Variance
Journal of Statistics and Its Applications
P-ISSN: 2685-8738 E-ISSN: 2685-872X

Variance: Journal of Statistics and Its Applications
Volume 7 Issue 2 | October 2025
10.30598/variancevol7iss2page135-146

COMPARISON OF SENTIMENT ANALYSIS USING NAÏVE BAYES CLASSIFICATION METHOD AND LEXICON BASED ON JIWA+ BY JANJI JIWA APPLICATION REVIEWS

Reyana Hilda Arti[1], Neva Satyahadewi[2*], Wirda Andani[3]
[1,2,3] Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Tanjungpura
Jl. Prof. Dr. H. Hadari Nawawi, Pontianak, 78124, West Kalimantan, Indonesia

Corresponding Author's Email: *neva.satya@math.untan.ac.id*

Abstract: *The coffee beverage industry in Indonesia is experiencing significant growth, intensifying competition among businesses striving to maintain quality for customer loyalty. E-commerce applications play a vital role in preserving business standards as they directly engage with consumers. Janji Jiwa is among the coffee brands leveraging an application named Jiwa+ in their operations. Analyzing reviews on this e-commerce platform provides valuable insights for business owners and app developers. In this study, sentiment analysis was conducted by classifying reviews into positive, neutral, and negative sentiments using two methods: Lexicon Based and Naïve Bayes. The Lexicon Based method uses a predefined dictionary as the basis for labeling, while Naïve Bayes relies on training data to provide new insights into how both methods handle this type of data. A total of 597 Jiwa+ application reviews from the Google Play Store were utilized, split into 90% training and 10% testing data sets. The study results indicate that Naïve Bayes produces a better model than the Lexicon-Based method, as shown by its higher accuracy, sensitivity, and specificity. This is because Lexicon-Based relies on labeling words from a dictionary, which may not cover all words in the reviews, leading to labeling errors and misclassification.*

Keywords: *application reviews, likelihood probability, Janji Jiwa, posterior probability, prior probability*

1. INTRODUCTION

Currently, the culinary business in Indonesia is experiencing rapid growth. It includes not only food products but also beverages. One example of a booming beverage business is the coffee industry. The proliferation of coffee shops has naturally intensified competition among them. Each coffee shop strives to provide the best service to customers, ensuring their satisfaction and loyalty towards the products. This loyalty will impact their sales and business sustainability. Customer satisfaction with a product or service will lead them to return, thereby promoting the continuity of the business itself [1].

One of the efforts that coffee shop owners can undertake to enhance their service and sales is by leveraging internet technology advancements, such as applications to facilitate user transactions. One such coffee shop that utilizes an application for its services is Janji Jiwa, with an application called Jiwa+ by Kopi Janji Jiwa. As mentioned earlier, maintaining customer satisfaction is crucial for business owners to gain customer loyalty. One way to achieve this is by paying attention to customer reviews of applications like Jiwa+ by Kopi Janji Jiwa. These reviews can be analyzed using sentiment analysis. Sentiment analysis aims to determine public or community perspectives on a particular subject or topic [2]. Sentiment analysis is performed by classifying public views into several categories, such as positive, negative, or neutral reviews of applications like Jiwa+ by Janji Jiwa. Several methods can be used to analyze sentiment, including Naïve Bayes and Lexicon Based methods.

Naïve Bayes is one of the commonly used methods in sentiment analysis due to its simplicity and speed in processing. Naïve Bayes can produce high accuracy [3]. However, it has the drawback of requiring initial labeling, which affects the accuracy of the model. In addition to using machine learning methods, sentiment review classification can be conducted using Lexicon Based methods. This method has the advantage over Naïve Bayes in that it does not require initial label information, enabling direct classification without the need for preparing initial review labels, which typically takes longer. However, this method also has its drawback, as it relies on

 Submitted: 2024-07-03 *Accepted: 2025-11-03*

sentiment classification decisions based on positive and negative word scores derived from dictionaries, making the completeness of the dictionary used a critical factor influencing the classification outcomes.

Previous studies analyzed tweets regarding the 2019 election by comparing Naïve Bayes and Lexicon-Based methods [4]. These studies typically used two labels, positive and negative, and found that the Naïve Bayes model outperformed the Lexicon-Based method. While previous studies employed only two sentiment labels, this study extends the analysis to three labels—positive, negative, and neutral—to capture more nuanced user opinions. In this study, both methods are applied to reviews of the Jiwa+ application by Janji Jiwa to compare the results obtained from each method.

With the respective strengths and weaknesses of both methods, this study aims to compare the results of sentiment analysis of Jiwa+ using the Kopi Janji Jiwa application reviews obtained from the Google Play Store, employing both Naïve Bayes and Lexicon Based methods.

2. METHODOLOGY

2.1. Data Description

In this study, 597 reviews of the Jiwa+ application by Janji Jiwa were collected from the Google Play Store using a web scraping method. The data collected consisted of user reviews of the application from January 24, 2023, to January 29, 2024. The reviews then underwent preprocessing, resulting in 545 remaining review data points. These reviews were classified into positive, neutral, and negative sentiment reviews using Naïve Bayes and Lexicon Based classification methods. Subsequently, the data were divided into training and testing datasets in a 90:10 ratio.

2.2. Research Method

The research commenced with the retrieval of Jiwa+ application reviews from the Google Play Store. Data retrieval was performed using a web scraping method, which employed the Python programming language and Jupyter Notebook software. Web scraping is defined as the process of gathering specific data or information from a system or document [5].

After acquiring the review data, the next step involves text preprocessing. Text preprocessing serves as the initial stage in text classification, aiming to transform text data into a more refined format to facilitate subsequent processing and ensure the quality of the resulting information [6]. Raw data obtained from the web scraping process cannot be directly utilized due to susceptibility to disruptive factors that may influence analysis outcomes, such as missing data, noise, or format inconsistencies, potentially caused by the data collection method [7]. The text preprocessing stages conducted are as follows:

1) Case folding: This step is used to convert the entire text content into lowercase letters.
2) Cleansing data: This stage involves cleaning the text from noise or disturbances such as symbols, characters, words, and punctuation marks that are not needed in the analysis.
3) Spelling normalization: This function is used to correct misspelled words, thus facilitating the analysis process.
4) Tokenizing: It is the process of breaking down a document into words.
5) Stopword removal: It is the process of removing unimportant words in a sentence, such as “the”, “and”, “from”, etc.

After undergoing the text preprocessing process, the data is visualized using word clouds and bar charts to display the most frequently occurring words in the reviews. The next step is sentiment classification of the reviews using the Naïve Bayes and Lexicon Based methods.

Naïve Bayes is a method based on Bayes’ theorem. The main characteristic of the Naïve Bayes method is the “naive” nature of the initial concept of Bayes’ theorem, which assumes independence of each event or condition, an assumption that is sometimes unrealistic in natural language processing, hence the name “Naïve

Bayes' [8]. In Bayes' theorem, analysis is conducted by combining two sources of information: prior probability and likelihood probability, which then produce posterior probability [9].

In sentiment classification using the Naïve Bayes method, the data will first be manually labeled with sentiment. This labeling is based on the researcher's subjectivity, where if the content of the review contains praise, advantages, or approval of Jiwa+ by the Janji Jiwa application, the review will be labeled as having a positive sentiment. Conversely, if not, it will be labeled as negative sentiment. If the review contains both the advantages and disadvantages of the application simultaneously and/or only contains suggestions for the Jiwa+ by Janji Jiwa application in the future, the review will be labeled as neutral. After labeling, a Naïve Bayes model will be formed from the training data. This model will then classify the sentiment of the test data.

The classification process in Naïve Bayes is performed by calculating the probability of a word belonging to a particular class (Posterior Probability). This posterior probability is obtained by multiplying the probability of a class (Prior Probability) by the probability of a word feature occurring in a particular class (Likelihood Probability), then dividing it by the probability of all features occurring (Evidence Probability), which is always a constant value [10].

1) Posterior probability can be formulated as follows [11]:

$$\boldsymbol{MAP = arg\,max\, P(C)\ \prod_{i=1}^{n} P(F_i|C)} \quad (1)$$

The MAP (Maximum A Posteriori) is obtained from the maximum value or argmax (argument of the maxima) of the multiplication of prior probabilities ($\boldsymbol{P(C)}$) with the multiplication of the likelihood probabilities of the data ($\boldsymbol{P(F_i|C)}$).

2) Prior probability can be formulated as follows [12]:

$$\boldsymbol{P(C) = \frac{|doc_c|}{N}} \quad (2)$$

Where $\boldsymbol{doc_c}$ is the number of reviews of the Jiwa+ application by Janji Jiwa that belong to the class $\boldsymbol{C}$ (Positive, Neutral, or Negative) and $\boldsymbol{N}$ is the total number of reviews of the Jiwa+ application by Janji Jiwa from all classes.

3) Likelihood probability can be formulated as follows[12]:

$$\boldsymbol{P(F_i|C) = \frac{n_i+1}{|n+kata|}} \quad (3)$$

Where $\boldsymbol{n}$ is the number of words in class $\boldsymbol{C}$ (Positive, Neutral, or Negative), $\boldsymbol{n_i}$ is the frequency of occurrence of the word $\boldsymbol{F_i}$ in the reviews of the Jiwa+ application by Janji Jiwa that belong to class C (Positive, Neutral, or Negative), and $\boldsymbol{kata}$ is the total number of words in all classes.

Next, for sentiment classification using the Lexicon Based method. Sentiment label classification with the Lexicon Based method is based on matching words in the sentence with the Lexicon dictionary. The labels consist of three categories: positive, neutral, and negative. Determining the sentiment label with this method is determined by counting the number of words that match the words in the Lexicon dictionary. If a sentence contains more positive words, it will be labeled as positive; if it contains more negative words, it will be labeled as negative. If the number of negative and positive words in the sentence is equal, the sentence will be labeled as neutral [13].

In this research, the model evaluation stage was conducted using a confusion matrix. The confusion matrix used in this study is a 3x3 matrix consisting of True Negative (TN), True Neutral (TNe), True Positive (TP), False Negative (FN), False Neutral (FNe), and False Positive (FP). True Negative represents the number of data classified into the Negative class and consistent with the actual data. True Neutral represents the number of data classified into the Neutral class and is consistent with the actual data. True Positive represents the number of data classified into the Positive class and consistent with the actual data. Meanwhile, False Negative, False Neutral, and False Positive are the numbers of data classified into Negative, Neutral, and Positive classes, respectively, but are inconsistent with the actual data. The illustration of the confusion matrix table for this study can be seen in the following table.

Table 1. Confusion Matrix

		Prediction		
		Negative	Neutral	Positive
Actual	Negative	TN	$FNe_{negative}$	$FP_{negative}$
	Neutral	$FN_{neutral}$	TNe	$FP_{neutral}$
	Positive	$FN_{positive}$	$FNe_{positive}$	TP

with:

TN : True Negative

TP : True Positive

TNe : True Neutral

FN : False Negative

FP : False Positive

FNe : False Neutral

$\mathrm{FP_{neutral}}$: False positives classified as neutral class

$\mathrm{FP_{negative}}$: False positives classified as negative class

$\mathrm{FN_{neutral}}$: False negatives classified as neutral class

$\mathrm{FN_{positive}}$: False negatives classified as positive class

$\mathrm{FNe_{negative}}$: False neutral classified as negative class

$\mathrm{FNe_{positive}}$: False neutral classified as positive class

After obtaining the sentiment classification results for each method, the performance of both methods will be compared by comparing the accuracy, sensitivity, and specificity values of the models.

1) Accuracy represents the percentage of correct classifications among all data points when classified using a model [14]. Accuracy can be formulated as follows[15]:

$$Accuracy = \frac{TN+TP+TNe}{TN+TP+TNe+FN+FP+FNe} \quad (4)$$

2) Recall or sensitivity serves to provide an overview of how many data points can be correctly predicted as positive by the model, compared to all data points that actually belong to the positive class [14]. Sensitivity can be formulated as follows [15]:

$$Recall = \frac{TP}{TP++FP_{neutral}+FP_{negative}} \quad (5)$$

In contrast to sensitivity, specificity depicts the accuracy of the model in predicting actual negative data points compared to all data points labeled as negative[16]. Specificity can be formulated as follows [16]:

$$Specificity = \frac{TN}{TN++FN_{neutral}+FN_{positive}} \quad (6)$$

3. RESULT AND DISCUSSION

3.1. Data Collection of Jiwa+ App Reviews by Janji Jiwa Using Web Scraping Method

Data for Jiwa+ app reviews by Janji Jiwa was extracted from the Google Play Store using web scraping methods with the Python programming language and Jupyter Notebook software. The dataset used in this study consists of 597 reviews. From this data extraction, the attributes obtained include the username or app user's

account name providing the review, the date of the review, and the user's actual review content. The Jiwa+ app reviews dataset, obtained through the web scraping process, is presented in Table 1 below.

Table 1. App Review Result of Web Scraping

	User Name	At	Preprocessing Process
1	Wava Aulia	2024-01-29 12:40:00	Sore2 lg pingin makan jiwa toast yang kari, trs…
2	Devina Amallia	2024-01-28 12:47:25	IUnya cakepp hrs diapresiasi! Ngeliatnya seger…
3	Rosyid Ridlo	2024-01-28 01:50:49	Kembali terjadi, bayar pake sopipai ditolak si…
4	oktafya aa	2024-01-27 14:42:34	No worries, lg ga perlu antri buat beli jiwa to…
5	Zella Ade	2024-01-27 14:33:31	Wiii mantap dah ga cuma2 download Jiwa+ menan…

3.2. Data Preprocessing

Next, the data will undergo preprocessing because the scraped data still contains unnecessary parts for analysis. In this preprocessing stage, the data will be standardized and cleaned of unnecessary characters other than letters to facilitate analysis. The preprocessing process consists of several stages: case folding, data cleansing, spelling normalization, tokenizing, and stopword removal. The results of the data preprocessing can be seen in Table 2 below.

Table 2. Data Preprocessing Process

Steps	Review	Preprocessing Process
1	IUnya cakepp hrs diapresiasi! ngeliatnya seger ga bikin user puyeng dan informatif jugak!	App Review Result of Web Scraping
2	iunya cakepp hrs diapresiasi! ngeliatnya seger ga bikin user puyeng dan informatif jugak!	Case folding
3	iunya cakepp hrs diapresiasi ngeliatnya seger ga bikin user puyeng dan informatif jugak	Cleansing data
4	iunya cakep harus diapresiasi ngeliatnya segar tidak bikin user pusing dan informatif juga	Spelling normalization
5	"iunya" "cakep" "harus" "diapresiasi" "ngeliatnya" "segar" "tidak" "bikin" "user" "pusing" "dan" "informatif "juga"	Tokenizing
7	iunya cakep diapresiasi ngeliatnya segar tidak bikin user pusing informatif	Stopword removal

3.3. Data Visualization

In text mining analysis, word clouds are commonly used to provide descriptive visualizations of text data. A word cloud visualizes words based on their frequency of occurrence. This visualization method can effectively illustrate the distribution and prominence of words in the data. The word cloud visualization for this research can be observed in Figure 2 below.

Figure 1. Word Cloud App Reviews

From Figure 2, it can be observed that the most prominent words in the word cloud are "promo" and "outlet." This indicates that "promo" and "outlet" are the most frequently mentioned words by users in their reviews. Both words suggest that many users pay particular attention to promotions and outlets available in the Jiwa+ app by Janji Jiwa. Considering the high user interest in promotions and outlets, as shown in the word cloud, this could be a consideration for Kopi Janji Jiwa owners to maximize their promotions and evaluate the service of their outlets, as well as the features available in the app. This could potentially enhance user satisfaction and attract new users to use the Jiwa+ app by Janji Jiwa.

In addition to using word clouds, data visualization can also be presented with bar charts to provide a clearer picture of word frequencies. While the word cloud already indicates the most frequently occurring words, it doesn't provide information on the actual proportion of these words relative to the total words. Therefore, the following bar chart displays the frequency of each word and the percentage of each word in the data. The bar chart is shown in Figure 3 below.

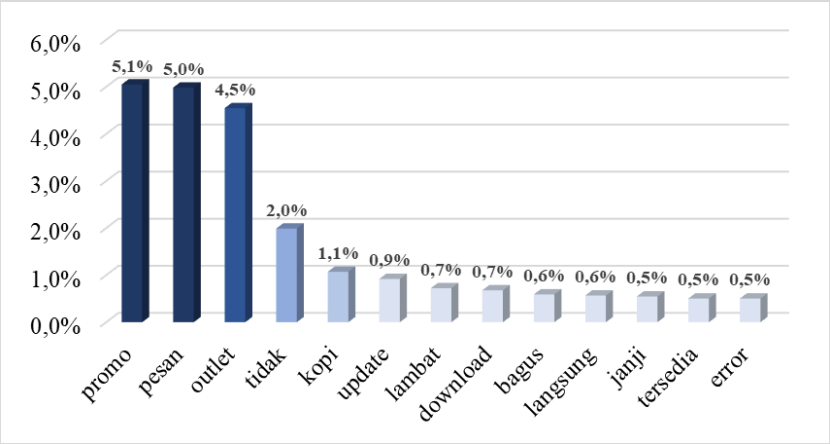

Figure 2. Most Frequently Occurring Words in the Comments

In Figure 3, we can see the top 13 words by frequency from all the reviews of the Jiwa+ app by Janji Jiwa. It is evident that the word "promo" ranks first as the most frequently used word in user reviews, accounting for 5,1% of the total. This indicates that promotions are the most highlighted feature or aspect of the Jiwa+ app by Janji Jiwa.

3.4. Sentiment Labeling Using Manual Method

Manual labeling involves subjectively assessing whether the sentences in Jiwa+ app reviews by Janji Jiwa are negative, positive, or neutral. This manual labeling is utilized for constructing the Naïve Bayes model and also serves as a comparison against the results of the Lexicon-Based classification in evaluating the model. The results of the manual labeling process can be seen in Table 3 below.

Table 3. Manual Labeling Result

No	Jiwa+ App Reviews by Janji Jiwa	Sentiment Classification
1	cakep banget aplikasinya kemaren dapet promo beli toast gratis minum	Positive
2	harap kedepannya mode gelap tampilan aplikasinya	Neutral
3	notif promo otw buka aplikasi cari yg promo an masukin keranjang wkwk tulisannya internet error mulu dahal dah ganti pake wifi data seluler aja niat kasi promo gak sih kemarin mulu	Negative
4	bayar pake sopipai ditolak sistem saldo giliran diulang tulisannya tagihan dibayar aneh aplikasinya ngaco	Negative

3.5. Sentiment Classification Using Naïve Bayes Method

Reviews of the Jiwa+ app by Janji Jiwa that have been manually labeled will undergo sentiment classification using the Naïve Bayes method. Before proceeding, the Jiwa+ app reviews data will be divided into training and testing sets with a ratio of 90%, where 491 reviews will be used for training and the remaining 10% (55 reviews) for testing. The training data will be used to build a model that will then be applied to classify the test data. From the 491 Jiwa+ app reviews, after processing with the `removeSparesterm` function, 83 words are identified for building the model. Once the probability values for each of these words are determined, they will be used as references in classifying the test data. Sentiment classification with Naïve Bayes involves calculating prior probabilities, likelihood probabilities, and posterior probabilities.

1) Prior Probability

After labeling is performed, we can determine the prior probabilities of each label using Equation (2), resulting in the probabilities for each label as shown in Table 4 below.

Table 4. Prior Probabilities of Each Class

Label	Prior Probabilities
Positive	0.35
Neutral	0.05
Negative	0.60

From Table 4, it is known that the prior probability for the negative sentiment label is 0,35. This means that before any additional information from the words used to build the model is considered, the probability of the negative label occurring in the data is 35%. Similarly, the neutral and positive labels have prior probabilities of 5% and 60%, respectively, indicating their likelihood before incorporating additional information from the model's words.

2) Likelihood Probability

After obtaining the prior probability values, the classification process proceeds by calculating the likelihood probabilities for each word to assess the likelihood that a word belongs to a particular class. For instance, we will calculate the likelihood probability for a Jiwa+ app review that reads "aplikasinya bagus". The likelihood probability can be calculated using Equation (2), and the results are presented in Table 5 below.

Table 5. Example Likelihood Probability for Each Word

Label	Word	
	aplikasinya	bagus
Positive	0.02	0.02
Neutral	0.02	0.02
Negative	0.02	0.004

From Table 5, the likelihood probability values for each word are known. The likelihood probability of the word "aplikasinya" for the positive label is 0.02. This means that in a sentence known to be labeled positive, the probability of the word "aplikasinya" appearing in that sentence is 0.02. Similarly, for other labels, if a sentence is known to be labeled neutral or negative, the probability of the word "aplikasinya" appearing in that sentence is also 0.02.

3) Posterior Probability

After determining the prior probabilities and likelihood values for each word, we can determine the sentiment label of a sentence by calculating the posterior probabilities for each label in the data. From Equation (1), the posterior probability values for each label in a given sentence are obtained as follows.

Table 6. Example Posterior Probability of Data

Label	Likelihood Probability		Prior Probability	Posterior Probability
	aplikasinya	bagus		
Positive	0.02	0.02	0.35	9.63×10^{-5}
Neutral	0.02	0.02	0.05	2.4×10^{-5}
Negative	0.02	0.004	0.06	4.29×10^{-5}

Based on Table 6, the highest posterior probability value is for the positive class, indicating that the sentence "aplikasinya bagus" is classified as positive. This step represents the classification process performed by Naïve Bayes in determining the sentiment of reviews for the Jiwa+ app by Janji Jiwa. Following the same approach, a model will be formed and used to label the sentiment of test data reviews for the Jiwa+ app by Janji Jiwa that do not yet have sentiment labels. From the classification results with Naïve Bayes, a confusion matrix for the test data is obtained as follows.

Table 7. Confusion Matrix for Classification with Naïve Bayes

		Classification Result Using Naïve Bayes		
		Negative	Neutral	Positive
Actual Data	Negative	24	5	0
	Neutral	0	2	0
	Positive	0	0	23

Then, the accuracy, sensitivity, and specificity of the Naïve Bayes model can be calculated. The accuracy, sensitivity, and specificity of the model can be seen in Table 8 below.

Table 8. Naïve Bayes Model Evaluation

Parameters	Value
Accuracy	90.74%
Sensitivity	100%
Specificity	82.76%

3.6. Sentiment Classification Using Lexicon Based Method

Sentiment classification with Lexicon Based utilizes matching words from the data with a lexical dictionary. Each word in the Jiwa+ app reviews data by Janji Jiwa will be analyzed and counted for how many words belong to positive and negative categories. The results of the sentiment classification process using the Lexicon Based method can be seen in Table 9 below.

Table 9. Sentiment Classification Result Using Lexicon Based

No	Jiwa+ by Janji Jiwa App Reviews Data After Preprocessing	Number of Positive Words	Number of Negative Words	Sentiment Classification
1	iunya cakep diapresiasi ngeliatnya segar tidak bikin user pusing informatif	3	2	Positive
2	notif promo otw buka aplikasi cari yg promo an masukin keranjang wkwk tulisannya internet *error* mulu dahal dah ganti pake wifi data seluler aja niat kasi promo gak sih kemarin mulu	3	3	Neutral
3	bayar pake sopipai ditolak sistem saldo giliran diulang tulisannya tagihan dibayar aneh aplikasinya ngaco	0	2	Negative

From the classification results with Lexicon Based, the confusion matrix for the test data is obtained as follows.

Table 10. Confusion Matrix for Classification with Lexicon Based

		Classification Result Using Lexicon Based		
		Negative	**Neutral**	**Positive**
Actual Data	Negative	14	7	8
	Neutral	1	0	1
	Positive	3	3	17

The accuracy, sensitivity, and specificity values for the classification method using Lexicon Based can be seen in Table 11 below.

Table 11. Lexicon Based Model Evaluation

Parameters	Value
Accuracy	57.41%
Sensitivity	65.38%
Specificity	82.35%

Looking at previous studies [4], the LexiconBased model achieved an accuracy, sensitivity, and specificity of 64.49%, 71.57%, and 87.88%, respectively. These values are slightly higher than the results of this study, although the difference is not substantial, which may be due to variations in the number of words in the dictionaries used. However, both studies show the same trend: the Lexicon-Based model has lower accuracy, sensitivity, and specificity compared to the Naïve Bayes model.

3.7. Analysis of The Classification Result

From the manual labeling method and Lexicon based, the classification results are as shown in the Figure below.

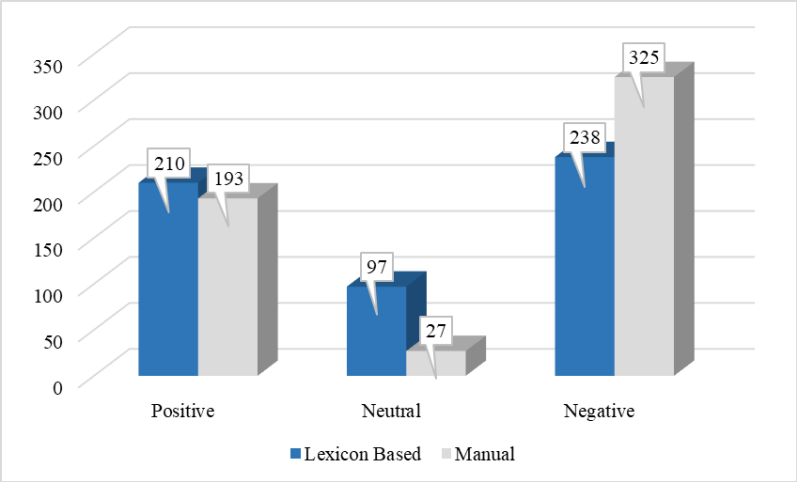

Figure 3. Classification Results Using Manual and Lexicon Based Methods

From the Figure 4, it can be seen that the classification results using the Lexicon Based and manual methods generally lead to the same conclusion, which is that most application reviews are negative in sentiment. Therefore, it can be concluded that users tend to give more negative reviews to the application. However, there is an aspect worth highlighting here, where the classification results using the Lexicon Based method exhibit a lower imbalance between classes compared to the results obtained from the manual classification method. This occurs due to the different classification processes employed between manual classification and the Lexicon Based method.

Based on the comparison of classification accuracy between Naïve Bayes and Lexicon-Based methods, it is evident that the Lexicon Based classification method has lower accuracy than Naïve Bayes. Therefore, in this study, Naïve Bayes exhibits a better model for classifying reviews of the Jiwa+ application by Janji Jiwa. Upon reviewing the confusion matrix in Table 10, it is found that 23 data points have differing labels between the Lexicon-Based method and the manual method. This discrepancy arises from differences in evaluation criteria between the two labeling methods.

The Lexicon-Based labeling method, which relies on the completeness of the dictionary and counts of positive and negative words in sentences, may yield different decisions compared to the manual labeling method based on subjective judgment. Sometimes, sentences with an equal number of positive and negative words might be labeled as neutral by the Lexicon-Based method, whereas subjective evaluation might classify such sentences as having negative or positive sentiment.

Another scenario occurs when a sentence contains more positive or negative words; the Lexicon-Based method might label it as positive or negative, while subjective evaluation could suggest the opposite sentiment or even neutrality. The table below provides examples of these differences between manual labeling and Lexicon-Based methods.

Table 12. Classification Results Between Manual and Lexicon-Based Method

App Review	Manual	Lexicon Based
pembelian promo lewat aplikasi sering di tolak	Negative	Neutral

From Table 12, we can observe the differences in classification results between manual and Lexicon-Based methods. In the manual labeling method, the review "pembelian promo lewat aplikasi sering di tolak" is classified as a negative sentiment review because it reflects user disappointment when attempting to order, and their order is rejected via the app. This clearly indicates a drawback or negative aspect of the application that needs attention to maintain customer satisfaction.

On the other hand, based on the Lexicon-Based classification method, this review is classified as having neutral sentiment because it contains an equal number of positive and negative words. The review has one positive word, "promo", and one negative word, "tolak"; thus, according to Lexicon-Based analysis, it is classified as neutral sentiment. Such differences also occur in other reviews, causing significant disparities between the classification results of the Lexicon-Based method and the manual method, which leads to lower accuracy in models using this classification approach.

In conclusion, the analysis suggests that while the Lexicon Based classification method is faster due to not requiring initial information for classification, it performs less effectively compared to Naïve Bayes when evaluated against actual data or manual labeling results. This limitation stems from its reliance on a dictionary-based scoring system, which may not fully capture the nuances of sentence meanings, leading to significant discrepancies in classification results compared to manual labeling.

4. CONCLUSION

Based on the results and discussion of sentiment analysis of Jiwa+ by Janji Jiwa app reviews using Naïve Bayes method and employing manual labeling and Lexicon Based approaches, the Naïve Bayes classification method has a superior model compared to Lexicon Based when considering accuracy, sensitivity, and specificity. Therefore, in the sentiment analysis research of Jiwa+ by Janji Jiwa app reviews, the method that performs best is Naïve Bayes. The drawback of Lexicon Based lies in its sentiment scoring system, which focuses on the counts of positive and negative words in sentences. In reality, the actual counts of positive and negative words may not accurately reflect the true intent of sentences. From this analysis, it can be concluded that Lexicon Based has limitations in understanding the intended meaning of sentences, leading to classifications that differ from the actual data.

REFERENCES

[1] V. R. Prasetyo, I. A. Ryanda, and Delta Ardy Prima, "Sentiment Analysis and Categorization of Customer Reviews on Kopi Paste Cafe Using Naive Bayes and K-nearest Neighbor Methods," *NERO Netw. Eng. Res. Oper.*, vol. 8, no. 1, 2024.

[2] Dania Siregar, Faroh Ladayya, Naufal Zhafran Albaqi, and Bintang Mahesa Wardana, "Penerapan Metode Support Vector Machines (SVM) dan Metode Naïve Bayes Classifier (NBC) dalam Analisis Sentimen

Publik terhadap Konsep Child-free di Media Sosial Twitter," *J. Stat. dan Apl.*, vol. 7, no. 1, pp. 93–104, 2023, doi: 10.21009/jsa.07109.

[3] M. Ikhsan Yusuf, N. Ransi, A. Arman, A. Tenriawaru, L. O. Saidi, and L. S. La Surimi, "Analisis Sentimen Komentar Netizen Instagram Terhadap Racism Di Sepak Bola Indonesia Dengan Metode Naive Bayes," *J. Mat. Komputasi dan Stat.*, vol. 2, no. 3, pp. 136–143, 2022, doi: 10.33772/jmks.v2i3.10.

[4] G. N. Aulia and E. Patriya, "IMPLEMENTASI LEXICON BASED DAN NAIVE BAYES PADA ANALISIS SENTIMEN PENGGUNA TWITTER TOPIK PEMILIHAN PRESIDEN 2019," *J. Ilm. Inform. Komput.*, vol. 24, no. 2, pp. 140–153, Aug. 2019, doi: 10.35760/ik.2019.v24i2.2369.

[5] D. Chrisinta and J. E. Simarmata, "Eksplorasi Teknik Web Scraping pada Data Mining: Pendekatan Pencarian Data Berbasis Python," *Fakt. Exacta*, vol. 17, no. 1, pp. 58–68, 2024, doi: 10.30998/faktorexacta.v17i1.22393.

[6] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.

[7] Suad A. Alasadi and Wesem S. Bhaya, "Review of Data Preprocessing Techniques," 2017.

[8] M. Rinestu, I. P. Made Indra, B. Marsanto, and S. Trisakti, "Classification Of Investment Decisions During Covid-19 Pandemic Using Naive Bayes Klasifikasi Keputusan Investasi Di Masa Pandemi Covid-19 Dengan Menggunakan Naive Bayes," *Manag. Stud. Entrepreneur. J.*, vol. 3, no. 4, pp. 1784–1796, 2022, [Online]. Available: http://journal.yrpipku.com/index.php/msej

[9] Riska Agustin, Vera Maya Santi, and Bagus Sumargo, "Metode Naive Bayes dalam Mendeteksi Sel Kanker Payudara," *J. Stat. dan Apl.*, vol. 3, no. 1, pp. 30–38, 2019, doi: 10.21009/jsa.03104.

[10] A. Basit, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Hasil Panen Padi," *JTIK (Jurnal Tek. Inform. Kaputama)*, vol. 4, no. 2, pp. 208–213, 2020, doi: 10.59697/jtik.v4i2.610.

[11] H. Perdana, N. Satyahadewi, R. Tamtama, and S. Anggriani, "The Use of Naïve Bayes Classifier to Predict the National Selection for State University Entrance (SNMPTN) Acceptance Status at Statistics Study Program in Tanjungpura University," *AIP Conf. Proc.*, vol. 2588, no. January, 2023, doi: 10.1063/5.0111969.

[12] V. Rizqiyani, A. Mulwinda, and R. D. M. Putri, "Klasifikasi Judul Buku dengan Algoritma Naive Bayes dan Pencarian Buku pada Perpustakaan Jurusan Teknik Elektro," *J. Tek. Elektro*, vol. 9, no. 2, pp. 60–65, 2017, [Online]. Available: https://journal.unnes.ac.id/nju/index.php/jte/article/view/11658/7332

[13] A. Priadana and A. A. Rizal, "Sentiment Analysis on Government Performance in Tourism During The COVID-19 Pandemic Period With Lexicon Based," *CAUCHY J. Mat. Murni dan Apl.*, vol. 7, no. 1, pp. 28–39, 2021, doi: 10.18860/ca.v7i1.12488.

[14] J. N. Permana, R. Goejantoro, and S. Prangga, "Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Prediksi Ketepatan Waktu Studi Mahasiswa (Studi Kasus: Program Studi Statistika Universitas Mulawarman)," *J. EKSPONENSIAL*, vol. 13, pp. 161–170, 2022, [Online]. Available: https://jurnal.fmipa.unmul.ac.id/index.php/exponensial/article/view/947

[15] E. I. Elmurngi and A. Gherbi, "Unfair reviews detection on Amazon reviews using sentiment analysis with supervised learning techniques," *J. Comput. Sci.*, vol. 14, no. 5, pp. 714–726, 2018, doi: 10.3844/jcssp.2018.714.726.

[16] A. C. Putri, F. E. Hariyanto, N. L. E. Andini, and Z. C. S. Zulkarnaen, "Klasifikasi Rumah Tangga Miskin di Provinsi Papua Tahun 2017 Menggunakan Metode Naive Bayes," *J. Sains Mat. dan Stat.*, vol. 7, no. 1, p. 89, 2021, doi: 10.24014/jsms.v7i1.11924.