Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2025. Vol. Sec. 5. No. 1; e-ISSN: 2808-4519

Analysis of Visitor Sentiment to Matayangu Waterfall Tourism in Central Sumba Regency Using Naïve Bayes

Apriani May Nggiri[1]*, Fajar Hariadi[2], Novem Berlian Uly[3]

[1,2,3] Informatics Engineering Study Program, Wira Wacana Christian University Sumba
rinyparanda@gmail.com[1*]

Abstract

Advances in digital technology allow the use of social media as a source of public opinion, including in the tourism sector. This study analyzed visitor sentiment towards Matayangu Waterfall in Central Sumba with a combined *lexicon-based* approach and Naive Bayes algorithm. Comment data was taken from TikTok through *the web crawling* method for the period January 1, 2024 to April 30, 2025. Data is processed through *text preprocessing* stages such as *cleaning*, *normalization*, and *stemming*. The initial sentiment label was determined using the SentiWord-ID dictionary, and then converted to numerical forms using TF-IDF before being classified by Naive Bayes. Evaluations are carried out to measure the performance of the model. The results of the evaluation showed that the Naïve Bayes method had an accuracy of 75.81% in predicting sentiment. Most of the comments exhibited a neutral sentiment (61.69%), with positive comments accounting for 20.67% and negative ones for 17.63%.

Keywords: sentiment analysis, Naive Bayes, lexicon-based, TF-IDF, TikTok, Matayangu Waterfall, social media, digital tourism

1. Introduction

The tourism sector is a promising field to be developed as one of the significant sources of regional income[1]. Central Sumba Regency, located on Sumba Island, East Nusa Tenggara Province, has extraordinary natural wealth and has not been fully explored for its potential. One of the leading natural tourist destinations owned by this district is Matayangu Waterfall. Located in Manurara Village, South Katikutana District, this waterfall offers an exotic view with a natural atmosphere that is still beautiful and far from the crowds. Its natural beauty is a special attraction for domestic and foreign tourists, especially for nature and adventure lovers.

Despite having promising tourism potential, the development and management of tourist destinations such as Matayangu Waterfall still faces challenges, namely low understanding of visitor needs and satisfaction. In this context, it is important to conduct periodic evaluations to improve the quality of services and facilities, so that this destination can compete with other tourist attractions. If not handled properly, existing tourism potential could be in danger of being lost, and this area risks losing opportunities to develop. For this reason, stakeholders need to conduct periodic evaluations to improve the quality of existing services and facilities. One way that can be used is to analyze the sentiments or opinions of visitors which are expressed in the form of reviews on social media, tourism platforms, or response forms.

In today's digital era, visitor reviews and comments on social media are very accessible and can provide a direct picture of the tourist experience, one of which is TikTok. TikTok has become a popular social media platform[2]. TikTok is a very relevant platform in this context due to its visual, interactive, and popular nature, especially among the younger generation [3]. Additionally, TikTok's algorithm allows *viral* content to spread widely in a short period of time, making it a dynamic and representative source of public opinion on tourism perception trends. The platform allows travelers to share photos, videos, and stories, which not only captures the attention of potential visitors but also provides valuable insights for destination managers [4]. Analysis of these opinions can help the management to understand public perception objectively. By leveraging data from this platform, managers can identify trends, preferences, and areas that need improvement. However, given the large amount of text data available and unstructured, a systematic and technology-based approach is needed to process the information efficiently. This is where the role of sentiment analysis becomes very relevant, because it is able to group public opinion into categories such as positive, negative, or neutral, which can be used as an indicator of the quality of a tourist destination.

To conduct sentiment analysis on visitor reviews, it was carried out by combining *lexicon-based* and *Naive Bayes* techniques. The *lexicon-based* method is a dictionary-based approach that does not require training data to be directly used to analyze sentiment based on words that have been given a sentiment score. This approach is considered suitable for data that does not have a label, and is easy to understand and implement [5]. Meanwhile, *Naive Bayes* is a probabilistic-based classification algorithm and is included in a simple but effective machine learning method. This method is able to learn from labeled data and generate a model that can classify new data fairly accurately [6]. By utilizing *the Naive Bayes* algorithm and *the Lexicon Based* technique in sentiment analysis on visitor reviews of Matayangu Waterfall, it is hoped that a better understanding of tourists' perceptions and expectations of the destination can be obtained. The results of this analysis can be the basis for formulating a tourism development strategy that is more on target, both in terms of services, facilities, and promotion. Therefore, this research is important to support data- and technology-based tourism management, as well as encourage the improvement of tourism quality in remote areas such as Central Sumba

2. Library Studies

2.1. Tourism

The tourism sector is a collection of production units from various industries that provide goods and services that are specifically needed by visitors, while economic growth is the process of increasing *output* in the long term [7]. In Indonesia, the tourism sector has become one of the national development priorities, with a focus on the sustainable development of leading tourist destinations.

2.2. Sentiment Analysis

Sentiment analysis is one of the interesting fields to be developed in today's digital era. This is due to the increasingly open space for the public to express their opinions and views through online media in the form of text. Sentiment analysis is a method used to extract opinion data, as well as understand and process textual data automatically to identify sentiments contained in an opinion[8]. This technique allows the grouping of sentiments into categories such as positive or negative. Thus, sentiment analysis aims to identify the expression of a person's feelings or opinions through writing and classify them based on the emotional polarity contained in those opinions.

2.3. Natural Language Preprocessing (NLP)

Natural Language Processing (*NLP*) is a field in computer science and artificial intelligence (*AI)* that focuses on the interaction between computers and human language. The goal of *NLP* is for machines to be able to "understand", "interpret", and "generate" human natural language in a meaningful way [9].

2.4. Lexicon-Based

The *lexicon-based* approach is one of the methods in sentiment analysis that utilizes a list of words or a sentiment lexicon, where each word is given a certain weight or score that indicates the tendency of its sentiment, whether positive, negative, or neutral [10]. These dictionaries are usually created manually by linguists or generated through automated processes from big data (*corpus*). Examples of popular lexicons include SentiWordNet, AFINN, VADER, and for Indonesian there is the SentiWord-ID Dictionary.

2.5. Naïve Bayes

Naive Bayes is a classification method based on *Bayes' theorem,* which utilizes a probability approach in data classification [11]. In the classification of opinions, this method estimates the probability that a text falls into a category based on the words that appear in it. This algorithm is considered "naïve" because it assumes that each word in a document is independent of the other words. Although these assumptions are often unrealistic, *Naive Bayes* has proven to be quite effective and efficient in text classification tasks such as spam *filtering* and sentiment analysis [12]. The advantage of this algorithm is its ability to provide good results even though the training data used is not very large.

The basic *formula of Naive Bayes* is as follows:

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

Where:

a. P(C|X) is a C-class probability given the feature X
b. P(X|C) is the probability that feature X appears in class C
c. P(C) is the initial probability of class C
d. P(X) is the probability of the feature X.

3. Research Methodology

In the study, there are stages of analysis consisting of seven stages which can be seen in the following figure.

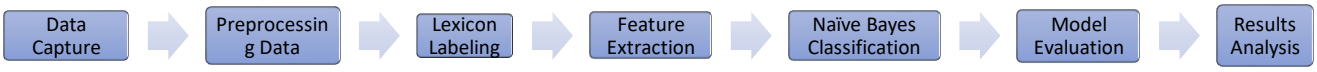

Figure 1: Stages of Analysis

3.1. Data Capture

Data retrieval in this stage is carried out by accessing *files* that have been stored on *Google Drive* through the *Google Colab platform.* The data is *crawled data* from *the Python library.* This process uses *a script* to connect *Google Colab* with a *Google Drive account,* so that the data can be downloaded or accessed directly. The data is *crawled data* from *the Python library.* The data taken will then be used in the next stage, namely *the preprocessing process,* to prepare the data before further analysis.

3.2. Preprocessing Data

Data preprocessing is the initial stage in text processing that aims to clean and prepare data to be ready for use by classification models. This stage is especially important because review data from social media is generally unstructured, containing non-standard language, symbols, or abbreviations. At this stage there are 5 processes, namely cleaning, case folding, normalization, tokenization, stopword removal and stemming.

3.3. Lexicon Labeling

This stage aims to give an initial label (positive, negative, neutral) to the review data based on *the Lexicon Based approach,* before being used in the training of *the Naive Bayes* model. The *lexicon dictionary* used in this study is the SentiWord-ID sentiment dictionary. This dictionary consists of words that commonly appear in the context of sentiment analysis. Dictionaries are divided into two categories, namely positive sentiments and negative sentiments, where each word is given its own weight.

3.4. Feature Extraction

After the lexicon labeling process, features are extracted from visitor comments that have gone through the *Preprocessing* stage. The method used for feature extraction is TF-IDF (Term Frequency - Inverse Document Frequency). This process aims to transform each document (comment) into a numerical representation that reflects the importance of a word in the document relative to the entire corpus TF-IDF formula:

$$TF\text{-}IDF(t, d) = TF(t, d) \times IDF(t) \quad (2)$$

Where:
TF(t,d) = how often the word t appears in the review d
IDF(t) = rare occurrence *of the t-word* throughout the review

3.5. Naïve Bayes Classification

After the review data is converted into numerical form using the TF-IDF method, the next process is to perform a classification using the *Naive Bayes* algorithm. *The Naive Bayes* model calculates the probabilities of each class based on the distribution of TF-IDF values from the words contained in the training data. The probability of a class is calculated by considering the *probability of prior* and the probability of the occurrence of words in each class. The trained model is then used to classify the test data. Naive *Bayes classification* begins with the separation of training data and test data. The processed data will be divided into 80% as training data to train the model and 20% as test data to test the model. For example, the training data and test data are taken from examples of lexicon labeling results, namely d1, d2, and d4 reviews as training data and d3 as test data.

3.6. Model Evaluation

Once the *Naive Bayes* classification is performed, the next step is to evaluate the model's performance to find out how well the model is able to correctly classify sentiment. This evaluation was carried out using standard metrics in the classification model test, namely *accuracy, precision* and *recall.* The calculation was made based on the model's prediction results on the test data and compared with the initial label. Through this evaluation, it can be seen how effective the *Naive Bayes* model is in classifying review sentiment into positive, negative and neutral classes. The results of this evaluation are the basis for assessing the work and reliability of the model for real data use. Because the data used as an example is still limited, another example is used for manual calculation.

3.7. Results Analysis

The results analysis stage was carried out after the sentiment classification model was successfully created and applied to visitor comment data related to Matayangu Waterfall. The analysis is carried out by referring to the results of sentiment predictions displayed through data visualization.

4. Results and Discussion

4.1. Data Capture

Data collection is carried out through the crawling method of the TikTok platform, by capturing visitor comments within a certain period of time. This crawling process uses Python-based scripts that rely on libraries and frameworks such as TikTok Scraper to find video links

based on *hashtags*, as well as TikTok Comments Scraper to extract comments from the video. The crawled data is then stored in a structured CSV format in *Google Drive*. The following are the results of data collection which can be seen in Figure 4.1.

Figure 2: Dataset in CSV form

4.2. Preprocessing Data

In the initial stage of data processing, *a preprocessing* process is carried out which aims to clean and prepare the text data so that it is ready to be used in the classification process. This process is very important considering that review data obtained from social media, especially TikTok, is generally unstructured and often contains non-standard language, abbreviations, and irrelevant symbols. By preprocessing, the data becomes cleaner so that it can improve the accuracy of the model in recognizing sentiment patterns. The *preprocessing* process is as follows.

4.2.1. Cleaning

The first process in *preprocessing* is *cleaning*, which is removing all elements in the text that have no informative value or can interfere with the analysis process. These elements include numbers, symbols, punctuation marks (such as periods, commas, exclamation marks, question marks), emoticons, links, and special characters that often appear in social media comments. The main goal of this process is to simplify the text so that it can be analyzed consistently. By cleaning up these elements, the system can focus more on analyzing the meaning of the words that actually contribute to the sentiment they want to identify.

Table 1 : Cleaning

Before *Cleaning*	**After *Cleaning***
🙏😁 A visit to the Temple of the Blessed Virgin Mary is coming to an end. Thank you so much for coming to visit me in S.C🙏.	A visit to the Temple of the Blessed Virgin Mary is coming to an end. Thank you so much for coming to visit me in S.C.

4.2.2. Case Folding

After the cleaning process, *case folding* is carried out, which is changing all letters in the text to lowercase (*lowercase*). This aims to standardize the data, so that words like "Good" and "good" are considered the same word by the system. In the computational process, the system distinguishes uppercase and lowercase letters, so without *case folding,* the model can mistakenly assume the same words as different words. *Case folding* is especially important in the context of matching words with dictionaries or in the calculation of word frequencies during the feature extraction process. This standardization helps to lower the complexity of the data and improve processing consistency.

Table 2: Case Folding Results

Before *Case Folding*	After *Case Folding*
A visit to the Temple of the Blessed Virgin Mary is coming to an end. Thank you so much for coming to visit me in S.C.	A visit to the Temple of the Blessed Virgin Mary is coming to an end. Thank you so much for coming to visit me in S.C.

4.2.3. Normalization

Normalization is the process of converting non-standard or non-standard words into the form of standard words according to Indonesian rules. On social media, there is often the use of slang, abbreviations, or phonetic forms of writing such as "gk" (no), "bgt" (really), "fit bgt" (very suitable), or "rame" (crowded). Words like these would not be recognized by the lexicon dictionary if they were not normalized. The normalization process aims to ensure that all words are semantically processable and match the entries in the sentiment dictionary. In the context of this study, normalization is very important because the *lexicon* approach used relies on matching words with sentiment dictionaries that contain only standard words.

Table 3: Normalization Results

Sebelum *Normalization*	Sesudah *Normalization*
A visit to the Temple of the Blessed Virgin Mary is coming to an end. 400Entim Suh came to visit Matayangu K.K.	Thank you brother for coming to visit Matayangu Waterfall don't regret coming Thank you for coming to visit in Matayangu sister

4.2.4. Tokenization

Once the text has been cleaned up and normalized, the next stage is tokenization, which is the process of breaking down the text into smaller parts called tokens, usually in word form. The goal of tokenization is so that each word can be analyzed individually. Tokens are the basic unit in text processing that will be used in subsequent stages such as *lexicon* matching, word frequency counting, and feature extraction using TF-IDF. By dividing the text into tokens, 401the entim can more precisely analyze the contribution of each word to 401 the review entity.

Table 4: Tokenization Results

Sebelum *Tokenization*	After *Tokenization*
Thank you brother for coming to visit Matayangu Waterfall don't regret coming	['receive', 'love', 'sister', 'already', 'come', 'visit', 'in', 'water', 'plunge', 'matayangu', 'don't', 'repent', 'come']
Thank you for coming to visit in Matayangu sister	['receive', 'love', 'already', 'come', 'ber', 'visit', 'di', 'matayangu', 'sister']

4.2.5. Stopword Removal

The next stage is to remove *stopwords* or common words that have no meaning, sentiment, or significance in the context of the analysis. Words like "and", "which", "is", "in", "that", and "for" belong to the category of *stopwords*. These words, although they appear frequently, do not contribute to the identification of opinions or emotions. By removing *the stopwords,* the data becomes more focused on only words that are relevant and contain the meaning of the opinion, such as "good", "crowded", "expensive", or "beautiful". In addition*, stopword removal* helps reduce the dimensions of the data and speed up the computation process at the classification stage.

Table 5: Stopword Removal Results

Sebelum *stopword removal*	After *stopword removal*
['receive', 'love', 'sister', 'already', 'come', 'visit', 'in', 'water', 'plunge', 'matayangu', 'don't', 'repent', 'come']	['receive', 'love', 'sister', 'visit', 'water', 'plunge', 'matayangu', 'don't', 'repent']
['receive', 'love', 'already', 'come', 'ber', 'visit', 'di', 'matayangu', 'sister']	['receive', 'love', 'ber', 'visit', 'matayangu', 'sister']

4.2.6. Vote

The last stage in *preprocessing* is stemming, which is the process of changing a derivative word to its root *form*. For example, the words "play", "play", "play" will be returned to the basic form of "play". The *stemming* process in this study uses the Literary library which has been widely used in Indonesian language processing. *Stemming* is important so that the system can identify the basic meaning of each word and avoid recording the same word in different derivative forms as separate entities. Thus, words such as "clean", "clean", and "clean" can be recognized as the same concept, i.e. "clean". *The source code* for this process can be seen in the image below.

Table 6: Voting Results

Before *voting*	After *voting*
['receive', 'love', 'sister', 'visit', 'water', 'plunge', 'matayangu', 'don't', 'repent']	Thank you so much for visiting the waterfall, don't forget to check it out.
['receive', 'love', 'ber', 'visit', 'matayangu', 'sister']	Thank you for visiting my sister's eyes.

4.3. Lexicon Labeling

After the review data from social media has passed the *preprocessing* stage and is cleaned of irrelevant elements, the next process is to label the data using *a lexicon-based* approach. This stage is very important because it serves as an initial determinant of sentiment polarity in each comment before the data is trained and tested using *the Naive Bayes* classification algorithm. At this stage, any pre-processed comments or reviews will be analyzed based on the words contained in them. The labeling process is carried out by matching each *word of stemming* results with a sentiment lexicon, in this case using SentiWord-ID, which is an Indonesian sentiment dictionary that has been categorized based on its polarity value. This dictionary contains a list of words that have been given a score or sentiment weight, whether positive, negative, or neutral.

Table 7: Lexicon Labeling Results

Reviews	Score	Label
Thank you so much for visiting the waterfall, don't forget to check it out.	-1	Negative
Thank you for visiting my sister's eyes.	+1	Positive

4.4. Feature Extraction

After the comment data from social media users has been successfully processed and labeled with sentiment, the next step is to extract features. The purpose of this stage is to convert the preprocessed text data into a numerical representation form that can be understood and processed by the classification algorithm. One of the methods used in this study is TF-IDF (*Term Frequency-Inverse Document Frequency*). The TF-IDF method is a word weighting technique that takes into account how often a word appears in a document (TF) and how rarely the word appears in the entire document set (IDF). In other words, TF-IDF gives high weight to words that often appear in one review, but rarely appear in other reviews. This allows the system to filter out words that actually have significance or are unique to a particular document.

Table 8: Word TF-IDF

word	TF-IDF Results
don't	0.508
handsome	1
good	0,717

4.5. Naïve Bayes Classification

Once the text features are extracted using TF-IDF, the data is classified with a Naive Bayes algorithm that works on the principle of probability assuming independence between words. The dataset was divided 80% for training and 20% for testing. The model calculates the probability of words per sentiment class with Laplace Smoothing to avoid a zero value. When it receives new comments, the system selects the class with the highest probability as a result of classification: positive, negative, or neutral.

4.6. Model Evaluation

After the classification model is completed using the *Naive Bayes* algorithm, the final step in the analysis process is to evaluate the model. This evaluation aims to assess the performance of the model in correctly classifying the data, as well as to find out how accurate the model is in recognizing the sentiment of the analyzed comments. One of the evaluation methods used in this study is the *confusion matrix*, which is used to compare the model's prediction results against the actual labels in each sentiment class, namely positive, neutral, and negative. This matrix shows the number of true and false predictions for each class, helping to identify how well the model recognizes each category.

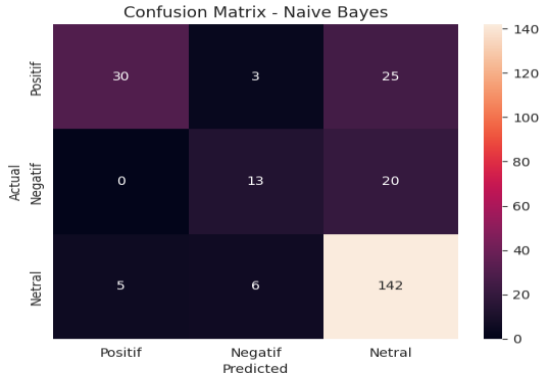

Figure 3: Confusion Matrix

Next, a calculation of model performance was carried out using evaluation metrics.

```
Akurasi: 0.7581967213114754

Classification Report:
              precision    recall  f1-score   support

     Negatif       0.59      0.39      0.47        33
      Netral       0.76      0.93      0.84       153
     Positif       0.86      0.52      0.65        58

    accuracy                           0.76       244
```

Figure 4: Model Evaluation

4.7. Results Analysis

The analysis of classification results is not only carried out through evaluation metrics, but also in the form of data visualization in order to provide a more intuitive and easy-to-understand picture. One form of visualization used is *the bar chart*, which shows the distribution of the number of comments by sentiment category and *word cloud* for each sentiment category, which aims to display the words that appear most often in the comments in each category.

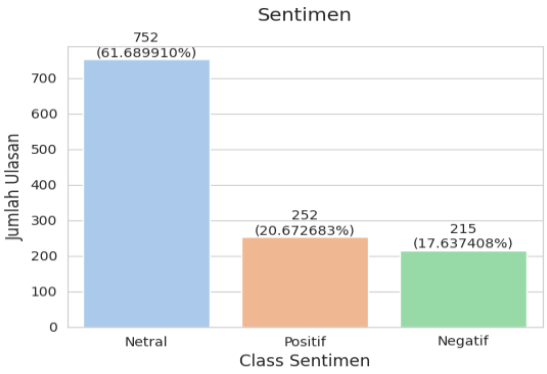

Figure 5: Distribution of Sentiment

Based on the bar chart displayed, it is known that of the total comments analyzed, neutral sentiment dominated with 752 comments (61.69%), followed by positive sentiment as many as 252 comments (20.67%), and negative sentiment as many as 215 comments (17.63%).

Figure 6: Word Cloud Positive Sentiment

On the positive sentiment word cloud, words such as *"beautiful*", *"steady*", *"good"*, *"beautiful"*, *"cool"*, *"thank you*", and *"invitation"* often appear, reflecting the satisfaction, admiration, and positive impression of visitors towards the natural beauty and their tourist experience.

Figure 7: Word Cloud Negative Sentiment

Meanwhile, word cloud negative sentiment shows words such as *"closed"*, *"difficult"*, *"road"*, *"tired"*, and *"child"*, which describe complaints about accessibility, road conditions to the location, and other obstacles during the visit. This is an important input for tourism managers to pay attention to infrastructure aspects and visitor comfort.

Figure 8: Word Cloud Neutral Sentiment

As for word cloud neutral sentiment, words such as *"cool"*, *"waterfall"*, *"sumba"*, *"jalan"*, *"kampung"*, and *"mandi"* dominate. These words tend to be informative and descriptive without a certain emotional content, for example to provide information about the location, physical condition of the place, or the activity being carried out.

5. Conclusion

The Naive Bayes algorithm is quite effective in analyzing the sentiment of comments from visitors to Matayangu Waterfall with an accuracy of 75.81%. The majority of comments were neutral sentiment (61.69%), followed by positive (20.67%) and negative (17.63%). The main obstacle lies in the classification of negative comments due to the unbalanced distribution of data. The use of the SentiWord-ID and TF-IDF lexicons helps the classification process, but accuracy can be improved through manual labeling and comparison with other algorithms such as SVM or Random Forest. These results are useful for tourism management decision-making

Acknowledgement

The author would like to thank Wira Wacana Christian University Sumba and the Informatics Engineering Study Program for the support and facilities provided during this research process. Thank you also to the supervisors and colleagues who have helped in collecting and processing data, so that this research can be completed properly.

References

[1] Pebriana, F., Mulyawan, R., & Sutrisno, B. (2021). Government strategy in tourism development to increase regional income. *The Journal of Government Administration (Janitra)*, *1*(1), 11–22. https://doi.org/10.24198/janitra.v1i1.33023

[2] Apriani, E., Hanif, I. F., Oktavianalisti, F., Monasari, L. D. H., & Winarni, I. (2024). Sentiment analysis of the use of TikTok as a learning medium using the Naïve Bayes Classifier algorithm. *Indonesian Institute of Research and Publications (IRPI)*, *4*(3), 1160–1168. https://journal.irpi.or.id/index.php/malcom

[3] Tarji, M. A., & Wiharjo, R. Y. (2025). The role of TikTok social media in tourism marketing in Indonesia. *Journal of Tourism Science Research*, 1(1), 53–58. https://doi.org/10.33862/jpip.v1i1.578

[4] Regita, G. M. N., Saputra, N. W., Anggara Giri, I. D. G. Y., & Ekasani, K. A. (2024). Gen Z tourists' interest in trendy videos through the TikTok application has increased visits to Trunyan Hill, Bali. *Journal of Tourism Studies*, 6(2), 99–108. https://doi.org/10.51977/jiip.v6i2.1794

[5] Rofiq, A., Nugroho, Y., & Santoso, R. D. (2021). Implementation of the lexicon-based method in the analysis of public opinion sentiment on social media. *RESTI Journal (Systems Engineering and Information Technology)*, 5(2), 276–282. https://doi.org/10.29207/resti.v5i2.2899

[6] Putra, H. D., & Abdurrahman, M. (2022). Sentiment analysis uses the Naïve Bayes method on user review data. *Journal of Information Technology and Systems*, 3(1), 25–32. https://doi.org/10.31294/jtsi.v3i1.12081

[7] Anggarini, D. R. (2021). Contribution of MSMEs in the Tourism Sector to Economic Growth in Lampung Province in 2020. *Scientific Journal of Economics and Business*, *9*(2), 345–355.https://jurnal.unived.ac.id/index.php/er/indexDOI:https://doi.org/10.37676/ekombis.v9i2.1462

[8] Sari, F. V., & Wibowo, A. (2019). Analysis of customer sentiment of online stores Jd.Id using the Naïve Bayes Classifier method based on the conversion of emotion icons. *SIMETRIS Journal*, *10*(2), 681–686.

[9] Epoka, B. E. (2023). Literature Review of Qualitative Data with Natural Language Processing. *Journal of Robotics Spectrum*, *1*, 56–65. https://doi.org/10.53759/9852/jrs202301006

[10] Rufaida, A. S., Permanasari, A. E., & Setiawan, N. A. (2022). *Lexicon-Based Sentiment Analysis Using InSet Dictionary: A Systematic Literature Review*. *Proceedings of ICAE 2022*, Engineering and Technology, 1(1), 363–367.

[11] Leong, J. Y., & Booma, P. M. (2020). Symptom-Based Disease Prediction System Using Machine Learning. *Journal of Theoretical and Applied Information Technology*, *98*(19), 3193–3210. https://doi.org/10.22214/ijraset.2024.59394

[12] Assiva, M. A. (2024). Analysis of Sentiment on Tourism in Grobogan Regency Based on Orange Using Naive Bayes. *Journal Of Social Science Research Volume*, *4*(6), 2351–2359. https://j-innovative.org/index.php/Innovative%0AAnalisis