

Sentiment Analysis on Google Reviews Using Naïve Bayes, K-Nearest Neighbors, and Logistic Regression to Improve Novotel Services

Yonathan Arya Dhamma^{1*}, Simon Prananta Barus^{2*}

* Informatika, Universitas Matana

yonathannews123@gmail.com¹, simonpbarus@gmail.com²

Article Info

Article history:

Received 2024-11-19

Revised 2024-12-25

Accepted 2025-01-15

Keyword:

K-NN,
Logistic Regression,
Machine Learning,
Naïve Bayes,
Sentiment Analysis.

ABSTRACT

The application of artificial intelligence (AI) has been widely used in various industrial sectors, including the hospitality industry. One of the applications that is widely used in the hospitality industry is sentiment analysis. Sentiment analysis is carried out by analyzing feedback data from hotel guests or customers. The results of this sentiment analysis are important for decision makers to improve and improve their services. This study aims to obtain sentiment analysis results from Novotel hotel Google reviews based on machine learning by comparing three algorithms, namely Naïve Bayes, K-Nearest Neighbors (KNN), and Logistic Regression. The stages carried out in this study are data collection, data labeling, exploratory data analysis (EDA), data preprocessing, text representation, data sharing, modeling, model training, model evaluation, selection of the most accurate model, visualization of the most accurate model, interpretation of results and writing research reports. The dataset used was 1200 reviews, only 1190 reviews were used in the analysis. From the training results, the model produced by the Logistic Regression algorithm was the most accurate, namely 94.54% with unigrams ($n = 1$). Here are the results of each category, positive as many as 723 reviews (60.76%), negative as many as 218 reviews (18.32%), and neutral as many as 249 reviews (20.92%). Thus, most of the sentiment towards the service is positive, but some services need to be fixed and improved for customer satisfaction. The next research, the research area is expanded and the use of Deep Learning.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Saat ini, *Machine Learning* (ML) merupakan bagian dari *artificial intelligent* (AI) dan *Deep Learning* (DL) merupakan bagian dari ML, bidang AI yang sedang tren. Tidak hanya ML dan DL yang tren, tapi juga *Natural Language Processing* (NLP) yang terimbas dari perkembangan ML dan DL. NLP memungkinkan AI memahami dan mengekstrak informasi dari bahasa manusia, yang sangat penting untuk analisis teks dan interaksi [1]-[2]. Analisis teks ini dapat dilakukan dari data di media sosial, email, messaging, e-commerce, layanan berita, dan lain-lain, di mana format datanya adalah teks. Analisis teks ini banyak dimanfaatkan oleh industri-industri, termasuk industri perhotelan.

Industri perhotelan mengaplikasikan AI secara luas untuk berbagai keperluan, seperti analisis sentimen ulasan

pelanggan, sistem rekomendasi, chatbot, dan manajemen prediksi pemesanan. Dengan penggunaan AI, hotel dapat mengidentifikasi tren dari umpan balik pelanggan, sehingga mampu meningkatkan kualitas layanan dan mengidentifikasi area yang perlu ditingkatkan untuk kepuasan pelanggan yang lebih baik [3].

Di era digital saat ini, media sosial telah menjadi salah satu faktor utama dalam memengaruhi keputusan pemilihan hotel. *Google Maps Review* memberikan kemudahan bagi penggunanya, memungkinkan wisatawan untuk membagikan pengalaman-pengalamannya, baik melalui unggahan di media sosial pribadi maupun ulasan atau penilaian yang tersedia di aplikasi tersebut. Unggahan dan ulasan ini bisa menjadi bahan pertimbangan dan memengaruhi keputusan calon tamu hotel lainnya yang melihatnya [4]. Peninjauan [5] menunjukkan bahwa saat ini 62.3% populasi dunia adalah pengguna aktif

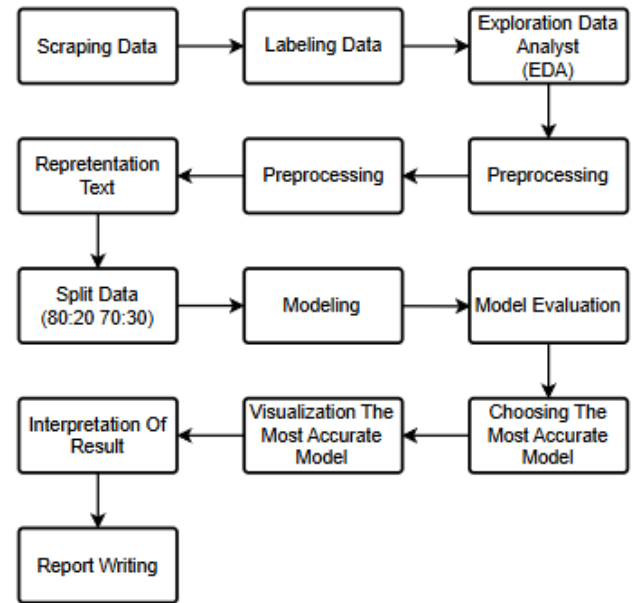
sosial media, sementara pengguna internet mencapai 66%. Penggunaan yang luas ini tentunya berdampak pada peningkatan jumlah unggahan dan *review*. *Big Data* yang tersedia secara publik memberikan wawasan baru bagi calon pengunjung hotel mengenai ulasan tentang fasilitas hotel [6]. Analisis sentimen dari ulasan pengunjung hotel menjadi sangat penting, karena tamu atau pelanggan mengharapkan pengalaman yang memuaskan saat berlibur atau berkunjung [7]. Analisis sentimen ini dapat mendukung proses perencanaan, implementasi langkah, dan strategi pemasaran para pemegang kepentingan dan mengambil keputusan dalam peningkatan fasilitas hotel. Analisis sentiment ini harus dilakukan dengan mencapai akurasi yang sangat tinggi agar bisa mendapatkan perspektif yang lebih beragam mengenai fasilitas hotel berdasarkan sentimen para pengunjung.

Berdasarkan masalah itu, penelitian ini dilakukan untuk mendukung pengembangan fasilitas hotel. Tujuan penelitian ini adalah untuk memahami sentimen pengunjung hotel terhadap fasilitas hotel dengan membandingkan tiga model ML berdasarkan akurasi. Temuan ini juga dapat digunakan sebagai bahan untuk dipertimbangkan dalam mengambil keputusan terkait memilih Novotel Tangerang untuk dijadikan tempat menginap. Sejumlah penelitian terdahulu telah menerapkan analisis sentimen *review* fasilitas hotel. Pada penelitian sebelumnya oleh [8], mengimplementasikan *Naive Bayes* untuk menganalisis sentimen *review* pengunjung hotel terhadap fasilitas dengan menggunakan lima variabel: *Review*, *Nama_Hotel*, *Nama*, *SENTIMEN*, *Rating*. Mencapai tingkat akurasi 75,66%. Penelitian lainnya [9] menggunakan metode serta variabel yang serupa untuk melakukan analisis sentimen *review* pengunjung hotel dengan mendapatkan hasil akurasi 90.71%. Kemudian terdapat penelitian terdahulu [10] yang menggunakan *Multinomial Naive Bayes* sebagai teknik analisis sentimen dengan menggunakan pembobotan *TF-IDF* mencapai tingkat akurasi 88% dengan nilai *recall* 88%, *precision* 88%, dan *f1-score* 87%. Terdapat juga penelitian lain [11] yang fokus pada menggabungkan metode *SVM* dan *Naive Bayes* guna mengetahui dimana model yang mendapatkan hasil akurasi yang paling tinggi, hasil akurasi tertinggi diperoleh menggunakan *SVM* dengan tingkat akurasi 95% dan tingkat akurasi terendah menggunakan *naive bayes* dengan tingkat akurasi 88%.

Mengingat penelitian-penelitian terdahulu paling banyak hanya menggunakan dua metode, penelitian ini mengusulkan pendekatan baru dengan menggunakan *Naive Bayes*, *Logistic Regression*, dan *K-NN* yang dikombinasikan dengan pembobotan *TF-IDF* dan *Unigram*(n=1) sampai *Fivegram*(n=5) untuk melakukan sentimen analisis *review* pengunjung hotel tentang fasilitas hotel. *Variabel* yang digunakan mencakup nama, ulasan, lokasi, dan sentimen. Kemudian akan dilakukan juga evaluasi menggunakan *Confusion Matrix* berdasarkan tiga algoritma yang digunakan dari *Unigram*(n=1) hingga *Fivegram*(n=5). Terdapat empat faktor utama dalam mengukur performa menggunakan *Confusion Matrix*, yang mencakup Akurasi, *Precision*, *Recall*, dan *F1-Score* [12].

II. METODE PENELITIAN

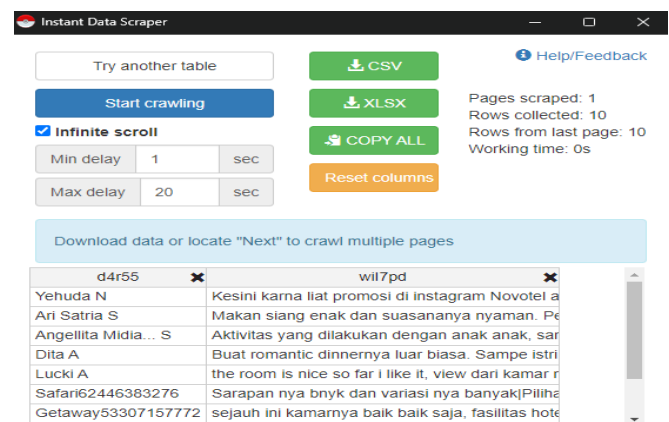
Metodologi penelitian disini berfungsi sebagai panduan yang mencakup rencana dan prosedur penelitian untuk mencapai tujuan yang diinginkan [13]. Pada penelitian ini, *Naive Bayes*, *K-NN*, dan *Logistic Regression* digunakan dalam melakukan sentimen analisis terhadap fasilitas hotel di Novotel Tangerang. Gambar 1 menunjukkan alur penelitian yang dilakukan.



Gambar 1. Tahapan Penelitian

A. Scraping Data

Scraping data adalah tahap mengumpulkan data dimana data diperoleh dari *google review*. Data yang dipakai dalam studi ini adalah *review* dari *google review* yang diambil di *google review* berjumlah 1200 data, diantara lain terdiri dari 733 sentimen yang positif 218 sentimen yang negatif, dan 249 sentimen yang netral.

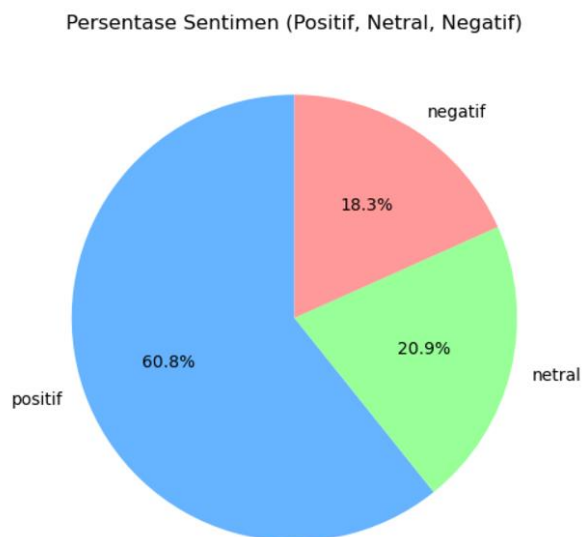


Gambar 2. Instant Data Scraper

Terdapat beberapa kategori ulasan, diantaranya mengenai layanan, fasilitas, kebersihan, makanan, dan lainnya. Data ini diperoleh menggunakan ekstensi *Instant Data Scrapper* di *Google Chrome* yang dilakukan pada 1 Oktober 2024 dan memiliki kolom nama, ulasan, lokasi, dan sentimen.

B. Labeling Data

Pelabelan data adalah proses untuk memberi tag pada data. Proses ini penting untuk menentukan kelas suatu ulasan pada analisis sentimen, apakah ulasan ini termasuk dalam label yang positif, yang negatif atau yang netral [14]. Label positif mencakup sanjungan dan perasaan positif seperti kesenangan, kepuasan, serta kebahagiaan oleh pengunjung hotel mengenai fasilitas yang diterima. Sebaliknya, label negatif terdiri dari keluhan, keberatan, protes, kecaman, dan perasaan negatif seperti kemarahan, kekecewaan, dan ketidakpuasan yang didapat. Sementara itu label netral berisi ulasan yang bersifat informatif atau deskriptif tanpa indikasi emosi tertentu, seperti laporan masalah tanpa keluhan emosional, saran tanpa ekspresi kepuasan atau ketidakpuasan, serta tanggapan yang bersifat umum. Penelitian ini menerapkan pelabelan dengan melibatkan beberapa pelabel untuk meningkatkan akurasi hasil pelabelan, sesuai dengan sentimen dari ulasan pada *Google Review* pengunjung Novotel Tangerang.



Gambar 3. Persentase Sentimen

C. Exploratory Data Analysis (EDA)

EDA adalah proses memperoleh pengetahuan awal mengenai data yang ada sebelum melanjutkan ke proses penelitian yang lebih lanjut dan kompleks [15]. *EDA* berguna untuk mengekstraksi berbagai informasi awal terkait data yang dianalisis, seperti jumlah data secara keseluruhan, kelengkapan data, jumlah data berlabel positif, negatif, dan netral.

D. Pre-processing

Preprocessing adalah proses mengubah data mentah menjadi bentuk yang terstruktur dan dapat diproses [10]. Berikut ini beberapa tahapan dalam *preprocessing* data:

1) Cleaning

Tahapan ini bertujuan untuk menyingkirkan atribut data yang tidak diperlukan untuk proses *classification*, seperti tanda baca, spasi kosong, dan emoji [16].

2) Case Folding

Tahapan ini bertujuan untuk mengubah semua teks menjadi *lower case* [17].

3) Remove Stopwords

Tahapan ini bertujuan untuk menyingkirkan kata yang tidak berguna bagi makna kalimat, seperti kata penghubung atau kata henti [16]. *Library* yang digunakan untuk *remove stopwords* menggunakan NTLK dari *python*.

4) Stemming

Proses ini bertujuan untuk *mentransformasi* kata dalam dataset menjadi bentuk dasar dan menyingkirkan imbuhan pada kata-kata tersebut [18]. *Library* yang digunakan untuk melakukan *stemming* bahasa indonesia yaitu, Sastrawi dari *python*.

5) Tokenization

Proses ini bertujuan untuk memecah kalimat menjadi kata-kata terpisah untuk mempermudah klasifikasi [16].

E. Transformation Data

Transformation Data adalah data yang diubah memakai *Term Frequency-Inverse Document Frequency (TF-IDF)* yang kemudian akan mengkombinasikan frekuensi dari n-gram yang berbeda dari *Unigram* hingga *Fivegram* untuk menghasilkan representasi numerik dari data teks dengan metode *labelEncoder* untuk mempermudah pembelajaran untuk menangkap lebih banyak konteks dari teks [19]. Sebelum transformasi, data dipecah terlebih dahulu dengan dua skenario perbandingan yang dilakukan secara *random sampling* sebagai berikut:

- *Scenario 1*: data 80:20
- *Scenario 2*: data 70:30

Kedua skenario ini dipilih karena merupakan skenario yang secara umum digunakan untuk menguji konsistensi performa model dengan ukuran data latih dan data uji yang berbeda.

F. Naive Bayes

Naive Bayes merupakan model klasifikasi probabilistik sederhana, terdapat pada Teori Bayes. Teori ini menyatakan bahwa probabilitas terjadinya suatu peristiwa sama dengan probabilitas dasarnya (dihitung dari data saat ini) dikalikan dengan probabilitas bahwa kejadian serupa akan terjadi lagi di masa depan (berdasarkan kejadian sebelumnya). *Naive Bayes* adalah metode klasifikasi dokumen berbasis

Supervised Learning, yang memerlukan data pelatihan sebelum melakukan klasifikasi [20]. Rumus *Naive Bayes* adalah sebagai berikut:

$$P(C_i|X) = \frac{P(X|C_i) \cdot P(C_i)}{P(X)} \quad (1)$$

- X: Data kelas yang tidak diketahui;
- C_i: Kemungkinan bahwa data X berada di suatu kelas tertentu;
- P(C_i|X): Peluang bahwa C_i sesuai dengan kondisi X;
- P(C_i): Probabilitas dari C_i;
- P(X|C_i): Peluang X sesuai dengan kondisi pada C_i;
- P(X): Probabilitas dari X;

G. K-NN

K-Nearest Neighbor merupakan model klasifikasi yang bekerja dengan mengelompokkan data baru berdasarkan jarak terdekatnya terhadap data tetangga terdekat. Algoritma ini termasuk metode sederhana namun efektif dalam menangani klasifikasi, mampu menghasilkan performa yang kompetitif [21]. Proses klasifikasi dengan *KNN* terdiri dari beberapa tahap:

- Menentukan nilai k;
- Menghitung jarak antara objek baru dan data; training yang sudah ada menggunakan rumus Euclidean Distance;
- $D(a, b) = \sqrt{\sum_{k=1}^d (a_k - b_k)^2}$ (2)
- Menghasilkan *output* klasifikasi akhir, sebelum data digunakan dalam proses klasifikasi, data tersebut terlebih dahulu dipreproses untuk memastikan bahwa data yang digunakan sudah bersih dan siap pakai.

H. Logistic Regression

Logistic Regression adalah model statistik yang bertujuan untuk mengetahui apakah suatu variabel independen memengaruhi variabel dependen biner. Dengan memanfaatkan fungsi *sigmoid*, *logistic regression* menghasilkan keluaran berupa probabilitas dengan rentang nilai antara 0 dan 1 [22]. *Logistic Regression* menggunakan fungsi *sigmoid*, yang dinyatakan sebagai berikut:

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \quad (3)$$

Dengan:

- $P(y = 1|x)$: probabilitas variabel dependen bernilai 1
- β_0 : konstanta atau intercept
- $\beta_1, \beta_2, \dots, \beta_n$: koefisien yang menunjukkan pengaruh variabel independen
- x_1, x_2, \dots, x_n terhadap probabilitas

I. Evaluation Model

Evaluasi model adalah proses penggunaan matriks konfusi sebagai alat ukur performa model *machine learning*. Pada tahap akhir, dilakukan evaluasi untuk menghitung validitas model yang dibentuk dari model yang telah diterapkan pada tahapan sebelumnya. Memperlihatkan model yang paling akurat dengan mempertimbangkan beberapa faktor. Matriks Konfusi ini berbentuk dua dimensi yang menunjukkan perbandingan antara hasil prediksi dan kondisi aktual, seperti yang terlihat pada tabel berikut [23]:

TABEL I
MATRIKS KONFUSI

Classification	Group	
	+	-
+	True Positives	False Negatives
-	False Positives	True Negatives

Rumus dari *accuracy*, *precision*, *recall*, dan *f1-score*:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (7)$$

J. Visualization Data

Visualisasi Data adalah tahap yang bertujuan untuk menggambarkan hasil evaluasi model secara grafis, sehingga mudah dipahami dan dapat memberikan wawasan yang lebih mendalam mengenai performa model dalam analisis sentiment [24]. Visualisasi ini mencakup penggunaan berbagai grafik dan diagram yang membantu dalam menyajikan kinerja model, distribusi sentimen, dan perbandingan antara model yang digunakan dalam penelitian ini. Berikut adalah kegunaan dari visualisasi data:

- $P(y = 1|x)$: probabilitas variabel dependen bernilai 1
- Mempermudah proses analisis data
- Memfasilitasi developer dalam menyampaikan data dengan lebih efektif
- Memudahkan developer dalam memahami data

K. Interpretation of Results

Tahap interpretasi hasil adalah tahapan untuk memberikan pemahaman yang lebih mendalam mengenai performa model yang terbaik dalam mengklasifikasikan sentimen ulasan [25]. Proses interpretasi hasil mencakup analisis performa model terakurat, hasil dari matriks konfusi, analisis distribusi sentimen, dan manfaat dari penelitian ini yang dapat digunakan sebagai rekomendasi untuk pihak hotel.

L. Report Writing

Penulisan laporan penelitian adalah tahap akhir yang bertujuan untuk menyampaikan hasil penelitian secara jelas,

sistematis, dan mudah dipahami. Laporan ini dirancang untuk memberikan gambaran komprehensif mengenai seluruh proses penelitian mulai dari pendahuluan, metodologi, dan hasil dan kesimpulan.

III. HASIL DAN PEMBAHASAN

Data yang dikumpulkan dari ulasan pengunjung Novotel Tangerang yang tersedia di *Google Review* digunakan dalam penelitian ini. Ulasan teks pengunjung dikumpulkan pada 1 Oktober 2024 yang mencakup ulasan tentang layanan Novotel Tangerang. Tabel II adalah sebagian dari hasil *Scraping Data*.

TABEL II
HASIL SCRAPING DATA

Nomor	Nama	Ulasan	Lokasi
1	Ilmi Ilmi	Pengalaman yang menakjubkan, kamar executive suite di lantai 1719 sangat mewah sekali.	Novotel Tangerang
2	Yohan Sasmita	Kamarnya biasa saja. Cocok dengan harganya.	Novotel Tangerang
3	Dekaes	Penurunan kualitas dibanding kunjungan sebelumnya. Semoga di kunjungan berikutnya lebih baik lagi. Sekarang hanya ada 2 kali sajian (Breakfast dan High Tea) di Premier Lounge	Novotel Tangerang

Sebanyak 1200 data telah berhasil dikumpulkan, yang kemudian akan melalui proses pelabelan manual yang dibantu oleh beberapa pelabel untuk memastikan hasil yang optimal. Tujuan dari pelabelan yang dilakukan adalah untuk mengkategorikan data ke dalam tiga kategori, yaitu label positif, label negatif, dan label netral. Tabel III adalah sebagian dari hasil pelabelan data.

TABEL III
HASIL PELABELAN DATA

Nomor	Nama	Ulasan	Lokasi	Sentimen
1	Ilmi Ilmi	Pengalaman yang menakjubkan, kamar executive suite di lantai 1719 sangat mewah sekali.	Novotel Tangerang	Positif

2	Yohan Sasmita	Kamarnya biasa saja. Cocok dengan harganya.	Novotel Tangerang	Netral
3	Dekaes	Penurunan kualitas dibanding kunjungan sebelumnya. Semoga di kunjungan berikutnya lebih baik lagi. Sekarang hanya ada 2 kali sajian (Breakfast dan High Tea) di Premier Lounge	Novotel Tangerang	Negatif

Setelah proses pelabelan, dilakukan analisis data eksploratori. Dari pengecekan tersebut, tidak ditemukan nilai kosong pada dataset. Dari 1200 data yang sudah diproses, sebanyak 733 data positif, 218 data negatif, dan 249 data netral. Tabel II adalah sebagian dari hasil *EDA*.

TABEL IV
HASIL EDA

Nomor	Column	NN Count	Dtype
1	Nama	1200 nn	Object
2	Ulasan	1200 nn	Object
3	Lokasi	1200 nn	Object
4	Sentimen	1200 nn	Object

Tahap berikutnya adalah *Preprocessing Data*, yang merupakan langkah yang penting untuk proses teks *mining*. Tabel V adalah sebagian dari hasil *Preprocessing*.

TABEL V
HASIL PREPROCESSING

Nomor	Tahapan	Hasil
1	Data Awal	Pengalaman yang menakjubkan, kamar executive suite di lantai 1719 sangat mewah sekali.
2	Cleaning	Pengalaman yang menakjubkan kamar executive suite di lantai 1719 sangat mewah sekali
3	Case Folding	pengalaman yang menakjubkan kamar

		executive suite di lantai 1719 sangat mewah sekali
4	Remove Stopwords	pengalaman menakjubkan kamar executive suite lantai 1719 mewah
5	Stemming	alam takjub kamar executive suite lantai 1719 mewah
6	Tokenization	['alam', 'takjub', 'kamar', 'executive', 'suite', 'lantai', '1719', 'mewah']

Tahap berikutnya adalah membagi data *training* dan data *testing*, diberikan dua scenario, *scenario* pertama 80:20 dan *scenario* kedua 70:30, kemudian dilakukan pembobotan kata, pada tahap ini, metode *Term Frequency-Inverse Document Frequency* dipakai untuk menentukan bobot dari setiap term atau kata dalam data teks, sehingga menghasilkan representasi yang lebih informatif bagi model. Pembobotan dilakukan dengan menghitung frekuensi kemunculan setiap term tertentu, tidak hanya dalam bentuk kata tunggal (*unigram*), tetapi juga dengan menggunakan kombinasi n-gram hingga lima kata berurutan (*fivegram*). Pendekatan ini memungkinkan analisis yang lebih mendalam karena mempertimbangkan konteks yang lebih luas dari pola kata yang muncul dalam data.

Berikutnya dilakukan pengujian model, bertujuan untuk mengevaluasi kemampuan model dalam membuat prediksi yang benar terhadap data baru yang belum pernah diproses model sebelumnya. Hasil pengujian model dengan Tingkat akurasi tertinggi dari setiap model *Supervised Learning* dan *ngram* bisa dilihat pada Tabel VI.

TABEL VI
HASIL UJI MODEL DENGAN METODE 80:20

Nomor	Model	Skenario	Akurasi
1	Naïve Bayes	80:20	93.70%
2	K-NN	80:20	91.6%
3	Logistic Regression	80:20	94.54%

TABEL VI
HASIL UJI MODEL DENGAN METODE 70:30

Nomor	Model	Skenario	Akurasi
1	Naïve Bayes	70:30	90.48%
2	K-NN	70:30	87.96%
3	Logistic Regression	70:30	91.32%

Dari pengujian model yang sudah dilakukan, *Logistic Regression* yang menggunakan data *training* 80% dan data *testing* 20% mendapatkan tingkat akurasi tertinggi di angka 94.54%, kemudian di posisi kedua *Naïve Bayes* di angka 93.70% dan di posisi ketiga *K-NN* di angka 91.6%, kemudian untuk data latih 70% dan data uji 30%, *Logistic Regression* masih menduduki posisi pertama dengan tingkat akurasi di angka 91.32%, selanjutnya *Naive Bayes* dengan 90.48% dan yang terakhir *K-NN* dengan 87.96%.

Tahapan yang terakhir adalah evaluasi model. Berikut adalah hasil dari evaluasi ketiga model *Supervised Learning* yang digunakan bisa dilihat pada Tabel VII, VIII, IX, X, XI, dan XII.

TABEL VII
CONFUSION MATRIX NAÏVE BAYES 80:20

N-Gram	Accuracy	Precision	Recall	F1-Score
1	90.76%	94%	97%	95%
2	93.70%	95%	99%	97%
3	92.86%	95%	99%	97%
4	92.02%	94%	99%	96%
5	92.44%	94%	99%	96%

TABEL VIII
CONFUSION MATRIX NAÏVE BAYES 70:30

N-Gram	Accuracy	Precision	Recall	F1-Score
1	87.39%	88%	96%	92%
2	90.48%	91%	99%	95%
3	88.80%	89%	98%	93%
4	89.08%	90%	98%	94%
5	88.80%	90%	97%	94%

Dari Tabel VII (80:20) dan Tabel VIII (70:30), *Naïve Bayes* menunjukkan performa terbaik pada *Bigram* (n=2) dengan akurasi 93,70%, *precision* 95%, *recall* 99%, dan *F1-Score* 97%. Penggunaan n-gram yang lebih besar (*trigram* hingga *fivegram*) tidak memberikan peningkatan akurasi yang signifikan dan bahkan sedikit menurunkan performa model. Hal ini menunjukkan bahwa *bigram* mampu memberikan representasi yang optimal untuk *Naïve Bayes*, karena mampu menangkap kombinasi kata yang relevan tanpa meningkatkan kompleksitas secara berlebihan. Sementara *trigram* hingga *fivegram* menambah konteks, tetapi juga berpotensi memperkenalkan noise.

TABEL IX
CONFUSION MATRIX K-NN 80:20

N-Gram	Accuracy	Precision	Recall	F1-Score
1	88.66%	91%	97%	94%
2	88.24%	90%	97%	93%
3	88.66%	95%	93%	94%
4	91.18%	94%	95%	94%
5	91.60%	95%	96%	95%

TABEL X
CONFUSION MATRIX K-NN 70:30

N-Gram	Accuracy	Precision	Recall	F1-Score
1	83.75%	86%	97%	91%
2	85.43%	89%	96%	92%
3	85.15%	92%	92%	92%
4	87.96%	90%	96%	93%
5	87.68%	91%	95%	93%

Dari Tabel IX (80:20) dan Tabel X (70:30), *K-NN* mencapai akurasi tertinggi pada *Fivegram* (n=5) sebesar 91,60%, dengan *precision* 95%, *recall* 96%, dan *F1-Score* 95%. Penggunaan n-gram yang lebih besar memungkinkan model *K-NN* untuk memanfaatkan konteks kata yang lebih

luas, yang membantu dalam klasifikasi berbasis jarak. Namun, pada skenario 70:30, peningkatan n-gram hingga n=5 hanya memberikan sedikit peningkatan akurasi, menunjukkan potensi *overfitting* pada dataset yang lebih kecil.

TABEL XI
CONFUSION MATRIX LOGISTIC REGRESSION 80:20

N-Gram	Accuracy	Precision	Recall	F1-Score
1	94.54%	100%	100%	96%
2	92.86%	97%	100%	96%
3	91.18%	95%	100%	94%
4	91.18%	95%	100%	94%
5	92.02%	95%	100%	95%

TABEL XII
CONFUSION MATRIX LOGISTIC REGRESSION 70:30

N-Gram	Accuracy	Precision	Recall	F1-Score
1	91.32%	97%	100%	94%
2	88.52%	94%	100%	92%
3	88.52%	94%	100%	92%
4	88.52%	94%	100%	92%
5	88.80%	94%	100%	92%

Menurut Tabel XI (80:20) dan Tabel XII (70:30), *Logistic Regression* mencapai akurasi tertinggi pada *Unigram* (n=1) sebesar 94,54%, dengan *precision* 100%, *recall* 100%, dan *F1-Score* 96%. Penggunaan n-gram yang lebih besar (n=2 hingga n=5) tidak meningkatkan akurasi dan justru menurunkan performa model. Hal ini menunjukkan bahwa *Logistic Regression* lebih efektif dengan representasi teks yang sederhana, dimana *Unigram* sudah cukup untuk memaksimalkan efisiensi model linear tanpa kehilangan informasi penting.

Penelitian ini menunjukkan bahwa pemilihan ukuran n-gram sangat bergantung pada algoritma yang digunakan serta karakteristik dataset. *Naïve Bayes* lebih optimal dengan *Bigram* yang mampu menangkap kombinasi kata tanpa menambah kompleksitas yang berlebihan. K-NN memanfaatkan *Fivegram* untuk konteks yang lebih luas, namun perlu diperhatikan potensi *overfitting* pada dataset yang lebih kecil. Sedangkan *Logistic Regression* menunjukkan performa terbaik dengan *Unigram*, mengindikasikan bahwa representasi sederhana cukup untuk model *linear* ini.

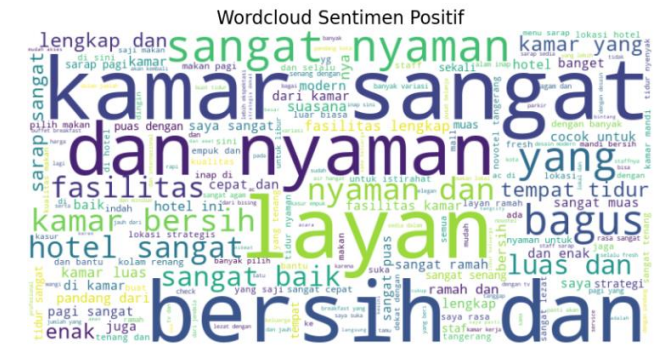
TABEL XIII
RINGKASAN DAN PERBANDINGAN KETIGA MODEL

Model	N-Gram Optimal	Akurasi	Precision	Recall	F1-Score
Naïve Bayes	Bigram (n=2)	93.70%	95%	99%	97%
K-NN	Fivegram (n=5)	91.60%	95%	96%	95%
Logistic Regression	Unigram (n=1)	94.54%	100%	100%	96%

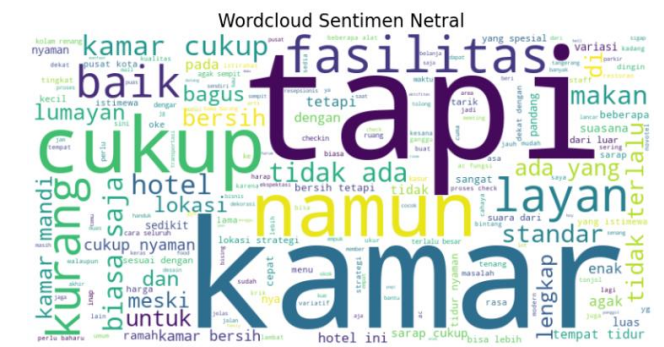
Kemudian penelitian ini juga menghasilkan temuan ilmiah bahwa algoritma *Logistic Regression* dengan menggunakan

pembobotan gabungan antara *TF-IDF* dan n-gram yang memakai 80% data latih dan 20% data uji memiliki *accuracy* tertinggi dalam melakukan analisis sentimen. Hasil analisis menunjukkan bahwa menggunakan *Logistic Regression* mendapatkan tingkat akurasi 94,5%, *Naïve Bayes* 93,7%, dan *K-Nearest-Neighbors* 91,6%.

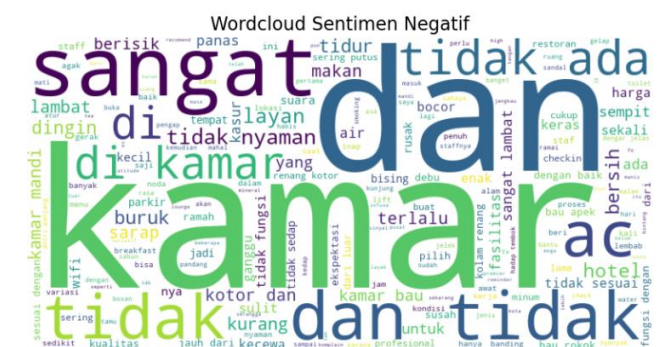
Visualisasi hasil berupa *word cloud* dapat menyajikan representasi visual dari kata-kata yang paling sering disebutkan dalam ulasan, visualisasi hasil *word cloud* dapat dilihat pada Gambar 3, 4, dan 5.



Gambar 4. World Cloud Positif



Gambar 5. World Cloud Negatif



Gambar 6. World Cloud Netral

Gambar 4, untuk *word cloud* positif, lima kata yang paling banyak muncul yaitu kata “bersih”, “nyaman”, “kamar”, “sangat”, dan “layan”. Yang menegaskan dimana fokus utama ulasan positif ada pada kenyamanan kamar dan keramahan staff.

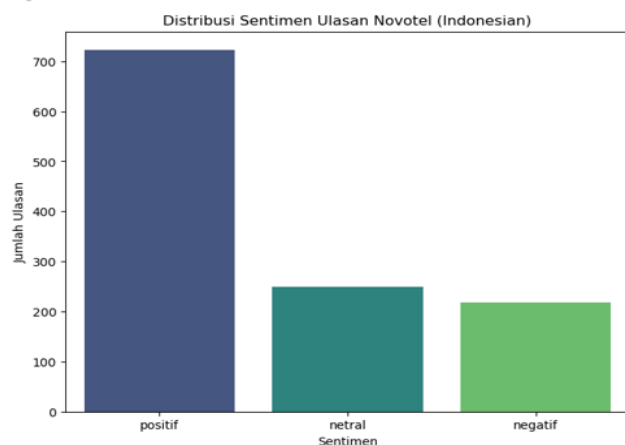
Gambar 5, untuk *world cloud* netral, lima kata yang paling banyak muncul yaitu kata “kamar”, “tapi”, “cukup”, “namun”, dan “fasilitas”. Yang menegaskan dimana fokus utama ulasan negatif ada pada kebersihan kamar dan kualitas makanan.

Gambar 6, untuk *world cloud* negatif, lima kata yang paling banyak muncul yaitu kata “kamar”, “sangat”, “tidak”, “dan”, dan “ac”.

Secara keseluruhan, hasil analisis sentimen menggunakan tiga algoritma *Supervised Learning* menunjukkan dominasi sentimen positif. Ini mengindikasikan bahwa persepsi pengunjung Novotel Tangerang terhadap fasilitas hotel dapat dianggap baik. Meskipun begitu, sentimen negatif dan netral tetap menjadi acuan penting untuk perbaikan dan peningkatan fasilitas di Novotel Tangerang.

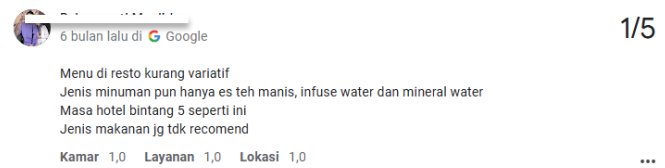
Hasil analisis sentimen pengunjung Novotel Tangerang yang dapat dilihat pada Gambar 7. Hasil tersebut, terlihat bahwa para pengunjung Novotel Tangerang menunjukkan variasi persentase sentimen yang beragam. Secara keseluruhan, sentimen positif dengan kategori layanan dan kamar lebih mendominasi, yang mengindikasikan bahwa persepsi pengunjung terhadap fasilitas hotel di Novotel Tangerang umumnya baik.

sentimen
positif 723
netral 249
negatif 218



Gambar 7. Analisis Sentimen

Meskipun demikian, adanya sentimen negatif menunjukkan bahwa terdapat beberapa masalah yang perlu diperhatikan. Keluhan yang sering muncul adalah terkait bau yang tidak sedap di beberapa kamar, kerusakan *air conditioner* (AC), rasa makanan yang dianggap kurang memuaskan, serta sikap dari *front office* (FO). Selain itu, kehadiran sentimen netral juga menandakan masih adanya ruang untuk memperbaiki layanan hotel. Salah satu contoh ulasan negatif dapat dilihat pada Gambar 4.



Gambar 8. Contoh Ulasan Negatif

Oleh karena itu, pihak Novotel Tangerang perlu fokus pada peningkatan fasilitas dan layanan hotel dengan lebih konsisten. Hal ini mencakup perbaikan kualitas udara dan kebersihan di kamar, memastikan peralatan seperti AC selalu berfungsi dengan baik, serta meningkatkan kualitas rasa pada makanan. Selain itu, peningkatan sikap dan profesionalitas staf *front office* juga penting untuk memastikan pengalaman tamu terjaga dengan baik. Implikasi manajerial yang dapat diambil dari hasil analisis meliputi:

- Mengintegrasikan analisis sentimen ke dalam sistem pemantauan ulasan harian.
- Fokuskan pada area layanan dengan keluhan terbanyak berdasarkan tren sentimen.
- Gunakan hasil analisis untuk mendesain ulang program pelatihan kepada staff.
- Manfaatkan ulasan positif untuk meningkatkan citra Novotel Tangerang melalui media sosial.

IV. KESIMPULAN

Hasil pengujian menunjukkan bahwa algoritma *Logistic Regression* memberikan tingkat akurasi tertinggi di antara algoritma lainnya, dengan akurasi mencapai 94,54% dengan *precision* 100%, *recall* 100%, dan *F1-Score* 96% menggunakan *unigram* (n=1). Analisis sentimen secara umum memperlihatkan pandangan positif dari para pengunjung terhadap fasilitas di Novotel Tangerang. Dari 1.190 data yang dianalisis, terdapat 723 ulasan (60,76%) yang positif, menunjukkan kepuasan pengunjung terhadap fasilitas, kenyamanan kamar, dan suasana hotel. Ini perlu ditingkatkan agar pelanggan tetap merasa puas.

Namun, terdapat pula 218 ulasan (18,32%) yang negatif dan 249 ulasan (20,92%) yang netral, yang mengindikasikan beberapa aspek perlu diperbaiki. Hasil analisis menunjukkan masalah utama pada:

- Beberapa ulasan negative menyoroti adanya bau tidak sedap yang mengurangi kenyamanan tamu. Maka diharuskan melakukan inspeksi kebersihan rutin dan peningkatan kualitas ventilasi di setiap kamar.
- Keluhan pada sistem pendingin udara (AC) sering muncul. Maka jadwalkan *maintenance preventif* dan response pat terhadap laporan kerusakan AC.
- Beberapa tamu merasa makanan kurang memuaskan. Maka lakukan survey preferensi makanan tamu dan evaluasi kualitas dapur secara berkala.
- Sikap kurang profesional dari *Front Office* menjadi perhatian. Maka perlu ditingkatkan pelatihan *soft skill* staff, khususnya di bagian FO, untuk memastikan keramahan dan profesionalitas.

Penelitian lebih lanjut diharapkan dapat menambah jumlah hotel Novotel yang berada diluar Tangerang, analisis sentimen dilakukan dengan berbasis *Deep Learning* atau menggunakan algoritma lain seperti *Support Vector Machine*, menguji representasi n-gram yang lebih fleksibel sesuai dengan karakteristik dataset yang berbeda, menerapkan teknik regularisasi pada model K-NN untuk mengurangi potensi *overfitting* saat menggunakan n-gram yang lebih besar.

DAFTAR PUSTAKA

- [1] "AI vs. machine learning vs. deep learning vs. neural networks: What's the difference?," IBM. Accessed: Nov. 02, 2024. [Online]. Available: <https://www.ibm.com/think/topics/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>
- [2] S. P. Barus and E. Surijadi, "Chatbot with Dialogflow for FAQ Services in Matana University Library," *International Journal of Informatics and Computation (IJICOM)*, vol. 3, no. 2, 2021, doi: 10.35842/ijicom.
- [3] T. P. Nagarhalli, V. Vaze, and N. K. Rana, "Impact of Machine Learning in Natural Language Processing: A Review," in *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 2021, pp. 1529–1534. doi: 10.1109/ICICV50876.2021.9388380.
- [4] H. Tuhutueru, L. P. Refialy, M. Laturake, and S. G. Pattirane, "Tourist Perceptions Through Sentiment Analysis to Support Tourism Development in Maluku Province," *Journal of Applied Informatics and Computing (JAIC)*, vol. 8, no. 1, pp. 119–126, 2024, doi: <https://doi.org/10.30871/jaic.v8i1.6989>.
- [5] S. Kemp, "Digital 2024: Global Overview Report," DataReportal – Global Digital Insights. Accessed: Nov. 02, 2024. [Online]. Available: <https://datareportal.com/reports/digital-2024-global-overview-report>
- [6] Y. Li, Z. Lin, and S. Xiao, "Using social media big data for tourist demand forecasting: A new machine learning analytical approach," *Journal of Digital Economy*, vol. 1, no. 1, pp. 32–43, Jun. 2022, doi: 10.1016/j.jdec.2022.08.006.
- [7] K. Puh and M. B. Babac, "Predicting sentiment and rating of tourist reviews using machine learning," *Journal of Hospitality and Tourism Insights*, vol. 6, no. 3, pp. 1188–1204, Jun. 2023, doi: 10.1108/JHTI-02-2022-0078.
- [8] M. J. Andiana, M. Martanto, and U. Hayati, "Analisis Sentimen Review Hotel Menggunakan Metode Naïve Bayes pada Hotel di Wilayah Kota Cirebon," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 284–264, 2024.
- [9] M. I. Tuku, E. Umar, and A. Adis, "Analisis Sentimen Review Hotel Menggunakan Algoritma Naive Bayes Pada Ella Hotel Tambolaka," *Journal Of Informatics And Business*, vol. 02, no. 03, pp. 346–356, 2024.
- [10] D. F. P. Rachmawan, M. H. P. Swari, and C. A. Putra, "Analisis Sentimen Review Hotel Favehotel Kusumanegara Menggunakan Algoritma Multinomial Naive Bayes," *Seminar Nasional Informatika Bela Negara (SANTIKA)*, vol. 3, pp. 155–159, 2023, doi: <https://doi.org/10.33005/santika.v3i>.
- [11] A. Munir, E. P. Atika, and A. D. Indraswari, "Analisis sentimen pada review hotel menggunakan metode pembobotan dan klasifikasi," *JNANALOKA*, vol. 3, no. 1, pp. 33–38, 2022, doi: 10.36802/jnanaloka.v3-no1-33-38.
- [12] T. Safitri, Y. Umaidah, and I. Maulana, "Analisis Sentimen Pengguna Twitter Terhadap BTS Menggunakan Algoritma Support Vector Machine," *Journal of Applied Informatics and Computing (JAIC)*, vol. 7, no. 1, pp. 34–41, 2023, doi: <https://doi.org/10.30871/jaic.v7i1.5039>.
- [13] J. Setyanto and T. B. Sasongko, "Sentiment Analysis of Sirekap Application Users Using the Support Vector Machine Algorithm," *Journal of Applied Informatics and Computing (JAIC)*, vol. 8, no. 1, pp. 71–76, 2024, doi: <https://doi.org/10.30871/jaic.v8i1.7772>.
- [14] H. P. Doloksaribu and Y. T. Samuel, "Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi," *Jurnal Teknologi Informasi*, vol. 16, no. 1, 2022, doi: 10.47111/JTI.
- [15] R. Merdiansah, S. Siska, and A. A. Ridha, "Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT," *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, vol. 7, no. 1, pp. 221–228, 2024.
- [16] M. D. Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 5, no. 1, pp. 1–7, 2022.
- [17] G. R. Ditami., E. F. Ripanti, and H. Sujaini, "Implementasi Support Vector Machine untuk Analisis Sentimen Terhadap Pengaruh Program Promosi Event Belanja pada Marketplace," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 8, no. 3, pp. 508–516, 2022, doi: <https://doi.org/10.26418/jp.v8i3.56478>.
- [18] D. S. Putri and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi Pospay dengan Algoritma Support Vector Machine," *Jurnal Ilmiah Informatika*, vol. 11, no. 1, pp. 32–40, 2023, doi: <https://doi.org/10.33884/jif.v11i01.6611>.
- [19] Z. Drus and H. Khalid, "Sentiment analysis in social media and its application: Systematic literature review," in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 707–714. doi: 10.1016/j.procs.2019.11.174.
- [20] A. S. Munir, A. B. Saputra, A. Aziz, and M. A. Barata, "Perbandingan Akurasi Algoritma Naive Bayes dan Algoritma Decision Tree dalam Pengklasifikasian Penyakit Kanker Payudara," *Jurnal Ilmiah Informatika Global*, vol. 15, no. 1, pp. 23–29, 2024.
- [21] N. A. Jiana and B. Hartono, "Sentimen Analisa Ulasan Aplikasi Access by KAI pada Google Play Store menggunakan Algoritma K-NN," *Jurnal Media Informatika Budidarma*, vol. 8, no. 3, pp. 1388–1397, Jul. 2024, doi: 10.30865/mib.v8i3.7730.
- [22] K. Kelvin, J. Banjarnahor, E. Indra, M. NK. Nababan, and S. H. Sinurat, "Analisis perbandingan sentimen Corona Virus Disease-2019 (Covid19) pada Twitter Menggunakan Metode Logistic Regression Dan Support Vector Machine (SVM)," *JUSIKOM PRIMA*, vol. 5, no. 2, pp. 47–52, 2022, doi: <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2365>.
- [23] R. N. J. S.Intam, Wulandari, A. A. N. Risal, and D. F. Surianto, "Klasifikasi Mahasiswa Berprestasi Menggunakan Fuzzy C-Means Dan Naive Bayes," *Jurnal Ilmiah Informatika Global*, vol. 15, no. 1, pp. 9–16, 2024, doi: <https://doi.org/10.36982/jiig.v15i1.3666>.
- [24] "Belajar Dasar Visualisasi Data," Dicoding Indonesia. Accessed: Nov. 02, 2024. [Online]. Available: <https://www.dicoding.com/academies/177>
- [25] A. Munna and E. Zuliarso, "Interpretasi model Stacking Ensemble untuk analisis sentimen ulasan aplikasi pinjaman online menggunakan LIME," *Jurnal Teknologi Informasi*, vol. 21, no. 2, pp. 183–196, 2024, doi: <https://doi.org/10.24246/aiti.v21i2.183-196>.