Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

June 2026. Vol. 5. No. 3; e-ISSN: 2808-4519

Sentiment Analysis of Lau Berte Waterfall Using the Naïve Bayes Algorithm

Nadifa Syachdini[1]*, Hotler Manurung[2], Kristina Annatasia Br Sitepu[3]

[1,2,3]information System, STMIK Kaputama
nadifasyc@gmail.com[1]*

Abstract

Social media platforms have become important sources of public opinion, allowing users to share experiences and perceptions regarding tourist destinations. One of the platforms widely used for this purpose is TikTok, where visitors frequently express their opinions through comments on video content. This study aims to implement the Naïve Bayes algorithm for sentiment analysis of TikTok comments related to Lau Berte Waterfall and evaluate its performance through a Streamlit-based web application. A total of 634 comments were collected and processed through several text preprocessing stages, including case folding, cleaning, normalization, tokenization, stopword removal, and stemming. Sentiment labeling was performed using a lexicon-based approach, followed by TF-IDF weighting and Naïve Bayes classification. The developed application provides functionalities for sentiment prediction, performance evaluation, and visualization of sentiment analysis results. Model evaluation in the Streamlit environment achieved an accuracy of 77%, a precision of 80%, a recall of 77%, and an F1-score of 76%. Sentiment distribution analysis revealed that positive sentiment dominated with 64.6% of the comments, followed by negative sentiment at 22.0% and neutral sentiment at 13.4%. These findings indicate that the proposed system is capable of effectively classifying public sentiment and providing valuable insights into visitor perceptions of Lau Berte Waterfall, which may support tourism evaluation and development efforts.

***Keywords**: Lau Berte Waterfall, Naïve Bayes, Sentiment Analysis, Streamlit, TikTok*

1. Introduction

Social media platforms have become one of the primary channels through which people share opinions, experiences, and recommendations about tourist destinations. The increasing amount of user-generated content provides valuable information that can be analyzed to understand public perceptions and support tourism development [1]. Among various social media platforms, TikTok enables users to express their travel experiences through comments on video content, creating a rich source of textual data.

Lau Berte Waterfall, located in Langkat Regency, North Sumatra, is one of the natural tourist attractions frequently discussed on TikTok. Visitors often share their impressions regarding the scenery, travel experience, and accessibility through comments. However, the large volume of textual data makes manual analysis inefficient and time-consuming. Therefore, an automated approach is required to process and classify public opinions effectively [2].

Sentiment analysis is a Natural Language Processing (NLP) technique used to identify and categorize opinions expressed in textual data into positive, negative, or neutral sentiments. In tourism studies, sentiment analysis can help evaluate visitor perceptions and identify aspects that influence public satisfaction. Among various classification algorithms, the Naïve Bayes algorithm is widely adopted because of its computational efficiency and its ability to produce reliable performance in text classification tasks [3].

[4] applied the Naïve Bayes algorithm to analyze visitor sentiment toward Kejawanan Beach using 998 reviews and achieved an accuracy of 78%, demonstrating that the algorithm can effectively identify public perceptions and provide useful information for tourism development.

Based on this background, this study implements the Naïve Bayes algorithm to classify TikTok user comments related to Lau Berte Waterfall. The collected comments undergo text preprocessing, lexicon-based sentiment labeling, and TF-IDF feature weighting before classification. Furthermore, the developed model is deployed in a Streamlit-based web application to facilitate sentiment analysis and visualize the classification results. The model performance is evaluated using accuracy, precision, recall, and F1-score metrics to assess its effectiveness in classifying user sentiments.

2. Research Methodology

This study adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, which consists of business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The overall research workflow includes data collection, text preprocessing, sentiment labeling, TF-IDF feature weighting, Naïve Bayes classification, deployment using Streamlit, and model evaluation.

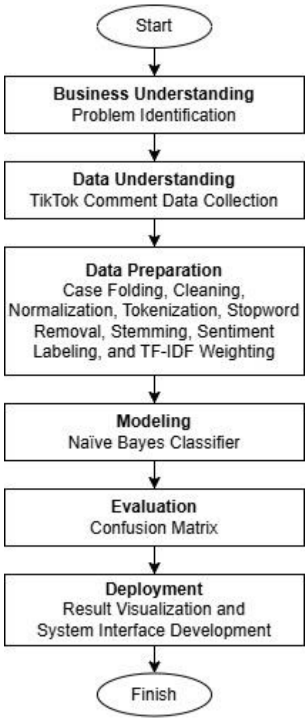

Fig. 1: Methodology

2.1. Data Collection

The data used in this study consisted of user comments related to Lau Berte Waterfall collected from the TikTok platform. Data collection was performed using Apify, a web scraping service that enables automated extraction of publicly available TikTok comments. The collected data were exported in CSV format and subsequently used as the dataset for sentiment analysis. A total of 634 comments were obtained and processed in this study. These comments represent public opinions and experiences regarding Lau Berte Waterfall, providing the primary data source for sentiment classification.

Table. 1: Sample of Data Collected from TikTok Comments

Username	Comment
TikTok User	cantikkk banget emang
TikTok User	aku org langkat kok baru tau yach??
TikTok User	woi ayokk
TikTok User	serius ada kek gini dilangkat??
TikTok User	airnya jernih dn dingin bngt itu

2.2. Text Preprocessing

Before the sentiment classification process, the collected TikTok comments were subjected to a series of text preprocessing steps to improve data quality and ensure consistency. This stage aims to eliminate irrelevant information and transform the raw textual data into a standardized format suitable for machine learning [5].

The preprocessing pipeline consists of case folding, cleaning, normalization, tokenization, stopword removal, and stemming. Case folding converts all characters into lowercase letters, while cleaning removes URLs, mentions, hashtags, punctuation marks, numbers, emojis, and unnecessary whitespace. Normalization is applied using a predefined normalization dictionary to convert non-standard words into their standard forms and reduce repeated characters. Subsequently, the text is tokenized into individual words, followed by stopword removal to eliminate common words that do not contribute significantly to sentiment classification. Finally, stemming is performed using the Sastrawi library to transform words into their root forms, producing the final text representation used in the subsequent analysis [6].

2.3. Sentiment Labeling

After the text preprocessing stage, sentiment labeling was performed using a lexicon-based approach. In this study, the SentiWords ID lexicon was employed, where each word is associated with a positive or negative sentiment score. The overall sentiment score of a comment was obtained by summing the scores of all matched words in the processed text.

The labeling process also considers negation words, such as *tidak*, *ga*, *gak*, *nggak*, and *tak*, which reverse the polarity of the subsequent sentiment-bearing word. Based on the final sentiment score, each comment was assigned to one of three sentiment categories: positive, negative, or neutral. Comments with a score greater than zero were labeled as positive, scores below zero were labeled as negative, and scores equal to zero were classified as neutral.

2.4. Split Dataset

After the sentiment labeling process, the dataset was divided into training and testing sets before model development. The preprocessed text was used as the input feature, while the sentiment label served as the target variable. An 80:20 ratio was applied to divide the dataset into 509 training data and 127 testing data while maintaining the distribution of each sentiment class.

The training data were used to build the Naïve Bayes classification model, whereas the testing data were used to evaluate the model performance on previously unseen data.

2.5. Term Frequency-Inverse Document Frequency Weighting (TF-IDF)

After sentiment labeling, the preprocessed text was transformed into numerical feature vectors using the Term Frequency–Inverse Document Frequency (TF-IDF) method. TF-IDF is a statistical technique that assigns weights to terms based on their frequency within a document and their importance across the entire document collection. As a result, words that frequently appear in a specific document but occur less often in other documents receive higher weights [7].

In this study, the TF-IDF representation was generated using the TfidfVectorizer class from the Scikit-learn library. The resulting feature vectors were subsequently used as input for the Naïve Bayes classifier during the model training and testing processes.

2.6. Naïve Bayes Model Training

After the TF-IDF weighting process, the resulting feature vectors were used to train the Naïve Bayes classification model. The Multinomial Naïve Bayes algorithm was selected because it is widely used for text classification tasks and performs effectively on data represented as term frequencies or TF-IDF weights.

During the training stage, the model learns the relationship between the TF-IDF features and their corresponding sentiment labels from the training dataset. The trained model is then capable of predicting the sentiment category of previously unseen comments by estimating the probability of each sentiment class and assigning the class with the highest probability.

2.7. Model Evaluation

The performance of the Naïve Bayes classifier was evaluated using four commonly adopted evaluation metrics: accuracy, precision, recall, and F1-score. These metrics provide a comprehensive assessment of the model's ability to correctly classify sentiment categories and measure the balance between prediction accuracy and class identification performance.

In addition to these metrics, a confusion matrix was employed to analyze the classification results in greater detail by comparing the predicted labels with the actual sentiment labels. The confusion matrix provides insights into the number of correctly classified instances as well as misclassified samples for each sentiment category, enabling a more detailed interpretation of the model's performance.

3. Result and Discussion

This section presents the implementation results of the proposed sentiment analysis model and discusses its performance in classifying TikTok user comments related to Lau Berte Waterfall. The developed Streamlit-based application is capable of performing sentiment prediction automatically through a series of text preprocessing steps and Naïve Bayes classification. Furthermore, the model performance is evaluated using accuracy, precision, recall, F1-score, and confusion matrix analysis. To provide a more comprehensive interpretation of public opinion, the classification results are also visualized through sentiment distribution and word cloud representations.

3.1. Sentiment Prediction

The developed Streamlit-based application provides an interface for automatically predicting the sentiment of TikTok comments related to Lau Berte Waterfall. After the uploaded dataset is processed, the system generates sentiment classification results using the Naïve Bayes algorithm and displays them in a tabular format.

	final_text	label_aktual	prediksi_sentimen	Log Posterior Positif
0	indah memang lau berte	positif	positif	-0.1451
1	desa adik aku betul jadi pemandu lau berte kabarin kalau mau kesana lagi	netral	netral	-1.2779
2	minggu lalu abis dari sini	netral	positif	-0.8856
3	dimana tempat aku orang langkat belum pernah tahu	netral	netral	-1.0027
4	keren amat	positif	positif	-0.3433
5	jalan uji mental tapi seru banget	positif	positif	-0.2772
6	masuk dalam jauh perjalanan	negatif	negatif	-1.1374
7	izin mau tanya kalau misal kesana datang sendiri bisa ga atau memang han	netral	netral	-0.8833
8	mantap	positif	positif	-0.0907
9	mantap mau kemari lagi kabari	positif	positif	-0.2916

Download Hasil Prediksi

Fig. 2: Label sentiment prediction page with Streamlit

As illustrated in Figure 2, the prediction results include the preprocessed text (final_text), the actual sentiment label (label_aktual), the predicted sentiment label (prediksi_sentimen), and the corresponding log posterior probability. The application also provides a download feature that allows users to save the prediction results for further analysis. This functionality enables users to efficiently review and interpret the sentiment classification outcomes.

3.2. Model Evaluation Metrics

The performance of the proposed Naïve Bayes classifier was evaluated using four evaluation metrics: accuracy, precision, recall, and F1-score. The evaluation results, presented in Figure 3, indicate that the model achieved an accuracy of 77%, a precision of 80%, a recall of 77%, and an F1-score of 76%.

Accuracy	Precision	Recall	F1-Score
0.77	0.80	0.77	0.76

Fig. 3: Sentiment prediction metrics evaluation model page with Streamlit

	Metric	Value
0	Accuracy	0.771654
1	Precision	0.798274
2	Recall	0.771654
3	F1-Score	0.757749

Fig. 4: Evaluation metrics model with Google Colab

Figure 4 presents the performance comparison between the Naïve Bayes model evaluated in Google Colaboratory and its implementation in the Streamlit-based application. The evaluation results show that both environments produced nearly identical performance. In Google Colaboratory, the model achieved an accuracy of 77.72%, a precision of 79.83%, a recall of 77.72%, and an F1-score of 75.77%. Meanwhile, the Streamlit implementation reported an accuracy of 77%, a precision of 80%, a recall of 77%, and an F1-score of 76%. The slight differences are caused by the rounding of decimal values in the application interface.

3.3. Counfusion Matrix

Figure 5 presents the confusion matrix of the Naïve Bayes model, which compares the actual sentiment labels with the predicted sentiment labels across three categories: negative, neutral, and positive. The confusion matrix provides a detailed overview of the model's classification performance and highlights both correctly classified and misclassified instances.

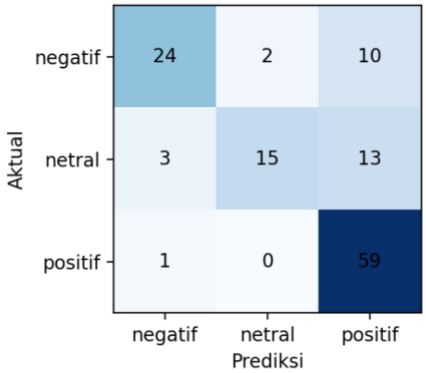

Fig. 5: Confusion Matrix with Streamlit

For the neutral sentiment class, 15 comments were correctly classified as neutral, whereas 3 comments were predicted as negative and 13 comments were predicted as positive. This represents the highest number of misclassifications among the three sentiment categories, suggesting that the model has difficulty distinguishing neutral comments from positive ones.

The best performance was observed in the positive sentiment class, where 59 comments were correctly classified as positive. Only 1 comment was misclassified as negative, and no positive comments were predicted as neutral. These findings demonstrate that the model has a strong ability to recognize positive sentiment patterns within the dataset.

Overall, the confusion matrix indicates that the Naïve Bayes model performs better in classifying positive sentiment than negative and neutral sentiment. This tendency may be influenced by the distribution of the dataset, where positive comments constitute the majority of the data. Consequently, the model learns positive sentiment patterns more effectively, leading to a higher classification accuracy for the positive class.

3.4. Plot Visualization

The Plot Visualization module serves as the primary interface for presenting sentiment analysis results of TikTok comments related to Lau Berte Waterfall. It incorporates sentiment distribution and word cloud visualizations to facilitate the interpretation of public opinions. These visual representations provide valuable insights into overall visitor perceptions and the most frequently discussed topics within each sentiment category.

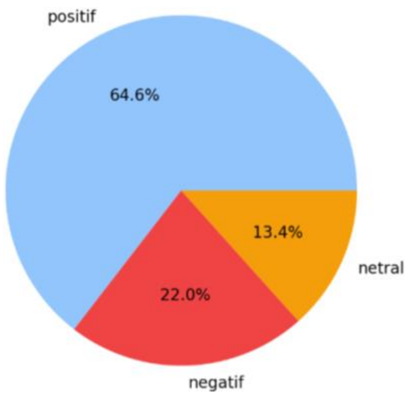

Fig. 6: Pie chart sentiment

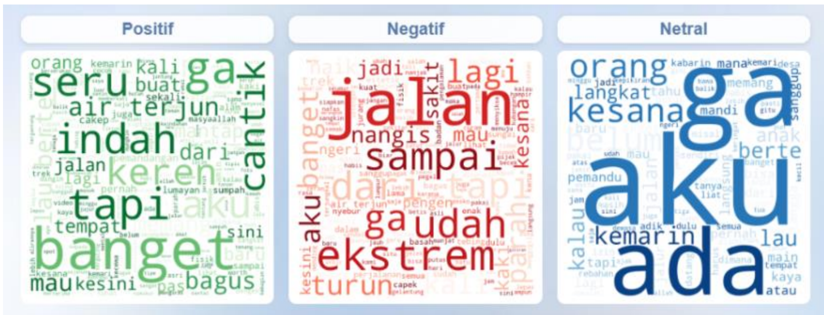

Fig. 7: Wordcloud sentiment positive, neutral, and negative

Figure 6 illustrates that positive sentiment dominates the classification results with 64.6% of the total comments, followed by negative sentiment at 22.0% and neutral sentiment at 13.4%. This finding indicates that most visitors express favorable opinions toward Lau Berte Waterfall.

Figure 7 presents the word clouds for positive, negative, and neutral sentiments. The positive sentiment word cloud highlights terms such as “air terjun,” “indah,” “seru,” “keren,” and “bagus,” reflecting visitors’ appreciation of the natural scenery and overall tourism experience. Meanwhile, the negative sentiment word cloud contains words such as “jalan,” “ekstrem,” “turun,” and “nangis,” suggesting that accessibility and challenging terrain are common concerns among visitors. The neutral sentiment word cloud is dominated by terms such as “aku,” “ada,” “dari,” and “orang,” which generally represent informational statements and personal experiences without expressing strong positive or negative opinions.

4. Conclusion

In conclusion, the proposed sentiment analysis system was successfully implemented using the Naïve Bayes algorithm and deployed through a Streamlit-based web application. The system provides an interactive interface for sentiment prediction, performance evaluation, and visualization of TikTok comments related to Lau Berte Waterfall. The implementation achieved an accuracy of 77%, precision of 80%, recall of 77%, and an F1-score of 76%. Comparatively, the evaluation conducted in Google Colaboratory produced an accuracy of 77.72%, precision of 79.83%, recall of 77.72%, and an F1-score of 75.77%, indicating that the deployed model maintained consistent performance with only minor differences caused by numerical rounding.

The sentiment prediction feature successfully classified comments into positive, negative, and neutral categories. Performance evaluation using confusion matrix analysis showed that the model performed best in identifying positive sentiment, while several neutral and negative comments were misclassified as positive. Furthermore, the visualization results provided valuable insights into visitor perceptions. The sentiment distribution revealed that positive sentiment dominated with 64.6% of the comments, followed by negative sentiment at 22.0% and neutral sentiment at 13.4%. Word cloud analysis indicated that positive comments were primarily associated with the natural beauty and visitor experience, whereas negative comments mainly reflected concerns regarding accessibility and challenging terrain.

Although the proposed system demonstrated satisfactory performance, several limitations remain, particularly in distinguishing neutral and negative sentiments. Future studies may utilize larger and more balanced datasets, explore alternative machine learning and deep learning algorithms, and enhance text preprocessing techniques to further improve sentiment classification performance and system reliability.

References

[1] H. Tuhuteru, L. P. Refialy, M. Laturake, and S. Gildion Pattirane, “Tourist Perceptions Through Sentiment Analysis to Support Tourism Development in Maluku Province,” *J. Appl. Informatics Comput.*, vol. 8, no. 1, p. 119, 2024, [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[2] A. Satria, “Implementasi Text Mining untuk Analisis Review pada Aplikasi Crowdfunding LX dan ST Menggunakan Metode Sentiment Analysis,” *LANCAH J. Inov. Dan Tren*, vol. 2, no. 1, pp. 124-130file:///C:/Users/User/Pictures/PALEMBANG/JU, 2024.

[3] M. Y. Urrochman, H. Asy, and A. Ro, “Aspect-Based Sentiment Analysis of Tumpak Sewu Waterfall Tourist Reviews Using the Naive Bayes Classifier (NBC) Method,” vol. 4, no. October, pp. 18–26, 2025.

[4] R. D. R. Apriliansyah, R. Astuti, W. Prihartono, and R. Hamonangan, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pengunjung Di Pantai Kejawanan,” *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 1, 2025, doi: 10.23960/jitet.v13i1.5774.

[5] A. Haris, T. Tukino, A. Hananto, and F. Nurapriani, “Penerapan Algoritma Naive Bayes Untuk Klasifikasi Sentimen Ulasan Pengguna Aplikasi Pidi Umkm Dengan Pembobotan Fitur Tf-Idf,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 5, pp. 8457–8462, 2025

[6] F. Aziz, A. R. Thaha, and N. Abdan, “Analisis Sentimen Destinasi Wisata Geopark Ciletuh,” *J. Ilmi*, vol. 27, no. 1, pp. 60–68, 2022.

[7] A. Rifa, H. Sujaini, and D. Prawira, “Sentiment Analysis of West Kalimantan Tourist Attractions on Google Maps Using the Naive Bayes Method,” *JEPIN (Jurnal Edukasi dan Penelit. Inform.*, vol. 4, no. 8, pp. 230–235, 2021.