Comparison of Naïve Bayes and Logistic Regression in Sentiment Analysis on Marketplace Reviews Using Rating-Based Labeling

Satya Abdul Halim Bahtiar [1], Chandra Kusuma Dewa [2], Ahmad Luthfi [3]

[1,2,3]Magister Informatics Department, Faculty of Industrial Technology, Universitas Islam Indonesia, Sleman, Yogyakarta, Indonesia
Email: [1]satya.bahtiar@students.uii.ac.id, [2]chandra.kusuma@uii.ac.id, [3]ahmad.luthfi @uii.ac.id

Abstract

This research focuses on sentiment analysis in the marketplace reviews in Google Play Store, a platform for downloading Android applications and providing reviews. Sentiment analysis is essential for understanding user responses to applications, particularly in the app marketplace. In this study, two machine learning algorithms, Naïve Bayes and Logistic Regression, are employed to classify user reviews. The application rating is used as a reference to determine the sentiment of each comment. The dataset is divided into two conditions: using 2 labels (positive & negative) and 3 labels (positive, neutral, & negative). The test results indicate that the highest performance is achieved by classifying with Logistic Regression on the Shopee dataset with 2 labels. The accuracy reaches 84.58%, precision reaches 84.66%, and recall reaches 84.63%. Additionally, the fastest processing time occurs when testing the Lazada 2-label dataset with Naïve Bayes, taking only 0.038 seconds. Overall, the research suggests that datasets with 2 labels tend to yield higher accuracy compared to datasets with 3 labels.

Keywords: Naïve Bayes, Logistic Regression, Marketplace, Google Play Store Review, Rating-based Labeling

1. INTRODUCTION

The marketplace in Indonesia has experienced rapid growth in recent years. The increase in internet users and the adoption of mobile devices, along with changes in the shopping patterns of Indonesian society, have driven the dynamic development of marketplaces. Currently, many users in Indonesia rely on marketplace applications for online purchases, price comparisons, product searches, and providing reviews or feedback about the products they buy.

In the context of the Indonesian app market, Google Play Store is also one of the main platforms for downloading applications on Android devices. Indonesia is a potential market for mobile applications, with a continuously growing number of smartphone users. Users can provide reviews or feedback on the applications they use, whether it's a positive, negative, or neutral experience. Therefore, sentiment

915

analysis on Google Play Store reviews becomes crucial for application developers and business owners to understand user responses to their applications and take appropriate action.

Conversely, reviews for applications on Google Play utilize a five-point rating scale, which allows for a broad interpretation. Thus, it becomes essential to investigate the most suitable sentiment labeling approach for these reviews. The question arises whether the data should be categorized into three labels: positive, negative, and neutral, or simply into two labels: positive and negative. In sentiment analysis of the marketplace in Indonesia on the Google Play Store reviews, machine learning algorithms are used to classify user reviews into positive, negative, or neutral sentiments. Two machine learning algorithms utilized for this task are Naïve Bayes and Logistic Regression.

The Naïve Bayes classification method is a technique employed in sentiment analysis. This approach has theoretical advantages in terms of data consistency and classification computation. Naïve Bayes is commonly used in classification techniques, especially in social media platforms, employing variations such as Unigram Naïve Bayes, Multinomial Naïve Bayes, and Maximum Entropy Classification. A key feature of Naïve Bayes classification is its ability to generate strong hypotheses about various conditions or events. The calculation of group probabilities in Naïve Bayes relies on the Bayesian algorithm, utilizing specialized equations [1]. In the previous study, the use of Naïve Bayes successfully achieved a high accuracy for sentiment analysis on Shopee application reviews on Google Play. The accuracy was obtained by employing data partitioning technique on the Shopee application data from Google Play Store, consisting of 200 review data, with 100 positive reviews and 100 negative reviews. The researcher divided the data into two partitions. The first data partition, known as training data, comprised 140 data, while the second data partition, known as testing data, consisted of 60 data [2].

On the other hand, Logistic Regression is one of the classification algorithms in machine learning used to predict the probabilities of categorical dependent variables. This method is a general form of linear regression used to study the relationship between several variables with binary or probabilistic variables [3]. The classification method of logistic regression utilizes a logistic function to model the probabilities of different classes in the given data. It produces predictions in the form of probabilities and then uses a specific threshold to classify data into positive, negative, or neutral classes. Logistic Regression is also a commonly used classification method in sentiment analysis due to its ability to generate useful probability predictions for determining the sentiment of texts.

In another research, the authors have reviewed the application of logistic regression in addressing text categorization problems. They state that ridge-style

logistic regression performs comparably to the Support Vector Machine (SVM) algorithm. However, the advantage of using logistic regression lies in its ability to compute probability values instead of scores. They also introduce a new selection method that approximates the ridge solution through a sparse solution. This selection method first computes the ridge solution and then performs feature selection. The final results demonstrate that this method provides a well-balanced compromise between the ridge and LASSO solutions [4]. Naïve Bayes is a probability-based method based on Bayes' theorem, while Logistic Regression is a regression method that employs the logistic function to model the relationship between independent and dependent variables. The difference in these approaches allows for an interesting comparison of their performance in terms of accuracy, precision, recall, and processing time for sentiment classification and prediction [5].

The research gap discussed in this paper compares the performance of Naïve Bayes and Logistic Regression in the sentiment analysis of marketplace application reviews using Rating-Based Labeling. In comparison, several previous studies have explored the application of naive Bayes algorithms or logistic regression to tweet data under imbalanced class conditions and manual labeling conditions. This research provides insights regarding which model is better when applied to the marketplace review dataset with rating-based labeling conditions. This study also aims to determine which labeling conditions are better for the marketplace review dataset.

2. METHODS

The research method conducted by the author includes literature review, data collection, data labeling, data preprocessing, modeling, and evaluation to compare Naïve Bayes and Logistic Regression in sentiment analysis on the reviews of the marketplace on the Google Play Store. The research process flow can be seen in Figure 1.
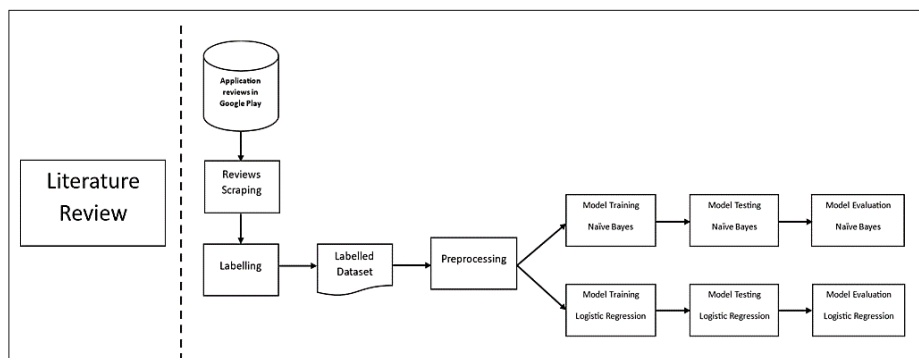

Figure 1. Research Method

2.1. Literature Review

The SLR method, which stands for Systematic Literature Review, is meticulously conducted with carefully structured steps and adheres to well-defined protocols, making it a robust and reliable approach to gather and analyze existing literature. The procedures performed in SLR are planning, conducting and reporting [6]. By following these systematic procedures, SLR minimizes the chances of introducing biases and subjective interpretations that might influence the outcome of the research. This method ensures that the review process remains objective and impartial, allowing for a comprehensive and unbiased synthesis of relevant information from various sources. As a result, the findings and conclusions drawn from an SLR are regarded with greater credibility and can provide valuable insights for further research and decision-making in the respective field of study. In conducting the literature review, it was found that both Naïve Bayes and Logistic Regression methods can be used for sentiment analysis. Whether used in sentiment classification in application reviews on Google Play Store[2], or in other cases such as sentiment analysis on tweets[7].

2.2. Reviews Scraping

During the data collection phase, user review data from marketplace applications on Google Play Store was gathered using the scraping method with a Python library called Google-Play-Scraper[8]. The data collected consists of reviews from 4 marketplace applications, namely Shopee, Tokopedia, Lazada, and Blibli.

The obtained data consisted of text with 10 variables, including 'reviewId', 'userName', 'userImage', 'content', 'score', 'thumbsUpCount', 'reviewCreatedVersion', 'at', 'replyContent', and 'repliedAt'. Among these variables, 'content' and 'score' were considered indicative of the sentiment level of an application, 'content' is user feedback about the apps and 'score' is the rating the user gives to the apps. The dataset comprised 9.000 data entries for each label in each marketplace.

Tabel 1. Results of data scraping marketplace reviews on Google Play Store

Marketplace	'content'	'content' in English
Tokopedia	bagus memudahkan, cepat dan banyak promo	good easy, fast and lots of promos
Shopee	Banyak biaya lain lain yg asal harganya 10k jadi 15k	Many other other costs that originally cost 10k to 15k
Lazada	Kenapa sekarang pengiriman tidak bisa cod	Why now shipping can not cod
Blibli	aplikasinya user frendly, bagus	the application is user frendly, good

2.3.　　Rating-Based Labeling

From the Google Play Store, where users can provide ratings and reviews for applications [9], the labeling process was conducted. The labels or categories represent pre-defined classes or categories that align with the sentiment classification purpose. In this research, rating-based labeling was performed with two conditions: the first one being 2 labels (positive & negative), where positive labels were assigned to comments with a rating of five, while the rest were categorized as negative. The second condition was 3 labels (positive, neutral, & negative), with ratings 5 and 4 considered positive, rating 3 as neutral, and ratings 2 and 1 as negative. Both conditions are applied to each marketplace dataset.

Tabel 2. Sample labelling dataset with 3 labels

'content'	'content' in English	'score'	label
bagus memudahkan, cepat dan banyak promo	good easy, fast and lots of promos	5	positive
lumayan bagus appnya	pretty good app	4	positive
berubah ubah harganya	Fluctuating its price	3	neutral
Tidak ada fitur cari produk pakai foto	There is no product search feature using photos	2	negative
Sejak gabung sama gojek jadi ga menarik lagi.	Since joining GoJek, it's no longer interesting.	1	negative

Tabel 3. Sample labelling dataset with 2 labels

'content'	'content' in English	'score'	label
Tokopedia NOW emang nolong bgt kalo mendadak malem2 perlu sesuatu. Mantap deh.	Tokopedia NOW is really helpful when suddenly you need something at night. That's great.	5	positive
e-commerce favourite dan utama kalo belanja, tapi sayangnya sekarang ada biaya layanan dan aplikasi	favorite and main e-commerce when shopping, but unfortunately now there are service and application fees	4	negative
Tolong dong gratis ongkirnya ngak konsisten, sebelum dimasukin keranjang nominalnya beda sama setelah check out	Please, the free shipping is not consistent, before it is put in the basket, the nominal is different from the one after checking out	3	negative
knp feed nya jd video smua?	why is the feed all video?	2	negative
Gagal koneksi terus	Connection failure continues	1	negative

2.4.　　　Preprocessing

Preprocessing on the 'content' of Google Play Store reviews is carried out to prepare the data for effective and efficient processing by machine learning algorithms [10]. The steps involved in this process include case folding, cleansing, tokenizing, removing stop words, and stemming [11].

Tabel 4. Text Preprocessing

case folding	The main purpose of converting text to lowercase is to ensure that words like "Bagus" and "bagus" are not considered different, as they are actually identical words. This is beneficial in reducing the number of words that need to be stored in the dictionary.
cleansing	This action is taken with the aim of removing punctuation or non-alphanumeric characters from the original text, including deleting URLs that start with HTTP, HTTPS, and the like.
tokenizing	In this process, sentences are broken down into text.
remove stop words	Removing common words that contribute little to revealing sentiment or meaning.
stemming	Process of reducing inflected (or sometimes derived) words to their word stem, base, or root form

2.4.1　　NLTK

NLTK is a Python library that provides a range of natural language processing algorithms. This open-source library is user-friendly, supported by a large community, and comes with comprehensive documentation. NLTK includes various algorithms, such as tokenization, part-of-speech tagging, stemming, and sentiment analysis [12]. The NLTK library and its documentation can be accessed at https://www.nltk.org/.

2.4.2　　Sastrawi

Sastrawi is a Python library that functions to convert various inflected words in the Indonesian language into their base form (stem). The use of Sastrawi stemming is employed to address the limitations of NLTK in performing stemming for Indonesian words. The Sastrawi library incorporates several core algorithms, including the Nazief and Adriani Algorithm, Enhanced Confix Stripping, and Modified Enhanced Confix Stripping. It serves as a stemmer library designed to handle word replacements with their base forms in the Indonesian language. Sastrawi implements algorithms based on Nazief and Adriani, which are then enhanced with CS (Confix Stripping) algorithm, ECS (Enhanced Confix Stripping) algorithm, and continuously improved with Modified ECS [13].

2.5. Modeling

In this phase, the models used for prediction or classification are Multinomial Naïve Bayes and Logistic Regression, and the entire data is tested with these two designed models. Creating the model requires training data to teach or train the machine learning model to recognize patterns and identify relationships between variables in the data. During the training phase, the machine learning model uses the training data to adjust parameters and rules needed to understand data characteristics and make accurate predictions. Once the model is trained using the training data, it is then tested for performance using testing data that has not been used in the model's training before. In this study, 80% of the dataset is used as training data, and the remaining 20% is used as testing data.

2.6. Evaluation

The final step in this research is to perform an evaluation measurement of the performance of the machine learning classification modeling conducted in the previous stages. The purpose of this evaluation measurement is to compare the performance and effectiveness of the two machine learning classification models used. Model evaluation is a crucial stage in the model creation process to ensure good performance on unseen data.

One of the techniques used to evaluate and summarize the performance of the machine learning classification models is the confusion matrix. The confusion matrix is a matrix that summarizes the total correct and incorrect classification results. By considering the true positive (TP), false positive (FP), false negative (FN), and true negative (TN) values, we can analyze the performance of the built classification model. The model evaluation is performed using the confusion matrix to determine accuracy, precision, recall [14], and also evaluates the processing time required for the data. Accuracy is an evaluation measure that shows how well a model can make correct predictions overall on all tested data. Accuracy is calculated using Equation 1.

$$ACC = \frac{TP+TN}{TP+TN+FN+FP} = \frac{TP+TN}{P+N} \tag{1}$$

Precision is an evaluation that measures how well the model can accurately identify positive data from the total predictions declared as positive. Precision is measured using Equation 2.

$$PREC = \frac{TP}{TP+FP} \tag{2}$$

Recall is an evaluation that assesses the model's ability to recognize or find positive data from all available positive data. This metric measures the percentage of positive data successfully identified or found by the model compared to the total actual positive data. Recall is calculated using Equation 3.

$$REC = \frac{TP+TN}{TP+FN} = \frac{TP}{P} \tag{3}$$

Meanwhile, the processing time measured in this research is the time it takes for the model to be trained using the training data until the model is tested using the training data and produces predictions.

3. RESULTS AND DISCUSSION

The data obtained from the results of scraping reviews on each marketplace application is then labeled with 2 rating-based labeling conditions. Subsequently, text preprocessing is performed, and the results are divided into 80% training data and 20% testing data. From this process, the following results are obtained. The results of the model testing using the 2-label marketplace dataset obtained the accuracy, precision, recall, and processing time as shown in Table 5.

Tabel 5. Results of Naïve Bayes testing with 2-label dataset

Dataset	Accuracy	Precission	Recall	Time
Shopee	**84.33%**	**84.32%**	**84.34%**	0.067s
Tokopedia	80.88%	81.13%	81.03%	0.067s
Lazada	77.44%	78.03%	77.34%	**0.038s**
Blibli	79.11%	79.35%	79.02%	0.042s

In table 5, the highest accuracy, precision, and recall values were obtained for the Shopee dataset with accuracy of 84.33%, precision of 84.32%, and recall of 84.34%. Although not significant, the highest score of 84.34% is obtained in recall, indicating that the model exhibits the highest capability to identify all objects in the target class. Meanwhile, the fastest processing time was found in the testing of the Lazada dataset, which took 0.038 seconds.

Tabel 6. Results of Logistic Regression testing with 2-label dataset

Dataset	Akurasi	Precission	Recall	Time
Shopee	**84.58%**	**84.66%**	**84.63%**	1.33s
Tokopedia	81.47%	81.75%	81.63%	1.27s
Lazada	77.86%	77.97%	77.81%	**0.47s**
Blibli	78.83%	78.82%	78.82%	0.65s

In Table 6, the highest accuracy, precision, and recall values were still obtained for the Shopee dataset, with slightly higher values: accuracy of 84.58%, precision of 84.66%, and recall of 84.63%. Highest precision score which means model has highest ability to correctly predicting target class. Although the processing time in this test was not faster than the previous test, the highest processing time was still found in the Lazada dataset, with a value of 0.47 seconds. Tables 7 and 8 present the testing results of the 3-label dataset.

**Tabel 7.** Results of Naïve Bayes testing with 3-label dataset

Dataset	Akurasi	Precission	Recall	Time
Shopee	66.22%	66.39%	66.31%	**0.055s**
Tokopedia	67.22%	66.74%	67.42%	0.057s
Lazada	69.51%	69.59%	69.59%	0.058s
Blibli	**70.75%**	**70.69%**	**70.87%**	0.065s

The 3-label dataset testing in Table 7 obtained lower accuracy, precision, and recall values compared to tables 5 and 6. The highest values were found in the Blibli dataset, with the highest achievable accuracy of 70.75%, precision of 70.69%, and recall of 70.87%. In this dataset, the highest ability of the model is in recall, which means it can find objects in the target class with a score of 70.87%. The fastest processing time was found in the Shopee dataset, with a value of 0.055 seconds, faster than overall time in Table 7.

Tabel 8. Results of Logistic Regression testing with 3-label dataset

Dataset	Akurasi	Precission	Recall	Time
Shopee	66.46%	65.95%	66.60%	3.60s
Tokopedia	69.20%	68.79%	69.38%	2.58s
Lazada	69.05%	69.05%	69.10%	**2.35s**
Blibli	**73.05%**	**72.89%**	**73.11%**	2.61s

In Table 8, the accuracy, precision, and recall values were higher than those in table 7, with accuracy of 73.05%, precision of 72.89%, and recall of 73.11%. These values were obtained in the testing of the Blibli dataset. The fastest processing time was found in the Lazada dataset, with a value of 2.35 seconds.
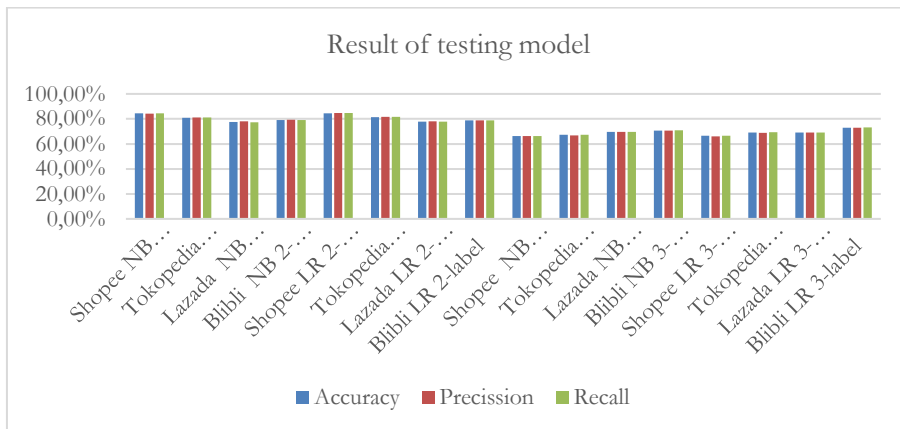
Figure 2. Result of testing model

The research findings indicate that the highest accuracy, precision, and recall scores were achieved when testing the Shopee dataset using logistic regression with 2 labels as shown in Figure 2. The obtained values were 84.58% accuracy, 84.66% precision, and 84.63% recall.

Tabel 9. Results of Processing Time

Dataset	Time
Shopee NB 2-label	0,06
Tokopedia NB 2-label	0,067
Lazada NB 2-label	0,038
Blibli NB 2-label	0,042
Shopee LR 2-label	1,33
Tokopedia LR 2-label	1,27
Lazada LR 2-label	0,47
Blibli LR 2-label	0,65
Shopee NB 3-label	0,055
Tokopedia NB 3-label	0,057
Lazada NB 3-label	0,058
Blibli NB 3-label	0,065
Shopee LR 3-label	3,6
Tokopedia LR 3-label	2,58
Lazada LR 3-label	2,35
Blibli LR 3-label	2,61

In the process time testing, it was found that the Naive Bayes algorithm has the ability to process faster than Logistic Regression. All the results of Naive Bayes algorithm testing were below 0.067 seconds, with the fastest time recorded at 0.038 seconds, while Logistic Regression only achieved the fastest processing time of 0.47 seconds, and the longest processing time was 3.6 seconds. Based on the testing on the marketplace dataset, both Naïve Bayes and Logistic Regression demonstrated comparable accuracy, precision, and recall rates ranging from 66% to 84%. Although both algorithms exhibited similar performance, Logistic Regression had a slight edge. Logistic regression achieved the highest results in accuracy with a value of 84.58%, precision of 84.66%, and recall of 84.63%.

Nevertheless, it is essential to emphasize that Naïve Bayes exhibited superior performance in terms of processing speed, achieving the shortest time during the evaluation of the Lazada dataset with two labels, registering a mere 0.038 seconds. All results in the processing time test, Naïve Bayes obtained values below 1 second, and the results obtained were faster than the processing time of logistic regression.

As for labeled datasets, the research findings indicate that datasets with 2 labels tend to yield higher accuracy compared to datasets with 3 labels. These results provide valuable insights for optimizing sentiment analysis in marketplace applications and suggest that selecting an appropriate algorithm can significantly impact both performance and processing efficiency. Looking at the results obtained, the combination of rating-based labeling with both Logistic Regression and Naive Bayes algorithms can achieve a relatively high level of accuracy, precision, and recall, similar to previous studies. In future research, other algorithms or optimizations in rating-based labeling could also be explored.

4.　CONCLUSION

Based on these findings, it is recommended to consider implementing Logistic Regression in sentiment analysis of marketplace application reviews, overall, the results of Logistic Regression experiments surpass those of Naive Bayes. while also taking into account the factor of more efficient data processing speed using Naïve Bayes. Moreover, adopting datasets with 2 labels facilitates better performance compared to datasets with 3 labels, ensuring a more effective categorization of sentiment in marketplace application reviews. These considerations can lead to improved user experience and better-informed business decisions in the competitive marketplace landscape.

REFERENCES

[1]　Pristiyono, M. Ritonga, M. A. Al Ihsan, A. Anjar, and F. H. Rambe, "Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes

Algorithm," *IOP Conf Ser Mater Sci Eng*, vol. 1088, no. 1, p. 12045, Feb. 2021, doi: 10.1088/1757-899X/1088/1/012045.

[2]   D. Pratmanto, R. Rousyati, F. F. Wati, A. E. Widodo, S. Suleman, and R. Wijianto, "App Review Sentiment Analysis Shopee Application In Google Play Store Using Naive Bayes Algorithm," *J Phys Conf Ser*, vol. 1641, no. 1, p. 12043, Nov. 2020, doi: 10.1088/1742-6596/1641/1/012043.

[3]   H. Hasanli and S. Rustamov, "Sentiment Analysis of Azerbaijani twits Using Logistic Regression, Naive Bayes and SVM," in *2019 IEEE 13th International Conference on Application of Information and Communication Technologies (AICT)*, Baku, Azerbaijan: IEEE, Oct. 2019, pp. 1–7. doi: 10.1109/AICT47866.2019.8981793.

[4]   K. Shah, H. Patel, D. Sanghvi, and M. Shah, "A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification," *Augmented Human Research*, vol. 5, no. 1, p. 12, Dec. 2020, doi: 10.1007/s41133-020-00032-0.

[5]   T. H. Jaya Hidayat, Y. Ruldeviyani, A. R. Aditama, G. R. Madya, A. W. Nugraha, and M. W. Adisaputra, "Sentiment analysis of twitter data related to Rinca Island development using Doc2Vec and SVM and logistic regression as classifier," *Procedia Comput Sci*, vol. 197, pp. 660–667, 2022, doi: 10.1016/j.procs.2021.12.187.

[6]   Y. A. Singgalen, "Pemilihan Metode dan Algoritma dalam Analisis Sentimen di Media Sosial: Sistematic Literature Review," *Journal of Information Systems and Informatics*, vol. 3, no. 2, 2021, doi: 10.33557/journalisi.v3i2.125

[7]   A. Poornima and K. S. Priya, "A Comparative Sentiment Analysis Of Sentence Embedding Using Machine Learning Techniques," in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India: IEEE, Mar. 2020, pp. 493–496. doi: 10.1109/ICACCS48705.2020.9074312.

[8]   Raksaka Indra Alhaqq, I Made Kurniawan Putra, and Yova Ruldeviyani, "Analisis Sentimen terhadap Penggunaan Aplikasi MySAPK BKN di Google Play Store," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 11, no. 2, pp. 105–113, May 2022, doi: 10.22146/jnteti.v11i2.3528.

[9]   H. Aldabbas, A. Bajahzar, M. Alruily, A. A. Qureshi, R. M. Amir Latif, and M. Farhan, "Google Play Content Scraping and Knowledge Engineering using Natural Language Processing Techniques with the Analysis of User Reviews," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 192–208, Jul. 2020, doi: 10.1515/jisys-2019-0197.

[10]   S. Pradha, M. N. Halgamuge, and N. Tran Quoc Vinh, "Effective Text Data Preprocessing Technique for Sentiment Analysis in Social Media Data," in *2019 11th International Conference on Knowledge and Systems Engineering (KSE)*, Da Nang, Vietnam: IEEE, Oct. 2019, pp. 1–8. doi: 10.1109/KSE.2019.8919368.

[11]   Y. A. Singgalen, "Sentiment Analysis on Customer Perception towards Products and Services of Restaurant in Labuan Bajo," *Journal of Information Systems and Informatics*, vol. 4, no. 3, 2022, doi: 10.51519/journalisi.v4i3.276

[12]   B. Khemani and A. Adgaonkar, "A Review on Reddit News Headlines with NLTK tool," *SSRN Electronic Journal*, 2021, doi: 10.2139/ssrn.3834240.

[13]   M. A. Rosid, A. S. Fitrani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, "Improving Text Preprocessing For Student Complaint Document Classification Using Sastrawi," *IOP Conf Ser Mater Sci Eng*, vol. 874, no. 1, p. 12017, Jun. 2020, doi: 10.1088/1757-899X/874/1/012017.

[14]   Ž. Đ. Vujovic, "Classification Model Evaluation Metrics," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, 2021, doi: 10.14569/IJACSA.2021.0120670.