SPRINGER NATURE SNAPP | Information Technology & Tourism

Notifications Account

Adaptive Boosting for Sentiment Analysis in Tourism Destination Recommendations

CURRENT STATUS

Your submission has been rejected

We've sent an email to imam@upm.ac.id with more details.

Need help?

If you have any questions about this submission, you can email the Editorial Office (priyadarshini.vijayasekar@springernature.com).

For general enquiries, please look at our support information.

How was your experience today?

Awful Bad OK Good Great

Send feedback

Progress so far

Show history

- Submission received
- Technical check
- Editorial assignment
- With editor
- Submission rejected

Learn about our submission process

Your submission

Title

Adaptive Boosting for Sentiment Analysis in Tourism Destination Recommendations

Type

Research

Journal

Information Technology & Tourism

Adaptive Boosting for Sentiment Analysis in Tourism Destination Recommendations

Imam Marzuki [1*] Nurhidayati [2] Rona Riantini[3]

Yunifa Miftachul Arif[4] Reza Fuad Rachmadi [5] Mochamad Hariadi [5]

[1]*Department of Electrical Engineering, Universitas Panca Marga, Probolinggo, Indonesia*

[2]*Department of Mathematics, Institut Ahmad Dahlan Probolinggo, Probolinggo, Indonesia*

[3]*Department of Ship Electrical Engineering, Politeknik Perkapalan Negeri Surabaya, Surabaya, Indonesia*

[4]*Department of Informatics Engineering, Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia*

[5]*Department of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*

* Corresponding author's Email: imam@upm.ac.id

Abstract

The rapid growth of user-generated content on online platforms, such as YouTube, has unlocked new opportunities to enhance tourism destination recommendation systems through sentiment analysis. However, existing studies primarily utilize conventional machine learning techniques such as K-Nearest Neighbors, Naïve Bayes, CART, and Random Forest—which often exhibit limitations in balancing classification metrics like precision, recall, and F1-score, especially in imbalanced datasets. Despite the proven effectiveness of Adaptive Boosting in other domains, its application in sentiment analysis for tourism destination recommendations remains underexplored. To fill this gap, this study proposes a sentiment-based recommendation framework leveraging AdaBoost to improve classification robustness and accuracy. Tourist comments on YouTube from ten destinations were collected, preprocessed, and transformed into numerical features using the Bag of Words model. These were classified into positive, neutral, and negative sentiments. The model was validated using 5-fold cross-validation and evaluated via confusion matrix metrics, yielding a precision of 83.31%, recall of 83.02%, and F1-score of 81.59%—outperforming other baseline models. F1-score values were then used to rank destinations for more balanced and objective recommendations. This research demonstrates the potential of AdaBoost to bridge existing limitations in tourism sentiment analysis and contribute to more reliable, data-driven recommendation systems.

Keywords: Tourism Sentiment Analysis, Adaptive Boosting, Tourism Recommendation System, Tourism.

I. INTRODUCTION

The growth of the tourism industry in the digital era has significantly increased user-generated content, especially in the form of reviews and comments on online platforms like YouTube (Damiasih et al., 2021). These reviews not only reflect individual experiences but also offer valuable insights into public perceptions of tourist destinations. However, as this data is typically unstructured, sentiment analysis provides a strategic approach to extract structured insights by classifying reviews into positive, neutral, or negative categories (Abbasi-Moud et al., 2021).

Prior studies have applied machine learning methods such as K-Nearest Neighbors (KNN), Naïve Bayes, and Random Forest for sentiment analysis in tourism applications (Fan & Lu, 2022; Kurniawan et al., 2023; Marzuki et al., 2025). However, these methods often struggle to balance precision, recall, and F1-score, especially when dealing with imbalanced datasets (Abushahla & Pala, 2024).

To address this issue, this study proposes an Adaptive Boosting-based sentiment analysis framework. Adaptive Boosting's ability to combine multiple weak learners into a strong classifier offers potential improvements in both accuracy and model robustness when processing real-world tourism data (Tsiapoki et al., 2021).

This research contributes a tourism recommendation system that utilizes F1-score as a ranking metric to handle data imbalance effectively, producing more objective and representative recommendations. The proposed system aims to enhance the accuracy and reliability of data-driven destination recommendations for tourism stakeholders.

II. RELATED WORK

Tourism destination recommender systems play a vital role in guiding tourists toward destinations matching their preferences (Abbasi-Moud et al., 2021; Marzuki et al., 2025). However, many systems struggle with managing complex user data and leveraging sentiment analysis to enhance recommendation accuracy. Techniques such as KNN-based stacking have improved prediction metrics across accuracy, precision, recall, and F1-score (Marzuki et al., 2025). Similarly, Neural Collaborative Filtering (NCF) has demonstrated superior accuracy in tourism destination recommendation (Marzuki et al., 2024).

The growth of user-generated content on social media has enabled sentiment analysis to extract user preferences for recommender systems (Abbasi-Moud et al., 2021; Matrutty et al., 2023; Ng et al., 2023; Yadav & Vishwakarma, 2020). Techniques employing weighted word representations and linguistic constraints have shown high accuracy in pharmaceutical reviews (Yadav & Vishwakarma, 2020), while multimodal sentiment analysis leveraging image-text data has enhanced sentiment extraction in tourism contexts (Monsalve-Pulido et al., 2024; K. Zhang et al., 2023). Additionally, Latent Dirichlet Allocation (LDA) has been applied to YouTube comments for understanding tourist sentiments (Thilakarathne et al., 2021).

Ensemble learning techniques have been extensively studied across machine learning, AI, and pattern recognition for improving classification performance (Rincy & Gupta, 2020). Adaptive boosting, in particular, has been effective in enhancing detection accuracy in structural health monitoring (Tsiapoki et al., 2021) and improving classification by integrating weighted feature selection and category confidence (Wang & Feng, 2021). However, its potential in sentiment analysis for tourism data remains underexplored.

Various machine learning models, including SVM, Naïve Bayes, KNN, Random Forest, LSTM, and BERT, have been applied to sentiment analysis with varying degrees of success (Ajhari, 2023; Chamorro-Atalaya et al., 2023; Ng et al., 2023; Nurkholis et al., 2025). While some studies achieved high accuracy using SVM in classifying sentiment in education and COVID-19 vaccination discourse (Chamorro-Atalaya et al., 2023; Nurkholis et al., 2025) others applied Naïve Bayes for hotel reviews classification with moderate performance (Matrutty et al., 2023). Additionally, multimodal sentiment models using SenticNet for tourism data in Spanish have improved F1-scores compared to manually labeled datasets (Monsalve-Pulido et al., 2024).

Tourism recommender systems have also explored multi-criteria approaches to improve recommendations beyond single-criterion systems (Arif et al., 2022). Integration of cosine similarity in multi-criteria systems has shown effects on accuracy, precision, recall, and F1-scores in halal tourism games. Other approaches have used collaborative and content-based filtering for personalized recommendations in movies (Husin et al., 2023) and telecom packages using optimized deep forest models, which increased F1-scores by 5% after parameter tuning (Y. Zhang et al., 2021).

The importance of clustering and feature extraction is highlighted in tourism destination classification using k-means (Muhammad & Saputra, 2021) and improved random forest algorithms for clustering tourism data (Fan & Lu, 2022). Meanwhile, the steady rise of smart tourism has led to systematic reviews identifying trends, themes, and challenges in implementing smart tourism destinations globally (Elda Hiererra et al., 2022).

Additional studies have examined complex network theory for personalized recommendations in airline crew management (Su et al., 2024) and word networks with context-aware techniques for English speaking recommendations (Ding, 2025), showing performance improvements in recommendation accuracy and system responsiveness. Dataset recommendation systems using domain-specific text classification have also achieved promising F1-scores, aiding researchers in finding relevant datasets (Färber & Leisinger, 2021).

In summary, while sentiment analysis and ensemble learning have been widely studied across domains, there is a clear research gap in utilizing adaptive boosting-based sentiment analysis of tourist reviews specifically for enhancing tourism destination recommender systems. Existing research has demonstrated the importance of accurate sentiment extraction and the benefits of ensemble learning, but the integration of Adaptive Boosting for refining sentiment feature extraction to improve recommendation accuracy and personalization in tourism remains insufficiently explored.

This study aims to address this gap by proposing an adaptive boosting-based sentiment analysis framework on tourism reviews, enhancing the personalization and accuracy of tourism destination recommender systems through enriched sentiment feature extraction.

III. RESEARCH METHOD

This study proposes a destination recommendation system framework based on sentiment analysis, utilizing traveler comments from YouTube as the primary data source. The methodology is structured into four main stages: data collection, data preprocessing, sentiment classification using the Adaptive Boosting algorithm, and the integration of sentiment analysis results into the recommendation system. This framework is designed to improve sentiment classification accuracy and address data imbalance issues, thereby supporting more reliable and objective tourism destination recommendations. The proposed system architecture, based on the research methodology, is presented in Figure 1.

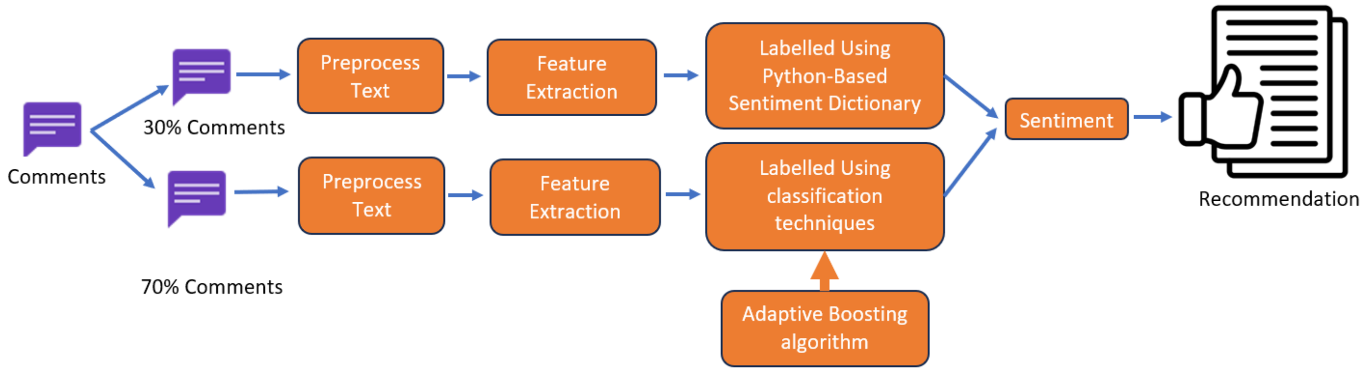

Fig1. Proposed System Architecture

A. Data Collection

The data used in this study is in the form of comments or reviews of tourists taken from the YouTube platform. Data collection was carried out using web scraping techniques, which focused on comments from videos of 10 popular tourism destinations in Batu City. These destinations include Jatim Park 1, Jatim Park 2, Jatim Park 3, Angkut Museum, Selecta, Eco Green Park, Kusuma Agro, Cangar, Pemandian Songgoriti, and Predator Fun Park (Marzuki, 2025). The dataset of this study is presented in table 1.

Table 1. Youtube URL Address, Tourism Destination and Number of Comments

No	Tourism Destination	URL	Number of Comments
1	Jatim Park 1	https://www.youtube.com/watch?v=lR-V6LyKES8	331
2	Jatim Park 2	https://www.youtube.com/watch?v=mURtBdYruDU	263
3	Jatim Park 3	https://www.youtube.com/watch?v=SNl1Yxj8H3A	185
4	Museum Angkut	https://www.youtube.com/watch?v=eAXpPX96nyQ	1048
5	Selecta	https://www.youtube.com/watch?v=BfsipVuzfE0	187
6	Eco Green Park	https://www.youtube.com/watch?v=GBq-dEh1i5A	110
7	Cangar	https://www.youtube.com/watch?v=FBRMiPc_I7c	43
8	Pemandian Songgoriti	https://www.youtube.com/watch?v=2aob-cg6c9M	99
9	Coban Rais	https://www.youtube.com/watch?v=h0szbzo3t88	56

10	Predator Fun Park	https://www.youtube.com/watch?v=piPlavtlkyo	65

As many as 30% of the comments that were successfully collected were manually labeled using a Python-based sentiment dictionary, while the remaining 70% were automatically labeled with classification techniques using the Adaptive Boosting algorithm. The labeling of this sentiment is carried out into three categories, namely positive, negative, and neutral, to facilitate the analysis of tourists' perception of tourist destinations in Batu City. The results of this labeling are expected to support the preparation of tourism development recommendations in a more targeted manner.

B. Preprocessing Text

The preprocessing stage is performed to clean and prepare the data before it is used in the training of the sentiment analysis model (Vibha & Singh, 2018). This process includes tokenization to break text into tokens, the removal of stopwords so that common words that have no significant contribution can be eliminated, as well as stemming to change the word into its basic form. In addition, the text is normalized by standardizing writing, such as changing capital letters to lowercase letters, as well as the removal of non-alphabetic characters and irrelevant numbers. All of these steps aim to ensure that the data is clean, consistent, and ready to use so that it can improve the accuracy of the model training stage.

C. Feature Extraction

Text feature extraction is performed using the Bag of Words (BoW) method to convert text data into numerical representations that can be processed by machine learning algorithms (Putrada et al., 2023). Through this method, each word in the text corpus is calculated in frequency and represented in vector form, where each vector element indicates the number of occurrences of a particular word on the document. This approach allows the model to recognize patterns and relationships between words based on their frequency, so that the information in the text can be processed into appropriate inputs for sentiment analysis. With the use of Bag of Words, text data that was previously unstructured can be compiled in a consistent numerical format and facilitate the process of training classification models.

In this method, each document is represented as a dimensional vector, where V is the unique number of words (vocabulary size) in the corpus. Mathematically, document d is represented as:

$$x^{(d)} = \left[{x_1}^{(d)}, {x_1}^{(d)}, \ldots, {x_V}^{(d)} \right]$$

With:

${x_i}^{(d)}$ = Number of occurrences of the word *i* in document *d*

Through this approach, each word is counted in frequency, and the result is arranged into a vector that represents the document in vector space. All documents in the corpus will then form a document-term matrix as follows:

$$\begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_V^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_V^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(D)} & x_2^{(D)} & \dots & x_V^{(D)} \end{bmatrix}$$

Where *D* is the sum of documents in the corpus. In this study, the Bag of Words approach was used to convert text data that was originally unstructured into a structured numerical format, thereby facilitating the process of training sentiment analysis models at the next stage.

D. Model Training

The model in this study was trained using the Adaptive Boosting algorithm by utilizing review data that has gone through the preprocessing and feature extraction stages. Adaptive Boosting is an ensemble method used to improve the accuracy of classification models by combining several weak learners into strong learners (Wang & Feng, 2021) In sentiment classification, this algorithm is used to predict sentiment categories (positive, negative, neutral) based on numerical features extracted from text data (e.g. using Bag of Words or TF-IDF). Here are the detailed steps for the implementation of the algorithm:

1. Data Weight Initialization

   The dataset of preprocessing and feature extraction results (Bag of Words) consists of *N* reviews of travellers who have been labeled sentiment. Each sample (x_i, y_i) is assigned the same initial weight:

   $$w_i^{(1)} = \frac{1}{N}, i = 1,2, \dots, N$$

   so that all reviews have an equal contribution in the early stages of training.

2. Weak Learner Training Iteration

   Perform as many iterations with the following step details:

   a. Train Weak Learners

      In each t-iteration, weak learners (e.g. decision stump or shallow decision tree) are trained on the dataset taking into account the sample weight w_i^t

   b. Menghitung Error Weak Learner

      Weak learner errors on the t-iteration are calculated using:

      $$e_t = \sum_{i=1}^{N} w_i^{(t)} \cdot I(h_t(x_i) \neq y_i)$$

      With is an indicator function that values 1 if the prediction is wrong, 0 if the prediction is correct.I(·)

   c. Calculating the Weight of a Weak Learner

      If the weight of the weak learner is calculated by 0<e_t<0.5:

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-e_t}{e_t}\right)$$

This value indicates the learner's contribution to the final prediction.

d. Update Data Weights

Sample weights are updated so that the model focuses on reviews that are difficult to predict correctly:

$$w_i^{(t+1)} = w_i^{(t)} \cdot \exp\left(-\alpha_t y_i h_t(x_i)\right)$$

- If the prediction is correct, the weight will decrease$(y_i = h_t(x_i))$.
- If the prediction is wrong, the weight will increase.$(y_i \neq h_t(x_i))$

e. Weight Normalization

The weight of the update result is normalized so that the total weight remains 1:

$$w_i^{(t+1)} = \frac{w_i^{(t+1)}}{\sum_{j=1}^{N} w_j^{(t+1)}}$$

3. Final Classification

After the entire iteration is complete, the ensemble model is formed from a combination of weak learners with the final prediction:

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$$

In this study, a one-vs-rest approach was applied to handle multiclass classifications (positive, negative, neutral).

E. Model Validation

To ensure the generalizability of the sentiment classification model while minimizing the risk of overfitting, this study applied the k-fold cross-validation technique with a value of *k*=5 (Chamorro-Atalaya et al., 2023). Dataset *D* is divided into five subsets of approximately the same size, namely:

$$D = \{F_1, F_2, F_3, F_4, F_5\}$$

In each ith i, a subset is used as F_i a validation set, while the other four subsets are used as training data:

$$Training_i = D \backslash F_i$$

$$Validation_i = F_i$$

This process is carried out five times (*i*=1,2,3,4,5). For each iteration, a model evaluation value such as precision, recall, or F1-score is obtained, which is symbolized as M_i. The average performance value is calculated as follows:

$$M_{avg} = \frac{1}{k}\sum_{i=1}^{k} M\_i$$

Since *k*=5, then:

$$M_{avg} = 1/5(M_1 + M_2 + M_3 + M_4 + M_5)$$

This average value (M_avg) used to represent the overall performance of the model, so that it can be assessed whether the model has a good generalization of data that has never been seen before.

F. Performance Evaluation

The evaluation of model performance in this study was carried out using a confusion matrix, which functions as the basis for calculating the main evaluation metrics, namely precision (P), recall (R), and F1-score ($F1$). These three evaluation metrics were used to provide a comprehensive picture of the model's ability to accurately classify traveler review sentiment.

Mathematically, the confusion matrix produces four main components, namely True Positives (TP), which indicates the number of positive reviews that have been successfully correctly classified as positive; True Negatives (TN), which is the number of negative reviews that are successfully predicted as negative; False Positives (FP), which indicates the number of negative reviews that are mistakenly predicted as positive reviews; and False Negatives (FN), which is the number of positive reviews that are incorrectly classified as negative. These four components are the basis for the calculation of the performance evaluation metrics of the classification model (Krstinić et al., 2020).

These components are used to compute the following evaluation metrics:

- Precision (P)

  Precision measures the proportion of correctly predicted positive observations to the total predicted positive observations:

$$P = \frac{TP}{TP + FP}$$

- Recall (R)

  Recall, or sensitivity, measures the proportion of correctly predicted positive observations to all actual positive observations:

$$R = \frac{TP}{TP + FN}$$

- F1-Score ($F1$)

  F1-score is the harmonic mean of precision and recall, providing a single measure that balances both concerns:

$$F1 = 2 \times \frac{P \times R}{P + R}$$

G. Integration to the Recommendation Syatem

Furthermore, the results of the F1-Score calculation are used as the basis for the ranking process, where each rating represents the level of recommendation given (Su et al., 2024). The use of F1-Score was chosen specifically because it provides an optimal balance between precision and recall, so that it can handle unbalanced data distribution more effectively. With this approach, the system can generate more accurate and fair recommendations, especially in situations where the amount of data from each class is not equal.

IV. RESULTS AND DISCUSSION

This section presents the main findings from testing and evaluating the proposed sentiment analysis model. The performance of the Adaptive Boosting algorithm in classifying tourist review sentiments is analyzed and compared with alternative models, including K-Nearest Neighbors, Naïve Bayes, CART, and Random Forest. The integration of sentiment analysis results into the tourism recommendation system is also discussed, highlighting how this approach generates more accurate and relevant destination recommendations. Finally, the results are compared with previous studies to demonstrate the advantages and contributions of the proposed method.

A. Experimental Results

The performance of the Adaptive Boosting model was evaluated using tourist review data that had undergone preprocessing and feature extraction. The evaluation focused on assessing the model's ability to classify sentiments into positive, neutral, and negative categories. Table 2 summarizes the distribution of sentiment classifications across the ten tourist destinations, providing an overview of visitor perceptions derived from the analyzed reviews.

Table 2. Distribution of Visitor Comment Sentiment in 10 Tourism Destinations

No	Destinations	Positive	Neutral	Negative	Number of Comments
1	Jatim Park 1	83	225	23	331
2	Jatim Park 2	57	182	24	263
3	Jatim Park 3	38	136	11	185
4	Museum Angkut	33	255	43	1048
5	Selecta	45	137	5	187
6	Eco Green Park	26	73	11	110
7	Cangar	23	20	0	43
8	Pemandian Songgoriti	50	44	5	99
9	Coban Rais	22	29	5	56
10	Predator Fun Park	17	45	3	65

Table 2 summarizes the distribution of positive, neutral, and negative visitor comments across ten tourism destinations in Batu City. The Angkut Museum received the highest number of comments (1048), with a majority expressing neutral sentiment. Jatim Park 1 and Jatim Park 2 also attracted significant comment volumes, totaling 331 and 263 comments, respectively. In contrast, destinations such as Cangar and Coban Rais received relatively few comments. Notably, Jatim Park 1 recorded the highest number of positive comments, while Angkut Museum received the most negative feedback. These findings provide an initial indication of each destination’s relative popularity and visitor perception, forming a foundation for sentiment analysis and the recommendation system development.

B. Model Evaluation and Comparison

Model evaluation and comparison were conducted to assess the effectiveness and performance of the model in accurately classifying the sentiment of traveler reviews. The evaluation process uses the confusion matrix as the basis for calculating evaluation metrics such as precision, recall, and F1-score, which provides a comprehensive picture of the model's capabilities. Furthermore, the results of the evaluation of the main model, namely Adaptive Boosting, were compared with several comparative models such as K-Nearest Neighbors (KNN), Naïve Bayes, Classification and Regression Tree (CART), and Random Forest. The results of the evaluation and comparison of the models are presented in table 3.

Table 3. Comparison of the results of the evaluation of five sentiment classification models

No	**Model**	**Evaluation**		
		Precision	**Recall**	**F1-Score**
1	Adaptive Boosting	83.31%	83.02%	81.59%
2	KNN	67.82%	67.30%	58.52%
3	Naïve Bayes	78.32%	68.10%	69.80%
4	CART	80.84%	80.52%	79.19%
5	Random Forest	82.11%	80.03%	76.92%

Table 3 presents the results of the evaluation of five sentiment classification models based on precision. recall. and F1-score metrics. From the table. it can be seen that the Adaptive Boosting model produces the best performance with precision of 83.31%. recall of 83.02%. and the highest F1-score of 81.59%. The CART and Random Forest models took the next position with F1 scores of 79.19% and 76.92%. respectively. although the resulting precision and recall were relatively lower than Adaptive Boosting. Meanwhile. KNN showed the lowest performance with an F1-score of only 58.52%. Overall. this table shows that Adaptive Boosting is the most optimal model to use in the classification of sentiment in the review of tourism destinations in this study.

C. Integration to the Recommendation System

Based on the results of the F1-Score evaluation of 14 tourism destinations obtained using the Adaptive Boosting model. a ranking process was carried out to determine the level of recommendations for each destination. F1-Score was chosen as the basis for the ranking because it is able to represent a balance between precision and recall. so that the results obtained are fairer in unbalanced data conditions. The tourism destinations with the highest F1-Score are placed at the top as the most recommended destinations to tourists. while the destinations with lower scores are ranked next. The results of the ranking are presented systematically in Table 4. thus facilitating the interpretation and decision-making process in the developed recommendation system.

Table 4. Ranking results of 10 tourism destinations

Rank	Tourism Destination	F1-Score
1	Museum Angkut	91.4%
2	Cangar	90.7%
3	Jatim Park 2	86.0%
4	Jatim Park 1	85.4%
5	Jatim Park 3	79.8%
6	Selecta	79.8%
7	Pemandian Songgoriti	78.6%
8	Eco Green Park	78.4%
9	Coban Rais	74.6%
10	Predator Fun Park	71.2%

Table 4 shows the ranking results of 10 tourism destinations based on the F1-Score obtained from the results of sentiment classification using the Adaptive Boosting model. Museum Angkut ranks first with the highest F1-Score of 91.4% followed by Cangar in second place with a score of 90.7% and Jatim Park 2 and Jatim Park 1 which occupy the third and fourth positions with scores of 86.0% and 85.4%. respectively. Meanwhile. destinations such as Coban Rais and Predator Fun Park occupy the bottom position with F1-Scores of 74.6% and 71.2%. This ranking shows that top-ranked destinations are more recommended in the recommended system developed. because they have a more positive and balanced perception of sentiment from visitors than other destinations.

D. Comparison with Previous Studies

To provide a more comprehensive picture of the advantages of the proposed method. the results of this study were compared with previous studies. This comparison aims to assess the extent to which the Adaptive Boosting model used is able to provide performance improvement in sentiment classification tasks. especially in the context of tourism destination recommendations. By comparing the results of F1-Score. precision. recall. and methodological approaches used in previous studies. it is hoped that a clearer understanding of the contributions and advantages of the approach proposed in this study can be obtained. Comparisons with previous studies are presented in Table 5.

Table 5. Comparison with previous studies

No	Study	Focus	Method	Evaluation Results		
				Precision	Recall	F1-Score
1	(K. Zhang et al., 2023)	Sentiment analysis of tourism reviews	Image-Text Multimodal (VGG19 + TinyBERT + BiGRU-Attention + Dual Linear Fusion)	78.61%	78.01%	78.34%

2	(Monsalve -Pulido et al., 2024)	Multimodal sentiment analysis on tourism data (Spanish. TASS dataset)	SenticNet 5 + Random Forest & SVM + Fusion (text & images)	71%	71%	71%
3	(Matrutty et al., 2023)	Sentiment analysis of hotel reviews (TripAdvisor. Manado. North Sulawesi)	Naïve Bayes	70.57%	99.85%	70.55%
4	Ours	Sentiment analysis of tourism destinations reviews	Adaptive Boosting	83.31%	83.02%	81.59%

Table 5 presents a comparison between this study and some previous studies related to the sentiment analysis of tourism reviews. The first study used a multimodal approach based on the combination of VGG19. TinyBERT. and BiGRU-Attention which resulted in a precision of 78.61%. recall of 78.01%. and an F1-score of 78.34%. The second study utilized a combination of SenticNet 5. Random Forest. and SVM on text and image data. resulting in a uniform evaluation score of 71%. Meanwhile. a third study that applied the Naïve Bayes method to hotel reviews on TripAdvisor resulted in high recalls (99.85%) but lower accuracy and F1-scores. namely 70.57% and 70.55%. In this context. the approach proposed in this study. namely Adaptive Boosting. shows the best performance with a precision of 83.31%. recall of 83.02%. and an F1-score of 81.59%. indicating the advantages of the model in classifying the sentiment of tourist destination reviews more accurately and balanced.

V. CONCLUSIONS AND FUTURE WORK

Based on the results of the study. the Adaptive Boosting-based sentiment analysis model developed is proven to be able to increase accuracy and resilience in sentiment classification in tourism destination reviews. The use of comment data from the YouTube platform as the main source provides a more realistic picture of tourist perception. Compared to the KNN. Naïve Bayes. CART. and Random Forest methods. the Adaptive Boosting model showed the best performance with an F1-score of 81.59%. as well as an adequate balance of precision and recall. The integration of this model into the tourism destination recommendation system allows for a more objective. transparent. and data-based destination ranking. The resulting recommendation system is also considered effective in handling data imbalances. so that it is able to provide fairer and more representative recommendations to tourist preferences.

As a suggestion for further research. development can be directed towards the application of more complex feature extraction methods. such as TF-IDF. word embeddings (Word2Vec or GloVe). to BERT-based contextual embeddings to capture richer meaning of words and contexts. In addition. the integration of multimodal data such as images and videos with text comments can be explored to improve the accuracy of sentiment detection. The implementation of deep learning models such as LSTM. BiGRU. or attention mechanisms is also recommended to improve sentiment classification performance. The expansion of data from various social media platforms as well as multilingual support can expand the scope and generalization of the model.

On the application side. the development of a recommendation system that can operate in real-time is expected to be a practical solution for stakeholders in supporting decision-making in the tourism sector.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization, I. Marzuki, and M. Hariadi; methodology, I. Marzuki, Y.M. Arif, R.F. Rachmadi, and M. Hariadi; software, I. Marzuki, D. Hindarto, F.N. Fiqri; validation, I. Marzuki, Y.M. Arif, R.F. Rachmadi, and M. Hariadi; formal analysis, I. Marzuki, F.N. Fiqri, Nurhidayati, and M. Hariadi; investigation, I. Marzuki, Y.M. Arif, R.F. Rachmadi, and M. Hariadi; resources, F.N. Fiqri and Nurhidayati; data curation, I. Marzuki; writing—original draft preparation, I. Marzuki, Y.M. Arif, R.F. Rachamdi, and M. Hariadi; writing—review and editing, I. Marzuki, Y.M. Arif, R.F. Rachmadi, and M. Hariadi; visualization, I. Marzuki, D. Hindarto, and F.N. Fiqri; supervision, Y.M. Arif, R.F. Rachmadi, and M. Hariadi; project administration, I. Marzuki and Nurhidayati; funding acquisition, I. Marzuki.

ACKNOWLEDGMENT

The authors express sincere gratitude for the financial assistance provided by the Indonesian Education Scholarship (BPI), the Education Fund Management Institute (LPDP), and the Center for Higher Education Financing and Assessment (PPAPT) under the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia. This support has been instrumental in enabling the research process, securing essential resources, and ensuring the successful completion of this study. The ongoing commitment of BPI, LPDP, and PPAPT reflects their dedication to fostering research development and promoting academic excellence. Their invaluable contribution has greatly enhanced the overall quality and impact of this research.

REFERENCES

Abbasi-Moud, Z., Vahdat-Nejad, H., & Sadri, J. (2021). Tourism recommendation system based on semantic clustering and sentiment analysis. *Expert Systems with Applications*, *167*, 114324. https://doi.org/10.1016/j.eswa.2020.114324

Abushahla, K. H., & Pala, M. A. (2024). Optimizing Diabetes Prediction: Addressing Data Imbalance with Machine Learning Algorithms. *ADBA Computer Science*, 1. https://doi.org/10.69882/adba.cs.2024075

Ajhari, A. A. (2023). The Comparison of Sentiment Analysis of Moon Knight Movie Reviews between Multinomial Naive Bayes and Support Vector Machine. *Applied Information System and Management (AISM)*, *6*(1), 13–20. https://doi.org/10.15408/aism.v6i1.26045

Arif, Y. M., Nurhayati, H., Nugroho, S. M. S., & Hariadi, M. (2022). Destinations Ratings Based Multi-Criteria Recommender System for Indonesian Halal Tourism Game. *International Journal of Intelligent Engineering and Systems*, *15*(1). https://doi.org/10.22266/ijies2022.0228.26

Chamorro-Atalaya, O., Arévalo-Tuesta, J., Balarezo-Mares, D., Gonzáles-Pacheco, A., Mendoza-León, O., Quipuscoa-Silvestre, M., Tomás-Quispe, G., & Suarez-Bazalar, R. (2023). K-Fold Cross-Validation through Identification of the Opinion Classification Algorithm for the Satisfaction of University Students. *International Journal of Online and Biomedical Engineering (iJOE)*, *19*(11). https://doi.org/10.3991/ijoe.v19i11.39887

Damiasih, D., Agustina, A. R., Azahari, A., & Susanto, D. R. (2021). A Review of YouTube for Tourism Promotion in Indonesia. *IJISTECH (International Journal of Information System & Technology)*, *4*(2), 615. https://doi.org/10.30645/ijistech.v4i2.102

Ding, Z. (2025). Design and evaluation of an English speaking recommender system using word networks and context-aware techniques. *Entertainment Computing*, *52*, 100920. https://doi.org/10.1016/j.entcom.2024.100920

Elda Hiererra, S., Meyliana, Ramadhan, A., & Purnomo, F. (2022). Smart Tourism Destinations Research Trend: A Systematic Literature Review. *2022 International Conference on ICT for Smart Society (ICISS)*, 1–11. https://doi.org/10.1109/ICISS55894.2022.9915164

Fan, Z., & Lu, L. (2022). Tourism Data Clustering Analysis Based on Random Forest Algorithm. *2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, 1211–1214. https://doi.org/10.1109/IPEC54454.2022.9777525

Färber, M., & Leisinger, A.-K. (2021). Recommending Datasets for Scientific Problem Descriptions. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 3014–3018. https://doi.org/10.1145/3459637.3482166

Husin, M. R. M., Razak, T. R., Malik, A. M. A., Nordin, S., & Abdul-Rahman, S. (2023). Hybrid Collaborative Movie Recommendation System. *2023 4th International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, 274–280. https://doi.org/10.1109/aidas60501.2023.10284679

Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020, June 27). Multi-label Classifier Performance Evaluation with Confusion Matrix. *Computer Science & Information Technology*. International Conference on Soft Computing, Artificial Intelligence and Machine Learning (SAIM 2020). https://doi.org/10.5121/csit.2020.100801

Kurniawan, R., Iskandar, I., Lestari, F., Harpizon, H. A. R., & Husti, I. (2023). Exploring Tourist Feedback on Riau Attractions Through Indonesian Language YouTube Opinion Using Naïve Bayes Algorithm. *2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, 415–419. https://doi.org/10.1109/iaict59002.2023.10205795

Marzuki, I. (2025). *Alamat URL Youtube Destinasi Wisata dan Jumlah Komentar* [Dataset]. https://imammarzuki.github.io/dataset/publikasi4/

Marzuki, I., Hariadi, M., Rachmadi, R. F., & Arif, Y. M. (2024). Neural Collaborative Filtering for Improved Tourism Destination Recommendation. *2024 8th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 481–486. https://doi.org/10.1109/icitisee63424.2024.10730542

Marzuki, I., Hindarto, D., Fiqri, F. N., Nurhidayati, N., Arif, Y. M., Rachmadi, R. F., & Hariadi, M. (2025). A KNN-Based Stacking for Tourism Destination Recommender System Based on User Demographic Data. *International Journal of Intelligent Engineering and Systems*, *18*(4), 219–232. https://doi.org/10.22266/ijies2025.0531.14

Matrutty, J. P., Adrian, A. M., & Angdresey, A. (2023). Sentiment Analysis of Visitor Reviews on Star Hotels in Manado City. *Journal of Information Technology and Computer Science*, *8*(1), 21–32. https://doi.org/10.25126/jitecs.202381403

Monsalve-Pulido, J., Parra, C. A., & Aguilar, J. (2024). Multimodal model for the Spanish sentiment analysis in a tourism domain. *Social Network Analysis and Mining*, *14*(1). https://doi.org/10.1007/s13278-024-01202-3

Muhammad, A., & Saputra, D. (2021). Machine Learning in the Development of Integrated Tourist Destinations. *2021 International Conference on Computer Science and Engineering (IC2SE)*, 1–6. https://doi.org/10.1109/IC2SE52832.2021.9791976

Ng, J. X., Lim, K. M., Lee, C. P., Lim, Q. Z., Ooi, E. K. H., & Loh, N. K. N. (2023). Sentiment Analysis Using Learning-based Approaches: A Comparative Study. *2023 11th International Conference on Information and Communication Technology (ICoICT)*, 469–474. https://doi.org/10.1109/icoict58202.2023.10262604

Nurkholis, A., Styawati, S., Alim, S., Saputra, H., & Ferriyan, A. (2025). Sentiment Analysis of COVID-19 Booster Vaccines on Twitter Using Multi-Class Support Vector Machine. *Applied Information System and Management (AISM)*, *8*(1), 29–36. https://doi.org/10.15408/aism.v8i1.42911

Putrada, A. G., Alamsyah, N., Pane, S. F., & Nurkamal Fauzan, M. (2023). Feature Importance on Text Analysis for a Novel Indonesian Movie Recommender System. *2023 11th International Conference on Information and Communication Technology (ICoICT)*, 34–39. https://doi.org/10.1109/icoict58202.2023.10262504

Rincy, T. N., & Gupta, R. (2020). Ensemble Learning Techniques and its Efficiency in Machine Learning: A Survey. *2nd International Conference on Data, Engineering and Applications (IDEA)*, 1–6. https://doi.org/10.1109/IDEA49133.2020.9170675

Su, S., Lu, S., & Liu, Y. (2024). Algorithm Design and Optimization of Personalized Recommendation System for Air Crew Based on Complex Network. *2024 3rd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS)*, 728–733. https://doi.org/10.1109/aiars63200.2024.00137

Thilakarathne, K. R. R. S., Kumara, B. T. G. S., & Kuhaneswaran, B. (2021). Analyzing Tourists' Perceptions of Tourism Destinations using YouTube Comments. *2021 International Conference on Data Analytics for Business and Industry (ICDABI)*, 301–305. https://doi.org/10.1109/ICDABI53623.2021.9655833

Tsiapoki, S., Bahrami, O., Häckell, M. W., Lynch, J. P., & Rolfes, R. (2021). Combination of damage feature decisions with adaptive boosting for improving the detection performance of a structural health monitoring framework: Validation on an operating wind turbine. *Structural Health Monitoring*, *20*(2), 637–660. https://doi.org/10.1177/1475921720909379

Vibha, & Singh, A. P. (2018). Analysis of Variants of KNN Algorithm based on Preprocessing Techniques. *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 186–191. https://doi.org/10.1109/ICACCCN.2018.8748429

Wang, Y., & Feng, L. (2021). An adaptive boosting algorithm based on weighted feature selection and category classification confidence. *Applied Intelligence*, *51*(10), 6837–6858. https://doi.org/10.1007/s10489-020-02184-3

Yadav, A., & Vishwakarma, D. K. (2020). A Weighted Text Representation framework for Sentiment Analysis of Medical Drug Reviews. *2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)*, 326–332. https://doi.org/10.1109/bigmm50055.2020.00057

Zhang, K., Wang, S., & Yu, Y. (2023). A TBGAV-Based Image-Text Multimodal Sentiment Analysis Method for Tourism Reviews. *International Journal of Information Technology and Web Engineering*, *18*(1), 1–17. https://doi.org/10.4018/ijitwe.334595

Zhang, Y., Wang, M., & Yu, Y. (2021). Telecommunications package recommendation algorithm based on Deep forest. *Journal of Physics: Conference Series*, *2078*(1), 012014. https://doi.org/10.1088/1742-6596/2078/1/012014