

Perbandingan Model Machine Learning dalam Analisis Sentimen Ulasan Pengunjung Keraton Yogyakarta pada Google Maps

Cahyo Prakoso^{*}, Arief Hermawan

Fakultas Sains dan Teknologi, Program Studi Informatika, Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia

Email: ^{1,*}cahyo.5200411379@student.uty.ac.id, ²ariefdb@uty.ac.id

Email Penulis Korespondensi: cahyo.5200411379@student.uty.ac.id

Abstrak—Keraton Yogyakarta, sebagai pusat budaya Jawa yang kaya akan sejarah, seni, dan tradisi, menarik perhatian dunia dan menjadi subjek berbagai ulasan wisata. Dalam era digital, ulasan-ulasan ini menjadi penting untuk pengelolaan dan peningkatan layanan pariwisata. Khususnya, teknologi informasi seperti Google Maps memfasilitasi partisipasi publik dengan memberikan platform bagi pengunjung untuk berbagi pengalaman mereka melalui ulasan. Ulasan ini menyediakan umpan balik penting bagi pengelola dan calon pengunjung, seraya mendorong peningkatan layanan dan inovasi. Oleh karena itu, penelitian ini melakukan analisis sentimen ulasan dengan model Machine Learning yang digunakan untuk memahami pandangan positif dan negatif pengunjung serta melakukan perbandingan pada tiga model populer dalam analisis sentimen, yaitu Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM). Pengumpulan data ulasan menggunakan metode web scraping dengan Selenium, diikuti oleh tahap preprocessing yang meliputi translasi, cleaning, normalisasi, tokenizing, remove stopwords, dan stemming untuk mempersiapkan data untuk analisis sentimen. Labeling dilakukan berdasarkan rating ulasan, dengan threshold di bawah 4 untuk sentimen negatif dan 4 atau 5 untuk positif. Berdasarkan pemrosesan dan analisis data, model Support Vector Machine (SVM) terbukti menjadi model terbaik dengan akurasi 87.12% dan F1-Score 90.91%, diikuti oleh Logistic Regression dan Naive Bayes.

Kata Kunci: Keraton Yogyakarta; Analisis Sentimen; Naive Bayes; Logistic Regression; Support Vector Machine

Abstract—The Yogyakarta Palace, as a center of Javanese culture rich in history, art and tradition, attracts world attention and is the subject of various tourist reviews. In the digital era, these reviews are important for managing and improving tourism services. In particular, information technologies such as Google Maps facilitate public participation by providing a platform for visitors to share their experiences through reviews. These reviews provide important feedback for managers and potential visitors, while driving service improvements and innovation. Therefore, this research conducts review sentiment analysis using a Machine Learning model which is used to understand visitors' positive and negative views and compares three popular models in sentiment analysis, namely Naive Bayes, Logistic Regression, and Support Vector Machine (SVM). Review data collection uses the web scraping method with Selenium, followed by a preprocessing stage which includes translation, cleaning, normalization, tokenizing, removing stopwords, and stemming to prepare the data for sentiment analysis. Labeling is done based on review ratings, with a threshold below 4 for negative sentiment and 4 or 5 for positive. Based on data processing and analysis, the Support Vector Machine (SVM) model was proven to be the best model with an accuracy of 87.12% and an F1-Score of 90.91%, followed by Logistic Regression and Naive Bayes.

Keywords: Yogyakarta Palace; Sentiment Analysis; Naive Bayes; Logistic Regression; Support Vector Machine

1. PENDAHULUAN

Keraton Yogyakarta adalah kompleks istana yang terletak di Yogyakarta. Keraton Yogyakarta menjadi tempat tinggal bagi Sultan Yogyakarta dan merupakan salah satu tempat wisata dan budaya terpenting di Indonesia. Keraton ini memiliki sejarah yang kaya dan menjadi pusat budaya, seni, dan tradisi Jawa yang penting[1]. Di dalam kompleks Keraton Yogyakarta, pengunjung dapat menemukan berbagai bangunan bersejarah, museum, dan area terbuka yang digunakan untuk berbagai acara dan upacara tradisional[2]. Keraton Yogyakarta telah lama menjadi ikon pariwisata menarik ribuan pengunjung dari berbagai penjuru dunia setiap tahunnya, tidak hanya karena keindahan arsitektur klasiknya, tetapi juga karena kekayaan budaya yang tersimpan di setiap sudutnya. Sejarah panjang yang melibatkan para raja dan ratu, serta tradisi keraton yang masih lestari, menjadikan Keraton Yogyakarta sebagai salah satu tempat yang paling menarik dan penuh makna di Indonesia.

Perkembangan teknologi informasi, terutama melalui platform Google Maps, telah mengubah cara pengunjung berinteraksi dengan tempat-tempat wisata seperti Keraton Yogyakarta. Google Maps memberi peluang bagi pengunjung untuk lebih berpartisipasi dengan memberikan ulasan serta feedback langsung mengenai pengalaman mereka di tempat tersebut. Keterbukaan informasi melalui Google Maps telah membuka pintu bagi pengunjung Keraton Yogyakarta untuk berbagi pandangan dan pendapat mereka tentang pengalaman yang mereka alami. Ini adalah langkah positif dalam menjalin keterlibatan yang lebih aktif dan terbuka antara pengelola tempat wisata dan pengunjung. Ulasan-ulasan ini mencakup berbagai aspek, seperti kualitas pelayanan yang diberikan oleh petugas, fasilitas yang tersedia, serta nuansa keseluruhan yang dapat dirasakan di Keraton[3]. Pengunjung dapat dengan mudah memberikan umpan balik positif atau kritik konstruktif melalui platform ini, yang dapat membantu pengelola tempat wisata untuk terus meningkatkan kualitas pengalaman yang mereka tawarkan[4]. Dengan demikian, ulasan-ulasan ini tidak hanya memberikan panduan berharga bagi calon pengunjung yang ingin merencanakan kunjungan mereka, tetapi juga memberikan dorongan bagi Keraton Yogyakarta untuk terus berinovasi dan meningkatkan kualitas layanan mereka.

Namun, untuk memahami sentimen di balik ulasan-ulasan ini, diperlukan analisis lebih lanjut salah satunya dengan melakukan analisis sentimen. Analisis sentimen adalah langkah penting dalam mengungkapkan apakah ulasan tersebut bersifat positif atau negatif. Penggunaan berbagai model Machine Learning akan memungkinkan kita untuk mengidentifikasi pola-pola dalam ulasan, termasuk kata-kata kunci yang menggambarkan pengalaman positif atau

negatif. Beberapa model machine learning yang sering digunakan dalam konteks ini adalah Naive Bayes[5], [6], Logistic Regression[6] dan Support Vector Machine (SVM)[7]. Naive Bayes adalah model probabilistik yang memanfaatkan prinsip probabilitas untuk mengklasifikasikan sentimen, sementara Logistic Regression adalah model yang memodelkan hubungan logistik antara fitur-fitur dan sentimen, dan SVM adalah model yang mencari pemisah linier terbaik antara sentimen positif dan negatif.

Berdasarkan penelitian yang akan dilakukan oleh [8] menggunakan data ulasan lima perguruan tinggi termewah yang diperoleh dari Google Maps. Dataset terdiri dari 50 data pada setiap perguruan tinggi dan membaginya menjadi 40 untuk data latih dan 10 untuk data uji. Penelitian ini menggunakan algoritma Support Vector Machine sebagai algoritma dalam penentuan akurasi. Hasil penelitian menunjukkan bahwa Universitas Ciputra menempati peringkat pertama dengan tingkat keakuratan sebesar 85,00%, diikuti oleh Presiden University dengan 75,00%. Universitas Tri Sakti menduduki peringkat ketiga dengan tingkat keakuratan sebesar 67,50%, sementara Universitas Pelita Harapan berada di peringkat keempat dengan 62,50%, dan Binus University di peringkat kelima juga dengan tingkat keakuratan 62,50%. Penelitian lain yang dilakukan [9] menggunakan data ulasan pada aplikasi video conference Google Meet. Data yang diperoleh dalam penelitian ini mencakup tanggal ulasan, nama pengguna, dan konten review. Total data yang berhasil dikumpulkan melalui proses scraping pada platform Google Meet sejumlah 5160 ulasan. Analisis sentimen dilakukan dengan memanfaatkan dua algoritma, yaitu algoritma Support Vector Machine (SVM) dan algoritma Logistic Regression. Pada algoritma SVM diterapkan dengan empat jenis kernel yang berbeda, yaitu kernel linear, kernel basis radial (RBF), kernel sigmoid, dan kernel polinomial. Hasil akurasi pada data ulasan aplikasi Google Meet untuk masing-masing kernel adalah 87,02%, 84,59%, 86,63%, dan 71,12%, sementara metode Logistic Regression mencapai akurasi sebesar 85,17%.

Penelitian [10] melakukan analisis sentimen ulasan pada e-commerce shopee. Pada penelitian tersebut, menggunakan 2000 dataset ulasan yang diambil dari aplikasi E-commerce Shopee pada website google playstore, kemudian dikelompokkan menjadi dua kelas yaitu postingan positif dan postingan negatif. Analisis sentimen ini menggunakan metode Naive Bayes dan Support Vector Machine. Hasil dalam penelitian ini, metode Naive Bayes menghasilkan akurasi sebesar 85%, menunjukan hasil lebih baik daripada metode Support Vector Machine yang hanya mencapai akurasi 81%. elain itu, metode Naive Bayes juga memberikan hasil yang lebih baik dalam hal akurasi dan Area Under the Curve (AUC) dibandingkan dengan metode Support Vector Machine. Lalu penelitian [11] melakukan analisis sentimen penerbangan Lion Air berdasarkan ulasan pengguna platform online. Penelitian ini dimulai dengan proses pengumpulan data melalui metode web scraping, menghasilkan 1955 ulasan yang disimpan dalam format CSV. Data tersebut kemudian dianalisis secara manual dengan dikategorikan berdasarkan Jejaring Penelitian dan Pengabdian Masyarakat (JPPM), ulasan mengenai rute perjalanan, serta kota terkait. Implementasi pada penelitian ini menggunakan tiga algoritma, yaitu Naive Bayes, Random Forest, dan Logistic Regression. Hasil penelitian menunjukkan tingkat akurasi klasifikasi menggunakan Logistic Regression sebesar 0.82, dengan presisi 0.82, recall 0.78, F1-score 0.80, dan akurasi 0.82, sementara metode Naive Bayes dan Random Forest masing-masing menghasilkan akurasi 0.47 dan 0.39.

Terakhir penelitian [12] melakukan analisis sentimen terhadap ulasan pada aplikasi Kredivo. Data penelitian ini diambil dari website Google Play Store, data yang diambil yaitu data teks ulasan dengan jumlah 10000 ulasan. Proses pengklasifikasian dilakukan dengan menggunakan algoritma klasifikasi Naive Bayes dan Mesin Vector Pendukung (Support Vector Machine). Hasil pengujian menggunakan algoritma Support Vector Machine menunjukkan akurasi sebesar 83,3% dengan presisi 77% untuk kelas positif dan 87% untuk kelas negatif, serta recall 89% untuk kelas positif dan 73% untuk kelas negatif. Sementara itu, algoritma Naive Bayes Classifier mencapai akurasi sebesar 80,8% dengan presisi 81% untuk kelas positif dan 87% untuk kelas negatif, serta recall 88% untuk kelas positif dan 79% untuk kelas negatif.

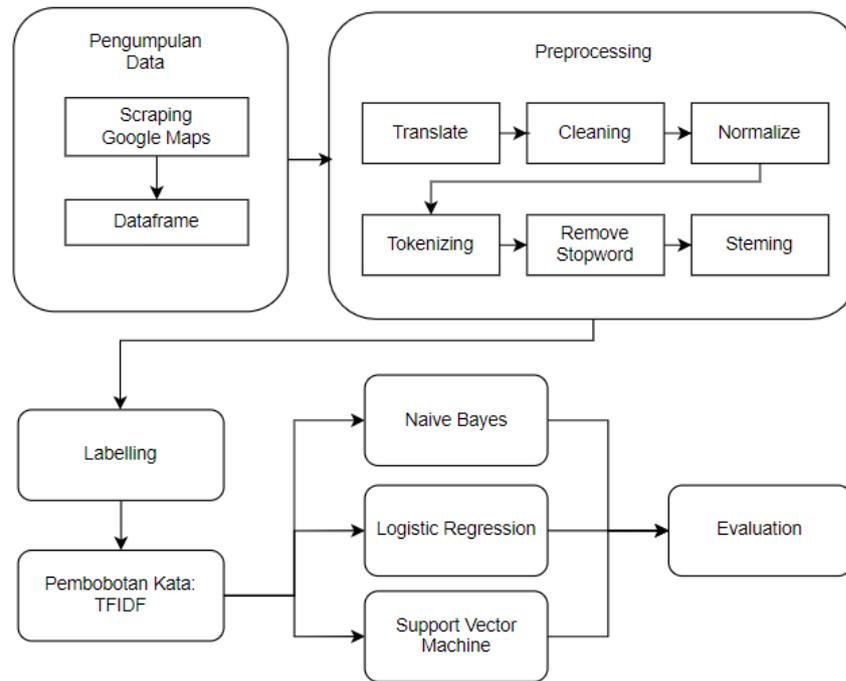
Sebagaimana telah diuraikan sebelumnya, penelitian ini akan memfokuskan perbandingan pada tiga metode populer dalam analisis sentimen, yaitu Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM). Masing-masing metode memiliki kelebihan dan kekurangan sendiri-sendiri, dan tujuan kami adalah untuk mengidentifikasi model yang paling efektif dalam mengklasifikasikan sentimen ulasan. Dengan demikian, penelitian ini akan memberikan wawasan yang berharga tentang bagaimana teknik-teknik Machine Learning dapat digunakan untuk menggali pemahaman lebih dalam terkait pengalaman pengunjung Keraton Yogyakarta, yang pada gilirannya dapat mendukung pengelolaan dan peningkatan layanan wisata di destinasi ini. Selain itu, hasil penelitian ini juga dapat bermanfaat bagi pemangku kepentingan lainnya, seperti pihak berwenang, pengelola wisata, dan pihak terkait dalam mengambil keputusan yang lebih informasional dan data-driven dalam meningkatkan kualitas layanan di Keraton Yogyakarta. Dengan demikian, penelitian ini memberikan kontribusi positif dalam menggali potensi peningkatan wisata dan pemberdayaan sumber daya lokal.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada Gambar 1, disajikan gambaran umum proses dari penelitian ini. Langkah awal melibatkan pengumpulan review dari Google Maps menggunakan teknik web scraping. Pada langkah selanjutnya, data menjalani pra-pemrosesan yang meliputi beberapa rangkaian proses untuk menghilangkan noise. Data tersebut kemudian pembobotan menggunakan TF-IDF (Term Frekuensi-Inverse Document Frekuensi). Untuk proses pemodelan klasifikasi, kami menggunakan Naive Bayes,

Logistic Regressio, dan Support Vector Machine (SVM). Terakhir, kami mengevaluasi performa model menggunakan matriks konfusi dan membandingkan ketiga model.



Gambar 1. Tahapan Penelitian

2.2 Pengumpulan Data

Data ulasan pengunjung Istana Yogyakarta diperoleh dari Google Maps dengan menggunakan metode web scraping. Web scraping adalah teknik ekstraksi otomatis data dari halaman web[13]. Dalam konteks ini, pustaka Selenium digunakan untuk mengotomatisasi browser web, memungkinkan pengambilan data dari halaman web yang dinamis seperti Google Maps. Proses pengumpulan data dimulai dengan mengakses halaman Google Maps yang berisi ulasan Istana Yogyakarta. Kemudian, kami menggunakan Selenium untuk mengotomatisasi tindakan seperti menggulir halaman, mengklik tombol "Tampilkan lebih banyak ulasan" jika diperlukan, dan mengekstrak teks ulasan, peringkat, dan tanggal ulasan dari setiap entri. Kami juga memastikan bahwa kami mengumpulkan data dari berbagai ulasan untuk mendapatkan representasi yang lebih baik dari pandangan pengunjung.

2.4 Preprocessing

2.4.1 Translasi

Langkah ini menjadi awal dalam memastikan kualitas dataset. Jika terdapat ulasan yang berbahasa asing, langkah pertama yang kami ambil adalah menerjemahkannya ke dalam Bahasa Indonesia. Kami menggunakan bantuan library mtranslate untuk melakukan proses translasi. Tujuan utamanya adalah menjaga konsistensi bahasa dalam dataset, sehingga seluruh data memiliki format yang seragam. Hasil dari proses ini adalah dataset yang homogen dalam bahasa dan siap untuk tahap berikutnya dalam analisis sentimen.

2.4.2 Cleaning

Pada tahap cleaning, karakter yang tidak penting atau tidak relevan seperti tanda baca, angka, atau simbol khusus dapat dihapus[14]. Hal ini dilakukan untuk membersihkan teks dan mempertahankan hanya informasi yang relevan dalam ulasan.

2.4.3 Normalize

Tahap normalisasi bertujuan untuk menangani kata-kata yang salah ketik atau typo. Ini melibatkan pemrosesan kata-kata yang tidak standar atau tidak sesuai dengan kamus ke bentuk yang tepat agar data lebih konsisten. Kami mencari kata-kata yang mungkin tidak sesuai dengan ejaan yang benar, dan jika ditemukan, kami mengoreksi mereka ke bentuk yang sesuai.

2.4.4 Tokenizing

Langkah tokenizing melibatkan pembagian teks ulasan menjadi unit-unit yang lebih kecil, seperti kata-kata[15]. Tujuan utama dari tokenisasi adalah menciptakan struktur yang lebih terperinci dalam teks, di mana masing-masing kata menjadi unit analisis yang terpisah.

2.4.5 Remove Stopwords

Tahap penghapusan stopwords melibatkan tindakan untuk menghilangkan kata-kata yang tidak penting yang sering muncul dalam teks ulasan, namun kurang memberikan informasi yang signifikan dalam analisis sentimen. Contoh kata-kata ini mungkin termasuk kata-kata seperti "dan," "atau," "yang," dan lainnya[16]. Kata-kata ini sering digunakan dalam bahasa sehari-hari tetapi cenderung tidak memiliki nilai sentimen yang kuat.

2.4.6 Stemming

Stemming adalah proses mengubah kata-kata menjadi bentuk dasar atau kata dasar[17]. Tujuannya adalah untuk mengurangi variasi kata yang sama ke bentuk dasar mereka, sehingga mengurangi dimensi kata yang digunakan dalam analisis. Dengan mengurangi variasi kata, kami meminimalkan dimensi kata yang digunakan dalam analisis. Ini membuat analisis lebih efisien dan membantu dalam mengidentifikasi pola-pola sentimen tanpa harus mempertimbangkan setiap bentuk variasi kata secara terpisah.

2.3 Labeling

Proses labeling dalam penelitian ini melibatkan penggunaan rating yang diberikan oleh pengunjung Istana Yogyakarta pada ulasan mereka sebagai dasar untuk mengkategorikan ulasan ke dalam dua kategori sentimen utama: "positif" dan "negatif." Rating yang biasanya berupa angka 1 hingga 5 akan dikonversi menjadi sentimen positif atau negatif sesuai dengan ambang batas yang ditentukan. Rating yang kurang dari 4 untuk mengkategorikan ulasan sebagai sentimen negatif. Ini berarti bahwa ulasan dengan rating 1 hingga 3 akan dianggap memiliki sentimen negatif terhadap pengalaman mereka di Istana Yogyakarta. Sebaliknya, ulasan dengan rating 4 dan 5 akan dianggap memiliki sentimen positif.

2.5 Pembobotan TF-IDF

Pembobotan TF-IDF adalah metode yang digunakan dalam analisis teks untuk memberikan skor pada kata-kata berdasarkan seberapa pentingnya kata tersebut dalam sebuah dokumen[18]. Dengan TF-IDF dapat membantu mengidentifikasi kata-kata kunci yang paling relevan dalam suatu dokumen. Berikut adalah rumus dasar TF-IDF:

$$TFIDF(i, j) = TF(i, j) \times IDF(i, j) \tag{1}$$

2.6 Naive Bayes

Naive Bayes adalah metode klasifikasi yang menghitung probabilitas kelas sentimen (misalnya, positif, negatif, atau netral) berdasarkan fitur-fitur (kata-kata) dalam ulasan[19]. Algoritma ini menggunakan Teorema Bayes dan mengasumsikan independensi fitur-fitur. Berikut rumus sederhana Naive Bayes:

$$P(A|B) = \frac{P(A) \times P(B|A)}{P(B)} \tag{2}$$

2.7 Logistic Regression

Logistic Regression adalah metode statistik yang digunakan untuk klasifikasi data ke dalam dua atau lebih kategori[11]. Ini bekerja dengan memodelkan probabilitas bahwa data akan masuk ke dalam salah satu dari dua kategori. Berikut rumus sederhana Logistic Regression:

$$P(Y = 1|X) = \frac{1}{1 + e^{-z}} \tag{3}$$

2.8 Support Vector Machine

SVM (Support Vector Machine) adalah algoritma pembelajaran mesin yang digunakan untuk mengklasifikasikan data ke dalam kelompok atau kategori. Tujuannya adalah untuk menemukan garis terbaik atau hyperplane yang memisahkan dua kelompok data dengan margin maksimum[20]. Berikut rumus sederhana SVM:

$$f(x) = w \times x + b \tag{4}$$

2.9 Evaluasi

Dalam tahap evaluasi model, kita akan menggunakan beberapa metrik klasifikasi penting bersama dengan Confusion Matrix dan ROC Curve untuk mengukur kinerja model SVM dalam mengklasifikasikan sentimen ulasan pengunjung Istana Yogyakarta. Confusion Matrix adalah komponen awal dalam evaluasi yang memungkinkan kita untuk melihat seberapa baik model mengklasifikasikan sentimen.

Tabel 1. Confusion Matrix

	Positif	Negatif
Positif	True Positives (TP)	True Negatives (TN)
Negatif	False Positives (FP)	False Negatives (FN)

Terlihat pada Tabel 1 terdiri dari True Positives (TP), True Negatives (TN), False Positives (FP), dan False Negatives (FN), yang membantu kita menghitung Akurasi (Accuracy), Presisi (Precision), Recall (Sensitivitas atau True

Positive Rate), dan F1-score untuk memberikan ukuran komprehensif tentang kinerja model. Selain itu, penggunaan Receiver Operating Characteristic (ROC) Curve yang menggambarkan kemampuan model dalam membedakan kelas positif dan negatif dengan memvariasikan ambang batas keputusan. Area di bawah kurva ROC (AUC-ROC) juga digunakan untuk mengukur kualitas model, dengan nilai yang lebih tinggi menunjukkan kinerja yang lebih baik dalam membedakan kedua kelas.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Hasil pengumpulan data menghasilkan total 1,607 data ulasan pengunjung Istana Yogyakarta. Data ini terdiri dari beberapa kolom, termasuk nama pengunjung, durasi memberikan ulasan, peringkat, dan ulasan dalam bentuk teks yang dapat dilihat pada Tabel 2. Namun, dalam analisis ini kolom yang digunakan hanya dua kolom, yaitu peringkat dan ulasan.

Tabel 2. Sampel Data Ulasan

<i>name</i>	<i>duration</i>	<i>rating</i>	<i>review</i>
Imam Rustadi	4 months ago	5 stars	<i>Great heritage of Java kingdom which still exist till now.</i>
Arini Zulfa	a month ago	5 stars	<i>Beautiful place with good guide</i> ❤️
Nathan Oxley	10 months ago	2 stars	<i>A little boring and old.</i>
Restu Saharan	4 years ago	1 star	<i>Nothing much to seehere. The ticket guy was pretty rude to me.</i>

3.2 Preprocessing

3.2.1 Translate

Proses translate dilakukan untuk mengubah ulasan yang awalnya dalam bahasa asing menjadi bahasa Indonesia. Translasi ini dilakukan menggunakan pustaka mtranslate untuk memastikan bahwa teks yang diterjemahkan tetap akurat dan benar. Pada Tabel 3. menunjukan bahwa data ulasan yang berupa text telah berhasil diterjemahkan ke bahasa Indonesia.

Tabel 3. Proses Translate

<i>review</i>	<i>translate</i>
<i>Great heritage of Java kingdom which still exist till now.</i>	Peninggalan Agung Kerajaan Jawa yang masih eksis hingga saat ini.
<i>Beautiful place with good guide</i> ❤️	Tempat indah dengan panduan bagus ❤️
<i>A little boring and old.</i>	Sedikit membosankan dan tua.
<i>Nothing much to seehere. The ticket guy was pretty rude to me.</i>	Tidak banyak yang bisa dilihat di sini. Petugas tiket itu sangat kasar kepada saya.

3.2.2 Cleaning

Pembersihan data dilakukan untuk menghilangkan karakter yang tidak relevan atau mengganggu dalam ulasan sehingga data menjadi lebih rapi dan mudah untuk dianalisis. Salah satu tindakan pembersihan yang umum adalah menghapus karakter selain huruf. Pada Tabel 4 karakter-karakter selain huruf seperti "love" dalam bentuk emoji telah dihapus dari ulasan.

Tabel 4. Proses Cleaning

<i>translate</i>	<i>cleaning</i>
peninggalan agung kerajaan jawa yang masih eksis hingga saat ini.	peninggalan agung kerajaan jawa yang masih eksis hingga saat ini
tempat indah dengan panduan bagus ❤️	tempat indah dengan panduan bagus
sedikit membosankan dan tua.	sedikit membosankan dan tua
tidak banyak yang bisa dilihat di sini. petugas tiket itu sangat kasar kepada saya.	tidak banyak yang bisa dilihat di sini petugas tiket itu sangat kasar kepada saya

3.2.3 Normalize

Dalam Tabel 5, hasil normalisasi menunjukkan bahwa tidak ada perubahan yang signifikan dalam ulasan. Dalam beberapa kasus, normalisasi mungkin tidak diperlukan jika data ulasan telah ditulis dengan benar dan tidak mengandung penyingkatan atau kesalahan typo yang signifikan. Normalisasi diperlukan jika terdapat variasi penulisan kata-kata yang dapat mengganggu analisis teks, seperti penyingkatan kata "bgt" menjadi "banget" atau perbaikan typo seperti "bagussss" menjadi "bagus".

Tabel 5. Proses *Normalize*

<i>cleaning</i>	<i>normalize</i>
peninggalan agung kerajaan jawa yang masih eksis hingga saat ini	peninggalan agung kerajaan jawa yang masih eksis hingga saat ini
tempat indah dengan panduan bagus	tempat indah dengan panduan bagus
sedikit membosankan dan tua	sedikit membosankan dan tua
tidak banyak yang bisa dilihat di sini petugas tiket itu sangat kasar kepada saya	tidak banyak yang bisa dilihat di sini petugas tiket itu sangat kasar kepada saya

3.2.4 Tokenizing

Dalam Tabel 6, teks ulasan telah dipecah menjadi token-token yang merupakan kata-kata atau frasa-frasa. Proses tokenisasi ini memungkinkan data untuk menjadi lebih terstruktur dan siap untuk pengolahan lebih lanjut dalam analisis teks.

Tabel 6. Proses *Tokenizing*

<i>normalize</i>	<i>tokenize</i>
peninggalan agung kerajaan jawa yang masih eksis hingga saat ini	[peninggalan, agung, kerajaan, jawa, yang, masih, eksis, hingga, saat, ini]
tempat indah dengan panduan bagus	[tempat, indah, dengan, panduan, bagus]
sedikit membosankan dan tua	[sedikit, membosankan, dan, tua]
tidak banyak yang bisa dilihat di sini. petugas tiket itu sangat kasar kepada saya.	[tidak, banyak, yang, bisa, dilihat, di, sini, petugas, tiket, itu, sangat, kasar, kepada, saya]

3.2.5 Remove Stopword

Remove stopwords adalah kata-kata umum yang sering muncul dalam teks tetapi memiliki sedikit makna atau kontribusi dalam analisis teks. Dalam Tabel 7, *stopword* seperti "yang," "saat," "ini," "masih," "dengan," "dan," "di," "kepada," "itu," dan lainnya telah berhasil dihapus. Data ulasan yang telah dibersihkan dengan *stopword* ini dapat membantu fokus pada kata-kata kunci yang lebih penting. Proses ini juga menyebabkan pengurangan jumlah data menjadi 1317. Hal itu dikarenakan pada 260 data keseluruhan kata dalam kalimatnya tidak memiliki makna.

Tabel 7. Proses *Remove Stopword*

<i>tokenize</i>	<i>remove_stopword</i>
[peninggalan, agung, kerajaan, jawa, yang, masih, eksis, hingga, saat, ini]	[peninggalan, agung, kerajaan, jawa, hingga]
[tempat, indah, dengan, panduan, bagus]	[tempat, indah, panduan, bagus]
[sedikit, membosankan, dan, tua]	[sedikit, membosankan, tua]
[tidak banyak yang bisa dilihat di sini. petugas tiket itu sangat kasar kepada saya]	[tidak, banyak, dilihat, sini, petugas, tiket, sangat, kasar]

3.2.6 Stemming

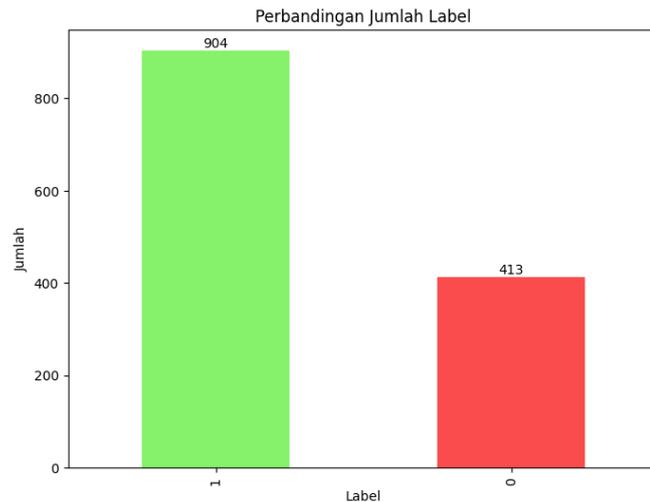
Dalam Tabel 8, terlihat bahwa kata-kata seperti "peninggalan" diubah menjadi "tinggal," "membosankan" menjadi "bosan," dan "dilihat" menjadi "lihat." Proses *stemming* membantu mengurangi variasi kata sehingga kata-kata yang memiliki akar yang sama dianggap sama. Ini mempermudah analisis teks, terutama ketika kita ingin menghitung frekuensi kata atau mengidentifikasi kata kunci dalam data ulasan.

Tabel 8. Proses *Stemming*

<i>remove_stopword</i>	<i>stemming</i>
[peninggalan, agung, kerajaan, jawa, hingga]	[tinggal, agung, raja, jawa, hingga]
[tempat, indah, panduan, bagus]	[tempat, indah, pandu, bagus]
[sedikit, membosankan, tua]	[sedikit, bosan, tua]
[tidak, banyak, dilihat, sini, petugas, tiket, sangat, kasar]	[tidak, banyak, lihat, sini, tugas, tiket, sangat, kasar]

3.3 Labelling

Hasil dari proses *labeling* ini adalah pembagian data ulasan menjadi dua kategori sentimen berdasarkan rating. ulasan-ulasan dapat diberi label sebagai "1" jika ratingnya lebih dari 3, dan label "0" jika ratingnya kurang dari atau sama dengan 3. Pada Gambar 2, terdapat 904 ulasan yang diberi label "1" (sentimen positif) dan 413 ulasan yang diberi label "0" (sentimen negatif). Terdapat perbedaan jumlah yang signifikan antara ulasan yang diberi label "1" (sentimen positif) dan ulasan yang diberi label "0" (sentimen negatif).



Gambar 2. Perbandingan Label

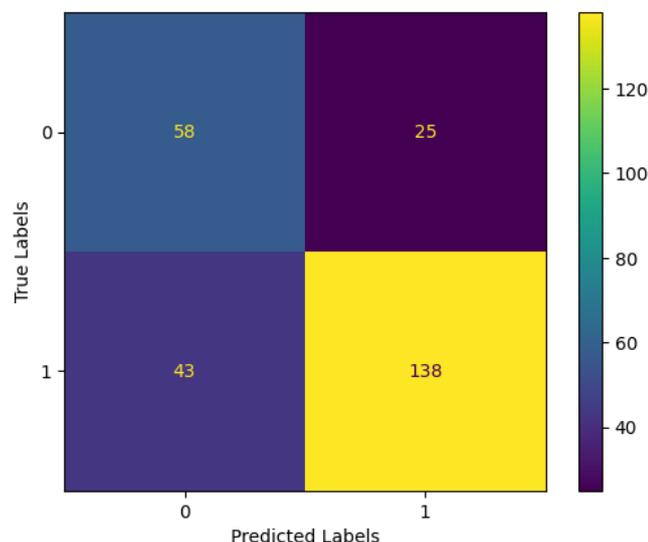
3.4 Naïve Bayes

Hasil evaluasi kinerja model *Naïve Bayes* pada empat percobaan dengan berbagai pembagian data latihan dan uji direkap dalam Tabel 9. Dalam tabel ini, dapat dilihat bahwa Percobaan 3 (80% data latihan, 20% data uji) menghasilkan kinerja terbaik. Dalam percobaan ini, model mencapai akurasi sebesar 74,24% dengan *F1-Score* tertinggi sebesar 80,23%. Pelatihan dengan 1053 ulasan mampu menjaga keseimbangan antara presisi dan recall dengan cukup baik, sehingga mencapai *F1-Score* yang masih terbilang optimal.

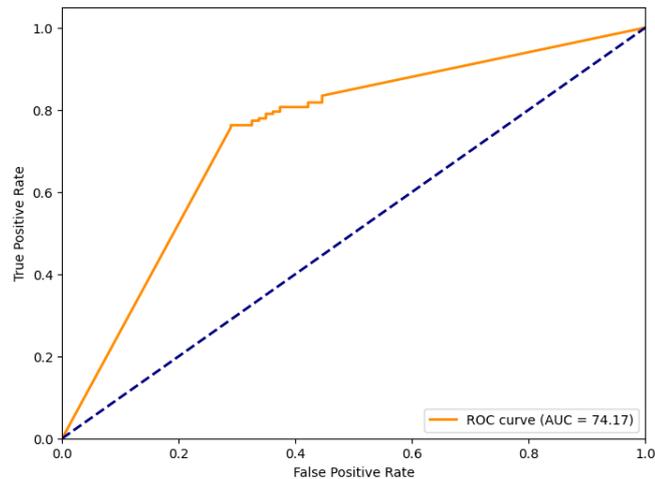
Tabel 9. Pengujian *Naïve Bayes*

Split Data		Metrics				
Train	Test	Accuracy	Precision	Recall	<i>F1-Score</i>	<i>AUC</i>
60%	40%	73,24%	83,79%	75,69%	79,54%	73,51%
70%	30%	73,23%	87,73%	70,96%	78,46%	77,27%
80%	20%	74,24%	84,66%	76,24%	80,23%	74,17%
90%	10%	70,26%	88,24%	65,93%	75,47%	77,37%

Hasil dari konfusi matriks pada split terbaik yang dapat dilihat pada Gambar 3, menunjukkan bahwa model *Naïve Bayes* dalam pengujian sebanyak 264 data uji, berhasil memprediksi dengan benar 58 ulasan dengan sentimen negatif (*True Negative*), dan 138 ulasan dengan sentimen positif (*True Positive*). Namun, terdapat 25 ulasan dengan sentimen negatif yang salah diklasifikasikan sebagai sentimen positif (*False Positive*) dan 43 ulasan dengan sentimen positif yang salah diklasifikasikan sebagai sentimen negatif (*False Negative*). Lalu hasil *ROC curve* yang dapat dilihat pada Gambar 4 menunjukkan hasil yang kurang optimal dalam membagi dua kelas dengan nilai *AUC* 77,37%.



Gambar 3. Confusion Matrix *Naïve Bayes*



Gambar 4. ROC Curve Naïve Bayes

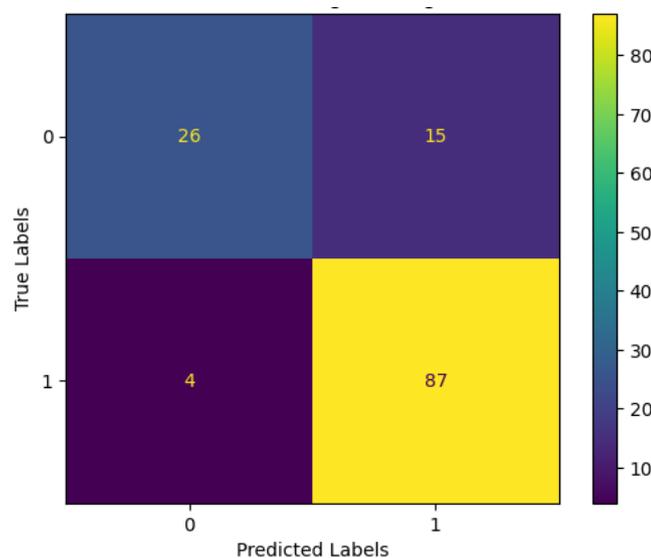
3.5 Logistic Regression

Hasil evaluasi kinerja model *Logistic Regression* pada empat percobaan dengan berbagai pembagian data latihan dan uji direkap dalam Tabel 10. Dalam tabel ini, dapat dilihat bahwa percobaan dengan 90% data latihan dan 10% data uji menghasilkan kinerja terbaik. Dalam percobaan ini, model mencapai akurasi sebesar 85,61% dengan *F1-Score* tertinggi sebesar 90,16%. Pelatihan dengan 1185 ulasan mampu menjaga keseimbangan antara presisi dan recall dengan sangat baik, sehingga mencapai *F1-Score* yang sangat optimal.

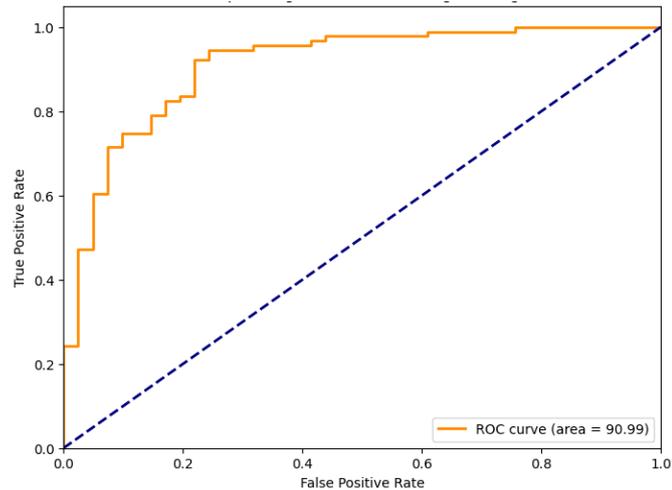
Tabel 10. Pengujian *Logistic Regression*

Split Data		Metrics				
Train	Test	Accuracy	Precision	Recall	F1-Score	AUC
60%	40%	82,93%	82,08%	96,13%	88,55%	90,1%
70%	30%	82,83%	82,48%	95,22%	88,4%	90,86%
80%	20%	83,71%	83,17%	95,58%	88,95%	91,01%
90%	10%	85,61%	85,29%	95,6%	90,16%	90,99%

Hasil dari konfusi matriks pada split terbaik yang dapat dilihat pada Gambar 5, menunjukkan bahwa model *Logistic Regression* dalam pengujian sebanyak 132 data uji, berhasil memprediksi dengan benar 26 ulasan dengan sentimen negatif (*True Negative*), dan 87 ulasan dengan sentimen positif (*True Positive*). Namun, terdapat 4 ulasan dengan sentimen negatif yang salah diklasifikasikan sebagai sentimen positif (*False Positive*) dan 15 ulasan dengan sentimen positif yang salah diklasifikasikan sebagai sentimen negatif (*False Negative*). Lalu hasil *ROC curve* yang dapat dilihat pada Gambar 6 menunjukan hasil yang optimal dalam membagi dua kelas dengan nilai *AUC* 90,99%.



Gambar 5. Confusion Matrix Logistic Regression



Gambar 6. ROC Curve Logistic Regression

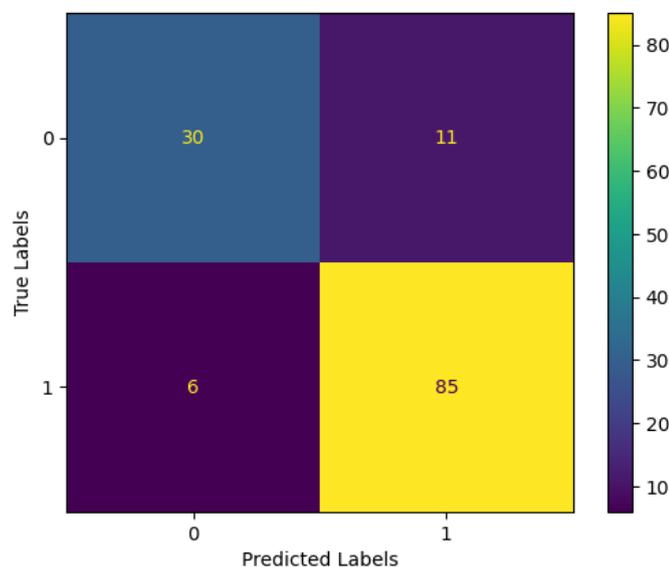
3.6 Support Vector Machine

Hasil evaluasi kinerja model SVM pada empat percobaan dengan berbagai pembagian data latihan dan uji direkap dalam Tabel 11. Sama dengan model logistic regression, model SVM dengan percobaan dengan 90% data latihan dan 10% data uji menghasilkan kinerja terbaik. Dalam percobaan ini, model mencapai akurasi sebesar 87,12% dengan *F1-Score* tertinggi sebesar 90,91%. Pelatihan dengan 1185 ulasan mampu menjaga keseimbangan antara presisi dan recall dengan sangat baik, sehingga mencapai *F1-Score* yang sangat optimal.

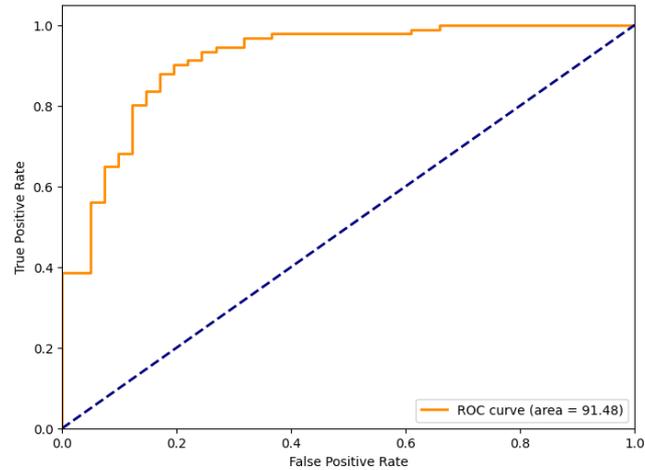
Tabel 11. Pengujian SVM

Split Data		Metrics				
Train	Test	Accuracy	Precision	Recall	<i>F1-Score</i>	<i>AUC</i>
60%	40%	84,44%	85,9%	92,54%	89,1%	90,48%
70%	30%	85,61%	86,69%	93,38%	89,91%	91,03%
80%	20%	85,63%	86,36%	94,48%	90,24%	90,76%
90%	10%	87,12%	88,54%	93,41%	90,91%	91,48%

Hasil dari konfusi matriks pada split terbaik yang dapat dilihat pada Gambar 7, menunjukkan bahwa model SVM dalam pengujian sebanyak 132 data uji, berhasil memprediksi dengan benar 30 ulasan dengan sentimen negatif (*True Negative*), dan 85 ulasan dengan sentimen positif (*True Positive*). Namun, terdapat 6 ulasan dengan sentimen negatif yang salah diklasifikasikan sebagai sentimen positif (*False Positive*) dan 11 ulasan dengan sentimen positif yang salah diklasifikasikan sebagai sentimen negatif (*False Negative*). Lalu hasil *ROC curve* yang dapat dilihat pada Gambar 8 menunjukan hasil yang optimal dalam membagi dua kelas dengan nilai *AUC* 91,48%.



Gambar 7. Confusion Matrix SVM



Gambar 8. ROC Curve SVM

3.7 Perbandingan Model

Hasil dari perbandingan kinerja tiga model, yaitu *Naïve Bayes*, *Logistic Regression*, dan *SVM*, terlihat pada Tabel 12. Perbandingan model ini didasari dari hasil terbaik dari perbandingan split data pada setiap model. *Naïve Bayes*, meskipun memiliki kinerja yang lebih rendah dibandingkan dengan model lainnya, masih memiliki nilai akurasi yang cukup baik dan *F1-Score* sebesar 80,23%. *Logistic Regression* dan *SVM* menunjukkan kinerja yang lebih baik dibandingkan dengan *Naïve Bayes* dalam hal akurasi, presisi, recall, *F1-Score*, dan *AUC*. *Logistic Regression* memiliki nilai tertinggi pada recall, sedangkan *SVM* memiliki nilai tertinggi pada *AUC*, menunjukkan bahwa keduanya baik dalam mengidentifikasi sentimen positif dan negatif. *SVM* memiliki performa yang baik secara keseluruhan, dengan akurasi 87,12% dan nilai *AUC* tertinggi, mencapai 91,48%. Dengan kinerja yang lebih baik dalam mayoritas metrik evaluasi, *SVM* menjadi pilihan yang unggul dalam melakukan analisis sentimen pada dataset yang digunakan. Oleh karena itu, dari data yang diberikan, *SVM* adalah model yang memberikan hasil terbaik dalam memprediksi sentimen ulasan.

Tabel 12. Perbandingan Model

Model	Split Data		Metrics				
	Train	Test	Accuracy	Precision	Recall	F1-Score	AUC
<i>Naïve Bayes</i>	80%	20%	74,24%	84,66%	76,24%	80,23%	74,17%
<i>Logistic Regression</i>	90%	10%	85,61%	85,29%	95,6%	90,16%	90,99%
<i>SVM</i>	90%	10%	87,12%	88,54%	93,41%	90,91%	91,48%

3.8 Perbandingan Klasifikasi

Tabel 13 menyajikan contoh hasil klasifikasi sentimen untuk empat ulasan dengan menggunakan tiga model yang berbeda, yaitu *Naïve Bayes*, *Logistic Regression*, dan *SVM*, serta label sentimen aktualnya. Secara keseluruhan, hasil ini memberikan keyakinan bahwa ketiga model yang digunakan dalam analisis sentimen mampu mengklasifikasikan ulasan dengan benar sesuai dengan sentimen yang diungkapkan dalam ulasan tersebut. Meskipun model *Naïve Bayes* memiliki performa yang terendah dari ketiga model yang dibandingkan, tetap dapat mengklasifikasikan ulasan dengan akurasi yang baik dan dengan benar, seperti yang dilakukan oleh *SVM* dan *Logistic Regression*.

Tabel 13. Percobaan Klasifikasi

review	Aktual	<i>Naïve Bayes</i>	<i>Logistic Regression</i>	<i>SVM</i>
<i>Great heritage of Java kingdom which still exist till now.</i>	1	1	1	1
<i>Beautiful place with good guide</i> ❤️.	1	1	1	1
<i>A little boring and old.</i>	0	0	0	0
<i>Nothing much to seehere. The ticket guy was pretty rude to me.</i>	0	0	0	0

4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah disampaikan dalam analisis sentimen terhadap ulasan pengunjung Istana Yogyakarta, dapat ditarik beberapa kesimpulan penting. Pertama, proses pengumpulan data menghasilkan total 1,607 data ulasan pengunjung Istana Yogyakarta. Data ulasan telah melalui serangkaian tahap pengolahan data, termasuk *translate*, *cleaning*, *normalize*, *tokenize*, penghapusan *stopword*, dan *stemming*. Dalam tahap *preprocessing*, khususnya dalam

proses penghapusan *stopword*, terjadi pengurangan jumlah data menjadi 1,317. Hal ini terjadi karena sebanyak 260 data ulasan mengandung kata-kata yang secara individual tidak memiliki makna yang signifikan dalam analisis sentimen. Kesimpulan kedua yang dapat diambil adalah pemilihan pembagian data yang tepat sangat berpengaruh terhadap kinerja model klasifikasi. Pembagian data 80% data latih dan 20% data uji, misalnya, menghasilkan kinerja terbaik untuk model *Naïve Bayes*, dengan akurasi mencapai 74.24% dan *F1-Score* 80.23%. Namun, model *Logistic Regression* dan *SVM* mencapai akurasi yang lebih tinggi, yakni 85.61% dan 87.12%, saat data dibagi dengan perbandingan yang sama. Dalam konteks pemilihan model terbaik, hasil menunjukkan bahwa *SVM* adalah model terbaik dalam analisis sentimen terhadap ulasan pengunjung Istana Yogyakarta. Model *SVM* mampu mencapai akurasi sebesar 87.12% dan *F1-Score* tertinggi, yaitu 90.91%. Ini mengindikasikan bahwa *SVM* memiliki kemampuan yang lebih baik dalam mengklasifikasikan sentimen ulasan dibandingkan dengan model-model lainnya. Meskipun begitu, dalam pengujian klasifikasi empat ulasan, hasil menunjukan bahwa ketiga model yang digunakan dalam analisis sentimen mampu mengklasifikasikan ulasan dengan benar sesuai dengan sentimen yang diungkapkan dalam ulasan tersebut.

REFERENCES

- [1] Dinas Kebudayaan Kota Yogyakarta, "Kawasan Kraton." Diakses: 8 November 2023. [Daring]. Tersedia pada: <https://kebudayaan.jogjakota.go.id/page/index/kawasan-kraton>
- [2] W. Ningsih, "Sejarah Keraton Yogyakarta." Diakses: 1 November 2023. [Daring]. Tersedia pada: <https://www.kompas.com/stori/read/2023/07/20/090000479/sejarah-keraton-yogyakarta?page=all>
- [3] H. Herlawati, R. T. Handayanto, P. D. Atika, F. N. Khasanah, A. Y. P. Yusuf, dan D. Y. Septia, "Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 2, hlm. 153–163, Nov 2021, doi: 10.31603/komtika.v5i2.6280.
- [4] W. Khofifah, D. N. Rahayu, dan A. M. Yusuf, "Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps," *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 16, no. 4, hlm. 28–38, Jan 2022, doi: 10.35969/interkom.v16i4.192.
- [5] R. Situmorang, U. Tamyis, dan L. Muni, "ANALISIS SENTIMEN DESTINASI WISATA DI JAWABARAT PADA TWITTER MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER," *JURNAL SISTEM INFORMASI DAN TEKNIK KOMPUTER*, vol. 8, hlm. 339–442, 2023.
- [6] A. A. Arifiyanti, M. F. Pandji, dan B. Utomo, "Analisis Sentimen Ulasan Pengunjung Objek Wisata Gunung Bromo pada Situs Tripadvisor," *Explore: Jurnal Sistem Informasi dan Telematika*, vol. 13, no. 1, hlm. 32, Jun 2022, doi: 10.36448/jsit.v13i1.2539.
- [7] G. Fikri Baihaqi, D. E. Ratnawati, dan B. T. Hanggara, "Analisis Sentimen Wisata Alun-Alun Kota Batu menggunakan Algoritma Support Vector Machine," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 12, hlm. 6010–6018, 2022, [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id>
- [8] A. Erfina, "ANALISIS SENTIMEN PERGURUAN TINGGI TERMEWAH DI INDONESIA MENURUT ULASAN GOOGLE MAPS MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM)," *Jurnal Manajemen Informatika & Sistem Informasi (MISI)*, vol. 5, no. 1, 2022, doi: 10.36595/misi.v5i1.
- [9] A. Novantika, "Analisis Sentimen Ulasan Pengguna Aplikasi Video Conference Google Meet menggunakan Metode SVM dan Logistic Regression," *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 5, hlm. 808–813, 2022, [Daring]. Tersedia pada: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [10] Tania Puspa Rahayu Sanjaya, Ahmad Fauzi, dan Anis Fitri Nur Masruriyah, "Analisis sentimen ulasan pada e-commerce shopee menggunakan algoritma naive bayes dan support vector machine," *INFOTECH : Jurnal Informatika & Teknologi*, vol. 4, no. 1, hlm. 16–26, Jun 2023, doi: 10.37373/infotech.v4i1.422.
- [11] I. Rahmawati dan T. R. Fitriani, "Analisis Sentimen Menggunakan Algoritma Logistic Regression Pada Penerbangan Lion Air berdasarkan Ulasan Pengguna Platform Online," 2023.
- [12] A. Muhammadin dan I. A. Sobari, "Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma SVM Dan NBC," *Jurnal Rekayasa Perangkat Lunak*, vol. 2, no. 2, 2021, [Daring]. Tersedia pada: <http://jurnal.bsi.ac.id/index.php/reputasi>
- [13] M. Kopravi dan W. S. Putra, "Implementasi Web Scraping pada Google Cendekia sebagai Sarana Profiling Penelitian Dosen," *Jurnal Ilmu Pengetahuan dan Teknologi*, hlm. 59–72, 2023, doi: 10.30738/st.
- [14] M. Khoirul, U. Hayati, dan O. Nurdiawan, "ANALISIS SENTIMEN APLIKASI BRIMO PADA ULASAN PENGGUNA DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES," 2023.
- [15] J. A. Aryadi, Y. A. Aidil Basith, M. Munawir, dan D. A. Rimadhani Agustini, "Analisis Data Review Hotel di Google Maps Melalui Text Mining (Studi Kasus : Kabupaten Bandung)," *JIKO (Jurnal Informatika dan Komputer)*, vol. 7, no. 2, hlm. 312, Sep 2023, doi: 10.26798/jiko.v7i2.938.
- [16] L. B. Ilmawan dan M. A. Mude, "Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, hlm. 154–161, Agu 2020, doi: 10.33096/ilkom.v12i2.597.154-161.
- [17] A. M. Sidabutar, I. M. Sarkis, dan E. Rajagukguk, "Analisis Sentimen Cuitan di Terhadap Kawasan Wisata Danau Toba Dengan Metode Naïve Bayes," 2023. [Daring]. Tersedia pada: <http://ojs.fikom-methodist.net/index.php/methodika>
- [18] R. Fitriansyah, "Pemanfaatan Vector Space Model Algoritma Nazief Andriani Pembobotan TfIdf Pada Prototipe Klasifikasi Teks Bahasa Indonesia," *Jurnal Teknologi Informasi*, vol. 2, no. 1, hlm. 37–47, 2021.
- [19] B. Hakim, "Analisa Sentimen Data Text Preprocessing Pada Data Mining Dengan Menggunakan Machine Learning," *JBASE - Journal of Business and Audit Information Systems*, vol. 4, no. 2, Agu 2021, doi: 10.30813/jbase.v4i2.3000.
- [20] M. R. Nadhif, D. Wisnu Brata, dan B. Rahayudi, "Analisis Sentimen Data Ulasan Pengguna Aplikasi TIX ID di Indonesia pada Google Play Store menggunakan Support Vector Machine," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 8, hlm. 3932–3937, 2022, [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id>