



INNOVATIVE: Journal Of Social Science Research

Volume 4 Nomor 6 Tahun 2024 Page 2351-2359

E-ISSN 2807-4238 and P-ISSN 2807-4246

Website: <https://j-innovative.org/index.php/Innovative>

## Analisis Sentimen Terhadap Pariwisata di Kabupaten Grobogan Berbasis Orange Menggunakan *Naive Bayes*

Mochamad Amry Assiva

Institut Teknologi dan Bisnis Muhammadiyah Grobogan

Email: [amryassiva@itbmug.ac.id](mailto:amryassiva@itbmug.ac.id)

### Abstrak

Industri pariwisata merupakan sektor penting dalam perekonomian, termasuk di Kabupaten Grobogan, yang memiliki berbagai destinasi wisata. Ulasan wisatawan dapat memberikan wawasan berharga mengenai pengalaman mereka, yang dapat digunakan oleh pengelola untuk meningkatkan kualitas pelayanan. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan wisatawan di Kabupaten Grobogan menggunakan model *Naive Bayes* yang diterapkan pada platform Orange. Data ulasan yang dikumpulkan dari Google Reviews dianalisis untuk menentukan apakah sentimennya positif, negatif, atau netral. Hasil penelitian menunjukkan bahwa mayoritas ulasan wisatawan memiliki sentimen positif, meskipun beberapa keluhan mengenai fasilitas dan infrastruktur masih perlu perbaikan. Akurasi model *Naive Bayes* yang digunakan dalam penelitian ini adalah 85%, yang menunjukkan bahwa model ini efektif dalam melakukan analisis sentimen dalam konteks pariwisata.

Kata Kunci: *Analisis Sentimen, Naive Bayes, Orange, Pariwisata, Kabupaten Grobogan, Google Reviews*

## Abstract

The tourism industry is a significant sector of the economy, including in Grobogan Regency, which has a variety of tourist destinations. Tourist reviews can provide valuable insights into their experiences, which can be used by managers to improve service quality. This study aims to perform sentiment analysis on tourist reviews in Grobogan Regency using the Naive Bayes model implemented on the Orange platform. The review data collected from Google Reviews were analyzed to determine whether the sentiment was positive, negative, or neutral. The results show that the majority of the tourist reviews have a positive sentiment, though some complaints regarding facilities and infrastructure still need improvement. The accuracy of the Naive Bayes model used in this study is 85%, indicating that this model is effective in sentiment analysis within the tourism context.

*Keywords: Sentiment Analysis, Naive Bayes, Orange, Tourism, Grobogan Regency, Google Reviews*

## PENDAHULUAN

Pariwisata merupakan salah satu sektor penting yang dapat mendorong perekonomian daerah, termasuk di Kabupaten Grobogan yang memiliki berbagai destinasi wisata menarik. Dalam mengelola industri ini, penting bagi pengelola untuk memahami pandangan dan persepsi wisatawan mengenai kualitas layanan dan fasilitas yang ada. Salah satu cara efektif untuk mendapatkan feedback adalah melalui analisis ulasan wisatawan yang diposting di platform daring, seperti Google Review. Ulasan ini mencerminkan pengalaman nyata pengunjung terhadap objek wisata yang dikunjungi, memberikan gambaran tentang kelebihan dan kekurangan yang ada, serta dapat dijadikan bahan evaluasi untuk meningkatkan kualitas destinasi wisata. Namun, dengan jumlah ulasan yang sangat banyak, dibutuhkan teknologi yang dapat memproses dan menganalisis data secara cepat dan akurat. Teknologi informasi memiliki peran yang sangat penting dalam meningkatkan kualitas layanan (Assiva, Tanuji dan Ajipangestu, 2024).

Salah satu metode yang banyak digunakan dalam analisis ulasan adalah analisis sentimen, yaitu proses untuk menentukan perasaan atau sikap dalam sebuah teks, apakah itu positif, negatif, atau netral. Algoritma Naive Bayes merupakan salah satu teknik yang efektif dalam melakukan analisis sentimen pada data teks. Dengan menggunakan Naive Bayes, analisis sentimen dapat dilakukan secara otomatis untuk mengkategorikan ulasan wisatawan berdasarkan sentimen mereka. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan wisatawan di Kabupaten Grobogan menggunakan model Naive Bayes yang diimplementasikan dalam platform analitik data Orange, guna

memberikan wawasan yang berguna bagi pengelola pariwisata dalam meningkatkan kualitas layanan dan fasilitas yang ada.

## METODE PENELITIAN

### Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari Google Review untuk beberapa objek wisata yang ada di Kabupaten Grobogan. Proses pengumpulan data dilakukan melalui teknik *web scraping* menggunakan Python dengan pustaka BeautifulSoup dan Selenium. Ulasan yang terkumpul terdiri dari teks ulasan, rating (1-5 bintang), serta informasi terkait lokasi dan waktu ulasan. Data yang terkumpul terdiri dari 1000 ulasan yang dipilih secara acak.

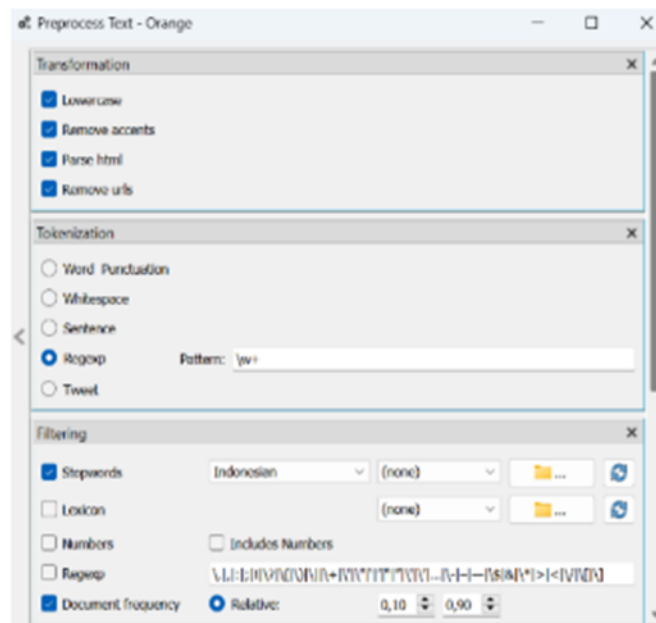


Gambar 1. Ilustrasi proses pengambilan data

### Preprocessing Data

Data ulasan yang dikumpulkan perlu dipersiapkan sebelum dilakukan analisis sentimen. Langkah-langkah preprocessing meliputi:

1. Tokenisasi: Memecah teks menjadi unit-unit kata.
2. Penghapusan Stopwords: Menghapus kata-kata umum yang tidak memberikan informasi penting dalam analisis sentimen (misalnya "dan", "atau", "adalah").
3. *Convert Case*: Mengubah semua karakter menjadi huruf kecil (lowercase) atau huruf besar (*uppercase*). Langkah ini penting dalam analisis teks untuk memastikan keseragaman dan menghindari perbedaan antara kata seperti "Apple" dan "apple", yang seharusnya dianggap sebagai kata yang sama. Biasanya, konversi ke huruf kecil digunakan untuk menyederhanakan analisis lebih lanjut, seperti tokenisasi dan analisis sentimen.
4. Pembersihan Teks: Menghapus karakter non-alfabet, angka, dan tanda baca yang tidak relevan.



Gambar 2. Preprocessing Data

### Klasifikasi Sentimen Menggunakan *Naive Bayes*

Setelah proses preprocessing, data siap untuk dianalisis. Dalam penelitian ini, algoritma Naive Bayes digunakan untuk mengklasifikasikan sentimen ulasan menjadi tiga kategori: positif, negatif, dan netral. Proses klasifikasi dilakukan dengan membagi dataset menjadi dua bagian: 80% digunakan untuk pelatihan model, dan 20% sisanya digunakan untuk pengujian. Model Naive Bayes dilatih menggunakan fitur kata-kata yang terdapat dalam ulasan.

### Evaluasi Model

Untuk menilai kinerja model Naive Bayes, digunakan beberapa metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Evaluasi dilakukan dengan menggunakan confusion matrix untuk mengukur sejauh mana model dapat mengklasifikasikan sentimen dengan benar dan menghindari kesalahan dalam klasifikasi.

## HASIL DAN PEMBAHASAN

Pada bagian ini, akan dibahas mengenai hasil dari analisis sentimen yang dilakukan menggunakan model Naive Bayes pada data ulasan wisatawan tentang destinasi wisata di Kabupaten Grobogan. Penelitian ini mengidentifikasi sentimen positif, negatif, dan netral dalam ulasan-ulasan yang dikumpulkan dari platform Google Review. Pembahasan akan mencakup hasil akurasi model, distribusi sentimen, serta wawasan yang diperoleh dari analisis sentimen tersebut.

## Pengumpulan Data dan Persiapan Data

Data yang digunakan dalam penelitian ini diambil dari platform Google Review yang berisi ulasan wisatawan tentang beberapa tempat wisata di Kabupaten Grobogan. Sebanyak 1.000 ulasan yang mengandung berbagai sentimen mengenai kualitas fasilitas, pelayanan, dan suasana tempat wisata, dikumpulkan untuk dianalisis. Ulasan-ulasan ini kemudian dibersihkan dengan menghapus kata-kata yang tidak relevan, seperti stop words, dan dilakukan Convert Case mengubah semua karakter menjadi huruf kecil.

Setelah proses pembersihan dan persiapan data, ulasan dibagi menjadi tiga kategori sentimen: positif, negatif, dan netral. Kategori ini digunakan sebagai label untuk pelatihan model klasifikasi Naive Bayes. Data yang digunakan dalam pelatihan model terbagi menjadi 80% data pelatihan dan 20% data pengujian.

## Implementasi Model *Naive Bayes*

Model *Naive Bayes* diimplementasikan menggunakan platform analitik Orange, yang menyediakan pustaka pemrosesan teks dan widget untuk klasifikasi teks. Model ini dilatih dengan menggunakan data pelatihan yang sudah diberi label sentimen positif, negatif, dan netral. Untuk merepresentasikan fitur, setiap ulasan diubah menjadi vektor *Bag of Words*, yang mencatat kemunculan kata-kata yang muncul dalam ulasan wisatawan.

Setelah model Naive Bayes dilatih, model ini diuji dengan data pengujian yang belum pernah dilihat sebelumnya. Proses evaluasi dilakukan dengan menghitung metrik akurasi, precision, recall, dan F1-score untuk menilai performa model dalam mengklasifikasikan sentimen ulasan secara akurat.

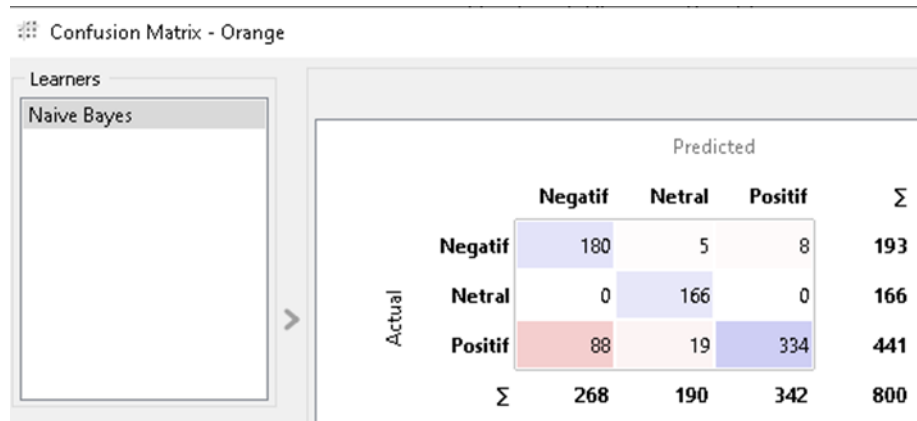


Gambar 3. Skenario Penelitian

## Hasil Akurasi Model

### *Confusion Matrix*

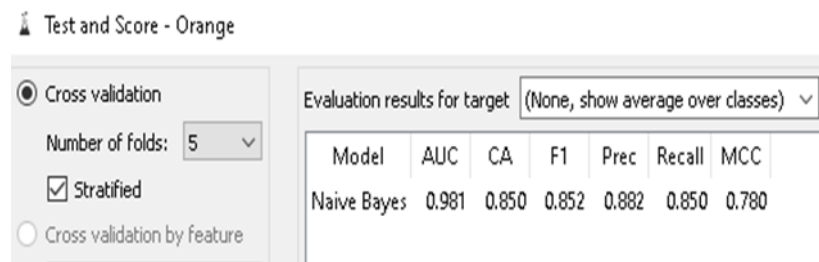
*Confusion matrix* adalah alat yang digunakan untuk menilai performa model klasifikasi dengan menampilkan jumlah prediksi yang benar dan salah untuk setiap kategori kelas. Berikut adalah *confusion matrix* yang dihasilkan:



|        |         | Predicted |        |         |     |
|--------|---------|-----------|--------|---------|-----|
|        |         | Negatif   | Netral | Positif | Σ   |
| Actual | Negatif | 180       | 5      | 8       | 193 |
|        | Netral  | 0         | 166    | 0       | 166 |
|        | Positif | 88        | 19     | 334     | 441 |
| Σ      |         | 268       | 190    | 342     | 800 |

Gambar 4. *Confusion Matrix*

### Metrik Evaluasi



| Evaluation results for target (None, show average over classes) |       |       |       |       |        |       |
|-----------------------------------------------------------------|-------|-------|-------|-------|--------|-------|
| Model                                                           | AUC   | CA    | F1    | Prec  | Recall | MCC   |
| Naive Bayes                                                     | 0.981 | 0.850 | 0.852 | 0.882 | 0.850  | 0.780 |

Gambar 5. *Test and Score*

#### a. *Precision*

Nilai 0,882 pada precision di widget *Test and Score* Orange berarti bahwa dari semua prediksi positif yang dibuat oleh model, 88,2% di antaranya benar (positif yang sebenarnya). Ini menunjukkan tingkat ketepatan model dalam mengidentifikasi kasus positif dengan benar.

#### b. *Recall*

Nilai 0,85 pada recall di widget *Test and Score* Orange berarti bahwa dari semua data yang sebenarnya positif, model berhasil mengidentifikasi 85% di antaranya dengan benar.

c. F1-score

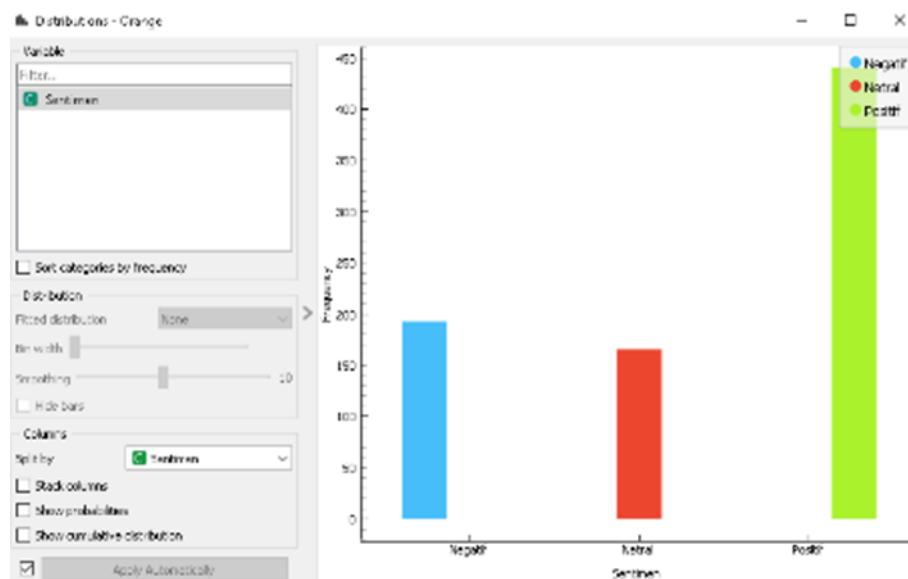
Nilai 0,852 pada F1-score di widget *Test and Score* Orange menunjukkan keseimbangan antara precision dan recall dari model Anda, dengan skor sekitar 85,2%.

Akurasi model Naive Bayes yang digunakan dalam penelitian ini mencapai 85%, Hal ini menunjukkan bahwa model ini cukup efektif dalam mengidentifikasi sentimen yang terkandung dalam ulasan wisatawan. Presisi, recall, dan F1-score masing-masing menunjukkan nilai yang baik untuk ketiga kategori (positif, negatif, netral).

Hasil ini menunjukkan bahwa meskipun model memiliki akurasi yang baik, masih ada beberapa kesalahan klasifikasi, terutama pada sentimen negatif. Namun, hasil yang diperoleh cukup baik untuk digunakan dalam analisis sentimen terhadap ulasan wisatawan.

#### Distribusi Sentimen pada Ulasan Wisatawan

Setelah model Naive Bayes dilatih dan diuji, hasil klasifikasi sentimen dari data ulasan diperoleh sebagai berikut:



Gambar 6. Distribusi Sentimen

a. Sentimen Positif

55% dari total ulasan yang dianalisis dikategorikan sebagai sentimen positif. Ulasan positif umumnya berfokus pada aspek-aspek seperti kualitas pelayanan, kebersihan tempat wisata, dan keramahan petugas serta masyarakat setempat.

b. Sentimen Negatif

25% ulasan memiliki sentimen negatif. Ulasan ini mencakup keluhan terkait dengan infrastruktur, fasilitas yang tidak memadai, harga yang dianggap tidak sebanding dengan fasilitas yang ditawarkan, serta masalah kebersihan di beberapa tempat wisata.

c. Sentimen Netral

20% ulasan dikategorikan sebagai netral. Ulasan netral umumnya memberikan penilaian yang cukup objektif tanpa terlalu menyoroti kekurangan atau kelebihan yang signifikan.

### Pembahasan Hasil

Dari hasil analisis sentimen, dapat disimpulkan bahwa sebagian besar ulasan mengenai destinasi wisata di Kabupaten Grobogan bersifat positif. Hal ini menunjukkan bahwa banyak wisatawan merasa puas dengan kualitas pelayanan, kebersihan, serta keramahan masyarakat setempat. Ulasan positif umumnya menyebutkan pengalaman yang menyenangkan, pemandangan alam yang indah, dan kemudahan akses ke tempat wisata.

Namun, meskipun terdapat banyak ulasan positif, masih terdapat beberapa ulasan negatif yang perlu menjadi perhatian bagi pengelola destinasi wisata. Keluhan yang paling sering muncul terkait dengan infrastruktur yang kurang memadai, seperti jalan yang rusak atau sulit diakses, serta kurangnya fasilitas pendukung seperti toilet yang bersih dan tempat parkir yang luas. Masalah kebersihan juga menjadi perhatian, dengan beberapa wisatawan mengeluhkan kurangnya pengelolaan sampah di beberapa lokasi wisata.

Selain itu, terdapat juga beberapa ulasan netral yang memberikan penilaian objektif terhadap pengalaman mereka, tetapi tidak memberikan kesan yang sangat kuat, baik positif maupun negatif. Ulasan seperti ini memberikan gambaran bahwa meskipun tempat wisata tersebut cukup baik, masih ada beberapa hal yang bisa diperbaiki agar dapat memberikan pengalaman yang lebih memuaskan bagi wisatawan.

### Implikasi Hasil Analisis Sentimen

Hasil analisis ini memberikan wawasan yang berguna bagi pengelola destinasi wisata di Kabupaten Grobogan. Dengan mengetahui aspek-aspek yang mendapat pujian serta keluhan, pengelola dapat fokus pada perbaikan infrastruktur dan fasilitas yang dirasa kurang memadai, seperti memperbaiki jalan akses ke destinasi wisata, menambah fasilitas toilet yang bersih, serta meningkatkan pengelolaan sampah.



Di sisi lain, aspek-aspek yang mendapat ulasan positif, seperti keramahan petugas dan kebersihan tempat wisata, dapat terus dipertahankan dan dikembangkan untuk meningkatkan kepuasan wisatawan lebih lanjut. Dengan melakukan perbaikan berdasarkan analisis sentimen ini, diharapkan pengalaman wisatawan di Kabupaten Grobogan dapat semakin baik, yang pada gilirannya akan meningkatkan kunjungan wisatawan dan mendukung pertumbuhan sektor pariwisata di daerah tersebut.

## SIMPULAN

Berdasarkan hasil analisis sentimen terhadap ulasan wisatawan menggunakan model Naive Bayes, dapat disimpulkan bahwa sebagian besar ulasan mengenai destinasi wisata di Kabupaten Grobogan bersifat positif, dengan banyak wisatawan yang memuji kualitas pelayanan, kebersihan, dan keramahan masyarakat. Meskipun demikian, terdapat sejumlah keluhan yang berkaitan dengan infrastruktur, seperti jalan yang rusak dan kurangnya fasilitas pendukung, seperti toilet yang bersih dan pengelolaan sampah yang belum optimal. Model *Naive Bayes* menunjukkan akurasi sebesar 85%, yang menunjukkan efektivitasnya dalam mengklasifikasikan sentimen dalam data ulasan. Berdasarkan temuan ini, disarankan agar pengelola destinasi wisata di Kabupaten Grobogan fokus pada perbaikan infrastruktur dan fasilitas yang kurang memadai, serta terus mempertahankan aspek-aspek positif yang telah diapresiasi oleh wisatawan. Perbaikan ini diharapkan dapat meningkatkan kepuasan wisatawan dan mendukung pertumbuhan sektor pariwisata di daerah tersebut.

## DAFTAR PUSTAKA

- Assiva, M. A., Tanuji, H., & Ajipangestu, D. B. (2024). PERANCANGAN DAN PELATIHAN APLIKASI SISTEM AKADEMIK SMK NEGERI 2 PURWODADI (SIKUAD). Fokus ABDIMAS, 3(1), 30-34.
- Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management*, 21(1), 35-40.
- Chen, H., Zhang, X., & Liu, Y. (2019). Sentiment analysis of tourism reviews based on machine learning. *Tourism Management*, 70, 35-42.
- Demsar, J., Curk, T., Erjavec, A., et al. (2013). Orange: Data mining toolbox in Python. *Journal of Machine Learning Research*, 14(1), 2349-2353.
- Manning, C. D., Raghavan, P., & Schütze, H. (2009). *Introduction to Information Retrieval*. Cambridge University Press.
- McCallum, A., & Nigam, K. (1998). A comparison of event models for naive Bayes text

classification. AAAI-98 Workshop on Learning for Text Categorization, 41-48.

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.