



## Analisis Sentimen Ulasan pada Google Review di Sebuah Penginapan Menggunakan Algoritma Naïve Bayes (Studi Kasus : Grand Jatra Hotel Pekanbaru)

Muhammad Azlan<sup>1\*</sup>, Elvi Rahmi<sup>2</sup>

<sup>1-2</sup> Politeknik Negeri Bengkalis, Indonesia

Alamat: Jl. Bathin Alam, Desa Sungai Alam, Kecamatan Bengkalis Kabupaten Bengkalis,  
Provinsi Riau, Indonesia Kode Pos 28712

Korespondensi penulis: [azlan14082003@gmail.com](mailto:azlan14082003@gmail.com)

**Abstract.** *This study aims to analyze the sentiment of customer reviews of the Grand Jatra Hotel Pekanbaru on the Google Review platform using the Naïve Bayes algorithm. Social media and online review platforms are increasingly becoming the primary source of information for potential customers in making purchasing decisions, particularly in the hospitality sector. Therefore, sentiment analysis of customer reviews is crucial for understanding consumer perceptions and providing strategic input for hotels in improving service quality. The research data was collected using web scraping techniques to obtain publicly available customer reviews. The obtained data was then processed through text preprocessing stages including case folding, tokenizing, normalization, stopword removal, and stemming. The Term Frequency-Inverse Document Frequency (TF-IDF) method was then used to weight each word, so that more relevant words have a greater influence in the classification process. The sentiment classification process was carried out into two main categories, namely positive and negative. The Naïve Bayes model was trained using training data and then tested with test data to measure the algorithm's performance in classifying sentiment. The evaluation results show that the model built is able to achieve an accuracy level of 98%, with a precision value of 97% and a recall of 100% in the positive class, and 92% in the negative class. These findings confirm that the Naïve Bayes algorithm can be effectively used in analyzing customer sentiment towards hotel services and facilities. Practically, the results of this study are expected to provide insight for the management of Grand Jatra Hotel Pekanbaru in understanding customer perceptions, identifying service strengths and weaknesses, and formulating more targeted marketing strategies. In addition, this study can also be a reference for the development of similar studies in the hotel industry and other service sectors.*

**Keywords:** *Sentiment Analysis, Google Reviews, Naïve Bayes, Hospitality, TF-IDF.*

**Abstrak.** Penelitian ini bertujuan untuk menganalisis sentimen ulasan pelanggan terhadap Grand Jatra Hotel Pekanbaru yang terdapat pada platform Google Review dengan menggunakan algoritma Naïve Bayes. Media sosial dan platform ulasan daring semakin menjadi sumber utama informasi bagi calon pelanggan dalam menentukan keputusan pembelian, khususnya dalam bidang perhotelan. Oleh karena itu, analisis sentimen terhadap ulasan pelanggan menjadi penting untuk memahami persepsi konsumen serta memberikan masukan strategis bagi pihak hotel dalam meningkatkan kualitas layanan. Data penelitian dikumpulkan menggunakan teknik web scraping untuk memperoleh ulasan pelanggan yang tersedia secara publik. Selanjutnya, data yang diperoleh diproses melalui tahapan text preprocessing yang meliputi case folding, tokenizing, normalisasi, stopword removal, serta stemming. Setelah itu, metode Term Frequency-Inverse Document Frequency (TF-IDF) digunakan untuk memberikan bobot pada setiap kata, sehingga kata-kata yang lebih relevan memiliki pengaruh lebih besar dalam proses klasifikasi. Proses klasifikasi sentimen dilakukan ke dalam dua kategori utama, yaitu positif dan negatif. Model Naïve Bayes dilatih menggunakan data latih dan kemudian diuji dengan data uji untuk mengukur performa algoritma dalam mengklasifikasikan sentimen. Hasil evaluasi menunjukkan bahwa model yang dibangun mampu mencapai tingkat akurasi sebesar 98%, dengan nilai presisi 97% dan recall 100% pada kelas positif, serta 92% pada kelas negatif. Temuan ini menegaskan bahwa algoritma Naïve Bayes dapat secara efektif digunakan dalam menganalisis sentimen pelanggan terhadap layanan dan fasilitas hotel. Secara praktis, hasil penelitian ini diharapkan dapat memberikan wawasan bagi pihak manajemen Grand Jatra Hotel Pekanbaru dalam memahami persepsi pelanggan, mengidentifikasi kelebihan maupun kekurangan layanan, serta merumuskan strategi pemasaran yang lebih tepat sasaran. Selain itu, penelitian ini juga dapat menjadi referensi bagi pengembangan studi serupa di bidang industri perhotelan maupun sektor jasa lainnya.

**Kata kunci:** Analisis Sentimen, Google Review, Naïve Bayes, Perhotelan, TF-IDF.

## 1. LATAR BELAKANG

Digitalisasi data yang ada sekarang berkembang pesat di semua lini, tidak terkecuali dalam dunia pariwisata. Masyarakat sekarang ini dimudahkan dalam memberikan suatu ulasan tentang sesuatu hal sehingga orang lain akan sangat terbantu yaitu para pelancong untuk menemukan tempat yang mereka inginkan hanya dengan koneksi internet saja. Para pelancong sangat dimudahkan untuk mendapatkan gambaran suatu tempat penginapan hanya dengan membaca ulasan yang banyak beredar baik di jejaring sosial maupun di website. Pemberian hak akses kepada pelanggan yang mudah dan nyaman dalam menyampaikan opini, kritik dan saran yang nantinya digunakan terhadap suatu layanan dalam pengukuran tingkat kepuasan pelanggan. *Google Review* merupakan salah satu jenis situs pemesanan hotel yang menawarkan berbagai fitur untuk memudahkan pengunjung dalam menentukan hotel yang akan dipilih. Salah satu fitur yang ditawarkan adalah adanya ulasan yang menampilkan berbagai komentar pengunjung tentang suatu hotel. Namun, semakin banyak komentar atau ulasan tentang suatu hotel maka pengunjung memerlukan waktu yang lama untuk membaca dan memilih hotel yang diinginkan. Algoritma *Naive Bayes* adalah metode klasifikasi yang sederhana namun efektif berdasarkan *Teorema Bayes* dengan asumsi kuat (*naif*) bahwa semua fitur bersifat independen. Algoritma ini sering digunakan dalam pengolahan teks, termasuk analisis sentimen dari ulasan, karena kinerjanya yang baik dengan teks dan kemampuannya menangani data besar. Penelitian akan berfokus pada pengukuran sentimen pelanggan hotel dari ulasan yang diunggah oleh pelanggan hotel pada situs *Google Review* yang kemudian akan diproses dan diolah untuk dapat menghasilkan kelas tingkatan kepuasan pelanggan Grand Jatra Hotel Pekanbaru. Penelitian ini juga akan melakukan analisis berdasarkan hasil-hasil yang didapatkan untuk dapat menjadi saran kepada pihak hotel dalam upaya peningkatan dan perbaikan kualitas. Oleh karena itu, penulis mengangkat penelitian dengan judul “Analisis Sentimen Ulasan Pada *Google Review* Di Sebuah Penginapan Menggunakan Algoritma *Naive Bayes*” dengan melakukan ekstraksi data review pengguna dari *Google Review*. *Google Review* menjadi platform utama bagi pelanggan dalam memberikan ulasan terhadap layanan hotel. Dengan meningkatnya jumlah ulasan, diperlukan metode yang mampu mengklasifikasikan opini pelanggan dengan akurat dan cepat.

## 2. KAJIAN TEORITIS

**Table 1. Kajian Teoritis**

NO	Penulis	Judul	Perbandingan
1.	(Athira Luqyana dkk., 2018)	Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine	Penelitian ini berfokus pada analisis sentimen cyberbullying di Instagram menggunakan metode Support Vector Machine (SVM). Berbeda dengan penelitian peneliti yang fokus pada analisis sentimen terkait keputusan MK pada hasil Pilpres 2024 di Twitter dengan algoritma Naïve Bayes.
2.	(Anam dkk., 2021)	Penerapan Naïve Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen dan Pemeritah	Penelitian ini membandingkan tiga metode klasifikasi (Naïve Bayes, KNN, dan Decision Tree) untuk analisis sentimen interaksi netizen dengan pemerintah. Sementara penelitian ini hanya menggunakan Naïve Bayes untuk menganalisis sentimen terkait keputusan MK di Twitter.
3.	(Atika dkk., 2022)	Term Frequency-Inverse Document Frequency Support Vector Machine untuk Analisis Sentimen Opini Masyarakat Terhadap Tekanan Twitter	Penelitian ini menggunakan metode TF-IDF dan SVM untuk analisis sentimen di Twitter, tetapi fokusnya adalah pada opini masyarakat terhadap tekanan di Twitter. Penelitian ini lebih spesifik pada sentimen terhadap keputusan MK terkait Pilpres 2024.
4.	(Fres, 2022)	Analisis sentimen komentar kebijakan full day school dari Facebook page Kemendikbud RI menggunakan algoritma naive bayes classifier	Penelitian ini melakukan analisis sentimen terhadap kebijakan pendidikan (full day school) di Facebook menggunakan Naïve Bayes. Penelitian ini hanya berfokus pada Twitter dan isu politik terkait keputusan MK.
5.	(Aldean, dkk., 2022)	Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 di Twitter Menggunakan Metode Random Forest Classifier (Studi Kasus: Vaksin Sinovac)	Penelitian ini menganalisis sentimen masyarakat terhadap vaksinasi Covid-19 di Twitter menggunakan Random Forest. Perbedaan utama dengan penelitian peneliti adalah topik yang dibahas dan algoritma yang digunakan.
6.	(Wildan, 2023)	Sentimen negatif netizen dalam kolom komentar detik.com terhadap pemberitaan kasus Ferdy Sambo	Penelitian ini fokus pada sentimen negatif netizen terhadap pemberitaan kasus kriminal di detik.com, berbeda dengan penelitian ini, penelitian ini hanya berfokus pada politik di Twitter.
7.	(Kaka dkk., 2023)	Analisis Sentimen Dampak Perkembangan Teknologi Informasi dan Komunikasi Terhadap Kemajuan Belajar Siswa SMK Rada Pamba dengan Metode Naive Baiyes	Penelitian ini menganalisis dampak teknologi terhadap kemajuan belajar siswa menggunakan Naïve Bayes, sedangkan penelitian ini menganalisis sentimen politik di media sosial.
8.	(Halim dkk., 2023)	Klasifikasi Sentimen Masyarakat Terhadap Prabowo Subianto Bakal Calon Presiden 2024 di Twitter Menggunakan	Penelitian ini menganalisis sentimen masyarakat terhadap calon presiden di Twitter menggunakan Naïve Bayes. Meskipun menggunakan media dan algoritma yang sama, topik dan fokus

NO	Penulis	Judul	Perbandingan
		Naïve Bayes Classifier	peristiwa politik yang dianalisis berbeda dengan penelitian ini.
9.	(Wildan, 2023)	Analisis Sentimen Respons Twitter Terhadap Persyaratan Badan Penyelenggara Jaminan Sosial (BPJS) dikantor Pertahanan	Penelitian ini fokusnya pada sentimen Twitter terkait persyaratan BPJS di kantor pertahanan. Penelitian ini berbeda dengan penelitian ini yang fokus pada keputusan MK.
10.	(Rita Apriani dkk.,)	Analisis Sentimen Dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia	Penelitian ini menganalisis sentimen di aplikasi Tokopedia menggunakan Naïve Bayes. Penelitian peneliti lebih spesifik pada platform Twitter dan isu politik nasional.

### 3. METODE PENELITIAN

Metode penelitian ini bersifat kuantitatif dan memanfaatkan algoritma Naïve Bayes dalam klasifikasi teks untuk menganalisis sentimen ulasan pelanggan di Google Review terhadap layanan sebuah penginapan. Desain penelitian ini merupakan penelitian terapan di bidang text mining dan machine learning dengan pendekatan analisis sentimen. Populasi dalam penelitian ini adalah seluruh ulasan pelanggan di Google Review yang membahas pengalaman menginap di penginapan tersebut. Sampel diambil dengan teknik purposive sampling menggunakan data scraping untuk mengumpulkan ulasan yang relevan dengan topik penelitian.

Data diperoleh melalui scraping dengan menggunakan bahasa pemrograman Python. Instrumen utama adalah script scraping yang dijalankan di Google Colab untuk mengumpulkan komentar terkait topik.

#### Alat Analisis Data:

Data dianalisis melalui beberapa tahapan pemrosesan teks, yaitu:

- Case folding
- Tokenizing
- Normalisasi
- Stopword removal
- Stemming

Pembobotan dilakukan dengan metode TF-IDF dan diklasifikasikan menggunakan *Naïve Bayes Classifier*. Evaluasi model dilakukan dengan menghitung akurasi, presisi, recall, dan F1-score, serta menggunakan confusion matrix.

## **Model Penelitian**

Model klasifikasi dilakukan dengan algoritma *Naïve Bayes* yang dilatih dari data latih (80%) dan diuji dengan data uji (20%). Model dievaluasi menggunakan metrik standar klasifikasi. Simbol dan formula yang digunakan mengacu pada teori *Bayes* dan disajikan secara sederhana tanpa menguraikan detail matematika rumus dasar yang umum.

## **4. HASIL DAN PEMBAHASAN**

### **4.1 Proses Pengumpulan Data**

Proses pengumpulan data dilakukan dengan metode scraping ulasan pelanggan pada platform Google Review, menggunakan library Python melalui *Google Colab*. Scraping dilakukan berdasarkan kata kunci yang relevan dengan topik, yaitu nama penginapan dan aspek-aspek layanan yang sering dibahas oleh pelanggan.

Rentang waktu pengumpulan data dilakukan pada bulan Mei–Juni 2024, sedangkan lokasi penelitian dilaksanakan secara daring (*online*) dari Bengkalis menggunakan perangkat laptop dengan sistem operasi Windows dan koneksi internet.

Data yang diperoleh kemudian melalui proses validasi oleh ahli linguistik guna memastikan kesesuaian makna sebelum dilakukan pelabelan sentimen.

### **4.2 Hasil Analisis Data**

#### **4.2.1 Proses Preprocessing Data**

Proses preprocessing dilakukan melalui beberapa tahapan:

- a) Case folding
- b) Tokenizing
- c) Normalisasi
- d) Stopword removal
- e) Stemming

Setiap tahapan ini bertujuan untuk membersihkan dan menstandarisasi data komentar sehingga dapat dianalisis secara akurat oleh algoritma klasifikasi.

Contoh hasil preprocessing dapat dilihat pada **Tabel 2**.

**Table 2** Hasil Preprocessing Ulasan Hotel

No	Komentar Asli	Hasil Stemming
1	“kasurnya gerak dikit bunyi2 air nya kuning poolnya oke sarapannya oke banyak varian dan rasanya okeee”	“kasur gerak dikit bunyi air kuning’ pool sarapan varian rasa”
2	“kamarnya nyaman sekali makanan enak semua namun mungkin lebih ditambah variasi nya event disini berjalan lancar dengan pelayanan semua staff yang sangat memuaskan”	“kamar aman makan enak senua variasi event jalan lancar layan staff sangat puas”

#### 4.2.2 Pembobotan dan Klasifikasi

Setelah preprocessing, komentar dianalisis dengan pembobotan TF-IDF lalu diproses menggunakan metode klasifikasi Naïve Bayes. Data dibagi menjadi data latih (80%) dan data uji (20%). Kinerja model dievaluasi menggunakan confusion matrix, dan hasil evaluasi ditampilkan pada Tabel 3

*Hasil Evaluasi Model Klasifikasi*

**Table 3.** Hasil Evluasi Model Klasifikasi

Metrik	Nilai
Akurasi	98%
Presisi	97%
Recall	100%
F1-Score	92%

#### 4.3 Interpretasi Dan Pembahasan

##### 4.3.1 Keterkaitan hasil dengan teori

Hasil penelitian ini memperkuat efektivitas algoritma Naïve Bayes dalam analisis sentimen berbasis teks pendek di Google Review. Kinerja model cukup baik dalam mengklasifikasikan ulasan ke dalam kategori positif dan negatif, sebagaimana dijelaskan dalam teori bahwa Naïve Bayes cocok untuk kasus klasifikasi teks yang sederhana namun efisien.

##### 4.3.2 Perbandingan dengan penelitian sebelumnya

Hasil penelitian ini sejalan dengan penelitian sebelumnya Halim dkk. (2023) yang juga menggunakan Naïve Bayes untuk menganalisis pendapat pengguna tentang layanan publik. Kedua penelitian sama-sama menunjukkan tingkat akurasi yang tinggi (lebih dari 80%), sehingga membuktikan bahwa metode ini cukup baik untuk mengenali apakah ulasan bersifat positif atau negatif, baik dalam layanan publik maupun ulasan hotel/penginapan.

### 4.3.3 Implikasi Penelitian

- a) Secara teoritis, penelitian ini memperkuat literatur bahwa algoritma probabilistik seperti Naïve Bayes tetap relevan untuk klasifikasi sentimen di era big data.
- b) Secara terapan, hasil analisis dapat digunakan oleh pengambilan kebijakan untuk memetakan opini secara dan efisien melalui Google Review.

## 5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menganalisis pendapat pelanggan tentang Grand Jatra Hotel Pekanbaru melalui ulasan di Google Review dengan menggunakan algoritma Naïve Bayes. Dari hasil pengujian, algoritma ini mampu membedakan ulasan menjadi dua kategori, yaitu positif dan negatif, dengan accuracy 98%, presisi 97%, recall 100% dan F1-Score 92% untuk ulasan positif. Hasil ini menunjukkan bahwa Naïve Bayes sangat efektif digunakan untuk mengenali sentimen dari teks ulasan pelanggan. Penelitian ini juga menunjukkan bahwa analisis sentimen bisa menjadi alat yang berguna untuk mengetahui pendapat pelanggan dan membantu pihak hotel dalam meningkatkan pelayanan dengan bantuan teknologi kecerdasan buatan.

## DAFTAR REFERENSI

- Aldean, F., Ramadhan, H., & Subhan, A. (2022). Analisis sentimen masyarakat terhadap vaksinasi Covid-19 di Twitter menggunakan metode Random Forest Classifier (Studi kasus: Vaksin Sinovac). *Jurnal Teknik Informatika*, 6(2), 85–92.
- Anam, M., Rafiq, M. A., & Siregar, A. (2021). Penerapan Naïve Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk menganalisis sentimen pada interaksi netizen dan pemerintah. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(3), 297–305.
- Apriani, R., & Gustian, D. (2020). Analisis sentimen dengan Naïve Bayes terhadap komentar aplikasi Tokopedia. *Jurnal Sistem Informasi dan Komputer*, 9(2), 112–119.
- Athira, L., Nabilah, S., & Hardiyanti, R. (2018). Analisis sentimen cyberbullying pada komentar Instagram dengan metode klasifikasi Support Vector Machine. *Seminar Nasional Teknologi Informasi dan Komunikasi*, 2(1), 55–61.
- Atika, R., Rahmah, A., & Widodo, P. (2022). Term Frequency-Inverse Document Frequency Support Vector Machine untuk analisis sentimen opini masyarakat terhadap tekanan mental pada media sosial Twitter. *Jurnal Ilmu Komputer*, 10(1), 45–53.

- Darman, R. (2020). Analisis sentimen respons Twitter terhadap persyaratan Badan Penyelenggara Jaminan Sosial (BPJS) di Kantor Pertahanan. *Jurnal Teknologi dan Sistem Informasi*, 7(2), 131–138.
- Fath, R. A. (2020). Analisis sentimen komentar kebijakan full day school menggunakan algoritma Naïve Bayes Classifier. *Jurnal Ilmu Komputer*, 8(1), 40–48.
- Firmansyah, D., & Yusuf, A. (2020). Perbandingan kinerja algoritma Naïve Bayes dan Logistic Regression untuk klasifikasi sentimen di Twitter. *Jurnal Informatika dan Sains Data*, 6(1), 55–63.
- Halim, A. R., Prasetyo, A., & Nursamsi, R. (2023). Klasifikasi sentimen masyarakat terhadap Prabowo Subianto bakal calon presiden 2024 di Twitter menggunakan Naïve Bayes Classifier. *Jurnal Teknik Komputer*, 11(1), 55–63.
- Hidayatullah, A., & Sari, D. P. (2021). Implementasi Naïve Bayes dalam analisis sentimen opini publik terhadap kebijakan pemerintah di Twitter. *Jurnal Teknologi Informasi dan Komputer*, 10(2), 66–74.
- Kaka, M., Fahmi, M., & Lestari, N. (2023). Analisis sentimen dampak perkembangan teknologi informasi dan komunikasi terhadap kemajuan belajar siswa SMK Rada Pamba dengan metode Naïve Bayes. *Jurnal Informatika dan Sistem Informasi*, 7(2), 78–84.
- Maulana, H., & Pratiwi, Y. (2022). Pemanfaatan media sosial untuk deteksi opini publik terkait isu nasional. *Jurnal Komunikasi Digital*, 4(1), 40–50.
- Putri, N., & Prasetya, H. (2021). Text mining untuk analisis sentimen netizen terhadap figur publik menggunakan Naïve Bayes dan SVM. *Jurnal Teknologi dan Sains Informasi*, 5(3), 98–106.
- Sembiring, R., & Yulianti, A. (2023). Analisis sentimen Twitter mengenai Pemilu 2024 menggunakan algoritma machine learning. *Jurnal Sistem Informasi dan Teknologi*, 9(1), 25–33.
- Wildan, R. (2023). Sentimen negatif netizen dalam kolom komentar detik.com terhadap pemberitaan kasus Ferdy Sambo. *Jurnal Komunikasi Massa*, 6(2), 101–109.