IRPI PUBLISHER
Institut Riset dan Publikasi Indonesia (IRPI)
IJATIS: Indonesian Journal of Applied Technology and Innovation Science
Journal Homepage: https://journal.irpi.or.id/index.php/ijatis
Vol. 2 Iss. 1 February 2025, pp: 23-30
ISSN(P): 3032-7466 | ISSN(E): 3032-7474

Analyzing Customer Sentiment Towards Marketplace Reviews Using Classification Algorithms

Nabiilah[1], Siti Rohimah[2*], Septi Kenia Pita Loka[3]

[1,2,3]Department of Information System, Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

E-Mail: [1]12050322475@students.uin-suska.ac.id,
[2]12050323845@students.uin-suska.ac.id, [3]12050323230@students.uin-suska.ac.id

Received Nov 17th 2024; Revised Feb 14th 2025; Accepted Feb 19th 2025; Available Online Feb 28th 2025, Published Feb 28th 2025
Corresponding Author: Siti Rohimah

Abstract

Numerous online marketplaces like Shopee and Lazada have been developed in Indonesia due to the rapid growth of e-commerce. The Shopee and Lazada apps link buyers and sellers in transactions to purchase and sell products and services. About 100 million users have downloaded both applications as of this writing. Since releasing these programs, the community has voiced various thoughts and complaints. Based on this, user sentiment regarding the Shopee and Lazada applications on the Google Play Store is determined using sentiment analysis using the K-Nearest Neighbor (KNN), Nave Bayes, and Support Vector Machine (SVM) algorithms. Data selection, pre-processing, transformation, data mining, and assessment are the five stages of the Knowledge Discovery in Databases (KDD) approach. For each E-commerce application, 2000 reviews were used as the data. With an accuracy of 85.71% for Gaussian-NB modeling for the Lazada dataset and an accuracy of 85.67% for Bernoulli-NB modeling for the Shopee dataset, the Naive Bayes algorithm has the highest accuracy in experiments on each dataset.

Keyword: Classification, Customer, Lazada, Sentiment Analysis, Shopee

1. INTRODUCTION

The rapid development of E-commerce has resulted in the emergence of many marketplaces in Indonesia, such as Shopee and Lazada. Marketplace is a business model allowing merchants to sell their products online [1]. With the existence of e-commerce, it can further facilitate transactions between sellers and buyers; apart from e-commerce websites, it can now also be accessed via mobile phones to make it easier for users to access wherever and whenever [2]. Shopee and Lazada are the favorite e-commerce applications that are most in demand by Indonesians because they have many advantages, including facilitating the buying and selling process from a variety of stores, designed with an intuitive and easy-to-use interface, varied and guaranteed safe payment methods, and offering many discounts, promos and free shipping [3].

Shopee and Lazada are popular online stores; many are downloaded through the Google Play store. An app ranking in the Google Play Store is followed by user reviews [4]. User reviews also discuss their opinions about the two apps and usually make users consider the reviews before using them[5]. Techniques are needed to determine user reviews of the Shopee and Lazada apps due to the large number of unorganized reviews in the Google Play Store [6]. Therefore, user review data must be evaluated. Sentiment analysis methods can be used to process text data to extract information about the text [7]. By using classification, words are organized according to classes and opinions. Naïve Bayes is a pre-processing technology for feature classification that improves the text classification process's efficiency, scalability, and accuracy [8].

Sentiment analysis uses algorithms such as K-Nearest Neighbor (K-NN), Naive Bayes, and Support Vector Machine (SVM). SVM produces the most accurate sentiment analysis compared to other classification algorithms in some references. The study by Ilmawan and Mude in 2020 found that SVM had an accuracy of 81.46% compared to Naive Bayes of 75.41% when used for sentiment analysis of Indonesian reviews on the Google Play Store. [9]. In another study by Iskandar and Nataliani in 2021, which compared SVM, Naïve Bayes, and K-Nearest Neighbor for device touch analysis, it was found that SVM had the best accuracy with a value of 96.43% [10]. SVM can transform data into larger dimensions, such as kernel tricks, so that data can be better separated than other classification algorithms [11]. SVM has a solid theoretical foundation and can

DOI: https://doi.org/10.57152/ijatis.v2i1.1774

perform classification with a higher degree of accuracy than most other algorithms in various applications. Many studies have shown that SVM is the most accurate method for performing text classification [12]. A study that conducted sentiment analysis of Ruangguru app reviews by comparing 3 SVM algorithm kernels found that the linear kernel had an accuracy of up to 89.7% [13]. On the other hand, in another study, the RBF kernel had the best accuracy in sentiment analysis on airlines [14].

In previous studies, sentiment analysis was conducted on a single object; however, this study experiments with two objects, namely the marketplaces Lazada and Shopee, and employs three classification algorithms K-NN, Naïve Bayes Classifier (NBC), and SVM.

Based on the explanation above, this research will classify positive and negative sentiments on Shopee and Lazada e-commerce application reviews on the Google Play Store to determine user sentiment towards these applications.

2. MATERIAL AND METHOD

The research method consists of 5 stages of research that begin with collecting Shopee and Lazada comment data from the Google Play Store. The data used consists of 2000 comments for each Shopee and Lazada. Next, the data is preprocessed through stages of data scraping, stemming, lexicon-based weighting, and variable selection. Subsequently, three algorithms are implemented, namely K-NN, NBC, and SVM. The final stage involves visualizing the processed data. Figure 1 depicts the flowchart of the research method.

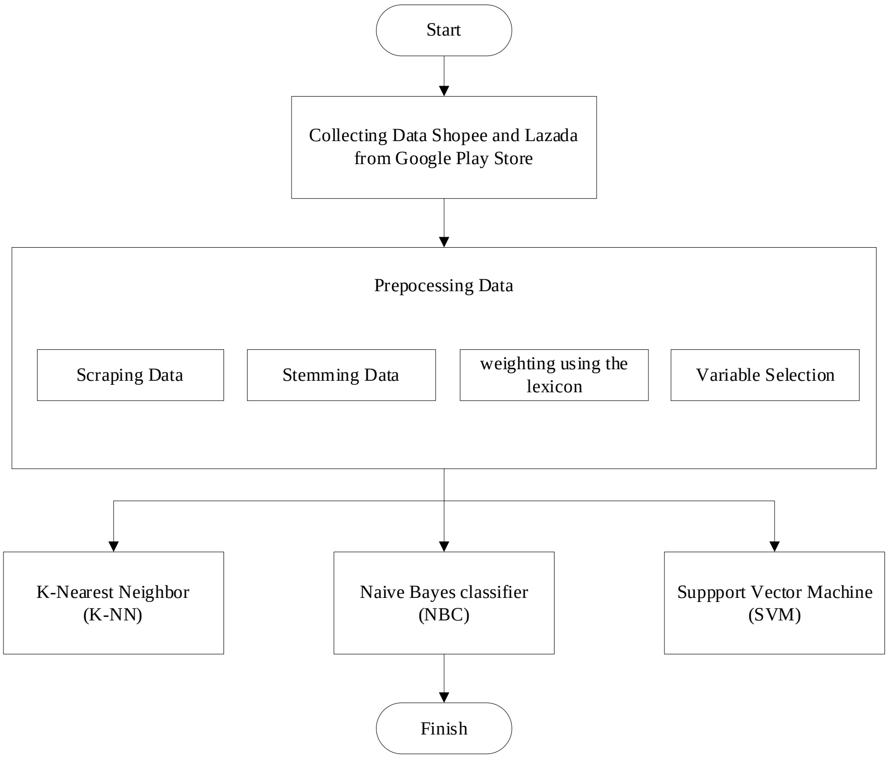

Start
Collecting Data Shopee and Lazada from Google Play Store
Prepocessing Data
Scraping Data
Stemming Data
weighting using the lexicon
Variable Selection
K-Nearest Neighbor (K-NN)
Naive Bayes classifier (NBC)
Suppport Vector Machine (SVM)
Finish

Figure 1. Research Methodology

2.1 Data Collecting

This research utilizes data from the Google Play Store regarding reviews of the Shopee and Lazada marketplaces. Data is taken by entering the application URL using Google Colab and then saved in excel to facilitate further analysis and structured storage. The data taken is the latest review from each application. The data that has been collected will be labeled using a lexicon-based dictionary.

2.2 Data Preprocessing

Data preprocessing is a series of steps taken to clean and prepare data before use in analysis and model training. Preprocessing data aims to produce calculation results. It will be more optimal and produce more accurate data. There are several stages of data preprocessing in this research:

1. Scraping Data
2. Stemming Data
3. Weighting Using Lexicon
4. Variable Selection

2.3 Sentiment Analysis

Sentiment analysis helps analyze the tendency of opinions toward an object or problem [12]. In unstructured data, sentiment analysis is instrumental in finding valuable information from the text and determining the negative, positive, or neutral emotional tone [15].

2.4 Support Vector Machine (SVM)

Support Vector Machine is a learning technique developed based on statistical learning theory. The SVM algorithm builds a model of assigning new examples into one category or another and makes a non-probabilistic binary linear classifier [16]. In applying the SVM algorithm, equation (1) is used [17].

$$x = \frac{1}{2}||w||^2 = \frac{1}{2}({w_1}^2 + {w_2}^2) \tag{1}$$

The vector w is the weight that determines the direction and orientation of the hyperplane, while b defines the margin, which is the shortest distance between the data points of each class and the hyperplane. In this equation a_i represents the Lagrange coefficient, y_i is the class label of the i-th sample (with values of +1 or -1 in binary classification cases), and x_i is the feature vector of the i-th sample.

This equation is utilized in the kernel trick method, which enables SVM to handle data that is not linearly separable by mapping it into a higher-dimensional space.

2.3 Naïve Bayes Classifier

Based on Bayes' theorem, Naïve Bayes Classifier is a statistical classification algorithm. It provides high predictability and achieves results comparable to other algorithms. Its classification techniques include decision tree and neural network construction [18]. Equation (2) is used to apply the NBC algorithm. [19].

$$P(i) = \frac{Nc}{N} \tag{2}$$

In this formula, P(i) represents the probability of the occurrence of category or class i, Nc denotes the number of samples in category i, and N is the total number of samples in the dataset. This formula is commonly used in statistics and machine learning, particularly in probability-based algorithms such as Naïve Bayes, to calculate the likelihood of a class in data classification. Then, calculate the probability of a probability using equation (3).

$$P\ (w \parallel c) = \frac{\text{count }(w,c)+1}{\text{count }(c) + |v|} \tag{3}$$

It is used to calculate the conditional probability in the context of machine learning models, particularly in Naïve Bayes for text analysis or word-based classification. In this formula, P (w||c) represents the probability of the word w occurring in class c, count (w, c) is the number of occurrences of the word w in documents belonging to class c, count (c) is the total number of words in documents belonging to class c , and |v| is the size of the vocabulary, which is the total number of unique words present in the entire dataset.

2.4 K-Nearest Neighbor (KNN)

The KNN algorithm uses neighbor classification as the predictive value of a new query instance. This algorithm classifies objects based on learning data that is the closest distance to the object. The presence or absence of irrelevant features strongly influences the accuracy of the KNN algorithm. [20]. In the application of the KNN algorithm, equation (4) is used [21].

$$d_i = \sqrt{\sum_1^p (x_{2i} - x_{1i})^2} \tag{4}$$

In this formula, d_i represents the distance between two data points, where one point is x_{1i} (the feature coordinates of the first point) and the other point is x_{2i} (the feature coordinates of the second point). This formula measures how far apart two points are in a p-dimensional space based on the differences in the values of their respective features.

3. RESULTS AND DISCUSSION

3.1 Data Collection

This research data was taken from the Google Play Store with the latest review data on the Shopee and Lazada applications. Amount the data used was 2000 reviews for Shopee and 2000 reviews also for Lazada. The results of data collection can be seen in Table 1 and Table 2.

Table 1. Data Collection Result for Lazada

No.	Content	Score
1	Mantap ahhh sikattt belanja online tapi... Tapi kasih kesbek atth kaya potongan belanja pocher gk selalu potongan Ongki biar makin bnyk yg belanja di online shop na🤣	5
2	Sekarang banyak yg mahal ongkirnya	3
3	bagus pelayanannya tapi gila anjirr ongkirnya	4
…	...	...
1999	Bercuma ngasih limit paylater kalau gak bisa di pake mah buruk sekali padahal selalu bayar tepat waktu sebelumnyh.	1
2000	Banyak iklan kalo main game	1

Table 2. Data Collecting Result for Shopee

No.	Content	Score
1	sangat berkesan dengan update terbaru, semakin nyaman untuk memakai shopee seterusnya	5
2	Mantap	5
3	Barang sangat bagus memuaskan	5
...	...	...
1999	sangat baik	5
2000	Top markotop	5

3.2 Data Pre-Processing

Before being used for classification data, the dataset will be processed at the Pre-processing stage, which is defined in Table 3 and Table 4.

Table 3. Data Preprocessing Result for Shopee

Process	Content
Data	Mantap ahhh sikattt belanja online tapi... Tapi kasih kesbek atth kaya potongan belanja pocher gk selalu potongan Ongki biar makin bnyk yg belanja di online shop na🤣
Cleaning	mantap belanja online tapi tapi kasih kesbek kaya potongan belanja pocher gk selalu potongan ongki biar makin bnyk yg belanja di online shop
Tokenizing	[‘mantap’, ‘belanja’, ‘online’, ‘tapi’, ‘tapi’, ‘kasih’, ‘kesbek’, ‘kaya’, ‘potongan’, ‘belanja’, ‘pocher’, ‘gk’, ‘selalu’, ‘potongan’, ‘ongki’, ‘biar’, ‘makin’, ‘bnyk’, ‘yg’, ‘belanja’, ‘di’, ‘online’, ‘shop’]
Stopwords	[‘mantap’, ‘belanja’, ‘online’, ‘kasih’, ‘kesbek’, ‘kaya’, ‘potongan’, ‘belanja’, ‘pocher’, ‘selalu’, ‘potongan’, ‘ongki’, ‘makin’, ‘bnyk’, ‘belanja’, ‘online’, ‘shop’]
Stemming	[‘mantap’, ‘belanja’, ‘online’, ‘kasih’, ‘kesbek’, ‘kaya’, ‘potong’, ‘belanja’, ‘pocher’, ‘selalu’, ‘potong’, ‘ongki’, ‘makin’, ‘bnyk’, ‘belanja’, ‘online’, ‘shop’]

Table 4. Data Preprocessing Result for Lazada

Process	Content
Data	sangat berkesan dengan update terbaru, semakin nyaman untuk memakai shopee seterusnya
Cleaning	sangat berkesan dengan update terbaru semakin nyaman untuk memakai shopee seterusnya
Tokenizing	[‘sangat’, ‘berkesan’, ‘dengan’, ‘update’, ‘terbaru’, ‘semakin’, ‘nyaman’, ‘untuk’, ‘memakai’, ‘shopee’, ‘seterusnya’]
Stopwords	[‘sangat’, ‘berkesan’, ‘dengan’, ‘update’, ‘terbaru’, ‘semakin’, ‘nyaman’, ‘memakai’, ‘shopee’, ‘seterusnya’]
Stemming	[‘sangat’, ‘kesan’, ‘dengan’, ‘update’, ‘baru’, ‘semakin’, ‘nyaman’, ‘pakai’, ‘shopee’, ‘terus’]

Furthermore, to perform labeling, the data is feature selected to only use the content column before pre-processing. Sentiment labeling uses the Lexicon Based method to determine the score of the reviews from the application, with a positive sentiment if the score has a value >0, a negative sentiment with a score <0, and a neutral sentiment with a score of 0 as in the following table.

3.3 Labelling Data

Data labeling using the lexicon-based approach is done by comparing words in the text with a predefined sentiment lexicon. Each word in the text is assigned a sentiment score based on the lexicon, and then the overall text score is calculated to determine the sentiment label, such as positive, negative, or neutral. The result of labelling data defined in Table 5 and Table 6.

Table 5. Shopee Review Sentiment Result

Score Sentiment	Sentiment	Content
2	Positive	terima bank transaksi emas shopee tolong upgrade bank bsinya ngk bank mandiri bni konvensional
1	Positive	ganti nomor telepon aja susah banget ribet
2	Positive	limit paylatter spinjam udh gak telat yaudah sgitu gitu aja
-6	Negative	emg hp sy yg butut apk nya lg gmna ya susah bgt ya chekout ga buka lemoot bgt sekali buka keranjang aja lg gmna sih berat bgt apk nya prasaan udh d fresin dlu dah klo rada lot d fresh lg tp ttp aja lot ga neh tolong bnrn yak bintang aja dl nnti klo dah bagus d upgrade

Table 6. Lazada Review Sentiment Result

Score Sentiment	Sentiment	Content
2	Positive	sy suka jarang dpt gratis ongkir ongkirnya mahal mahal
7	Positive	layan nya muas
-2	Negative	bagus aplikasi layan nya
4	Positive	moga amanah jujur yg utama jual percaya ragu tipu

After labeling, the word cloud function is used to find 100 words that often appear or visualize data in negative, positive, and neutral labels.

3.4 Term Frequency-Inverse Document Frequency (TF-IDF)

To further refine the sentiment analysis process, the Term Frequency-Inverse Document Frequency (TF-IDF) technique was employed. TF-IDF is a numerical statistic intended to reflect how important a word is to a document in a collection or corpus. This method helps to highlight the words that are most distinctive and relevant within the Shopee and Lazada reviews, while diminishing the weight of commonly used words that may not contribute significantly to sentiment classification. The following Tables 7 illustrate the TF-IDF scores for a selection of terms extracted from the pre-processed review data. These scores provide insight into the relative importance of each term in distinguishing between positive, negative, and neutral sentiments.

Table 7. TF-IDF from Lazada Dataset

	acara	aduh	…	aja	barang	event	ganggu	promo	zonk	zoom
0	0	0	…	0	0	0	0,143489	0	0	0
1	0	0	…	0	0	0	0	0	0	0
2	0	0	…	0,09284	0	0,205529	0	0,149703	0	0
3	0	0	…	0,09578	0	0	0	0	0	0
4	0	0	…	0,109596	0	0	0	0	0	0
…	…	…	…	…	…	…	…	…	…	…
1999	0	0	…	0	0	0	0	0	0	0
2000	0	0	…	0	0	0	0	0	0	0

The use of the Term Frequency-Inverse Document Frequency (TF-IDF) technique in the sentiment analysis process represents a key strategy to improve the accuracy and efficiency of sentiment classification. TF-IDF is a statistical method that aims to evaluate the importance of a word within a document relative to a collection or corpus of documents. By focusing on words that appear frequently within a specific document but are rare across the entire corpus, TF-IDF helps to emphasize the most relevant terms that contribute meaningfully to the classification task.

In the context of the sentiment analysis on Shopee and Lazada reviews, applying TF-IDF allows for a more nuanced understanding of the reviews by identifying words that are distinctive and critical to the sentiment conveyed by the reviewers. For example, words such as "excellent," "fast," or "disappointed" might carry a stronger weight in determining sentiment when they appear in a review, while common words like "the," "is," or "and" are likely given less importance due to their frequent occurrence across the entire corpus.

By diminishing the weight of commonly used words (also known as stop words), TF-IDF ensures that the sentiment analysis process focuses on the terms that provide unique insights into the user's experience. This approach is essential in accurately distinguishing between positive, negative, and neutral sentiments in the reviews. Words that are frequently used across both positive and negative reviews, but don’t contribute specific sentiment, are down-weighted, allowing the model to focus on words that offer more substantial cues about the overall tone of the review.

The TF-IDF scores presented in Table 7 illustrate the relative importance of specific terms extracted from the pre-processed review data. These scores help to highlight which terms play a pivotal role in shaping sentiment, offering valuable insights into how the words influence sentiment classification. For instance, terms with higher TF-IDF scores are likely more distinctive and carry greater significance in differentiating positive from negative sentiments, aiding the model in its classification tasks.

Overall, employing TF-IDF enhances the precision of sentiment analysis by ensuring that the model is focused on the most relevant words in the reviews, improving the model’s ability to effectively classify sentiment and offering a more robust analysis of user feedback.

3.5 Classification Model Building

After this, research can be concluded that there are four classification algorithms being compared and the result for each algorithm is SVM (65% for Shopee and 63% for Lazada), KNN (78% for Shopee and 76% for Lazada), Gaussian Naïve Bayes (80% for Shopee and 79% for Lazada), and Multinomial Naïve Bayes (72% for Shopee and 70% for Lazada). Based on the results, the Gaussian Naïve Bayes algorithm shows the highest accuracy for both platforms, followed by KNN, which performs quite well with a slight accuracy difference. Meanwhile, SVM has lower accuracy compared to Gaussian Naïve Bayes and KNN but remains higher than Multinomial Naïve Bayes, which has the lowest accuracy among the four algorithms. Overall, the classification accuracy for Shopee and Lazada exhibits a similar pattern, with minor variations between the algorithms used. This indicates that Gaussian Naïve Bayes and KNN are the best choices for data classification on both platforms. Comparison of Classification Algorithm Accuracy Diagram can view figure 2.

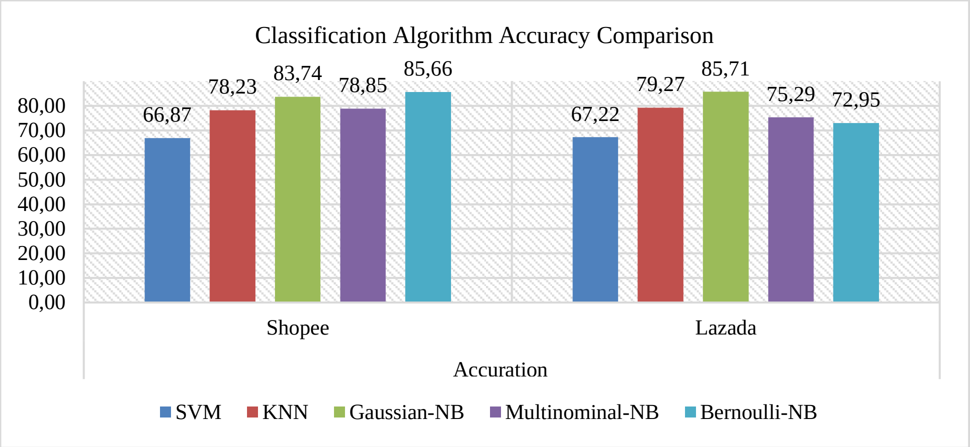
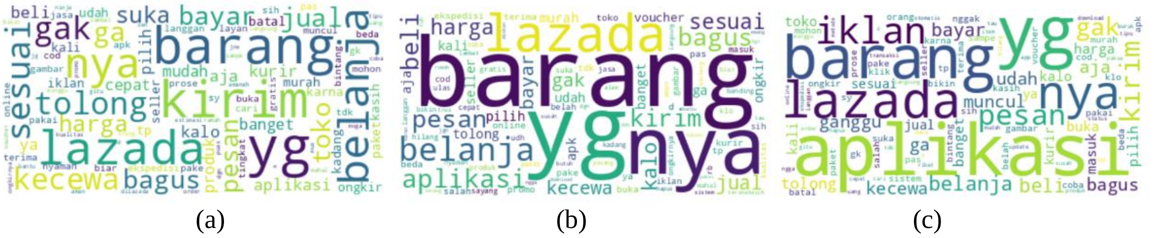


Figure 2. Classification Algorithm Accuracy Diagram

3.6 Data Visualization

Data visualization using a word cloud aims to represent the dataset graphically, making it easier to identify the most frequently appearing data in the document. The process of creating this word cloud is carried out using the matplotlib library in the Python programming language. The result of the word cloud visualization can be seen in Figure 3(a), Figure 3(b), and Figure 3(c).



(a) (b) (c)

Figure 3. Data Visualization Sentiment on Lazada (a) Positive, (b) Neutral, (c) Negative

4. DISCUSSION

The study aimed to classify sentiments from Google Play Store reviews of the Shopee and Lazada e-commerce applications using K-Nearest Neighbor (KNN), Naïve Bayes Classifier (NBC), and Support Vector Machine (SVM) algorithms. The Naïve Bayes algorithm, specifically Gaussian-NB for the Lazada dataset (85.71% accuracy) and Bernoulli-NB for the Shopee dataset (85.67% accuracy), demonstrated the highest accuracy [1]. These findings align with the general understanding of Naïve Bayes as an effective method for text classification due to its efficiency, scalability, and accuracy.

The results showed that while SVM can transform data into larger dimensions and has a solid theoretical foundation, it showed lower accuracy compared to Gaussian Naive Bayes and KNN. This contrasts with other studies that found SVM to be the most accurate method for text classification. This difference may be attributed to the specific characteristics of the datasets used in this study, such as the prevalence of mixed languages, slang words, and local languages, which posed challenges for sentiment analysis with lexicon-based approaches.

The challenges encountered due to the nature of user reviews suggest a need for further development of sentiment analysis methods. Future research could explore techniques to better capture the actual meaning of comments, potentially through advanced natural language processing techniques or the incorporation of contextual information. Additionally, expanding the study to include more e-commerce platforms or incorporating user demographic data could provide a more comprehensive understanding of customer sentiment in the e-commerce landscape.

5. CONCLUSION

The Naïve Bayes algorithm has the highest accuracy in experiments on each dataset with an accuracy of 85.71% on Gaussian-NB modeling for the Lazada dataset and 85.67% on Bernoulli-NB modeling for the Shopee dataset. Analysis of sentiment from app reviews primarily uses mixed language, slang words, and local languages so that some comments have sentiment scores that are different from the actual meaning of the comments, which makes sentiment analysis with lexicon-based more challenging. So the need to develop the method used so that sentiment assessment is by the actual meaning of comments can improve the accuracy and accuracy of sentiment analysis results

REFERENCES

[1] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning," IEEE Access, vol. 8, pp. 23522–23530, 2020, doi: 10.1109/ACCESS.2020.2969854.

[2] F. Xu, Z. Pan, and R. Xia, "E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework," Inf. Process. Manag., vol. 57, no. 5, p. 102221, 2020, doi: 10.1016/j.ipm.2020.102221.

[3] X. Lin, "Sentiment Analysis of E-commerce Customer Reviews Based on Natural Language Processing," ACM Int. Conf. Proceeding Ser., pp. 32–36, 2020, doi: 10.1145/3436286.3436293.

[4] S. W. Iriananda, R. P. Putra, and K. S. Nugroho, "Analisis Sentimen Dan Analisis Data Eksploratif Ulasan Aplikasi Marketplace Google Playstore," 4th Conf. Innov. Appl. Sci. Technol. (CIASTECH 2021), no. Ciastech, pp. 473–482, 2021.

[5] S. A. Aaputra, Didi Rosiyadi, Windu Gata, and Syepry Maulana Husain, "Sentiment Analysis Analysis of E-Wallet Sentiments on Google Play Using the Naive Bayes Algorithm Based on Particle Swarm Optimization," J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 3, no. 3, pp. 377–382, 2019, doi: 10.29207/resti.v3i3.1118.

[6] S. Fransiska and A. Irham Gufroni, "Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method," Sci. J. Informatics, vol. 7, no. 2, pp. 2407–7658, 2020, [Online]. Available: http://journal.unnes.ac.id/nju/index.php/sji

[7] M. I. Ahmadi, F. Apriani, M. Kurniasari, S. Handayani, and D. Gustian, "Sentiment Analysis Online Shop on the Play Store Using Method Support Vector Machine (Svm," Semin. Nas. …, vol. 2020, no.

Semnasif, pp. 196–203, 2020, [Online]. Available: http://jurnal.upnyk.ac.id/index.php/semnasif/article/view/4101

[8] E. H. Muktafin, K. Kusrini, and E. T. Luthfi, “Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing,” J. Eksplora Inform., vol. 10, no. 1, pp. 32–42, 2020, doi: 10.30864/eksplora.v10i1.390.

[9] L. B. Ilmawan and M. A. Mude, “Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store,” Ilk. J. Ilm., vol. 12, no. 2, pp. 154–161, 2020, doi: 10.33096/ilkom.v12i2.597.154-161.

[10] J. W. Iskandar and Y. Nataliani, “Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek,” J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 5, no. 6, pp. 1120–1126, 2021, doi: 10.29207/resti.v5i6.3588.

[11] R. Mukarramah, D. Atmajaya, and L. B. Ilmawan, “Performance comparison of support vector machine (SVM) with linear kernel and polynomial kernel for multiclass sentiment analysis on twitter,” Ilk. J. Ilm., vol. 13, no. 2, pp. 168–174, 2021, doi: 10.33096/ilkom.v13i2.851.168-174.

[12] H. P. P. Zuriel and A. Fahrurozi, “Implementasi Algoritma Klasifikasi Support Vector Machine Untuk Analisa Sentimen Pengguna Twitter Terhadap Kebijakan Psbb,” J. Ilm. Inform. Komput., vol. 26, no. 2, pp. 149–162, 2021, doi: 10.35760/ik.2021.v26i2.4289.

[13] F. F. Irfani, “Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine,” JBMI (Jurnal Bisnis, Manajemen, dan Inform., vol. 16, no. 3, pp. 258–266, 2020, doi: 10.26487/jbmi.v16i3.8607.

[14] H. C. Husada and A. S. Paramita, “Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM),” Teknika, vol. 10, no. 1, pp. 18–26, 2021, doi: 10.34148/teknika.v10i1.311.

[15] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, “Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi,” J. Teknoinfo, vol. 14, no. 2, p. 115, 2020, doi: 10.33365/jti.v14i2.679.

[16] F. Luo, “Affective-feature-based Sentiment Analysis using SVM Classifier”.

[17] J. Jabbar, I. Urooj, W. Junsheng, and N. Azeem, “Real-time sentiment analysis on E-Commerce application,” Proc. 2019 IEEE 16th Int. Conf. Networking, Sens. Control. ICNSC 2019, pp. 391–396, 2019, doi: 10.1109/ICNSC.2019.8743331.

[18] S. Rana, “Comparative Analysis of Sentiment Orientation Using SVM and Naïve Bayes Techniques,” no. October, pp. 106–111, 2016.

[19] F. S. Fitri, M. N. S. Si, and C. Setianingsih, “Sentiment analysis on the level of customer satisfaction to data cellular services using the naive bayes classifier algorithm,” Proc. - 2018 IEEE Int. Conf. Internet Things Intell. Syst. IOTAIS 2018, pp. 201–206, 2019, doi: 10.1109/IOTAIS.2018.8600870.

[20] M. T. Akter, M. Begum, and R. Mustafa, “Bengali Sentiment Analysis of E-commerce Product Reviews using K-Nearest Neighbors,” 2021 Int. Conf. Inf. Commun. Technol. Sustain. Dev. ICICT4SD 2021 - Proc., pp. 440–444, 2021, doi: 10.1109/ICICT4SD50815.2021.9396910.

[21] F. Firmansyah et al., “Comparing Sentiment Analysis of Indonesian Presidential Election 2019 with Support Vector Machine and K-Nearest Neighbor Algorithm,” 6th Int. Conf. Comput. Eng. Des. ICCED 2020, pp. 5–10, 2020, doi: 10.1109/ICCED51276.2020.9415767.