https://doi.org/10.37817/ikraith-informatika.v10i2
P-ISSN : 2580-4316
E-ISSN : 2654-8054

Perbandingan Kinerja Multinomial Naïve Bayes pada Variasi Rasio Data Training dan Testing dalam Analisis Sentimen Ulasan Taman Rakyat Slawi Ayu

[1]Moh. Syaogi, [2]Nur Ariesanto Ramdhan, [3]Otong Saeful Bachri
[1,2,3]Teknik Informatika, Universitas Muhadi Setiabudi, Brebes

E-mail: [1]mohammadsyaogi301204@gmail.com, [2]ariesantoramdhan@gmail.com, [3]otongsaifulbahriumus@gmail.com

ABSTRAK

Analisis sentimen merupakan salah satu teknik *text mining* yang digunakan untuk mengidentifikasi opini masyarakat terhadap suatu objek atau layanan. Taman Rakyat Slawi Ayu sebagai salah satu ruang terbuka hijau di Kabupaten Tegal menerima berbagai ulasan pengunjung melalui *Google Maps* yang dapat dimanfaatkan sebagai sumber informasi untuk mengevaluasi kualitas layanan dan fasilitas. Penelitian ini bertujuan membandingkan kinerja algoritma *Multinomial Naïve Bayes* pada variasi rasio pembagian data *training* dan *testing* dalam analisis sentimen ulasan Taman Rakyat Slawi Ayu. Data diperoleh melalui teknik *web crawling* pada *Google Maps* dan menghasilkan 4.153 ulasan. Setelah proses seleksi data, diperoleh 1.794 ulasan yang memenuhi kriteria analisis. Tahap *preprocessing* meliputi *cleaning, case folding, tokenizing, stopword removal*, dan *stemming*, sehingga diperoleh 1.789 data yang siap diklasifikasikan. Pelabelan sentimen dilakukan berdasarkan nilai rating, yaitu rating 1–3 sebagai sentimen negatif dan rating 4–5 sebagai sentimen positif. Pengujian dilakukan menggunakan tiga variasi rasio pembagian data, yaitu 70:30, 80:20, dan 90:10. Hasil penelitian menunjukkan bahwa rasio 90:10 menghasilkan nilai *accuracy* tertinggi sebesar 87,71%, *precision* sebesar 0,88, *recall* sebesar 0,99, dan *f1-score* sebesar 0,93. Meskipun demikian, rasio tersebut menggunakan data uji yang lebih sedikit sehingga hasil evaluasi memiliki representativitas yang lebih terbatas dibandingkan rasio lainnya. Secara keseluruhan, penelitian ini menunjukkan bahwa variasi rasio pembagian data memengaruhi performa klasifikasi algoritma *Multinomial Naïve Bayes* dalam analisis sentimen ulasan Taman Rakyat Slawi Ayu.

Kata kunci : Analisis Sentimen, *Multinomial Naïve Bayes, Google Maps,* Taman Rakyat Slawi Ayu, Rasio Pembagian Data.

ABSTRACT

Sentiment analysis is one of the text mining techniques used to identify public opinions toward a particular object or service. Taman Rakyat Slawi Ayu, as one of the public green spaces in Tegal Regency, receives numerous visitor reviews through Google Maps, which can be utilized as a source of information to evaluate the quality of its services and facilities. This study aims to compare the performance of the Multinomial Naïve Bayes algorithm using different training and testing data split ratios for sentiment analysis of Taman Rakyat Slawi Ayu reviews. The data were collected through a web crawling technique on Google Maps, resulting in 4,153 reviews. After the data selection process, 1,794 reviews met the analysis criteria. The preprocessing stage consisted of cleaning, case folding, tokenizing, stopword removal, and stemming, resulting in 1,789 reviews ready for classification. Sentiment labeling was performed based on user ratings, where ratings of 1–3 were categorized as negative sentiment and ratings of 4–5 as positive sentiment. The

P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

experiments were conducted using three data split ratios: 70:30, 80:20, and 90:10. The results showed that the 90:10 ratio achieved the highest accuracy of 87.71%, with a precision of 0.88, recall of 0.99, and F1-score of 0.93. However, this ratio used a smaller testing dataset, making the evaluation results less representative than those obtained using the other data split ratios. Overall, the study demonstrates that variations in the training and testing data split ratio affect the classification performance of the Multinomial Naïve Bayes algorithm in sentiment analysis of Taman Rakyat Slawi Ayu reviews.

Keyword : ***Sentiment Analysis, Multinomial Naïve Bayes, Google Maps, Taman Rakyat Slawi Ayu, Data Split Ratio.***

1. PENDAHULUAN

Kemajuan teknologi informasi serta pesatnya penggunaan layanan berbasis lokasi telah mentransformasi pola masyarakat dalam mengutarakan pendapat dan pengalaman mereka terhadap berbagai aspek kehidupan, tidak terkecuali dalam konteks pelayanan publik dan sarana perkotaan (Permana et al., 2023). Taman Rakyat Slawi Ayu merupakan ruang terbuka hijau yang berfungsi sebagai ikon sekaligus tujuan rekreasi utama bagi warga Kabupaten Tegal, Jawa Tengah. Sebagai fasilitas publik yang berada di bawah pengelolaan pemerintah daerah, taman ini memiliki peran strategis dalam menyediakan kenyamanan serta kepuasan bagi para pengunjung. Dengan demikian, pemahaman yang mendalam mengenai opini masyarakat terhadap kualitas, fasilitas, dan pelayanan di Taman Rakyat Slawi Ayu menjadi aspek krusial dalam upaya peningkatan dan pengembangan taman tersebut.

Seiring dengan meningkatnya penggunaan *platform* pemetaan digital seperti *Google Maps*, masyarakat cenderung mengekspresikan pandangan mereka secara terbuka melalui *rating* maupun komentar yang disertakan pada fitur ulasan pengguna. Data yang dihasilkan dari aktivitas ini termasuk dalam kategori *big data* tidak terstruktur sehingga diperlukan teknik pengolahan khusus untuk mengekstrak informasi secara efektif. Analisis sentimen merupakan salah satu pendekatan dalam *text mining* yang dapat dimanfaatkan untuk mengidentifikasi, mengekstrak, dan mengklasifikasikan opini publik ke dalam kategori seperti positif dan negatif (Widyadhana et al., 2023).

Metode *Naive Bayes* menjadi salah satu algoritma yang banyak diadopsi dalam analisis sentimen. Metode ini didasarkan pada teorema *Bayes* dengan asumsi independensi antar fitur, sehingga mampu melakukan klasifikasi teks dengan tingkat akurasi yang memadai meskipun menggunakan data latih dalam jumlah terbatas. Keunggulan metode ini terletak pada kesederhanaannya, efisiensi komputasi, serta kemampuannya dalam memproses data berdimensi tinggi yang umumnya ditemui pada data teks (Rozam & Sari, 2024).

Dalam proses pembangunan model klasifikasi, salah satu faktor yang dapat memengaruhi kinerja algoritma adalah pembagian *dataset* menjadi data *training* dan data *testing*. Data *training* digunakan untuk melatih model agar mampu mengenali pola pada data, sedangkan data *testing* digunakan untuk mengevaluasi kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya. Perbedaan proporsi pembagian data dapat menghasilkan performa model yang berbeda karena jumlah data yang digunakan untuk pelatihan dan pengujian turut memengaruhi proses pembelajaran algoritma.

Meskipun berbagai penelitian terkait analisis sentimen telah banyak dilakukan, kajian yang secara spesifik membahas opini publik terhadap Taman Rakyat Slawi Ayu berdasarkan ulasan pada *Google Maps* masih terbatas. Selain itu,

P-ISSN : 2580-4316
https://doi.org/10.37817/ikraith-informatika.v10i2 E-ISSN : 2654-8054

penelitian yang mengkaji pengaruh variasi rasio pembagian data *training* dan *testing* terhadap kinerja algoritma *Multinomial Naïve Bayes* pada analisis sentimen ulasan Taman Rakyat Slawi Ayu juga masih jarang ditemukan. Oleh karena itu, diperlukan pengujian terhadap beberapa skenario pembagian data untuk mengetahui rasio yang mampu menghasilkan performa klasifikasi terbaik.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membandingkan kinerja algoritma *Multinomial Naïve Bayes* pada variasi rasio pembagian data *training* dan *testing*, yaitu 70:30, 80:20, dan 90:10 dalam analisis sentimen ulasan Taman Rakyat Slawi Ayu yang diperoleh dari *Google Maps*. Kinerja model dievaluasi menggunakan nilai *accuracy, precision, recall,* dan *f1-score*. Hasil penelitian diharapkan dapat memberikan informasi mengenai rasio pembagian data yang paling optimal dalam menghasilkan performa klasifikasi yang baik serta menjadi referensi bagi penelitian analisis sentimen selanjutnya yang menggunakan algoritma *Multinomial Naïve Bayes.*

2. LANDASAN TEORI

2.1 Analisis Sentimen

Analisis sentimen merupakan teknik *text mining* yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini dalam data teks ke dalam kategori sentimen, seperti positif dan negatif. Teknik ini banyak dimanfaatkan untuk menganalisis ulasan pengguna pada berbagai *platform digital* sehingga dapat memberikan gambaran mengenai persepsi masyarakat terhadap suatu objek atau layanan. Pada penelitian ini, analisis sentimen digunakan untuk mengklasifikasikan ulasan pengunjung Taman Rakyat Slawi Ayu berdasarkan sentimen positif dan negatif (Novantika & Sugiman, 2022).

2.2 TF-IDF

Term Frequency–Inverse Document Frequency (TF-IDF) merupakan metode pembobotan kata yang digunakan untuk mengubah data teks menjadi representasi numerik. Metode ini memberikan bobot yang lebih tinggi pada kata yang penting dalam suatu dokumen sehingga dapat digunakan sebagai fitur dalam proses klasifikasi. Pada penelitian ini, TF-IDF digunakan sebagai metode ekstraksi fitur sebelum proses klasifikasi menggunakan algoritma *Multinomial Naïve Bayes* (Helmayanti et al., 2023).

2.3 Multinomial Naïve Bayes

Multinomial Naïve Bayes merupakan algoritma klasifikasi yang dikembangkan berdasarkan *Teorema Bayes* dengan asumsi bahwa setiap fitur bersifat independen. Algoritma ini banyak digunakan pada analisis sentimen karena mampu mengolah data teks hasil pembobotan TF-IDF secara efisien. Pada penelitian ini, algoritma *Multinomial Naïve Bayes* digunakan untuk mengklasifikasikan ulasan ke dalam kategori sentimen positif dan negatif (Khoerunnisa et al., 2025).

2.4 Confusion Matrix

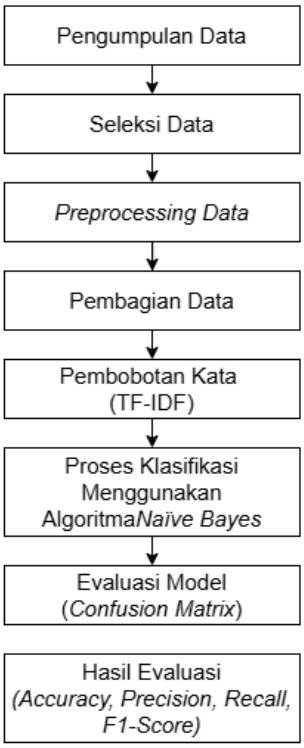
Confusion matrix merupakan metode evaluasi yang digunakan untuk membandingkan hasil prediksi model dengan label sebenarnya. Berdasarkan *confusion matrix* dapat dihitung nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebagai indikator performa model klasifikasi. Pada penelitian ini, metrik tersebut digunakan untuk membandingkan kinerja algoritma *Multinomial Naïve Bayes* pada setiap variasi rasio pembagian data (Humaidi & Maulani, 2023).

3. METODOLOGI

Metode penelitian yang diterapkan dalam penelitian ini terdiri atas beberapa tahapan yang disusun secara sistematis untuk melakukan analisis sentimen terhadap ulasan pengunjung Taman Rakyat Slawi Ayu. Tahapan penelitian diawali dengan pengumpulan data ulasan dari *platform Google Maps*, kemudian dilanjutkan dengan proses seleksi data, *preprocessing data*, pembagian data menjadi data latih dan data uji, pembobotan kata menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF), serta proses klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*. Selanjutnya, performa model dievaluasi menggunakan *confusion matrix* berdasarkan nilai *accuracy*, *precision*, *recall*,

P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

dan *f1-score*. Alur tahapan penelitian secara keseluruhan dapat dilihat pada Gambar 1.



Gambar 1. Alur Tahapan Penelitian

3.1 Pengumpulan Data

Tahapan pengumpulan data pada penelitian ini dilakukan dengan memperoleh ulasan pengunjung Taman Rakyat Slawi Ayu (TRASA) melalui fitur ulasan (*review*) pada *Google Maps*. Proses pengambilan data dilakukan secara otomatis menggunakan teknik *web crawling* berbasis bahasa pemrograman *Python* yang dijalankan pada *Google Colab* dengan bantuan beberapa *library* pendukung. Data yang dikumpulkan terdiri atas teks ulasan dan rating pengguna yang berkaitan dengan fasilitas, kebersihan, kenyamanan, pelayanan, serta kondisi lingkungan taman. Pengambilan data dilakukan pada tanggal 17 Januari 2026. Hasil proses pengumpulan data menghasilkan sebanyak 4.153 ulasan yang selanjutnya digunakan sebagai *dataset* awal dalam penelitian ini.

3.2 Seleksi Data

Setelah proses pengumpulan data selesai dilakukan, tahapan berikutnya adalah seleksi data yang bertujuan untuk memastikan kualitas *dataset* sebelum memasuki proses pengolahan lebih lanjut. Pada tahap ini dilakukan penyaringan terhadap ulasan yang tidak memiliki teks komentar *(missing value)* dan hanya berisi pemberian rating tanpa disertai ulasan tertulis. Karena penelitian ini berfokus pada analisis sentimen berbasis teks, data yang tidak mengandung komentar tidak dapat digunakan dalam proses ekstraksi fitur maupun klasifikasi sentimen. Oleh karena itu, data tersebut dihapus dari *dataset* penelitian. Dari total 4.153 ulasan yang diperoleh pada tahap pengumpulan data, sebanyak 1.794 ulasan dinyatakan memenuhi kriteria dan selanjutnya digunakan pada proses pelabelan sentimen sebelum memasuki tahap *preprocessing*.

3.3 Pelabelan Data

Setelah proses seleksi data selesai dilakukan, tahapan berikutnya adalah pelabelan data yang bertujuan untuk menentukan kelas sentimen pada setiap ulasan. Pada penelitian ini, pelabelan dilakukan berdasarkan nilai rating yang diberikan oleh pengguna pada *Google Maps*. Pendekatan ini digunakan karena rating dapat merepresentasikan tingkat kepuasan pengunjung terhadap Taman Rakyat Slawi Ayu. Pelabelan data dibagi menjadi dua kategori sentimen, yaitu sentimen positif dan sentimen negatif. Ulasan dengan rating 1, 2, dan 3 dikategorikan sebagai sentimen negatif, sedangkan ulasan dengan rating 4 dan 5 dikategorikan sebagai sentimen positif. Hasil pelabelan kemudian digunakan sebagai kelas target dalam proses klasifikasi menggunakan algoritma Multinomial Naïve Bayes. Tabel 1 menunjukkan aturan pelabelan data yang digunakan pada penelitian ini.

Tabel 1. Aturan Pelabelan Sentimen

Rating	Label Sentimen
1 – 3	Negatif
4 – 5	Positif

Data yang telah diberi label selanjutnya diproses pada tahap *preprocessing* untuk menghasilkan data teks yang siap digunakan dalam proses klasifikasi.

3.4 Preprocessing Data

Preprocessing data merupakan tahapan yang dilakukan untuk membersihkan dan menyiapkan data teks agar dapat diproses pada tahap analisis selanjutnya. Tahap ini bertujuan untuk mengurangi *noise*, menyeragamkan bentuk kata, serta meningkatkan kualitas data sehingga proses

klasifikasi dapat menghasilkan performa yang lebih baik. Pada penelitian ini, *preprocessing data* dilakukan melalui beberapa tahapan, yaitu *cleaning, case folding, tokenizing, stopword removal,* dan *stemming* sebagaimana ditunjukkan pada Gambar 2.

Gambar 2. Alur Preprocessing

a. Cleaning

Cleaning merupakan proses pembersihan data teks dari karakter yang tidak diperlukan dalam proses analisis sentimen. Pada tahap ini dilakukan penghapusan tanda baca, angka, simbol, karakter khusus, hyperlink, serta elemen lain yang tidak memiliki kontribusi terhadap proses klasifikasi sentimen (Masykur et al., 2026).

b. Case Folding

Case folding merupakan proses menyeragamkan seluruh huruf pada teks menjadi huruf kecil (*lowercase*). Tahapan ini dilakukan untuk menghindari perbedaan penulisan huruf kapital dan huruf kecil yang dapat menyebabkan kata dengan makna yang sama dianggap sebagai kata yang berbeda oleh sistem (Mubaroroh et al., 2022).

c. Tokenizing

Tokenizing merupakan proses pemecahan kalimat atau dokumen menjadi unit-unit kata yang lebih kecil yang disebut token. Tahapan ini bertujuan untuk memudahkan proses pengolahan teks pada tahap berikutnya karena setiap kata akan diperlakukan sebagai fitur yang dapat dianalisis oleh model klasifikasi (Prasetyo & Fitrani, 2023).

d. Stopword Removal

Stopword removal merupakan proses penghapusan kata-kata umum yang sering muncul dalam dokumen tetapi memiliki kontribusi yang rendah terhadap makna sentimen, seperti kata "dan", "yang", "di", "ke", dan kata umum lainnya. Penghapusan *stopword* dilakukan untuk mengurangi jumlah fitur yang tidak relevan dan meningkatkan efisiensi proses klasifikasi (Misrun et al., 2023).

e. Stemming

Stemming merupakan proses mengubah kata berimbuhan menjadi bentuk kata dasarnya. Sebagai contoh, kata "bermain", "memainkan", dan "permainan" akan diubah menjadi kata dasar "main". Tahapan ini bertujuan untuk mengurangi variasi kata sehingga kata yang memiliki makna serupa dapat direpresentasikan dalam bentuk yang sama (Surbakti et al., 2021).

3.5 Pembagian Data

Setelah melalui tahap *preprocessing*, sebanyak 1.789 data yang siap digunakan kemudian dibagi menjadi data latih (*training data*) dan data uji (*testing data*). Pembagian data dilakukan untuk membentuk model klasifikasi sekaligus mengukur kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya. Data latih digunakan untuk melatih algoritma *Multinomial Naïve Bayes* dalam mengenali pola sentimen pada data ulasan, sedangkan data uji digunakan untuk mengevaluasi performa model yang dihasilkan. Pada penelitian ini digunakan tiga skenario pembagian data *training* dan *testing*, yaitu 70:30, 80:20, dan 90:10. Variasi rasio tersebut digunakan untuk mengetahui pengaruh proporsi data latih dan data uji terhadap performa klasifikasi yang dihasilkan oleh algoritma *Multinomial Naïve Bayes*. Semakin besar proporsi data latih yang digunakan, semakin banyak informasi yang dapat dipelajari oleh model. Namun demikian, proporsi data uji yang lebih kecil dapat memengaruhi representativitas hasil evaluasi. Oleh karena itu, diperlukan pengujian pada beberapa rasio pembagian data untuk memperoleh rasio yang menghasilkan performa terbaik. Skenario pembagian data yang digunakan pada penelitian ini ditunjukkan pada Tabel 2.

Tabel 2. Skenario Pembagian Data

Skenario	Data Latih	Data Uji
S1	70%	30%
S2	80%	20%
S3	90%	10%

Setiap skenario pembagian data digunakan pada proses pelatihan dan pengujian model *Multinomial Naïve Bayes*. Hasil dari masing-masing skenario kemudian dibandingkan berdasarkan nilai *accuracy, precision, recall,* dan *f1-score* untuk menentukan rasio pembagian data yang menghasilkan performa klasifikasi terbaik.

3.6 Pembobotan Kata (TF-IDF)

Metode *Term Frequency-Inverse Document Frequency* (TF-IDF) digunakan untuk mengubah data teks hasil *preprocessing*

P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi (Gautama, 2022). Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen serta tingkat kemunculannya pada seluruh dokumen dalam dataset. Dengan demikian, kata-kata yang dianggap lebih representatif terhadap isi dokumen akan memperoleh bobot yang lebih tinggi dibandingkan kata-kata yang sering muncul pada banyak dokumen. Perhitungan bobot TF-IDF dilakukan menggunakan Persamaan berikut.

$$W_{t,d} = TF_{t,d} \times \log\left(\frac{N}{df_t}\right)$$

Keterangan:

- $W_{t,d}$ = bobot TF-IDF kata t pada dokumen d
- $TF_{t,d}$ = frekuensi kemunculan kata t pada dokumen d
- N = jumlah seluruh dokumen
- df_t = jumlah dokumen yang mengandung kata t

Pada penelitian ini, proses pembobotan kata dilakukan menggunakan *TfidfVectorizer* dari pustaka *scikit-learn*. Pembobotan diterapkan pada data latih dan data uji yang telah melalui tahap *preprocessing*. Hasil pembobotan berupa matriks fitur numerik yang selanjutnya digunakan sebagai masukan pada proses klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*.

3.7 *Multinomial Naïve Bayes*

Multinomial Naïve Bayes merupakan salah satu algoritma klasifikasi yang banyak digunakan dalam analisis sentimen dan pengolahan teks (Prakoso et al., 2022). Algoritma ini merupakan pengembangan dari metode *Naïve Bayes* yang didasarkan pada *Teorema Bayes* dengan asumsi bahwa setiap fitur bersifat independen terhadap fitur lainnya. *Multinomial Naïve Bayes* dirancang untuk menangani data diskrit berupa frekuensi kemunculan kata sehingga sesuai digunakan pada data teks yang telah direpresentasikan menggunakan pembobotan TF-IDF.

Proses klasifikasi dilakukan dengan menghitung probabilitas suatu dokumen termasuk ke dalam kelas tertentu berdasarkan kemunculan kata-kata yang terdapat pada dokumen tersebut. Kelas dengan nilai probabilitas terbesar akan dipilih sebagai hasil klasifikasi. Persamaan dasar *Teorema Bayes* ditunjukkan pada Persamaan berikut.

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

Keterangan:

- $P(C|X)$ = probabilitas dokumen X termasuk ke dalam kelas C
- $P(X|C)$ = probabilitas kemunculan dokumen X pada kelas C
- $P(C)$ = probabilitas awal (prior probability) kelas C
- $P(X)$ = probabilitas kemunculan dokumen X

Pada penelitian ini, algoritma *Multinomial Naïve Bayes* digunakan untuk mengklasifikasikan sentimen ulasan Taman Rakyat Slawi Ayu ke dalam dua kategori, yaitu sentimen positif dan sentimen negatif. Model dibangun menggunakan data latih yang telah melalui proses *preprocessing* dan pembobotan TF-IDF, kemudian dievaluasi menggunakan data uji berdasarkan beberapa skenario pembagian data yang telah ditentukan.

3.8 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja algoritma *Multinomial Naïve Bayes* dalam mengklasifikasikan sentimen ulasan Taman Rakyat Slawi Ayu. Pada penelitian ini, proses evaluasi dilakukan menggunakan *confusion matrix*, yaitu metode evaluasi yang membandingkan hasil prediksi model dengan label sebenarnya pada data uji. Melalui confusion matrix, dapat diketahui jumlah prediksi yang benar maupun salah untuk setiap kelas sentimen. Struktur confusion matrix yang digunakan pada penelitian ini ditunjukkan pada Tabel 3.

Tabel 3. *Confusion Matrix*

	Prediksi Positif	**Prediksi Negatif**
Aktual Positif	*True Positive* (TP)	*False Negative* (FN)
Aktual Negatif	*False Positive* (FP)	*True Negative* (TN)

Berdasarkan nilai yang diperoleh dari *confusion matrix*, performa model dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *f1-score*. Metrik-metrik tersebut dipilih karena mampu memberikan gambaran

https://doi.org/10.37817/ikraith-informatika.v10i2
P-ISSN : 2580-4316
E-ISSN : 2654-8054

menyeluruh mengenai performa model dalam mengklasifikasikan sentimen positif dan negatif.

a) *Accuracy*

Accuracy digunakan untuk mengukur tingkat ketepatan model secara keseluruhan dengan membandingkan jumlah prediksi yang benar terhadap total data yang diuji. Semakin tinggi nilai *accuracy*, semakin baik kemampuan model dalam melakukan klasifikasi. Perhitungan *accuracy* ditunjukkan pada Persamaan berikut.

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

b) *Precision*

Precision digunakan untuk mengukur tingkat ketepatan prediksi pada kelas positif. Metrik ini menunjukkan seberapa banyak data yang diprediksi positif oleh model benar-benar termasuk ke dalam kelas positif. Nilai *precision* yang tinggi menunjukkan rendahnya kesalahan *False Positive* (FP). Perhitungan *precision* ditunjukkan pada Persamaan berikut.

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

c) *Recall*

Recall digunakan untuk mengukur kemampuan model dalam menemukan seluruh data positif yang terdapat pada dataset. Nilai *recall* yang tinggi menunjukkan bahwa model mampu meminimalkan kesalahan *False Negative* (FN). Perhitungan *recall* ditunjukkan pada Persamaan berikut.

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

d) *F1-Score*

F1-Score merupakan rata-rata harmonis (*harmonic mean*) dari nilai *precision* dan *recall*. Metrik ini digunakan untuk memberikan keseimbangan antara ketepatan prediksi dan kemampuan model dalam menemukan data positif, sehingga sangat sesuai digunakan pada dataset yang tidak seimbang. Perhitungan *F1-Score* ditunjukkan pada Persamaan berikut.

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

Keterangan:

- TP (*True Positive*) = jumlah data positif yang diprediksi positif.
- TN (*True Negative*) = jumlah data negatif yang diprediksi negatif.
- FP (*False Positive*) = jumlah data negatif yang diprediksi positif.
- FN (*False Negative*) = jumlah data positif yang diprediksi negatif.

Nilai *accuracy*, *precision*, *recall*, dan *f1-score* yang diperoleh dari setiap skenario pembagian data kemudian dibandingkan untuk menentukan rasio data *training* dan *testing* yang menghasilkan performa klasifikasi terbaik pada algoritma *Multinomial Naïve Bayes*.

4. HASIL DAN PEMBAHASAN

4.1 Hasil Pengumpulan Data

Tahap pengumpulan data menghasilkan sebanyak 4.153 ulasan pengunjung Taman Rakyat Slawi Ayu yang diperoleh dari *Google Maps*. Dataset yang diperoleh terdiri atas informasi rating dan teks ulasan yang diberikan oleh pengguna. Data tersebut selanjutnya digunakan sebagai dataset awal pada penelitian ini sebelum memasuki tahap seleksi data.

4.2 Hasil Seleksi Data

Tahap seleksi data dilakukan dengan menyaring ulasan yang tidak memiliki teks komentar dan hanya berisi rating. Hasil seleksi menunjukkan bahwa tidak seluruh data hasil pengumpulan dapat digunakan pada proses analisis sentimen karena sebagian data tidak mengandung informasi teks yang dapat diolah. Berdasarkan proses seleksi yang dilakukan, jumlah data berkurang dari 4.153 ulasan menjadi 1.794 ulasan yang memenuhi kriteria untuk dianalisis lebih lanjut. Sebanyak 2.359 data tidak digunakan karena tidak memiliki teks ulasan atau hanya berisi rating tanpa komentar. Sebanyak 1.794 data yang memenuhi kriteria kemudian digunakan pada tahap pelabelan data untuk menentukan kategori sentimen positif dan negatif.

4.3 Hasil Pelabelan Data

Tahap pelabelan data dilakukan berdasarkan nilai rating yang diberikan oleh pengguna pada *Google Maps*. Ulasan dengan rating 1, 2, dan 3 dikategorikan sebagai sentimen negatif, sedangkan ulasan dengan

P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

rating 4 dan 5 dikategorikan sebagai sentimen positif. Pelabelan dilakukan terhadap seluruh data hasil seleksi yang berjumlah 1.794 ulasan. Hasil pelabelan menunjukkan bahwa dataset didominasi oleh sentimen positif. Distribusi data berdasarkan kategori sentimen ditunjukkan pada Tabel 4.

Tabel 4. Distribusi Label Sentimen

Label Sentimen	Jumlah Data
Positif	1527
Negatif	267
Total	1794

Berdasarkan hasil pelabelan, sebanyak 1.527 data (85,12%) termasuk ke dalam kategori sentimen positif, sedangkan 267 data (14,88%) termasuk ke dalam kategori sentimen negatif. Hasil tersebut menunjukkan bahwa dataset didominasi oleh sentimen positif. Data yang telah diberikan label sentimen kemudian digunakan pada tahap *preprocessing* untuk mempersiapkan data sebelum dilakukan proses klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*.

4.4 Hasil Preprocessing Data

Tahap *preprocessing* menghasilkan data teks yang lebih bersih dan terstruktur melalui proses *cleaning, case folding, tokenizing, stopword removal*, dan *stemming*. Proses ini bertujuan untuk mengurangi *noise* pada data sehingga kata-kata yang tidak relevan dapat dihilangkan dan teks lebih siap untuk diolah pada tahap pembobotan kata menggunakan metode TF-IDF.

Selain menghasilkan data yang lebih terstruktur, tahap *preprocessing* juga menyebabkan berkurangnya jumlah data yang dapat digunakan. Dari total 1.794 data hasil pelabelan, ditemukan 5 data yang menghasilkan nilai kosong (*missing value*) setelah proses *preprocessing*. Data tersebut kemudian dihapus karena tidak dapat digunakan dalam proses pembobotan TF-IDF maupun klasifikasi sentimen. Dengan demikian, jumlah data yang digunakan pada tahap selanjutnya menjadi 1.789 data.

4.5 Hasil Pembagian Data

Setelah melalui tahap *preprocessing*, diperoleh sebanyak 1.789 data yang siap digunakan pada proses klasifikasi. Data tersebut kemudian dibagi menjadi data latih (*training data*) dan data uji (*testing data*) menggunakan tiga skenario pembagian data, yaitu 70:30, 80:20, dan 90:10. Pembagian data dilakukan untuk mengetahui pengaruh proporsi data latih dan data uji terhadap performa klasifikasi yang dihasilkan oleh algoritma *Multinomial Naïve Bayes*. Hasil pembagian data pada masing-masing skenario ditunjukkan pada Tabel 5.

Tabel 5. Hasil Pembagian Data

Skenario	Data Latih	Data Uji
S1 (70:30)	1.252	537
S2 (80:20)	1.431	358
S3 (90:10)	1.610	179

Berdasarkan Tabel 5, semakin besar proporsi data latih yang digunakan maka semakin banyak data yang tersedia untuk proses pembelajaran model. Sebaliknya, jumlah data uji akan semakin berkurang sehingga proses evaluasi dilakukan pada jumlah data yang lebih sedikit. Ketiga skenario pembagian data tersebut selanjutnya digunakan untuk membangun dan mengevaluasi model *Multinomial Naïve Bayes*.

4.6 Hasil Klasifikasi *Multinomial Naïve Bayes*

Proses klasifikasi dilakukan menggunakan algoritma *Multinomial Naïve Bayes* terhadap data yang telah melalui tahap *preprocessing*, pembagian data dan pembobotan TF-IDF. Pengujian dilakukan menggunakan tiga skenario pembagian data, yaitu 70:30, 80:20, dan 90:10. Hasil klasifikasi dievaluasi menggunakan nilai *accuracy*, *precision*, *recall*, dan *f1-score* pada data uji.

a) Skenario S1 (70:30)

Pada skenario pertama, dataset dibagi menjadi 70% data latih dan 30% data uji. Hasil pengujian menunjukkan bahwa model memperoleh akurasi data latih (*training data*) sebesar 91,21% dan akurasi data uji (*testing data*) sebesar 85,66%. Hasil *classification* report untuk skenario S1 ditunjukkan pada Gambar 3.

```
Akurasi Data Training: 91.21%
Akurasi Data Testing: 85.66%

=== Classification Report (Data Testing) ===
              precision    recall  f1-score   support

     negatif       0.57      0.15      0.24        80
     positif       0.87      0.98      0.92       457

    accuracy                           0.86       537
   macro avg       0.72      0.57      0.58       537
weighted avg       0.82      0.86      0.82       537
```

Gambar 3. Hasil Klasifikasi pada Rasio 70:30

https://doi.org/10.37817/ikraith-informatika.v10i2
P-ISSN : 2580-4316
E-ISSN : 2654-8054

Berdasarkan Gambar 3, kelas positif memperoleh nilai *precision* sebesar 0,87, *recall* sebesar 0,98, dan *f1-score* sebesar 0,92. Sementara itu, kelas negatif memperoleh nilai *precision* sebesar 0,57, *recall* sebesar 0,15, dan *f1-score* sebesar 0,24.

b) Skenario S2 (80:20)

Pada skenario kedua, dataset dibagi menjadi 80% data latih dan 20% data uji. Hasil pengujian menunjukkan bahwa model memperoleh akurasi data latih (*training data*) sebesar 91,19% dan akurasi data uji (*testing data*) sebesar 86,59%. Hasil *classification report* untuk skenario S2 ditunjukkan pada Gambar 4.

```
Akurasi Data Training: 91.19%
Akurasi Data Testing: 86.59%

=== Classification Report (Data Testing) ===
              precision    recall  f1-score   support

     negatif       0.65      0.21      0.31        53
     positif       0.88      0.98      0.93       305

    accuracy                           0.87       358
   macro avg       0.76      0.59      0.62       358
weighted avg       0.84      0.87      0.84       358
```

Gambar 4. Hasil Klasifikasi pada Rasio 80:20

Berdasarkan Gambar 4, kelas positif memperoleh nilai *precision* sebesar 0,88, *recall* sebesar 0,98, dan *f1-score* sebesar 0,93. Adapun kelas negatif memperoleh nilai *precision* sebesar 0,65, *recall* sebesar 0,21, dan *f1-score* sebesar 0,31.

c) Skenario S3 (90:10)

Pada skenario ketiga, dataset dibagi menjadi 90% data latih dan 10% data uji. Hasil pengujian menunjukkan bahwa model memperoleh akurasi data latih (*training data*) sebesar 90,75% dan akurasi data uji (*testing data*) sebesar 87,71%. Hasil *classification report* untuk skenario S3 ditunjukkan pada Gambar 5.

```
Akurasi Data Training: 90.75%
Akurasi Data Testing: 87.71%

=== Classification Report (Data Testing) ===
              precision    recall  f1-score   support

     negatif       0.86      0.22      0.35        27
     positif       0.88      0.99      0.93       152

    accuracy                           0.88       179
   macro avg       0.87      0.61      0.64       179
weighted avg       0.87      0.88      0.84       179
```

Gambar 5. Hasil Klasifikasi pada Rasio 90:10

Berdasarkan Gambar 5, kelas positif memperoleh nilai *precision* sebesar 0,88, *recall* sebesar 0,99, dan *f1-score* sebesar 0,93. Sementara itu, kelas negatif memperoleh nilai *precision* sebesar 0,86, *recall* sebesar 0,22, dan *f1-score* sebesar 0,35.

Berdasarkan hasil pengujian pada ketiga skenario pembagian data tersebut, selanjutnya dilakukan perbandingan kinerja model untuk mengetahui rasio pembagian data yang menghasilkan performa klasifikasi terbaik.

4.7 Perbandingan Kinerja Model

Untuk mengetahui pengaruh variasi rasio pembagian data terhadap performa klasifikasi, dilakukan perbandingan hasil pengujian pada ketiga skenario pembagian data yang telah diterapkan. Perbandingan dilakukan berdasarkan nilai *accuracy*, *precision*, *recall*, dan *f1-score* yang diperoleh dari hasil klasifikasi pada data uji. Hasil perbandingan kinerja model ditunjukkan pada Gambar 6.

Rasio Data	*Accuracy* (%)	*Precision*	*Recall*	*F1-Score*
70:30	85,66	0,87	0,98	0,92
80:20	86,59	0,88	0,98	0,93
90:10	87,71	0,88	0,99	0,93

Gambar 6. Perbandingan Kinerja Model pada Berbagai Rasio Pembagian Data

Berdasarkan Gambar 6, nilai *accuracy* mengalami peningkatan pada setiap skenario pembagian data. Rasio 70:30 menghasilkan nilai *accuracy* sebesar 85,66%, kemudian meningkat menjadi 86,59% pada rasio 80:20 dan mencapai nilai tertinggi sebesar 87,71% pada rasio 90:10. Hasil ini menunjukkan bahwa semakin besar proporsi data latih yang digunakan, semakin banyak pola yang dapat dipelajari oleh algoritma *Multinomial Naïve Bayes* sehingga performa klasifikasi cenderung meningkat.

Nilai *precision* juga mengalami peningkatan dari 0,87 pada rasio 70:30 menjadi 0,88 pada rasio 80:20 dan tetap berada pada nilai yang sama pada rasio 90:10. Sementara itu, nilai *recall* meningkat dari 0,98 menjadi 0,99 pada rasio 90:10, yang menunjukkan bahwa model semakin baik dalam mengenali ulasan yang termasuk ke dalam kategori sentimen positif.

Nilai *recall* yang relatif tinggi pada seluruh skenario menunjukkan bahwa model mampu mengenali sebagian besar ulasan yang termasuk ke dalam kategori sentimen positif. Hal ini juga dipengaruhi oleh distribusi dataset yang didominasi oleh sentimen positif dibandingkan sentimen negatif.

Nilai f1-score mengalami peningkatan dari 0,92 pada rasio 70:30 menjadi 0,93 pada rasio 80:20 dan tetap berada pada nilai yang

sama pada rasio 90:10. Hal ini menunjukkan bahwa model mampu mempertahankan keseimbangan antara nilai *precision* dan *recall* pada seluruh skenario pengujian.

Berdasarkan hasil perbandingan yang diperoleh, rasio pembagian data 90:10 menghasilkan performa terbaik dengan nilai *accuracy* sebesar 87,71%, *precision* sebesar 0,88, *recall* sebesar 0,99, dan *f1-score* sebesar 0,93. Hasil tersebut menunjukkan bahwa penggunaan proporsi data latih yang lebih besar memberikan kemampuan pembelajaran yang lebih baik bagi model dalam melakukan klasifikasi sentimen ulasan Taman Rakyat Slawi Ayu.

5. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma *Multinomial Naïve Bayes* untuk analisis sentimen ulasan Taman Rakyat Slawi Ayu yang diperoleh dari *Google Maps*. Dari 4.153 data ulasan yang berhasil dikumpulkan, diperoleh 1.794 data yang memenuhi kriteria setelah proses seleksi data. Selanjutnya, setelah melalui tahap *preprocessing*, jumlah data yang digunakan dalam proses klasifikasi menjadi 1.789 data.

Pengujian dilakukan menggunakan tiga variasi rasio pembagian data, yaitu 70:30, 80:20, dan 90:10. Hasil pengujian menunjukkan bahwa variasi rasio pembagian data memberikan pengaruh terhadap performa klasifikasi yang dihasilkan oleh algoritma *Multinomial Naïve Bayes*. Rasio 70:30 menghasilkan nilai *accuracy* sebesar 85,66%, *precision* sebesar 0,87, *recall* sebesar 0,98, dan *f1-score* sebesar 0,92. Rasio 80:20 menghasilkan nilai *accuracy* sebesar 86,59%, *precision* sebesar 0,88, *recall* sebesar 0,98, dan *f1-score* sebesar 0,93. Sementara itu, rasio 90:10 menghasilkan performa terbaik dengan nilai *accuracy* sebesar 87,71%, *precision* sebesar 0,88, *recall* sebesar 0,99, dan *f1-score* sebesar 0,93.

Berdasarkan hasil pengujian yang dilakukan, rasio pembagian data 90:10 menghasilkan performa terbaik pada dataset penelitian ini dengan nilai accuracy sebesar 87,71%, *precision* sebesar 0,88, *recall* sebesar 0,99, dan *f1-score* sebesar 0,93. Hasil tersebut menunjukkan bahwa penggunaan proporsi data latih yang lebih besar mampu meningkatkan kemampuan model dalam mempelajari pola sentimen pada data ulasan. Namun demikian, penggunaan rasio 90:10 juga memiliki keterbatasan karena jumlah data uji yang digunakan relatif lebih sedikit dibandingkan skenario lainnya, yaitu hanya 10% dari keseluruhan dataset. Kondisi ini dapat menyebabkan hasil evaluasi kurang merepresentasikan variasi data yang lebih luas. Oleh karena itu, pemilihan rasio pembagian data perlu mempertimbangkan keseimbangan antara jumlah data latih yang digunakan untuk pembelajaran model dan jumlah data uji yang digunakan untuk evaluasi performa.

DAFTAR PUSTAKA

Gautama, I. M. B. (2022). Klasifikasi Data Saran Pemustaka di Perpustakaan STIKOM Bali Menggunakan TF-IDF dan Multinomial Naive Bayes. *JURNAL SISTEM DAN INFORMATIKA*, 62–72.

Helmayanti, S. A., Hamami, F., & Fa'rifah, R. Y. (2023). PENERAPAN ALGORITMA TF-IDF DAN NAÏVE BAYES Jurnal Indonesia : Manajemen Informatika dan Komunikasi. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, *4*(3), 1822–1834.

Humaidi, M. R., & Maulani, A. (2023). KLASIFIKASI NAÏVE BAYES DAN CONFUSION MATRIX PADA PENGGUNA APLIKASI E-COMMERCE DI PLAY STORE. *Jurnal Ilmiah Informatika*, *8*(2), 132–139.

Khoerunnisa, S., Shiddiq, D. F., & Nurhayati, D. (2025). Application of the Naive Bayes Algorithm with TF-IDF and Cross Validation Techniques for Sentiment Analysis Towards Starlink. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *5*(April), 566–577.

Masykur, A. A., Gumilang, R. R., & Rosyid, H. Al. (2026). Analisis Sentimen X : Kegagalan Timnas ke Piala Dunia 2026 dengan Naive Bayes. *ELKOM (Jurnal Elektronika Dan Komputer)*, *18*(2), 340–351.

Misrun, C. A., Haerani, E., Fikry, M., & Budianita, E. (2023). Analisis sentimen komentar youtube terhadap Anies Baswedan sebagai bakal calon presiden

2024 menggunakan metode naive bayes classifier. *Jurnal Computer Science and Information Technology (CoSciTech)*.

Mubaroroh, H. H., Yasin, H., & Rusgiyono, A. (2022). ANALISIS SENTIMEN DATA ULASAN APLIKASI RUANGGURU PADA SITUS GOOGLE PLAY MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER DENGAN NORMALISASI KATA LEVENSHTEIN DISTANCE. *JURNAL GAUSSIAN*, *11*(1), 257–266.

Novantika, A., & Sugiman. (2022). Analisis Sentimen Ulasan Pengguna Aplikasi Video Conference Google Meet menggunaka n Metode SVM dan Logistic Regression. *PRISMA, Prosiding Seminar Nasional Matematika*, *5*, 808–813.

Permana, A. P., Chamidy, T., & Crysdian, C. (2023). Klasifikasi Ulasan Fasilitas Publik Menggunakan Metode Naïve Bayes dengan Seleksi Fitur Chi-Square. *JISKA (Jurnal Informatika Sunan Kalijaga)*, *8*(2), 112–124.

Prakoso, Q. A. N., Muliawati, A., & Isnainiyah, I. N. (2022). Analisis Sentimen terhadap Produk Skin Game di Forum Review Female Daily Menggunakan Metode Multinomial Naïve Bayes dan TF-IDF. *JURNAL INFORMATIK*, *4221*, 198–207.

Prasetyo, H., & Fitrani, A. S. (2023). Sentiment Analysis Before Presidential Election 2024 Using Naïve Bayes Classifier Based On Public Opinion In Twitter. *Procedia of Engineering and Life Science*, *4*(June 2023).

Rozam, N. F., & Sari, T. N. (2024). Analisis Sentimen Penilaian Masyarakat Terhadap Rumah Sakit Berdasarkan Ulasan Pada Google Maps Menggunakan Naive Bayes. *Jurnal Teknologi Informasi*, *8*(2).

Surbakti, A. Q., Hayami, R., & Al-amien, J. (2021). Analisa tanggapan terhadap PSBB di indonesia dengan algoritma decision tree pada twitter. *Jurnal Computer Science and Information Technology (CoSciTech)*, *2*(2), 91–97.

Widyadhana, F. K., Setiawan, N. Y., & Rahayudi, B. (2023). Sentimen Analysis pada Opini Masyarakat terhadap Pelayanan Publik Polres Ponorogo menggunakan Metode Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *7*(7), 3047–3056.