Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.11 n.52, p. 58-67. March / April., 2025.

DOI: https://doi.org/10.5935/jetia.v11i52.1533

RESEARCH ARTICLE OPEN ACCESS

ENHANCING SENTIMENT ANALYSIS IN TOURISM REVIEWS: A COMPARATIVE STUDY OF ALGORITHMS IN ASPECT-BASED SENTIMENT ANALYSIS AND EMOTION DETECTION

Viktor Handrianus Pranatawijaya[1], Putu Bagus Adidyana Anugrah Putra[2], Ressa Priskila[3], and Novera Kristianti[4]

[1,2,3,4] Informatics Engineering Department, Faculty of Engineering, University of Palangka Raya, Indonesia.2, 3 Example Institute - EI. Toronto, Canada.

[1]http://orcid.org/0000-0002-3301-0702, [2]http://orcid.org/0000-0003-1986-125X, [3]http://orcid.org/0009-0000-6606-5120, [4]http://orcid.org/0000-0003-1536-2875

Email: viktorhp@it.upr.ac.id, putubagus@it.upr.ac.id, ressa@it.upr.ac.id, noverakristianti@eng.upr.ac.id

ARTICLE INFO

Article History
Received: January 06, 2025
Revised: February 20, 2025
Accepted: March 15, 2025
Published: March 31, 2025

Keywords:
Aspect-Based Sentiment Analysis (ABSA),
Emotion Detection,
Machine Learning,
Deep Learning,
Tourism reviews.

ABSTRACT

Information technology now enables utilizing online review data to support the Sustainable Development Goals (SDGs). However, traditional sentiment analysis often cannot capture the complexity of sentiment. This research aims to combine Aspect-Based Sentiment Analysis (ABSA) and emotion detection for a more in-depth analysis of tourism reviews in Palangka Raya City and compare the performance of various algorithms. Review data was taken from Google Maps and analyzed using BoW, LDA, NRC Emotion Lexicon, machine learning, and deep learning algorithms such as Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gradient Boosting (GB), Decision Tree (DT), and BERT. Most of the reviews are positive, with the dominance of the emotions of anticipation and joy. The combination of cross-validation with the best parameters from GridSearchCV resulted in the most significant increase in model accuracy. The SVM model performed better than other machine learning and deep learning algorithms, with accuracy and F1-score reaching 99.86%. The combination of ABSA and emotion detection improves the understanding of sentiment and emotion to support strategic decisions in tourism.

I. INTRODUCTION

Developments in information technology have expanded the utilization of online review data to support the Sustainable Development Goals (SDGs) in increasingly significant ways. Sentiment analysis, which traditionally classifies reviews as positive, negative, or neutral [1], [2] often falls short of capturing the more complex nuances of sentiment [3]. To overcome these limitations, this research aims to adopt a more comprehensive approach by combining aspect-based sentiment analysis (ABSA) and emotion detection. This approach enables a more in-depth analysis by categorizing review data based on aspects and identifying sentiments on each aspect [4], [5]. Thus, this research not only provides an overview of sentiment but also enables the identification of specific strengths and weaknesses of the reviewed product or service, ultimately supporting more effective strategic decisions [6], [7].

In line with SDG 8, which encourages inclusive and sustainable economic growth and job creation, a better understanding of sentiment and emotions in tourism reviews can help businesses improve services and attract more tourists [8]. This in turn can increase employment and income for local communities. In addition, SDG 9 emphasizes the importance of sustainable infrastructure development and the promotion of innovation [9]. By utilizing data analysis technologies such as those used in this study, the tourism sector can innovate marketing and service strategies, thereby promoting higher competitiveness [10]. In this study, aspect analysis was conducted through topic modeling using Bag of Words (BoW) and Latent Dirichlet Allocation (LDA) methods, while emotion detection used the NRC Emotion Lexicon based on Plutchik's emotion model. This model identifies basic emotions in text with two main dimensions: valence and arousal. Valence indicates whether the emotion is positive or negative, while arousal measures emotional intensity [11]. An understanding of valence and arousal helps determine whether the text reflects satisfaction or dissatisfaction as well as the emotional intensity involved. Combining aspect-based sentiment analysis and emotion detection enables better decision-making, increases user satisfaction, and improves service quality [12].

This research focuses on analyzing sentiments and emotions towards various aspects of tourism in Palangka Raya City based on reviews on Google Maps. To achieve this goal, this research uses various Machine Learning and Deep Learning methods such as Logistic Regression (LR), Naive Bayes (NB), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gradient Boosting (GD), Decision Tree (DT), and BERT [13]. This method was chosen for its ability to effectively handle various aspects of review data. The main objectives of this research are to analyze sentiment, identify relevant aspects of tourism, analyze valence and arousal in reviews, and compare the performance of various machine learning and deep learning algorithms in classifying sentiment and emotions.

The main innovation of this research lies in the combined application of aspect-based sentiment analysis and emotion detection in the context of tourism in Palangka Raya City, Indonesia. This research compares the performance of various machine learning and deep learning algorithms on a specific dataset to contribute to developing sentiment analysis. To improve the performance of the analysis, labeling is performed using VADER, TextBlob, Flair, and BERT, with the final sentiment determined through the majority voting method. VADER extracts sentiment features from unlabeled datasets, while TextBlob provides high accuracy through sentiment dictionaries and statistical models. Flair shows effectiveness in handling various language nuances, while BERT improves accuracy by deeply understanding language context and nuances [14].

To improve the performance of machine learning models, this research also conducted several experiments, including oversampling, cross-validation, GridSearchCV, and cross-validation using the best parameters from GridSearchCV [15]. Oversampling is used to overcome the problem of class imbalance, cross-validation to estimate model performance more accurately, GridSearchCV to find the best combination of parameters, and cross-validation with the best parameters to get more reliable results [16]. The results of this study are expected to provide a deeper understanding of tourists' perceptions of various aspects of tourism in Palangka Raya City and provide policy recommendations to improve the attractiveness and quality of tourism in the city. This research is limited to analyzing reviews on Google Maps regarding tourist attractions in Palangka Raya City and focuses on sentiments and emotions that are explicitly stated in the review text. Thus, it is expected that this research can contribute to the development of sentiment analysis and support data-driven decision-making in the tourism sector.

II. RELATED WORKS

Sentiment analysis research has rapidly grown, emphasizing the importance of emotional understanding in user interactions. Early research focused on text polarity, but aspect-based approaches are crucial for deeper business insights. Studies by [17] and [18] support this. [19] applied sentiment analysis to Twitter using Naïve Bayes and SVM, with SVM achieving 89% accuracy. [20] evaluated algorithms like Random Forest, KNN, and Naïve Bayes on social media, finding Naïve Bayes most accurate at 92.01%.

Research by [21] explored the polarization of positive and negative sentiments in user reviews on Twitter, using Naïve Bayes (NB), SVM, and logistic regression (LR). The findings showed an LR model accuracy of 77%, higher than SVM (76%) and NB (70%). [22] also compared the performance of Naïve Bayes, SVM, and Random Forest, finding that Naïve Bayes achieved the highest accuracy of 81%, while SVM and Random Forest recorded accuracies of 80% and 76%, respectively. Research by [23] that BERT outperformed traditional machine learning algorithms in sentiment analysis, with accuracy, precision, recall, and F-scores of 85%, 84%, 87%, and 85% on social media datasets, respectively. The importance of the integration between sentiment analysis and emotion detection is clear, as emotion detection can provide more comprehensive insights into user opinions. With the ability to identify specific emotions such as happy, sad, angry, or fearful, understanding the background of the emerging sentiment can be improved [24].

Research by [12] applied various machine learning and deep learning methods to detect emotions in textual content, where the BERT model showed the highest accuracy on the AIT-2018 and ISEAR datasets. Another study analyzed hate speech against Asians on Twitter by building prediction models using machine learning algorithms and deep learning methods such as Long Short-Term Memory (LSTM) and Bidirectional LSTM. Results showed that logistic regression achieved the highest F1 score of 0.72, while BERT recorded the best F1 score among deep learning models with a value of 0.85 [25]Overall, significant advances in sentiment analysis and emotion detection demonstrate the effectiveness of these methods. However, there are opportunities for innovation through the integration of both approaches. This research focuses on developing techniques that combine sentiment analysis and emotion detection to improve understanding of user interaction dynamics and provide more relevant recommendations for product and service development.

III. MATERIALS AND METHODS

The method used in this research is shown in Figure 1.

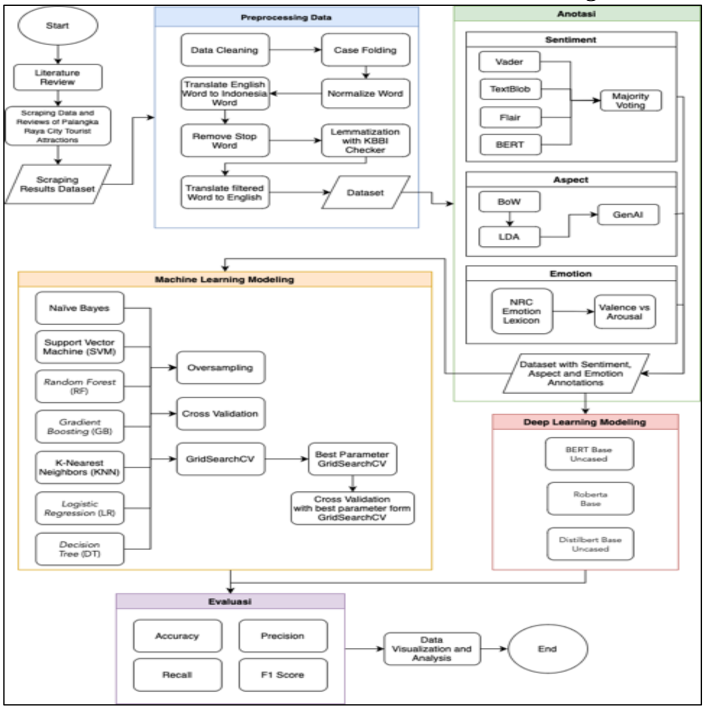

Figure 1: Research Method.
Source: Authors, (2025).

3.1 Literature Review

This research begins with an in-depth literature study on sentiment analysis, specifically on the aspects and emotions of user reviews on Google Maps using machine learning and deep learning.

3.2 Data collection

The dataset in this research is taken from reviews of tourist attractions in Palangka Raya City found on Google Maps during the period 2018 to 2024 through the scraping method. The data collected includes location, user, time, and comments. The amount of data collected is 5,221 data. Examples of data collected can be seen in Table 1.

Table 1: Sample of scraped reviews for tourist attractions in Palangkaraya from Google Maps.

Location	User	Time	Review
bukit baranahu	PRATAMA RAHMATU LLAH	setahun lalu	Lokasi buat pendakian cukup terjal Pamandangan di atas tower sangat bagus 👍👍 Tapi sayang jalannya banyak yang licin dan terjal dan diatas nggak ada tempat buat berlindung kalau hujan nya kencang …
bukit baranahu	Yosua Krisopras Kaharap	5 tahun lalu	bukit yang bagus dijadikan spot camping dgn puncak yg datar dan luas

Source: Authors, (2025).

3.3 Data Preprocessing

The collected dataset will go through a comprehensive preprocessing process to prepare it for subsequent analysis [26]. The preprocessing steps include data cleaning, case folding, word normalization, translation from English to Bahasa Indonesia, stopword removal, lemmatization, and translation back to English [27]. Data cleaning is performed to remove noise and errors, such as excess spaces, emojis, hashtags, mentions, URLs, non-alphanumeric characters, dates, numbers, date time entries, extra punctuation, and null-valued reviews. Case folding is applied to equalize all text characters to lowercase [28]. Word normalization was performed by converting nonstandard words and abbreviations to standard forms using a list from an Excel file [29]. Text translation from English to Indonesian was performed using the deep_translator library. Stopword removal was performed with the literature library. Lemmatization is applied using the stanza library to convert words into their basic form, with verification from the Big Indonesian Dictionary (KBBI) contained in the Excel file [29]. The normalized text is then translated back to English. This preprocessing process resulted in a clean dataset consisting of 4,113 data points. The processing results are shown in Table 2.

Table 2: Sample Data Processing.

Preprocessing	Input	Output
Data Cleaning	good views,sunrisenya mantap jiwa⍰	good views sunrisenya mantap jiwa
Case Folding	good views sunrisenya mantap jiwa	good views sunrisenya mantap jiwa
Normalize Word	good views sunrisenya mantap jiwa	bagus pemandangan matahari terbit mantap jiwa
Translate English Word to Indonesia Word	bagus pemandangan matahari terbit mantap jiwa	pemandangan matahari terbit yang indah menenangkan jiwa
Removal Stopword	pemandangan matahari terbit yang indah menenangkan jiwa	pemandangan matahari terbit indah menenangkan jiwa
Lemmatization	pemandangan matahari terbit indah menenangkan jiwa	pandang matahari terbit indah tenang jiwa
Translate to English	pandang matahari terbit indah tenang jiwa	look beautiful sunrise calm soul

Source: Authors, (2025).

3.4 Annotation

The processed dataset is then labeled or annotated based on sentiment, aspect, and emotion.

a. Sentiment Annotation

Sentiment analysis is performed using a combination of four libraries, namely VADER, TextBlob, Flair, and BERT. Each library has a different approach to determining sentiment; VADER uses composite scores, TextBlob uses polarity scores, and Flair and BERT use probability scores [14]. The results from the four libraries are then consolidated using majority voting to get more accurate results and reduce the bias of the single method [30]. While sentiment is generally categorized into positive, negative, and neutral, this study only labeled the data as positive or negative to simplify the analysis and focus on extreme emoticons [31]. The sentiment annotation results are shown in Table 3.

Table 3: Examples of sentiment annotation result.

Content	Sentiment_VADER	Sentiment_TextBlob	Sentiment_Flair	Sentiment_BERT	Sentiment_majority voting
look beautiful sunrise calm soul	Positive	Positive	Positive	Positive	Positive
good hill camp wide flat peak	Positive	Positive	Negative	Negative	Negative

Source: Authors, (2025).

b. Aspect Annotation

The aspects in the text or reviews in the dataset are analyzed using the Bag of Words (BoW) method to convert the text into numerical vectors [32]. These vectors are then analyzed using the gensim library. To determine the optimal number of topics, the Latent Dirichlet Allocation (LDA) algorithm was applied, focusing on coherence measures that assess the relevance and relationship of words within each topic [33], [34]. Important parameters in the LDA model include corpus, num_topics, id2word, random_state, update_every, chunksize, and passes [35]. In this study, the pass parameter is set at 10 to ensure stable model convergence. Once the topics are identified, the text is categorized based on the contribution of the words to the topics. To deepen the analysis, the Generative AI (GenAI) model with the Gemini-Pro approach was used to extract relevant information according to the topics identified by LDA [36]. This method provides an in-depth analysis of the relationship between words and topic relevance, which is useful for applications such as sentiment analysis [37].

The results identified two main aspects of the topics extracted using LDA and GenAI, namely city and attractions and nature and tourism. In the city and attractions aspect, this topic includes words and phrases related to cities, landmarks, and places of interest. Words such as city, park, bridge, road, and kahayan appear frequently, highlighting how cities promote their attractions and how travelers experience urban locations. In contrast, the Nature and Tourism aspect deals with natural elements and tourism activities related to nature. Words such as place, good, relax, and view appear frequently in this topic, illustrating how tourism focuses on natural beauty and experiences related to the natural environment. The results of the aspect annotation are shown in Table 4.

Table 4: Aspect Based Sentiment Analysis (ABSA).

Dominant topic	Perc_Contrib	Topic_Keywords	Review	Aspect
1	0.92000001 67	place, good, beautiful, nice, relax, family, cool, view, need, comfortable	look beautiful sunrise calm soul	Nature and Tourism
0	0.50700002 91	palangka, raya, city, tourist, park, bridge, national, road, kahayan, attraction	good hill camp wide flat peak	City and Attractions

Source: Authors, (2025).

c. Emotion Annotation

Emotion annotation plays an important role in sentiment analysis by providing deep insight into the sentiment in the text. At this stage, the review text was analyzed using the NRC Emotion Lexicon, which allows the identification of dominant emotions and the classification of emotions into valence and arousal categories. Based on Plutchik's emotion model, this lexicon classifies emotions into eight basic categories: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust [13]. This lexicon also enables the classification of emotions with high accuracy [11]. The results of emotion annotation can be seen in Table 5.

Table 5: Emotions Annotation Results.

Steps	Result
Review	look beautiful sunrise calm soul
Sentiment	Positive
Aspect	Nature and Tourism
Emotions	{'anticipation': 0.0, 'joy': 0.08333333333333333, 'surprise': 0.0, 'trust': 0.0}
Dominant_Emotion	joy
Valence	Positive
Arousal	High

Source: Authors, (2025).

The level of customer satisfaction is determined based on the combination of valence, arousal, and dominant emotion in the dataset. Feedback with positive valence usually indicates high satisfaction, while negative valence indicates dissatisfaction [38]. Common sentiment categories in customer feedback analysis include satisfied, very satisfied, dissatisfied, and very dissatisfied [39]. These sentiment categories help businesses understand customers, make better decisions, and increase sales and loyalty. The mapping of the combination of valence, dominant emotion, and arousal to the level of customer satisfaction can be seen in Table 6.

3.5 Classification Model

The sentiment, aspect, and emotion annotated datasets were combined to build classification models using machine learning and deep learning algorithms. The goal was to identify the best model to classify the sentiment aspects and emotions in the dataset. For the machine learning approach, several methods are used, namely Logistic Regression (LR), Naive Bayes (NB), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gradient Boosting (GB), and Decision Tree (DT) [40]. In these experiments, several techniques are applied to improve model performance, such as random oversampling, cross-validation, GridSearchCV, and cross-validation using the best parameters from GridSearchCV [16]. Random oversampling is used to balance the classes by adding samples from minority classes, thereby reducing bias towards the majority class and improving model accuracy on underrepresented classes [41]. Cross-validation is performed to evaluate model accuracy and reduce the risk of overfitting by dividing the data into subsets (folds) for training and testing [42]. GridSearchCV is used to find the best combination of parameters for each algorithm so that the most suitable parameters can be found to improve model performance [43]. After the optimal parameters are found with GridSearchCV, cross-validation is performed again to ensure more reliable and consistent results [44].

For the deep learning approach, pre-trained BERT models were tested, namely BERT-base-uncased, Roberta-base, and Distilbert-base-uncased [45]. The models were trained with a predefined number of parameters. The training was conducted for 50 epochs, with early stopping using patience 5, which stops training if there is no improvement in model performance for 5 consecutive epochs [46]. This helps prevent overfitting and saves training time. The learning rate is set at 2e-5 to ensure the model can learn stably without getting stuck at a local minimum, while the batch size is set at 128 for efficiency in data processing and optimization of GPU memory usage. The maximum input length is set at 512 tokens, which is the maximum length of input that can be processed by the BERT model, to ensure that important information in the text is not truncated. Gradient accumulation is applied to overcome the GPU memory limitation by updating the parameters after several batches, thus enabling model training with a large batch size without requiring a very large GPU memory [47]. With this approach, it is expected that the model can utilize the advantages of both machine learning and deep learning techniques to achieve optimal sentiment classification results.

3.6 Evaluation and Visualization

The resulting model is evaluated using accuracy, recall, precision, and F1 score metrics to determine the best performance [48]. The final stage involves visualization of the results in graphical form as well as analysis of aspects, sentiments, and emotions from user reviews in relation to their relevance to sustainable tourism [49].

IV. RESULTS AND DISCUSSIONS

4.1 Sentiment Annotation Results

The sentiment distribution based on the dataset of 4,113 reviews using VADER shows that 95.4% of the reviews are categorized as positive and 4.6% as negative. TextBlob showed similar results to VADER, albeit with slightly more negative reviews, at 6.9%, and positive reviews at 93.1%. Flair categorized

86.1% of reviews as positive and 13.9% as negative, while BERT categorized 84.4% of reviews as positive and 15.6% as negative. The result of the sentiment analysis was determined through majority voting, which combines the judgments from all libraries and results in a more balanced distribution in text classification. In this process, 88% of the reviews were categorized as positive and 12% as negative. Thus, although there are variations in sentiment analysis results between different methods, all tools show that most reviews of tourist attractions in Palangka Raya city are categorized as positive. The comparison of the sentiment annotation results using VADER, TextBlob, Flair, BERT, and majority voting, shown in Table 6, provides an overall picture of the sentiment classification of the text.

Table 6: Sentiment Distribution by Method.

Labeling technique	Sentiment label	
	Positive	Negative
VADER	3925	188
TextBlob	3828	285
Flair	3541	572
BERT	3469	644
Majority voting	3619	494

Source: Authors, (2025).

The aspect distribution of tourist attractions reviews in Palangka Raya City consists of 3,015 reviews for nature and tourism and 1,098 reviews for city and attractions. The sentiment distribution for the City and Attractions aspect includes 974 positive sentiments and 124 negative sentiments. Meanwhile, the nature and tourism aspect have 2,645 positive sentiments and 370 negative sentiments. The aspect distribution is shown in Figure 2, while the sentiment distribution per aspect is shown in Figure 3.

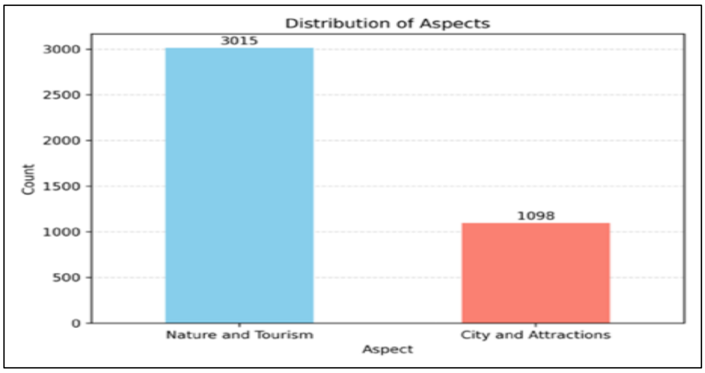

Figure 2: Distribution of Aspects.
Source: Authors, (2025).

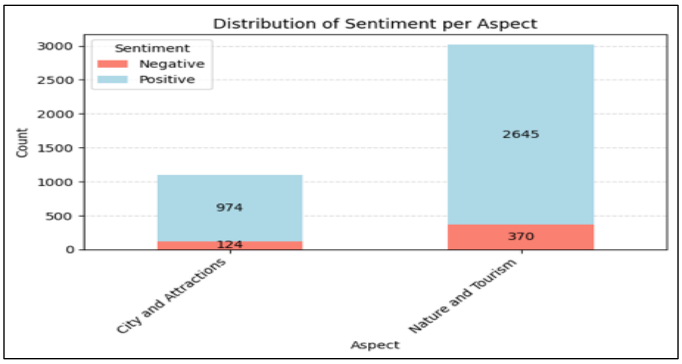

Figure 3: Sentiment per Aspect.
Source: Authors, (2025).

The frequency distribution of emotions in the dataset shows that anticipation is the most frequent emotion, with 2,405, followed by joy with 857. Emotions such as anger at 326, trust at 20, fear at 69, disgust at 58, sadness at 41, and surprise at 37 appear less frequently compared to anticipation and joy. The frequency distribution of these emotions can be seen in Figure 4. Sentiment analysis of the nature and tourism aspect shows the dominance of joy and anticipation emotions, reflecting high enthusiasm for nature and tourism. In contrast, in the city and attractions aspect, the distribution of emotions is more diverse. In addition to joy and anticipation, the emotion of trust is also significant, indicating trust in information related to cities and attractions. Details of sentiment distribution and dominant emotions by aspect are shown in Figure 5.

Dominant emotions are mapped into the dimensions of valence (positive or negative) as well as arousal (high or low). Positive emotions with high arousal, such as excitement and anticipation, appear frequently, indicating a pleasant and arousing experience. Negative emotions with high arousal, such as anger, also arise frequently but with high intensity. In contrast, low-arousal negative emotions, such as sadness and disgust, appear with lower intensity. Finally, positive emotions with low arousal, such as trust, tend to appear with a calmer intensity. This mapping can be seen in Figure 6.

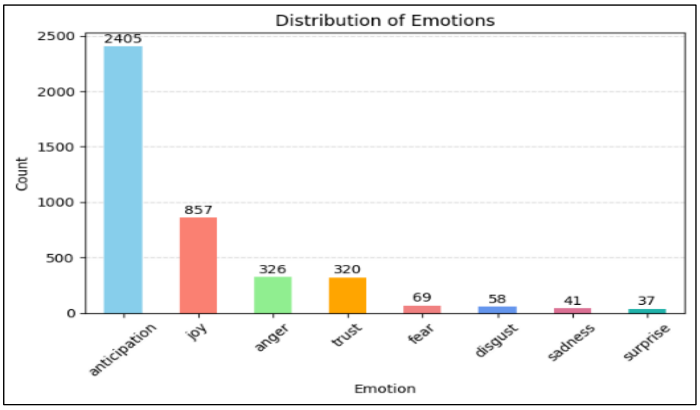

Figure 4**:** Distribution of emoticons.
Source: Authors, (2025).

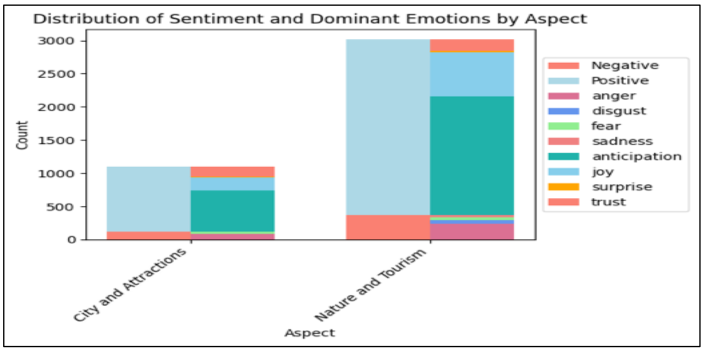

Figure 5: Sentiment and Emotion Distribution by Aspect.
Source: Authors, (2025).

The level of customer satisfaction is determined by analyzing the combination of valence, arousal, and dominant emotions contained in the dataset. The analysis results show that 2,405 customers rated tourist attractions in Palangka Raya City as very satisfying, while 1,214 customers were satisfied. In contrast, the number of dissatisfied and very dissatisfied customers is relatively small, at 168 and 326, respectively. This finding indicates that despite some complaints, the majority of customers experience high levels of satisfaction, reflecting good performance in meeting visitors' expectations. In addition, the level of customer satisfaction by aspect was analyzed. For the City and Attractions aspect, the number of very satisfied visitors was 619, satisfied was 355, dissatisfied was 40, and very dissatisfied was 84. Meanwhile, for the nature and tourism aspect, the number of very satisfied visitors reached 1,786, 859 satisfied, 128 dissatisfied, and 242 very dissatisfied. Details of these results can be seen in Figure 7.

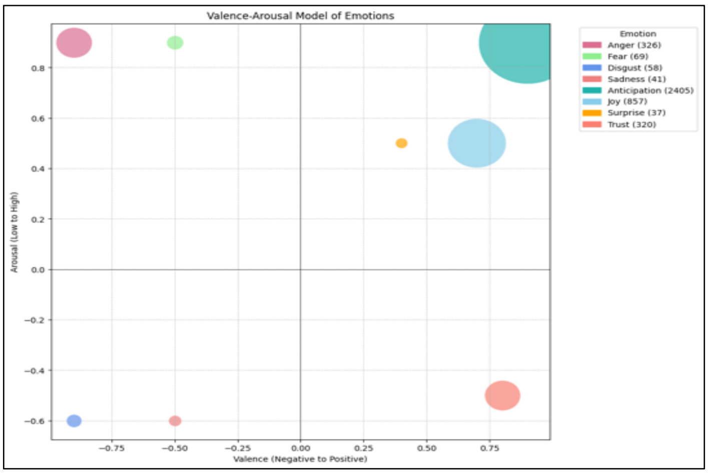

Figure 6: Emotion Distribution by Valence and Arousal.
Source: Authors, (2025).

Figure 7: Distribution of Customer Satisfaction.
Source: Authors, (2025).

4.2 Comparison of Classification Model Results

This section presents a comparative analysis of the performance of various machine learning (ML) and deep learning (DL) algorithms in classifying datasets that have been annotated for sentiment, aspect, and emotion. The purpose of this analysis is to identify the best model for each of these classification tasks. To improve the performance of machine learning models, several techniques were applied, namely random oversampling, cross-validation, GridSearchCV, and cross-validation with the best parameters from GridSearchCV.

The results of aspect classification show that the Support Vector Machine (SVM) algorithm provides excellent performance compared to other algorithms, especially when the cross-validation technique, GridSearchCV, and cross-validation with optimal parameters from GridSearchCV are applied. With the cross-validation technique, SVM achieved 94.42% accuracy, 94.44% precision, 94.42% recall, and 94.31% F1-score. The use of GridSearchCV further improved the performance of SVM, with an accuracy of 98.10%, precision of 98.13%, recall of 98.10%, and F1-score of 98.10%. The application of cross-validation with optimal parameters from GridSearchCV resulted in the highest accuracy value for SVM, which was 98.60%, with 98.62% precision, 98.60% recall, and 98.60% F1-score. On the other hand, when applying the Random Oversampling technique, the Logistic Regression algorithm showed better performance compared to SVM and other algorithms, with accuracy, precision, recall, and an F1-score of 95.15% each. Complete data regarding the results of this evaluation can be seen in Table 7.

Table 7: Comparing Machine Learning Algorithms for Aspect-Based Sentiment.

Aspect					
Algorithm	**Technique**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
LR	Random Oversampling	95,15	95,15	95,15	95,15
	Cross-validation	96,81	96,86	96,82	96,81
	GridSearchCV	96,82	96,86	96,82	96,81
	Cross-validation with best parameter from GridSearchCV	98,38	98,41	98,38	98,38
NB	Random Oversampling	93,20	93,59	93,20	93,31
	Cross-validation	96,12	96,26	96,12	96,11
	GridSearchCV	96,12	96,26	96,12	96,11
	Cross-validation with best parameter from GridSearchCV	96,28	96,40	96,28	96,28
RF	Random Oversampling	90,53	90,69	90,5	90,60
	Cross-validation	96,56	96,56	96,54	96,56
	GridSearchCV	96,56	96,56	96,54	96,56
	Cross-validation with best parameter from GridSearchCV	96,89	96,82	96,65	96,57
SVM	Random Oversampling	94,42	94,44	94,42	94,31
	Cross-validation	98,10	98,13	98,10	98,10
	GridSearchCV	98,10	98,13	98,10	98,10
	Cross-validation with best parameter from GridSearchCV	98,60	98,62	98,60	98,60
KNN	Random Oversampling	87,38	87,62	87,38	87,48
	Cross-validation	91,68	91,86	91,68	91,67
	GridSearchCV	91,68	91,86	91,68	91,67
	Cross-validation with best parameter from GridSearchCV	95,07	98,62	95,07	95,06
GB	Random Oversampling	87,86	87,86	87,86	87,86
	Cross-validation	88,81	89,23	88,86	88,85
	GridSearchCV	88,81	89,23	88,86	88,85
	Cross-validation with best parameter from GridSearchCV	96,54	96,66	96,45	96,54
DT	Random Oversampling	86,65	86,95	86,65	86,77
	Cross-validation	94,42	94,73	94,66	94,39
	GridSearchCV	94,42	94,73	94,66	94,39
	Cross-validation with best parameter from GridSearchCV	94,42	94,69	94,37	94,37

Source: Authors, (2025).

The results of the sentiment classification evaluation using machine learning show that the Random Forest algorithm provides better performance compared to other algorithms in the Random Oversampling and GridSearchCV techniques. In the Random

Oversampling technique, Random Forest achieved 91.26% accuracy, 90.18% precision, 91.26% recall, and 90.20% F1-score. For the GridSearchCV technique, the algorithm recorded 95.92% accuracy, 95.91% precision, 95.92% recall, and 95.91% F1-score. In contrast, in the cross-validation technique and cross-validation with the optimal parameters of GridSearchCV, the support vector machine (SVM) algorithm showed superior performance with accuracy, recall, and F1-score of 99.09% and precision of 99.11% for the cross-validation technique, respectively. For cross-validation with the best parameters from GridSearchCV, SVM achieved the highest accuracy, precision, recall, and F1-score of 99.86%. Details of these results can be seen more clearly in Table 8.

Table 8: Comparing Machine Learning Algorithms for Sentiment Classification.

Emotion					
Algorithm	**Technique**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
LR	Random Oversampling	78,16	81,73	78,16	79,55
	Cross-validation	97,29	97,31	97,29	97,24
	GridSearchCV	72,80	73,11	72,80	72,95
	Cross-validation with best parameter from GridSearchCV	98,87	98,89	98,87	98,84
NB	Random Oversampling	57,04	73,68	57,04	61,24
	Cross-validation	89,94	90,18	89,94	89,25
	GridSearchCV	69,94	69,19	69,94	69,07
	Cross-validation with best parameter from GridSearchCV	92,07	92,25	92,07	91,58
RF	Random Oversampling	77,43	75,80	77,43	74,98
	Cross-validation	99,14	99,19	99,18	99,20
	GridSearchCV	95,14	95,18	95,14	95,14
	Cross-validation with best parameter from GridSearchCV	99,24	99,28	99,26	99,23
SVM	Random Oversampling	78,64	76,57	78,64	75,40
	Cross-validation	99,41	99,41	99,41	99,41
	GridSearchCV	94,12	94,42	94,12	94,12
	Cross-validation with best parameter from GridSearchCV	99,59	99,59	99,59	99,59
KNN	Random Oversampling	61,41	63,80	61,41	62,32
	Cross-validation	93,72	93,57	93,72	93,48
	GridSearchCV	93,72	93,62	93,72	93,72
	Cross-validation with best parameter from GridSearchCV	96,487	96,65	96,48	96,31
GB	Random Oversampling	78,40	79,71	78,40	78,81
	Cross-validation	94,21	94,48	94,17	94,18
	GridSearchCV	85,17	87,20	85,06	85,00
	Cross-validation with best parameter from GridSearchCV	98,63	98,60	98,61	98,56
DT	Random Oversampling	72,82	73,91	72,82	73,17
	Cross-validation	98,06	98,11	98,04	98,10
	GridSearchCV	94,48	95,30	94,75	94,62
	Cross-validation with best parameter from GridSearchCV	98,22	98,20	98,10	98,12

Source: Authors, (2025).

For the deep learning model, the evaluation results show that Roberta-base outperforms BERT-base-uncased and DistilBERT-base-uncased in aspect and sentiment classification. Roberta-base achieved 91.26% accuracy, 91.38% precision, 91.26% recall, and 91.31% F1-score for aspect classification, and 88.63% accuracy, 87.84% precision, 88.83% recall, and 88.25% F1-score for sentiment classification. In contrast, for emotion classification, DistilBERT-base-uncased shows better performance compared to BERT-base-uncased and Roberta-base, with an accuracy value of 68.20%, precision of 72.23%, recall of 68.20%, and F1-score of 69.75%. The detailed performance of the deep learning model can be seen in Table 9.

Table. 9: Deep Learning Algorithms Performance Comparison

Category	**Algorithm (%)**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F-1 Score (%)**
Aspect	BERT-base-uncased	88,35	89,90	88,35	88,74
	Roberta-base	91,26	91,38	91,26	91,31
	DistilBERT-base-uncased	89,32	91,16	89,32	89,73
Sentiment	BERT-base-uncased	86,65	87,26	86,65	86,92
	Roberta-base	88,83	87,84	88,83	88,25
	DistilBERT-base-uncased	88,11	87,64	88,11	87,85
Emotion	BERT-base-uncased	66,75	68,68	66,75	67,55
	Roberta-base	67,72	69,71	67,72	63,38
	DistilBERT-base-uncased	68,20	72,33	68,20	69,75

Source: Authors, (2025).

Sentiment analysis of various methods such as VADER, TextBlob, Flair, and BERT shows that most reviews about tourist attractions in Palangka Raya City are positive. While VADER and TextBlob generally generate a lot of positive reviews, TextBlob records slightly more negative reviews than VADER. Flair and BERT also reflect a similar trend with a higher proportion of negative reviews [50]. Majority voting gave a sentiment distribution of 88% positive and 12% negative, confirming the general trend of positive reviews [30], [50].

Aspect analysis revealed that the Nature and Tourism category received more positive reviews than City and Attractions, although negative reviews were also more frequent in that category. The frequency of emotions in the dataset was dominated by anticipation and joy, while negative emotions such as anger and fear appeared less frequently [51]. The level of customer satisfaction is very high, with 2,405 visitors very satisfied and 1,214 satisfied, indicating that tourist attractions in Palangka Raya successfully meet visitors' expectations. Nature and tourism visitors tend to be very satisfied more often than city and attraction visitors, possibly because experiences in nature are considered more satisfying than urban environments [52].

In terms of machine learning classification models, Random Oversampling, Cross-Validation, and GridSearchCV techniques have been shown to improve the performance of various algorithms, such as Logistic Regression, Naive Bayes, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gradient Boosting, and Decision Tree [44], [53], [54]. In particular, the combination of cross-validation with the best parameters of GridSearchCV resulted in the most significant improvement in model accuracy [16], [55]. Support Vector Machine (SVM) showed superiority over other machine learning algorithms as well as deep learning models, achieving 98.60% accuracy and 98.60% F1-score for aspect classification, 99.86% accuracy and 98.86% F1-score for sentiment analysis, and 99.59% accuracy and 99.59% F1-score for emotion classification [56], [57] However, it should be noted that this study has some limitations. Firstly, the data used is limited to the Google Maps platform, so it may not cover all visitor’s perspectives. In addition, the relatively small amount of data may affect the accuracy of the model. These limitations need to be considered in the interpretation of the results and the development of future research need to be considered in the interpretation of the results and the development of future research.

V. CONCLUSIONS

This study found that the majority of reviews regarding tourist attractions in Palangka Raya City are positive, with 88% positive and 12% negative reviews according to majority voting. The analysis methods used, including VADER, TextBlob, Flair, and BERT, were consistent in showing a positive trend, although there were variations in the proportion of negative reviews. The nature and tourism aspect received more positive reviews compared to city and attractions but also had more negative reviews. This reflects that visitors may enjoy the nature experience more, despite some criticism. Emotion frequencies show a predominance of anticipation and joy, indicating that the visitor experience is generally uplifting. Customer satisfaction levels are also high, with the majority of visitors feeling very satisfied or satisfied. The combination of cross-validation with the best parameters from GridSearchCV resulted in the most significant improvement in model accuracy. In classification, SVM showed the best performance for aspects, while Random Forest excelled in sentiment and Logistic Regression in emotion on specific techniques. Roberta-base and DistilBERT-base-uncased deep learning models showed the best performance in aspect and emotion classification, respectively. These findings provide valuable insights for tourism destination management and marketing strategies, with an emphasis on positive experiences and visitor preferences.

Based on the research results, several strategic steps can be implemented to improve the management of tourist destinations in Palangka Raya City. Improving the quality of experience in the Nature and Tourism category needs to be focused on improving facilities and addressing criticism to reduce complaints. Positive sentiment data should be used in marketing strategies, highlighting natural advantages to attract visitors. In addition, proven effective analysis algorithms, such as support vector machine (SVM), random forest, and logistic regression, can be applied to improve classification accuracy. The use of deep learning models such as Roberta-base is also recommended for more in-depth analysis. Data collection from various sources and continued research on visitor preferences will provide valuable additional insights. These steps are expected to improve visitor experience, optimize marketing, and make effective use of analytics technology for long-term success.

This research opens opportunities for further exploration of sentiment and emotion analysis in tourism. Future research could include analysis in other cities in Indonesia to understand traveler sentiment patterns more broadly. In addition, the development and application of new deep-learning models can improve the accuracy of emotional classification. Integration of multi-source data, such as review data, social media, satisfaction surveys, and transactions, is also important for more comprehensive analysis. This approach is expected to improve the service quality and attractiveness of tourist destinations and strengthen the traveler experience.

VI. AUTHOR’S CONTRIBUTION

Conceptualization: Author One, Author Two, Author Three, Author Four.
Methodology: Author One, Author Two, Author Three, Author Four.
Investigation: Author One, Author Two, Author Three, Author Four.
Discussion of results: Author One, Author Two, Author Three, Author Four.
Writing – Review and Editing: Author One and Author Two.
Resources: Author One, Author Two, Author Three, Author Four.
Supervision: Author One, Author Two, Author Three, Author Four.
Approval of the final text: Author One, Author Two, Author Three, Author Four.

VII. ACKNOWLEDGMENTS

The authors gratefully acknowledge all useful suggestions regarding this work.

VII. REFERENCES

[1] M. A. AL-Barrak and A. I. Al-Alawi, “Sentiment Analysis on Customer Feedback for Improved Decision Making: A Literature Review,” in 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS), IEEE, Jan. 2024, pp. 207–212. doi: 10.1109/ICETSIS61505.2024.10459452.

[2] R. K. Sharma and A. Dagur, “Various methods to classify the polarity of text based customer reviews using sentiment analysis,” in Artificial Intelligence, Blockchain, Computing and Security Volume 2, London: CRC Press, 2023, pp. 107–114. doi: 10.1201/9781032684994-18.

[3] G. Aivatoglou, A. Fytili, G. Arampatzis, D. Zaikis, N. Stylianou, and I. Vlahavas, “End-to-End Aspect Extraction and Aspect-Based Sentiment Analysis Framework for Low-Resource Languages,” 2024, pp. 841–858. doi: 10.1007/978-3-031-47715-7_56.

[4] C. A. Bahri and L. H. Suadaa, “Aspect-Based Sentiment Analysis in Bromo Tengger Semeru National Park Indonesia Based on Google Maps User Reviews,” IJCCS (Indonesian Journal of Computing and Cybernetics Systems), vol. 17, no. 1, p. 79, Feb. 2023, doi: 10.22146/ijccs.77354.

[5] A. Chowanda, R. Sutoyo, Meiliana, and S. Tanachutiwat, “Exploring Text-based Emotions Recognition Machine Learning Techniques on Social Media Conversation,” in Procedia Computer Science, Elsevier B.V., 2021, pp. 821–828. doi: 10.1016/j.procs.2021.01.099.

[6] J. Lee, S. Benjamin, and M. Childs, “Unpacking the Emotions behind TripAdvisor Travel Reviews: The Case Study of Gatlinburg, Tennessee,” International Journal of Hospitality & Tourism Administration, vol. 23, no. 2, pp. 347–364, Mar. 2022, doi: 10.1080/15256480.2020.1746219.

[7] M.-N. Truțescu and D. Nicolaie, “Sentiment Analysis: a Comparative Study of Online Booking Platforms Used for Spa Tourism in Northern Oltenia, Romania,”

Balneo and PRM Research Journal, vol. 15, no. Vol.15, 1, pp. 651–651, Mar. 2024, doi: 10.12680/balneo.2024.651.

[8] J. Vaid, "Sustainable Tourism: A Roadmap for the 2030 Sustainable Development Agenda," 2024, pp. 21–26. doi: 10.1108/S1569-37592024000113B002.

[9] S. Hannafin, "Building a better future through industry, innovation and infrastructure (SDG 9)," in Teaching the Sustainable Development Goals to Young Citizens (10-16 years), London: Routledge, 2024, pp. 259–274. doi: 10.4324/9781003232001-15.

[10] A. K. E. V. S. Stroumpoulis, "Role of Artificial Intelligence and Big Data Analytics In Smart Tourism: A Resource-Based View Approach," Aug. 2022, pp. 99–108. doi: 10.2495/ST220091.

[11] S. Zad, J. Jimenez, and M. A. Finlayson, "Hell Hath No Fury? Correcting Bias in the NRC Emotion Lexicon," in WOAH 2021 - 5th Workshop on Online Abuse and Harms, Proceedings of the Workshop, Association for Computational Linguistics (ACL), 2021, pp. 102–113. doi: 10.18653/v1/2021.woah-1.11.

[12] A. A. Sallam, W. Q. Saif, T. H. Rassem, and B. A. Mohammed, "Detecting Emotional Sentiments in Textual Data Using Various Machine Learning and Deep Learning Techniques," 2024, doi: 10.21203/rs.3.rs-4558175/v1.

[13] T. Chutia and N. Baruah, "A review on emotion detection by using deep learning techniques," Artif Intell Rev, vol. 57, no. 8, Aug. 2024, doi: 10.1007/s10462-024-10831-1.

[14] K. Barik and S. Misra, "Analysis of customer reviews with an improved VADER lexicon classifier," J Big Data, vol. 11, no. 1, p. 10, Jan. 2024, doi: 10.1186/s40537-023-00861-x.

[15] A. M. Shetty, M. F. Aljunid, D. H. Manjaiah, and A. M. S. Shaik Afzal, "Hyperparameter Optimization of Machine Learning Models Using Grid Search for Amazon Review Sentiment Analysis," 2024, pp. 451–474. doi: 10.1007/978-981-99-7814-4_36.

[16] Y. Zhao, W. Zhang, and X. Liu, "Grid search with a weighted error function: Hyper-parameter optimization for financial time series forecasting," Appl Soft Comput, vol. 154, p. 111362, Mar. 2024, doi: 10.1016/j.asoc.2024.111362.

[17] V. V. Tadkal, S. Patil, and S. A. Patil, "Aspect Based Sentiment Analysis," in 2024 5th International Conference for Emerging Technology (INCET), IEEE, May 2024, pp. 1–8. doi: 10.1109/INCET61516.2024.10593116.

[18] A. Siddiqua, V. Bindumathi, G. Raghu, and Y. S. Vamsi Bhargav, "Aspect-based Sentiment Analysis (ABSA) using Machine Learning Algorithms," in 2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), IEEE, Apr. 2024, pp. 1–6. doi: 10.1109/ICDCECE60827.2024.10549140.

[19] M. Shenify, "Sentiment analysis of Saudi e-commerce using naïve bayes algorithm and support vector machine," International Journal of Data and Network Science, vol. 8, no. 3, pp. 1607–1612, Jun. 2024, doi: 10.5267/j.ijdns.2024.3.006.

[20] B. Mridula, A. H. Juliet, and N. Legapriyadharshini, "Deciphering Social Media Sentiment for Enhanced Analytical Accuracy: Leveraging Random Forest, KNN, and Naive Bayes," in 2024 10th International Conference on Communication and Signal Processing (ICCSP), IEEE, Apr. 2024, pp. 1410–1415. doi: 10.1109/ICCSP60870.2024.10543836.

[21] P. Assiroj, A. Kurnia, and S. Alam, "The performance of Naïve Bayes, support vector machine, and logistic regression on Indonesia immigration sentiment analysis," Bulletin of Electrical Engineering and Informatics, vol. 12, no. 6, pp. 3843–3852, Dec. 2023, doi: 10.11591/eei.v12i6.5688.

[22] E. D. Madyatmadja, H. Candra, J. Nathaniel, M. R. Jonathan, and Rudy, "Sentiment Analysis on User Reviews of Threads Applications in Indonesia," Journal Européen des Systèmes Automatisés, vol. 57, no. 4, pp. 1165–1171, Aug. 2024, doi: 10.18280/jesa.570423.

[23] T. Bikku, J. Jarugula, L. Kongala, N. D. Tummala, and N. Vardhani Donthiboina, "Exploring the Effectiveness of BERT for Sentiment Analysis on Large-Scale Social Media Data," in 2023 3rd International Conference on Intelligent Technologies (CONIT), IEEE, Jun. 2023, pp. 1–4. doi: 10.1109/CONIT59222.2023.10205600.

[24] W. S. Ismail, "Emotion Detection in Text: Advances in Sentiment Analysis Using Deep Learning," J Wirel Mob Netw Ubiquitous Comput Dependable Appl, vol. 15, no. 1, pp. 17–26, Mar. 2024, doi: 10.58346/JOWUA.2024.I1.002.

[25] A. Toliyat, S. I. Levitan, Z. Peng, and R. Etemadpour, "Asian hate speech detection on Twitter during COVID-19," Front Artif Intell, vol. 5, Aug. 2022, doi: 10.3389/frai.2022.932381.

[26] V. Handrianus Pranatawijaya, W. Widiatry, and D. J. Lestari, "ENHANCING OPTICAL DISTRIBUTION POINT PLACEMENT: A DECISION SUPPORT SYSTEM INTEGRATING WEIGHTED PRODUCT METHOD, CONTENT-BASED FILTERING, AND LOCATION-BASED SERVICES," Journal of Engineering and Technology for Industrial Applications ITEGAM-JETIA Manaus, pp. 206–213, doi: 10.5935/jetia.v10i50.1329.

[27] C. El Morr, M. Jammal, H. Ali-Hassan, and W. El-Hallak, "Data Preprocessing," 2022, pp. 117–163. doi: 10.1007/978-3-031-16990-8_4.

[28] Md. S. Hossen, "Data Preprocess," in Machine Learning and Big Data, Wiley, 2020, pp. 71–103. doi: 10.1002/9781119654834.ch4.

[29] V. V. Bozhenko and T. M. Tatarnikova, "Application of Data Preprocessing in Medical Research," in 2023 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), IEEE, May 2023, pp. 1–4. doi: 10.1109/WECONF57201.2023.10148004.

[30] Zakaria and A. Sunyoto, "Hybrid Sentiment Analysis: Majority Voting with Multinomial Naive Bayes and Logistic Regression on IMDB Dataset," in 2023 6th International Conference on Information and Communications Technology (ICOIACT), IEEE, Nov. 2023, pp. 448–453. doi: 10.1109/ICOIACT59844.2023.10455937.

[31] A. A. Eponon and K. Dimililer, "A Bert Model with Deep Learning Approach in Natural Language Processing (NLP)," 2024, pp. 23–30. doi: 10.1007/978-3-031-53488-1_3.

[32] U. K. Singh, B. Prabhu Shankar, R. Chinnaiyan, and N. Jain, "Machine Learning-Based Text Categorization with Bag of Words," 2024, pp. 577–587. doi: 10.1007/978-981-97-2839-8_40.

[33] I. D. Sudirman and D. Y. Nugraha, "Understanding Tourist Perspectives of Borobudur through LDA Topic Modelling Using Elbow Method," in 2023 International Conference on Informatics, Multimedia, Cyber and Informations System (ICIMCIS), IEEE, Nov. 2023, pp. 374–379. doi: 10.1109/ICIMCIS60089.2023.10349022.

[34] V. H. Pranatawijaya, N. N. K. Sari, R. A. Rahman, E. Christian, and S. Geges, "Unveiling User Sentiment: Aspect-Based Analysis and Topic Modeling of Ride-Hailing and Google Play App Reviews," Journal of Information Systems Engineering and Business Intelligence, vol. 10, no. 3, pp. 328–339, Oct. 2024, doi: 10.20473/jisebi.10.3.328-339.

[35] J. M. Wheeler, J. Xiong, C. Mardones-Segovia, H.-J. Choi, and A. S. Cohen, "An Investigation of Prior Specification on Parameter Recovery for Latent Dirichlet Allocation of Constructed-Response Items," 2022, pp. 203–215. doi: 10.1007/978-3-031-04572-1_15.

[36] O. M. Al-Janabi, N. H. A. H. Malim, and Y.-N. Cheah, "Aspect Categorization Using Domain-Trained Word Embedding and Topic Modelling," 2020, pp. 191–198. doi: 10.1007/978-981-15-1289-6_18.

[37] C. Daase, C. Haertel, A. Nahhas, A. Zeier, A. Ramesohl, and K. Turowski, "On the Current State of Generative Artificial Intelligence: A Conceptual Model of Potentials and Challenges," in Proceedings of the 26th International Conference on Enterprise Information Systems, SCITEPRESS - Science and Technology Publications, 2024, pp. 845–856. doi: 10.5220/0012707500003690.

[38] G. Meirovich, M. M. Jeon, and L. J. Coleman, "Interaction of normative and predictive expectations in customer satisfaction and emotions," Journal of Marketing Analytics, vol. 8, no. 2, pp. 69–84, Jun. 2020, doi: 10.1057/s41270-020-00078-4.

[39] K. Saroha, M. Sehrawat, and V. Jain, "Machine Learning and Sentiment Analysis for Analyzing Customer Feedback," 2024, pp. 411–440. doi: 10.4018/979-8-3693-0413-6.ch017.

[40] Salsabila, S. M. P. Tyas, Y. Romadhona, and D. Purwitasari, "Aspect-based Sentiment and Correlation-based Emotion Detection on Tweets for Understanding Public Opinion of Covid-19.," Journal of Information Systems Engineering and Business Intelligence, vol. 9, no. 1, pp. 84–94, 2023, doi: 10.20473/jisebi.9.1.84-94.

[41] S. El Kafhali and M. Tayebi, "Generative Adversarial Neural Networks based Oversampling Technique for Imbalanced Credit Card Dataset," in 2022 6th SLAAI

International Conference on Artificial Intelligence (SLAAI-ICAI), IEEE, Dec. 2022, pp. 1–5. doi: 10.1109/SLAAI-ICAI56923.2022.10002630.

[42] J. Li et al., "Quantum k-fold cross-validation for nearest neighbor classification algorithm," Physica A: Statistical Mechanics and its Applications, vol. 611, p. 128435, Feb. 2023, doi: 10.1016/j.physa.2022.128435.

[43] P. J. Forcadilla, "Indoor Pollutant Classification Modeling using Relevant Sensors under Thermodynamic Conditions with Multilayer Perceptron Hyperparameter Tuning," International Journal of Advanced Computer Science and Applications, vol. 14, no. 2, 2023, doi: 10.14569/IJACSA.2023.01402103.

[44] Sukamto, Hadiyanto, and Kurnianingsih, "KNN Optimization Using Grid Search Algorithm for Preeclampsia Imbalance Class," E3S Web of Conferences, vol. 448, p. 02057, Nov. 2023, doi: 10.1051/e3sconf/202344802057.

[45] N. Akhila, S. J. P, and S. K. P, "Comparative Study of Bert Models and Roberta in Transformer based Question Answering," in 2023 3rd International Conference on Intelligent Technologies (CONIT), IEEE, Jun. 2023, pp. 1–5. doi: 10.1109/CONIT59222.2023.10205622.

[46] Z. Zhang, W. Zhu, J. Zhang, P. Wang, R. Jin, and T.-S. Chung, "PCEE-BERT: Accelerating BERT Inference via Patient and Confident Early Exiting," in Findings of the Association for Computational Linguistics: NAACL 2022, Stroudsburg, PA, USA: Association for Computational Linguistics, 2022, pp. 327–338. doi: 10.18653/v1/2022.findings-naacl.25.

[47] Z. Huang, B. Jiang, T. Guo, and Y. Liu, "Measuring the Impact of Gradient Accumulation on Cloud-based Distributed Training," in 2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing (CCGrid), IEEE, May 2023, pp. 344–354. doi: 10.1109/CCGrid57682.2023.00040.

[48] A. Ahmadzadeh, D. J. Kempton, P. C. Martens, and R. A. Angryk, "Contingency Space: A Semimetric Space for Classification Evaluation," IEEE Trans Pattern Anal Mach Intell, vol. 45, no. 2, pp. 1501–1513, Feb. 2023, doi: 10.1109/TPAMI.2022.3167007.

[49] F. Rajani and K. A. Boluk, "A Critical Commentary on the SDGs and the Role of Tourism," Tourism and Hospitality, vol. 3, no. 4, pp. 855–860, Oct. 2022, doi: 10.3390/tourhosp3040053.

[50] A. Mahgoub et al., "Sentiment Analysis: Amazon Electronics Reviews Using BERT and Textblob," in 2022 20th International Conference on Language Engineering (ESOLEC), IEEE, Oct. 2022, pp. 6–10. doi: 10.1109/ESOLEC54569.2022.10009176.

[51] D. Grgić, V. Podobnik, and A. Carvalho, "Deriving and validating emotional dimensions from textual data," Expert Syst Appl, vol. 198, p. 116721, Jul. 2022, doi: 10.1016/j.eswa.2022.116721.

[52] P. Štumpf and T. Kubalová, "Tangible or intangible satisfiers? Comparative study of visitor satisfaction in a nature-based tourism destination in the pre- and during-COVID pandemic," Journal of Outdoor Recreation and Tourism, vol. 46, p. 100777, Jun. 2024, doi: 10.1016/j.jort.2024.100777.

[53] M. Muntasir Nishat et al., "A Comprehensive Investigation of the Performances of Different Machine Learning Classifiers with SMOTE-ENN Oversampling Technique and Hyperparameter Optimization for Imbalanced Heart Failure Dataset," Sci Program, vol. 2022, pp. 1–17, Mar. 2022, doi: 10.1155/2022/3649406.

[54] R. Wijaya, F. Saeed, P. Samimi, A. M. Albarrak, and S. N. Qasem, "An Ensemble Machine Learning and Data Mining Approach to Enhance Stroke Prediction," Bioengineering, vol. 11, no. 7, p. 672, Jul. 2024, doi: 10.3390/bioengineering11070672.

[55] G. Czarnek and D. Stillwell, "Two is better than one: Using a single emotion lexicon can lead to unreliable conclusions," PLoS One, vol. 17, no. 10, p. e0275910, Oct. 2022, doi: 10.1371/journal.pone.0275910.

[56] M. H. I. Ahmad Hapez, N. L. Adam, and Z. Ibrahim, "Performance Analysis of Machine Learning Techniques for Sentiment Analysis," 2021, pp. 205–213. doi: 10.1007/978-3-030-90235-3_18.

[57] A. Lathish and T. Devi, "Comparison of Support Vector Machine and K-Nearest Neighbour Algorithm for Accurate text Classification," in 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA), IEEE, Sep. 2022, pp. 1110–1113. doi: 10.1109/ICIRCA54612.2022.9985594.