
Analisis Sentimen Ulasan *Game Harry Potter: Hogwarts Mystery* Pada Situs Google Play Menggunakan *Naïve Bayes Classifier*

Puti Ayu Andhini Rahman¹, Ermatita², Helena Nurramdhani Irmanda³

Fakultas Ilmu Komputer

^{1,3}Universitas Pembangunan Nasional Veteran Jakarta

²Universitas Sriwijaya

Email: putiaar@upnvj.ac.id¹, ermatitaz@yahoo.com², irmandahelena@gmail.com³

Jalan RS. Fatmawati Pondok Labu Jakarta Selatan, DKI Jakarta 12450, Indonesia

Abstrak

Salah satu dampak dari pandemi Covid-19 adalah semua orang harus berada di rumah untuk meminimalisir penyebaran virus dan mereka merasa jenuh karena taman rekreasi atau bermain tutup, untuk menghilangkan rasa jenuh tersebut banyak dari mereka yang mencari hiburan melalui permainan video baik *online* maupun *offline*. *Harry Potter: Hogwarts Mystery* merupakan permainan RPG bermain peran yang berlatarkan cerita *Harry Potter* berdasarkan seri novel JK Rowling yang dikembangkan oleh Jam City dengan lisensi dari Portkey Games. Ulasan pengguna merupakan salah satu hal penting untuk dijadikan pertimbangan developer. Untuk memantau ulasan tersebut, metode yang dilakukan pada penelitian ini menyangkut kegiatan pengumpulan data menggunakan *web scrapper*, melalui *preprocessing* data, pembobotan kata, dan teknik yang digunakan untuk klasifikasi sentimen menggunakan *Naïve Bayes Classifier* yang dapat mengkategorikan ulasan pengguna tanpa harus melakukannya secara manual. Tujuan dari penelitian ini adalah untuk melakukan klasifikasi terhadap ulasan tersebut dan mendapatkan informasi yang berguna bagi pihak yang berkepentingan.

Kata Kunci: Analisis Sentimen, Google Play, *Naïve Bayes Classifier*

1 PENDAHULUAN

Di masa pandemi Covid-19, permainan video game mengalami kecenderungan kenaikan minat yang pesat, menurut Chris Bullen dan Jinsong Chen (2020) dilansir dari auckland.ac.nz laman University of Auckland dengan judul “*Video games and mental health during Covid-19: Opportunities and precautions*” bahwa, “Twitch, yang merupakan platform streaming video game yang sedang naik daun, mencatat 1,49 miliar jam *gaming hours* pada April 2020, meningkat tajam sebesar 50% selama Maret”. Adapun Newzoo yang merupakan perusahaan yang memiliki basis data lengkap mengenai industri *gaming* dan *e-sports* telah mengeluarkan Laporan Pasar *Game Global Platform Analytics* dan memperkirakan bahwa pasar game global tahun 2020 akan menghasilkan pendapatan sebesar \$ 159,3 miliar dengan pertumbuhan tahun ke tahun sebesar + 9,3%. Menurut laporan Newzoo, semua segmen game mengalami peningkatan dalam keterlibatan dan pendapatan yang merupakan dampak Covid-19, namun segmen game seluler adalah yang mengalami peningkatan terbesar. Secara total, game seluler diperkirakan akan menghasilkan pendapatan sebesar \$ 77,2 miliar pada tahun 2020, tumbuh sekitar +13,3% dari tahun ke tahunnya.

Dikarenakan menjamurnya game dan berakibat menimbulkan persaingan ketat, maka penulis mengambil topik tentang analisis sentimen pengguna terhadap game. Game yang akan dilakukan analisis sentiment adalah *Harry Potter: Hogwarts Mystery*, yang merupakan sebuah game pilihan editor pada Google Play Store. Karena game ini RPG (*Role Playing Game*), para pengguna biasanya sudah memiliki ekspektasi tertentu terhadap game ini, sehingga pemenuhan ekspektasi pengguna dan *user experience* menjadi poin penting untuk diperhatikan dengan harapan developer dapat mengetahui kelemahan yang ada di dalam gamenya, sehingga game ini dapat bertahan dan bersaing dengan kompetitor game lain.

Untuk proses analisisnya, pada penelitian ini dilakukan klasifikasi data ulasan pengguna untuk mengidentifikasi ulasan-ulasan baik yang bersifat positif maupun negatif menggunakan metode naïve bayes classifier. Penulis memilih metode ini karena naïve bayes merupakan metode klasifikasi yang dapat menghasilkan keakuratan klasifikasi yang tinggi dan independen, sebagaimana menurut [1] *naïve bayes classifier* adalah contoh paling sederhana dari pengklasifikasi probabilistik. Hasil dari $P(C|D)$ dari pengklasifikasi probabilistik adalah bahwa dokumen d milik kelas C. Setiap dokumen berisi istilah-istilah yang diberi probabilitas berdasarkan jumlah kejadian dalam dokumen tertentu. Menurut [2] kekhasan dari *naïve bayes classifier* ini adalah anggapan atau dugaan yang sangat kuat akan independensi dari masing-masing kondisi atau kejadian.

Berdasarkan penelitian-penelitian sebelumnya yang menggunakan *naïve bayes classifier*, hasil evaluasi algoritma yang dilakukan dapat dikatakan sudah baik, kondisi ini diperkuat dengan adanya penelitian dari [3] yang mengimplementasikan algoritma *naïve bayes* pada analisis sentimen terhadap opini film di twitter yang menggunakan 500 tweet dan menghasilkan akurasi sebesar 90%, presisi 92% dan recall 90% yang jika dibandingkan dengan klasifikasi *supervised learning* lain, contohnya KNN pada penelitian [4] yang meneliti tentang opini masyarakat pada media sosial twitter dengan 400 tweet menghasilkan uji akurasi menggunakan pembobotan tekstual sebesar 82,50%, pembobotan non-tekstual 60% dan penggabungan keduanya sebesar 83,33%.

Untuk itu, penulis dapat menarik kesimpulan bahwa *naïve bayes classifier* dapat bekerja dengan lebih optimal sebagai teknik pengklasifikasi dan penelitian ini diharapkan mampu mengklasifikasikan teks dengan baik untuk mendapatkan informasi yang dapat berguna bagi pihak yang berkepentingan.

2 TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Menurut [5] Analisis sentimen memiliki tugas pokok yaitu mengelompokkan polaritas dari teks yang ada didalam kalimat suatu dokumen, kalimat, atau pendapat. Polaritas ini memiliki arti dimana teks tersebut memuat lebih banyak aspek positif atau negatif. Dengan melakukan analisis sentimen, diharapkan pengujian analisis dapat mengetahui gambaran besar sentimen publik terhadap suatu produk atau jasa, apakah opini pengguna lebih cenderung kepada positif atau negatif, dimana hasil analisis yang dilakukan dapat membantu pengembang produk atau jasa dalam melakukan perbaikan dan juga sebagai bahan pertimbangan para pemangku kepentingan untuk melakukan inovasi terhadap produk mereka.

2.2 Game

Menurut [6] game dapat diartikan secara garis besar oleh Suits (1967b: 156) yaitu “aktivitas yang diarahkan menuju suatu keadaan tertentu menggunakan cara-cara yang telah diatur oleh aturan khusus. Dimana aturan tersebut memiliki cakupan terbatas untuk memungkinkan aktivitas tersebut”. Menurut [7] Permainan video adalah permainan yang kami mainkan berkat alat audiovisual dan dapat didasarkan melalui sebuah cerita. RPG (*Role Playing Game*) adalah salah satu genre dimana para pemainnya dapat memainkan peran karakter yang berlatar fiksi, dimana pemain memiliki tanggung jawab (*task*) yang sudah tercantum dalam aturan permainan sebagai acuan pedoman formal game.

2.3 Text Mining

Menurut [8] *Text Mining* adalah menemukan informasi baru dari data tekstual yang sebelumnya tidak teridentifikasi atau informasi yang bersifat rahasia dengan mengekstraknya menggunakan teknik yang berbeda. Tantangan utama dalam *text mining* adalah ketika dokumen memiliki ratusan atau bahkan ribuan istilah atau kata, yang mana hal ini akan meningkatkan kompleksitas dalam klasifikasi teks.

2.4 Preprocessing

Menurut [9] *Preprocessing* adalah proses dari menyiapkan data mentah yang bertujuan untuk mengubah data teks yang tidak beraturan menjadi kumpulan data yang terstruktur. Tahapan *preprocessing* ini dilakukan dengan menyeleksi data-data yang akan diproses menjadi bahan yang lebih kecil dan terstruktur, seperti *data cleaning*, *tokenization*, *filtering*, *case folding*, *stemming*, pembobotan kata, dan TF-IDF.

2.3 Naïve Bayes Classifier

Menurut NBC [2] adalah suatu metode yang berguna untuk memprediksi kemungkinan di masa depan berdasarkan pengalaman pada masa sebelumnya, berdasarkan hal tersebut maka metode ini dikenal sebagai Teorema Bayes. Kekhasan dari *Naïve Bayes Classifier* ini adalah anggapan atau dugaan yang sangat kuat akan independensi dari masing-masing kondisi atau kejadian. Konsep Naive Bayes Classifier menggunakan konsep Teorema Bayes, nilai probabilitas dinyatakan sebagai berikut menurut Pop [5],

$$p(C / D) = \frac{p(D|C)p(C)}{p(D)}. \quad (2.1)$$

Dari rumus 2.1 diatas dapat diketahui bahwa:

$p(C | D)$ menunjukkan seberapa sering C terjadi setelah D terjadi

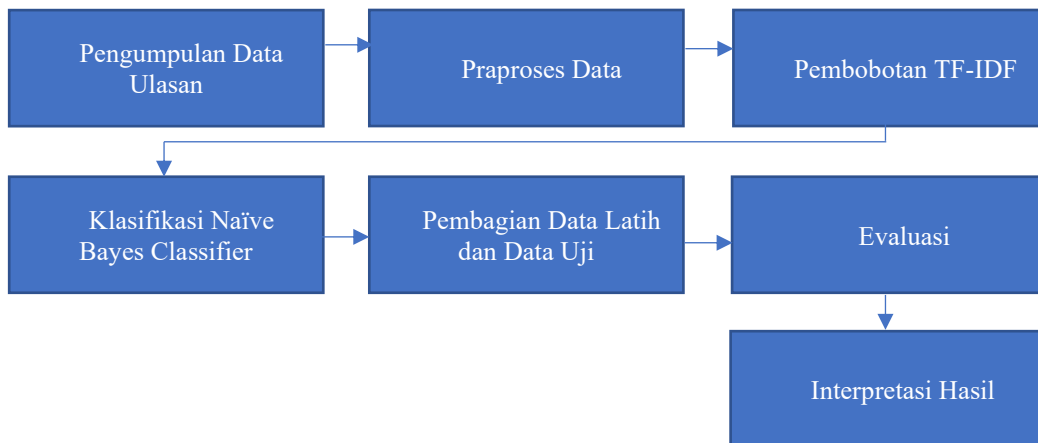
$p(D | C)$ menunjukkan seberapa sering D terjadi setelah C terjadi

$p(C)$ menunjukkan seberapa besar kemungkinan C adalah benar

$p(D)$ menunjukkan seberapa besar kemungkinan D adalah benar

3 METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah klasifikasi menggunakan *Naïve Bayes Classifier*, yang mana pada tahapan penelitian ini terdiri dari beberapa langkah, yaitu pengumpulan data ulasan, praproses data, pembobotan TF-IDF, klasifikasi *Naïve Bayes Classifier*, pembagian data latih dan data uji, evaluasi, dan interpretasi hasil. Berikut adalah penggambaran tahap penelitian.



Gambar. 1. Tahap Penelitian

2.1. Pengumpulan Data

Penelitian ini dilakukan proses *Web Scraping* untuk pengumpulan data ulasan dikarenakan kemudahan yang diperoleh, menurut [10] *web scraping* memiliki keuntungan karena memungkinkan peneliti untuk memperoleh kumpulan data baru yang belum tersentuh tanpa perlu penelitian hibah untuk mendanai pembelian peralatan mahal atau biaya kompensasi peserta. Pokok dari *web scraping* adalah mendapatkan data menggunakan pengambilan dan ekstraksi data dengan ukuran data yang beragam. Untuk mendapatkan data ulasan permainan ini pada basis data Google Play, penulis menggunakan teknik *scraping* data yang implementasinya menggunakan salah satu *extention* pada Chrome bernama Data Miner Pro, *extention* ini berfungsi sebagai alat bantu untuk memudahkan pengambilan

data ulasan tanpa harus melakukannya secara manual. Pengumpulan data ulasan yang akan digunakan dalam penelitian ini adalah dari *website Google Playstore* pada ulasan game Harry Potter: Hogwarts Mystery yaitu <https://play.google.com/store/apps/details?id=com.tinyco.potter&hl=in&gl=US>. Setelah selesai, data-data yang telah di *scrap* menggunakan *dataminer.io* tadi dapat di download pada laman tersebut. Data ulasan dapat disimpan ke dalam komputer baik dalam bentuk *csv* maupun *excel*. Pada proses analisisnya, penelitian ini dilakukan klasifikasi data ulasan pengguna untuk mengidentifikasi ulasan-ulasan baik yang bersifat positif maupun negatif dengan menggunakan metode *Naïve Bayes Classifier*.

2.2 Praproses

2.2.1 Terjemah ulasan Bahasa asing

Dalam 1.193 data tersebut, masih terdapat ulasan berbahasa Inggris, baik yang bersifat *full* Bahasa Inggris, maupun campuran Bahasa Indonesia – Bahasa Inggris. Untuk itu, dilakukan terjemah ulasan agar semua data ulasan berbahasa Indonesia. Contoh data *excel* sebelum di terjemah bisa dilihat pada tabel 3 berikut.

Tabel 1. Proses terjemah ulasan

Input ulasan bahasa asing	Output bahasa Indonesia
<i>The game is good but when I play I don't know why but my quest is kinda bug can you fix with that btw I'm still stuck on chapter 3</i>	Game ini bagus tetapi ketika saya bermain saya tidak tahu mengapa tapi tugas saya agak bug dapatkah Anda memperbaiki itu btw saya masih macet di bab 3

2.2.2 Pelabelan data

Pelabelan data dilakukan dengan memisahkan data menjadi dua kelas/label yaitu, label positif dan label negatif. Proses pelabelan dengan otomatis menggunakan *python*, sesuai dengan rating (*star*) yang diberikan oleh pengguna game yang telah di *scrap* pada halaman *website google play Harry Potter: Hogwarts Mystery*.

	nama_pengguna	tanggal_ulasan	ulasan	label
0	raikhan faqih	31 Januari 2021	Loading nya lama Gak bisa di mainin Game burik...	0
1	Tutorial Ngasal	31 Januari 2021	Ini gmn sih sya ngga bisa masuk ke dlm game da...	0
2	Zahra fitria Ramadhani	31 Januari 2021	Menurut gw game ini udh bagus, dari grafik nya...	1
3	Nn riini	31 Januari 2021	Halo, cuman mau bilang, game bagus banget dan ...	1
4	gilang aga	31 Januari 2021	Game nya terlalu bergantung sama energi.. mala...	0
5	Hanser Alja Nenobais	31 Januari 2021	Saya memiliki sedikit kendala dalam bermain ga...	0
6	Rawi Addinda D	31 Januari 2021	Gamenya bagus banget!!!!!!!!!!!!!!	1
7	Lila Nabilla11	30 Januari 2021	Assalamuallaikum nama saya neeta.aku kn suka b...	1
8	Janeeta erbilia	30 Januari 2021	Tolong energinya!!!!	1
9	Sindy Sindy	30 Januari 2021	Ngga bisa bahasa Inggris 🙄 Kebanyakan loading, ...	0

Gambar. 3. Proses pelabelan data

Dilakukan proses perubahan data rating menjadi label, pada *python* data rating yang berbunyi, “Diberi rating 5 bintang dari 5 bintang” akan diubah menjadi label 1, dan selain itu label 0.

2.2.3 Data Cleaning

Pada tahap ini data-data ulasan yang sifatnya ambigu dihapus karena terdapat tidak konsistensinya dalam memberikan ulasan, sehingga akan mengakibatkan keambiguan pada *machine learning* jika tetap dipakai. Contohnya, pada ulasan di *website* ini, ada beberapa ulasan yang berbunyi, “game ini jelek... tapi boong, ayo guys di download gamenya” dan sejenis lainnya. Ulasan-ulasan yang seperti ini dianggap tidak perlu dan dihapus dari korpus.

2.2.4 Tokenization

Tahap tokenisasi dilakukan untuk memisahkan teks, yang dalam konteks ini adalah data ulasan menjadi potongan-potongan kata yang tidak memiliki ikatan atau pengaruh satu sama lain, sehingga kata-kata tersebut menjadi sebuah kata yang independen. Proses tokenisasi ini membantu praproses untuk menghitung banyaknya kata atau frekuensi

kemunculan kata di dalam dataset karena tidak memiliki pengaruh dalam pemrosesan teks. Contoh proses tokenisasi sebagai berikut.

Loadingnya lama gak bisa di mainin Game burik kok di buat

Gambar. 4. Kalimat sebelum tokenisasi

Setelah melewati proses tokenisasi, maka setiap katanya akan terpisah dan menjadi seperti gambar berikut.

['Loading', 'nya', 'lama', 'Gak', 'bisa', 'di', 'mainin', 'Game', 'burik', 'kok', 'di', 'buat']

Gambar. 5. Proses tokenisasi

2.2.5 Filtering(stopwords)

Pada tahap *filtering* atau yang biasa disebut stopwords dilakukan penghapusan kata-kata umum pada korpus yang tidak akan menambah atau mengurangi bobot saat dilakukan pembobotan kata. Pada penelitian ini, stopwords yang digunakan adalah stopword berbahasa Indonesia dari *library* nltk, juga dilakukan *filtering* kata yang memiliki huruf kurang dari tiga. Sebagai contoh, kata: 'yg', 'sm', 'km' karena kata-kata tersebut tidak ada didalam *library* dan diasumsikan kata yang kurang dari tiga huruf tersebut tidak memiliki makna sehingga dapat dihapus dengan aman.

2.2.6 Stemming

Proses *stemming* dilakukan untuk mengolah kata yang memiliki imbuhan menjadi kata dasar dengan diberi asumsi bahwa kata tersebut memiliki arti kata yang sama sehingga proses pembobotan kata menjadi lebih mudah dilakukan. Proses ini memakan waktu yang lama pada saat *running* menggunakan Python namun hal tersebut tergantung seberapa banyak kalimat/kata di dalam dokumen yang diolah. Proses *stemming* yang dilakukan pada tahap ini dilakukan menggunakan *library* NLP (*Natural Language Programming*) bahasa Indonesia, yaitu *library* Sastrawi. Digunakan *library* ini karena Sastrawi walaupun sederhana, namun *library* ini tetap berkualitas dan terdokumentasi dengan baik.

2.3 Pembobotan TF-IDF

Pembobotan TF-IDF (*Term Frequency Inverse Document Frequency*) merupakan metode pembobotan kata yang berfungsi untuk menentukan hubungan antar kata pada dokumen dengan memberikan bobot pada setiap katanya. Metode pembobotan kata ini menjadi populer karena perhitungan akurasi yang cukup tinggi sehingga dapat melakukan pembobotan kata secara efisien dan mudah dengan tetap memiliki hasil yang baik. Pada penelitian ini, penulis menggunakan *library* yang tersedia pada Python yang memfasilitasi rumus TF-IDF yaitu *library* TfidfVectorizer.

2.4 Klasifikasi NBC

Pada proses klasifikasi NBC (*Naïve Bayes Classifier*) dilakukan proses pengklasifikasian berdasarkan metode probabilitas dan statistik untuk mendapatkan hasil akurasi klasifikasi dokumen yang telah diproses. Penelitian ini menggunakan algoritma Naïve Bayes Classifier dengan menggunakan metode Multinomial Naïve Bayes, dimana metode ini merupakan salah satu metode klasifikasi yang berguna untuk menghitung frekuensi masing-masing kemunculan kata pada sebuah dokumen. Multinomial Naïve Bayes digunakan karena algoritma ini memiliki beberapa keunggulan, yaitu memiliki tingkat akurasi yang tinggi, mudah untuk diimplementasikan, memakan waktu komputasi yang rendah, dan *error state* yang minimum [11].

2.5 Pembagian Data Latih dan Data Uji

Pada proses pembagian data *training* dan data *test* dilakukan proses pembentukan model klasifikasi agar nantinya hasil data latih dan data uji tersebut dapat menghasilkan informasi yang dapat digunakan atau berguna bagi prediksi kelas data baru.

Tabel 2. Pembagian data latih dan data uji

Persentase Data Training	Persentase Data Test	Jumlah Data Training	Jumlah Data Test	Total Data Keseluruhan
70%	30%	639	275	914
80%	20%	731	183	914

Berdasarkan Tabel 10 pembagian data latih dan data uji dilakukan sebanyak dua kali dan dengan total keseluruhan data menjadi 914 data yang sebelumnya sebanyak 1193 data yang telah tereduksi oleh tahap *preprocessing*.

Tabel 3. Normalisasi Data Sampel

Label	Ulasan_clean3
0	457
1	457

Data ini sudah dilakukan normalisasi data label 1 (positif) dan label 0 (negatif) dengan perbandingan 0.5: 0.5 sebanyak 457 data sampel. Dengan ini, data yang digunakan sudah seimbang.

2.6 Evaluasi

Pada tahap evaluasi, ada beberapa langkah yang dijalankan, yang pertama yaitu penggambaran kinerja dari NBC menggunakan *Confusion Matrix*. Hal ini sangat berguna untuk mengukur *Recall*, *Precision*, *Accuracy*, dan *F-Measure*. Pada penelitian ini *Confusion Matrix* yang digunakan adalah 2x2. Berikut ini adalah persamaan dari performansi masing-masing penilaian.

Recall: $\frac{\text{true positive}}{\text{true positive} + \text{false negative}}$

Precision: $\frac{\text{true positive}}{\text{true positive} + \text{false positive}}$

F1-Score: $2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$

Accuracy: $\frac{\text{true positive} + \text{true negative}}{\text{tp} + \text{fn} + \text{fp} + \text{tn}} \times 100\%$

Gambar.6. Rumus Evaluasi

2.7 Interpretasi Hasil

Berdasarkan hasil yang didapat dari klasifikasi Naïve Bayes Classifier pada game Harry Potter: Hogwarts Mystery, analisa sentiment pada produk game ini dapat dikatakan bahwa game ini sudah cukup baik namun ada beberapa masukan yang dapat diambil, seperti penambahan Bahasa Indonesia, kecepatan dalam loading game, maupun energi yang dalam hal ini dapat disimpulkan bahwa untuk memainkan game ini lebih lama dibutuhkan energi yang cukup banyak, sehingga tidak sebanding dengan permintaan/request tugas-tugas yang harus dijalankan dalam permainan.

4 HASIL DAN PEMBAHASAN

4.1 Evaluasi Algoritma Naïve Bayes

Penelitian ini dilakukan dua kali rangkaian percobaan pembagian data uji dan data latih, yaitu 70:30 dan 80:20.

4.1.1 Confusion Matrix

Dengan mengimport library pada python: `from sklearn.metrics import confusion_matrix` dapat diketahui:

- a. Confusion Matrix pada rangkaian pertama, yaitu 70% data latih berbanding 30% data uji.

Tabel 4. Confusion Matrix rangkaian pertama

Actual	Prediction	
	Positif	Negatif
Positif	123	15
Negatif	22	115

- b. Confusion Matrix pada rangkaian kedua, yaitu 80% data latih berbanding 20% data uji.

Tabel 5. Confusion Matrix rangkaian kedua

Actual	Prediction	
	Positif	Negatif
Positif	82	10
Negatif	12	79

Pada rangkaian uji pertama yang merupakan perbandingan dari 70% data latih dan 30% data tes, menghasilkan total 275 data tes. Data tes tersebut menghasilkan sentimen positif yang diprediksi dengan benar dan memiliki nilai

kebenaran sesungguhnya adalah 123 data, kemudian sebanyak 22 data ulasan positif yang diprediksi salah oleh model. Sedangkan untuk data ulasan yang negative, model prediksi memprediksi sebanyak 115 data ulasan yang diprediksi dengan benar bahwa ulasan tersebut negatif dan sebanyak 15 data yang diprediksi dengan salah oleh model.

Sedangkan pada rangkaian uji kedua, yang merupakan perbandingan data latih sebanyak 80% dan data tes sebanyak 20% menghasilkan total jumlah data tes keseluruhan yaitu 183 data, terdapat 82 data ulasan yang memiliki sentimen positif diprediksi dengan benar yang memiliki nilai kebenaran yang sebenarnya, kemudian sebanyak 12 data ulasan positif yang diprediksi salah oleh model. Sedangkan untuk data ulasan yang negatif, model prediksi memprediksi ada 79 data ulasan yang diprediksi dengan benar bahwa ulasan tersebut negatif dan sebanyak 10 data yang diprediksi dengan salah oleh model.

4.1.2 Recall

Pada tahap ini, recall digunakan untuk mendapatkan perbandingan antara jumlah prediksi positif data ulasan yang benar dibandingkan dengan keseluruhan data yang diprediksi positif.

- a. Recall pada rangkaian percobaan data latih dan data uji 70:30

Tabel 6. Recall pada rangkaian pertama

Data Test pada Recall	Positif (label 1)	Negatif (label 0)
TP: 123 FN: 15	0.84	0.89

Pada tabel diatas menghasilkan TP sebanyak 123 data, FN sebanyak 15 data, dan recall pada rangkaian pertama ini menghasilkan nilai recall positif sebanyak 0.84 dan recall negatif sebanyak 0.89.

- b. Recall pada rangkaian percobaan data latih dan data uji 80:20

Tabel 7. Recall pada rangkaian kedua

Data Test pada Recall	Positif (label 1)	Negatif (label 0)
TP: 82 FN: 10	0.87	0.89

Pada tabel diatas menghasilkan TP sebanyak 82 data, FN sebanyak 10 data, dan recall pada rangkaian kedua ini menghasilkan nilai recall positif sebanyak 0.87 dan recall negatif sebanyak 0.89.

4.1.3 Precision

Tahap evaluasi Precision dilakukan untuk mendapatkan perbandingan dari jumlah data diprediksi positif benar oleh model dibandingkan dengan keseluruhan data yang positif.

- a. Precision pada rangkaian percobaan data latih dan data uji 70:30

Tabel 8. Precision pada rangkaian pertama

Data Test pada Precision	Positif (label 1)	Negatif (label 0)
TP: 123 FP: 22	0.88	0.85

Pada tabel diatas menghasilkan TP sebanyak 123 data, FN sebanyak 22 data, dan precision pada rangkaian pertama ini menghasilkan nilai recall positif sebanyak 0.88 dan recall negatif sebanyak 0.85.

- b. Precision pada rangkaian percobaan data latih dan data uji 80:20

Tabel 9. Precision pada rangkaian kedua

Data Test pada Precision	Positif (label 1)	Negatif (label 0)
TP: 82 FP: 12	0.89	0.87

Pada tabel diatas menghasilkan TP sebanyak 82 data, FN sebanyak 12 data, dan precision pada rangkaian kedua ini menghasilkan nilai recall positif sebanyak 0.89 dan recall negatif sebanyak 0.87.

4.1.4 F1-Score

Pada tahap evaluasi F1-Score, perhitungan evaluasi ini digunakan untuk membandingkan rata-rata precison dan recall yang telah dibobotkan.

- a. F1-Score pada rangkaian percobaan data latih dan data uji 70:30

Tabel 10. F1-Score pada rangkaian pertama

Data Test pada F1-Score	Positif (label 1)	Negatif (label 0)
TP: 123 FP: 22	0.86	0.87

Pada tabel diatas menghasilkan TP sebanyak 123 data, FN sebanyak 22 data, dan F1-Score pada rangkaian pertama ini menghasilkan nilai recall positif sebanyak 0.86 dan recall negatif sebanyak 0.87.

- b. F1-Score pada rangkaian percobaan data latih dan data uji 80:20

Tabel 11. F1-Score pada rangkaian kedua

Data Test pada F1-Score	Positif (label 1)	Negatif (label 0)
TP: 82 FP: 12	0.89	0.87

Pada tabel diatas menghasilkan TP sebanyak 82 data, FN sebanyak 12 data, dan F1-Score pada rangkaian kedua ini menghasilkan nilai recall positif sebanyak 0.89 dan recall negatif sebanyak 0.87.

Berdasarkan kedua tabel diatas, dapat disimpulkan bahwa perhitungan pada python menggunakan rangkaian F1-Score pada perbandingan data latih dan data uji 80:20 menghasilkan F1-Score yang lebih baik, yaitu MutlinomialNB f1_score: 0.881, atau bisa disimpulkan sebanyak 88%, hal ini lebih tinggi dari hasil F1-Score pada rangkaian F1-Score perbandingan data latih dan data uji 70:30 yang menghasilkan F1-Score sebesar 0.869 atau sebesar 86%. Sehingga dapat dikatakan bahwa perhitungan evaluasi rata-rata precision dan recall setelah dibobotkan sudah baik.

4.1.5 Accuracy

Pada tahap ini akan dilakukan perbandingan pada akurasi kedua rangkaian untuk menilai akurasi dari hasil algoritma yang dilakukan.

- a. Skor akurasi pada rangkaian percobaan data latih dan data uji 70:30

Tabel 13. Hasil Accuracy pada rangkaian pertama

Data Test pada Accuracy	Hasil Accuracy
TP: 123 TN: 115 FN: 15 FP: 22	0.87

Pada tabel diatas menghasilkan TP sebanyak 123 data, TN sebanyak 115 data, FN sebanyak 15 data, dan FP sebanyak 22 data. Hasil akurasi yang dilakukan pada rangkaian pertama ini menghasilkan 0.87.

- b. Skor akurasi pada rangkaian percobaan data latih dan data uji 80:20

Tabel 14. Hasil Accuracy pada rangkaian kedua

Data Test pada Accuracy	Hasil Accuracy
TP: 82 TN: 79 FN: 10 FP: 12	0.88

5. KESIMPULAN

Hasil penelitian yang dapat disimpulkan adalah sebagian besar pengguna game ini sudah merasa puas dengan beberapa saran yang perlu ditingkatkan, yaitu penambahan Bahasa Indonesia, penambahan kuota energi, dan peningkatan kecepatan *loading* pada game. Kemudian, hasil penerapan klasifikasi NBC dengan pembagian data latih sebesar 80% dan data uji sebesar 20% menghasilkan performa yang baik. Hal ini ditunjukkan dengan uji evaluasi menggunakan *confussion matrix* yaitu dari total data keseluruhan setelah dilakukan pra-proses dan pembersihan data, menghasilkan data total sebanyak 914 ulasan, dimana pada rangkaian uji digunakan sebanyak 183 data uji yang menghasilkan nilai *recall* sebesar 89%, *precision* 87%, *F1-Score* 88%, dan akurasi sebesar 88%. Saran untuk penelitian berikutnya proses filtering dilakukan dengan melakukan penambahan filtering kata secara manual, sehingga kata yang seharusnya tidak penting seperti kata ‘banget’ dan lainnya dapat terhapus dan menghasilkan word cloud yang lebih baik.

Daftar Pustaka

- [1] S. L. Ting, W. H. Ip, and A. H. C. Tsang, “Is Naïve bayes a good classifier for document classification?,” *Int. J. Softw. Eng. its Appl.*, vol. 5, no. 3, pp. 37–46, 2011.
- [2] M. S. Mustafa, M. R. Ramadhan, and A. P. Thenata, “Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier,” *Creat. Inf. Technol. J.*, vol. 4, no. 2, p. 151, 2018, doi: 10.24076/citec.2017v4i2.106.
- [3] F. Ratnawati, “Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter,” *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018, doi: 10.35314/isi.v3i1.335.
- [4] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, “Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1750–1757, 2017.
- [5] J. LING, I. P. E. N. KENCANA, and T. B. OKA, “Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier Dengan Seleksi Fitur Chi Square,” *E-Jurnal Mat.*, vol. 3, no. 3, p. 92, 2014, doi: 10.24843/mtk.2014.v03.i03.p070.
- [6] J. Arjoranta, “How to Define Games and Why We Need to,” *Comput. Games J.*, vol. 8, no. 3–4, pp. 109–120, 2019, doi: 10.1007/s40869-019-00080-6.
- [7] N. Esposito, “A short and simple definition of what a videogame is,” *Proc. DiGRA 2005 Conf. Chang. Views - Worlds Play*, no. January 2005, 2005.
- [8] S. Dang and P. H. Ahmad, “Text Mining : Techniques and its Application,” *Int. J. Eng. Technol. Innov.*, vol. 1, no. 4, pp. 22–25, 2014.
- [9] P. Dellia and A. Tjahyanto, “Tax Complaints Classification on Twitter Using Text Mining,” *IPTEK J. Sci.*, vol. 2, no. 1, p. 11, May 2017, doi: 10.12962/j23378530.v2i1.a2254.
- [10] A. Bradley and R. J. E. James, “Web Scraping Using R,” *Adv. Methods Pract. Psychol. Sci.*, vol. 2, no. 3, pp. 264–270, 2019, doi: 10.1177/2515245919859535.
- [11] N. A. Vidya, “Opinion Mining Dengan Menggunakan Multinomial Naive Bayes Classifier Pada Blog,” *Tek. Inform.*, pp. 1–55, 2012.