Sentiment Analysis of Lazada Product Reviews using Convolutional Neural Network and Naïve Bayes Models

Ikhsan Maulana Siddiq, Kemas Muslim Lhaksmana*

School of Computing and Informatics, Telkom University, Bandung, Indonesia
Email: 1ikhsanmaulanas@student.telkomuniversity.ac.id, 2 kemasmuslim@telkomuniversity.ac.id
Correspondence Author Email: kemasmuslim@telkomuniversity.ac.id

Abstract—Lazada is one of the biggest marketplaces in Southeast Asia. One of the main features of Lazada is product reviews, any customer who has purchased and used a product from Lazada can provide reviews and ratings on the ones that have been purchased. Sentiment analysis on product reviews can help improve product and service quality, increase customer satisfaction, and improve purchasing decisions. Doing sentiment analysis of product reviews aims to help customers how to feel about the product, reading and analyzing each review manually is very not efficient. Sentiment analysis can automate handling large volumes of data quickly and accurately. In this research, using Lazada product review dataset to analyze sentiment by comparing Convolutional Neural Network (CNN) and Naïve Bayes. CNN and Naïve Bayes are two common methods used in text analysis and a comparison of their performance can provide the effectiveness of each in analyzing product sentiment. In this study, the authors propose to analyze the sentiment of product reviews using deep learning algorithm with CNN method. The results of this study explain that the CNN method can provide satisfactory results than Naïve Bayes. Based on the overall evaluation, CNN gets an accuracy value of 99.31%, precision of 99.31%, recall 99.31%, and f1-score of 99.31%, while Naïve Bayes gets the highest accuracy rate of 96.16%, precision 96.34%, recall 96.16%, and f1-score 96.16%.

Keywords: Sentiment Analysis; Lazada; CNN; Naïve Bayes; Accuracy

1. INTRODUCTION

Since 2020 there has been a Covid-19 pandemic that has attacked and spread to various countries in the world including Indonesia, which has an impact on community activities, especially economic activities [1]. To reduce the spread of the virus, the Indonesian government provides a social restriction policy for the public. With the making of this policy, it reduces income for people who buy and sell, and results in some sellers starting to leave their places to sell because the store is quiet, due to social restrictions implemented by the government [2].

The implementation of this social restriction policy has had an impact on the Indonesian economy. One of the impacts is the change in the way people shop, from direct shopping to indirect shopping (online), to fulfill basic needs and other interests. This way of shopping also has an impact on sellers, in which conventional sellers switch to trading online by joining marketplaces such as Lazada, to market the goods that want to be sold [3].

Lazada is one of the biggest marketplaces in Southeast Asia. Founded in 2012 by Rocket Internet, Lazada has grown into one of the leading online stores, serving millions of customers every day. Lazada provides a wide range of products, from electronics, clothing, home appliances, beauty, even food and beverages [4]. One of the main features of Lazada is product reviews, any customer who has purchased and used a product from Lazada can provide reviews and ratings about the one that has been purchased. Customer reviews have a significant role in online purchase decision making. As one of the leading marketplaces in Southeast Asia, Lazada brings thousands of product reviews every day [5].

Sentiment analysis and product reviews on Lazada have an important role in impacting purchasing decisions for consumers. In the digital era, product reviews are one of the most important factors in influencing purchasing decisions. And sentiment analysis can help in identifying sentiments related to existing product reviews. Therefore sentiment analysis and product reviews on Lazada can help improve product and service quality, increase customer experience, and improve purchasing decisions [6]. Sentiment analysis can help in identifying sentiments related to product reviews, such as positive or negative reviews. Positive sentiments can indicate that consumers are satisfied with products and services, while negative sentiments indicate that consumers are unsatisfied with products and services [7].

Customer insights play a crucial role in shaping purchasing decisions by offering valuable information about how other consumers perceive and interact with a product. These insights are typically gathered through various means, such as customer reviews, rating, and testimonials. By accessing this information, potential buyers can gain a comprehensive understanding of the product strengths, weaknesses, and overall user satisfaction. When customers see positive feedback and high ratings, customers are more likely to trust the product and feel confident in their purchasing decision. Conversely, if customers encounter negative reviews or consistent complaints, customer reconsider or look for alternative options.

Customer feedback gives companies direct insights into the customer experience, which is a crucial tool for improving the quality of their goods and services. By systemically collecting and analyzing feedback through various channels such as review and rating. Positive feedback highlights aspects that resonate well with customers, helping to reinforce successful features, while negative feedback pinpoints areas needing improvement, such as product defects, service inefficiencies or unmet customer expectations. This urgency in this matter lies in the competitive between marketplaces where businesses must quick to adapt on customer feedback to remain relevant.
and successful and the motivation behind this research is to address the increasing volume of product reviews and the need for efficient analysis method, where robust sentiment analysis is crucial for understanding customer sentiments and making informed business decisions.

By doing sentiment analysis of product reviews it aims to help customers how to feel about the product, by analyzing reviews it can also identify whether customer sentiment towards the product is positive, negative or neutral. In the digital world, product often get thousands of reviews. Reading and analyzing each review manually is very not efficient. Sentiment analysis can automate to handle large volumes of data quickly and accurately.

Manually analyzing thousands of customer reviews is inefficient process that can lead to delayed insights and inconsistent. Sentiment analysis, utilizing natural language processing and machine learning, automates the task, enabling businesses to swiftly and accurately process vast amounts of feedback. The rapid processing of data through sentiment analysis not only save time but also provides actionable insights by identifying recurring specific areas of concern or satisfaction. This allows businesses to respond promptly to customer needs, make informed decisions and maintain a competitive edge in a fast paced market environment.

However, to efficiently understand the sentiment of such reviews, robust analysis methods are required. Convolutional Neural Network and Naive Bayes methods have proven effective in analyzing sentiment on various types of data, including product reviews [8]. The use of these two methods in the context of sentiment analysis of Lazada reviews can provide a point of view for sellers and potential buyers. Comparing Convolutional Neural Network and Naive Bayes, sentiment analysis of Lazada product reviews can be done efficiently and accurately. Thereby providing a better understanding of customer sentiment towards the products offered by Lazada [9].

The related research in supporting this research is conducted by Muhammad Naufal Humam and colleagues [10] titled “Comparison of CNN and Naive Bayes Work on Sentiment Analysis of Manchester United Performance on Twitter”, this study uses CNN and Naive Bayes methods to analyze the sentiments of Manchester United supporters. Based on this research, different classification values are obtained. In the English dataset CNN is better than Naive Bayes with an accuracy value of 94%, precision 94%, recall 93%, and f1-score 94%, and Naive Bayes with accuracy 79%, precision 87%, recall 71%, and f1-score 72%. On the Indonesian dataset, CNN is also better than Naive Bayes with accuracy 91%, precision 91%, recall 91%, and f1-score 91%, and Naive Bayes with accuracy 75%, precision 85%, recall 68%, and f1-score 68%. In [11], author obtained that the system can categorize text sentiment in the class of positive, negative, and neutral to the condition of Covid-19. Predictions made with the Multinomial Naive Bayes method resulted in an accuracy of 74%, precision of 74%, and also recall of 74%. Research from [12] obtained the accuracy of sentiment classification using Random Forest method with 70% training data and 30% testing data is 97.38%. In [13], author gets the best accuracy 80.18%, the recall of 72.49%, precision 77.25%, and f1-score 74.73%. Research [14] of Sulindawaty and colleagues got the results of accuracy 99.5%, precision 99.94%, and recall 100%. Naive Bayes algorithm classification method is quite relevant even though the accuracy is not 100%.

2. RESEARCH METHODOLOGY

2.1 Research Stages

In this research, a system was formed that has several stages aimed at analyzing sentiments related to product reviews on Lazada. The system flowchart design shown in Figure 1.

![Figure 1. System Flowchart](image)

As shown in Figure 1, the first step in planning sentiment analysis of product reviews on Lazada is taking datasets from Kaggle. After the datasets obtained, it will be implemented for further processing. The next step is...
to preprocess the data to convert the raw data into ready to use data. This preprocessing process has several steps, including data cleaning, case folding, tokenization, feature extraction, stemming, stopword removal, and labeling. The next step is feature extraction to reduce the complexity of large dataset. One of the feature extraction methods used is TF-IDF. The next stage, split data is to make two parts, training data and test data. There is an imbalance between positive and negative classes, oversampling is done to overcome this problem. After reaching a balanced data class, CNN and Naive Bayes model training is carried out to produce sentiment analysis predictions. After that, the accuracy, precision, recall, and F1-score will be evaluated for both models.

2.2 Dataset

The dataset used in this research is a dataset obtained from product reviews on Lazada. The data used is a product review by customers totaling 203,788, after cleaning it gets 35,057 data and is labeled. Data labeling uses positive, negative, and neutral. The following labeling result are shown in Table 1.

<table>
<thead>
<tr>
<th>Label</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>8732</td>
</tr>
<tr>
<td>Negative</td>
<td>506</td>
</tr>
<tr>
<td>Neutral</td>
<td>25817</td>
</tr>
</tbody>
</table>

Based on Table 1, positive sentiment in this research was 8,732, the negative sentiment was 506, and the neutral sentiment was 25,817. In this research focuses on using positive and negative data.

2.3 Preprocessing

Preprocessing is the process of transforming raw data into a more easily understood version. This process is needed to correct errors in raw data that is incomplete and has an irregular format [15]. Several steps are required for data preprocessing, as follows:

a. Data Cleaning

Data cleaning is a process of data quality analysis that involves changing, deleting, or correcting data that is incorrect, corrupt, incomplete, or out of format. The purpose of data cleaning is to ensure the correctness, consistency, and usefulness of the data in the dataset, resulting in more accurate analysis results [16]. The data cleaning process shown in Table 2.

<table>
<thead>
<tr>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saya minat tapi kalau dikerim ke Gresik Jawa Timur bisa tidak? Kalau dari info pengiriman bebas tapi takut tidak sampai? Dan tanya apa ini sudah full HD atau belum? Terima kasih (I’m interested but if sent to Gresik East Java can you? If from the free shipping info but afraid of not arriving. And ask if this is full HD or not? Thank You)</td>
<td>Saya minat tapi kalau dikerim ke Gresik Jawa Timur bisa tidak Kalau dari info pengiriman bebas tapi takut tidak sampai Dan tanya apa ini sudah full HD atau belum Terima Kasih (Im interested but if sent to Gresik East Java can you if from the free shipping info but afraid of not arriving and ask if this is full HD or not thank you)</td>
</tr>
</tbody>
</table>

Based on Table 2, the data is cleaned by removing punctuation marks, special characters, and symbols, which aim make to the text easier to understand.

b. Case Folding

Case folding is the process of converting capital letters into lowercase letters in a text. This process is done to simplify sentiment analysis by eliminating capital letter differences that can affect the analysis results [17]. The outcomes of case folding in this research are shown in Table 3.

<table>
<thead>
<tr>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saya minat tapi kalau dikerim ke gresik jawa timur bisa tidak kalau dari info pengiriman bebas tapi takut tidak sampai dan tanya apa ini sudah full hd atau belum terima kasih (I'm interested but if sent to gresik east java can you? If from the free shipping info but afraid of not arriving and ask if this is full HD or not Thank You)</td>
<td>saya minat tapi kalau dikerim ke gresik jawa timur bisa tidak kalau dari info pengiriman bebas tapi takut tidak sampai dan tanya apa ini sudah full hd atau belum terima kasih (im interested but if sent to gresik east java can you if from the free shipping info but afraid of not arriving and ask if this is full hd or not thank you)</td>
</tr>
</tbody>
</table>

Based on Table 3, the data before the case folding process still used capital letters. At this step, capital letters were changed to lowercase, it can reduce the risk of problem from the sentiment analysis in this research.
Tokenization

Tokenization is the technique of dividing sentences in a collection of text documents into individual words. During the tokenization stage, special characters like as punctuation are eliminated, and all words are transformed to lowercase. These bits are known as tokens [18]. The tokenization process in this research are shown in Table 4.

<table>
<thead>
<tr>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>(im interested but if sent to gresik east java can you if from the free shipping info but afraid of not arriving and ask if this is full hd or not thank you)</td>
<td>(['im', 'interest', 'but', 'if', 'sent', 'to', 'gresik', 'java', 'east', 'can', 'you', 'if', 'from', 'the', 'free', 'shipping', 'info', 'but', 'afraid', 'of', 'not', 'arriving', 'and', 'ask', 'if', 'this', 'is', 'full', 'hd', 'or', 'not', 'thank', 'love'])</td>
</tr>
</tbody>
</table>

Based on Table 4, we can conclude on this process to divide the text into smaller units.

d. Stemming

Stemming is the process of transforming a word into basic form. The main objective of stemming is to remove affixes, so that words originating from the same word stem can be converted into the same form [19]. The stemming stage shown in Table 5.

<table>
<thead>
<tr>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>(['im', 'interest', 'but', 'if', 'sent', 'to', 'gresik', 'java', 'east', 'can', 'you', 'if', 'from', 'the', 'free', 'shipping', 'info', 'but', 'afraid', 'of', 'not', 'arriving', 'and', 'ask', 'if', 'this', 'is', 'full', 'hd', 'or', 'not', 'thank', 'love'])</td>
<td>(['im', 'interest', 'but', 'if', 'sent', 'to', 'gresik', 'java', 'east', 'can', 'you', 'if', 'from', 'the', 'free', 'shipping', 'info', 'but', 'afraid', 'of', 'not', 'arriving', 'and', 'ask', 'if', 'this', 'is', 'full', 'hd', 'or', 'not', 'thank', 'love'])</td>
</tr>
</tbody>
</table>

Based on Table 5, stemming stage is process removes every affix into their basic forms.

e. Stopword Removal

Stopword removal is part of the preprocessing stage which aims to remove irrelevant words in a sentence based on a stopword list. The stopword list that is usually used is in the form of a digital library whose list is already available [20].

2.4 Feature Extraction

Feature extraction is the process of transforming raw data into more useful information for further analysis, especially in the context of machine learning and data analytics. The main goal of feature extraction is to simplify the number of data sources required to accurately describe a phenomenon. In this process, important features of the data are identified and isolated so that machine learning algorithms can work more efficiently and effectively. One of the most popular and frequently used feature extraction in the text analysis is Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF is a method used in natural language processing and information retrieval. This method calculates the value of each word in a document based on the frequency of occurrence of the word and the frequency of documents containing the word. Term-Frequency (TF) is a statistic that measures the frequency with which a word appears in a document. This frequency can be calculated by counting the number of occurrences of the word in the document. Inverse Document Frequency (IDF) is a statistic that measures how rarely a word appears in all documents. IDF counts how many documents contain the word [21].

2.5 Split Data

Split Data is the process of splitting the dataset into several parts to train and test the model. The purpose is to make sure that the model can be tested correctly and its performance can be measured after training. The data is divided into two parts, test data and training data. Test data is used to evaluate the performance of the model and ensure prediction accuracy.
2.6 Oversampling

Oversampling is a data processing technique to balance and increase the amount of minority data. In this research, the data obtained is not balanced, this research uses the Synthetic Minority Over-sampling Technique (SMOTE) method to balance [22]. The process before and after using the SMOTE method can be shown in Figure 2.

![Figure 2. Before and After Using SMOTE](image)

Based on Figure 2, before doing the SMOTE method, there are 8,732 positive data and 508 negative data and after doing the SMOTE method, the positive and negative data are balanced.

2.7 One Hot Encoding

One Hot Encoding (OHE) is an encoding technique used to convert categorical data into binary form. In this research OHE is used to convert categorical data such as words or sentences into binary form that can be processed by CNN [23]. This method is particularly valuable in sentiment analysis, where text data is often transformed into numerical form to be processed by machine learning models. In sentiment analysis, One Hot Encoding can be applied to words or features extracted from the text, such as unigrams or bigram, converting them into a sparse matrix where each column represents a unique word or feature. This transformation enables algorithms to understand and process the text data effectively, facilitating the classification of sentiment. By representing text data in this binary format, One Hot Encoding helps models capture the presence or absence of specific words or features, which is crucial for accurate sentiment prediction and analysis.

2.7 Convolutional Neural Network

Convolutional Neural Network (CNN) is a type of artificial neural network used in various applications, including sentiment analysis. CNN used to recognize predictive information from objects such as text, sound bites, and images. CNN can automatically recognize relevant features and classify data into suitable categories.

![Figure 3. Convolutional Neural Network Architecture](image)

There are several layers of the CNN. First, Convolutional Layer, this layer uses convolution techniques to identify patterns in the data. Second, Pooling Layer, this is layer uses pooling techniques to reduce the size of data. Third, Connection Layer, this layer connects the output of the convolution and pooling layers. Fourth, Classification Layer, this layer uses a fully connected technique to connect the outputs. Fifth, Output Layer, this layer produces the final output of the Convolutional Neural Network, such as image or text classification [24].

2.8 Naive Bayes

Naive Bayes is classification method used in sentiment analysis to classify text as positive or negative based on word associated with sentiment. According to Bustomi [25] titled “Application of Naive Bayes Algorithm to
Classify Insurance Customer Data”, in this research discusses the Naive Bayes Algorithm, which is one of the algorithms used in classification approaches. Naive Bayes is a categorization approach that uses probability and statistical methods proposed by English scientist Thomas Bayes. It predicts future opportunities based on existing experience, thus the name Bayes Theorem. The theorem is combined with Naive, which assumes that the characteristics requirements are mutually independent. Naive Bayes classification assumes that the existence or absence of specific qualities of one class has no bearing on characteristics of other classes Equation 1 shows the formula of Naive Bayes.

\[
P(C|X) = \frac{P(X|C) \times P(C)}{P(X)}
\]  

Based on Equation 1 of formula Naive Bayes explained as follows:

1. P(C|X): Posterior probability, this is the probability of class C after observing X data. This is the value to be calculated in Naive Bayes classification. C is a particular class. X is the observed data or feature.
2. P(X|C): Likelihood, this is probability of getting data X if know that the class is C. Likelihood describes how likely it is that data X is generated from class C.
3. P(C): Class Prior Probability, this is the prior probability of class C without considering X data. Prior Probability describes how likely a class is to appear before any data. It is usually calculated based on the frequency of occurrence of the class in the training data.
4. P(X): Predictor Prior Probability, these are the probabilities of the data X in all classes. This is also referred to as the normalizing factor.

2.9 Evaluation

The evaluation step is to determine how well the processed model is performing. This procedure uses the data it analyzes to predict sentiment. The algorithms performance is visualized through the evaluation procedure as well. The matrix consists of:

1. True Positive (TP): Accurately, the data expected to be positive and is correctly predicted as positive.
2. True Negative (TN): Accurately, the data expected to be negative and is correctly predicted as negative.
3. False Positive (FP): Accurately, the data expected to be positive and is correctly predicted as negative.
4. False Negative (FN): Accurately, the data expected to be negative and is correctly predicted as negative.

Based on these values, such as Accuracy, Precision, Recall, and F1-Score will be calculated. The formulation used are explained as follows:

1. Accuracy

   Accuracy shows how accurate the applied model is to be able to classify correctly or can be referred to as the ratio of correct predictions to the entire existing data. Equation 2 shows the formula for calculating the accuracy values.

   \[
   \text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}
   \]  

2. Precision

   Precision shows the accuracy between the requested data and the predicted results given by the model or can be called the ratios of true positive predictions compared to all positive predicted results. Equation 3 shows the formula for calculating the precision values.

   \[
   \text{Precision} = \frac{TP}{TP+FP}
   \]  

3. Recall

   Recall indicates the success of the model to retrieve information or can be referred to as the ratio of positive true predictions compared to all positive true data. Equation 4 shows the formula for calculating the recall values.

   \[
   \text{Recall} = \frac{TP}{TP+FN}
   \]  

4. F1-Score

   F1-score shows the average comparison of precision and recall. F1-score is used as a performance reference or algorithm performance when the number of false negative and false positive data is not close to the number of false data. Equation 5 shows the formula for calculating the f1-score values.

   \[
   F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
   \]  

3. RESULT AND DISCUSSION

This section discusses the results of the experiments of all the models that have been created. This research builds various scenarios for each existing model, aiming to analyze the performance and effectiveness of each model.
The methods used in this research are CNN and Naive Bayes. Further explanation will be presented in the following section of test results and analysis of test results.

3.1 Test Result

Tests have been carried out to test the performance of the CNN and Naive Bayes methods. With the ratio of test data used, which is (60:40, 70:30, 80:20, 90:10). The results are reviewed through an evaluation matrix that includes accuracy, precision, recall, and f1-score. The following test results from these methods are presented in Table 6.

<table>
<thead>
<tr>
<th>Model</th>
<th>Ratio</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>60:40</td>
<td>97.75%</td>
<td>97.84%</td>
<td>97.75%</td>
<td>97.75%</td>
</tr>
<tr>
<td></td>
<td>70:30</td>
<td>99.29%</td>
<td>99.30%</td>
<td>99.29%</td>
<td>99.29%</td>
</tr>
<tr>
<td></td>
<td>80:20</td>
<td>99.31%</td>
<td>99.31%</td>
<td>99.31%</td>
<td>99.31%</td>
</tr>
<tr>
<td></td>
<td>90:10</td>
<td>98.56%</td>
<td>98.60%</td>
<td>98.56%</td>
<td>98.56%</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>60:40</td>
<td>94.77%</td>
<td>95.12%</td>
<td>94.77%</td>
<td>94.76%</td>
</tr>
<tr>
<td></td>
<td>70:30</td>
<td>95.51%</td>
<td>95.79%</td>
<td>95.51%</td>
<td>95.50%</td>
</tr>
<tr>
<td></td>
<td>80:20</td>
<td>95.84%</td>
<td>96.05%</td>
<td>95.84%</td>
<td>95.84%</td>
</tr>
<tr>
<td></td>
<td>90:10</td>
<td>96.16%</td>
<td>96.34%</td>
<td>96.16%</td>
<td>96.16%</td>
</tr>
</tbody>
</table>

From the test results that have been obtained using the CNN and Naive Bayes methods through four different data ratio, it appears that the data ratio has very important effect on the performance of the two models. The performance of both models increases with oversampling. Overall, the CNN method shows better performance when compared to Naive Bayes. At a data ratio of 80:20, the CNN method with TF-IDF feature extraction gives a high percentage in terms of accuracy, precision, recall, and f1-score. With an accuracy value of 99.31%, precision of 99.31%, recall of 99.31%, and f1-score of 99.31%.

3.2 Analysis of The Results

The Confusion Matrix is a table used to assess the effectiveness of a classification model. This table provides more information about how much data is successfully or wrongly classified. The Confusion Matrix compares the actual and projected values of the model, allowing it to create assessment metrics such as accuracy, precision, recall, and f1-score. The Confusion Matrix of the best performing Convolutional Neural Network model is displayed in Table 7.

<table>
<thead>
<tr>
<th>Predict Negative</th>
<th>Predict Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Negative</td>
<td>1741</td>
</tr>
<tr>
<td>Actual Positive</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>1728</td>
</tr>
</tbody>
</table>

Based on Table 7, the results of Confusion Matrix is:
1. True Positive (TP): 1728 positive predicted value that is truly positive corresponds to the actual value. There are 1728 cases where the model predicted a positive result and the result was indeed positive.
2. True Negative (TN): 1741 negative predicted value that is truly negative corresponds to the actual value. There are 1741 cases where the model predicted a positive result and the result was indeed negative.
3. False Positive (FP): 5 positive predicted value that negatively matched the actual values. There were 5 cases where the model predicted a positive outcome but was actually negative.
4. False Negative (FN): 19 negative predicted value that positively matched the actual values. There were 19 cases where the model predicted a negative outcome but was actually positive.

3.3 Method Discussion

a. Convolutional Neural Network

CNNs are well known for their outstanding efficiency in image and text categorization uses. In this study, the CNN model performed exceptionally well, particularly at the 80:20 data ratio, with high accuracy, precision, recall, and f1-score. CNNs capacity to catch complicated patterns in data is largely responsible for their effectiveness in sentiment analysis of product reviews.

b. Naive Bayes

Naive Bayes is a probabilistic classifier that works well with large datasets and is particularly useful for text classification. Although it performed slightly lower than the CNN model, Naive Bayes still achieved good results, demonstrating its reliability and efficiency. It performed consistently across different data ratios, though it was outperformed by CNN in most scenarios.

c. Comparison and Effectiveness
The comparative analysis of the two models indicates that while both models are effective for sentiment analysis, CNN consistently outperforms Naïve Bayes in terms of accuracy, precision, recall, and f1-score. The superior performance of CNN can be attributed to its capability to capture intricate patterns and dependencies in the data, which is essential for sentiment analysis. Naïve Bayes, on the other hand, relies on the assumption of feature independence, which may not hold true for complex datasets.

In conclusion, this research demonstrates that CNNs are highly effective for sentiment analysis of Lazada product reviews, providing accurate and reliable results. Naïve Bayes, while also effective, is slightly less accurate but still a viable option for sentiment analysis tasks. The choice of method may depend on the specific requirements and constraints of the application, with CNN being preferable for scenarios demanding higher accuracy and Naïve Bayes being suitable for simpler, more straightforward tasks.

4. CONCLUSION

Doing sentiment analysis of product reviews, it aims to help how to feel about the product, reading and analyzing each review manually is very not efficient. Sentiment analysis can automate handling large volumes of data quickly and accurately. In this research titled “Sentiment Analysis of Lazada Product Reviews Using the Convolutional Neural Network and Naïve Bayes Models”, the use of the Convolutional Neural Network method results in higher accuracy, precision, recall, and f1-score compared to the use of the Naïve Bayes method, with a data split ratio of 80:20. Convolutional Neural Network gets an accuracy value 99.31%, precision 99.31%, recall 99.31%, and f1-score 99.31%. Naïve Bayes only gets the highest accuracy level of 96.16%, precision 96.34%, recall 96.16%, and f1-score 96.16%.

REFERENCES


Ikhsan Maulana Siddiq, Copyright © 2024, MIB, Page 1584
Submitted: 25/06/2024; Accepted: 23/07/2024; Published: 27/07/2024