ILKOM Jurnal Ilmiah Vol. 13, No. 2, August 2021, pp. 101-109
Accredited 2nd by RISTEKBRIN No. 200/M/KPT/2020; E-ISSN 2548-7779 | P-ISSN 2087-1716

Research Article

Sentiment analysis of game product on shopee using the TF-IDF method and naive bayes classifier

Rifki Kosasih [a, 1, *]; **Anggi Alberto** [a, 2]
[a] *Universitas Gunadarma, Jl. Margonda Raya No. 100, Depok, 16424, Indonesia*
[1] *rifki_kosasih@staff.gunadarma.ac.id;* [2] *albertoanggi@gmail.com*
** Corresponding author*

Article history: Received January 10, 2021; Revised April 23, 2021; Accepted July 04, 2021; Available online August 07, 2021

Abstract

In every product sold on the E-commerce platform, there is a review column from consumers who have made transactions on the products. These reviews are in the form of comments and ratings (stars from one to five) written and given by consumers based on their assessment of the products purchased. With the product evaluation feature based on the rating, the consumer can find out how good or bad the quality of the product is. However, a problem arises when some consumers give negative comments with five stars or vice versa. This causes the product assessment feature based on the rating to be less good so that it does not represent the real value. Therefore, to determine the quality of the product, sentiment analysis was carried out using the TF-IDF method and the Naive Bayes Classifier based on reviews from buyers. The data collected is 1000 reviews which are divided into 700 training data and 300 test data. The next stage is the preprocessing text such as case folding (converting uppercase letters to lowercase), tokenizing (separating sentences into single words), stopwords (removing tokenizing conjunctions that have nothing to do with sentiment analysis), stemming (changing words into basic word forms), and word weighting with TF-IDF. The last step is to classify. Based on the classification results obtained an accuracy rate of 80.2223%.

Keywords: Feature; Product; Sentiment Analysis; TF-IDF; Naive Bayes Classifier

Introduction

In this era, the rapid growth of technology can turn conventional economic principles into the principles of the digital economy. The digital economy emerged due to an e-marketplace used by companies to sell their products online. Electronic Marketplace (E-Marketplace) is where buyers and sellers meet to make a transaction virtually [1]. Parties who sell products on e-marketplace sites are called sellers. For every product sold by the seller, there is a review column from buyers who have made transactions on the product. These reviews are in the form of comments and ratings (stars from one to five) written and given by consumers based on their assessment of the purchased products. Each product can have tens, hundreds, even thousands of reviews. This rating-based product assessment feature enables new buyers to see the product rating from one to five. Through this feature, new consumers can find out how good or bad the product is. However, there are problems such as, some consumers leave negative comments but give five-stars rating or vice versa. This can make the product assessment feature based on star ratings less good so that a product does not represent its true value.

Therefore, sentiment analysis is needed to conclude product reviews accurately and efficiently (apart from the rating feature). In this study, sentiment analysis was made on the products found on the e-marketplace site based on product reviews from customers who had made transactions with these products using the TF-IDF and Naive Bayes Classifier methods. The products used are game products available on the e-marketplace called Shopee.

Several studies have been conducted in sentiment analysis on e-marketplaces using Naive Bayes, such as Fiarni et al. who did sentiment analysis on an online store based on store reviews using the Naive Bayes method [2]. The sentiment is categorized into 3 classes, called positive class, neutral class and negative class. The data were taken from the results of online store reviews provided by customers through Facebook. The number of review data obtained was 1442 reviews. The objects reviewed were materials, products, prices, quality, design, service, exhibition space and the general public. The results of the study obtained an accuracy rate of 89.21% [2]. However, at the preprocessing text stage of her research, the stopword and word weighting stages were not carried out, so that in this study, those stages were added to the preprocessing text process.

Muthia conducted a sentiment analysis on a restaurant in Indonesia based on customer reviews using Naive Bayes. The review data was obtained from the restaurant website and the Zomato website. From the results of the study obtained an accuracy rate of 86.50% [3]. However, at the preprocessing text stage of her research, she only did tokenize.

http://dx.doi.org/10.33096/ilkom.v13i2.721.101-109

Furthermore, Sari conducted a sentiment analysis on the e-marketplace called JD.id, based on the results of a review from Twitter social media using the Naive Bayes Classifier. Sentiment is categorized into three classes, called positive class, neutral class and negative class. In her research, there is a feature of converting emoticons into sentiment classes in tweets. Her research obtained an accuracy rate of 96.44% [4]. However, the data used is limited because it has to use emoticons.

Based on the results of previous studies, it was found that the Naive Bayes method is a good method used in classifying sentiments. Therefore, in this study, the Naive Bayes method was used in classifying game product sentiment on the Shopee e-marketplace.

Method

This study used the Naive Bayes method to classify the results of game product reviews in the Shopee marketplace comments.

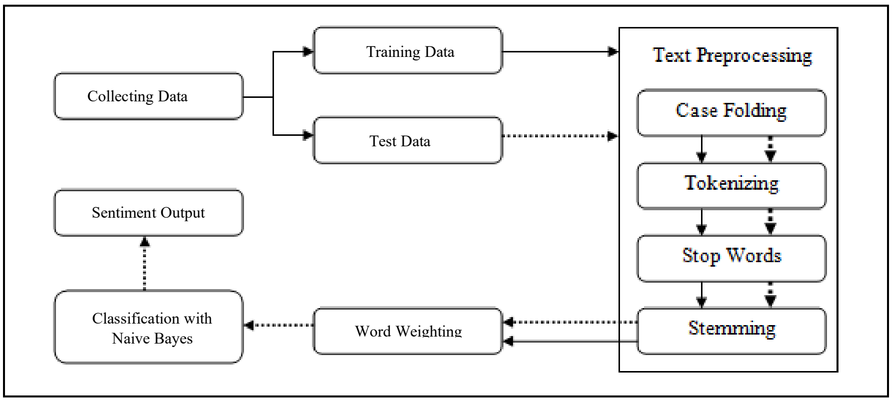

Figure 1. Research Framework

Based on **Figure 1**, the first stage is to collect data used as training and test data. The next stage is the Text Preprocessing stage. After the Text Preprocessing process is carried out, word weighting is carried out using the TF-IDF calculation, the TF-IDF calculation results from the word weighting stage are used to classify sentiments using the Naive Bayes Classifier algorithm.

A. Data Collection

At this stage, the data used comes from reviews on products on the Shopee site as many as 1000 review data. After collecting data, then the data is divided into 700 training data and 300 test data. The next step is to preprocess the text data to be processed at the next stage.

B. Text Preprocessing

At this stage, preprocessing is carried out to clean the data from noise and inconsistent data. The preprocessing text stages used are case folding, tokenizing, stopword and stemming [5]–[7]. At the case folding stage, the text that contains uppercase letters is converted into lowercase letters. After that, the tokenizing stage is carried out, that is separating sentences into single words.

After a single word is formed from tokenizing, the next step is to do stopword removal on the word resulting from tokenizing, which is checking every word resulting from tokenizing, such as if there are conjunctions or other words that are not related to sentiment analysis, they will be removed. The last stage in preprocessing is to do stemming, which changes the stopword results into the basic word form. After the preprocessing stage is complete, the word weighting is carried out.

C. Word Weighting

Word weighting is the process of assigning values to all words in the review data that have passed the preprocessing text process. The weighting is carried out using the TF-IDF method as in equations (1) and (2). The value (weight) is entered in the classification stage [8]–[10]. The steps in calculating word weights that must be carried out are as follows:

- Counting the number of Term Frequency (TF) per word

At this stage, the number of Term Frequency (TF) is calculated by separating sentences into one word and each word is given a value of 1.

- Counting the number of document frequency (DF) per word

At this stage, the document frequency (DF) is calculated by adding up the TF values for each word.

- Calculate the number of inverse document frequency (IDF)

After calculating the TF, the next step is to calculate the inverse document frequency. The following is the formula for Inverse Document Frequency (IDF).

$$IDF(w) = \log\left(\frac{N}{DF(w)}\right) \quad (1)$$

- Calculating the weight

The next step in the TF-IDF stage is to calculate the weight of each word. To get the weight, multiplication of the TF value with IDF is the same as the equation 2.

$$W_{ij} = \mathrm{tf_{ij}} \log\left(\frac{\mathrm{D}}{\mathrm{df_j}}\right) \quad (2)$$

After the weighting is done, then the test data classification is carried out based on the word-weighted data using the Naive Bayes Classifier method.

D. Naive Bayes Classifier

Naïve Bayes is a classification algorithm using conditional probabilities to calculate the probability of an event being selected into a certain class [11], [12]. To calculate the conditional probability, the TF-IDF method is used to use the word weighting result feature.

To calculate the probability depends on the number of classes and the number of features. For example, there are three categories; positive (C_1), neutral (C_2) and negative (C_3) with n word features G1, G2, … Gn. If there is an event x, then the probability of x to enter the three classes can be calculated as in equations (3), (4) and (5).

$$Pr(C_1|x) = Pr(C_1) * Pr(G_1|C_1) * Pr(G_2|C_1) * \cdots * Pr(G_n|C_1) \quad (3)$$
$$Pr(C_2|x) = Pr(C_2)\ Pr(G_1|C_2) * Pr(G_2|C_2) * \cdots * Pr(G_n|C_2) \quad (4)$$
$$Pr(C_3|x) = Pr(C_3)\ Pr(G_1|C_3) * Pr(G_2|C_3) * \cdots * Pr(G_n|C_3) \quad (5)$$

Where:

$$Pr(C_i) = \frac{|Doc_i|}{|sample|}$$
$$Pr(G_n|C_i) = \frac{n_k + 1}{n + |v|}$$

$Pr(C_i)$ is the probability that class ith will appear, $|Doc_i|$ is the number of documents that go to class i, $|sample|$ is the number of documents in the sample used for training, $Pr(G_n|C_i)$ is the probability of the appearance of the nth word feature if given class i, n_k is the frequency of occurrence of the word G_n in documents that have class i, n is the number of all words that include in in class i,$|v|$ is the number of words in the training sample.

To get the value of the conditional probability, a probability density function (pdf) is used in each category [13]. The results of the Naive Bayes classifier are vectoring whose contents are category probability values for each test data.

Results and Discussion

The data collected in this study is a data review of game products on Shopee about 1000 data and divided into 700 training data and 300 test data. The sample data is shown in **Table 1**.

Table 1. Example of Shopee Customer Review Data (5 out of 1000 data)

No.	Comments (Review)
1	The quality of the real product is as good as the description, thanks to the Seller and Shopee.
2	The product is nice. The delivery is fast and safe with extra bubble wrap package.
3	The quality of the real product is not as good as the description, already texted the seller for many times but still not received a response. not a recommended seller
4	The quality of the product is acceptable, but not packaged sufficiently.
5	Thank you, the package has arrived safely. Product works normally.

After that, to process the review data, preprocessing text was carried out first. The preprocessing stage is divided into 4 parts: the case folding stage, the tokenizing stage, the stopword removal stage, and the stemming stage. The results of case-folding are shown in **Table 2**.

Table 2. Results of Case Folding Processing (5 out of 1000 Data)

Kode	Comments (Review)	Case Folding
D1	The quality of the real product is as good as the description, thanks to the Seller and Shopee.	The quality of the real product is as good as the description, thanks to the Seller and Shopee.
D2	The the product is nice. The delivery is fast and safe with extra bubble wrap package.	The product is nice. The delivery is fast and safe with extra bubble wrap package.
D3	The quality of the real product is not as good as the description, already texted the seller for many times but still not received a response. not a recommended seller	The quality of the real product is not as good as the description, already texted the seller for many timesbut still not received a response. not a recommended seller
D4	The quality of the product is acceptable, but not packaged sufficiently.	The quality of the product is acceptable, but not packaged sufficiently.
D5	Thank you, the package has arrived safely. Product works normally.	Thank you, the package has arrived safely. Product works normally.

An example of tokenizing results is shown in **Table 3**.

Table 3. Case Tokenizing Results (5 out of 1000 Data)

Code	Case Folding	Tokenizing
D1	The quality of the real product is as good as the description, thanks to the Seller and Shopee.	[The, quality, of, the, real, product, is, as, good, as, the, description, thanks, to, the, Seller, and, Shopee.]
D2	The product is nice. The delivery is fast and safe with extra bubble wrap package.	[The, product, is, nice, The, delivery, is, fast, and, safe, with, extra, bubble, wrap, package]
D3	The quality of the real product is not as good as the description, already texted the seller for many times but still not received a response. not a recommended seller	[The, quality, of, the, real, product, is, not, as, good, as, the, description, already, texted, the, seller, for, many, times, but, still, not, received, a, response, not, a, recommended, seller]
D4	The quality of the product is acceptable, but not packaged sufficiently.	[The, quality, of, the, product, is, acceptable, but, not, packaged, sufficiently]
D5	Thank you, the complete package has arrived safely. Product works normally.	[Thank, you, the, complete, package, has, arrived, safely, Product ,works, normally]

Examples of stopword removal results are shown in **Table 4**.

Table 4. Results of Stopword Removal (5 out of 1000 Data)

Code	Tokenizing	Stopword Removal
D1	[The, quality, of, the, real, product, is, as, good, as, the, description, thanks, to, the, Seller, and, Shopee.]	[package, as good as, description, thank you, seller, shopee]
D2	[The, product, is, nice, The, delivery, is, fast, and, safe, with, extra, bubble, wrap, package]	[the product, nice, delivery, fast, safe, bubble, wrap, extra]
D3	[The, quality, of, the, real, product, is, not, as, good, as, the, description, already, texted, the, seller, for, many, times, but, still, not, received, a, response, not, a, recommended, seller]	[package, as good as, description, seller, texted, many, times, response, not received, not, recommended, seller]
D4	[The, quality, of, the, product, is, acceptable, but, not, packaged, sufficiently]	[acceptable, product, but, not, packaged, sufficiently]
D5	[Thank, you, the, complete, package, has, arrived, safely, Product, works, normally]	[thank you, package, complete, safe, product, works, normally]

Examples of stemming results are shown in **Table 5**.

Table 5. Stemming Results (5 out of 1000 data)

Code	Stopword Removal	Stemming
D1	[package, as good as, description, thank you, seller, shopee]	[package, as good as, description, thank you, seller, shopee]
D2	[the product, nice, delivery, fast, safe, bubble, wrap, extra]	[the product, nice, delivery, fast, safe, bubble, wrap, extra]
D3	[package, as good as, description, seller, texted, response, many, times, not received, not, recommended, seller]	[package, as good as, description, seller, texted, many, times, response, not received, not, recommended, seller]
D4	[acceptable, product, but, not, packaged, sufficiently]	[acceptable, product, but, not, packaged, sufficiently]
D5	[thank you, package, complete, safe, product, works, normally]	[thank you, package, complete, safe, product, works, normally]

After the preprocessing stage is complete, then labels are made on the data, which called positive labels (1), neutral labels (0) and negative labels (-1). The examples of training data and test data after preprocessing and labeled are shown in **Table 6** and **Table 7**.

Table 6. Example of Training Data (4 out of 700 Data)

Code	Before Preprocessing	After Preprocessing	Label
D1	The quality of the real product is as good as the description, thanks to the Seller and Shopee.	[package, as good as, description, thank you, seller, shopee]	Positiv e
D2	The product is nice. The delivery is fast and safe with extra bubble wrap package.	[the product, nice, delivery, fast, safe, bubble, wrap, extra]	Positiv e
D3	The quality of the real product is not as good as the description, already texted the seller for many times but still not received a response. not a recommended seller	[package, as good as, description, seller, texted, many, times, response, not received, not, recommended, seller]	Negati ve
D4	The quality of the product is acceptable, but not packaged sufficiently.	[acceptable, product, but, not, packaged, sufficiently]	Neutral

Table 7. Sample Test Data (1 out of 300 Data)

Code	Before Preprocessing	After Preprocessing	Label
D5	Thank you, the complete package has arrived safely. Product works normally.	[thank you, package, complete, safe, product, works, normally]	?

Table 6 is 4 of 700 training data that has been collected, while Table 7 is test data that still have no label. After that, word weighting was carried out on the training data and the preprocessing test data using the TF-IDF method as shown in **Table 8**.

Table 8. Example of Making Frequency Term (TF)

Words	TF				DF
	D1	D2	D3	D4	
Safe		1			1
Good		1			1
Nice	1				1
Product		1	1	1	3
bubble		1			1
Package				1	1
Fast		1			1
Texted			1		1
Try					1
But				1	1
Description	1		1		2
Extra		1			1
Times			1		1
Send		1			1
Condition	1				1
acceptable				1	1
not			1		1
Package	1				1
Order	1				1
recommended			1		1
response			1		1
seller	1		1		2
As good as	1		1		2
Shopee	1				1
sufficient				1	1
Received	1				1
Thank you	1				1
Wrap		1			1

The next step is to classify the test data using the Naive Bayes classifier method. The Naive Bayes calculation process can be seen in **Table 9** and **Table 10**.

Table 9. Example of the formation of $Pr(G_n|C_i)$

C_i	**Positive Sentiment**		**Neutral Sentimen**		**Negative Sentiment**	
	n_k	n	n_k	n	n_k	n
	1	14	1	5	1	10
|v|	29		29		29	
$Pr(G_n\|C_i)$	$\frac{1}{43}$		$\frac{1}{34}$		$\frac{1}{39}$	

Based on **Table 9**, for example, if the number of k^th words in the positive category document (n_k) is 1, all the words in the document that have a positive category (n) are 14 and the number of words in the training sample (|v|) is 29, then the probability of the appearance of the word G_n is given a positive class $\left(Pr(G_n|C_i)\right)$ is $\frac{1}{43}$. In the same way, it is also carried out on the neutral and negative sentiments as in **Table 9.** The next step is to calculate the probability of each word in the training data for each class, as shown in **Table 10**.

Table 10. Probability of each word if given the i^th class

G_n	$Pr(G_n\|C_i)$		
	Positif $P(C_1) = \frac{2}{4}$	**Netral** $P(C_2) = \frac{1}{4}$	**Negatif** $P(C_3) = \frac{1}{4}$
Safe	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Good	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Nice	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Product	$\frac{2}{43}$	$\frac{2}{34}$	$\frac{2}{39}$
bubble	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Package	$\frac{1}{43}$	$\frac{2}{34}$	$\frac{1}{39}$
Fast	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Texted	$\frac{1}{43}$	$\frac{1}{34}$	$\frac{2}{39}$
Try	$\frac{1}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
But	$\frac{1}{43}$	$\frac{2}{34}$	$\frac{1}{39}$
Description	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{2}{39}$
Extra	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Times	$\frac{1}{43}$	$\frac{1}{34}$	$\frac{2}{39}$
Send	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Condition	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
acceptable	$\frac{1}{43}$	$\frac{2}{34}$	$\frac{1}{39}$
not	$\frac{1}{43}$	$\frac{1}{34}$	$\frac{2}{39}$
Package	$\frac{3}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Order	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
recommended	$\frac{1}{43}$	$\frac{1}{34}$	$\frac{2}{39}$
response	$\frac{1}{43}$	$\frac{1}{34}$	$\frac{2}{39}$
seller	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{3}{39}$
As good as	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{2}{39}$
Shopee	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
sufficient	$\frac{1}{43}$	$\frac{2}{34}$	$\frac{1}{39}$
Received	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$

G_n	$Pr(G_n\|C_i)$		
	Positif $P(C_1) = \frac{2}{4}$	**Netral** $P(C_2) = \frac{1}{4}$	**Negatif** $P(C_3) = \frac{1}{4}$
Thank you	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$
Wrap	$\frac{2}{43}$	$\frac{1}{34}$	$\frac{1}{39}$

The calculation of the probability of each word if given the i^th class, in Table 10 is calculated using the calculations as in Table 9. The next step is to classify using the test data in **Table 7**. The calculation of the probability with a positive class if given test data (**Table 7**) using equation (3) is as follows:

$$\mathrm{Pr}(\mathrm{C_1}|\mathrm{x}) = \mathrm{Pr}\ (\mathrm{C_1}) \times \mathrm{Pr}(\text{Thank you}|\mathrm{C_1}) \times \mathrm{Pr}(\text{Package}|\mathrm{C_1}) \times \mathrm{Pr}(\text{Complete}|\mathrm{C_1}) \times \mathrm{Pr}(\text{Safe}|\mathrm{C_1}) \times \mathrm{Pr}(\text{Product}|\mathrm{C_1}) \times \mathrm{Pr}(\text{Works}|\mathrm{C_1}) .\times\ \mathrm{Pr}\ (\text{normally}|\mathrm{C_1})$$

$$\mathrm{Pr}(C_1|x) = \frac{2}{4}.\frac{3}{43}.\frac{3}{43}.\frac{1}{43}.\frac{1}{43}.\frac{1}{43}.\frac{1}{43}.\frac{1}{43} = 1{,}6555\,.10^{-11}.$$

The calculation of probability with neutral class if given test data (Table 7) using equation (4) is as follows:

$$\mathrm{Pr}(C_2|x) = \mathrm{Pr}\ (C_2) \times \mathrm{Pr}(\text{Thank you}|C_2) \times \mathrm{Pr}(\text{Package}|C_2) \times \mathrm{Pr}(\text{Complete}|C_2) \times \mathrm{Pr}(\text{Safe}|C_2) \times \mathrm{Pr}(\text{Product}|C_2) \times \mathrm{Pr}(\text{Works}|C_2) .\times\ \mathrm{Pr}\ (\text{normally}|C_2)$$

$$\mathrm{Pr}(C_2|x) = \frac{1}{4}.\frac{1}{34}.\frac{1}{34}.\frac{1}{34}.\frac{1}{34}.\frac{1}{34}.\frac{1}{34}.\frac{1}{34} = 0{,}47597.10^{-11}.$$

The calculation of the probability with a negative class if given test data (Table 7) using equation (5) is as follows:

$$\mathrm{Pr}(C_3|x) = \mathrm{Pr}\ (C_3) \times \mathrm{Pr}(\text{Thank you}|C_3) \times \mathrm{Pr}(\text{Package}|C_3) \times \mathrm{Pr}(\text{Complete}|C_3) \times \mathrm{Pr}(\text{Safe}|C_3) \times \mathrm{Pr}(\text{Product}|C_3) \times \mathrm{Pr}(\text{Works}|C_3) .\times\ \mathrm{Pr}\ (\text{normally}|C_3)$$

$$\mathrm{Pr}(C_3|x) = \frac{1}{4}.\frac{1}{39}.\frac{1}{39}.\frac{1}{39}.\frac{1}{39}.\frac{1}{39}.\frac{1}{39}.\frac{1}{39} = 0{,}18217.10^{-11}.$$

From these results it can be seen that $\mathrm{Pr}(C_1|x)$ as the greatest probability value compared to the other classes so that the test data given is classified into a positive class. The results of the classification with Naive Bayes as a whole are shown in **Table 11**.

Table 11. Prediction results of test data

		Prediction		
		Negative	**Neutral**	**Positive**
	Negative	74	20	4
Actual	Neutral	0	3	0
	Positive	19	46	134

Based on **Table 11**, negative sentiments in actual conditions predicted to be negative are 74, neutral predictions are 20 and positive predictions are 4. Positive sentiments in the actual situation predicted to be negative are 19, predicted neutral are 46, and positive prediction are 134. To evaluate, the next step is to first transform the confusion matrix in **Table 11** into a confusion matrix with 2 classes shown in **Tables 12, 13 and 14**.

Table 12. Prediction Results for Negative Class

		Prediction	
	Label	**Negative**	**Not Negative**
Actual	Negative	74	24
	Not Negative	19	183

Table 13. Prediction Results for Neutral Class

		Prediction	
	Label	**Neutral**	**Not Neutral**
Actual	Neutral	3	0
	Not Neutral	66	231

Table 14. Prediction Results for Positive Class

		Prediction	
	Label	**Positive**	**Not Positive**
Actual	Positive	134	65
	Not Positive	4	97

After that, the test results were evaluated by calculating the value of precision, recall, f1 score and accuracy in each category using (6), (7), (8) and (9) [14]–[16].

$$\text{precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \times 100\% \quad (6)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP+FN}} \times 100\% \tag{7}$$
$$\text{accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \times 100\% \tag{8}$$

$$\text{f1 score} = \frac{2\times\text{precision}\times\text{recall}}{\text{precision+recall}} \tag{9}$$

In **Table 9** for the negative class, TP (Negative predicted negative) = 74, FP (Negative predicted not negative) = 24, TN (Not negative predicted not negative) = 183, FN (Not Negative predicted Negative) = 19. Based on the results, the calculated precision, recall and accuracy for Table 9 are as follows:

$$\text{precision}_1 = \frac{\text{TP}}{\text{TP + FP}} \times 100\% = \frac{74}{74+24} \times 100\% = 75{,}51\%$$
$$\text{recall}_1 = \frac{\text{TP}}{\text{TP + FN}} \times 100\% = \frac{74}{74+19} \times 100\% = 79{,}569\%$$
$$\text{accuracy}_1 = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}} \times 100\% = \frac{74+183}{74+183+24+19} \times 100\% = 85{,}667\%$$

In **Table 10** for the neutral class, the value of TP (Neutral is predicted to be neutral) = 3, FP (Neutral is predicted to be neutral) = 0, TN (Not neutral is predicted to be not neutral) = 231, FN (Not Neutral is predicted to be Neutral) = 66. Based on the results, the calculated precision, recall and accuracy for Table 10 are as follows:

$$\text{precision}_2 = \frac{\text{TP}}{\text{TP + FP}} \times 100\% = \frac{3}{3+0} \times 100\% = 100\%$$
$$\text{recall}_2 = \frac{\text{TP}}{\text{TP + FN}} \times 100\% = \frac{3}{3+66} \times 100\% = 4{,}347\%$$
$$\text{accuracy}_2 = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}} \times 100\% = \% = \frac{3+231}{3+231+0+66} \times 100\% = 78\%$$

In **Table 11** for the positive class, the value of TP (Positive predicted positive) = 134, FP (Positive predicted not positive) = 65, TN (Not positive predicted not positive) = 97, FN (Not Positive predicted positive) = 4. Based on the results, the calculated precision, recall and accuracy for Table 11 are as follows:

$$\text{precision}_3 = \frac{\text{TP}}{\text{TP + FP}} \times 100\% = \frac{134}{134+65} \times 100\% = 67{,}336\%$$
$$\text{recall}_3 = \frac{\text{TP}}{\text{TP + FN}} \times 100\% = \frac{134}{134+4} \times 100\% = 97{,}101\%$$
$$\text{accuracy}_3 = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}} \times 100\% = \frac{134+97}{134+97+4+65} \times 100\% = 77\%$$

After obtaining the values of precision, recall and accuracy for each class, then the average of precision, recall and accuracy is calculated using equations (10), (11) and (12) [17].

$$\text{Average of Precision} = \frac{\sum_{i=1}^{k} \text{Precision}_i}{k} \tag{10}$$
$$\text{Average of Recall} = \frac{\sum_{i=1}^{k} \text{Recall}_i}{k} \tag{11}$$
$$\text{Average of accuracy} = \frac{\sum_{i=1}^{k} \text{accuracy}_i}{k} \tag{12}$$

Based on the equation (9), (10), (11) dan (12) then obtained

$$\text{Average of Precision} = \frac{75{,}51\% + 100\% + 67{,}336\%}{3} = 80{,}94867\%$$
$$\text{Average of Recall} = \frac{79{,}569\% + 4{,}347\% + 97{,}101\%}{3} = 60{,}339\%$$
$$\text{Average of accuracy} = \frac{85{,}667\% + 78\% + 77\%}{3} = 80{,}2223\%$$
$$\text{f1 score} = \frac{2 \times \text{Average of Precision} \times \text{Average of Recall}}{\text{Average of Precision} + \text{Average of Recall}} = \frac{2 \times 80{,}94867\% \times 60{,}339\%}{80{,}94867\% + 60{,}339\%} = 0{,}691372$$

Conclusion

Today, online shopping is very popular among people because it enables consumers to make purchases without going to the actual shop. The consumers usually see first the reviews and ratings given from previous buyers to know the quality of the product. However, errors often occur in rating, such as buyers leave negative comments but give five stars or vice versa. This makes the product rating feature based on star ratings less good. Therefore, another method is needed to find out the quality of these products, such as conducting sentiment analysis using the TF-IDF and Naive Bayes Classifier methods based on reviews from buyers to conclude reviews of the product appropriately efficiently (apart from the rating feature). The data collected were 1000 game product reviews on the online shopping site Shopee, divided into 700 training data and 300 test data. Based on the study results obtained an accuracy rate of 80.2223% and an f1 score of 0.691372. For further research as a comparison, sentiment analysis will be carried out using the K Nearest Neighbor (KNN) and Convolutional Neural Network (CNN) algorithms. The sentiment analysis dataset is available at http://rifki_kosasih.staff.gunadarma.ac.id/Downloads/files/91288/dataset-comment-product-category-gaming.csv.

References

[1] Y. Kurniawan, S. E. Hiererra, H. Nicholas, J. Setiawan, and Frendy, "The Effect Of Indonesian E-marketplace Reputation On Customer Trust In Online Shopping," *JATIT*, vol. 98, no. 03, pp. 505–516, 2020.

[2] C. Fiarni, H. Maharani, and R. Pratama, "Sentiment Analysis System for Indonesia Online Retail Shop Review Using Hierarchy Naive Bayes Technique," in *International Conference on Information and Communication Technologies (ICoICT)*, 2016, pp. 212–217.

[3] D. A. Muthia, "Analisis Sentimen Pada Review Restoran Dengan Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes," *J. Ilmu Pengetah. Dan Teknol. Komput.*, vol. 2, no. 2, pp. 39–45, 2017.

[4] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online JD.id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon," *SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.

[5] F. C. Permana, Y. Rosmansyah, and A. S. Abdullah, "Naive Bayes as Opinion Classifier to Evaluate Students Satisfaction Based on Student Sentiment in Twitter Social Media," in *The Asian Mathematical Conference*, 2017, pp. 1–9.

[6] A. Indriani, "Klasifikasi Data Forum dengan Menggunakan Metode Naïve Bayes Classifier," in *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 2014, pp. 5–10.

[7] Z. Efendi and Mustakim, "Text Mining Classification Sebagai Rekomendasi Dosen Pembimbing Tugas Akhir Program Studi Sistem Informasi," in *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI)*, 2017, pp. 235–242.

[8] F. Nurhuda, S. W. Sihwi, and A. Doewes, "Analisis Sentimen Masyarakat Terhadap Calon Presiden Indonesia 2014 Berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier," *IJurnal ITSMART*, vol. 2, no. 2, pp. 35–42, 2013.

[9] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat Pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 12, pp. 1750–1757, 2017.

[10] Normah, "Naïve Bayes Algorithm For Sentiment Analysis Windows Phone Store Application Reviews," *J. Publ. Informatics Eng. Res.*, vol. 3, no. 2, pp. 13–19, 2019.

[11] S. F. Rodiyansyah and E. Winarko, "Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naive Bayesian Classification," *IJCCS*, vol. 7, no. 1, pp. 13–22, 2013.

[12] A. P. Wibawa *et al.*, "Naïve Bayes Classifier for Journal Quartile Classification," *IJES*, vol. 7, no. 2, pp. 91–99, 2019.

[13] R. Kosasih and A. Fahrurozi, "Pengklasifikasian Bunga Dengan Menggunakan Metode Isomap dan Naive Bayes Classifier," *J. Ilm. Inform. dan Komput.*, vol. 22, no. 3, pp. 171–179, 2017.

[14] R. Kosasih, A. Fahrurozi, and I. Mardhiyah, "Vehicle Detection Using Principal Component Analysis," *J. Ilm. Komputasi*, vol. 19, no. 2, pp. 155–160, 2020.

[15] D. P. Lestari, R. Kosasih, T. Handhika, Murni, I. Sari, and A. Fahrurozi, "Fire Hotspots Detection System on CCTV Videos Using You only Look Once (YOLO) Method and Tiny YOLO Model for High Buildings Evacuation," in *2nd International Conference of Computer and Informatics Engineering IC2IE*, 2019, pp. 87–92.

[16] R. Kosasih and M. Arfiansyah, "Pendeteksian Kendaraan dengan Menggunakan Metode Running Average Background Substraction dan Morfologi Citra," *J. Media Inform. Budidarma*, vol. 4, no. 4, pp. 979–985, 2020.

[17] M. Sokolova and G. Lapalme, "A Systematic Analysis of Performance Measures for Classification Tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.