

Enhancing Aspect-based Sentiment Analysis in Visitor Review using Semantic Similarity

Ni Made Satvika Iswari^{1,*}, Nunik Afriliana², Eddy Muntina Dharma³, Ni Putu Widya Yuniari⁴

^{1,3}*Faculty of Information Technology and Design, Universitas Primakara, Denpasar 80226, Indonesia*

²*Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Tangerang 15810, Indonesia*

⁴*Faculty of Engineering and Planning, Universitas Warmadewa, Denpasar 80239, Indonesia*

(Received: February 19, 2024; Revised: March 25, 2024; Accepted: April 27, 2024; Available online: May 31, 2024)

Abstract

The global economy greatly depends on the tourism industry, which fosters job opportunities and stimulates economic development. With the growing reliance of tourists on online platforms for guidance, evaluations of tourist destinations have gained heightened significance. These assessments, frequently expressed through user-generated content, offer valuable perspectives on customer experiences, viewpoints, and levels of satisfaction. Nevertheless, analyzing and interpreting these reviews can pose difficulties because of the unstructured or semi-structured nature of user-generated content. Conventional sentiment analysis methods might not adequately grasp the intricacies and particular aspects of tourism encounters that user convey in their reviews. The efficacy of sentiment analysis can be augmented by integrating semantic similarity. This study explores methods to enhance aspect-based sentiment analysis within tourism reviews by utilizing semantic similarity approaches. Five aspects have been curated, representing keywords frequently reviewed by visitors to the tourist attraction. These aspects encompass scenery, dusk, surf, amenities, and sanitation. Based on the data analysis, F-Measure values with Semantic Similarity tend to increase for the scenery and dusk aspects. This is because in the sample data used, visitor reviews for the scenery and dusk categories may use other words that are semantically similar. The sample data used for these categories is also quite extensive, resulting in a better classification model for both categories. While it is valuable to analyze user-generated content data from visitor reviews, it's important to consider the limitations and potential biases associated with this data. The classification results per aspect need to be further reviewed in more depth. What aspects lead visitors to give positive reviews will certainly be maintained and even improved by stakeholders. Similarly, for negative review outcomes, it is necessary to investigate more deeply the factors contributing to visitor dissatisfaction so that they can be addressed by stakeholders.

Keywords: Aspect-Based Sentiment Analysis, Visitor Review, Semantic Similarity, User-Generated Content, Tourism

1. Introduction

The tourism sector is vital for the world economy, fostering employment opportunities and driving economic expansion. As more and more tourists rely on online media for information, reviews of tourist attractions have become increasingly important. User-generated content, frequently in the guise of reviews, offers valuable perspectives on customers' experiences, viewpoints, and levels of satisfaction [1], [2]. However, analyzing and interpreting these reviews can be challenging due to the unstructured or semi-structured format of user-generated content. Traditional sentiment analysis techniques may not be sufficient to capture the nuances and specific aspects of tourism experiences that users discuss in their reviews.

To address this challenge, researchers have proposed an approach for improvement of aspect-based sentiment analysis using Semantic Similarity. Aspect-based sentiment analysis aims to extract and analyze opinions and sentiments expressed towards specific aspects or features of a product or service, in this case, tourist attractions. This approach allows for a more detailed understanding of customer feedback and can provide valuable insights for businesses in the tourism industry to improve their offerings and enhance customer satisfaction. Despite the potential benefits of aspect-based sentiment analysis in tourism reviews, there are several challenges that need to be addressed for improved accuracy and effectiveness. One of the main challenges is the identification and extraction of relevant aspects or

*Corresponding author: Ni Made Satvika Iswari (iswari@primakara.ac.id)

 DOI: <https://doi.org/10.47738/jads.v5i2.249>

This is an open access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>).

© Authors retain all copyrights

features from user-generated content. Tourist reviews often contain a variety of information and opinions, making it difficult to identify the specific aspects that users are discussing. Another challenge is the ambiguity and context-dependency of sentiment expressions. Sentiments expressed in tourism reviews can be subjective and context-specific, making it challenging for sentiment analysis algorithms to accurately classify them as positive, negative, or neutral. Furthermore, sentiment analysis in tourism reviews faces the challenge of linguistic variations and cultural nuances.

To analyze sentiment detection in travel reviews, researchers have explored various techniques and approaches. Some of these approaches include:

- 1) Utilizing aspect-based sentiment analysis: Aspect-based sentiment analysis focuses on extracting and analyzing opinions and sentiments towards specific aspects or features of tourist attractions. This approach allows for a more fine-grained analysis of sentiment, enabling businesses to understand the specific aspects that customers love or dislike about their offerings[1].
- 2) Employing machine learning algorithms: Machine learning algorithms can be trained on large datasets of tourism reviews to accurately classify sentiments. These algorithms can learn patterns and relationships between words or phrases and their associated sentiments, improving the accuracy of sentiment detection in travel reviews[2].
- 3) Taking into account linguistic variations and cultural nuances: Tourism reviews often include opinions and feedback from users from different linguistic backgrounds and cultures. Taking into account these variations and nuances is essential for accurate sentiment detection[3].
- 4) Utilizing semantic similarity, the study conducts sentiment analysis on visitor reviews of tourist attractions. However, the algorithm developed in this research has constraints; it can only progress up to the semantic similarity stage and cannot discern which sentences are pertinent to the specific category being analyzed for sentiment. Put differently, the sentiment analysis is evaluated based on an entire review. Therefore, if a review encompasses multiple categories simultaneously, the sentiment analysis will consider the overall sentiment rather than categorizing it individually[4], [5], [6].
- 5) Survey of key sentiment-extraction approaches. This study provides an examination of the mining of feelings or opinion [7].

The effectiveness of sentiment analysis can be further enhanced through the incorporation of semantic similarity. This paper delves into strategies for improving aspect-based sentiment analysis in the context of tourism reviews by leveraging semantic similarity techniques. Through this research, we aim to contribute to the advancement of sentiment analysis in the tourism domain, facilitating more accurate and insightful interpretations of traveler feedback.

2. Literature Review

In this research, sentiment analysis was carried out using a machine learning approach, namely Naive Bayes, Decision Tree, and Maximum Entropy. Apart from that, improvement of sentiment analysis was carried out using Semantic Similarity. Further explanation regarding this approach will be explained in the following subsection.

2.1. Naïve Bayes

Naive Bayes classification is a statistical machine learning technique rooted in Bayes' theorem, assuming feature independence, hence the term "naive" assumption. [6]. Given a set of features and a target variable, Naive Bayes classification calculates the probability of each class label for a new observation based on its feature values. This is done by calculating the conditional probability of each class label given the observed features, and then selecting the class label with the highest probability as the predicted class label for the new observation. Naive Bayes classification is particularly useful when working with large datasets and high-dimensional feature spaces. It is commonly used in text classification tasks, such as spam filtering or sentiment analysis, where the features correspond to words or word frequencies.

Sentiment analysis, or opinion mining, holds significance within natural language processing, entailing the categorization of text documents or social media posts into positive, negative, or neutral sentiments. Several machine learning algorithms have been applied to sentiment analysis, with Naive Bayes classification being one of the most

commonly used techniques [7], [8]. Naive Bayes classification is based on the Bayes theorem and assumes that features are independent of each other. The effectiveness of Naive Bayes classification for sentiment analysis has been extensively studied and evaluated in various research studies. One such study conducted by Zhang, et al. explored the use of Naive Bayes classification for sentiment analysis on social media data. The study found that Naive Bayes classification achieved competitive performance in classifying sentiment on social media posts[9]. Another study by Pang and Lee demonstrated the effectiveness of Naive Bayes classification for sentiment analysis on movie reviews. The researchers compared Naive Bayes classification with other machine learning algorithms and found that Naive Bayes performed well in terms of accuracy [10].

2.2. Decision Tree

Decision tree classification is a popular and powerful machine learning algorithm that is widely used for solving classification problems. It is based on the concept of creating a tree-like model of decisions and their possible consequences. The decision tree algorithm works by partitioning the data into subsets based on different attributes, and making decisions at each node by selecting the attribute that best splits the data and maximizes the information gain or Gini index. The decision tree algorithm is appropriate when the objective is to generate easily understandable rules that can be expressed in natural language. Sentiment analysis, on the other hand, is a text mining technique that involves determining the sentiment or emotion expressed in a piece of text, such as a review or comment. Using decision tree classification in sentiment analysis involves training a decision tree model on a labeled dataset of reviews, where the sentiment is the class label, and the features are the words or phrases in the reviews [11].

Determining the best attribute for splitting data involves calculating the information gain or Gini index at each node. The information gain is a measure of the reduction in entropy or uncertainty achieved by splitting the data based on a particular attribute. To make decisions with decision trees, we start at the root node and follow the branches based on the values of the attributes. We continue down the tree until we reach a leaf node, which represents the predicted class label or sentiment for the input data. The decision tree algorithm determines the most critical attribute at each node, which helps in classifying the data and predicting the outputs.

2.3. Maximum Entropy

Maximum entropy classification is a machine learning algorithm that is widely used in natural language processing tasks, including sentiment analysis. It aims to assign a given input (such as a text document) to one of several predefined categories based on the probability of occurrence. In maximum entropy classification, the principle of maximum entropy is utilized to generate traffic matrices with minimal hidden assumptions about the data. The principle of maximum entropy classification can be applied to sentiment analysis by using keyword analysis and mathematical equations to determine the sentiment of a text. Keyword sentiment analysis involves analyzing the presence and frequency of certain words or phrases that are associated with positive or negative sentiments [12].

2.4. Semantic Similarity

Semantic similarity plays a crucial role in sentiment analysis by measuring the relatedness or similarity between texts based on their meanings. Researchers have been actively investigating various methods and techniques to incorporate semantic similarity into sentiment analysis models. Using semantic similarity approach, it is not only considering exact keyword matches but also identifies texts that have similar semantics to the sentiment being analyzed. This approach is particularly useful in cases where the sentiment category of a text is not explicitly mentioned. Several studies have explored the application of semantic similarity in sentiment analysis [13]. In a research endeavor, scholars investigated the utilization of lexical, syntactic, semantic, and discourse aspects of language analysis across various tasks, including automatic and semi-automatic text indexing, text retrieval, text summarization, generating thesauri from text corpora, and conceptual information retrieval. Additionally, the researchers deliberated on their own efforts regarding the application of syntactic analysis for matching and ranking phrases, employing structured text representations.

By considering the semantic relationships between different terms in customer reviews, sentiment analysis models can gain a deeper understanding of the sentiment expressed. This can help overcome the limitations of traditional sentiment analysis methods that only consider lexical-level information and fail to capture the nuances and subtleties of sentiment. Additionally, semantic similarity measures can enhance the effectiveness of text clustering techniques for sentiment

analysis. By clustering customer reviews based on semantic similarity, sentiment analysis models can identify, and group together reviews that discuss similar aspects or topics about the consumer experience. This allows for a more granular and detailed analysis of sentiment, preventing the loss of important insights that can occur when assigning a review to a single category based on lexical-level information alone [14], [15].

3. Methodology

This research proceeded through four consecutive stages: gathering data, labeling data, analyzing data, and evaluating & interpreting data, as illustrated in figure 1. Initially, user-generated content (UGC) data was gathered from the Google Maps website, to be processed in subsequent phases. User-generated content encompasses various types of content, including reviews, comments, photographs, videos, or posts on social media platforms, which are generated and shared by users or consumers of a product or service. This content is valuable for organizations because it provides authentic and unbiased information about their products or services. Moreover, user-generated content enables organizations to glean insights into customers' sentiments, inclinations, and encounters. This data can be leveraged to enhance products or services, pinpoint areas for enhancement, and customize marketing strategies to align more effectively with customer requirements.

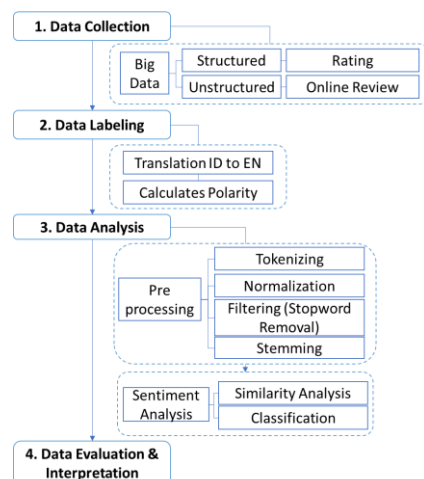


Figure 1. Research Methodology

Figure 2 (left) displays visitor reviews sourced from Google Maps, categorized under various headings such as massage, mall, traffic jam, tattoo, and so forth. Conversely, figure 2 (right) illustrates multiple user reviews categorized by word frequency within each category. The word occurrences must match exactly with the category name or its translated equivalent. For instance, under the massage category, the displayed reviews must contain the exact word "massage".

The second stage of this study is data labeling. Data labeling is a crucial step in sentiment analysis, which involves assigning sentiment labels (such as positive, negative, or neutral) to a given set of textual data. To label the data, the sentiment analysis task usually requires human annotators to read and analyze each text and determine its sentiment based on predefined guidelines or criteria. The process of data labeling in sentiment analysis starts by providing a representative sample of labeled data to human annotators. One popular tool for data labeling in Python is TextBlob (<https://textblob.readthedocs.io/>). TextBlob is a powerful library in Python that provides capabilities for text processing, including data labeling. To label data using TextBlob, a representative sample of labeled data are needed. This labeled data acts as a training set for TextBlob to learn patterns and relationships between the text and the labels.

3.1. Data Collection and Labeling

This study utilizes user-generated content data in the form of visitor reviews on Google Maps for Kuta Beach, Bali, as illustrated in figure 2 as an example. The review data used consists of Indonesian, English, and a combination of both languages. A total of 398 visitor review data were collected. After data collection, the next step is to label the data. The data annotation procedure utilizes the TextBlob Library, accessible at <https://pypi.org/project/textblob/>. TextBlob

represents a Python library employed for analyzing textual data. Subsequently, the visitor review data undergoes categorization according to sentiment analysis labels, which include positive, negative, or neutral classifications.

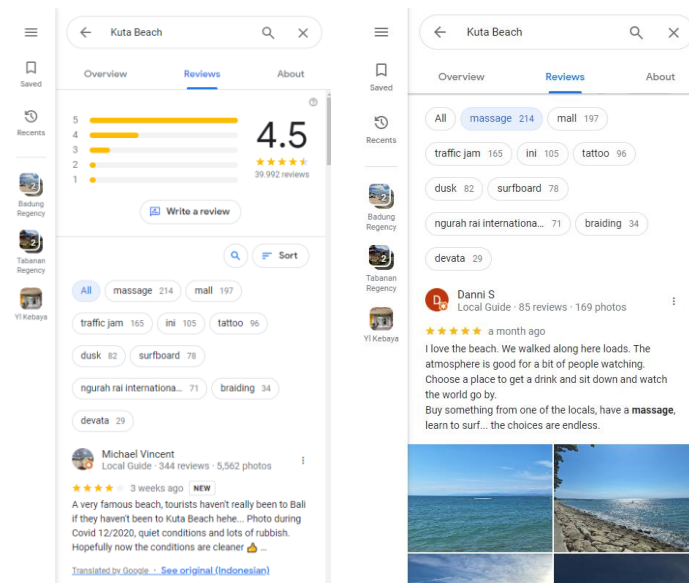


Figure 2. Google Maps visitor reviews categorized into various groups (on the left), and the categorization of reviews centered on the term "massage" (on the right).

Figure 3 illustrates an example of data labeling. At this stage, 5 aspects have been curated, representing keywords frequently reviewed by visitors to the tourist attraction. These aspects encompass scenery, dusk, surf, amenities, and sanitation. The semantic similarity analysis undertaken will generate reviews pertaining to the aforementioned 5 aspects. The utilized sentiment labels include "positive," "negative," and "neutral." Moreover, for review data lacking pre-defined aspects, the sentiment classification will be denoted as "none."

User-generated Content				
<p>"The most popular beach in Bali for domestic even international tourist. Kuta Beach also a good place for you all who wanna learn to surf. There are so many things that locals will offer you in this beach. Massage, surfing lesson, food & beverages, souvenir, and many more. Just be careful when you come to this beach at the first time. I suggest you to act like 'it's not your first time'. See u on my next review."</p>				
Aspect				
Scenery	Dusk	Surf	Amenities	Sanitation
none	none	positive	positive	none

Figure 3. Example of Data Labeling

3.2. Semantic Similarity Analysis

The Semantic Similarity method was employed to detect review results that aligned not only with the precise keywords of the categories but also identified reviews that, while not explicitly containing identical keywords, conveyed comparable meanings to the categories. In computer science, semantics refers to the logical interpretation of valid strings specified by programming languages, contrasting with syntax, which governs the structure of computer programs. Semantic analysis entails deriving significance from text, enabling computers to comprehend and interpret phrases, paragraphs, or entire documents by discerning grammatical structures and word associations within their contexts. Employing semantic analysis can benefit businesses in various ways, including gaining comprehensive insights from customer reviews. Particularly in the analysis of visitor reviews, semantic analysis can provide thorough information about tourist feedback, which is essential for reviews that do not explicitly mention the category of a tourist spot [16], [17]. The semantic similarity process utilizes the SBERT and Cosine Similarity algorithms. SBERT enhances BERT's performance with a different architecture, enabling quicker computation of cosine similarity. For instance,

while it might take 65 hours to search for sentence similarities using BERT, SBERT can accomplish the task in just 5 seconds [18].

Pre-processing is an initial step in data mining techniques to refine raw data for further analysis. This involves several stages, including tokenization, normalization to convert slang or colloquial words into proper language, filtering to remove stop words, and stemming to simplify words by removing affixes. Following pre-processing, the data is prepared for semantic analysis [19], [20].

3.3. Modeling

At this stage, the model is built using machine learning models, namely Naive Bayes, Decision Tree, and Max Entropy. For the machine learning model that has been built, the evaluation of the model was carried out by comparing the values of accuracy, precision, recall, and F1-Score of machine learning.

3.3.1. Machine Learning

After preprocessing the data, the next stage is feature extraction. Feature extraction involves transforming the text data into numerical features that can be used as input for the machine learning model. This can be done using techniques such as bag of words, where each word in the text is represented as a feature, and the frequency or presence of the word in the text is used as the value for that feature. Once the features have been extracted, the dataset is divided into a training set and a test set. The training set is used to train the machine learning model (Naïve Bayes, Decision Tree, and Max Entropy) on the labeled data, while the test set is used to evaluate the performance of the model on unseen data. During the training phase, the machine learning algorithm calculates the probabilities of each feature occurring in each class (positive, negative, or neutral). These probabilities are then used to classify new, unseen text data based on the likelihood of each feature occurring given the class.

3.3.2. Evaluation

This research assessed the model by examining the accuracy, precision, recall, and F1-Score of machine learning. The evaluation utilized the Cross Validation technique, which divides the data into two sets: one for model training and the other for evaluating the model's predictive performance, aiding in selecting the most suitable model for prediction.

3.3.2.1. Accuracy

Accuracy represents the proportion of accurately predicted data compared to the entire dataset. The formula for accuracy is:

$$\text{accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

Note:

- TP (True Positives) : The number of correctly predicted positive cases.
- TN (True Negatives) : The number of correctly predicted negative cases.
- FP (False Positives) : The number of incorrectly predicted positive cases.
- FN (False Negatives) : The number of incorrectly predicted negative cases.

3.3.2.2. Precision

Precision is a measure of the proportion of correctly predicted positive instances relative to the total number of instances predicted as positive. The formula for precision is:

$$\text{precision} = \frac{TP}{TP + FP} \quad (2)$$

3.3.2.3. Recall

Recall measures the proportion of correctly identified positive instances relative to the total number of actual positive instances. The formula for recall is:

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

3.3.2.4. F1-Score

The F1-Score represents the balanced average of precision and recall values. The formula for F1-Score is:

$$\text{F1 Score} = \frac{2 \times (\text{recall} \times \text{precision})}{\text{recall} + \text{precision}} \quad (4)$$

4. Result and Discussion

Based on the research stages that have been conducted, this study has produced sentiment analysis based on aspect analysis improvised with a semantic similarity approach. This chapter will discuss the comparison of the classification models that have been generated, namely the model without using the semantic similarity approach and the model using the semantic similarity approach.

4.1. UGC Sentiment Analysis

Figure 4 illustrates the comparison of confusion metrics for sentiment analysis classification results based solely on aspect compared to when semantic similarity improvement is added. Based on the data, F-Measure values with Semantic Similarity tend to increase for the scenery and dusk aspects. This is because in the sample data used, visitor reviews for the scenery and dusk categories may use other words that are semantically similar. The sample data used for these categories is also quite extensive, resulting in a better classification model for both categories. Meanwhile, F-Measure values with Semantic Similarity tend to decrease for the surf, amenities, and sanitation aspects. This is contrary to the previous two aspects, as in the sample data used, visitor reviews for the surf, amenities, and sanitation categories tend to use the same words as the category names and do not use other similar words. Additionally, the sample data used for these three categories is relatively smaller, resulting in classification models that are not better than the previous two categories.

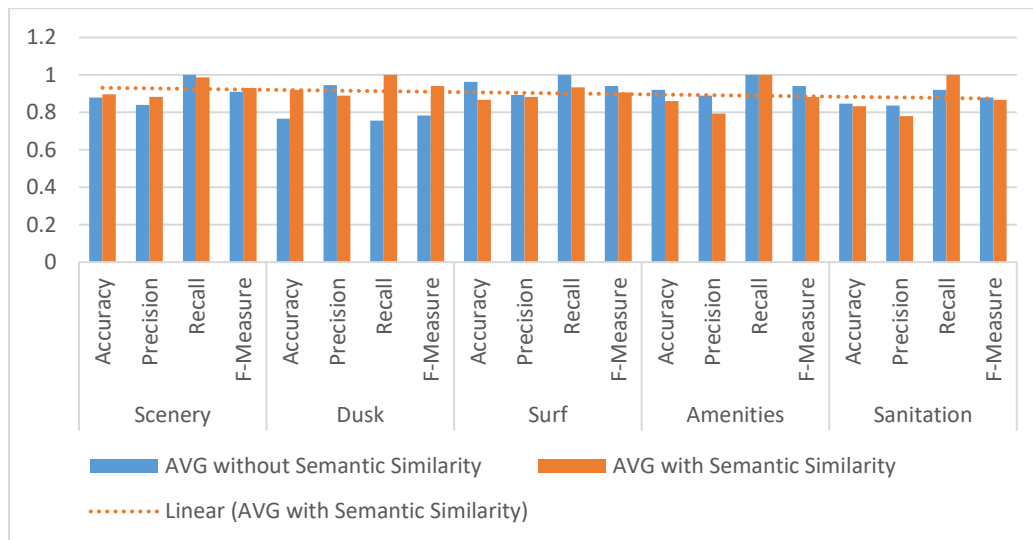


Figure 4. Comparison of Confusion Matrix

Figure 5 presents the results of sentiment analysis classification for the scenery, dusk, surf, amenities, and sanitation categories using the Naive Bayes, Decision Tree, and Max Ent models with Semantic Similarity improvement.

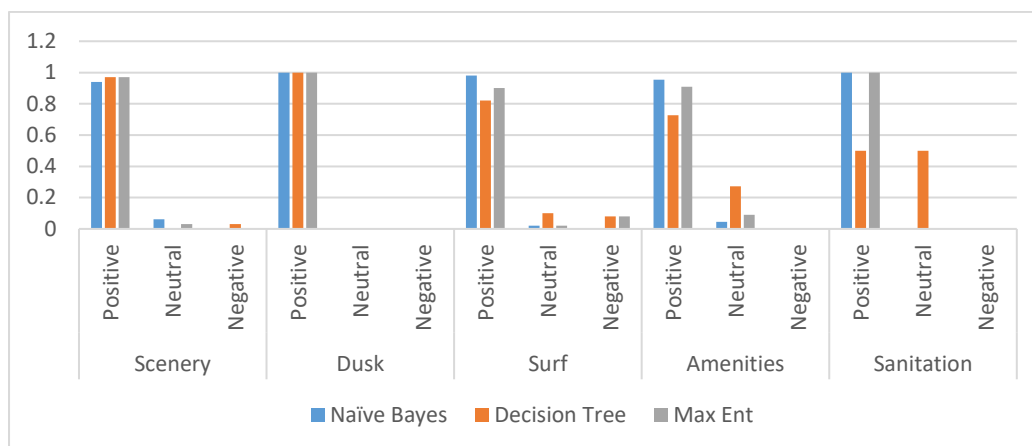


Figure 5. Review Analysis

Based on the classification data, it is found that the scenery, dusk, surf, amenities, and sanitation aspects generally receive positive reviews from visitors. This indicates that overall, visitors to Kuta Beach, Bali, are satisfied with their visit. However, there are also a few negative reviews for the scenery and surf aspects. This needs to be addressed by stakeholders in order to improve visitor satisfaction in the future. Table 1 show some examples of sentiment analysis result.

Table 1. Examples of Sentiment Analysis Result

Visitor Review	Aspect				
	Scenery	Dusk	Surf	Amenities	Sanitation
Trash everywhere.... In huge piles across the beach, lined along the shores as far as the eye can see, and in large floating patches. If you need an eye-opening experiencing about pollution, maybe go to this beach. The waves might be decent for beginner surfing but couldn't get over all the plastic and trash in the water. Do not recommend and would avoid.	none	none	negative	none	negative
Beautiful beach and good for surfers, there many shops who offer training and support with surfing boards and guides. There are a lot of shops also available for drinks and snacks, they provide chairs and shades.	positive	none	positive	neutral	none

4.2. Evaluation

In this subsection, the evaluation results conducted for the three classification models, namely Naive Bayes, Decision Tree, and Maximum Entropy, are explained. Table 2 shows the evaluation results of the models without the Semantic Similarity approach, while table 3 shows the evaluation results of the models with the Semantic Similarity approach. On average, the F-Measure value of Naive Bayes is the highest for classification without Semantic Similarity. Meanwhile, the F-Measure value of Decision Tree is the highest for classification with Semantic Similarity.

Table 2. Evaluation Result without Semantic Similarity

Aspect		Naïve Bayes	Decision Tree	Max Ent
Scenery	Accuracy	0.88	0.88	0.88
	Precision	0.84	0.84	0.84
	Recall	1.00	1.00	1.00
	F-Measure	0.91	0.91	0.91
Dusk	Accuracy	0.93	0.44	0.93
	Precision	0.92	1.00	0.92
	Recall	1.00	0.27	1.00

	F-Measure	0.96	0.43	0.96
Surf	Accuracy	0.99	0.96	0.94
	Precision	0.93	0.88	0.87
	Recall	1.00	1.00	1.00
	F-Measure	0.96	0.93	0.93
Amenities	Accuracy	0.92	0.92	0.92
	Precision	0.89	0.89	0.89
	Recall	1.00	1.00	1.00
	F-Measure	0.94	0.94	0.94
Sanitation	Accuracy	0.86	0.82	0.86
	Precision	0.84	0.83	0.84
	Recall	0.94	0.88	0.94
	F-Measure	0.89	0.86	0.89

Table 3. Evaluation Result with Semantic Similarity

Aspect		Naïve Bayes	Decision Tree	Max Entropy
Scenery	Accuracy	0.94	0.85	0.90
	Precision	0.92	0.85	0.88
	Recall	1.00	0.96	1.00
	F-Measure	0.96	0.89	0.94
Dusk	Accuracy	0.92	0.92	0.92
	Precision	0.89	0.89	0.89
	Recall	1.00	1.00	1.00
	F-Measure	0.94	0.94	0.94
Surf	Accuracy	0.9	0.84	0.86
	Precision	0.88	0.88	0.89
	Recall	1.00	0.88	0.92
	F-Measure	0.94	0.88	0.90
Amenities	Accuracy	0.86	0.86	0.86
	Precision	0.82	0.75	0.81
	Recall	1.00	1.00	1.00
	F-Measure	0.9	0.86	0.89
Sanitation	Accuracy	0.75	1.00	0.75
	Precision	0.67	1.00	0.67
	Recall	1.00	1.00	1.00
	F-Measure	0.80	1.00	0.80

The analysis of user-generated content data from visitor reviews on Google Maps for Kuta Beach, Bali has provided valuable insights into visitor experiences and sentiments at this popular tourist destination. The labeling and categorization process, utilizing advanced algorithms for semantic similarity and sentiment analysis, have allowed for a deep understanding of the content and sentiments expressed by the visitors. These insights can be instrumental in

informing tourism management strategies and marketing approaches to enhance the overall visitor experience at Kuta Beach. This research adds to the growing body of knowledge on the significance of user-generated content in understanding tourist preferences and behaviors, and it sets the stage for further exploration and application of advanced data analysis methods in the field of tourism research. The findings of this study contribute to the understanding of the role of user-generated content in shaping responsible environmental behavior in the tourism industry. Furthermore, this research highlights the potential for user-generated content to serve as persuasive triggers, influencing travelers' environmental concerns and attitudes, ultimately leading to responsible environmental behavior and the development of sustainable tourism practices.

While it is valuable to analyze user-generated content data from visitor reviews on Google Maps for Kuta Beach, it's important to consider the limitations and potential biases associated with this data. The classification results per aspect need to be further reviewed in more depth. What aspects lead visitors to give positive reviews will certainly be maintained and even improved by stakeholders. Similarly, for negative review outcomes, it is necessary to investigate more deeply the factors contributing to visitor dissatisfaction so that they can be addressed by stakeholders. User-generated content can be subjective and may not always reflect the overall visitor experience accurately. Some visitors may have specific expectations or biases that influence their reviews, and these individual opinions may not represent the majority of visitors. Additionally, sentiment analysis algorithms may struggle to accurately interpret the nuances of language and cultural context, leading to potential misinterpretation of the true sentiment expressed in the reviews.

Furthermore, categorizing the reviews based on keywords and phrases may oversimplify the complexity of visitor experiences at Kuta Beach. Visitors' experiences are multifaceted and cannot always be adequately captured by semantic analysis alone. It's essential to approach user-generated content analysis with a critical mindset and to supplement it with other forms of research, such as surveys or direct observational studies, to ensure a more comprehensive understanding of visitor experiences and sentiments at Kuta Beach.

5. Conclusion

The analysis of user-generated content data from visitor reviews on Google Maps for Kuta Beach, Bali has provided valuable insights into visitor experiences and sentiments at this popular tourist destination. The labeling and categorization process, utilizing advanced algorithms for semantic similarity and sentiment analysis, have allowed for a deep understanding of the content and sentiments expressed by the visitors. These insights can be instrumental in informing tourism management strategies and marketing approaches to enhance the overall visitor experience at Kuta Beach. This research adds to the growing body of knowledge on the significance of user-generated content in understanding tourist preferences and behaviors, and it sets the stage for further exploration and application of advanced data analysis methods in the field of tourism research. The findings of this study contribute to the understanding of the role of user-generated content in shaping responsible environmental behavior in the tourism industry. Furthermore, this research highlights the potential for user-generated content to serve as persuasive triggers, influencing travelers' environmental concerns and attitudes, ultimately leading to responsible environmental behavior and the development of sustainable tourism practices.

While it is valuable to analyze user-generated content data from visitor reviews on Google Maps for Kuta Beach, it's important to consider the limitations and potential biases associated with this data. The classification results per aspect need to be further reviewed in more depth. What aspects lead visitors to give positive reviews will certainly be maintained and even improved by stakeholders. Similarly, for negative review outcomes, it is necessary to investigate more deeply the factors contributing to visitor dissatisfaction so that they can be addressed by stakeholders. User-generated content can be subjective and may not always reflect the overall visitor experience accurately. Some visitors may have specific expectations or biases that influence their reviews, and these individual opinions may not represent the majority of visitors. Additionally, sentiment analysis algorithms may struggle to accurately interpret the nuances of language and cultural context, leading to potential misinterpretation of the true sentiment expressed in the reviews.

Furthermore, categorizing the reviews based on keywords and phrases may oversimplify the complexity of visitor experiences at Kuta Beach. Visitors' experiences are multifaceted and cannot always be adequately captured by semantic analysis alone. It's essential to approach user-generated content analysis with a critical mindset and to

supplement it with other forms of research, such as surveys or direct observational studies, to ensure a more comprehensive understanding of visitor experiences and sentiments at Kuta Beach.

6. Declarations

6.1. Author Contributions

Conceptualization: N.M.S.I., N.A., E.M.D., and N.P.W.Y.; Methodology: N.A.; Software: N.M.S.I.; Validation: N.M.S.I., N.A., E.M.D., and N.P.W.Y.; Formal Analysis: N.M.S.I., N.A., E.M.D., and N.P.W.Y.; Investigation: N.M.S.I.; Resources: N.A.; Data Curation: N.A.; Writing Original Draft Preparation: N.M.S.I. and E.M.D.; Writing Review and Editing: N.A. and N.M.S.I.; Visualization: N.M.S.I.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received financial support for the research, authorship, and/or publication of this article from Universitas Primakara and Universitas Multimedia Nusantara.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] M. Schuckert, X. Liu, and R. Law, "Hospitality and Tourism Online Reviews: Recent Trends and Future Directions," *Journal of Travel & Tourism Marketing*, vol. 32, no. 5, pp. 608–621, Jul. 2015, doi: 10.1080/10548408.2014.933154.
- [2] N. Afriliana, N. M. S. Iswari, and Suryasari, "Sentiment Analysis of User-Generated Content: A Bibliometric Analysis," *Journal of System and Management Sciences*, vol. 12, no. 6, 2022, doi: 10.33168/JSMS.2022.0634.
- [3] E. Marrese-Taylor, J. D. Velásquez, and F. Bravo-Marquez, "A novel deterministic approach for aspect-based opinion mining in tourism products reviews," *Expert Syst Appl*, vol. 41, no. 17, pp. 7764–7775, Dec. 2014, doi: 10.1016/j.eswa.2014.05.045.
- [4] K. Puh and M. Bagić 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.
- [5] Y. Xi, C. Ma, Q. Yang, and Y. Jiang, "A Cross-cultural Analysis of Tourists' Perceptions of Airbnb Attributes," *International Journal of Hospitality & Tourism Administration*, vol. 23, no. 4, pp. 754–787, Aug. 2022, doi: 10.1080/15256480.2020.1862014.
- [6] N. M. S. Iswari and I. G. J. E. Putra, "Analysis of User-generated Content in Visitor Reviews of Tourist Attractions Using Semantic Similarity," *Ultimatics : Jurnal Teknik Informatika*, vol. 15, no. 1, p. 59, 2023, [Online]. Available: <https://www.nltk.org/>
- [7] T. Hariguna, H. Teja Sukmana, and J. Il Kim, "Survey Opinion using Sentiment Analysis," *Journal of Applied Data Sciences*, vol. 1, no. 1, pp. 35–40, 2020.
- [8] R. Setiabudi, N. M. S. Iswari, A. Rusli, "Enhancing text classification performance by preprocessing misspelled words in Indonesian Language," *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, vol. 19, no. 4, pp. 1234–1241, Aug. 2021, doi: 10.12928/TELKOMNIKA.v19i4.20369.
- [9] D. R. Yehoshua, "Naive Bayes Classification." [Online]. Available: <https://towardsdatascience.com/naive-bayes->

classification-41d1fe802a1e?gi=ce85640dc30d [Accessed: Apr. 16, 2024]

- [10] B. Seref and E. Bostanci, "Sentiment Analysis using Naive Bayes and Complement Naive Bayes Classifier Algorithms on Hadoop Framework," in 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), IEEE, vol. 1, no. 1, pp. 1–7, Oct. 2018. doi: 10.1109/ISMSIT.2018.8567243.
- [11] Z. Zi-qiong, Y. Qiang, and Li Yi-jun, "Using Naïve Bayes Classifier to Distinguish Reviews from Non-review Documents in Chinese," in 2007 International Conference on Management Science and Engineering, IEEE, vol. 1, no. 1, pp. 115–121, Aug. 2007. doi: 10.1109/ICMSE.2007.4421834.
- [12] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques," EMNLP, 2002. [Online]. Available: <http://www.cs.cornell.edu/people/pabo/movie-review-data/>. [Accessed: Apr. 16, 2024]
- [13] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.
- [14] C. Yin, J. Xi, and J. Wang, "The Research of Text Classification Technology Based on Improved Maximum Entropy Model," in 2015 First International Conference on Computational Intelligence Theory, Systems and Applications (CCITSA), IEEE, vol. 1, no. Dec, pp. 142–145, Dec. 2015. doi: 10.1109/CCITSA.2015.12.
- [15] M. C. Lee, J. W. Chang, and T. C. Hsieh, "A Grammar-Based Semantic Similarity Algorithm for Natural Language Sentences," *The Scientific World Journal*, vol. 2014, no. 1, pp. 1–17, 2014, doi: 10.1155/2014/437162.
- [16] adminlp2m, "Analisis Semantik - Definisi, Cara Kerja dan Contohnya." [Online]. Available: <https://lp2m.uma.ac.id/2023/02/11/analisis-semantik-definisi-cara-kerja-dan-contohnya/#:~:text=Analisis%20Semantik%2C%20diungkapkan%2C%20adalah%20proses,paragraf%2C%20atau%20bahkan%20seluruh%20manuskrip.> [Accessed: May 04, 2023]
- [17] A. R. Alaei, S. Becken, and B. Stantic, "Sentiment Analysis in Tourism: Capitalizing on Big Data," *J Travel Res*, vol. 58, no. 2, pp. 175–191, Feb. 2019, doi: 10.1177/0047287517747753.
- [18] D. Kaplan, "What Does Semantic Mean In Computer Science." [Online]. Available: <https://enjoymachinelearning.com/blog/finding-semantic-similarity-between-sentences-in-python/>. [Accessed: Apr. 17, 2023]
- [19] J. Hariharakrishnan, S. Mohanavalli, Srividya, and K. B. Sundhara Kumar, "Survey of pre-processing techniques for mining big data," in 2017 International Conference on Computer, Communication and Signal Processing (ICCCSP), IEEE, vol. 2017, no. Jan, pp. 1–5, Jan. 2017. doi: 10.1109/ICCCSP.2017.7944072.
- [20] S. García, S. Ramírez-Gallego, J. Luengo, J. M. Benítez, and F. Herrera, "Big data preprocessing: methods and prospects," *Big Data Analytics*, vol. 1, no. 1, p. 9, Dec. 2016, doi: 10.1186/s41044-016-0014-0