

Analisis Sentimen Pada *Steam Review* Menggunakan Metode *Multinomial Naïve Bayes* dengan Seleksi Fitur *Gini Index Text*

1st Ragil Haditira
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia

ragilhaditira@students.telkomuniversity.ac.id

2nd Danang Triantoro Murdiansyah
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia

danangtri@telkomuniversity.ac.id

3rd Widi Astuti
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia

widiwdu@telkomuniversity.ac.id

Abstrak

Video game adalah salah satu media hiburan yang sering digunakan oleh kebanyakan orang saat ini, terutama pada perangkat komputer baik yang berbayar atau tidak, pada salah satu distribusi digital yaitu *Steam*. Akan tetapi *Steam* masih memiliki permasalahan yang berbagai macam jenisnya salah satunya *Steam Review*, untuk mengurangi atau mengantisipasi *review* yang kurang jelas bagi para pemain yang ingin mencoba dan atau membeli game tersebut, maka digunakan Analisis Sentimen pada ulasan. Tujuan tugas akhir ini untuk memberikan informasi sentimen di setiap game yang ada pada *Steam*. Informasi yang diberikan berupa klasifikasi analisis sentimen untuk memfilter *review* dengan *Multinomial Naïve Bayes* dan dikombinasikan dengan seleksi fitur *Gini Index Text* supaya mengklasifikasi dokumen menjadi kelas direkomendasikan dan tidak direkomendasikan. Pada penelitian ini untuk menguji sistem yang dibangun digunakan Dataset yang berisi ulasan-ulasan berupa kalimat, data akan dibagi menjadi tiga rasio akurasi sebagai training dan testing sebanyak 90:10; 80:20; dan 70:30 serta dilakukan 5 kali percobaan. Hasil pengujian *Gini Index Text* sebagai fitur seleksi dapat menghasilkan akurasi setiap rasionya dimana Rata-rata akurasi MNNB dengan GIT yang paling maksimal sebesar 60.29%.

Kata kunci : Analisis Sentimen, *Steam Review*, *Multinomial Naïve Bayes* (MNNB), *Gini Index Text* (GIT), Klasifikasi

Abstract

Video games are one of the entertainment media that are often used by most people today, especially on computer devices, whether paid or not, at digital distributions, namely *Steam*. However, *Steam* still has problems of various kinds, which is *Steam Reviews*, to reduce or anticipate reviews that are less clear for players who want to try or buy the game, *Sentiment Analysis* is used in reviews. The purpose of this final project is to provide sentiment information on every game on *Steam*. The information provided is in the form of sentiment analysis classification to filter reviews with *Multinomial Naïve Bayes* and combined with the *Gini Index Text* feature selection in order to classify documents into recommended and not recommended classes. In this study, to test the system built, a dataset containing comments in the form of sentences is used, the data will be divided into three accuracy ratios for training and testing of 90:10; 80:20; and 70:30 and carried out 5 times the experiment. The test results *Gini Index Text* as a selection feature can produce the accuracy of each ratio where the average of accuracy at MNNB with the maximum GIT is 60.29%.

Keywords: *Sentiment Analysis*, *Steam Review*, *Multinomial Naïve Bayes* (MNNB), *Gini Index Text* (GIT), Classification

I. PENDAHULUAN

Pada bagian pendahuluan berisi empat sub-bagian yang terdiri dari **Latar Belakang**, **Topik dan Batasannya**, **Tujuan**, **Organisasi Tulisan**, berikut penjelasan dari masing-masing sub bagian tersebut.]

A. Latar Belakang

Seiring perkembangan teknologi saat ini banyak platform penjualan dan pembelian yang menyediakan konten salah satunya adalah *video games* seperti *Steam*, *Origin*, *Uplay*, *Epic Games Store*, dan sebagainya pada perangkat komputer. Untuk mengenali apakah suatu game itu bagus atau buruk, lantas penting mengetahui ulasan dari para pemain game tersebut, menarik perhatian untuk dimainkan. Sebagian *reviewer* game menebarkan ulasan mereka pada kolom komentar terkait game yang diunduh dan sudah dicoba apakah layak dibeli atau tidak. Sebagian ulasan mungkin terlihat jelas tergolong dalam ulasan yang direkomen atau tidak direkomen namun masih ada sajah ulasan yang tidak jelas kategorinya

Salah satu teknik untuk mengklasifikasikan opini yaitu dengan analisis sentimen atau dikenal dengan *opinion mining*, dengan teknik ini dapat mengenali apakah ulasan itu termasuk positif atau negatif, ada 2 pendekatan cara mengklasifikasikannya yaitu yang pertama dengan pendekatan *machine learning* dan yang kedua pendekatan *lexicon*. Kedua pendekatan tersebut mengklasifikasikan teks kedalam kategori positif dan negatif tergantung dari polaritas kalimatnya, pendekatan *lexicon* umumnya bergantung dengan kamus dari kata-kata opini atau dikenal dengan kamus sentimen, untuk mendeterministikan orientasi sentimen sebagai positif atau negatif. Sedangkan pendekatan *machine learning* menggunakan klasifikasi data manual dari *dataset* dan melatih *classifier* dari contoh atau disebut *data training* yang nanti akan diuji pada *data testing* [1]. Berdasarkan hasil penelitian yang telah dilakukan [2] dapat dilihat hasil dari bermacam-macam metode untuk analisis sentimen dan terlihat bahwa dengan pendekatan *machine learning* menghasilkan akurasi yang tertinggi.

Untuk penelitian ini penulis mencoba mengabungkan dua metode yaitu *machine learning* dengan menitik beratkan pada penggunaan metode *naïve bayes multinomial* dengan seleksi fitur *gini index*, dengan alasan utama yang berdasar pada penelitian [3] penulis ingin mencoba apakah metode ini

dapat diimplementasikan dengan sempurna pada Dataset *Steam Review* supaya Akurasinya bisa mencapai 98.66%. Berdasarkan penelitian [4] dapat dilihat bermacam-macam metode untuk meningkatkan akurasi dari metode *naïve bayes* dan untuk metode yang terbaik adalah menggunakan seleksi fitur, dalam penelitian [5] menjelaskan bahwa metode *naïve bayes* sangat sensitif terhadap seleksi fitur sehingga jumlah fitur yang digunakan akan sangat berpengaruh pada metode ini. Pada [6] metode *Gini Index* dilakukan perubahan menjadi *Gini Index Text* dengan *K Nearest Neighbor* (KNN) dan *Support Vector Machines* (SVM) sebagai *classifier* hasil dari penelitian ini mendapatkan performansi yang sangat baik mencapai rata-rata akurasi 98.5% untuk KNN dan 93% pada SVM. Dengan menggunakan penggabungan seleksi fitur *Gini Index* dan teknik klasifikasi *Naïve Bayes* diharapkan dapat mengurangi waktu untuk klasifikasi manual data training dan juga menambah akurasi dibandingkan hanya menggunakan *Naïve Bayes* saja.

B. Topik dan Batasannya

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, maka perumusan masalah pada penelitian ini adalah bagaimana cara melakukan klasifikasi analisis sentimen untuk ulasan *steam review* menggunakan metode *Multinomial Naïve Bayes*. Lalu bagaimana performa yang dihasilkan dari system yang dibuat.

Dalam penelitian ini terdapat Batasan Masalah yaitu diantaranya, Data yang digunakan adalah dataset pada ulasan terlaris *steam* pada periode februari tahun 2019, dengan judul *Collection of Steam's best-selling game reviews*. Data telah diperbaharui pada <https://www.kaggle.com/luthfim/steam-reviews-dataset> berisikan 2000 data positif dan 2000 data negative yang akan diuji, dengan porsi rasio data akurasi sebanyak 90:10%, 80:20%, 70:30% untuk 5 kali percobaan, serta dalam penelitian ini menggunakan *benchmark* untuk melihat akurasi dari kelipatan 100 fitur sehingga akan menonjolkan performa dari fitur yang diseleksi.

C. Tujuan

Berdasarkan Topik dan Batasan yang telah dipaparkan sebelumnya tujuan utama dari penelitian ini yaitu. Memproses data serta melakukan seleksi fitur dengan metode *Gini Index* dan menggunakan metode *Multinomial Naïve Bayes* untuk melakukan sentiment analisis pada *steam review*, dan mengukur performa dari sistem yang dibuat dari sisi akurasi, recall, presisi, skore-F1, fitur dan jumlah total fitur.

D. Organisasi Tulisan

Berdasarkan penelitian yang sudah dilakukan oleh penulis, maka jurnal tugas akhir ini berisikan beberapa bagian yaitu pada bagian 1 menjelaskan Pendahuluan (Latar Belakang, Topik dan Batasannya, Tujuan serta Organisasi Tulisan). Selanjutnya pada bagian 2 menjelaskan studi dan teori yang terkait dalam penelitian yang digunakan dalam membangun system klasifikasi. Setelah itu pada bagian 4 menjelaskan evaluasi, pengujian dan analisis pengujian. Terakhir pada bagian 5, dilakukan penulisan kesimpulan dan saran untuk penelitian selanjutnya.

II. KAJIAN TEORI

Terdapat banyak studi mengenai analisis sentimen tentang topik-topik seperti komentar media sosial, produk, politik, dan banyak lagi. Terdapat banyak sekali teknik untuk mengklasifikasikan text, banyak peneliti yang mencoba untuk melakukan kombinasi teknik agar mencapai performansi yang lebih baik. Seperti pada penelitian [7] dimana peneliti membandingkan 6 teknik klasifikasi dengan menggunakan ekstrasi fitur *word*, *bigram*, dan *word+adjective* dengan seleksi fitur *Information Gain*. Dengan menggunakan data *movie review* IMDB mendapatkan hasil terbaik dengan teknik *Random Forest* yaitu 88.95 %, sedangkan *Naïve Bayes* mendapatkan akurasi terendah yaitu 54.77%, peneliti mengusulkan untuk menerapkan konsep *natural language processing* (NLP) lebih dalam untuk prediksi polaritas dokumen yang lebih baik.

Pada [8] membandingkan beragam seleksi fitur untuk algoritma *Naïve Bayes* tujuan penelitiannya adalah untuk menentukan seleksi fitur yang akurat serta stabil dengan menggunakan data *Newsgroups* dan *Reuters-21578*. Penelitian ini *pre-processing* data tidak ditampilkan sehingga proses data yang dilakukan tidak dijelaskan sampai ke seleksi fitur. Hasil akurasi tertinggi didapatkan dengan seleksi fitur *Chi* Sebesar 81% untuk data *Reuters-21578* dan 84% untuk *Newsgroups* dan untuk hasil yang stabil didapatkan dengan menggunakan teknik *Term Frequency* (TF) dan *Domain Frequency* (DF), namun tidak efisien dibandingkan dengan teknik lain. Dalam Penelitian [9] membandingkan seleksi fitur *Gini Index* (NG), *Weight formula* (NG), dan *Word Frequency Mutual Information* (MIDF) dengan teknik klasifikasi KNN dan *Naïve Bayes*, tujuan dari penelitian ini adalah untuk membandingkan dan memvalidasi dengan fitur bobot baru berdasarkan *Gini Index* pada performa dari teknik klasifikasi. Hasil dari penelitian ini *Gini Index* mendapatkan performansi yang sangat baik yaitu 75% dengan KNN dan 95% dengan *Naïve Bayes*. Pada penelitian [10] membandingkan teknik klasifikasi *Multinomial Naïve Bayes* (MNNB) dan SVM dengan beragam seleksi fitur sebanyak 18 seleksi fitur, untuk *pre-processing* data tidak dijelaskan dengan detail proses apa saja yang dilakukan. Hasil penelitian menunjukkan performa yang sama dengan menggunakan *Gini Index*, *Weighted Log Likelihood Ratio* (WLLR) dan *Cross Entropy for Text* (CET) pada MNNB. Berdasarkan hasil penelitian menghasilkan kesimpulan bahwa SVM memerlukan fitur yang lebih sedikit untuk mencapai hasil maximal sedangkan *Naïve Bayes* memerlukan lebih banyak fitur untuk mencapai hasil optimal dan lebih sensitif pada pemilihan metode seleksi fitur

Pada penelitian [11] peneliti membandingkan *Naïve Bayes*, *Maximum Entropy*, dan *Gradient Descent* dengan seleksi fitur *Gini Index*. Teori tidak dijelaskan dengan detail dan rumus yang ditampilkan hanya *Gini Index* saja dan data yang digunakan tidak dijelaskan. penelitian ini mendapatkan hasil akurasi tertinggi *Gradient Descent* 70.10% dan terendah *Naïve Bayes* 65.24%. Dalam [12] Peneliti membandingkan *Gini Index Text* dengan beragam seleksi fitur pada data *IMDB dataset*, *Spanish dataset*, dan *Portuguese Dataset* data diekstrasi terlebih dahulu dengan *n-gram* dan dengan menggunakan teknik klasifikasi MNNB, SVM, dan *Weighted SVM* (WSVM). Hasil yang didapatkan dijelaskan dengan detail namun waktu eksekusi tidak dijelaskan. Dari

kesimpulan didapatkan bahwa seleksi fitur Gini Index, Domain Frequency, CET, dan CHI mendapatkan performansi terbaik yaitu 90% dengan SVM dan 92% dengan MNNB. Untuk teknik klasifikasi SVM merupakan algoritma yang cepat dan dapat bekerja pada dimensi yang tinggi sedangkan MNNB bekerja lebih baik pada dokumen singkat atau untuk mengklasifikasikan kalimat, WSVM berkerja lebih baik dibandingkan SVM jika tidak menggunakan seleksi fitur.

A. Seleksi Fitur

Seleksi fitur adalah proses dimana akan memilih fitur yang relevan yang berguna untuk menyederhanakan model agar lebih mudah untuk ditafsirkan dan dapat mengurangi waktu training pada data. Teknik pada seleksi fitur diantaranya:

1. Gini Index Text

Gini Index umumnya digunakan untuk memisahkan atribut teknik ini pertama kali digunakan pada *decision tree* dan berhasil meningkatkan presisi dari klasifikasi ini. Telah banyak penelitian untuk meningkatkan metode gini index ini salah satunya adalah *Gini Index Text* (GIT) yang diperkenalkan pada [13], dibuat untuk bekerja pada dokumen dengan fitur yang sangat banyak dengan metode ini dapat mereduksi fitur dari subset fitur sementara dapat juga mempertahankan banyak fitur representatif. Gini index didefinisikan secara umum sebagai:

$$GiniText(w) = P(w|c_i)^2 P(c_i|w)^2 \quad (1)$$

Dengan:

$P(w|c_i)$ adalah kejadian dimana munculnya kata pada kategori c_i .

$P(c_i|w)$ adalah distribusi kategori c_i pada munculnya kata.

2. Complete Gini Index Text

Seleksi fitur GIT memiliki tiga persamaan yaitu $GIT_A(W, C_k)$, $GIT_B(W, C_k)$, dan $GIT_C(W, C_k)$. Didefinisikan W merupakan himpunan fitur $W = \{w_1, \dots, w_j\}$ dan C adalah kelas *negative* atau bukan *negative*. Persamaan GIT tersebut didefinisikan sebagai berikut:

$$GIT_A(W, C_k) = P(C_k|W)^2 \quad (2)$$

$$GIT_B(W, C_k) = \left| \frac{P(C_k|W)^2}{\log_2 P(W)} \right| \quad (3)$$

$$GIT_C(W, C_k) = \frac{P(C_k|W)^2}{|\log_2 P(C_k|W)|^2} \quad (4)$$

Keterangan:

C_k merupakan himpunan kelas *negative* dan *positive*

W merupakan himpunan seluruh fitur

$P(C_k|W)$ merupakan peluang bersyarat kelas k terhadap seluruh fitur

$P(W)$ merupakan peluang dari seluruh fitur W

B. Naïve Bayes

Pada analisis sentimen dan klasifikasi text, algoritma *naïve bayes* merupakan algoritma yang paling sering digunakan. Terdapat beberapa jenis *Naïve Bayes* diantaranya *Multinomial Naïve Bayes*, *Naïve Bayes Multivariate Bernoulli*, *Naïve Bayes Gaussian*. Berdasarkan varian *naïve bayes* tersebut diproses menggunakan bentuk data yang berbeda yang diperoleh pada tahap *pre-processing* [14] klasifikasi *naïve bayes* menduga suatu fitur tidak bergantung pada fitur yang lainnya, yang dinamakan dengan asumsi *naïve*. Naïve Bayes didefinisikan secara umum sebagai berikut:

$$P = \frac{P(X|C)P(C)}{P(X)} \quad (5)$$

Dimana X adalah himpunan dari fitur $X = \{x_1, \dots, x_j\}$ dan C menyatakan himpunan dari kelas atau kategori $C = \{c_1, \dots, c_k\}$.

1. Multinomial Naïve Bayes

Naïve Bayes Classifier merupakan sebuah pengklasifikasi probabilitas sederhana yang mengaplikasikan *Teorema Bayes* dengan asumsi ketidaktergantungan yang tinggi. *Teorema Bayes* adalah teorema yang dipakai dalam statistika untuk menghitung peluang untuk suatu hipotesis [15]. Umumnya terdapat tiga model distribusi yaitu *Bernoulli*, *Multinomial*, dan *Poisson* ketiga model ini kerap digunakan sebagai pengklasifikasi sebagai *Bernoulli Naïve Bayes*, *Multinomial Naïve Bayes* dan *Poisson Naïve Bayes*.

Multinomial Naïve Bayes memperhitungkan jumlah kata dalam dokumen sehingga mengasumsikan independensi kemunculan kata dalam dokumen. Dengan asumsi ini menunjukkan bahwa kemungkinan tiap kejadian kata dalam dokumen adalah bebas tidak memperhitungkan urutan kata dan konteks kata dalam dokumen [15].

Dengan persamaan:

$$P(w, c) = \frac{1 + \sum_{j=1}^N N_{ij}}{\delta_{jk}} \quad (6)$$

$$|V| + \sum_{i=1}^{|V|} \sum_{n=1}^N N_{in} \delta_{jk}$$

Dengan:

N adalah jumlah dari dokumen.

δ_{jk} bernilai 1 jika dokumen ke j milik kelas c_k dan bernilai 0 jika sebaliknya.

N_{ij} adalah jumlah fitur w_i yang terjadi pada dokumen ke j + 1 untuk menghindari nilai nol atau adanya satu kemungkinan.

$|V|$ adalah jumlah kosakata dari sebuah dokumen ulasan game

C. Pengukuran Kerja

Setelah mendapatkan hasil dari klasifikasi akan diukur seberapa baik mesin berkerja. Untuk mengetahui hal tersebut, maka digunakan *confusion matrix* sebagai berikut:

True Positives (TP) : jumlah ulasan game positif yang diklasifikasikan positif

True Negatif (TN) : jumlah ulasan game negatif yang diklasifikasikan negatif
 False Positif (FP) : jumlah ulasan game negatif yang diklasifikasikan positif
 False Negative (FN) : jumlah ulasan game positif yang diklasifikasikan negatif
 Akurasi adalah hasil dari mesin klasifikasi (pengukuran umum)

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \quad (7)$$

Presisi adalah ukuran keberhasilan suatu ulasan sebagai kelas direkomendasikan

$$Presisi = \frac{TP}{TP + FP} \quad (8)$$

Recall adalah tingkat keberhasilan suatu ulasan terindikasi positif (direkomendasikan) oleh mesin klasifikasi

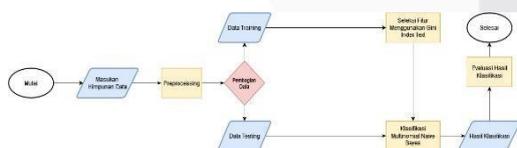
$$Recall = \frac{TP}{TP + FN} \quad (9)$$

F1 Score adalah untuk mengukur rata-rata recall dan presisi hasil klasifikasi

$$F1\ Score = \frac{2\ Presisi\ Recall}{Presisi + Recall} \quad (10)$$

III. METODE

Rancangan sistem yang digunakan untuk penelitian ini adalah data yang sudah didapatkan akan di proses untuk mengurangi jumlah fitur dengan *pre-processing* yaitu *tokenization*, *Casefolding*, *stopword removal*, *Stemming dan lemmatization*. Data akan dibagi untuk data *training* dan *data testing*, untuk data *training* fitur akan diseleksi dengan metode *complete gini index text* (GIT). Setelah fitur diseleksi, *Multinomial Naïve Bayes* (MNNB) akan dilatih pada fitur-fitur hasil seleksi. *Data testing* akan diklasifikasi dengan MNNB yang sudah dilatih lalu perfromansi akan dihitung. Alur rancangan sistem dapat dilihat sebagai berikut :



Gambar 1. Flowchart Klasifikasi Steam Review Menggunakan Multinomial Naïve Bayes dengan Seleksi

A. Dataset

Adapun data yang digunakan pada penelitian ini berjumlah 303594 data positif dan 131297 data negatif menggunakan sample data yang digunakan pada penelitian terkait, untuk jumlah data *training* yang digunakan yaitu 4000 data termasuk 2000 data positif dan 2000 data negatif dan untuk data *testing* diukur berdasarkan rasio akurasi 90:10% = 200 testing, 80:20% = 400 testing 70:30% = 600 testing.

B. Preprocessing

Data akan diproses untuk menghilangkan kata-kata yang tidak mengandung informasi penting untuk menghilangkan

sejumlah fitur agar lebih mudah untuk diseleksi. Seperti dokumen berikut:

I will be straight to the point. Buy the game for the Single Player the Multiplayer it's full of cheaters and no matter how many times they get banned they still come they are multiplying like damn rabbits

1. Tokenization

Data akan diproses dimana kalimat akan dirubah menjadi kata-kata:

I', 'will', 'be', 'straight', 'to', 'the', 'point', 'Buy', 'the', 'game', 'for', 'the', 'Single', 'Player', 'the', 'Multiplayer', 'it's', 'full', 'of', 'cheaters', 'and', 'no', 'matter', 'how', 'many', 'times', 'they', 'get', 'banned', 'they', 'still', 'come', 'they', 'are', 'multiplying', 'like', 'damn', 'rabbits'

2. Casefolding

Data akan diproses dimana huruf besar akan dirubah menjadi huruf kecil:

'i', 'will', 'be', 'straight', 'to', 'the', 'point', 'buy', 'the', 'game', 'for', 'the', 'single', 'player', 'the', 'multiplayer', 'it's', 'full', 'of', 'cheaters', 'and', 'no', 'matter', 'how', 'many', 'times', 'they', 'get', 'banned', 'they', 'still', 'come', 'they', 'are', 'multiplying', 'like', 'damn', 'rabbits'

3. Stopwords Removal

Data akan diproses dimana kata-kata yang tidak memiliki makna akan dihilangkan:

'straight', 'point', 'buy', 'game', 'single', 'player', 'multiplayer', 'full', 'cheaters', 'matter', 'times', 'get', 'banned', 'still', 'multiplying', 'damn', 'rabbits'

4. Stemming dan Lemmatization

Data akan diproses dimana kata akan dirubah ke bentuk dasar sesuai dengan kamus yang ada menjadi:

'straight', 'point', 'buy', 'game', 'single', 'player', 'multiplayer', 'full', 'cheater', 'matter', 'time', 'get', 'banned', 'still', 'multiply', 'damn', 'rabbit'

C. Menghitung Skor Gini Index Text

Fitur – fitur yang ada pada data training akan dihitung menggunakan rumus Gini Index Text (GIT) yaitu jumlah total fitur sebanyak “k” memiliki nilai skor yang nantinya akan digunakan untuk memilih fitur – fitur dengan nilai skor tertinggi.

1. Data Training dengan Seleksi fitur Gini Index

Tabel 1. Contoh Sampel Sebanyak 6 Dokumen untuk Kategori “like” dan “Kelas”

Dokumen	like	Kelas
1	1	Direkomendasikan
2	0	Direkomendasikan
3	1	Direkomendasikan
1001	0	Tidak direkomendasikan
1002	0	Tidak direkomendasikan
1003	1	Tidak direkomendasikan

Berupa contoh sampel dari pengambilan dokumen secara acak pada dataset awal, dengan kategori “like” akan bernilai 1 yang mana kategori “Kelas”-nya akan Direkomendasikan

serta untuk kategori “like” akan bernilai 0 dengan kategori “Kelas”-nya akan Tidak direkomendasikan.

$$P(\text{like}|\text{Direkomendasikan})^2 * P(\text{Direkomendasikan}|\text{like})^2$$

$$\left(\frac{2}{6}\right)^2 \times \left(\frac{3}{6}\right)^2 = 0.0275$$

$$P(\text{like}|\text{Tidak direkomendasikan})^2 \cdot P(\text{Tidak direkomendasikan}|\text{like})^2$$

$$\left(\frac{1}{6}\right)^2 \times \left(\frac{3}{6}\right)^2 = 0.006945$$

Untuk fitur *like* akan dikategorikan pada fitur Direkomendasikan dikarenakan mempunyai peluang lebih besar. Kemudian memilih fitur-fitur teratas sebanyak k dengan menggunakan *select by weight* hasil seleksi fitur akan membentuk *bag of words*.

Multinomial Training Naïve Bayes

Hasil *bag of words* akan *fitting* (dicocokkan) pada *classifier* untuk dikenali pada *Multinomial Naïve Bayes*.

E. Klasifikasi dengan Multinomial Naïve Bayes

Multinomial Naïve Bayes yang sudah dilatih akan mengklasifikasikan data testing kedalam kategori dire dan negatif.

F. F1-Measure

Untuk Mengukur performansi hasil klasifikasi dengan menggunakan pengukuran F1, teknik ini mempertimbangkan precision dan recall.

$$\frac{2 \times \text{Precision} \times \text{recall}}{\text{Precision} + \text{recall}} \quad (8)$$

IV. HASIL DAN PEMBAHASAN

Untuk rancangan pengujiannya, yang (1) adalah pengujian dengan menggunakan Metode Multinomial Naïve Bayes saja tanpa menggunakan seleksi fitur yang (2) ialah pengujian dengan menggunakan Multinomial Naïve Bayes dengan seleksi fitur Gini Index Text, Masing-masing dengan 5 kali percobaan setiap porsi rasio data akurasi berdasarkan validasi menggunakan fungsi random terhadap data *training* dan *testing*.

Pada penelitian ini penulis akan menggunakan 3 rasio porsi data akurasi dari dataset yaitu 90:10% ; 80:20% ; 70:30%, sebanyak 4000 data untuk tiap rasio training dan testing dengan menggunakan benchmark dipilih untuk melihat akurasi dari kelipatan 100 fitur agar dapat menonjolkan performa dari fitur yang diseleksi. Untuk lebih jelasnya 90:10% akan menjadi (1800 data *train* positif, 1800 data *train* negatif, 200 data *test* positif, 200 data *test* negatif), dan 80:20% akan menjadi (1600 data *train* positif, 1600 data *train* negatif, 400 data *test* positif, 400 data *test* negatif), serta 70:30% menjadi (1400 data *train* positif, 1400 data *train* negatif, 600 data *test* positif, 600 data *test* negatif).

A. Pengujian Tanpa Gini Index Text

Pada pengujian ini data akan diklasifikasi dengan menggunakan metode Multinomial Naïve Bayes saja tanpa menggunakan seleksi fitur untuk mengetahui nilai presentase pada Dataset. Kumpulan data yang dihasilkan dari pre-processing langsung diklasifikasi data akan dijalankan sebanyak 5 kali percobaan pada 4000 data dari kelipatan 100 fitur pada proses *benchmark*. Hasil dari pengujian ini adalah nilai Rata-rata akurasi (%) dari 5 kali percobaan untuk setiap kategori yaitu nilai *MNNB*, *precision*, *recall*, dan *f1-score*, serta jumlah total fitur. Untuk hasil pengujiannya ada pada tabel berikut.

Tabel 2. Hasil Pengujian Tanpa Gini Index Text untuk Rata-rata dari 5 kali percobaan

Akurasi MNB	Akurasi Precision	Akurasi Recall	Akurasi F1-score	Jumlah Total Fitur	Rasio Training Testing	Nilai Rata-rata Akurasi (%) dari 5 kali percobaan
81.12 %	81.95%	80.08 %	81.00 %	4694	90:10	
80.99 %	83.02%	78.07 %	80.46 %	4575	80:20	
81.38 %	83.30%	78.94 %	81.04 %	4392	70:30	

Berdasarkan tabel diatas Rata-rata hasil akurasi (%) terabaik dengan menggunakan 5 kali percobaan setiap porsi rasio akurasi (training dan testing) data dengan rata-rata akurasi sebesar 81.38%, rata – rata presisi sebesar 83.30%, rata-rata recall sebesar 80.08%, serta rata-rata skor-f1 sebesar 81.04%.

B. Pengujian dengan Gini Index Text

Pada pengujian kali ini data akan diklasifikasi dengan menggunakan metode Multinomial Naïve Bayes dengan seleksi fitur untuk mengetahui akurasi (%) terbaik yang akan dilakukan secara berkala sebanyak 5 kali percobaan setiap rasio porsi akurasi, Pengujian pada code program yang nantinya akan dilakukan pengecekan dan pengurangan fitur berdasarkan skor akurasi nilai *Gini Index Text*. Sekenario pengujian pada tahapan ini yaitu dengan menentukan nilai akurasi MNNB dengan GIT.

Tabel 3. Hasil Pengujian dengan Gini Index Text untuk Rata-rata dari 5 kali percobaan

Akurasi MNB	Akurasi Precision	Akurasi Recall	Akurasi F1-score	Total Fitur	Rasio Training Testing	Nilai Rata-rata Akurasi (%) dari 5 kali percobaan
60.29 %	56.65%	90.44 %	69.64 %	4480	90:10	
59.69 %	56.33%	91.46 %	69.71 %	4380	80:20	
58.33 %	55.57%	89.15 %	68.45 %	4200	70:30	

Berdasarkan tabel diatas diketahui bahwa hasil untuk menentukan akurasi (%) terbaik dengan Gini Index Text dari 5 kali percobaan setiap porsi rasio akurasi (*training* dan *testing*) data dengan rata-rata akurasi sebesar 60.29%, rata-rata presisi sebesar 56.65%, rata-rata recall sebesar 91.46%, serta rata-rata skor-f1 sebesar 69.71%

Pada penelitian ini dapat dilihat bahwa performa sistem dengan menggunakan seleksi fitur hasilnya lebih rendah daripada tanpa menggunakan seleksi fitur, misalnya dari sisi akurasi, dengan menggunakan seleksi fitur akurasi yang dihasilkan lebih rendah daripada dengan tanpa menggunakan seleksi fitur. Hal tersebut kemungkinan disebabkan karena seleksi fitur Gini Index Text kurang efektif dalam mengidentifikasi fitur / kata penting dalam kasus yang ada di penelitian ini.

V. KESIMPULAN

Pada Penelitian ini telah dibangun sistem klasifikasi untuk mengenal suatu data *record* menjadi bentuk data fitur berdasarkan ulasan atau tweet pada *games* yang ada di Steam, penggunaan metode *Multinomial Naive Bayes* dengan Seleksi Fitur Gini Index Text dapat menghasilkan akurasi, presisi, recall, dan skor-F1. Hasil klasifikasi pada 5 kali percobaan training dan testing, untuk porsi Rasio data 90:10, 80:20, 70:30, pada 4000 data dengan mengambil 100 fitur terbaik (skor tertinggi) dan dilakukan penambahan fitur setiap iterasinya sebagai akurasi terbaik, nilai Rata-rata akurasi MNNB Tanpa GIT didapat 81.38%, 83.30%, 80.08%, 81.04%. Sedangkan untuk MNNB dengan GIT didapat, 60.29%, 56.65%, 91.46%, 69.71%. Fitur seleksi terlihat dapat berpengaruh pada klasifikasi data menggunakan 5 kali percobaan terhadap 4000 data dengan validasi fungsi random, sehingga menunjukkan data yang terdistribusi normal.

Adapun beberapa saran yang bisa dilakukan untuk penelitian selanjutnya yaitu penambahan metode *sampling* yang berbeda, dan memperbanyak kelipatan data menjadi puluhan ribu agar dapat meningkatkan akurasi. Selain itu juga, disarankan menggunakan metode klasifikasi dan proses seleksi fitur dengan kombinasi *feature extraction* lain supaya nilainya lebih optimal dari penelitian sebelumnya.

REFERENSI

- [1] C. Dhaoui, C. Webster dan L. P. Tan, "Social media sentiment analysis: lexicon versus machine learning," *Journal of Consumer Marketing*, pp. 1-23, 2017.
- [2] V. B. Vilmakumar and J. M. Bhumika, "Analysis of Various Sentiment Classification Techniques," *International Journal of Computer Applications*, vol. 140, no. 3, pp. 0975 - 8887, 2016.
- [3] F. R. a. D. T. M. Imadudin, ""Implementation of Naive Bayes and Gini Index for Spam Email Classification."," *Indonesian Journal on Computing (Indo-JC)*, no. 6.1, pp. 85-94, 2021.
- [4] V. Narayanan, I. Arora and A. Bhatia, "Fast and accurate sentiment classification using an enhanced Naive Bayes model," in *International Conference on Intelligent Data Engineering and Automated Learning*, Varanasi, India, 2013.
- [5] J. Chen, H. Huang, S. Tian and Y. Qu, "Feature selection for text classification with Naive Bayes," *Expert Systems with Applications*, vol. 36, pp. 5432 - 5435, 2008.
- [6] W. Shang, H. Huang, H. Zhu, Y. Lin, Y. Qu and Z. Wang, "A novel feature selection algorithm for text categorization," *Expert Systems with Applications*, vol. 33, pp. 1 - 5, 2007.
- [7] T. P. Sahu and S. Ahuja, "Sentiment Analysis of Movie Reviews: A study on Feature Selection & Classification Algorithms," India, 2016.
- [8] E. Lux, Feature selection for text classification with Naive Bayes, 2012.
- [9] X. JIA and J. SUN, "AN IMPROVED TEXT CLASSIFICATION METHOD BASED," *Journal of Theoretical and Applied Information Technology*, vol. 43, no. 2, pp. 267 - 273, 2012.
- [10] P. Varela, P. M. Q. Aquiar, A. F. Martins and M. A. T. Figueiredo, "An Empirical Study of Feature Selection for Sentiment Analysis," in *9th Conference on Telecommunications*, Castelo Branco, 2013.
- [11] F. Iqbal, "Sentiment Analysis Using Ensemble Learners and Gini Index," *International Journal of Engineering and Techniques*, vol. 4, no. 2, pp. 586 - 591, 2018.
- [12] P. L. Varela, "Sentiment Analysis," Instituto Superior T'ecnico, 2012.
- [13] H. Park, S. Kwon and H.-C. Kwon, "Complete Gini-Index Text (GIT) Feature-Selection Algorithm for Text Classification," in *Software Engineering and Data Mining (SEDM). 2010 2nd International Conference*, Busan, Korea, 2010.
- [14] Z. Zuo, "Sentiment analysis of steam review datasets using naive bayes and decision tree classifier," 2018.
- [15] T. M. E. Yuridis, M. A. Fauzi and Indriati, "Penentuan Rating Review Film Menggunakan Metode Multinomial Naive Bayes Classifier dengan Feature Selection berbasis Chi-Square dan Galavotti-Sebastiani-Simi Coefficient," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 1, pp. 8034 - 8040, 2019.

