

PERBANDINGAN KINERJA NAÏVE BAYES DAN SVM DALAM ANALISIS SENTIMEN ULASAN FREE FIRE DI PLAY STORE

La Ode Alyandi

Teknologi Informasi, Institut Teknologi dan Bisnis Muhammadiyah Wakatobi
Jln. Adhyaksa. No. 37 Desa Numana Kec. Wangi-wangi Selatan Kab. Wakatobi
laodealyandi@gmail.com

ABSTRAK

Industri game mobile berkembang pesat, dengan Free Fire sebagai salah satu game yang memiliki banyak ulasan di Google Play Store. Analisis sentimen terhadap ulasan pengguna menjadi penting untuk memahami kepuasan serta kendala yang dihadapi pemain. Penelitian ini membandingkan kinerja algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen ulasan Free Fire. Data dikumpulkan menggunakan teknik web scraping dengan Google Colab dan Python, kemudian melalui tahap pra-pemrosesan seperti pembersihan teks, tokenisasi, stopword removal, stemming, serta ekstraksi fitur menggunakan TF-IDF. Evaluasi dilakukan dengan K-Fold Cross Validation (K=10) menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa SVM memiliki performa lebih baik dengan akurasi 76%, sementara Naïve Bayes hanya mencapai 70%. Dengan demikian, SVM direkomendasikan sebagai metode yang lebih optimal dalam klasifikasi sentimen ulasan pengguna Free Fire. Penelitian ini memberikan wawasan bagi pengembang dalam memahami opini pemain dan meningkatkan kualitas layanan dalam game.

Kata kunci : Analisis Sentimen, Naïve Bayes, Support Vector Machine, Free Fire, Google Play Store

1. PENDAHULUAN

Industri game mobile berkembang pesat dengan jutaan pengguna aktif di seluruh dunia. Salah satu game populer adalah Free Fire, yang dikembangkan oleh Garena dan banyak diulas di Google Play Store [1]. Ulasan ini mencerminkan pengalaman pemain, baik positif maupun negatif, serta memberikan wawasan bagi pengembang untuk meningkatkan kualitas permainan. Beragam faktor seperti gameplay, grafis, kestabilan server, serta fitur dalam game menjadi aspek yang sering disoroti dalam ulasan pengguna [2].

Analisis sentimen terhadap ulasan pengguna penting untuk memahami tingkat kepuasan serta permasalahan yang dihadapi pemain. Ulasan mencakup aspek teknis seperti stabilitas server dan performa game, serta aspek emosional seperti kepuasan terhadap fitur dan event dalam game [3]. Ulasan yang diberikan pengguna dapat membantu pengembang dalam mengidentifikasi permasalahan serta menyesuaikan strategi pengembangan agar lebih sesuai dengan kebutuhan pemain. Oleh karena itu, diperlukan metode yang efektif untuk mengolah ulasan ini agar dapat diinterpretasikan dengan akurat dan menghasilkan informasi yang bermanfaat [4].

Seiring meningkatnya partisipasi pengguna dalam memberikan ulasan pada platform digital seperti Google Play Store, ulasan tersebut tidak hanya menjadi bentuk ekspresi pengalaman pribadi, tetapi juga menjadi sumber data yang sangat bernilai bagi pengembang [5]. Dalam konteks game Free Fire, ulasan pengguna mencerminkan kepuasan, kritik, hingga harapan terhadap fitur-fitur permainan. Sayangnya, ulasan yang jumlahnya ribuan ini seringkali tidak dianalisis secara sistematis, sehingga potensi informasinya tidak tergali secara maksimal. Di

sinilah analisis sentimen menjadi penting sebagai pendekatan untuk menggali opini dan emosi yang terkandung dalam ulasan, dan mengubahnya menjadi informasi terstruktur [6]. Urgensi penelitian ini terletak pada kebutuhan untuk menyediakan metode yang efisien dalam memahami persepsi pengguna secara keseluruhan, agar pengembang dapat mengambil keputusan berbasis data yang lebih tepat sasaran. Selain itu, hasil dari klasifikasi sentimen dapat menjadi acuan dalam perbaikan fitur game dan pengembangan strategi bisnis yang lebih responsif terhadap kebutuhan komunitas pemain.

Machine learning menjadi pendekatan yang banyak digunakan dalam analisis sentimen karena kemampuannya dalam mengolah data dalam jumlah besar secara efisien. Dalam penelitian ini, dua algoritma populer, Naïve Bayes dan Support Vector Machine (SVM), dibandingkan untuk mengklasifikasikan sentimen pengguna terhadap Free Fire. Naïve Bayes, berbasis probabilistik, mampu menangani data dengan cepat dan efisien, serta sering digunakan dalam klasifikasi teks. Sementara itu, SVM dikenal karena kemampuannya dalam memisahkan kelas data secara optimal melalui hyperplane, sehingga menghasilkan klasifikasi yang lebih akurat dalam berbagai kasus analisis sentiment [7].

Penelitian sebelumnya menemukan bahwa menggunakan algoritma Support Vector Machine (SVM) dan Multinomial Naive Bayes (MNB) untuk menganalisis komentar masyarakat tentang pelayanan publik Pemerintah DKI Jakarta memberikan hasil yang positif. Algoritma SVM menunjukkan performa terbaik dengan tingkat akurasi 82%, sementara MNB menghasilkan akurasi 80%. Untuk SVM, pengaturan parameter yang paling efektif adalah $C=100$ dan $\gamma=0.1$. Sementara untuk MNB, pengaturan

terbaiknya adalah $\alpha=2.0$ dan $\text{fit_prior}=\text{True}$ [8]. Penelitian lainnya juga menggunakan algoritma naïve bayes dan svm dalam meneliti sentiment boyband BTS pada platform twitter dan berhasil mendapatkan dalam mengidentifikasi dan mengkategorikan sentimen dengan tingkat akurasi yang baik. Dari total 1648 komentar yang dianalisis, ditemukan bahwa 78,8% berasal dari sentimen positif, sedangkan 21,2% menunjukkan sentimen negatif. Algoritma Support Vector Machine menghasilkan akurasi tertinggi sebesar 81%, yang mengindikasikan bahwa penggemar BTS lebih banyak memberikan reaksi positif di platform media sosial [9].

Penelitian sebelumnya juga menerapkan algoritma Naive Bayes untuk menganalisis sentimen ulasan pengguna game Honor of Kings di Playstore berhasil menghasilkan model dengan akurasi sebesar 71%. Dari 500 ulasan yang dikaji, terdapat distribusi sentimen yang hampir seimbang, yaitu 50,5% positif dan 49,5% negatif. Kata-kata yang paling sering muncul dalam ulasan positif merefleksikan kepuasan pengguna, seperti "bagus" dan "seru," sedangkan ulasan negatif sering kali berisi keluhan tentang fitur dan masalah teknis. Temuan penelitian ini memberikan informasi yang berharga bagi pengembang game untuk meningkatkan kualitas produk mereka dan lebih responsif terhadap kebutuhan serta harapan pengguna, sambil menyoroti efektivitas metode Naive Bayes dalam klasifikasi sentiment [10].

Berdasarkan penelitian-penelitian sebelumnya, penelitian ini bertujuan untuk menentukan algoritma yang lebih efektif dalam analisis sentimen ulasan Free Fire di Play Store dengan menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan f1-score. Dengan membandingkan kinerja kedua algoritma ini, yakni algoritma naïve bayes dan SVM diharapkan dapat diperoleh pemahaman yang lebih mendalam mengenai pola sentimen pengguna terhadap game Free Fire. Hasil penelitian ini juga diharapkan dapat memberikan rekomendasi bagi pengembang dalam meningkatkan kualitas layanan berdasarkan opini pemain, sehingga Free Fire dapat terus berkembang dan memberikan pengalaman bermain yang lebih baik bagi penggunanya.

2. TINJAUAN PUSTAKA

2.1. Sentiment

Sentimen merupakan ekspresi perasaan atau pandangan seseorang terhadap suatu topik, baik dalam bentuk opini positif, negatif, maupun netral [11]. Dalam analisis sentimen, konsep ini digunakan untuk mengidentifikasi dan memahami bagaimana individu mengekspresikan pendapat mereka dalam teks, seperti ulasan produk, komentar media sosial, atau diskusi online [12].

Dalam konteks penelitian ini, sentimen merujuk pada opini pemain terhadap game Free Fire yang tercermin dalam ulasan di Google Play Store. Analisis

sentimen bertujuan untuk mengelompokkan ulasan tersebut ke dalam kategori tertentu berdasarkan kecenderungan emosionalnya [13]. Dengan menggunakan algoritma machine learning seperti Naïve Bayes dan Support Vector Machine (SVM), proses klasifikasi dapat dilakukan secara otomatis dengan tingkat akurasi yang lebih tinggi.

2.2. PlayStore dan Free Fire

Google Play Store adalah platform distribusi aplikasi digital yang dikembangkan oleh Google untuk perangkat berbasis Android. Platform ini menyediakan berbagai aplikasi, termasuk game, yang dapat diunduh dan dinilai oleh pengguna. Salah satu game yang memperoleh popularitas tinggi di Play Store adalah Free Fire, sebuah game battle royale yang dikembangkan oleh Garena [14].

Free Fire menawarkan mekanisme permainan yang kompetitif, di mana pemain bertahan hidup dalam arena pertempuran yang terus menyusut. Game ini memiliki berbagai fitur seperti karakter dengan kemampuan khusus, mode permainan beragam, serta pembaruan berkala untuk meningkatkan pengalaman bermain [15]. Ulasan pengguna di Google Play Store menjadi sumber informasi penting mengenai kepuasan pemain, mencakup aspek seperti performa game, stabilitas server, keseimbangan permainan, dan kualitas pembaruan fitur. Analisis terhadap ulasan ini dapat memberikan wawasan bagi pengembang dalam meningkatkan kualitas layanan serta menyesuaikan strategi pengembangan agar lebih sesuai dengan kebutuhan komunitas pemain [16].

2.3. Naïve Bayes dan Svm

Naïve Bayes merupakan algoritma berbasis probabilitas yang menggunakan Teorema Bayes untuk menentukan kemungkinan suatu data termasuk dalam kategori tertentu. Metode ini mengasumsikan bahwa setiap kata dalam teks bersifat independen terhadap kata lainnya, sehingga perhitungannya menjadi lebih sederhana dan cepat [17]. Naïve Bayes sering digunakan dalam analisis sentimen karena kemampuannya dalam mengolah data dalam jumlah besar dengan efisiensi tinggi. Selain itu, algoritma ini tetap dapat memberikan hasil yang cukup baik meskipun terdapat keterbatasan dalam jumlah data pelatihan [18].

Support Vector Machine (SVM) bekerja dengan mencari garis pemisah atau hyperplane yang paling optimal untuk membedakan dua kelompok data. Algoritma ini sangat efektif dalam menangani data dengan pola yang kompleks dan sering digunakan dalam berbagai jenis klasifikasi, termasuk analisis sentiment [19]. Dengan pendekatan berbasis vektor, SVM mampu mengenali pola dalam teks dan memisahkan sentimen positif, negatif, atau netral dengan akurasi yang lebih tinggi, terutama ketika data memiliki distribusi yang tidak merata [20].

3. METODE PENELITIAN

3.1. Metode Analisis

Langkah pertama yang dilakukan dalam penelitian ini adalah memahami konsep dasar analisis sentimen, termasuk tujuan, metode, serta penerapannya dalam berbagai bidang. Setelah memperoleh pemahaman yang cukup, tahap selanjutnya adalah melakukan pengumpulan data yang relevan, yakni berupa komentar atau ulasan pengguna yang mengandung ekspresi sentimen. Data ini kemudian digunakan sebagai bahan utama dalam proses pelatihan dan pengujian model klasifikasi yang akan diterapkan.

3.2. Analisis Sentimen

Analisis sentimen merupakan teknik yang digunakan untuk mengekstrak opini atau perasaan seseorang terhadap suatu isu atau kejadian. Teknik ini banyak *dimanfaatkan* untuk memahami pandangan publik terhadap suatu topik, menilai kepuasan pengguna terhadap layanan, mengevaluasi kebijakan, mengidentifikasi tindakan cyberbullying, memprediksi tren harga saham, hingga menganalisis persaingan bisnis melalui data berbasis teks. Tugas utama dalam analisis sentimen adalah mengklasifikasikan teks berdasarkan polaritasnya, baik dalam skala dokumen, kalimat, maupun opini. Polaritas ini merujuk pada apakah suatu teks memiliki kecenderungan positif atau negatif.

3.3. Pengumpulan Data

Langkah berikutnya adalah memperoleh data dari ulasan aplikasi di Play Store dengan menggunakan teknik web scraping. Web scraping adalah metode otomatis untuk mengekstrak data dalam jumlah besar dari suatu situs web. Dalam penelitian ini, data yang diambil berupa ulasan pengguna terhadap game Free Fire di Play Store. Setelah data mentah diperoleh, langkah selanjutnya adalah memproses data agar siap untuk dianalisis. Kualitas data menjadi faktor utama dalam tahap ini karena akan sangat berpengaruh terhadap akurasi serta validitas hasil analisis di tahap selanjutnya.

3.4. Pra-Pemrosesan

Data yang telah dikumpulkan tidak dapat langsung digunakan untuk analisis. Oleh karena itu, diperlukan serangkaian langkah pemrosesan awal agar data lebih terstruktur dan mudah dipahami oleh komputer. Langkah-langkah pra-pemrosesan data meliputi:

- Pembersihan Data (Cleaning): Data dibersihkan dari elemen-elemen yang tidak diperlukan, seperti tanda baca, simbol khusus, atau karakter yang tidak memiliki makna signifikan. Langkah ini bertujuan untuk memastikan data dalam kondisi yang lebih optimal untuk analisis.
- Normalisasi Huruf (Case Folding): Seluruh teks dikonversi menjadi huruf kecil untuk menghindari perbedaan akibat penggunaan huruf

kapital. Dengan cara ini, kata-kata seperti "Free Fire" dan "free fire" akan diperlakukan sebagai entitas yang sama.

- Tokenisasi (Tokenizing): Proses ini bertujuan untuk memecah teks menjadi unit-unit kata yang lebih kecil agar lebih mudah dianalisis secara terstruktur.
- Penghapusan Kata Umum (Stopword Removal): Kata-kata yang sering muncul tetapi memiliki kontribusi informasi rendah, seperti "yang", "atau", dan "dan", dihapus untuk meningkatkan fokus analisis pada kata-kata yang lebih bermakna.
- Stemming: Teknik ini mengubah kata menjadi bentuk dasarnya untuk menyamakan variasi kata yang memiliki makna serupa. Sebagai contoh, kata "bermain" akan diubah menjadi "main" agar analisis lebih akurat.

3.5. Ekstraksi Fitur

Setelah tahap pra-pemrosesan selesai, informasi penting dari teks diekstraksi menggunakan teknik berikut:

- Term Frequency-Inverse Document Frequency (TF-IDF): Teknik ini digunakan untuk memberikan bobot pada setiap kata dalam dokumen berdasarkan frekuensi kemunculannya dalam satu dokumen dibandingkan dengan keseluruhan kumpulan dokumen. Dengan metode ini, kata-kata yang memiliki bobot lebih tinggi dianggap lebih relevan dalam analisis.
- Pelabelan (Labeling): Data diberi label atau kategori untuk menentukan apakah suatu ulasan bersifat positif atau negatif. Proses ini penting dalam pembelajaran mesin berbasis klasifikasi (supervised learning). Labelisasi dilakukan berdasarkan ulasan pengguna terhadap game Free Fire di Play Store.

3.6. Pembagian Data

Setelah fitur diekstraksi, data dibagi menjadi beberapa subset untuk melatih dan menguji model. Pada penelitian ini, metode yang digunakan untuk membagi data adalah K-Fold Cross Validation dengan nilai $K = 10$. Pendekatan ini membagi data menjadi 10 bagian, di mana setiap bagian bergantian digunakan sebagai data uji, sementara bagian lainnya digunakan sebagai data latih. Teknik ini bertujuan untuk memastikan setiap bagian data ikut berkontribusi dalam proses pelatihan dan pengujian, sehingga hasil analisis lebih akurat dan stabil.

3.7. Implementasi Model

Setelah data dibagi, algoritma pembelajaran mesin diterapkan untuk mengklasifikasikan ulasan pengguna. Dalam penelitian ini, digunakan dua algoritma, yaitu Naive Bayes dan Support Vector Machine (SVM).

- Naive Bayes merupakan algoritma klasifikasi berbasis probabilitas yang mengasumsikan

bahwa setiap fitur (kata) dalam data bersifat independen satu sama lain. Proses penerapannya meliputi:

- Menghitung probabilitas setiap kata muncul pada masing-masing kelas (positif, negatif, dan netral).
- Menggunakan Teorema Bayes untuk memperkirakan kemungkinan suatu ulasan termasuk dalam kategori tertentu.
- Mengklasifikasikan data uji berdasarkan nilai probabilitas tertinggi yang diperoleh dari model.

Algoritma ini dikenal ringan secara komputasi dan efektif digunakan dalam klasifikasi teks karena kemampuannya memberikan hasil yang cepat dan akurat, bahkan dengan jumlah data besar.

- Support Vector Machine (SVM): SVM bekerja dengan mencari **garis pemisah (hyperplane)** terbaik yang dapat membedakan antar kelas secara optimal. Tahapan implementasi meliputi:

- Mentransformasikan data ulasan ke dalam bentuk vektor fitur melalui teknik seperti TF-IDF.
- Mencari hyperplane dengan margin terbesar antara dua kelas (positif dan negatif) untuk membangun model klasifikasi.
- Menggunakan model tersebut untuk mengklasifikasikan data uji berdasarkan posisi relatif terhadap hyperplane.

SVM sangat efektif dalam menangani data berdimensi tinggi seperti teks, serta mampu menghasilkan akurasi yang tinggi berkat pendekatan margin maksimal yang digunakan.

3.8. Evaluasi Model

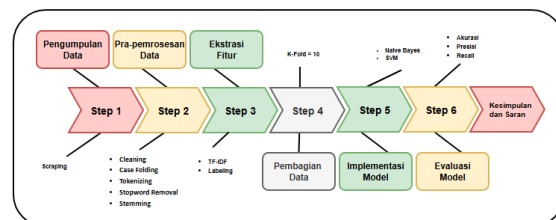
Setelah model selesai dilatih, langkah berikutnya adalah mengevaluasi kinerjanya. Evaluasi dilakukan menggunakan Google Colab dengan bahasa pemrograman Python. Beberapa metrik evaluasi yang digunakan dalam penelitian ini meliputi:

- Akurasi: Menunjukkan sejauh mana model dapat memprediksi sentimen dengan benar dibandingkan dengan total data uji.
- Presisi: Menghitung proporsi data yang diklasifikasikan sebagai positif dan benar-benar merupakan kategori positif.
- Recall: Mengukur seberapa baik model dalam mendeteksi semua data yang termasuk dalam kategori tertentu. Evaluasi ini bertujuan untuk menilai efektivitas model dalam mengklasifikasikan ulasan pengguna game Free Fire dengan tingkat akurasi yang optimal.
- F1-Score: Merupakan rata-rata harmonis dari presisi dan recall, yang digunakan untuk menyeimbangkan keduanya dalam satu metrik. F1-Score memberikan gambaran yang lebih baik tentang kinerja model, terutama jika terdapat kesinambungan kelas dalam dataset.

3.9. Kesimpulan dan Saran

Tahap terakhir dalam penelitian ini adalah menganalisis hasil yang diperoleh dan menarik kesimpulan terkait efektivitas model yang digunakan. Berdasarkan hasil evaluasi, penelitian ini memberikan saran untuk meningkatkan performa model di masa depan atau membahas bagaimana hasil penelitian dapat diterapkan dalam skala yang lebih luas. Kesimpulan ini diharapkan dapat menjadi dasar bagi penelitian lebih lanjut atau implementasi praktis dalam dunia industri.

Gambar 1 menampilkan alur penelitian yang digunakan, dimulai dari tahap awal pengumpulan data hingga proses analisis akhir. Setiap tahapan dalam alur tersebut disusun secara sistematis untuk memberikan pemahaman yang lebih komprehensif terkait topik yang diteliti. Dengan demikian, alur ini membantu menggambarkan keseluruhan proses penelitian.



Gambar 1. Alur Penelitian

Pada Gambar 1, alur penelitian dijelaskan secara sistematis, dimulai dari pengumpulan data ulasan pengguna terhadap game Free Fire di Play Store menggunakan teknik *scraping*. Setelah data diperoleh, dilakukan tahap pra-pemrosesan untuk membersihkan teks, termasuk menghapus karakter yang tidak relevan, mengubah teks menjadi huruf kecil, memisahkan kata, menghapus kata-kata umum, serta mengubah kata ke bentuk dasarnya. Proses ini bertujuan untuk meningkatkan kualitas data agar lebih mudah diolah dalam tahap berikutnya.

Setelah itu dilakukan ekstraksi fitur menggunakan metode *TF-IDF* untuk menentukan bobot kata dalam teks. Data yang telah diproses kemudian dibagi menggunakan teknik *K-Fold Cross Validation* dengan $K=10$, yang memastikan model diuji dengan lebih baik. Pada tahap implementasi, dua algoritma pembelajaran mesin, yaitu **Naive Bayes** dan **Support Vector Machine (SVM)**, digunakan untuk menganalisis sentimen dalam ulasan.

Proses evaluasi dilakukan menggunakan Google Colab untuk mengukur kinerja model berdasarkan pengukuran akurasi, presisi, dan recall, serta f1-score. Dengan pendekatan ini, penelitian diharapkan mampu memberikan gambaran yang lebih akurat mengenai sentimen pemain terhadap game Free Fire serta memberikan wawasan yang berguna bagi pengembang dalam meningkatkan pengalaman pengguna.

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan dengan teknik web scraping menggunakan Google Colab dan bahasa pemrograman Python. Tahap awal dimulai dengan mengakses halaman aplikasi Free Fire di Google Play Store menggunakan kata kunci yang sesuai. Selanjutnya, proses scraping dilakukan untuk mengambil data ulasan pengguna, termasuk teks ulasan, rating, dan informasi relevan lainnya. Data yang diperoleh kemudian diolah lebih lanjut untuk dianalisis guna memahami pola sentimen pemain terhadap game Free Fire.

```
from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'com.dts.freefireth',
    lang='id',
    country='id',
    sort=Sort.NEWEST,
    count=800,
    filter_score_with=None
)
```

Gambar 2. Scapping data

Gambar tersebut menampilkan cuplikan kode Python yang menggunakan pustaka `google_play_scraper` untuk mengambil ulasan aplikasi Free Fire dari Google Play Store. Kode ini memanggil fungsi `reviews()` dengan beberapa parameter penting, seperti `'com.dts.freefireth'`, yang merupakan package name dari aplikasi Free Fire. Parameter `lang='id'` dan `country='id'` memastikan bahwa ulasan yang diambil berasal dari pengguna berbahasa Indonesia dan berada di wilayah Indonesia. Selain itu, ulasan diurutkan berdasarkan yang terbaru dengan `sort=Sort.NEWEST`, dan jumlah ulasan yang dikumpulkan dibatasi hingga 800. Parameter `filter_score_with=None` memungkinkan pengambilan semua ulasan tanpa memfilter berdasarkan skor rating tertentu. Kode ini bertujuan untuk mengumpulkan data ulasan yang nantinya dapat dianalisis lebih lanjut, misalnya dalam penelitian analisis sentimen.

Tabel 1. Data hasil scrapping

Username	Score	At	Content
aqil ziyad	1	19/02/2025 17:38	kembali n cs elit sg2 onlyðŸ˜ˆ
Gideon Pananggun g	1	19/02/2025 17:31	â€½(‘Đ’İ¼)/
Taufik Kurrahman	5	19/02/2025 17:30	mantap
Sulastri We	5	19/02/2025 17:18	peak old nya harus di adakan lah nih
Sri Rahayu	5	19/02/2025 17:16	bagi aku terus hoki
.....			
Sakrawati Sakrawati	1	19/02/2025 05:57	Saya sangat kecewa kepada free

Username	Score	At	Content
			fire,karena telah mengganti menjadi desa Konoha,kembali kan suasa dulu please saya kangen

Tabel 1 menyajikan hasil scraping dari Google Play Store yang berisi berbagai informasi penting, seperti nama pengguna, rating yang diberikan, waktu ulasan, serta isi komentar.

4.2. Pra-Proses Data

Setelah data dikumpulkan, langkah selanjutnya adalah pra-pemrosesan untuk membersihkan dan menyusun data agar lebih mudah dianalisis. Tabel 2 menampilkan hasil dari beberapa tahap penting, seperti Cleaning (menghapus karakter tidak perlu), Case Folding (mengubah teks menjadi huruf kecil), Tokenizing (memecah teks menjadi kata-kata), Stopword Removal (menghapus kata umum yang kurang relevan), dan Stemming (mengubah kata ke bentuk dasar). Tahapan ini membantu meningkatkan akurasi dalam analisis sentimen.

Tabel 2. Hasil pra-proses data

Sebelum Pra-Pemrosesan	
Review	Saya sangat kecewa kepada free fire,karena telah mengganti menjadi desa Konoha,kembali kan suasa dulu please saya kangen
Sesudah Pra-Pemrosesan	
Cleaning	Saya sangat kecewa kepada free fire,karena telah mengganti menjadi desa Konoha,kembali kan suasa dulu please saya kangen
Case Folding	saya sangat kecewa kepada free fire,karena telah mengganti menjadi desa konoha,kembali kan suasa dulu please saya kangen
Tokenization	['saya', 'sangat', 'kecewa', 'kepada', 'free', 'fire', 'karena', 'telah', 'mengganti', 'menjadi', 'desa', 'konoha', 'kembali', 'suasana', 'dulu', 'please', 'saya', 'kangen']
Stopword Removal	['kecewa', 'free', 'fire', 'karena', 'ganti', 'desa', 'konoha', 'kembali', 'suasa', 'please', 'kangen']
Stemming	kecewa free fire karena ganti desa konoha kembali suasa please kangen

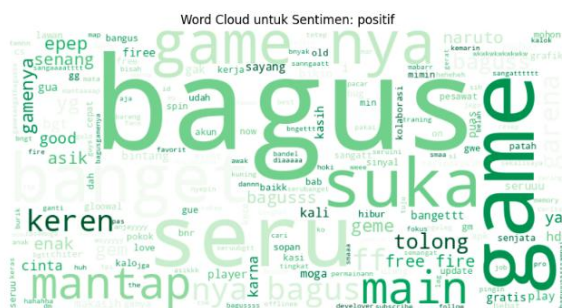
Tabel 2 menjelaskan secara rinci setiap tahap dalam pra-proses data, menunjukkan bagaimana data yang semula tidak terstruktur diubah menjadi lebih rapi dan konsisten. Proses ini mencakup berbagai langkah untuk membersihkan data dari elemen yang tidak diperlukan, seperti duplikasi, data kosong, dan inkonsistensi. Dengan tahapan ini, data mentah yang awalnya sulit dipahami menjadi lebih terorganisir, lebih mudah dianalisis, dan lebih akurat saat diproses oleh komputer.

Word cloud dari pelabelan dapat di lihat pada gambar 5,6, dan 7 yang terlihat jelas bahwa data telah di labeli netral, positif, dan negatif.



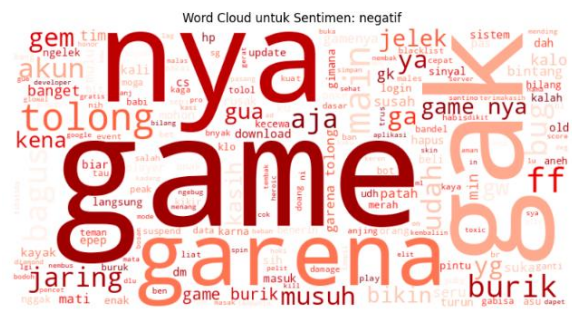
Gambar 5, Label netral

Gambar 5 menampilkan word cloud yang dihasilkan dari komentar dengan sentimen netral. Ukuran kata yang lebih besar menunjukkan frekuensi kemunculan yang lebih tinggi di dalam teks. Beberapa kata yang menonjol, seperti "game", "seru", dan "main", menunjukkan bahwa komentar netral seringkali membahas pengalaman bermain secara umum, tanpa muatan emosi positif atau negatif yang kuat. Kata-kata seperti "bagus", "tolong", dan "burik" juga muncul, menandakan pengguna kadang memberikan masukan atau opini yang tidak sepenuhnya positif maupun negatif. Dengan menggunakan word cloud ini, kita dapat lebih mudah mengenali topik yang paling sering dibicarakan dalam komentar netral.



Gambar 6. Label Positif

Gambar 6 menampilkan word cloud dari komentar yang memiliki sentimen positif. Kata-kata yang lebih besar menunjukkan frekuensi kemunculan yang lebih tinggi dalam ulasan pengguna. Kata-kata seperti "bagus", "game", "suka", "seru", dan "mantap" mendominasi, mencerminkan apresiasi dan kepuasan pemain terhadap game. Selain itu, kata "keren", "asik", dan "Senang" juga muncul, yang semakin menguatkan bahwa banyak komentar memberikan ulasan positif. Word cloud ini membantu dalam memahami bagaimana pemain mengekspresikan pengalaman positif mereka dan aspek apa saja yang mereka sukai dari game tersebut.



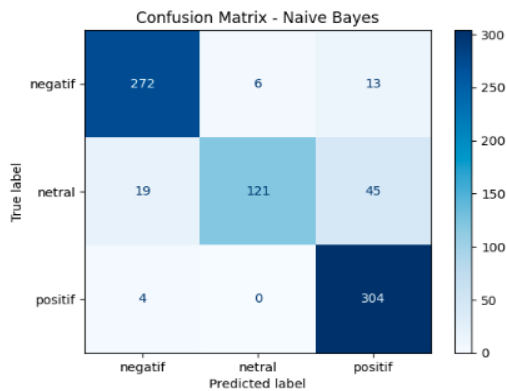
Gambar 7. Label negative

Gambar ini menampilkan word cloud yang dihasilkan dari komentar dengan sentimen negatif. Kata-kata yang berukuran lebih besar menandakan bahwa kata tersebut sering digunakan dalam ulasan pemain. Beberapa kata yang dominan, seperti "game", "nya", "garena", "burik", dan "tolong", menunjukkan banyaknya keluhan yang disampaikan pengguna. Selain itu, kata "jaring", "lag", "jelek", dan "musuh" juga muncul, yang mengindikasikan bahwa sebagian besar kritik berkaitan dengan masalah koneksi, performa permainan, atau pengalaman bermain yang kurang memuaskan. Melalui word cloud ini, dapat melihat aspek-aspek dalam game yang paling sering dikritik oleh pemain.

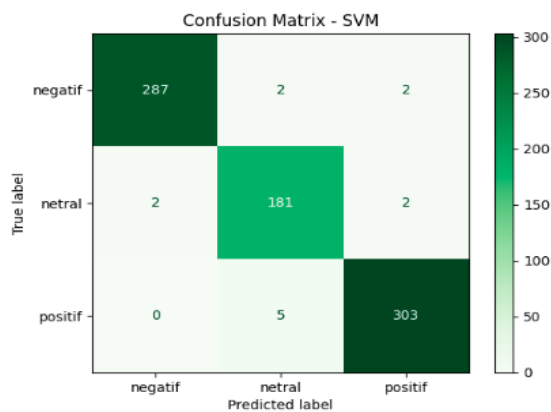
4.4. Hasil Implementasi Model

Proses implementasi model dilakukan dengan memanfaatkan google colab, dimulai dari membaca data dari file CSV yang berisi kolom komentar dan label sentimen. Komentar dalam bentuk teks kemudian diubah menjadi representasi numerik menggunakan teknik CountVectorizer agar bisa diproses oleh algoritma machine learning. Selanjutnya, dilakukan validasi model menggunakan K-Fold Cross Validation sebanyak 10 kali lipat untuk memastikan hasil evaluasi model lebih stabil dan tidak bias. Dua model kemudian dibangun, yaitu Multinomial Naive Bayes dan SVM dengan kernel linear. Masing-masing model dievaluasi berdasarkan akurasi, presisi, recall, dan nilai F1 untuk mengetahui performa klasifikasi.

Untuk mengevaluasi kemampuan klasifikasi dari masing-masing algoritma, dilakukan analisis visual menggunakan confusion matrix. Gambar 8 menunjukkan hasil prediksi model Naive Bayes terhadap data sentimen, sedangkan gambar 9 merupakan hasil evaluasi dari model Support Vector Machine (SVM). Melalui kedua confusion matrix tersebut, terlihat secara jelas jumlah prediksi yang sesuai dan tidak sesuai antara label asli dan label hasil prediksi dari masing-masing model. Visualisasi ini membantu dalam memahami distribusi kesalahan dan keberhasilan klasifikasi dari kedua algoritma secara lebih intuitif.

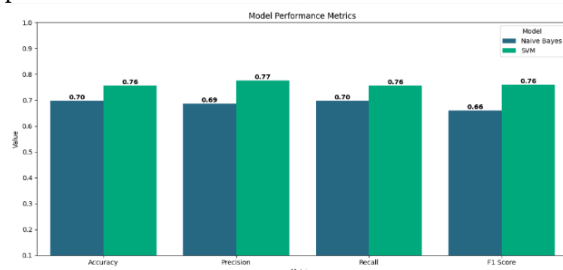


Gambar 8. Confusion matrix - naive bayes



Gambar 9. Confusion matrix - svm

Gambar 9 menunjukkan hasil klasifikasi sentimen menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM) dengan metode validasi K-Fold Cross Validation (K=10). Dalam metode ini, data dibagi menjadi 10 lipatan, di mana setiap iterasi menggunakan 9 bagian untuk pelatihan dan 1 bagian untuk pengujian secara bergantian. Klasifikasi dilakukan ke dalam tiga kategori utama, yaitu positif, negatif, dan netral. Naïve Bayes, sebagai algoritma berbasis probabilistik, mampu mengolah data teks dengan cepat dan efisien, sementara SVM bekerja dengan membangun hyperplane yang optimal untuk memisahkan kelas sentimen. Penelitian ini bertujuan untuk membandingkan performa kedua algoritma dalam mengklasifikasikan sentimen pengguna terhadap game Free Fire, guna menentukan model yang paling akurat dalam memahami opini pemain.



Gambar 10. Perbandingan hasil model

Gambar ini menampilkan perbandingan kinerja algoritma Naïve Bayes dan Support Vector Machine (SVM) berdasarkan empat metrik evaluasi, yaitu Akurasi, Presisi, Recall, dan F1 Score. Dari hasil yang ditampilkan, algoritma SVM menunjukkan kinerja yang lebih baik dibandingkan Naïve Bayes dalam semua metrik. Akurasi Naïve Bayes berada di angka 0.70, sedangkan SVM lebih tinggi dengan 0.76. Untuk Presisi, Naïve Bayes memperoleh nilai 0.69, sementara SVM lebih unggul dengan 0.77. Begitu pula pada Recall, Naïve Bayes memiliki nilai 0.70, sementara SVM mencapai 0.76. Pada metrik F1 Score, Naïve Bayes memperoleh 0.66, sedangkan SVM mencatat angka 0.76. Dengan hasil ini, dapat disimpulkan bahwa SVM memiliki kinerja yang lebih optimal dibandingkan Naïve Bayes dalam mengklasifikasikan data pada penelitian ini.

5. KESIMPULAN DAN SARAN

Penelitian ini membandingkan performa algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam analisis sentimen ulasan pengguna terhadap game Free Fire di Google Play Store. Hasil evaluasi menunjukkan bahwa algoritma SVM memiliki kinerja lebih baik dibandingkan Naïve Bayes dalam semua metrik evaluasi, dengan akurasi 76%, presisi 77%, recall 76%, dan F1-score 76%. Sementara itu, Naïve Bayes memperoleh akurasi sebesar 70%, presisi 69%, recall 70%, dan F1-score 66%. Dengan demikian, SVM lebih direkomendasikan dalam klasifikasi sentimen ulasan Free Fire karena kemampuannya yang lebih optimal dalam memisahkan kelas sentimen. Sebagai saran, penelitian selanjutnya dapat memperluas jumlah data ulasan yang digunakan untuk meningkatkan generalisasi model. Selain itu, eksplorasi terhadap metode deep learning seperti LSTM atau BERT dapat dilakukan untuk meningkatkan akurasi klasifikasi sentimen. Penggunaan teknik preprocessing yang lebih canggih juga dapat membantu dalam mengurangi noise pada data, sehingga hasil analisis menjadi lebih akurat..

DAFTAR PUSTAKA

- [1] Kariyamin, muh. ikhsan Alyakin, and la ode Alyandi, "Menganalisis Ulasan Mobile Legends : Analisis Kinerja Berdasarkan Opini Pengguna dengan Naive Bayes," *Bul. Sist. Inf. dan Teknol. Islam*, vol. 6, no. 1, pp. 1–8, 2025, doi: 10.33096/busiti.v6i1.2475.
- [2] M. Haikal, M. Martanto, and U. Hayati, "Analisis Sentimen Terhadap Penggunaan Aplikasi Game Online Pubg Mobile Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 6, pp. 3275–3281, 2024, doi: 10.36040/jati.v7i6.8174.
- [3] N. Agustina, D. H. Citra, W. Purnama, C. Nisa, and A. R. Kurnia, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 2, no.

- 1, pp. 47–54, 2022, doi: 10.57152/malcom.v2i1.195.
- [4] M. Safrudin, M. Martanto, and U. Hayati, “Perbandingan Kinerja Naïve Bayes Dan Support Vector Machine Untuk Klasifikasi Sentimen Ulasan Game Genshin Impact,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 3, pp. 3182–3188, 2024, doi: 10.36040/jati.v8i3.8415.
- [5] A. Primandani, S. Pungkas, and K. Bagus Adhi, “Analisis Sentimen Game Genshin Impact pada Play Store Menggunakan Naïve Bayes Clasifier,” *J. Ilm. Tek. Mesin, Elektro dan Komput.,* vol. 3, no. 1, pp. 161–170, 2023, doi: 10.51903/juritek.v3i1.1962.
- [6] N. Anizah, Y. Salim, and L. B. Ilmawan, “Analisis Sentimen Terhadap Event Big Sale 11.11 Shopee di Media Sosial Instagram menggunakan Metode Naïve Bayes,” *Bul. Sist. Inf. dan Teknol. Islam,* vol. 4, no. 1, pp. 25–34, 2023, doi: 10.33096/busiti.v4i1.1309.
- [7] Y. L. R. Rehatalanit, M. P. E. Longdong, and A. Ramadhany, “IMPLEMENTASI NAÏVE BAYES DAN SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN ULASAN PADA GOOGLE PLAY,” *JSI (Jurnal Sist. Informasi),* vol. 11, no. 2, pp. 223–236, 2024, doi: <https://doi.org/10.35968/jsi.v11i2.1256>.
- [8] R. R. Putri and N. Cahyono, “Analisis Sentimen Komentar Masyarakat Terhadap Pelayanan Publik Pemerintah Dki Jakarta Dengan Algoritma Super Vector Machine Dan Naive Bayes,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 2, pp. 2363–2371, 2024, doi: 10.36040/jati.v8i2.9472.
- [9] R. Noviana and I. Rasal, “Penerapan Algoritma Naive Bayes Dan Svm Untuk Analisis Sentimen Boy Band Bts Pada Media Sosial Twitter,” *J. Tek. dan Sci.,* vol. 2, no. 2, pp. 51–60, 2023, doi: 10.56127/jts.v2i2.791.
- [10] L. O. Alyandi, L. Hadiani, and Irma, “Implementasi Naive Bayes dalam Analisis Sentimen Ulasan Game Honor of Kings di Playstore Implementation of Naive Bayes in Sentiment Analysis of Comments on the,” *Bul. Sist. Inf. dan Teknol. Islam,* vol. 6, no. 1, pp. 29–37, 2025, doi: 10.33096/busiti.v6i1.2589.
- [11] V. Fazrian, T. Suprpti, and R. Narasati, “Penerapan Algoritma Naive Bayes Terhadap Analisis Sentimen Aplikasi Game Multiplayer Online Battle Arena,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 1, pp. 1005–1012, 2024, doi: 10.36040/jati.v8i1.8432.
- [12] D. Wijaya, R. A. Saputra, and F. Irwiensyah, “Analisis Sentimen Ulasan Aplikasi Samsat Digital Nasional Pada Google Playstore Menggunakan Algoritma Naïve Bayes,” *KLIK Kaji. Ilm. Inform. dan Komput.,* vol. 4, no. 4, pp. 2369–2380, 2024, doi: 10.30865/klik.v4i4.1738.
- [13] M. Apriliyani, M. I. Musyaffaq, S. Nur’Aini, M. R. Handayani, and K. Umam, “Implementasi analisis sentimen pada ulasan aplikasi Duolingo di Google Playstore menggunakan algoritma Naïve Bayes,” *AITI J. Teknol. Inf.,* vol. 21, no. 2, pp. 298–311, 2024, doi: 10.36040/jati.v8i1.8708.
- [14] R. Wahyudi and G. Kusumawardhana, “Analisis Sentimen Review Aplikasi LinkedIn di Google Play Store Menggunakan Support Vector Machine,” *J. Inform.,* vol. 20, no. 1, p. 454, 2024, doi: 10.35889/progresif.v20i1.1614.
- [15] Lisnawati, N. Ganda, and S. Hidayat, “Dampak Game Online Free Fire pada Siswa Sekolah Dasar,” *PEDADIDAKTIKA J. Ilm. Pendidik. GURU Sekol. DASAR,* vol. 8, no. 4, pp. 840–850, 2021, [Online]. Available: <http://ejournal.upi.edu/index.php/pedadidaktika/index>
- [16] T. Septiani Gumilar, R. Astuti, and Y. Arie Wijaya, “Analisis Sentimen Ulasan Aplikasi Lita Di Play Store Menggunakan Algoritma Naive Bayes,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 1, pp. 543–550, 2024, doi: 10.36040/jati.v8i1.8778.
- [17] R. T. H. Putra and A. R. Dzirkillah, “Penerapan Metode Naïve Bayes untuk Analisis Sentimen pada Ulasan Pengguna Aplikasi ChatGPT di Google Play Store,” *Build. Informatics, Technol. Sci.,* vol. 6, no. 1, pp. 430–439, 2024, doi: 10.47065/bits.v6i1.5400.
- [18] A. Pangestu, Y. Tajul Arifin, and R. Ade Safitri, “Analisis Sentimen Review Publik Pengguna Game Online Pada Platform Steam Menggunakan Algoritma Naïve Bayes,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 7, no. 6, pp. 3106–3113, 2024, doi: 10.36040/jati.v7i6.8829.
- [19] A. E. Widodo, F. F. Wati, and N. Hidayati, “Implementasi Algoritma SVM Dan Naive Bayes Untuk Analisis Sentimen Pada Ulasan Pengguna Aplikasi Onedrive,” *Indones. J. Softw. Eng.,* vol. 10, no. 2, pp. 131–138, 2024, doi: <https://doi.org/10.31294/ijse.v10i2.23523>.
- [20] M. D. Hendriyanto, A. A. Ridha, and U. Enri, “ANALISIS SENTIMEN ULASAN APLIKASI MOLA PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE,” *J. Inf. Technol. Comput. Sci.,* vol. 5, no. 1, pp. 1–7, 2022, doi: <https://doi.org/10.31539/intecom.v5i1.3708>.