipilih secara acak dari dataset dengan membandingkan secara langsung label sentimen yang dihasilkan oleh VADER dan hasil interpretasi manual pada ulasan yang sama**.** Hasil validasi menunjukkan bahwa sebagian besar label sentimen yang dihasilkan oleh VADER konsisten dengan interpretasi manual, dengan tingkat kesesuaian sebesar 76%. Perbedaan label umumnya terjadi pada ulasan yang mengandung bahasa informal, ekspresi campuran, atau konteks emosional yang ambigu. Oleh karena itu, hasil pelabelan sentimen dalam penelitian ini tetap mengacu pada VADER sebagai metode pelabelan utama.

3.4 Ekstraksi Fitur

Tahap selanjutnya yaitu ekstraksi fitur, yang dilakukan karena algoritma *machine learning* tidak dapat secara langsung memahami kata atau karakter. Oleh karena itu, teks perlu diubah menjadi representasi numerik agar dapat diproses oleh model. Pada tahap ini metode *Term Frequency–Inverse Document Frequency* (TF-IDF) diterapkan untuk menghitung signifikansi suatu kata dalam dokumen. TF menghitung frekuensi kemunculan suatu kata dalam sebuah dokumen relatif terhadap total kata dalam dokumen tersebut, sedangkan IDF mengevaluasi kelangkaan suatu kata di seluruh kumpulan dokumen. Nilai TF-IDF diperoleh dari hasil TF dikali IDF, yang menghasilkan bobot untuk merepresentasikan signifikansi suatu kata dalam dokumen relatif terhadap keseluruhan korpus teks [25].

3.5 Pembagian Data

Tahap selanjutnya yaitu pembagian data menjadi dua subset menggunakan metode *train–test split*, di mana 80% digunakan sebagai data latih dan 20% sebagai data uji. Pembagian ini adalah praktik standar dalam evaluasi kinerja algoritma *machine learning*, dengan tujuan agar model dapat diuji menggunakan data yang belum pernah diobservasi sebelumnya. Strategi tersebut berfungsi untuk mengukur kemampuan model dalam melakukan generalisasi terhadap data baru secara lebih akurat dan objektif [26]. Untuk menjaga keseimbangan distribusi kelas sentimen, penelitian ini menerapkan parameter *stratify* pada label saat melakukan pembagian data, serta menggunakan *random_state* = 42 agar proses pembagian data bersifat *reproducible*. Meskipun metode ini sudah memadai untuk penelitian ini, *cross-validation* sebenarnya dapat memberikan evaluasi yang lebih stabil karena menguji model pada beberapa skenario pembagian data yang berbeda; pendekatan tersebut tidak diterapkan pada penelitian ini, namun direkomendasikan untuk penelitian selanjutnya.

3.6 Klasifikasi Model

Tahap klasifikasi model dilakukan dengan menggunakan tiga algoritma *machine learning*, yaitu Naïve Bayes, Random Forest, dan SVM. Berikut merupakan persamaan dari ketiga algoritma yang diterapkan dalam penelitian ini:

 ISSN:2302-8149
e-ISSN:2540-9719

1. Naïve Bayes
   Algoritma Naïve Bayes merupakan metode klasifikasi berbasis probabilistik yang sederhana dan cepat. Dengan asumsi independensi antar fitur, algoritma ini menjadi efisien secara komputasi dan sering kali sudah memadai untuk berbagai tugas klasifikasi teks [27]. Perhitungan probabilitas pada algoritma Naïve Bayes dirumuskan dalam Persamaan (1) sebagai berikut:
   $$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$
   Keterangan:
   X : Data yang belum diketahui kelasnya
   H : Hipotesis atau dugaan terhadap kelas data
   $P(H|X)$: Probabilitas hipotesis H benar setelah adanya data X
   $P(H)$: Probabilitas awal (*prior probability*) dari hipotesis H
   $P(X|H)$: Peluang munculnya data X jika hipotesis H benar
2. Random Forest
   Algoritma Random Forest menggunakan teknik *ensemble learning* dengan melatih sejumlah pohon keputusan untuk meningkatkan performa model tanpa menyebabkan *overfitting*. Pendekatan ini efektif diterapkan pada data yang kompleks dan berdimensi tinggi [27]. Perhitungan model Random Forest dirumuskan dalam Persamaan (2) sebagai berikut:
   $$f\ (x)\ =\ Average(f_1(x), f_2(x), \ldots, f_n(x)) \quad (2)$$
   Keterangan:
   $f\ (x)$: Prediksi akhir yang dihasilkan oleh model
   $f_{1-n}(x)$: Prediksi dari masing-masing pohon keputusan ke-n
   x : Data yang digunakan sebagai input pada model
3. Support Vector Machine (SVM)
   SVM merupakan algoritma yang kuat karena mampu menemukan *hyperplane* optimal yang memisahkan data dalam ruang berdimensi tinggi. SVM sangat efektif untuk data yang bersifat linier maupun nonlinier, serta banyak digunakan pada kasus dengan dimensi fitur yang besar [27]. Formulasi pemisahan kelas pada SVM ditunjukkan pada Persamaan (3) dan Persamaan (4) sebagai berikut:
   $$(w.x_i + b) \leq 1, y_i = -1 \quad (3)$$
   $$(w.x_i + b) \geq 1, y_i = 1 \quad (4)$$
   Keterangan:
   x_i : data pada posisi ke-i.
   $w.x_i$: hasil perkalian bobot dengan data ke-i.
   b : nilai bias pada model.
   y_i : label atau kelas dari data ke-i.

3.7 Evaluasi Model

Tahap evaluasi bertujuan untuk mengukur kinerja model dalam melakukan klasifikasi sentimen. Proses ini memanfaatkan *confusion matrix* yang memuat empat komponen utama: *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN). TP merepresentasikan prediksi positif yang sesuai dengan kondisi sebenarnya, FP menunjukkan prediksi positif yang keliru, TN menggambarkan prediksi negatif yang tepat, sedangkan FN merupakan prediksi negatif yang tidak tepat. Keempat nilai tersebut digunakan untuk menganalisis hasil prediksi model melalui perhitungan sejumlah metrik evaluasi, yaitu akurasi, presisi, *recall*, dan skor F1 [28].

1. Akurasi digunakan untuk menghitung persentase keseluruhan sampel yang diprediksi dengan benar oleh model. Nilai akurasi dapat dihitung menggunakan Persamaan (5):
   $$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$
2. Presisi menggambarkan tingkat ketepatan model dalam memprediksi kelas positif, yaitu proporsi data yang benar-benar positif dari seluruh data yang diprediksi positif. Rumus presisi dapat dilihat pada Persamaan (6):
   $$\text{Presisi} = \frac{TP}{TP+FP} \quad (6)$$

 ISSN:2302-8149
e-ISSN:2540-9719

3. *Recall* menunjukkan kemampuan model untuk mengidentifikasi dan menghitung semua data asli yang termasuk dalam kelas positif, dan dihitung menggunakan Persamaan (7):
$$Recall = \frac{TP}{TP+FN} \quad (7)$$
4. Skor F1 merupakan metrik yang menggabungkan nilai presisi dan *recall* secara seimbang melalui rata-rata harmonis keduanya. Nilainya dapat dihitung menggunakan Persamaan (8):
$$\text{Skor F1} = 2x\frac{Presisi \times Recall}{Presisi+Recall} \quad (8)$$

4 Hasil dan Pembahasan

Bagian ini menjelaskan temuan penelitian tentang analisis sentimen aplikasi Weverse menggunakan algoritma Naïve Bayes, Random Forest, dan SVM. Pada tahap ini, hasil penelitian dipaparkan mulai dari pengumpulan data, tahap *preprocessing* data, hasil klasifikasi menggunakan ketiga algoritma *machine learning*, hingga evaluasi kinerja model berdasarkan metrik yang telah ditetapkan. Berikut adalah penjelasan mengenai hasil dan pembahasan dari setiap tahap penelitian yang sudah dilaksanakan.

4.1 Pengumpulan Data

Data ulasan aplikasi Weverse dikumpulkan melalui proses *web scraping* dari Google Play Store (*https://play.google.com/store/apps/details?id=co.benx.weverse&hl=id*). Pengumpulan data dilakukan pada periode Maret 2022 hingga Oktober 2025, dengan total sebanyak 5.000 ulasan yang merupakan ulasan terbaru yang tersedia pada rentang waktu tersebut. Pengambilan data tidak dibatasi oleh kata kunci tertentu, karena seluruh ulasan secara langsung merepresentasikan opini pengguna terhadap aplikasi Weverse. Data yang dikumpulkan mencakup atribut nama pengguna, teks ulasan, skor penilaian, dan tanggal ulasan. Mengingat aplikasi Weverse mendukung berbagai bahasa, proses pengumpulan data dibatasi pada ulasan berbahasa Indonesia dengan menerapkan parameter *lang*=id untuk menjaga konsistensi linguistik dalam proses analisis sentimen. Data ulasan yang diperoleh melalui *web scraping* disajikan dalam Tabel 1.

Tabel 1 Sampel hasil *web scraping*

No	Nama Pengguna	Isi Ulasan
1.	Pengguna Google	Weverse ini seru sih bisa liat idol post krna ga kebanyak mreka pnya akun instagram, jadi seru deh walupun bayar-bayar ttp bagus!! agak bingung si pakenya tp okelah ♡♡♡ rate: 100/10
2.	Pengguna Google	ini kenapa lagi yaa weverse baru update terus udah membership tapi.... kok konten yang membership itu malah nggak bisa ditonton?! sedangkan konten yang nggak membership bisa di tonton, tolong dongggg dengan sangat memohon TOLONG BANGET di benerin lagi🙏🙏🙏🙏klo kayak gitu aku merasa dirugikan udah beli membership digital tp malah nggak bisa digunakan!
3.	Pengguna Google	udah bagus sih apk nya, TAPI gak ada SUBTITLE NYA dan TRANSLATE NYA. walaupun ada di membership , tapi untuk kaum kaum yang modal quota ini begimana:((((,jadi tolong kasih translate nya ditambahin untuk kaum modal quota walaupun ada waktunya gapapa kok. walaupun seminggu sekali translate nya :)))

4.2 *Preprocessing* Data

Pada tahap ini dilakukan proses *preprocessing* data untuk membersihkan dan menyiapkan teks sebelum dilakukan analisis sentimen. Proses *preprocessing* yang dilaksanakan adalah sebagai berikut:

1. *Case Folding*
   Seluruh teks diubah menjadi huruf kecil untuk keseragaman. Contoh hasil penerapan *case folding* ditampilkan pada Tabel 2.

 ISSN:2302-8149
e-ISSN:2540-9719

Tabel 2 Hasil *case folding*

Sebelum	Sesudah
weverse udah baguss bangett,karna itu aplikasi kita untuk komunikasi sama idol kesayangan kita😍,TAPI MINUSNYA KENAPA TRANSLATE NYA HARUS BAYAR?!,KENAPAA,tolong lah di update gratis terjemahan🙏🙏🙏🙏🙏	weverse udah baguss bangett,karna itu aplikasi kita untuk komunikasi sama idol kesayangan kita😍,tapi minusnya kenapa translate nya harus bayar?!,kenapaa,tolong lah di update gratis terjemahan🙏🙏🙏🙏🙏

2. *Cleaning*
Karakter yang tidak diperlukan seperti angka, tanda baca, emoji, dan simbol dihapus sehingga hanya tersisa teks yang relevan. Contoh hasil penerapan *cleaning* ditampilkan pada Tabel 3.

Tabel 3 Hasil *cleaning*

Sebelum	Sesudah
weverse udah baguss bangett,karna itu aplikasi kita untuk komunikasi sama idol kesayangan kita😍,tapi minusnya kenapa translate nya harus bayar?!,kenapaa,tolong lah di update gratis terjemahan🙏🙏🙏🙏🙏	weverse udah baguss bangett karna itu aplikasi kita untuk komunikasi sama idol kesayangan kita tapi minusnya kenapa translate nya harus bayar kenapaa tolong lah di update gratis terjemahan

3. Normalisasi
Kata tidak baku, singkatan, dan slang diganti menjadi bentuk baku agar konsisten dengan kaidah bahasa Indonesia. Contoh hasil penerapan normalisasi ditampilkan pada Tabel 4.

Tabel 4 Hasil normalisasi

Sebelum	Sesudah
weverse udah baguss bangett karna itu aplikasi kita untuk komunikasi sama idol kesayangan kita tapi minusnya kenapa translate nya harus bayar kenapaa tolong lah di update gratis terjemahan	weverse sudah bagus banget karena itu aplikasi kita untuk komunikasi sama idol kesayangan kita tapi minusnya kenapa translate nya harus bayar kenapa tolong lah di update gratis terjemahan

4. *Tokenizing*
Setiap kalimat dipecah menjadi potongan kata (token) untuk memudahkan analisis lebih lanjut. Contoh hasil penerapan *tokenizing* ditampilkan pada Tabel 5.

Tabel 5 Hasil *tokenizing*

Sebelum	Sesudah
weverse sudah bagus banget karena itu aplikasi kita untuk komunikasi sama idol kesayangan kita tapi minusnya kenapa translate nya harus bayar kenapa tolong lah di update gratis terjemahan	['weverse', 'sudah', 'bagus', 'banget', 'karena', 'itu', 'aplikasi', 'kita', 'untuk', 'komunikasi', 'sama', 'idol', 'kesayangan', 'kita', 'tapi', 'minusnya', 'kenapa', 'translate', 'nya', 'harus', 'bayar', 'kenapa', 'tolong', 'lah', 'di', 'update', 'gratis', 'terjemahan']

5. *Stopword Removal*
Kata-kata umum yang tidak memiliki pengaruh terhadap makna, seperti “di”, “dan”, dan “yang”, dieliminasi dari data. Contoh hasil penerapan *stopword removal* ditampilkan pada Tabel 6.

Tabel 6 Hasil *stopword removal*

Sebelum	Sesudah
['weverse', 'sudah', 'bagus', 'banget', 'karena', 'itu', 'aplikasi', 'kita', 'untuk', 'komunikasi', 'sama', 'idol', 'kesayangan', 'kita', 'tapi', 'minusnya', 'kenapa', 'translate', 'nya', 'harus', 'bayar', 'kenapa', 'tolong', 'lah', 'di', 'update', 'gratis', 'terjemahan']	weverse bagus aplikasi komunikasi idol kesayangan minusnya translate bayar tolong update gratis terjemahan

 ISSN:2302-8149
e-ISSN:2540-9719

6. *Stemming*
   Setiap kata dikembalikan ke bentuk dasarnya menggunakan *library* Sastrawi, sehingga kata dengan makna sama tetapi bentuk berbeda dapat diseragamkan. Contoh hasil penerapan *stemming* ditampilkan pada Tabel 7.

Tabel 7 Hasil *stemming*

Sebelum	Sesudah
weverse bagus aplikasi komunikasi idol kesayangan minusnya translate bayar tolong update gratis terjemahan	weverse bagus aplikasi komunikasi idol sayang minus translate bayar tolong update gratis terjemah

4.3 *Labeling*

Setelah tahap *preprocessing* selesai, dataset dilabeli sentimennya secara otomatis menggunakan VADER. Sebelum proses pelabelan, teks ulasan dalam Bahasa Indonesia terlebih dahulu diterjemahkan ke Bahasa Inggris menggunakan *library* Googletrans, karena modul VADER hanya dapat mengenali teks berbahasa Inggris. Setelah hasil terjemahan diperoleh, proses pelabelan dilakukan dengan menghitung skor polaritas menggunakan VADER. Setiap teks diberi nilai sentimen berdasarkan skor yang dihasilkan, kemudian diklasifikasikan menjadi dua kategori utama:

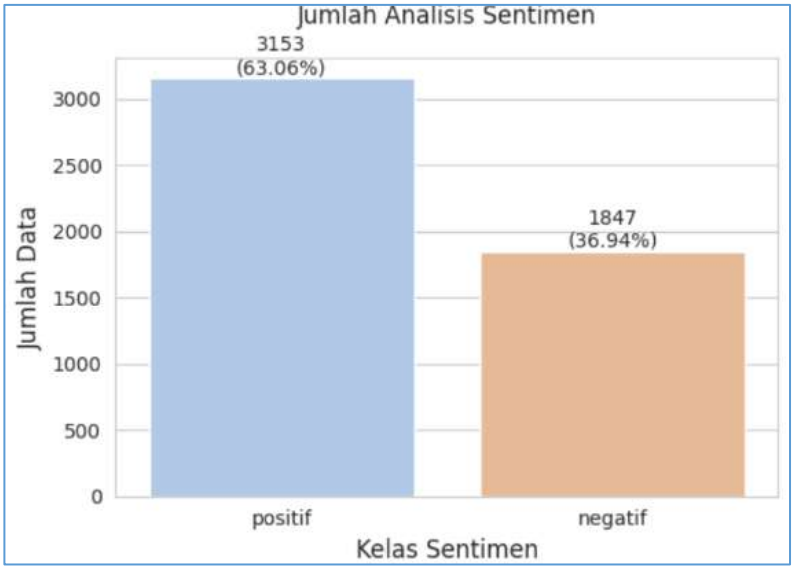
- Positif, jika skor polaritas > 0
- Negatif, jika skor polaritas < 0

Langkah ini menghasilkan dataset dengan dua kelas utama berdasarkan total skor polaritas dari setiap ulasan. Tabel 8 memperlihatkan beberapa contoh hasil pelabelan sentimen menggunakan VADER.

Tabel 8 Sampel *labeling*

Teks Ulasan (Setelah *Preprocessing*)	*Score*	Sentimen
bagus lihat live pop komentar bagus	0.6808	Positif
malas ah pelit subtitlenya live cuman pakai membership	-0.6249	Negatif
download bagus kadang ngonten enhypen tunggu live	0.4404	Positif
suruh update mulu capek memo penuh suruh update mulu ngapa	-0.7184	Negatif

Berdasarkan hasil pelabelan, sebagian besar ulasan menunjukkan kecenderungan positif terhadap aplikasi, dengan total 3.153 ulasan positif (63,06%), sedangkan 1.847 ulasan (36,94%) dikategorikan sebagai negatif. Perbandingan jumlah ulasan pada setiap kategori sentimen divisualisasikan pada Gambar 2.



Gambar 2 Distribusi sentimen

Selanjutnya, visualisasi *WordCloud* pada Gambar 3 dan 4 menunjukkan kata-kata yang paling sering muncul pada setiap kelas sentimen. Kata dominan pada ulasan positif antara lain bagus, *live*,

 ISSN:2302-8149
e-ISSN:2540-9719

idol, dan suka, yang menggambarkan apresiasi terhadap fitur-fitur aplikasi. Dalam ulasan negatif, kata-kata seperti notifikasi, *update*, dan *subtitle*, menandakan adanya keluhan terkait fitur penerjemahan serta pembaruan aplikasi.

Gambar 3 *WordCloud* sentimen positif

Gambar 4 *WordCloud* sentimen negatif

4.4 Ekstraksi Fitur

Tahap berikutnya yaitu pembobotan teks menggunakan TF-IDF. Teknik ini memberikan nilai numerik pada setiap kata untuk merepresentasikan tingkat kepentingannya dalam kumpulan data teks. Hasil pembobotan TF-IDF rata-rata ditunjukkan pada Tabel 9 berikut.

Tabel 9 Hasil pembobotan TF-IDF

Term	TF	TF-IDF (*Average*)
bagus	1854	0.098561
aplikasi	1184	0.050665
suka	757	0.047579
update	813	0.035392
live	886	0.029697

4.5 Pembagian Data

Dataset yang telah diberi label selanjutnya dibagi menjadi dua bagian, yaitu 80% data latih dan 20% data uji. Dari keseluruhan 5.000 ulasan, sebanyak 4.000 digunakan sebagai data latih dan 1.000 sisanya sebagai data uji, yang kemudian dimanfaatkan dalam proses pelatihan serta pengujian model analisis sentimen.

4.6 Klasifikasi Model

Dalam penelitian ini, proses klasifikasi model dilakukan menggunakan tiga algoritma machine learning, yaitu Naïve Bayes, Random Forest, dan SVM. Sebelum tahap klasifikasi dijalankan, data ulasan yang telah melalui tahap *preprocessing* dan pembobotan TF-IDF dibagi menjadi dua bagian, yaitu 80% data latih dan 20% data uji menggunakan metode *train-test split* dengan parameter *test_size* sebesar 0.2 dan *random_state* bernilai 42. Berikut merupakan hasil kinerja dari ketiga model klasifikasi yang diterapkan:

1. Naïve Bayes
   Hasil pengujian model Naïve Bayes terhadap data ulasan menghasilkan metrik performa seperti disajikan pada Tabel 10.

Tabel 10 *Classification report* model Naïve Bayes

Akurasi Naïve Bayes			**0.7420**	
Kelas	**Presisi**	***Recall***	**Skor F1**	***Support***
Negatif	0.74	0.46	0.57	369
Positif	0.74	0.90	0.82	631
Akurasi	–	–	0.74	1000
Macro Average	0.74	0.68	0.69	1000
Weighted Avg	0.74	0.74	0.73	1000

Model Naïve Bayes memperoleh akurasi sebesar 0.7420 (74,2%). Berdasarkan Tabel 10, model menunjukkan presisi yang sama pada kedua kelas, yaitu 0.74. Namun, *recall* pada

 ISSN:2302-8149
e-ISSN:2540-9719

kelas negatif (0.46) jauh lebih rendah dibandingkan kelas positif (0.90), sehingga skor F1 untuk kelas negatif (0.57) juga lebih rendah daripada kelas positif (0.82). Nilai *macro average* dan *weighted average* yang berada pada kisaran 0.68–0.74 mengindikasikan bahwa performa model masih belum seimbang antar kelas. Secara keseluruhan, model cenderung lebih baik dalam mendeteksi sentimen positif dibandingkan negatif, sehingga beberapa kesalahan klasifikasi masih terjadi pada ulasan yang berisi opini negatif atau berkonteks ambigu.

2. Random Forest
   Hasil pengujian model Random Forest terhadap data ulasan menghasilkan metrik performa seperti ditunjukkan pada Tabel 11.

Tabel 11 *Classification report* model *Random Forest*

Akurasi Random Forest			**0.8300**	
Kelas	**Presisi**	***Recall***	**Skor F1**	***Support***
Negatif	0.80	0.72	0.76	369
Positif	0.85	0.89	0.87	631
Akurasi	–	–	0.83	1000
Macro Average	0.82	0.81	0.81	1000
Weighted Avg	0.83	0.83	0.83	1000

Model Random Forest memberikan kinerja yang baik dengan akurasi sebesar 0.8300 (83%). Berdasarkan Tabel 11, model memiliki presisi yang tinggi pada kedua kelas, yaitu 0.80 untuk kelas negatif dan 0.85 untuk kelas positif. Nilai *recall* juga cukup seimbang, masing-masing 0.72 dan 0.89, sehingga menghasilkan skor F1 yang kuat, yaitu 0.76 pada kelas negatif dan 0.87 pada kelas positif. Nilai *macro average* dan *weighted average* yang berada pada kisaran 0.81–0.83 mengindikasikan performa model yang stabil dan konsisten. Secara keseluruhan, Random Forest mampu mengidentifikasi sentimen secara lebih efektif dibandingkan Naïve Bayes, terutama karena kemampuannya menangani variasi fitur TF-IDF dan mendeteksi pola sentimen yang lebih kompleks dalam data ulasan.

3. Support Vector Machine (SVM)
   Hasil pengujian model SVM terhadap data ulasan menghasilkan metrik performa seperti ditunjukkan pada Tabel 12.

Tabel 12 *Classification report* model SVM

Akurasi SVM			**0.8020**	
Kelas	**Presisi**	***Recall***	**Skor F1**	***Support***
Negatif	0.75	0.69	0.72	369
Positif	0.83	0.87	0.85	631
Akurasi	–	–	0.80	1000
Macro Average	0.79	0.78	0.78	1000
Weighted Avg	0.80	0.80	0.80	1000

Model SVM memperoleh akurasi 0.8020 (80,2%). Berdasarkan Tabel 12, model menunjukkan presisi yang cukup baik pada kedua kelas, yaitu 0.75 untuk kelas negatif dan 0.83 untuk kelas positif. Nilai *recall* masing-masing berada pada 0.69 dan 0.87, sehingga menghasilkan skor F1 sebesar 0.72 pada kelas negatif dan 0.85 pada kelas positif. *Macro average* dan *weighted average* yang konsisten pada kisaran 0.78–0.80 menunjukkan bahwa model mampu memberikan performa yang stabil dalam mengklasifikasikan sentimen. Hasil ini menunjukkan bahwa SVM mampu mengidentifikasi pola teks positif dan negatif dengan baik, serta memberikan hasil yang cukup stabil meskipun sedikit di bawah performa Random Forest.

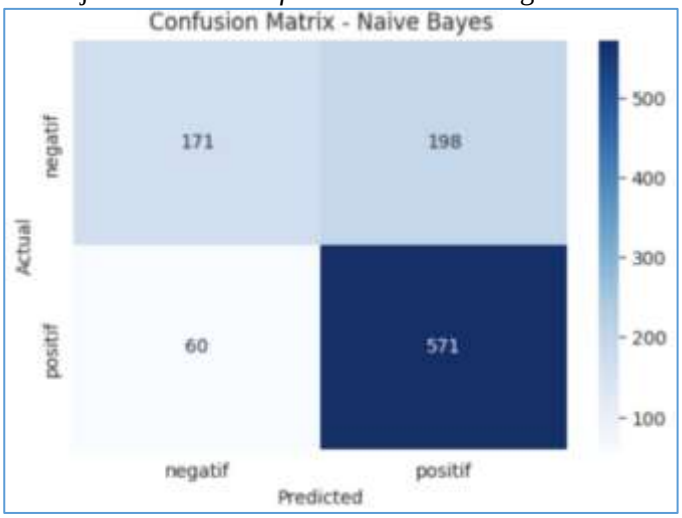
Secara keseluruhan, hasil klasifikasi menunjukkan bahwa algoritma Random Forest memberikan kinerja paling unggul dalam proses klasifikasi sentimen ulasan aplikasi Weverse dibandingkan SVM dan Naïve Bayes.

 ISSN:2302-8149
e-ISSN:2540-9719

4.7 Evaluasi Model

Tahap evaluasi model dilakukan untuk menilai kinerja dari ketiga algoritma klasifikasi yang diterapkan, yaitu Naïve Bayes, Random Forest, dan SVM. Evaluasi dilakukan menggunakan metrik akurasi, presisi, *recall*, dan skor F1, yang dihitung berdasarkan *Confusion Matrix*. Metrik ini menggambarkan sejauh mana model mampu memprediksi kelas sentimen dengan benar.

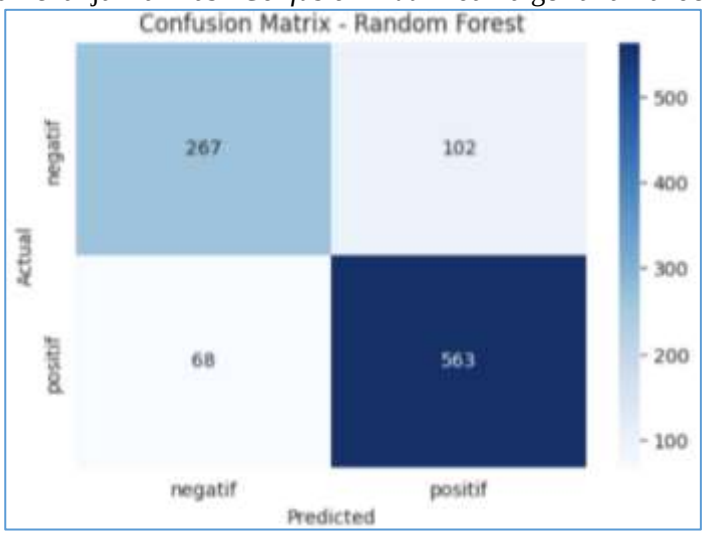
1. Naïve Bayes
   Gambar 5 berikut menunjukkan hasil *Confusion Matrix* dari algoritma Naïve Bayes:



Gambar 5 *Confusion matrix Naïve Bayes*

Berdasarkan Gambar 5, model Naïve Bayes menghasilkan jumlah prediksi benar sebanyak 171 data negatif dan 571 data positif, sementara terdapat 198 data negatif yang salah diprediksi sebagai positif dan 60 data positif yang salah diprediksi sebagai negatif. Temuan ini mengindikasikan bahwa model lebih mampu mengenali sentimen positif dibandingkan sentimen negatif. Namun, tingkat kesalahan pada kelas negatif masih cukup tinggi, yang menandakan bahwa model ini kurang optimal dalam menangani variasi konteks kalimat yang bersifat ambigu atau mengandung kata negasi. Secara keseluruhan, Naïve Bayes dapat melakukan klasifikasi dengan akurasi yang cukup baik, meskipun masih perlu peningkatan terutama pada pengenalan ulasan dengan nada negatif.

2. Random Forest
   Gambar 6 berikut menunjukkan hasil *Confusion Matrix* dari algoritma Random Forest:



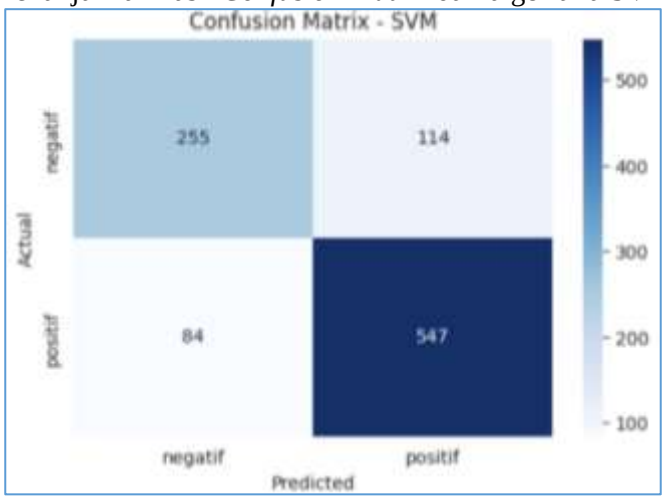
Gambar 6 *Confusion matrix Random Forest*

Berdasarkan Gambar 6, model Random Forest menghasilkan jumlah prediksi benar sebanyak 267 data negatif dan 563 data positif, sementara terdapat 102 data negatif yang salah terprediksi sebagai positif dan 68 data positif yang salah terprediksi sebagai negatif. Temuan ini mengindikasikan bahwa model Random Forest memiliki keseimbangan yang baik antara

 ISSN:2302-8149
e-ISSN:2540-9719

pengenalan sentimen positif dan negatif. Jumlah kesalahan prediksi yang relatif kecil menandakan bahwa model ini stabil dan konsisten dalam membedakan kelas sentimen. Secara keseluruhan, Random Forest menunjukkan performa terbaik di antara ketiga model karena mampu mengenali pola kata yang kompleks dengan lebih efektif, terutama pada data yang memiliki ekspresi emosional beragam.

3. Support Vector Machine (SVM)
   Gambar 7 berikut menunjukkan hasil *Confusion Matrix* dari algoritma SVM:



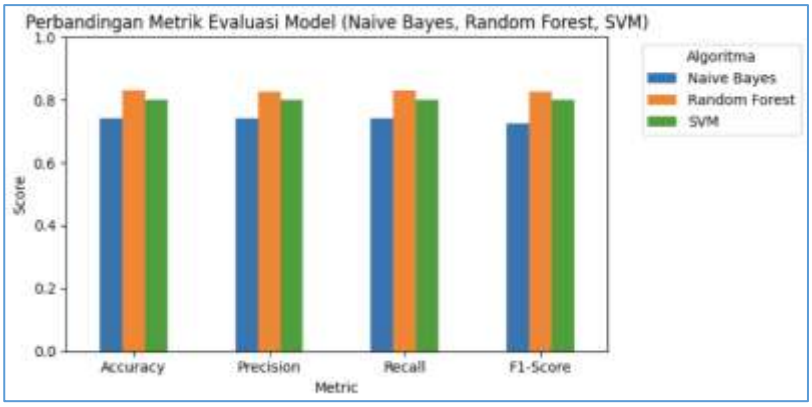
Gambar 7 *Confusion matrix* SVM

Berdasarkan Gambar 7, model SVM menghasilkan jumlah prediksi benar sebanyak 255 data negatif dan 547 data positif, sementara terdapat 114 data negatif yang salah terprediksi sebagai positif dan 84 data positif yang salah terprediksi sebagai negatif. Temuan ini mengindikasikan bahwa model SVM memiliki kemampuan yang sangat baik dalam memisahkan data berdasarkan kelas sentimen, namun masih sedikit di bawah performa Random Forest. Meskipun demikian, SVM tetap menunjukkan performa yang stabil dan akurasi tinggi, sehingga dapat dianggap sebagai model kompetitif kedua setelah Random Forest dalam klasifikasi sentimen pada ulasan aplikasi Weverse.

4. Perbandingan Metrik Evaluasi Antar Model
   Untuk melihat performa ketiga algoritma secara keseluruhan, dilakukan perbandingan terhadap nilai metrik evaluasi sebagaimana ditunjukkan pada Tabel 13 berikut.

Tabel 13 Perbandingan nilai metrik evaluasi model

Metric	Naïve Bayes	Random Forest	SVM
Akurasi	0.742000	0.830000	0.802000
Presisi	0.741688	0.828314	0.799740
Recall	0.742000	0.830000	0.802000
Skor F1	0.725046	0.828125	0.800104



Gambar 8 Perbandingan metrik evaluasi model

 ISSN:2302-8149
e-ISSN:2540-9719

Gambar 8 menampilkan perbandingan nilai metrik dalam bentuk grafik batang. Berdasarkan grafik tersebut, dapat dilihat bahwa algoritma Random Forest memiliki nilai akurasi, presisi, *recall*, dan skor F1 tertinggi dibandingkan Naïve Bayes dan SVM. Temuan tersebut mengindikasikan bahwa Random Forest mampu melakukan generalisasi dengan lebih baik dibandingkan dua algoritma lainnya dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Weverse.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa Random Forest merupakan algoritma dengan kinerja paling optimal, diikuti oleh SVM, sedangkan Naïve Bayes memiliki performa terendah namun masih cukup efektif dalam melakukan klasifikasi sentimen dasar. Hasil ini sejalan dengan penelitian yang dilakukan oleh Mola *et al*. [29] dalam analisis sentimen aplikasi Halo BCA, yang juga menunjukkan bahwa algoritma Random Forest memperoleh akurasi tertinggi dibandingkan SVM dan Naïve Bayes dalam klasifikasi sentimen berbasis pembobotan TF-IDF, dengan tingkat akurasi mencapai 91,28%.

Keunikan penelitian ini terletak pada penggunaan dataset ulasan aplikasi Weverse, yang hingga saat ini masih jarang dijadikan objek dalam studi analisis sentimen. Selain itu, penelitian ini menerapkan pembobotan TF-IDF untuk membandingkan tiga algoritma utama analisis sentimen, yaitu Naïve Bayes, Random Forest, dan SVM. Pendekatan ini memungkinkan evaluasi yang lebih komprehensif terhadap efektivitas ketiga algoritma dalam memproses data teks berdimensi tinggi.

Karakteristik ulasan Weverse yang telah melalui *preprocessing*—seperti adanya kosakata berulang, penggunaan bahasa campuran Indonesia–Inggris, frasa fandom yang tidak baku, serta ekspresi emosional yang cukup kuat—menciptakan representasi fitur TF-IDF yang sangat sparse dan bervariasi. Pola ini lebih mudah ditangani oleh model berbasis pohon.

Keunggulan algoritma Random Forest dalam penelitian ini disebabkan oleh kemampuannya dalam menggabungkan hasil dari banyak pohon keputusan (*decision tree*) sehingga mampu mengenali pola non-linear dan menangani fitur berjumlah besar yang dihasilkan oleh representasi TF-IDF, sebagaimana tercermin dari nilai akurasi, presisi, *recall*, dan skor F1 yang lebih tinggi dibandingkan algoritma pembanding pada Tabel 13 dan Gambar 8. Selain itu, Random Forest lebih toleran terhadap noise dan variasi bahasa pengguna, membuatnya lebih stabil ketika memproses teks fandom Weverse yang cenderung informal dan tidak konsisten.

Sementara itu, SVM tetap menunjukkan performa yang kompetitif dalam memisahkan kelas positif dan negatif melalui penentuan *hyperplane* optimal, sedangkan Naïve Bayes cenderung lebih sensitif terhadap asumsi independensi antarfitur sehingga akurasinya tidak setinggi dua model lainnya.

5 Kesimpulan

Dari keseluruhan proses penelitian, dapat disimpulkan bahwa analisis sentimen terhadap ulasan pengguna aplikasi Weverse dengan menerapkan tiga algoritma klasifikasi, yaitu Naïve Bayes, Random Forest, dan Support Vector Machine (SVM), berhasil mengindikasikan perbandingan kinerja yang signifikan. Setelah melalui tahap *preprocessing* dan pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF), hasil evaluasi mengindikasikan bahwa Random Forest merupakan algoritma dengan performa terbaik dengan nilai akurasi mencapai 83%, presisi 82,83%, *recall* 83%, dan skor F1 82,81%. Model SVM menempati urutan kedua dengan akurasi 80,2%, sedangkan Naïve Bayes memperoleh akurasi 74,2%. Random Forest unggul karena kemampuannya dalam menggabungkan hasil dari berbagai pohon keputusan (*decision tree*), sehingga mampu menghasilkan prediksi yang lebih stabil dan akurat. Model ini juga terbukti lebih efektif dalam menangani fitur berjumlah besar yang dihasilkan oleh representasi TF-IDF serta lebih toleran terhadap variasi bahasa dan ekspresi pengguna pada teks ulasan. Keunikan penelitian ini terletak pada penggunaan dataset ulasan aplikasi Weverse, yang hingga kini masih jarang dijadikan objek penelitian analisis sentimen, serta pada pembandingan langsung tiga algoritma populer dalam konteks analisis sentimen berbasis TF-IDF. Meskipun hasil yang diperoleh menunjukkan kinerja model yang baik, penelitian ini masih dibatasi oleh sumber data yang hanya terfokus pada ulasan dari Google Play Store, penggunaan dua kelas sentimen yaitu positif dan negatif, serta proses pelabelan otomatis menggunakan VADER yang tidak selalu akurat untuk teks informal atau campur kode. Pada studi berikutnya, peneliti dianjurkan untuk memperluas sumber data dengan melibatkan ulasan yang berasal dari beragam platform dan

 ISSN:2302-8149
e-ISSN:2540-9719

menambahkan kelas sentimen netral agar hasil analisis lebih komprehensif. Selain itu, penerapan pendekatan berbasis *deep learning* seperti LSTM atau BERT memperkuat pemahaman model terhadap konteks bahasa alami yang bersifat kompleks, serta mengintegrasikan analisis emosi (*emotion analysis*) untuk memberikan wawasan yang lebih mendalam terhadap persepsi pengguna.

Referensi

[1] L. Z. Damaika, "Apa Itu *Weverse*? *Platform* Populer di Kalangan Fans KPop," IDN Times. Accessed: Oct. 09, 2025. [Online]. Available: https://www.idntimes.com/tech/trend/weverse-adalah-00-2k3g7-ms22j5

[2] E. P. Nirwandani, Indriati, and R. C. Wihandika, "Analisis Sentimen pada Ulasan Pengguna Aplikasi Mandiri Online menggunakan Metode *Modified Term Frequency Scheme* dan *Naïve Bayes*," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Vol. 5, No. 3, pp. 1039–1047, 2021, [Online]. Available: http://j-ptiik.ub.ac.id

[3] M. F. Rahmatullah, P. L. L. Belluano, and H. Darwis, "Analisis Sentimen *Review* Aplikasi di *Google PlaStore* menggunakan *Random Forest*," *LINIER: Literatur Informatika dan Komputer*, Vol. 2, No. 3, pp. 380–389, 2025, DOI: 10.33096/linier.v2i3.3149.

[4] A. H. Nurdy, A. Rahim, and Arbansyah, "Analisis Sentimen Ulasan Game *Stumble Guys* pada *Playstore* menggunakan *Algoritma Naïve Bayes*," *Teknika*, Vol. 13, No. 3, pp. 388–395, 2024, DOI: 10.34148/teknika.v13i3.993.

[5] Rahmat, A. Rahim, and Arbansyah, "Perbandingan Metode *Naïve Bayes* dan *Support Vector Machine* untuk Analisis Sentimen pada Ulasan Pengguna Aplikasi Alibaba di *Google Play Store*," *Jurnal Mahasiswa Teknik Informatika)*, Vol. 9, No. 2, 2025.

[6] N. Fitriyah, B. Warsito, and D. A. I. Maruddani, "Analisis Sentimen Gojek pada Media Sosial *Twitter* dengan Klasifikasi *Support Vector Machine (SVM)*," *JURNAL GAUSSIAN*, Vol. 9, No. 3, pp. 376–390, 2020, [Online]. Available: https://ejournal3.undip.ac.id/index.php/gaussian/

[7] P. Anggraini and Winarsih, "Komparasi *Naïve Bayes*, *Support Vector Machine*, dan *Random Forest* dalam Analisis Sentimen Aplikasi Shopee di *Google Play Store*," *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 9, No. 3, pp. 4451–4457, 2025.

[8] J. Celvin and S. P. Barus, "Analisis Sentimen Ulasan Pengguna Aplikasi *Famiapps* menggunakan *Naïve Bayes, SVM,* dan *Random Forest*," *Jurnal PETISI*, Vol. 06, No. 02, 2025.

[9] G. Jeffson Sagala and Y. T. Samuel, "*Sentiment Analysis on ChatGPT App Reviews on Google Play Store using Random Forest Algorithm, Support Vector Machine and Naïve Bayes*," *International Journal of Engineering Business and Social Science*, Vol. 2, No. 04, pp. 2980–4272, 2024, [Online]. Available: https://ijebss.ph/index.php/ijebss

[10] N. R. Setiawan and E. R. Kaburuan, "Sentimen Analisis Review Aplikasi Digital Korlantas pada *Google Play Store* menggunakan Metode *SVM*," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, Vol. 12, No. 1, pp. 105–116, Mar. 2023, DOI: 10.32736/sisfokom.v12i1.1614.

[11] D. A. Tarigan, Z. Situmorang, and R. Rosnelly, "Analisis Sentimen Aplikasi *Playstore* Sirekap 2024 Pasca Pilpres dengan Perbandingan Metode *Support Vector Machine (SVM), Naïve Bayes Classifier* dan *Random Forest*," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, Vol. 11, No. 3, pp. 661–670, 2025.

[12] N. D. Kurniawan, P. R. Ferdian, and N. Hidayati, "Analisis Sentimen *Algoritma Naïve Bayes, Support Vector Machi*ne, dan *Random Forest* pada Ulasan Aplikasi Ajaib," *Jurnal Nasional Teknologi dan Sistem Informasi*, Vol. 11, No. 1, pp. 87–97, 2025, DOI: 10.25077/teknosi.v11i1.2025.87-97.

[13] I. Irawan, Wardianto, Mh. Wathan, and M. B. Prayogi, "Studi Perbandingan: *Algoritma Random Forest, Naive Bayes* dan *Support Vector Machine* dalam Analisis Sentimen pada Aplikasi *Capcut* di *Google Play Store*," *Jurnal Pengembangan Sistem Informasi dan Informatika*, Vol. 5, No. 4, pp. 152–164, 2024.

[14] Y. A. Mustofa and I. S. K. Idris, "Pendekatan *Ensemble* pada Analisis Sentimen Ulasan Aplikasi *Google Play Store*," *Jambura Journal of Electrical and Electronics Engineering*, Vol. 6, No. 2, pp. 181–188, 2024.

[15] A. Amelia, L. N. Hayati, and H. Darwis, "Analisis Sentimen Masyarakat terhadap Sistem Pembayaran Mypertamina dengan Metode *Random Forest, SVM*, dan *Naïve Bayes*," *Literatur Informatika & Komputer*, Vol. 1, No. 1, pp. 28–44, 2024, DOI: 10.33096/linier.v1i1.2269.

[16] D. S. A. Rizki, M. S. Khabib, N. Rahmayuna, and V. Gayuh Utomo, "Klasifikasi Sentimen Ulasan Pengguna Aplikasi Layanan Publik *Google Play Store* menggunakan NLP dan ML," *Jurnal TEKNO KOMPAK*, Vol. 20, No. 1, pp. 51–64, 2025.

[17] D. Nurwahidah, G. Dwilestari, N. D. Nuris, and R. Narasati, "Analisis Sentimen Data Ulasan Pengguna Aplikasi *Google* Kelas pada *Google Play Store* menggunakan *Algoritma Naïve Bayes*," *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 7, No. 6, pp. 3673–3678, 2023.

[18] D. Rudini, D. G. Purnama, and A. A. Khan, "Penggunaan Teknik *Web Scraping* dalam Aplikasi Pengambilan Data dari *Google Maps* untuk menunjang *Digital Marketing*," *Lentera: Multidisciplinary Studies*, Vol. 2, No. 1, 2023, [Online]. Available: https://lentera.publikasiku.id/index.php

[19] R. Hasanah *et al.*, "*Play Store Data Scrapping and Preprocessing done as Sentiment Analysis Material*," *Indonesian Journal of Modern Science and Technology (IJMST)*, Vol. 1, No. 1, pp. 16–21, 2025, [Online]. Available: https://journal.abhinaya.co.id/index.php/IJMST

[20] A. A. A. Daniswara and I. K. D. Nuryana, "Data *Preprocessing* Pola pada Penilaian Mahasiswa Program Profesi Guru," *Journal of Informatics and Computer Science*, Vol. 05, No. 01, 2023.

[21] A. B. Muzayyanah, R. E. Pawening, and Z. Arifin, "Analisis Sentimen pada Ulasan Aplikasi Ehadrah di *Google Playstore* menggunakan *Support Vector Machine*," *Idealis: Indonesia Journal Information System*, Vol. 7, No. 2, pp. 258–266, 2024.

[22] F. Nufairi, N. Pratiwi, and F. Herlando, "Analisis Sentimen pada Ulasan Aplikasi *Threads* di *Google Play Store* menggunakan *Algoritma Support Vector Machine*," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, Vol. 9, No. 1, pp. 339–348, 2024, DOI: 10.29100/jipi.v9i1.4929.

[23] P. Elisa and A. R. Isnain, "*Comparison of Random Forest, Support Vector Machine and Naive Bayes Algorithms to Analyze Sentiment Towards Mental Health Stigma*," *Jurnal Teknik Informatika (JUTIF)*, Vol. 5, No. 1, pp. 321–329, 2024, DOI: 10.52436/1.jutif.2024.5.1.1817.

[24] U. Khairani, V. Mutiawani, and H. Ahmadian, "Pengaruh Tahapan *Preprocessing* terhadap Model *Indobert* dan *Indobertweet* untuk mendeteksi Emosi pada Komentar Akun Berita Instagram," *Jurnal Teknologi Informasi dan Ilmu Komputer*, Vol. 11, No. 4, pp. 887–894, 2024, DOI: 10.25126/jtiik.1148315.

[25] R. R. Pratama and R. R. Suryono, "*Performance Comparison of Naive Bayes, Support Vector Machine and Random Forest Algorithms for Apple Vision Pro Sentiment Analysis*," *Jurnal Teknik Informatika (Jutif)*, Vol. 6, No. 1, pp. 31–40, Feb. 2025, DOI: 10.52436/1.jutif.2025.6.1.4035.

[26] D. Shafira Akbar Rizki, M. Syaiful Khabib, N. Rahmayuna, and V. Gayuh Utomo, "Klasifikasi Sentimen Ulasan Pengguna Aplikasi Layanan Publik *Google Play Store* menggunakan NLP dan ML," *Jurnal TEKNO KOMPAK*, Vol. 20, No. 1, pp. 51–64, 2025.

[27] A. M. Chaid, Z. A. Abdulrazzaq, R. N. Sadoon, and M. A. Aljabery, "*Comparative Analysis of Innovative Machine Learning Algorithms: Advancements in Natural Language Processing*," *Journal of Information Systems Engineering and Management*, Vol. 10, No. 14s, pp. 648–668, 2025, DOI: 10.52783/jisem.v10i14s.2373.

[28] H. Wang, L. Zhang, K. Yin, H. Luo, and J. Li, "*Landslide Identification using Machine Learning*," *Geoscience Frontiers*, Vol. 12, pp. 351–364, 2021, DOI: 10.1016/j.gsf.2020.02.012.

[29] 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, 2024, DOI: 10.52972/hoaq.vol15no2.p69-79.