

## Analisis Sentimen Ulasan Video Game My Lovely Daughter Menggunakan Metode Multinomial Naïve Bayes

### *Sentiment Analysis of My Lovely Daughter Video Game Reviews Using Multinomial Naïve Bayes Method*

Sagtya Rengga Surya Pratama<sup>1</sup>, Deni Arifianto<sup>2\*</sup>, Budi Satria Bakti<sup>3</sup>

<sup>1</sup> Mahasiswa Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember  
Email : [sagtyapratama@gmail.com](mailto:sagtyapratama@gmail.com)

<sup>2</sup> Dosen Fakultas Teknik, Universitas Muhammadiyah Jember \*Koresponden Author  
Email : [deniarifianto@unmuhjember.ac.id](mailto:deniarifianto@unmuhjember.ac.id)

<sup>3</sup> Dosen Fakultas Teknik, Universitas Muhammadiyah Jember  
Email : [satrio93@unmuhjember.ac.id](mailto:satrio93@unmuhjember.ac.id)

#### Abstrak

Dengan kemajuan teknologi informasi, platform untuk penjualan dan pembelian konten elektronik seperti *video game* semakin berkembang, contohnya adalah Steam, Epic Games Store, Origin, dan Uplay. Evaluasi kualitas suatu *game* seringkali melibatkan analisis ulasan dari para pemain, tetapi tidak semua ulasan dengan jelas dapat dikategorikan sebagai direkomendasikan atau tidak. Contohnya adalah *game My Lovely Daughter*, yang dikembangkan oleh developer Indonesia dan tersedia di platform Steam. Masalahnya adalah kejelasan kategori pada ulasan yang dapat mempengaruhi pengembangan *game* selanjutnya. Oleh karena itu, analisis sentimen diperlukan untuk mengidentifikasi apakah ulasan tersebut bersifat positif atau negatif. Dalam hal ini, metode klasifikasi Multinomial Naïve Bayes telah terbukti efisien. Implementasi Multinomial Naïve Bayes berhasil meningkatkan akurasi, presisi, dan *recall* dibandingkan dengan metode klasifikasi lainnya. Implementasi metode Multinomial Naïve Bayes mencapai akurasi tertinggi sebesar 70% pada K-Fold 10 bagian 8, presisi tertinggi sebesar 70% pada K-Fold 10 bagian 7, dan *recall* tertinggi sebesar 73% pada K-Fold 10 bagian 8. Berdasarkan hasil penelitian, dapat disimpulkan bahwa *game My Lovely Daughter* mendapat ulasan positif dari pengguna di platform Steam.

**Kata kunci** : *My Lovely Daughter, Steam, Analisis Sentimen, Multinomial Naïve Bayes*

#### Abstract

*With the advancement of information technology, platforms for the sale and purchase of electronic content such as video games are increasingly developing, examples of which include Steam, Epic Games Store, Origin, and Uplay. Evaluating the quality of a game often involves analyzing reviews from players, but not all reviews can clearly be categorized as recommended or not. An example is the game My Lovely Daughter, developed by an Indonesian developer and available on the Steam platform. The problem lies in the clarity of categories in reviews, which can affect the development of future games. Therefore, sentiment analysis is needed to identify whether the reviews are positive or negative. In this case, the Multinomial Naïve Bayes classification method has proven to be efficient. The implementation of Multinomial Naïve Bayes successfully increased accuracy, precision, and recall compared to other classification methods. The implementation of the Multinomial Naïve Bayes method achieved the highest accuracy of 70% in 8 parts of K-Fold 10, the highest precision of 70% in 7 parts of K-Fold 10, and the highest recall of 73% in 8 parts of K-Fold 10. Based on the research results, it can be concluded that the game My Lovely Daughter received positive reviews from users on the Steam platform.*

**Keywords** : *My Lovely Daughter, Steam, Sentiment Analysis, Multinomial Naïve Bayes*

## 1. PENDAHULUAN

Bersamaan dengan perkembangan teknologi informasi, banyak platform penjualan dan pembelian konten elektronik seperti *video game* yang tersedia, seperti Steam, Epic Games Store, Origin, dan Uplay. Untuk menilai kualitas suatu *game*, penting untuk memperhatikan ulasan dari para pemain untuk mengetahui apakah *game* tersebut menarik atau tidak. Salah satu contoh *game* yang dikembangkan oleh developer Indonesia dan tersedia di platform Steam adalah *My Lovely Daughter*. *Game* ini mendapatkan tanggapan positif dari pengguna Steam, dengan ulasan yang bagus dan jumlah unduhan yang cukup tinggi untuk *game* lokal Indonesia. Akan tetapi, terdapat ulasan yang kurang tepat dalam kategori ulasan, apakah direkomendasikan atau tidak, yang dapat mempengaruhi pengembangan *game* selanjutnya. Maka dari itu, diperlukan analisis sentimen untuk mengevaluasi ulasan-ulasan tersebut.

Analisis sentimen adalah proses menganalisis teks digital untuk menentukan apakah pesan tersebut bersifat positif atau negatif. Dalam konteks ulasan *game*, terkadang terdapat ulasan dengan kategori yang tidak jelas atau sulit dipahami. Untuk mengatasi hal ini, penelitian dilakukan untuk melihat bagaimana ulasan yang diberikan oleh pemain atau pengguna *game* *My Lovely Daughter*, sehingga dapat diperoleh informasi tentang sentimen positif atau negatif dari ulasan tersebut. Salah satu metode yang digunakan dalam analisis sentimen adalah klasifikasi menggunakan metode Multinomial Naïve Bayes.

Menurut Cindo Mona (2019), dalam penelitiannya memaparkan beberapa metode dalam klasifikasi sentimen, metode Multinomial Naïve Bayes terbukti lebih efisien dibanding metode yang lain. Penelitian ini mengimplementasikan metode Multinomial Naïve Bayes. Implementasi metode ini dilakukan pada *dataset* ulasan *game* *My Lovely Daughter* untuk mengetahui signifikansi peningkatan akurasinya. Algoritma Multinomial Naïve Bayes merupakan perkembangan dari metode Naïve Bayes yang digunakan dalam menganalisis sentimen. Berbagai penelitian menjelaskan bahwasannya

algoritma Multinomial Naïve Bayes dikenal memiliki tingkat akurasi yang tinggi (Verawati & Audit, 2022).

Pada penelitian Purnomoputra Riko B. (2019), Metode Multinomial Naïve Bayes dipilih karena mampu menghitung dengan efisien dalam proses klasifikasi data dan efektif dalam menangani masalah yang melibatkan banyak data. Menggunakan teknik klasifikasi Multinomial Naïve Bayes, pengklasifikasian manual data *training* diharapkan dapat berjalan lebih singkat dan efisien. Selain itu,

## 2. TINJAUAN PUSTKA

### A. *My Lovely Daughter*

*My Lovely Daughter* adalah sebuah permainan video yang dapat dimainkan di platform Steam. Permainan ini mengharuskan pemain untuk menjadi seorang alkemis yang harus menciptakan homunculus atau manusia buatan untuk menghidupkan kembali putrinya yang telah meninggal. Permainan ini menggabungkan simulasi, strategi, dan elemen permainan peran karena pemain perlu mengelola sumber daya, bereksperimen dengan berbagai bahan alkimia, dan membuat keputusan moral yang sulit untuk mencapai tujuan. Dengan visual yang indah, musik latar yang atmosferik, dan cerita yang mendalam, *My Lovely Daughter* menawarkan pengalaman bermain yang menarik (Toge Productions).

### B. Steam

Steam adalah platform digital yang memungkinkan pengguna untuk membeli, mengunduh, dan memainkan berbagai macam *game* secara *online*. Steam juga menyediakan fitur komunitas yang memungkinkan pengguna untuk berinteraksi, berbagi konten, dan membuat serta bergabung dalam grup dengan minat yang sama. Platform ini dikelola oleh Valve Corporation dan merupakan salah satu platform distribusi *game* digital terbesar dan paling populer di dunia (Jayana, 2022).

### C. Analisis Sentimen

Analisis sentimen adalah metode komputasi yang menggunakan teknologi *natural language processing* (NLP) dan algoritma *machine learning* untuk

mengidentifikasi, mengekstrak, dan mengukur informasi subjektif yang terkandung dalam teks. Tujuan utamanya adalah untuk memahami perasaan, opini, dan sentimen yang terkandung dalam teks, baik itu positif, negatif, maupun netral. Dengan menggunakan analisis sentimen, kita dapat mengeksplorasi pandangan dan tanggapan orang terhadap suatu topik atau produk berdasarkan teks yang mereka hasilkan, seperti ulasan, komentar, atau postingan media sosial. Metode ini memiliki berbagai aplikasi yang luas, termasuk dalam bidang bisnis untuk mengukur kepuasan pelanggan, dalam penelitian pasar untuk memahami tren dan preferensi konsumen, serta dalam analisis media sosial untuk memantau citra merek dan respons terhadap kampanye pemasaran (Hasri C F et al. 2022).

#### D. Text Mining

*Text mining*, juga dikenal sebagai *text data mining*, merujuk pada proses ekstraksi informasi yang signifikan dari teks yang tidak terstruktur. Tujuannya adalah untuk mengidentifikasi pola, tren, dan pengetahuan tersembunyi dalam teks. Teknik *text mining* melibatkan pemrosesan teks, analisis sentimen, klasifikasi dokumen, ekstraksi informasi, dan sebagainya. Teknologi ini umumnya digunakan di berbagai bidang seperti ilmu sosial, ekonomi, bisnis, kesehatan, dan lainnya untuk mendapatkan wawasan berharga dari teks yang besar. (Rismayani et al., 2022).

#### E. Preprocessing Text

*Preprocessing teks* adalah proses membersihkan dan mengubah data teks mentah menjadi format yang dapat dengan mudah dianalisis dan diproses oleh algoritma *machine learning*. Langkah-langkah ini membantu mengurangi gangguan dan mengekstrak informasi penting dari data teks (Sharma C et al., 2022). Tahapan pada *preprocessing text* adalah sebagai berikut:

##### 1. Cleaning.

Proses *cleaning* merupakan tahap awal dari *preprocessing*. Pada tahap ini, bagian tertentu dari ulasan dihapus seperti simbol, tanda baca dan mengubah huruf kapital menjadi huruf kecil.

##### 2. Tokenizing.

Tahap berikutnya adalah *tokenizing* yang merupakan tahapan pemecahan kalimat menjadi kata.

##### 3. Stopword removal.

*Stopword removal* adalah menghapus kata hubung atau kata yang tidak memiliki banyak arti.

##### 4. Stemming.

Proses *stemming* ini mengubah kata yang memiliki akhiran kata menjadi bentuk kata dasarnya, misalnya *dogs* menjadi *dog*, *gamer* menjadi *game*.

#### F. Naïve Bayes

Naïve Bayes adalah salah satu algoritma klasifikasi probabilistik yang sering digunakan dalam analisis data. Algoritma ini dikenal karena sederhana namun efektif dalam banyak kasus. Pendekatan ini mengandalkan teorema Bayes untuk menghitung probabilitas kondisional dari suatu kelas berdasarkan fitur-fitur yang terkait, dengan asumsi bahwa semua fitur tersebut saling independen. Dengan kata lain, Naïve Bayes menganggap bahwa setiap fitur berkontribusi terhadap probabilitas suatu kelas secara independen, meskipun pada kenyataannya fitur-fitur tersebut mungkin saling berkorelasi. Meskipun sederhana, Naïve Bayes telah terbukti cukup efektif dalam berbagai aplikasi klasifikasi, terutama dalam analisis teks dan klasifikasi dokumen. Metode ini sering digunakan dalam pengenalan pola, analisis sentimen, dan filter spam email (Syahputra R et al., 2022). Adapun persamaan dari Naïve Bayes sebagai berikut.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Keterangan:

- A = Data dengan kelas yang belum diidentifikasi.
- B = Merupakan kelas spesifik.
- $P(B|A)$  = Probabilitas kelas B berdasarkan A (*prior* probabilitas).
- $P(B)$  = Probabilitas B yang dicari dari keseluruhan data (*prior* Probabilitas).
- $P(A|B)$  = Probabilitas tiap A berdasarkan kondisi pada kelas.
- $P(A)$  = Probabilitas A.

### G. Multinomial Naïve Bayes

Metode Multinomial Naïve Bayes merupakan salah satu varian dari algoritma Naive Bayes yang dikhususkan untuk pengolahan teks, terutama dalam konteks klasifikasi teks. Algoritma ini sering digunakan dalam berbagai aplikasi seperti analisis sentimen, klasifikasi dokumen, dan tugas-tugas lain yang melibatkan data teks. Kelebihan dari metode ini terletak pada kemampuannya dalam mengatasi data teks yang memiliki dimensi yang besar dan jarang, serta kemampuannya dalam menangani masalah klasifikasi dengan cepat meskipun dengan asumsi sederhana yang dimiliki oleh algoritma Naive Bayes. Dalam konteks analisis sentimen, metode ini digunakan untuk mengklasifikasikan teks menjadi kategori sentimen yang berbeda, seperti positif, negatif, atau netral, berdasarkan pada kemunculan kata-kata tertentu dalam teks tersebut (Syahputra R et al., 2022). Adapun persamaan dari Multinomial Naïve Bayes sebagai berikut:

$$P(w_i, c) = \frac{1 + \text{term}(w_i, c)}{|V| + \text{term}(c)}$$

Keterangan:

- $\text{term}(w_i, c)$  = Jumlah kemunculan fkata  $w_i$  pada kelas  $c$  (positif dan negatif).
- $\text{term}(c)$  = Jumlah dari semua fkata  $w$  pada kelas  $c$  (positif dan negatif).
- $|V|$  = Semua kosakata (*vocabulary*) yang muncul di dalam dokumen.
- Angka 1 = Konstanta, yaitu nilai *smoothing* untuk menghindari *error*.

### H. K-Fold Cross Validation

K-Fold Cross Validation adalah teknik validasi model yang digunakan untuk mengevaluasi keakuratan model yang dibangun berdasarkan dataset. Metode ini bertujuan untuk mencari tingkat akurasi optimal. Dalam K-Fold Cross Validation, dataset dibagi menjadi  $k$  *subset* yang sama besar. Kemudian, model dibangun menggunakan  $k-1$  *subset* sebagai data latih dan *subset* yang tersisa sebagai data uji. Proses ini diulang  $k$  kali dengan setiap subset

digunakan satu kali sebagai data uji. Hasil akhir dari  $K$  iterasi ini adalah rata-rata akurasi dari setiap iterasi, yang memberikan perkiraan yang lebih baik tentang seberapa baik model akan berperforma pada data baru yang tidak terlihat sebelumnya (Putra J., 2019).

### I. Confusion Matrix

Confusion Matrix merupakan alat evaluasi yang digunakan dalam proses data mining untuk mengukur performa hasil klasifikasi. Metode ini menggunakan tabel yang memuat empat kombinasi yang berbeda antara nilai yang diprediksi oleh model dan nilai yang sebenarnya (Pamuji A et al., 2022). Tabel ini terdiri dari empat istilah yang mewakili hasil klasifikasi (Pratama L I et al., 2020). Tabel kombinasi untuk metode Confusion Matrix sebagai berikut.

**Tabel 1.** Confusion Matrix

Predict Values	Actual Values	
	Positif (+)	Negatif (-)
Positif (+)	TP	FP
Negatif (-)	FN	TN

Sumber: Pratama, 2020

Keterangan:

1. *True Positive* (TP) untuk jumlah data yang terklasifikasi dengan benar sebagai positif.
2. *True Negative* (TN) untuk jumlah data yang terklasifikasi dengan benar sebagai negatif.
3. *False Positive* (FP) digunakan ketika data yang sebenarnya negatif tetapi terklasifikasi sebagai positif.
4. *False Negative* (FN) digunakan ketika data yang sebenarnya positif terklasifikasi sebagai negatif.

Persamaan model yang digunakan untuk mengukur kinerja klasifikasi sebagai berikut:

1. Akurasi adalah perbandingan informasi yang diprediksi secara benar oleh sistem dari seluruh informasi yang dirumuskan sebagai berikut.

$$\text{akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

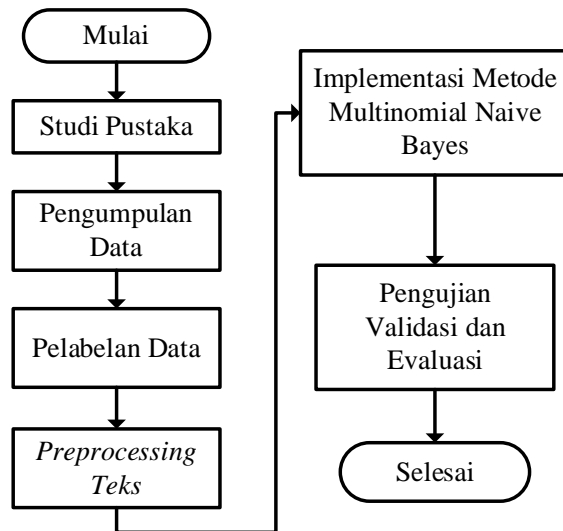
2. Presisi adalah suatu tingkat ketepatan dari data yang telah diuji dengan prediksi jawaban yang dilakukan oleh sistem yang dirumuskan sebagai berikut.

$$presisi = \frac{TP}{TP + FP} \times 100\%$$

3. *Recall* adalah suatu tingkat kedekatan dari sebuah nilai prediksi dengan nilai aktual yang dirumuskan sebagai berikut.

$$recall = \frac{TP}{TP + FN} \times 100\%$$

### 3. METODOLOGI PENELITIAN



**Gambar 1.** Diagram Alur Penelitian  
 Sumber: Laporan Hasil Penelitian (Bab 3)

#### A. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah sebanyak 1.000 data ulasan game My Lovely Daughter pada platform Steam, yang diambil pada tanggal 17 Februari 2023. Pengumpulan data dilakukan menggunakan teknik *scraping* dengan bahasa pemrograman Python melalui Jupyter Notebook pada Anaconda Terminal.

#### B. Pelabelan Data

Setelah pengumpulan data selesai, langkah selanjutnya adalah melakukan pelabelan pada data tersebut. Proses pelabelan dilakukan dengan membagi data menjadi dua kategori sentimen, yaitu sentimen positif dan sentimen negatif. Pelabelan data dilakukan secara manual dan divalidasi oleh dosen ahli bahasa Inggris yang kompeten dari Program Studi Pendidikan Bahasa Inggris Universitas Muhammadiyah Jember.

#### C. Preprocessing Text

Setelah data ulasan melalui proses pelabelan dan validasi, langkah berikutnya adalah melakukan *preprocessing teks*. Proses ini meliputi beberapa tahapan, termasuk pembersihan data (*cleaning*), pembentukan token (*tokenizing*), penghapusan kata hubung (*stopword removal*), dan penyederhanaan kata (*stemming*). Contoh dari *preprocessing text* untuk ulasan adalah sebagai berikut.

**Tabel 2.** Ulasan

Ulasan
<i>This game is very easy to play, and I really like it.</i>

Sumber: Laporan Hasil Penelitian (Bab 3)

##### 1. *Cleaning*

*Cleaning* merupakan tahap untuk mengubah huruf kapital menjadi huruf kecil dan menghapus tanda baca serta angka pada data.

**Tabel 3.** *Cleaning*

<i>Cleaning</i>
<i>this game is very easy to play and i really like it</i>

Sumber: Laporan Hasil Penelitian (Bab 3)

##### 2. *Tokenizing*

*Tokenizing* merupakan tahapan pembagian setiap kata yang ada pada kalimat.

**Tabel 4.** *Cleaning*

<i>Tokenizing</i>
<i>'this', 'game', 'is', 'very', 'easy', 'to', 'play', 'and', 'I', 'really', 'like', 'it'</i>

Sumber: Laporan Hasil Penelitian (Bab 3)

##### 3. *Stopword Removal*

*Stopword removal* berfungsi menghilangkan kata-kata umum yang tidak memiliki banyak arti dalam sebuah kalimat.

**Tabel 5.** *Stopword Removal*

<i>Stopword Removal</i>
<i>'game', 'easy', 'play', 'really', 'like'</i>

Sumber: Laporan Hasil Penelitian (Bab 3)

##### 4. *Stemming*

*Stemming* ialah proses mengubah kata yang memiliki akhiran kata menjadi bentuk kata dasarnya.

**Tabel 6.** *Stemming*

<i>Stemming</i>
<i>game easy play really like</i>

Sumber: Laporan Hasil Penelitian (Bab 3)

#### D. Multinomial Naïve Bayes

Langkah selanjutnya dalam analisis ini adalah melakukan klasifikasi menggunakan metode Multinomial Naïve Bayes. Pada metode ini, setiap kata dalam dokumen akan dihitung nilai kemunculannya. Dengan demikian, metode ini akan mempertimbangkan frekuensi kemunculan setiap kata dalam proses klasifikasi, sehingga dapat memberikan gambaran yang lebih akurat tentang sentimen yang terkandung dalam dokumen tersebut.

#### E. Evaluasi dan Validasi

Tahap terakhir yaitu, setelah dilakukannya pengklasifikasian menggunakan metode Multinomial Naïve Bayes akan dilakukan pengukuran dan pengujian hasil dengan persamaan Confusion Matrix dan K-Fold Cross Validation.

### 4. HASIL DAN PEMBAHASAN

#### A. Evaluasi dan Validasi Hasil

Evaluasi hasil pengujian klasifikasi data, pada metode Multinomial Naïve Bayes tanpa seleksi fitur dan dengan seleksi fitur Complete Gini Index Text, disusun dengan memilih nilai indikator akurasi, presisi, dan *recall* tertinggi. Berdasarkan Tabel 7, Multinomial Naïve Bayes tanpa seleksi fitur diperoleh nilai akurasi tertinggi sebesar 70% pada 10-Fold bagian 8, nilai presisi tertinggi sebesar 70% pada 10-Fold bagian 7, dan nilai *recall* tertinggi sebesar 73% ada 10-Fold bagian 8. Tabel 7 di bawah menampilkan nilai tertinggi untuk setiap metode yang diuji.

**Tabel 7.** Hasil Pengujian K-Fold

Metode	Nilai Tertinggi Pengujian		
	K-Fold Cross Validation		
	K	Bagian	Nilai Indikator Tertinggi
MNNB	10	8	Akurasi = 70%
	10	7	Presisi = 70%
	10	8	Recall = 73%

Sumber: Laporan Hasil Penelitian (Bab 4)

#### B. Hasil dan Analisis

Dari hasil pengujian evaluasi dan validasi menggunakan K-Fold 2, 5, 7, dan 10 menggunakan metode Multinomial Naïve Bayes tanpa seleksi fitur dan dengan seleksi fitur Complete Gini Index Text, diperoleh hasil

rangkaiannya rata-rata akurasi, presisi, dan *recall* yang ditampilkan pada Tabel 8 berikut.

**Tabel 8.** Hasil Rata-Rata

Metode yang Diuji	Rata-rata			
	FOLD	Akurasi	Presisi	Recall
MNNB	2	59%	59%	58%
	5	62%	62%	62%
	7	63%	63%	65%
	10	60%	60%	61%

Sumber: Laporan Hasil Penelitian (Bab 4)

Berdasarkan Tabel 8, Multinomial Naïve Bayes berhasil meningkatkan akurasi, presisi, dan *recall*. Multinomial Naïve Bayes mencapai rata-rata akurasi tertinggi sebesar 63% dengan pencapaian tertinggi terjadi pada 7-Fold. Multinomial Naïve Bayes mencapai rata-rata presisi tertinggi yang sama sebesar 63% dengan pencapaian tertinggi terjadi pada 7-Fold. Multinomial Naïve Bayes mencapai rata-rata *recall* tertinggi sebesar 65% pada 7-Fold. Penelitian ini menunjukkan bahwa performa sistem cukup baik dalam mencari nilai akurasi, presisi, *recall*.

### 5. KESIMPULAN DAN SARAN

#### A. KESIMPULAN

Berdasarkan penelitian yang sudah dilakukan, implementasi metode Multinomial Naïve Bayes untuk analisis sentimen ulasan video game My Lovely Daughter pada situs Steam, memberikan kesimpulan sebagai berikut.

1. Implementasi metode Multinomial Naïve Bayes menghasilkan performa yang baik untuk analisis sentimen ulasan game My Lovely Daughter.
2. Implementasi metode Multinomial Naïve Bayes diperoleh.
  - a. Nilai akurasi tertinggi sebesar 70% pada 10-Fold bagian 8.
  - b. Nilai presisi tertinggi sebesar 70% pada 10-Fold bagian 7.
  - c. Nilai *recall* tertinggi sebesar 73% pada 10-Fold bagian 8.
3. Hasil penelitian yang telah dilakukan memberikan kesimpulan bahwa game My Lovely Daughter memperoleh review yang baik bagi penggunaannya pada platform Steam.

## B. SARAN

Berdasarkan evaluasi yang telah dilakukan terhadap penelitian ini, rekomendasi untuk pengembangan sistem guna mencapai hasil yang lebih optimal adalah sebagai berikut.

1. Dalam melakukan analisis sentimen, selain metode klasifikasi Multinomial Naïve Bayes, juga perlu dieksplorasi metode klasifikasi lainnya seperti Support Vector Machine (SVM), Decision Tree, dan lain-lain. Penelitian lebih lanjut diperlukan untuk membandingkan efektivitas metode klasifikasi ini dengan Multinomial Naïve Bayes. Dengan demikian, dapat diperoleh hasil yang optimal melalui pemilihan metode klasifikasi yang paling sesuai dengan karakteristik data dan tujuan analisis sentimen yang diinginkan.
2. Membahakan metode pembobotan kata, seperti TF-IDF, PCA, Information Gain, dengan penambahan metode pembobotan kata diharapkan pada penelitian selanjutnya dapat meningkatkan akurasi dan efisiensi dari sistem yang akan dibuat.
3. Pada penelitian ini, digunakan sebanyak 1000 data ulasan sebagai sampel. Untuk penelitian mendatang, disarankan untuk menambah jumlah data ulasan guna meningkatkan performa sistem yang telah dikembangkan. Dengan penambahan data, diharapkan sistem dapat lebih akurat dan dapat memberikan hasil yang lebih dapat diandalkan dalam analisis sentimen.

## 6. DAFTAR PUSTAKA

- Arifianto, D., Umilasari, R., & Wahyudi, E. (2020). Pemanfaatan Media Sosial dan Marketplace Untuk Meningkatkan Pemasaran Produksi UMKM Bonpay dan Arumfood. *Jurnal Pengabdian Masyarakat IPTEKS*, 6(2), 155-160.
- Bakti, B. S., & Halim, M. (2016). Penerapan Fungsi Linier Untuk Penentuan Komponen Penilaian Kesehatan Koperasi. *JUSTINDO (Jurnal Sistem dan Teknologi Informasi Indonesia)*, 1(2).
- Cindo, M., Rini, D. P., & Ermatita, E. (2019). Literatur Review: Metode Klasifikasi pada Sentimen Analisis. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) (Vol. 1, No. 1)*.
- Hasri, C. F., & Alita, D. (2022). Penerapan Metode Naïve Bayes Classifier dan Support Vector Machine pada Analisis Sentimen Terhadap Dampak Virus Corona di Twitter. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 3(2), 145-160.
- Imadudin, F. R., & Murdiansyah, D. T. (2021). Implementation of Naïve Bayes and Gini Index for Spam Email Classification. *Indonesia Journal on Computing (Indo-JC)*, 6(1), 85-94.
- Jayana, M. A. P., & Priharsari, D. (2022). Analisis Pengalaman Pengguna pada Situs Distribusi Digital menggunakan Metode Usability Testing dan User Experience Questionnaire (Studi pada store.steampowered.com dan gog.com). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(2), 527-536.
- Pamuji, A. (2022). Prediksi Otorisasi Pengguna Sistem Berkas pada Algoritma Klasifikasi dengan Teknik Naïve Bayes. *INFOMATEK: Jurnal Informatika, Manajemen dan Teknologi*, 24(1), 35-44.
- Pratama, L. I., Lestanti, S., & Wulansari, Z. (2023). Penerapan Metode Weighted Sum Model pada Sistem Seleksi Supplier di UD. Sumber Besi Berbasis Web. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(2), 1157-1161.
- Purnomoputra, R. B., Adiwijaya, A., & Wisesty, U. N. (2019). Sentiment

Analysis of Movie Review Using Naïve Bayes Method with Gini Index Feature Selection. *Journal of Data Science and Its Applications*, 2(2), 85-94.

Putra, J. W. G. (2019). *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning*. Tokyo. Jepang.

Rintyarna, B. S., Sarno, R., & Fatichah, C. (2020). Enhancing the Performance of Sentiment Analysis Task on Product Reviews by Handling Both Local and Global context. *International Journal of Information and Decision Sciences*, 12(1), 75-101.

Rismayani, R., Hasyrif, S. Y., Darwansyah, T., & Mansyur, I. (2022). Implementasi Algoritma Text Mining dan Cosine Similarity untuk Desain Sistem Aspirasi Publik Berbasis Mobile. *Komputika: Jurnal Sistem Komputer*, 11(2), 169-176.

Sharma, C., Sakhuja, S., & Nijjer, S. (2022). Recent Trends of Green Human Resource Management: Text Mining and Network Analysis. *Environmental Science and Pollution Research*, 29(56), 84916-84935.

Syahputra, R., Yanris, G. J., & Irmayani, D. (2022). SVM and Naïve Bayes Algorithm Comparison for User Sentiment Analysis on Twitter. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 7(2), 671-678.

Toge Productions. (2018). *My Lovely Daughter* by Toge Productions. <https://togeproductions.itch.io/my-lovely-daughter>.