Aspect-Based Sentiment Analysis using Adaptive Aspect on Tourist Reviews in Jakarta

[1]Ade Rifqy Setyawan*, [2]Lya Hulliyyatus Suadaa, [3]Budi Yuniarto
[1,2,3]*Department of Statistical Computing, Politeknik Statistika STIS*
[1,2,3]Jl. Otto Iskandar Dinata 64C, Jatinegara, Jakarta Timur, Indonesia
*e-mail: *222011538@stis.ac.id*

(***received***: 6 September 2024, ***revised***: 26 September 2024, ***accepted***: 16 October 2024)

Abstract

Tourism is one of the business fields affected by the Covid-19 pandemic. The decline in the number of tourists, both domestic and foreign, has resulted in the contribution of the tourism business sector to Indonesia's GDP decreasing. The government is now preparing plans to restore and improve tourism in tourist destination areas, one of which is DKI Jakarta in order to increase visits by domestic and foreign tourists. In achieving these goals, this study propose to utilize reviews about tourist attractions in DKI Jakarta from Google Maps and extract public opinion by conducting aspect-based sentiment analysis. Multi-label classification is a common method that is often used in aspect-based sentiment analysis. However, the multi-label approach has limited flexibility in the aspects used. One alternative method that can be used is an adaptive aspect classification method which is more flexible if there are additional new aspects used. This research aims to automate sentiment classification of tourist reviews for each aspect by developing an aspect level sentiment analysis model with an adaptive aspect classification method which will be compared with multi-label classification as a baseline method. The models used in both methods are transfer learning IndoBERT. The adaptive aspect classification method with aspect level sentiment analysis has better performance in comparison to baseline method multi-label classification with accuracy values and F1-score respectively 0.90394 and 0.71504.

Keywords: *tourism, aspect-based sentiment analysis, adaptive aspects, multi-label classification, transfer learning*

1 Introduction

Indonesia has a great tourism potential. Based on data from the Ministry of Tourism and Creative Industries in 2022, the tourism business sector contributed around 3.6% to Indonesia's GDP. This figure has increased by 1.2% from the previous period. The same thing also happened to Indonesia's foreign exchange, where Indonesia's tourism foreign exchange in 2022 increased from IDR 1.191 trillion in 2021 to IDR 1.236 trillion [1].

DKI Jakarta, as one of the gateways for foreign tourists to enter Indonesia, holds a variety of tourism potential that can be utilized to attract tourists to Jakarta for holidays. During the COVID-19 pandemic, in 2020 and 2021, the tourism business in DKI Jakarta was sluggish, where the number of tourists decreased by 82.40% and 27%. This resulted in a decrease in the contribution of tourism to DKI Jakarta's PAD by 53.60% and 6.99% [2].

One of the platforms that provide reviews of a place is Google Maps. Google Maps is now not only used as a guide for visiting a place, but also includes reviews about a place. Since 2015, the number of reviews on Google Maps has grown significantly and quickly compared to other media containing reviews about tourist destinations, such as Tripadvisor, Yelp, and Facebook [3]. The review data in Google Maps can be used to see people's opinions after visiting a place [4].

Based on SemEval 2014 task 4, an international seminar on natural language processing (NLP), four subtasks can be done related to aspect-based sentiment analysis. Among the four subtasks are aspect category detection and sentiment category detection. The aspect here is the topic or feature being discussed in a sentence. Meanwhile, sentiment is an expression of human emotion contained in a sentence. In aspect category detection, the aspect category contained in each review is identified. After obtaining each review's aspects, each review continued with the sentiment classification [5]

ISSN:2302-8149
e-ISSN:2540-9719

A review can contain one or more pieces of information travellers would like to review regarding their travel experience. Reviews that contain more than one aspect will only be classified into one aspect in single-label classification. Therefore, multi-label classification is used to optimise the information obtained from a review so that all aspects contained in a review can be classified [6]. In multi-label classification, the aspects are limited to the defined aspects in the datasets. The resulted models are used only to predict the sentiment of aspects that appear in the training datasets. In this research, we propose a model with adaptive aspect that can learn the semantic of aspects in the datasets then predict sentiment of aspects that have not been seen in the dataset. The model will retrieve information at the aspect level according to the aspects that are input into the model [7].

In this research, we analyse aspect-based sentiment by classifying aspects and sentiments of tourist attraction reviews on Google Maps reviews in DKI Jakarta Province. The five selected attractions with the most visitors in 2022 in DKI Jakarta are Ancol, TMII, Ragunan Zoo, Monas, and Jakarta History Museum. This research is used to determine the condition of the tourist attraction based on the reviews of tourists who have visited the tourist attraction. Based on the analysis results it can be used as a reference for the DKI Jakarta Provincial Government to improve the quality of tourism in DKI Jakarta and attract more tourists to visit Jakarta.

This research aims to automate sentiment classification for each aspect by developing an aspect-based sentiment analysis model with multi-label and aspect-level sentiment analysis methods. Our contributions encompass multiple facets. First, build a dataset for aspect-based sentiment analysis from reviews of tourist attractions in DKI Jakarta on the Google Maps site. Second, aspect-based sentiment analysis was conducted on the reviews of tourist attractions in DKI Jakarta on the Google Maps site using the adaptive aspect, and its performance was compared with the multi-label classification method as a baseline method.

2 Literature Review

There are several related studies on aspect-based sentiment analysis on text data. Pontiki et al [5] researched SemEval-2014 task 4: aspect-based sentiment analysis. This research contains four subtasks in SemEval-2014 task 4 aspect-based sentiment analysis: aspect extraction, sentiment extraction, aspect category detection, and sentiment category detection. CRF model is the best model for aspect extraction and sentiment extraction, with an F1 value of 74.55%. At the same time, the SVM model is the best model for aspect and sentiment detection, with an F1 value of 82.92%.

Laily et al. [8] have conducted research to evaluate tourism in Lamongan Regency. This study applied multi-label classification and LDA to two tourist attractions, namely WBL and Mazoola. multi-label classification uses five tourism-related aspects: price, location, safety, service, and facilities. The NBSVM model is the best model for both scenarios based on the experimental results. Afzaal et al. [9] also conducted research on tourism reviews with a multi-aspect classification. By using the method, it was found that 87% of aspects were successfully identified correctly, with an average accuracy of 90%.

Liu and Zhao [10] conducted research with the title "A BERT-Based Aspect-Level Sentiment Analysis Algorithm for Cross-Domain Text." This study applied an aspect-level sentiment analysis algorithm using the BERT model. Compared to other algorithms in sentiment analysis, the proposed algorithm is proven to have better accuracy and F1 values than other algorithms. In addition, Liu et al. [11] also researched aspect category detection (ACD) and aspect category sentiment analysis using text generation tasks. BERT and BART models are proposed as models in this study. The BERT model became the best, with an F1 value of 90.43%.

Willie et al. [12] have developed an Indonesian version of the BERT model known as IndoBERT. Jayadianti et al. [13] researched aspect-based sentiment analysis before the 2024 election using the IndoBERT model. The data used is tweet data from the Twitter application with the keyword election. This research found that the IndoBERT large-p1 model became the best model with an accuracy value of 83.5% and an F1 value of 88.49%. In addition, Bahri and Suadaa [14] have conducted a multi-label classification on review Google Maps Bromo Tengger Semeru National Park. IndoBERT became the best model with accuracy and f1-score of 91.48% and 71.56%, respectively.

ISSN:2302-8149
e-ISSN:2540-9719

3 Research Method

This research uses an aspect-based sentiment analysis on tourism objects in DKI Jakarta with a multi-label classification approach as a baseline method and an adaptive aspect approach with aspect-level sentiment analysis. The model used in both methods uses one of the IndoBERT transfer learning models. The selected tourism objects are the five leading tourism objects with the highest number of tourist visits in DKI Jakarta in 2022, namely Taman Impian Jaya Ancol, TMII, Ragunan, Monas, and Jakarta History Museum. The stages of research carried out in this study are illustrated in Figure 1

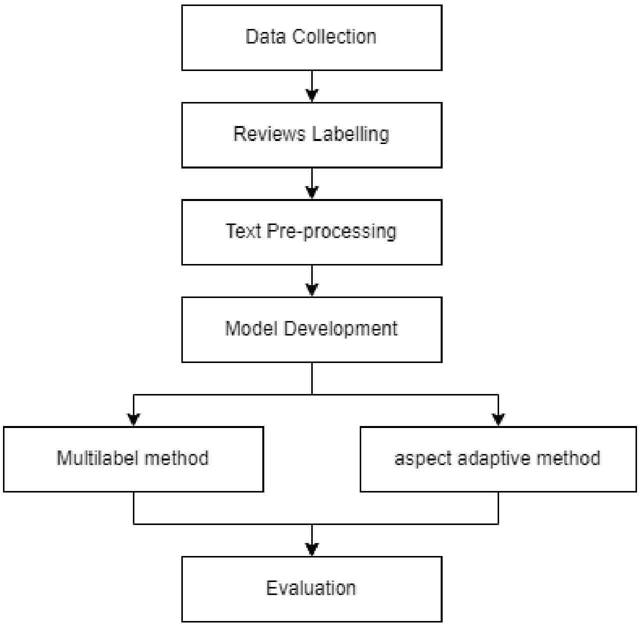

Figure 1. Research flow

3.1 Data Collection

The data used in this study are tourist reviews of attractions in DKI Jakarta, collected from Google Maps using a scraping technique on 26 November 2023. The list of attractions includes the five leading tourist spots in DKI Jakarta, with the most tourist visits in 2022. Python codes were created using the Selenium package to scrape the reviews. The attributes of the retrieved reviews include the name of the attraction, rating, review time, and review. An example of a review of Google Maps used in this study is illustrated in Figure 2

Figure 2 .Example of reviews in google maps

ISSN:2302-8149
e-ISSN:2540-9719

3.2 Data Labelling

After the review data from Google Maps was collected, three annotators labeled the reviews manually according to the guidelines compiled by the researcher. The majority voting mechanism determined the final label. The aspect labeling conducted for each review uses four labels, the main components that must be owned by a Tourist Destination Area (DTW) [15]. The aspect labels used are attractions, amenities, access, and price. For sentiment labeling, the sentiment labels used are positive, negative, and neutral. It will be labeled none for review data that does not contain any of the four aspects above and does not include any of the sentiment labels.

After the annotators label the data, the annotator agreement is calculated using the inter-rater reliability test. This calculation aims to determine the consistency of the annotators' labeling results. The measure used is Krippendorff's alpha, the alpha coefficient for multivalue data [16].

3.3 Text Pre-processing

After collecting data, the next step is data preprocessing. This stage aims to convert raw data into a more structured form and eliminate noise that has the potential to cause errors and provide less accurate results. The preprocessing stages in this study are case folding and removing redundant spaces.

Stopword removal is not included in our preprocessing phase since the process removes some words that are important for classifying sentiment. Punctuation mark elimination is also not conducted since it can remove emoji characters important in sentiment classification.

3.4 Model Development

The model was developed for classifying aspects and sentiments in Google Maps reviews in DKI Jakarta. The IndoBERT is fine-tuned for aspect-level sentiment analysis tasks using multi-label classification and the adaptive aspect method. IndoBERT is one of the BERT models trained on a sizeable Indonesian corpus (Indo4B). This corpus includes formal and informal language such as Indonesian Wikipedia, news articles, social media, blogs, websites, and video recordings with subtitles [12]. The classification model is adapted from IndoNLU with adjustments to the fine-tuning for multi-label classification methods and adaptive aspect methods. The fine-tuning architecture adjustments made to both the multi-label classification method and the adaptive aspect method are illustrated in Figure 3 and Figure 4.

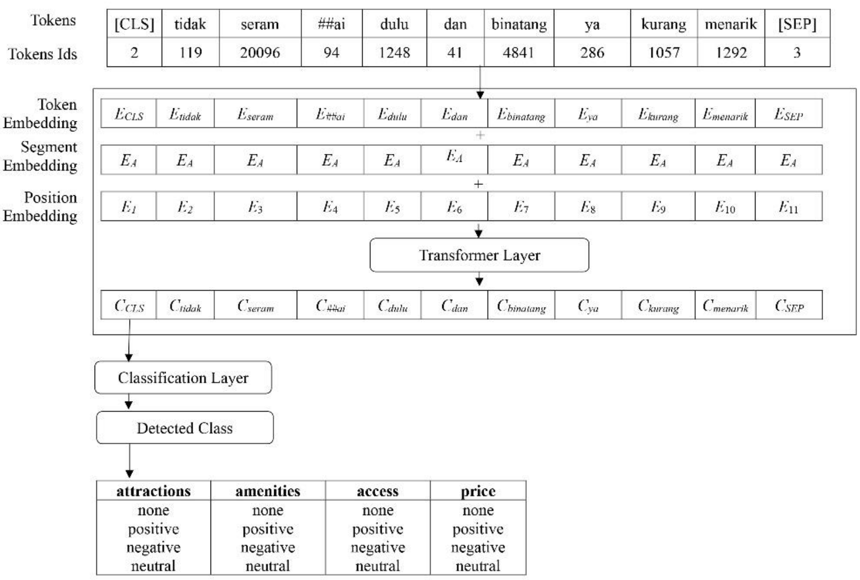

Figure 3. IndoBERT architecture for multi-label classification

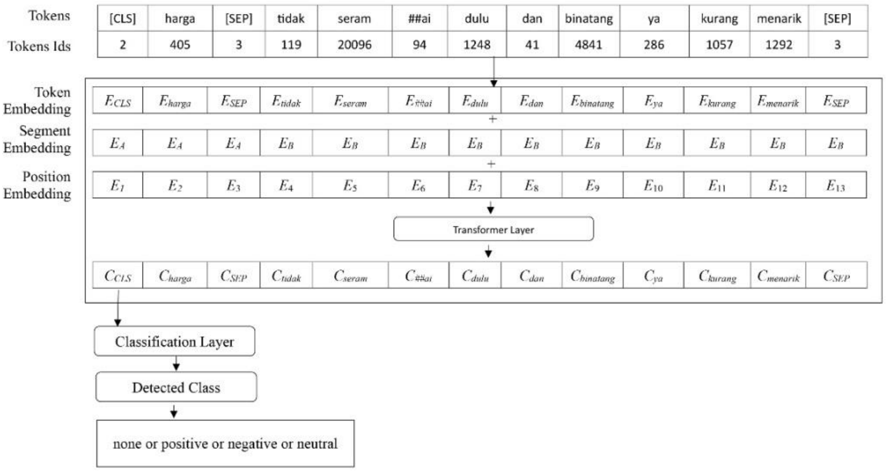

Figure 4. IndoBERT architecture for aspect level sentiment analysis

Figure 5 and Figure 6 illustrate the difference between multi-label classification and aspect-level sentiment analysis. In multi-label classification, the input is the review text, which produces the output of the review's aspects and sentiments. Meanwhile, in aspect-level sentiment analysis, the inputs to the model are aspects and reviews that produce sentiments in the reviews.

Figure 5. Aspect level sentiment analysis illustration

REVIEW			
murah meriah, tapi petunjuk arahnya sebagian ada yang ga bener *It's cheap, but some of the directions aren't right*			
Aspect Based Sentiment			
Attractions	**Amenities**	**Access**	**Price**
none	none	negative	positive

Figure 6. Multi-label classification illustration

3.5 Evaluation

After modeling the dataset, the classification results are evaluated to measure the model's performance on each aspect and sentiment classification. The evaluation will use 5-fold cross-validation. Cross-validation is a method of evaluating and comparing models by dividing data into two parts: training data for model building and testing data for model validation [17]. This approach yields more reliable and precise estimations, while also minimizing the chance of overfitting [18]. In this research, the dataset are divided with a proportion of 80% training data, 10% validation data, and 10\% testing data. Due to imbalanced data conditions, cross-validation is combined with stratified

ISSN:2302-8149
e-ISSN:2540-9719

random sampling called stratified k-fold cross-validation, where the training, testing, and validation datasets have the same class proportion [19]. The metrics used to measure the model's performance were built using a confusion matrix with calculation of accuracy, precision, recall, and F1-Score [20].

$$Accuracy = \frac{TP+TN}{Total} \tag{1}$$

$$Precision = \frac{TP}{FN+TP} \tag{2}$$

$$Recall = \frac{TP}{FN+TP} \tag{3}$$

$$F1 - Score = 2\ x\ \frac{precision\ x\ recall}{precision+recall} \tag{4}$$

4 Results and Analysis

In this section, we present and analyze the results of our sentiment analysis from reviews of the five most visited tourist attractions in DKI Jakarta during 2022, the performance comparison between classification methods, and the prediction results on new review dataset with the selected model.

4.1 Data Collection

The collected data is 1920 reviews from users of Google Maps on five tourist attractions with the most visitors in DKI Jakarta during 2022: Taman Impian Jaya Ancol, Ragunan, TMII, Monas, and Jakarta History Museum. Then, language detection was carried out using the langdetect library. However, due to limitations in the library, some reviews that use Indonesian are not standardly identified as other languages, so manual detection is necessary on reviews that are not identified as Indonesian and English. It can be seen that out of 1920 reviews, 83.4\% were in Indonesian, 14.8\% were in English, and 3.8\% were other than Indonesian and English.

4.2 Data Labelling

Three annotators labeled the data following the labeling guidelines. By using majority voting, data on which at least two annotators disagree will be eliminated, resulting in labeled data for 1866 reviews. In Figure 7, the most reviewed aspect is attraction.

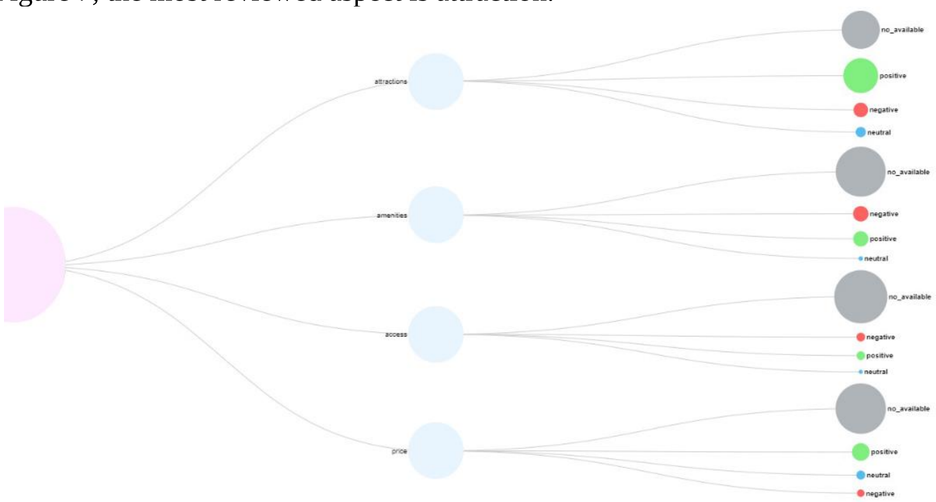

Figure 7. Sentiment distribution for each aspect

Then, the alpha value is compared to see the inter-rater reliability of the labeled data. The higher the alpha value, the higher the level of annotator agreement on the data. The results of the inter-rater reliability test are shown in Table 1.

ISSN:2302-8149
e-ISSN:2540-9719

Table 1. Inter-rater reability test

Aspect	Alpha
Attractions	0.600
Amenities	0.604
Access	0.607
Price	0.814

In Table 1, the attraction, amenities, and access labels have an alpha value of 0.6, which means that the labels on these aspects are quite consistent. The alpha value is quite high at 0.812, which means that the label on the price aspect is consistent.

4.3 Model Development

In the model development stage, transfer learning is used for multi-label classification and aspect-level sentiment analysis. The transfer learning is conducted from the pre-trained IndoBERT model using the PyTorch library and the transformer using Adam as the optimization algorithm with a batch size of 8 and a learning rate of 1×10^{-5} as the default parameters.

4.4 Evaluation

Table 2 shows the results of the experiments conducted using 5-fold cross-validation. Using the IndoBERT model, the adaptive aspect classification method with aspect-based sentiment analysis performs better than the multi-label classification method with accuracy and F1-score values of 90.39% and 71.20%, respectively. It indicates that the aspect adaptive method can capture the meaning of sentences up to the aspect level compared to the multi-label classification method. This result follows the findings of Liu & Zhao [10], which state that the aspect-level sentiment analysis method can extract meaning and information from sentences down to the aspect level better than using other methods. Thus, the method for predicting the new review data uses the adaptive aspect method with aspect-level sentiment analysis.

Table 2. Experimental result of classification model using 5-fold cross validation

Model	Accuracy	Precision	Recall	F1-score
Multi-Label Classification				
IndoBERT	0.8747	0.7323	0.6585	0.6813
Aspect-Level Sentimen Analysis Classification				
IndoBERT	**0.9039**	**0.7525**	**0.6913**	**0.7105**

Table 3 and Table 4 show the distribution of the classification scores of each sentiment for each aspect in the IndoBERT fine-tuned model with multi-label classification and aspect-level sentiment analysis. The F1-score generated from both methods varies because the available dataset is not balanced for all aspect labels and sentiments. The F1-score value for each label and aspect produced by the aspect-level sentiment analysis method is better than the multi-label classification method. When compared, the aspect-level sentiment analysis method performs better, especially in attractions, amenities, and access, than the multi-label classification results. Multi-label classification has better performance in classification on the aspect of price but with a slight difference of F1-score.

Table 3. Multi-label classification results for each label on indobert using 5-fold cross validation

Aspect		Precision	Recall	F1-Score	Support
Attractions	Positive	0.7483	0.8443	0.7802	77
	Negative	0.6327	0.5235	0.5695	12
	Neutral	0.5500	0.4166	0.4111	7
	None	0.8422	0.7849	0.8098	93
Amenities	Positive	0.7118	0.5460	0.6025	13

 ISSN:2302-8149
e-ISSN:2540-9719

	Negative	0.7548	0.6241	0.6741	16
	Neutral	0	0	0	1
	None	0.9277	0.9706	0.9482	156
Access	Positive	1	0,3428	0.4655	6
	Negative	0.3000	0.0650	0.1015	5
	Neutral	0	0	0	1
	None	0.9484	1	0.9735	176
Price	Positive	0.8584	0.8745	0.8610	19
	Negative	1	0.5242	0.6632	5
	Neutral	0.5700	0.5600	0.5644	4
	None	0.9623	0.9862	0.9802	160

Table 4. Adaptive aspect classification results for each label on indobert using 5-fold cross validation

Aspect		Precision	Recall	F1-Score	Support
Attractions	Positive	0.8112	0.8403	0.8239	77
	Negative	0.6947	0.6278	0.6489	12
	Neutral	0.6534	0.3701	0.3701	7
	None	0.8482	0.8515	0.8485	93
Amenities	Positive	0.7618	0.5586	0.6167	15
	Negative	0.7046	0.6423	0.6519	16
	Neutral	0	0	0	1
	None	0.9291	0.9578	0.9427	156
Access	Positive	0.9500	0.5600	0.6788	6
	Negative	0.8000	0.4266	0.5071	5
	Neutral	0.1000	0.2000	0.1334	1
	None	0.9692	0.9943	0.9815	176
Price	Positive	0.8480	0.8886	0.8607	19
	Negative	0.7733	0.5576	0.6158	5
	Neutral	0.5934	0.6900	0.6367	4
	None	0.9814	0.9775	0.9793	160

4.5 Prediction with Selected Model

Predictions are made on a new review dataset collected on May 9, 2024, with 592 reviews of the Pari Island tourist attraction in Kepulauan Seribu. The method used for prediction is the aspect-level sentiment analysis method. The aspect-level sentiment analysis method can adjust input aspects to what will be used. This prediction process uses tourism aspects sourced from the WTO: attractions, amenities, accessibility, price, image, and human resources [3]. Two additional aspects do not exist in the training data: image and human resource.

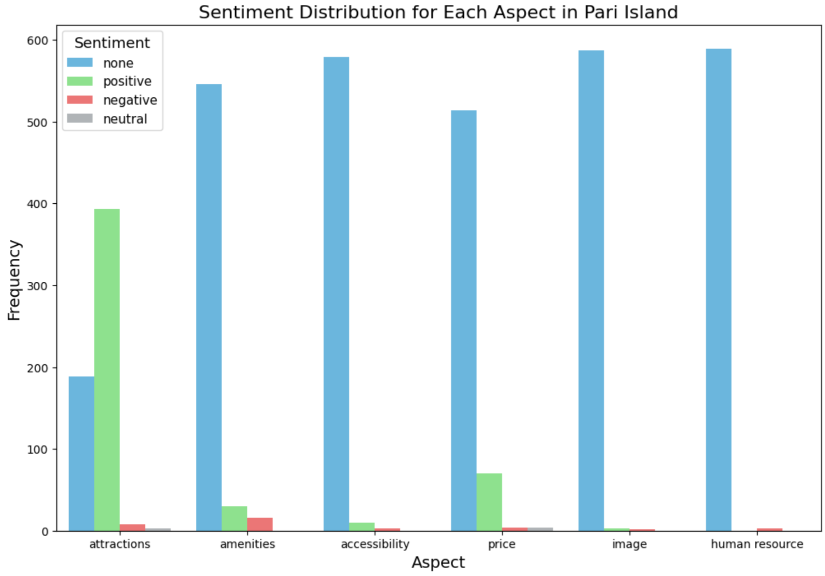

Figure 8. Sentiment distribution for each aspect in pari island

Figure 8 shows the distribution of sentiment predictions per aspect of the Pari Island tourist attraction. It can be seen that the aspect often discussed is the attraction aspect. Attraction, amenity, accessibility, and price have more positive sentiments than negative and neutral sentiments. Image and human resources tend to have "none" sentiments that are more dominant than others.

Table 5. Prediction Results of Adaptive Aspect Method with Aspect Level Sentiment Analysis

Aspect		Precision	Recall	F1-Score	Support
Attractions	Positive	0.5769	0.5556	0.5660	27
	Negative	0	0	0	1
	Neutral	0	0	0	0
	None	0.2000	0.2143	0.2069	14
Amenities	Positive	0	0	0	5
	Negative	0	0	0	2
	Neutral	0	0	0	0
	None	0.8250	0.9429	0.8800	35
Accessibility	Positive	0	0	0	1
	Negative	0	0	0	1
	Neutral	0	0	0	3
	None	0.8810	1	0.9367	37
Price	Positive	0.5000	0.5000	0.5000	6
	Negative	0	0	0	0
	Neutral	0	0	0	1
	None	0.8857	0.8857	0.8857	35
Image	Positive	0	0	0	33
	Negative	0	0	0	3
	Neutral	0	0	0	0
	None	0.1463	1	0.2533	6
Human Resource	Positive	0	0	0	3
	Negative	0	0	0	1
	Neutral	0	0	0	0
	None	0.9048	1	0.9500	38

ISSN:2302-8149
e-ISSN:2540-9719

After that, an evaluation was carried out by manually checking the sentiment labels on each aspect of the data for November, December, and January. Table 5 shows the distribution of the prediction scores of each sentiment for each aspect of the IndoBERT fine-tuned model with aspect-level sentiment analysis. The results show that the model can predict the aspects of attraction, amenity, and price because these aspects are the same as those entered into the model training process. In the image, the prediction results produce more "none" labels. This result can be caused by the definition of the image aspect, which is an aspect that is assessed based on reviews about the image of tourists at the tourist attraction [3]. In the labeling process, reviews that meet these criteria are mostly labeled by the "none" label so that the prediction results of the image aspect produce a dominant "none" label.

5 Conclusion

In this research, the construction of a dataset for aspect-based sentiment analysis of tourist attraction reviews in DKI Jakarta has been carried out by scraping from Google Maps reviews. Then, the IndoBERT model for aspect and sentiment classification was fine-tuned using the adaptive aspect method, compared to the multi-label classification method as a baseline. As a result, the model with the adaptive aspect method has a better F1 score and accuracy than the multi-label classification method with F1-score and accuracy values of 0.71054 and 0.90394, respectively. For further development of this research, other transfer learning models that can handle multilingualism, such as the mBERT model, can be used, and studies on the use of generative models that are more powerful for zero-shot and few-shot tasks, as in the research of J.Liu et al. for the application of aspects as input as in the adaptive aspect method.

Reference

[1] I. G. A. D. Hendriyani, "Siaran Pers: Menparekraf Paparkan Penyerapan Pagu Anggaran Tahun 2022 di Hadapan Komisi X DPR RI," Kementerian Pariwisata dan Ekonomi Kreatif/Badan Pariwisata dan Ekonomi Kreatif. Accessed: Oct. 27, 2023. [Online]. Available: https://kemenparekraf.go.id/berita/siaran-pers-menparekraf-paparkan-penyerapan-pagu-anggaran-tahun-2022-di-hadapan-komisi-x-dpr-ri

[2] BAPPEDA Jakarta, *Rencana Pembangunan Jangka Menengah Daerah 2023-2026*. BAPPEDA Jakarta, 2022.

[3] D. Arianto and I. Budi, "Aspect-based Sentiment Analysis on Indonesia's Tourism Destinations Based on Google Maps User Code-Mixed Reviews (Study Case: Borobudur and Prambanan Temples)," *Proceedings of the 34th Pacific Asia Conference on Language, Information and Computation*, vol. 2019, no. May 2021, 2020.

[4] Munawir, M. D. Koerniawan, and B. J. Dewancker, "Visitor perceptions and effectiveness of place branding strategies in thematic parks in Bandung City using text mining based on google maps user reviews," *Sustainability (Switzerland)*, vol. 11, no. 7, 2019, doi: 10.3390/SU11072123.

[5] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, "SemEval-2014 Task 4: Aspect Based Sentiment Analysis," in *8th International Workshop on Semantic Evaluation, SemEval 2014 - co-located with the 25th International Conference on Computational Linguistics, COLING 2014, Proceedings*, 2014. doi: 10.3115/v1/s14-2004.

[6] S. Li and J. Ou, "Multi-Label Classification of Research Papers Using Multi-Label K-Nearest Neighbour Algorithm," in *Journal of Physics: Conference Series*, 2021. doi: 10.1088/1742-6596/1994/1/012031.

[7] L. Zhu, M. Xu, Y. Bao, Y. Xu, and X. Kong, "Deep learning for aspect-based sentiment analysis: a review," *PeerJ Comput Sci*, vol. 8, 2022, doi: 10.7717/PEERJ-CS.1044.

ISSN:2302-8149
e-ISSN:2540-9719

[8] I. L. Laily, I. Budi, A. B. Santoso, and P. K. Putra, "Mining Indonesia Tourism's Reviews to Evaluate the Services through Multilabel Classification and LDA," in *Proceedings of the International Conference on Electrical Engineering and Informatics*, 2020. doi: 10.1109/ICELTICs50595.2020.9315392.

[9] M. Afzaal, M. Usman, A. C. M. Fong, and S. Fong, "Multiaspect-based opinion classification model for tourist reviews," *Expert Syst*, vol. 36, no. 2, 2019, doi: 10.1111/exsy.12371.

[10] N. Liu and J. Zhao, "A BERT-Based Aspect-Level Sentiment Analysis Algorithm for Cross-Domain Text," *Comput Intell Neurosci*, vol. 2022, pp. 1–11, Jun. 2022, doi: 10.1155/2022/8726621.

[11] J. Liu, Z. Teng, L. Cui, H. Liu, and Y. Zhang, "Solving Aspect Category Sentiment Analysis as a Text Generation Task," in *EMNLP 2021 - 2021 Conference on Empirical Methods in Natural Language Processing, Proceedings*, 2021. doi: 10.18653/v1/2021.emnlp-main.361.

[12] B. Wilie *et al.*, "IndoNLU: Benchmark and resources for evaluating Indonesian natural language understanding," *arXiv preprint arXiv:2009.05387*, 2020.

[13] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, S. Saifullah, F. A. Dwiyanto, and R. Drezewski, "Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN," *ILKOM Jurnal Ilmiah*, vol. 14, no. 3, pp. 348–354, Dec. 2022, doi: 10.33096/ilkom.v14i3.1505.348-354.

[14] 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, 2023, doi: 10.22146/ijccs.77354.

[15] I. G. N. Suwena, I Ketut dan Widyatmaja, *Pengetahuan Dasar Ilmu Pariwisata. Denpasar*. 2017.

[16] K. Krippendorff and R. Craggs, "The Reliability of Multi-Valued Coding of Data," *Commun Methods Meas*, vol. 10, no. 4, pp. 181–198, Oct. 2016, doi: 10.1080/19312458.2016.1228863.

[17] L. Liu and M. T. Ozsu, *Encyclopedia of Database Systems*. Boston, MA: Springer US, 2009. doi: 10.1007/978-0-387-39940-9.

[18] D. Setiawan, N. Umar, and M. A. Nur, "Feature Extraction Optimization to Improve Naïve Bayes Accuracy in Sentiment Analysis of Bulukumba Tourism Objects," *Sistemasi: Jurnal Sistem Informasi*, vol. 13, no. 5, pp. 2209–2221, 2024.

[19] S. Prusty, S. Patnaik, and S. K. Dash, "SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer," *Frontiers in Nanotechnology*, vol. 4, 2022, doi: 10.3389/fnano.2022.972421.

[20] S. Pramana, B. Yuniarto, S. Mariyah, I. Santoso, and R. Nooraeni, *Data mining dengan R konsep serta implementasi*. 2018.