

KLASIFIKASI SENTIMEN ULASAN MINECRAFT PADA GOOGLE PLAY STORE: STUDI KOMPARATIF NAÏVE BAYES DAN RANDOM FOREST

Luluk Asti Qomariah, Ade Eviyanti, Hamzah Setiawan, Novia Ariyanti

Program Studi Informatika S1, Universitas Muhammadiyah Sidoarjo

Jl. Mojopahit No.666 B, Sidowayah, Celep, Kec. Sidoarjo, Kabupaten Sidoarjo, Jawa Timur 61215, Indonesia

adeeviyanti@umsida.ac.id

ABSTRAK

Pertumbuhan aplikasi *mobile* mendorong peningkatan volume ulasan pengguna pada platform *Google Play Store*, khususnya untuk aplikasi *game*. Ulasan ini mengandung pandangan pengguna yang bermanfaat untuk memahami penilaian mereka terhadap mutu aplikasi. Riset ini fokus pada klasifikasi sentimen terhadap ulasan *game* Minecraft di *Google Play Store* dengan membandingkan performa algoritma *Naïve Bayes* dan *Random Forest*. Pengumpulan data dilakukan melalui teknik *scraping* yang menghasilkan 12.691 ulasan, kemudian setelah tahap *preprocessing* diperoleh 11.078 data valid. Klasifikasi dilakukan dengan menerapkan pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) dengan tiga skenario *split* data: 90:10, 80:20, dan 70:30. Penilaian model menerapkan metrik *accuracy*, *precision*, *recall*, *F1-score*, serta *Area Under Curve* (AUC). Temuan riset memperlihatkan bahwa *Naïve Bayes* secara stabil menunjukkan performa superior dibanding *Random Forest* di semua skenario uji. Performa optimal dicapai pada *split* 90:10 dengan *accuracy* 0,9052 dan *F1-score* 0,8993. Temuan ini mengonfirmasi bahwa *Naïve Bayes* lebih optimal untuk klasifikasi sentimen ulasan *game* berbasis teks.

Kata kunci : analisis sentimen, *Naïve Bayes*, *Random Forest*, *Google Play Store*, *Minecraft*

1. PENDAHULUAN

Perkembangan teknologi digital mendorong peningkatan penggunaan aplikasi berbasis *mobile*, termasuk aplikasi permainan (*Game*) yang tersedia pada platform *Google Play Store*. *Game* tidak lagi hanya dipandang sebagai sarana hiburan semata, tetapi telah berkembang menjadi bagian dari industri kreatif dan ekonomi digital dengan jumlah pengguna yang sangat besar [1]. Salah satu *game* yang memiliki popularitas tinggi adalah *Minecraft*, yang telah diunduh oleh jutaan pengguna dengan beragam latar belakang. Tingginya jumlah pengguna tersebut mendorong munculnya berbagai ulasan yang berisi pendapat, kritik, dan kesan dan interaksi pengguna selama bermain *game* tersebut.

Umpan balik pengguna yang dipublikasikan di *Google Play Store* merupakan sumber informasi yang bernilai penting bagi pengembang maupun calon pengguna. Ulasan tersebut dapat menggambarkan tingkat kepuasan, permasalahan teknis, kualitas fitur, serta pengalaman bermain secara keseluruhan [2]. Banyaknya ulasan pengguna dengan bentuk teks yang tidak terstruktur menjadikan proses analisis secara manual sulit dilakukan secara efektif. Kondisi tersebut mendorong perlunya metode otomatis yang dapat memproses serta mengelompokkan ulasan pengguna secara sistematis, salah satunya melalui penerapan analisis sentimen.

Dalam kajian pengolahan teks, analisis sentimen digunakan sebagai pendekatan untuk mengekstraksi kecenderungan opini dari data tekstual dengan mengelompokkan sentimen, seperti sentimen positif dan negatif [3]. Berbagai riset terdahulu menunjukkan bahwa analisis sentimen telah diaplikasikan secara luas untuk mengkaji umpan balik pengguna terhadap

berbagai aplikasi digital yang tersedia di *Google Play Store*, baik untuk kategori layanan keuangan, media sosial, maupun permainan [4], [5]. Temuan riset tersebut menunjukkan bahwa analisis sentimen dapat menyajikan gambaran objektif terkait persepsi pengguna terhadap suatu aplikasi sekaligus berfungsi sebagai landasan evaluasi bagi pengembang.

Algoritma *machine learning* kerap diaplikasikan dalam klasifikasi teks untuk keperluan analisis sentimen. Dalam klasifikasi teks, algoritma *Naïve Bayes* dan *Random Forest* termasuk metode yang umum diterapkan. *Naïve Bayes* dikenal sebagai metode klasifikasi probabilistik yang simpel, efisien, serta menunjukkan kinerja optimal pada dataset teks [6]. *Random Forest* dikembangkan sebagai pendekatan ensemble yang membangun sejumlah pohon keputusan secara bersamaan, sehingga mampu menangani kompleksitas data dan menekan risiko overfitting pada proses klasifikasi. [7]. Sejumlah riset terdahulu mengindikasikan bahwa *Random Forest* sering kali memberikan hasil klasifikasi yang lebih baik jika dibandingkan dengan *Naïve Bayes*, meskipun efektivitasnya sangat bergantung pada karakteristik data yang dianalisis. [8] [9].

Kendati analisis sentimen pada ulasan aplikasi di *Google Play Store* telah menjadi topik yang sering dikaji, kajian yang secara khusus mengkaji ulasan *game* Minecraft dengan melakukan perbandingan kinerja antara algoritma *Naïve Bayes* dan *Random Forest* masih terbatas. Beberapa penelitian terdahulu lebih banyak berfokus pada aplikasi layanan digital seperti aplikasi Linkaja dan Tantan [10], [11]. Selain itu, penelitian lain juga berfokus pada aplikasi permainan selain *Minecraft*, seperti *Mobile Legends*, *Genshin Impact*, *Growtopia*, dan *Wattpad* [12], [5],

[13], [14]. Padahal, *Minecraft* memiliki karakteristik pengguna dan gaya ulasan yang beragam, termasuk penggunaan bahasa informal, singkatan, dan ekspresi emosional, yang berpotensi memengaruhi kinerja model klasifikasi. Oleh karena itu, diperlukan penelitian yang mengkaji studi perbandingan performa algoritma *Naïve Bayes* dan *Random Forest* menggunakan ulasan game *Minecraft* dengan tujuan menghasilkan model klasifikasi sentimen yang lebih tepat serta mampu merepresentasikan data

Fokus utama penelitian ini adalah mengevaluasi kemampuan algoritma *Naïve Bayes* dan *Random Forest* dalam mengklasifikasikan sentimen ulasan pengguna game *Minecraft* pada platform *Google Play Store*. Pengujian dilakukan menggunakan tiga variasi proporsi data pelatihan dan pengujian, yakni 70:30, 80:20, dan 90:10. Kinerja model dianalisis berdasarkan metrik *accuracy*, *precision*, *recall*, serta *F1-score*. Hasil penelitian ini diharapkan dapat memperkaya kajian analisis sentimen dan menjadi bahan pertimbangan bagi pengembang *game* dalam meningkatkan mutu serta pengalaman pengguna.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Penerapan analisis sentimen dilakukan guna mengetahui kecenderungan opini dari pengguna dengan menelaah isi teks dan menentukan polaritas sentimen yang terkandung di dalamnya, seperti kecenderungan positif maupun negatif terhadap suatu objek tertentu. Analisis sentimen banyak diterapkan pada data tidak terstruktur seperti ulasan pengguna di platform digital karena mampu memberikan gambaran persepsi pengguna secara objektif dan sistematis [3]. Dalam konteks *Google Play Store*, analisis sentimen berperan penting sebagai alat evaluasi untuk menilai kualitas aplikasi berdasarkan pengalaman pengguna yang disampaikan melalui ulasan [4].

2.2. Analisis Sentimen pada Ulasan Aplikasi dan Game

Sejumlah riset sebelumnya telah mengaplikasikan analisis sentimen untuk mengkaji berbagai ulasan aplikasi digital di *Google Play Store*. Penelitian pada aplikasi layanan digital menunjukkan bahwa analisis sentimen mampu mengidentifikasi kelebihan dan kekurangan berdasarkan opini pengguna, sehingga dapat menjadi dasar pengambilan keputusan bagi pengembang [5], [8]. Di samping itu, analisis sentimen banyak dimanfaatkan pada ulasan game guna mengidentifikasi pengalaman bermain dan berbagai permasalahan yang sering dialami pengguna.

Algoritma *Naïve Bayes* digunakan oleh Setiawan dkk. dalam penelitian mereka untuk menganalisis sentimen pada ulasan *game Growtopia*, di mana hasil penelitian mengindikasikan bahwa metode tersebut memiliki tingkat efektivitas yang baik dalam membedakan sentimen positif dan negatif [13]. Selain itu, penelitian lain pada *game* seperti *Mobile Legends* dan *Genshin Impact* menunjukkan bahwa pembobotan

TF-IDF dan algoritma *machine learning* dapat meningkatkan akurasi klasifikasi sentimen [12], [14]. Hasil penelitian tersebut menegaskan bahwa karakteristik ulasan game yang beragam memerlukan pemilihan metode klasifikasi yang tepat.

2.3. Penelitian Terdahulu

Beberapa penelitian terdahulu telah menerapkan analisis sentimen pada ulasan aplikasi dan game di *Google Play Store* menggunakan algoritma klasifikasi berbasis *machine learning*. Setiawan dkk. Menganalisis sentimen, menganalisis sentimen ulasan game *Growtopia* menggunakan algoritma *Naïve Bayes* dengan dua kelas sentimen, yaitu positif dan negatif. Penelitian tersebut menerapkan tahapan *preprocessing* teks serta evaluasi model menggunakan metrik kinerja, dan menunjukkan bahwa *Naïve Bayes* mampu memberikan performa klasifikasi yang cukup baik pada data ulasan game [13].

Penelitian lain oleh Reinaldo dkk. menerapkan algoritma *Random Forest* untuk analisis sentimen pada data ulasan aplikasi *Ganshin impact*. Dengan memanfaatkan data teks yang telah diproses dan direpresentasikan dalam bentuk fitur numerik, hasil penelitian menunjukkan bahwa *Random Forest* mampu menghasilkan kinerja klasifikasi yang stabil dan efektif dalam menangani kompleksitas data teks [5].

Selain itu, Miftahusalam dkk. Melakukan perbandingan kinerja algoritma *Naïve Bayes* dan *Random Forest* pada analisis sentimen ulasan aplikasi *BCA Mobile* di *Google Play Store*. Penelitian tersebut menggunakan pembobotan TF-IDF dan mengevaluasi model menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa *Naïve Bayes* unggul dari sisi kesederhanaan dan efisiensi, sedangkan *Random Forest* memberikan performa yang lebih stabil pada data dengan karakteristik yang kompleks [6].

Berdasarkan penelitian-penelitian tersebut, dapat disimpulkan bahwa pemilihan algoritma klasifikasi berpengaruh terhadap hasil analisis sentimen dan sangat bergantung pada karakteristik data. Namun, kajian yang secara khusus membandingkan kinerja *Naïve Bayes* dan *Random Forest* pada ulasan game dengan karakteristik bahasa yang beragam, seperti *Minecraft*, masih relatif terbatas. Oleh karena itu, penelitian ini berfokus pada perbandingan performa kedua algoritma tersebut pada ulasan game *Minecraft*.

2.4. Algoritma Naïve Bayes

Algoritma *Naïve Bayes* dikenal luas dan sering diterapkan pada permasalahan klasifikasi teks karena menawarkan proses komputasi yang simpel, efisien, serta mampu menghasilkan performa yang baik pada *dataset* berukuran besar [6]. Dalam analisis sentimen, *Naïve Bayes* sering dikombinasikan dengan teknik pembobotan kata seperti TF-IDF untuk meningkatkan kinerja model klasifikasi.

Beberapa penelitian sebelumnya mengindikasikan bahwa *Naïve Bayes* mampu menghasilkan kinerja yang memadai pada klasifikasi sentimen terhadap ulasan aplikasi digital, meskipun pada beberapa kasus performanya masih kalah dibandingkan algoritma ensemble atau metode berbasis kernel [8].

2.5. Algoritma Random Forest

Algoritma *Random Forest* bekerja dengan mengonstruksi sejumlah *decision tree* secara independen, kemudian menentukan hasil klasifikasi berdasarkan keputusan yang paling banyak dipilih. Pendekatan ini mampu menangani kompleksitas data berdimensi tinggi secara efektif, menekan terjadinya *overfitting* dan menghasilkan performa klasifikasi yang cenderung stabil [7].

Dalam konteks penelitian analisis sentimen, *Random Forest* sering menunjukkan kinerja yang baik, khususnya pada dataset dengan tingkat kompleksitas tinggi dan distribusi kelas yang tidak seimbang. Beberapa penelitian sebelumnya menunjukkan bahwa *Random Forest* cenderung menghasilkan nilai *accuracy*, *presisi*, *recall*, dan *F1-score* yang lebih baik dibandingkan metode klasifikasi lain, termasuk *Naïve Bayes* [15], [9], [16].

2.6. Perbandingan Naïve Bayes dan Random Forest

Topik perbandingan performa algoritma *Naïve Bayes* dan *Random Forest* telah sering menjadi fokus dalam studi analisis sentimen. Secara umum, *Naïve Bayes* dikenal memiliki keunggulan dari sisi kecepatan komputasi dan kesederhanaan model, sementara *Random Forest* lebih efektif dalam menangani kompleksitas data serta menghasilkan kinerja klasifikasi yang lebih stabil [8], [9]. Tidak hanya pada analisis sentimen, kajian perbandingan algoritma klasifikasi seperti *Naïve Bayes* dan *Random Forest* juga sering dilakukan pada berbagai jenis data teks. Untuk melihat karakteristik dan keunggulan masing-masing metode [17].

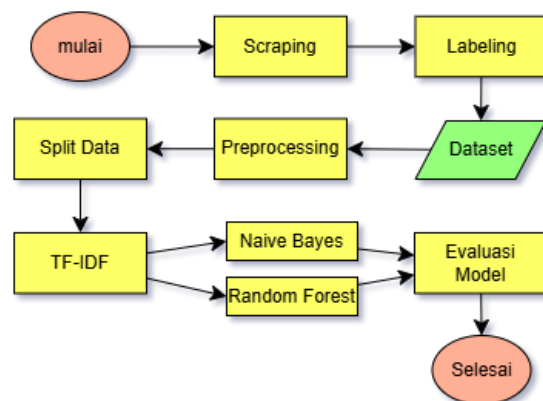
Hasil telaah terhadap penelitian-penelitian sebelumnya mengindikasikan bahwa tingkat efektivitas algoritma klasifikasi dipengaruhi oleh karakteristik data yang menjadi objek penelitian. Dengan demikian, diperlukan penelitian lanjutan untuk membandingkan kinerja *Naïve Bayes* dan *Random Forest* dalam ulasan *game Minecraft* yang memiliki variasi bahasa serta pola sentimen yang beragam.

Kajian terhadap sejumlah penelitian terdahulu memperlihatkan bahwa analisis sentimen pada ulasan aplikasi digital diterapkan menggunakan beragam algoritma klasifikasi dengan skema pengujian yang bervariasi. Hasil penelitian terdahulu menunjukkan bahwa pemilihan algoritma klasifikasi dan skenario pengujian dapat memengaruhi performa analisis sentimen secara signifikan [18]. Namun, kajian yang secara spesifik mengkomparasi performa *Naïve Bayes*

dan *Random Forest* pada konteks ulasan *game* masih terbatas dengan karakteristik bahasa yang unik seperti *Minecraft*. Oleh karena itu, penelitian ini membandingkan kinerja kedua algoritma melalui beberapa skenario rasio pembagian data guna menentukan model dengan performa yang paling optimal untuk klasifikasi sentimen ulasan *game Minecraft*.

3. METODE PENELITIAN

Gambar 1 menyajikan tahapan penelitian yang dilakukan, dimulai dari pengambilan data hingga penilaian performa model klasifikasi sentimen.



Gambar 1. Diagram Alir Penelitian

Alur penelitian pada studi ini menggambarkan rangkaian proses yang dimulai dari pengambilan ulasan pengguna *game Minecraft* melalui *web scraping*, dilanjutkan dengan pelabelan sentimen dan preprocessing teks. Data hasil preprocessing dikelompokkan menjadi data pelatihan dan pengujian, direpresentasikan menggunakan pembobotan TF-IDF, kemudian diklasifikasikan dengan algoritma *Naïve Bayes* dan algoritma *Random Forest* sebelum dilakukan evaluasi kinerja model.

3.1. Jenis dan Pendekatan Penelitian

Metode komparatif digunakan dalam penelitian ini dengan pendekatan kuantitatif, mengingat analisis yang dilakukan berfokus pada pengolahan data numerik hasil klasifikasi sentimen. Metode komparatif digunakan untuk mengkomparasi performa dua algoritma *machine learning*, yaitu *Naïve Bayes* dan *Random Forest*, saat mengklasifikasikan sentimen ulasan pada *game Minecraft* pada platform *Google Play Store*.

3.2. Sumber dan Pengumpulan Data

Ulasan pengguna *game Minecraft* pada *Google Play Store* digunakan sebagai data sekunder dalam penelitian ini, yang dihimpun melalui proses *web scraping* dengan memanfaatkan library *google-play-scraper*. Komentar ulasan berbahasa Indonesia digunakan sebagai data dalam penelitian ini yang mencerminkan pengalaman pengguna terhadap *game Minecraft*. Hasil data yang diperoleh kemudian

disimpan dalam format CSV untuk keperluan sehingga dapat dimanfaatkan pada tahap pengolahan dan analisis data berikutnya.

3.3. Pelabelan Data

Pelabelan data dilaksanakan secara manual dengan menetapkan kategori sentimen pada setiap ulasan berdasarkan kecenderungan opini yang terkandung, yaitu sentimen positif dan negatif. Label yang dihasilkan digunakan sebagai *ground truth* untuk keperluan pelatihan dan evaluasi model klasifikasi sentimen

3.4. Preprocessing Data Teks

Tahap preprocessing diterapkan untuk menyiapkan data ulasan sebelum proses pemodelan dilakukan. Pada tahap ini, data teks diproses melalui beberapa langkah, antara lain *cleaning*, *case folding*, *tokenization*, *normalization*, *stopword removal*, serta *stemming* dengan menggunakan *library Sastrawi*. Tahap tersebut dilakukan untuk meminimalkan noise pada data teks dan meningkatkan kualitas fitur yang digunakan dalam klasifikasi.

Selain itu, pada tahap ini juga dilakukan *exploratory data analysis* (EDA) untuk melihat kondisi dan distribusi data ulasan sebelum dan sesudah preprocessing. Melalui EDA, dipastikan bahwa data berada dalam kondisi bersih, terstruktur, dan siap digunakan dalam pembagian data serta proses pemodelan

3.5. Pembagian Data

Data yang terkumpul dari hasil preprocessing selanjutnya dibagi menjadi data pelatihan dan data pengujian menggunakan tiga macam rasio pembagian, yaitu 70:30, 80:20, dan 90:10. Variasi rasio tersebut digunakan untuk mengkaji pengaruh proporsi data pelatihan dan data pengujian terhadap kinerja model klasifikasi.

3.6. Pembobotan TF-IDF

Secara matematis, pembobotan TF-IDF dirumuskan sebagai berikut:

$$\text{TF-IDF}(t,d) = \text{TF}(t,d) \times \text{IDF}(t)$$

dengan:

- $\text{TF}(t,d)$ merepresentasikan frekuensi kemunculan term t dalam dokumen d
- $\text{IDF}(t)$ mempresentasikan inverse document frequency dari term t

Nilai IDF dihitung menggunakan persamaan:

$$\text{IDF}(t) = \log(N/\text{df}(t))$$

Keterangan :

- N adalah total jumlah dokumen
- $\text{df}(t)$ adalah jumlah dokumen yang memuat term t

Pada tahap pembobotan fitur, data teks pada *data training* dan *data testing* direpresentasikan dalam bentuk vektor numerik dengan menerapkan metode

TF-IDF. Pembobotan TF-IDF diterapkan untuk menyajikan bobot kepentingan suatu kata pada dokumen, sehingga dapat berfungsi sebagai fitur masukan bagi algoritma klasifikasi.

3.7. Pemodelan Klasifikasi

Tahap pemodelan klasifikasi dilakukan dengan menerapkan dua metode pembelajaran mesin berupa *Naïve Bayes* dan *Random Forest*. *Data training* yang telah direpresentasikan menggunakan pembobotan TF-IDF dimanfaatkan untuk membangun model, sementara *Data testing* dimanfaatkan untuk menilai kemampuan model dalam mengklasifikasikan sentimen ulasan pengguna.

3.8. Algoritma Naïve Bayes

Algoritma *Naïve Bayes* menggunakan pendekatan probabilistik dalam proses klasifikasi dengan mengacu pada *Teorema Bayes* serta mengasumsikan di mana setiap atribut dianggap independen terhadap atribut lainnya. Secara matematis, *Teorema Bayes* dapat dinyatakan dalam persamaan berikut:

$$P(C|X) = (P(X|C) \times P(C)) / P(X)$$

Keterangan:

- $P(C|X)$ merupakan probabilitas kelas C dengan kondisi fitur X
- $P(X|C)$ merupakan probabilitas fitur X pada kondisi kelas C
- $P(C)$ merupakan probabilitas *prior* dari kelas C
- $P(X)$ merupakan probabilitas dari fitur X

Pada proses klasifikasi, nilai $P(X)$ bersifat konstan sehingga dapat diabaikan. Sentimen ditentukan berdasarkan kelas yang memiliki nilai probabilitas *posterior* tertinggi.

3.9. Random Forest

Algoritma *Random Forest* menerapkan pendekatan *ensemble learning* dengan mendirikan sejumlah *decision tree* secara independen. Setiap pohon menghasilkan prediksi kelas, kemudian keputusan akhir ditentukan berdasarkan mekanisme *majority voting* dari seluruh pohon yang terbentuk.

Secara umum, prediksi *Random Forest* dapat dirumuskan sebagai berikut:

$$\hat{y} = \text{mode} \{ h_1(x), h_2(x), \dots, h_n(x) \}$$

Keterangan:

- merupakan output prediksi akhir
- $h_i(x)$ merupakan output dari pohon keputusan ke- i
- n merupakan total pohon keputusan yang digunakan

Pendekatan ensemble pada *Random Forest* memungkinkan model memiliki performa yang lebih stabil dan mampu mengurangi risiko *overfitting* dibandingkan dengan satu pohon keputusan tunggal.

3.10. Evaluasi Model

Penilaian performa model bertujuan untuk mengukur kapabilitas dua algoritma yang digunakan dalam melakukan klasifikasi sentimen ulasan aplikasi. Penilaian ini menggunakan *confusion matrix* serta sejumlah metrik evaluasi meliputi *accuracy*, *precision*, *recall*, dan *F1-score*.

Accuracy dimanfaatkan untuk mengukur tingkat ketepatan model pada saat melakukan klasifikasi dan dirumuskan sebagai:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Nilai *precision* menunjukkan tingkat ketepatan model saat menghasilkan prediksi positif, yang bisa dihitung dengan persamaan berikut:

$$Precision = \frac{TP}{TP + FP}$$

Nilai *recall* mencerminkan kemampuan model dalam mendeteksi data kelas positif, dengan perhitungan menggunakan persamaan berikut:

$$Recall = \frac{TP}{TP + FN}$$

sebagai metrik evaluasi, *F1-score* menggabungkan *precision* dan *recall* dalam bentuk rata-rata harmonis yang dirumuskan sebagai:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Keterangan:

- TP (*True Positive*)
- TN (*True Negative*)
- FP (*False Positive*)
- FN (*False Negative*)

Kemampuan model dalam mengklasifikasikan kelas positif dan negatif pada berbagai tingkat ambang dinilai berdasarkan nilai AUC yang diperoleh dari kurva ROC [19].

Hasil evaluasi dari setiap algoritma pada masing-masing skenario pembagian data selanjutnya dianalisis untuk menentukan model dengan kinerja terbaik.

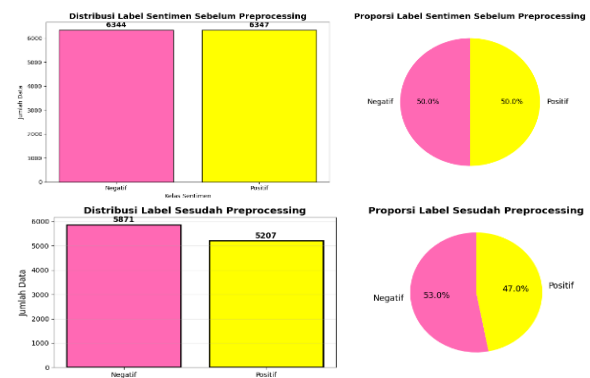
4. HASIL DAN PEMBAHASAN

4.1. Deskripsi Dataset dan Analisis Awal Data

Pengumpulan dataset dilakukan dengan melakukan scraping terhadap ulasan pengguna game Minecraft di *Google Play Store*. Dataset mentah yang berhasil dikumpulkan totalnya 12.691 ulasan, dengan distribusi sentimen yang relatif seimbang, yaitu 6.347 ulasan positif dan 6.344 ulasan negatif. Kondisi ini menunjukkan bahwa data mentah tidak mengalami ketimpangan kelas (*class imbalance*) yang signifikan.

Setelah dilakukan tahap preprocessing yang mencakup *cleaning*, *case folding*, *tokenization*, *normalization*, *stopword removal*, dan *stemming*, serta penghapusan data duplikat, data kosong (*NaN*), dan

komentar dengan jumlah kata kurang dari tiga, diperoleh dataset akhir sebanyak 11.078 ulasan. Dataset tersebut terdiri dari 5.207 ulasan positif (47%) dan 5.871 ulasan negatif (53%).

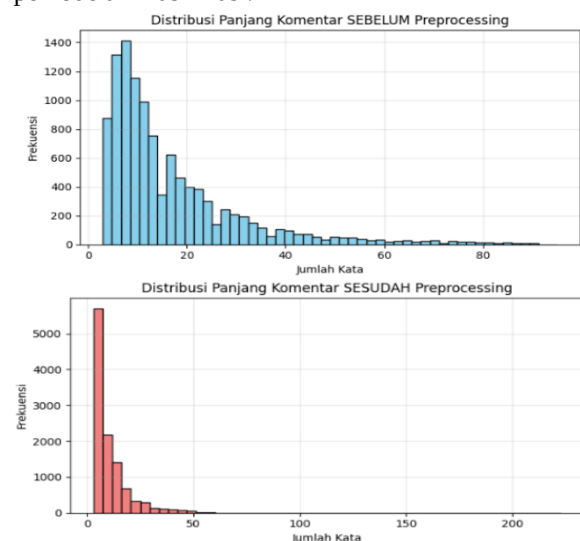


Gambar 2. Distribusi label sentimen sebelum dan sesudah preprocessing

4.2. Analisis Panjang Komentar dan Karakteristik Data

Analisis panjang komentar dilakukan untuk melihat pengaruh preprocessing terhadap struktur data teks. Berdasarkan Gambar 3, terlihat bahwa sebelum preprocessing, data memiliki variasi panjang komentar yang cukup besar dengan keberadaan komentar sangat panjang. Hal ini menunjukkan adanya noise berupa kata tidak relevan dan pengulangan kata pada data mentah.

Pasca *preprocessing*, distribusi panjang komentar menunjukkan pemusatan yang lebih baik dengan berkurangnya nilai-nilai ekstrem pada data. Hal ini mengindikasikan bahwa tahap pembersihan teks efektif dalam mengeliminasi kata-kata yang kurang relevan serta membentuk representasi teks yang lebih kompak. Dengan demikian, data yang diperoleh menjadi lebih terorganisasi dan siap digunakan pada tahap pembobotan TF-IDF serta pemodelan klasifikasi.



Gambar 3. Distribusi panjang komentar sebelum dan sesudah preprocessing

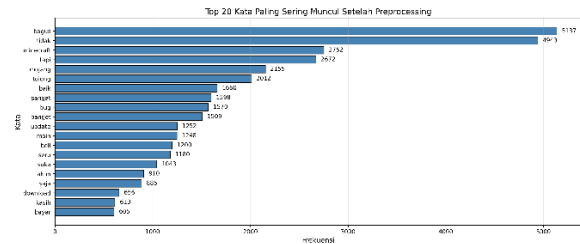
4.3. Analisis Kata Dominan dan Wordcloud

Untuk memahami karakteristik kata yang sering digunakan oleh pengguna, dilakukan visualisasi *wordcloud* terhadap data hasil preprocessing. *Wordcloud* menunjukkan bahwa kata-kata seperti *bagus*, *tidak*, *minecraft*, *update*, dan *bug* memiliki frekuensi kemunculan yang tinggi. Hal ini menunjukkan bahwa ulasan pengguna banyak berisi evaluasi langsung terhadap kualitas permainan dan pembaruan aplikasi.



Gambar 4. *Wordcloud* data ulasan setelah preprocessing

Selain analisis terhadap kata yang paling sering muncul memperkuat temuan bahwa kata evaluatif dan kontekstual memiliki peran penting dalam membentuk sentimen ulasan.



Gambar 5. Top 20 kata paling sering muncul setelah preprocessing

4.4. Hasil Pemodelan Klasifikasi pada Berbagai Rasio Data

Pemodelan klasifikasi sentimen memanfaatkan algoritma *Naïve Bayes* dan *Random Forest* dengan tiga konfigurasi rasio pembagian data pelatihan dan pengujian, yakni 90:10, 80:20, dan 70:30. Penilaian performa model menerapkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

Hasil pengujian mengindikasikan bahwa kedua algoritma mampu mempertahankan performa yang stabil pada berbagai rasio pembagian data. *Naïve Bayes* memperoleh nilai *accuracy* sebesar 0,9052 pada rasio 90:10, 0,8935 pada rasio 80:20, dan 0,8908 pada rasio 70:30. Sementara itu, *Random Forest* menghasilkan nilai *accuracy* sebesar 0,8944, 0,8840, dan 0,8806 pada masing-masing rasio.

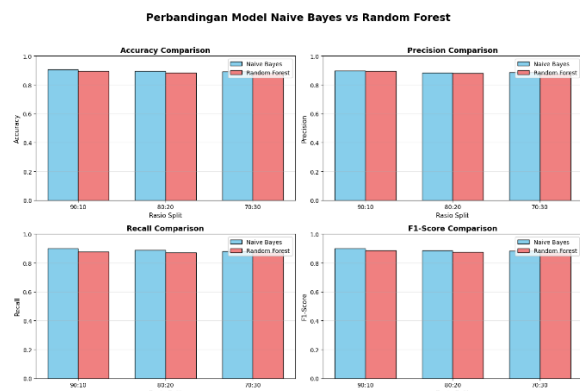
Hasil tersebut menunjukkan bahwa perubahan rasio pembagian data mempengaruhi performa model, namun *Naïve Bayes* secara konsisten menghasilkan nilai evaluasi yang lebih tinggi dibandingkan *Random Forest* pada seluruh skenario pengujian.

Tabel 1. Ikhtisar performa Naïve Bayes dan *Random Forest* di seluruh rasio pembagian data

Rasio	Model	Accuracy	Precision	Recall	F1-score
90:10	Naïve Bayes	0.9052	0.8985	0.9002	0.8993
90:10	Random Forest	0.8944	0.8961	0.8772	0.8865
80:20	Naïve Bayes	0.8935	0.8845	0.8896	0.8871
80:20	Random Forest	0.8840	0.8807	0.8714	0.8760
70:30	Naïve Bayes	0.8908	0.8865	0.8803	0.8834
70:30	Random Forest	0.8806	0.8741	0.8713	0.8727

4.5. Perbandingan Kinerja Naïve Bayes dan Random Forest Analisis Confusion Matrix Rasio Terbaik

Perbandingan kinerja kedua algoritma ditampilkan melalui visualisasi grafik untuk melihat perbedaan performa secara lebih jelas. Grafik menunjukkan bahwa *Naïve Bayes* unggul pada metrik *accuracy*, recall, dan F1-score di seluruh rasio pembagian data, sedangkan *Random Forest* menunjukkan performa yang relatif stabil namun sedikit lebih rendah.

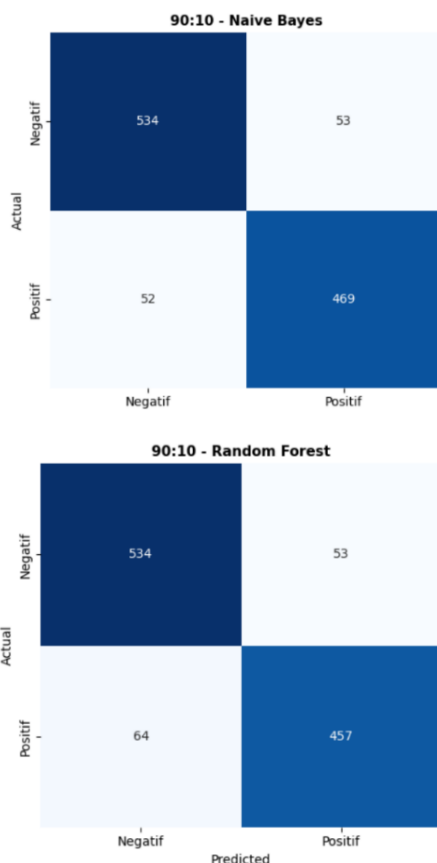


Gambar 6. Perbandingan performa *Naïve Bayes* dan *Random Forest* pada seluruh rasio pengujian

Keunggulan *Naïve Bayes* disebabkan oleh karakteristik data teks yang bersifat berdimensi tinggi dan *sparse*, sehingga lebih sesuai dengan pendekatan probabilistik yang digunakan oleh algoritma tersebut.

4.6. Analisis Confusion Matrix Rasio Terbaik

Untuk analisis lebih mendalam, *confusion matrix* ditampilkan pada rasio pembagian data 90:10 sebagai skenario dengan performa terbaik. Berdasarkan *confusion matrix*, *Naïve Bayes* menunjukkan jumlah *false negative* yang lebih sedikit dibanding *Random Forest*, sehingga memiliki kemampuan lebih baik dalam mendeteksi sentimen positif.



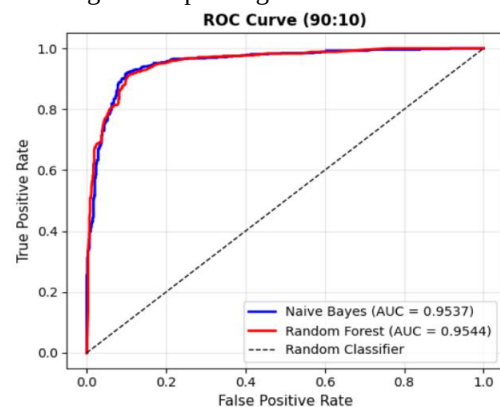
Gambar 7. *Confusion matrix Naïve Bayes dan Random Forest pada rasio 90:10*

Temuan ini mengindikasikan bahwa *Naïve Bayes* memiliki tingkat sensitivitas (*recall*) yang lebih superior, yang menjadi aspek krusial dalam analisis sentimen ulasan pengguna.

4.7. Analisis ROC Curve dan AUC

Evaluasi lanjutan dilakukan dengan menggunakan kurva ROC dan nilai *Area Under Curve* (AUC) sebagai indikator untuk menilai efektivitas model dalam mengklasifikasikan sentimen positif dan negatif pada berbagai nilai ambang (*threshold*). Temuan pengujian memperlihatkan bahwa kedua model mencapai nilai AUC yang tinggi, yakni melebihi 0,95, yang mengindikasikan kinerja klasifikasi yang sangat optimal.

Meskipun *Random Forest* memiliki nilai AUC sedikit lebih tinggi pada salah satu skenario, perbedaan tersebut tidak signifikan. Secara umum, *Naïve Bayes* memperlihatkan tingkat konsistensi yang lebih baik pada berbagai rasio pembagian data.



Gambar 8. Kurva ROC *Naïve Bayes* dan *Random Forest* pada rasio 90:10

4.8. Pembahasan Umum

Berdasarkan seluruh hasil pengujian pada tiga skenario pembagian data, dapat disimpulkan bahwa rasio 90:10 menghasilkan performa terbaik bagi kedua algoritma. Namun demikian, *Naïve Bayes* secara konsisten menunjukkan kinerja yang lebih unggul dibandingkan *Random Forest* pada seluruh rasio, baik dari sisi *accuracy*, *recall*, maupun *F1-score*.

Dengan mempertimbangkan konsistensi kinerja dan karakteristik data teks ulasan, *Naïve Bayes* bisa direkomendasikan selaku metode yang lebih optimal untuk klasifikasi sentimen ulasan *game Minecraft* di *Google Play Store*. Temuan ini selaras dengan beberapa riset terdahulu yang mengungkapkan bahwa *Naïve Bayes* cenderung lebih efektif pada data teks berdimensi tinggi dan distribusi kata yang jarang (*sparse*) [20].

5. KESIMPULAN DAN SARAN

Mengacu pada temuan penelitian, algoritma *Naïve Bayes* dan *Random Forest* mampu diterapkan dengan baik dalam proses klasifikasi sentimen pada ulasan pengguna *game Minecraft* di *Google Play Store*. Dataset yang digunakan dikumpulkan melalui teknik *scraping* dan telah melewati tahap *preprocessing* sehingga menghasilkan data yang lebih bersih dan terstruktur untuk keperluan pemodelan.

Hasil evaluasi memperlihatkan bahwa kedua algoritma mampu melakukan klasifikasi sentimen dengan kinerja yang baik di semua skenario *split* data. Temuan pengujian menyatakan bahwa algoritma *Naïve Bayes* memperlihatkan nilai kinerja yang lebih superior dibandingkan *Random Forest* pada keseluruhan skenario pembagian data yang diterapkan. Dari ketiga rasio pengujian tersebut, konfigurasi 90:10 mencapai hasil kinerja terbaik dengan nilai *accuracy* sebesar 0,9052 dan *F1-score* sebesar 0,8993.

Selain itu, penilaian menggunakan *confusion matrix* dan ROC-AUC mengungkapkan bahwa *Naïve*

Bayes mempunyai kapabilitas yang lebih superior saat mengidentifikasi sentimen positif dengan tingkat sensitivitas yang lebih optimal, serta konsistensi performa yang baik pada data teks berdimensi tinggi dan *sparse*. Dengan demikian, *Naïve Bayes* merupakan metode yang efektif dan dapat direkomendasikan untuk klasifikasi sentimen ulasan *game* pada *Google Play Store*.

Berdasarkan temuan riset ini, diperlukan pengembangan lebih lanjut yang dapat dijadikan rekomendasi untuk penelitian selanjutnya. Riset berikutnya dapat memanfaatkan volume data yang lebih besar serta mengintegrasikan variasi sumber ulasan dari platform lain guna meningkatkan generalisasi model. Selain itu, penerapan metode klasifikasi lain seperti pendekatan *deep learning* atau *Support Vector Machine* dapat dipertimbangkan bagi memperoleh komparasi performa yang lebih komprehensif.

Penelitian selanjutnya juga dapat mengembangkan teknik preprocessing dan ekstraksi fitur, misalnya dengan menggunakan *word embedding* atau mempertimbangkan konteks kalimat, sehingga diharapkan mampu meningkatkan akurasi dan kemampuan model dalam memahami sentimen secara lebih mendalam.

DAFTAR PUSTAKA

- [1] Grand View Research, “*Video Game Market Size, Share And Growth Report, 2030*,” *Grand View Research*, 2025. [Online]. Available: <https://www.grandviewresearch.com/industry-analysis/video-game-market>
- [2] F. H. H. K. N. Putra, and G. E. Buji, “Analisis Pengembangan Media Pembelajaran Berbasis Game Interaktif Edunomic Adventure Pada Pembelajaran Ekonomi,” vol. 5, no. 2, pp. 148–159, 2024.
- [3] A. Suryana, A. I. Purnamasari, and I. Ali, “MENGOPTIMALKAN KEPUASAN PENGGUNA : ANALISIS SENTIMEN REVIEW APLIKASI GRAB DI INDONESIA,” vol. 8, no. 3, pp. 3396–3404, 2024.
- [4] S. Mujilawati and M. A. Ubaydillah, “Analisis Sentimen Berbasis Aspek pada Ulasan Game Mobile Legends dengan Pendekatan Decision Tree untuk Evaluasi Pengalaman Pengguna Teknik Informatika , Fakultas Teknik , Universitas Islam Lamongan , Indonesia Aspect-Based Sentiment Analysis on Mobile Legends Game Reviews with Decision Tree Approach for User Experience Evaluation,” vol. 4, no. 11, pp. 325–333, 2024.
- [5] S. Nasional, T. Elektro, S. Informasi, and T. Informatika, “Seminat Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika,” 2024.
- [6] A. Miftahusalam, H. Pratiwi, and I. Slamet, “Perbandingan Metode Random Forest dan Naive Bayes pada Analisis Sentimen Review Aplikasi BCA Mobile,” *SIPTEK Semin. Nas. Inov. dan Pengemb. Teknol. Pendidik.*, vol. 1, no. 1, pp. 1–8, 2023.
- [7] C. G. Indrayanto, D. E. Ratnawati, and B. Rahayudi, “Analisis Sentimen Data Ulasan Pengguna Aplikasi MyPertamina di Indonesia pada Google Play Store menggunakan Metode Random Forest,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, pp. 1131–1139, 2023, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [8] Fadhilah R, “Perbandingan algoritma Random Forest dan Naive Bayes dalam analisis sentimen ulasan Spotify,” *J. Sist. dan Inform.*, vol. 15, no. 1, pp. 78–90, 2024.
- [9] N. D. Kurniawan, P. R. Ferdian, and N. Hidayati, “Analisis Sentimen Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest Pada Ulasan Aplikasi Ajaib,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 11, no. 1, pp. 87–97, 2025, doi: 10.25077/teknosi.v11i1.2025.87-97.
- [10] K. Kaeren and A. Andrianingsih, “Analisis Sentimen Aplikasi LinkAja di Google Play Store Menggunakan Algoritma Naïve Bayes dan Random Forest,” *J. Ris. dan Apl. Mhs. Inform.*, vol. 6, no. 02, pp. 438–447, 2025, doi: 10.30998/jrami.v6i02.13821.
- [11] S. T. Andini, “Analisis Sentimen Pengguna Aplikasi Tantan : Perbandingan Kinerja Metode Naive Bayes dan SVM,” vol. 15, no. 2, pp. 396–407, 2025.
- [12] T. Informatika, F. Teknik, U. Nusantara, and P. Kediri, “Analisis Sentimen Ulasan Pengguna Aplikasi Mobile Legends Pada Google Playstore Menggunakan Naïve Bayes,” vol. 4, pp. 641–646, 2025.
- [13] R. P. Setiawan, B. Irawan, and W. P. Prihartono, “Analisis Sentimen Ulasan Growtopia Di Google Play Store Menggunakan Naïve Bayes Classifier Untuk Identifikasi Kebutuhan Pengguna,” *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 2, 2025, doi: 10.23960/jitet.v13i2.6415.
- [14] I. Deasiva, O. Nurdiawan, and M. F. Basysyar, “Model Sentimen Analisis Berdasarkan Ulasan Aplikasi Webtoon pada Goggle Play Store Ditingkatkan dengan Algoritma Random Forest,” pp. 22–35.
- [15] S. Nur Adhan, G. N. A. Wibawa, D. C. Arisona, I. Yahya, and R. Ruslan, “Analisis Sentimen Ulasan Aplikasi Wattpad Di Google Play Store Dengan Metode Random Forest,” *AnoaTIK J. Teknol. Inf. dan Komput.*, vol. 2, no. 1, pp. 6–15, 2024, doi: 10.33772/anoatik.v2i1.32.
- [16] A. Gitacahyani, A. I. Purnamasari, and I. Ali, “NAÏVE BAYES CLASSIFIER,” vol. 8, no. 1, pp. 176–181, 2024.
- [17] K. Algoritma, R. Forest, and N. Bayes, “METODE RANDOM FOREST DAN NAÏVE BAYES,” vol. 7, no. 1, pp. 498–504, 2023.
- [18] M. S. Hamid, “Analisis Sentimen Tingkat Kepuasan Aplikasi WordPress Menggunakan Metode K-Nearest Neighbor dan Naive Bayes,” vol. 15, no. 1, pp. 102–114, 2025.
- [19] M. K. Insan, U. Hayati, and O. Nurdiawan, “ANALISIS SENTIMEN APLIKASI BRIMO PADA ULASAN PENGGUNA DI,” vol. 7, no. 1, pp. 478–483, 2023.
- [20] I. Artikel and A. Info, “ANALISIS SENTIMEN PENGGUNA APLIKASI SHOPEE PADA GOOGLE PLAY STORE MENGGUNAKAN METODE RANDOM FOREST,” vol. 4, no. 3, pp. 265–276, 2025.