JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1070-1078
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6311

Sentiment Analysis of Practo Application Reviews Using Naïve Bayes and TF-IDF Methods

Rizal Adi Putranto*, Mahendra Dwifebri Purbolaksono, Widi Astuti

Fakultas Informatika, Informatika, Telkom University, Bandung, Indonesia
Email: [1,*]rizaladip@student.telkomuniversity.ac.id, [2]mahendradp@telkomuniversity.ac.id, [3]widiwdu@telkomuniversity.ac.id
Correspondence Author Email: rizaladip33@gmail.com

Abstract−Entering the 4.0 era, it seems that the healthcare industry is the one most likely to benefit from the combination of physical, digital and biological systems. Digital health applications or telemedicine have experienced significant growth in recent years. In the current era, the development of telemedicine is accelerating, one of which is the Practo application. As the number of users using this app increases, it is important to get their opinions in order to improve the health services provided by the app. Therefore, sentiment analysis of the comments regarding the health services on the app is necessary to find out the users' opinions. By utilizing sentiment analysis, it is possible to use the sentiment analysis results obtained as a sample that corresponds to both positive and negative comments. In addition, it can be revealed that there is a mismatch between the ratings and comments given by users. This information has the benefit of being able to improve the Practo application and improve the health services provided to more effectively meet the needs and expectations of users. This research employs the Naïve Bayes approach for sentiment analysis, utilizing TF-IDF feature extraction. Naïve Bayes was chosen because it is known as an efficient classification algorithm but has a high level of accuracy. This approach involves utilizing the Bayes rule formula to calculate probabilities and make classifications. It is applicable for solving classification problems that involve either numeric or nominal feature data. Meanwhile, TF-IDF was chosen because it can associate each word in a document with a numerical value that reflects its level of relevance to the document. TF-IDF is used to measure the weighting of words as features in the summary. In this study, the best model achieved a performance with an f1-score of 85.50%.

Keywords: Sentiment Analysis; Practo Application Review; Naïve Bayes; TF-IDF

1. INTRODUCTION

Along with the development of technology, especially the internet, the internet is becoming an increasingly important part, especially for mankind [1]. Entering the 4.0 era, it seems that the health industry is the industry most likely to benefit from the combination of physical, digital, and biological systems [2]. In recent times, there has been a significant expansion in the use and development of digital health applications and telemedicine. In the present era, telemedicine is advancing at an accelerated pace, with one notable example being the Practo application [3] To enhance the healthcare services within the application, it becomes imperative to analyze the sentiments expressed in the user comments as the user base continues to grow. This sentiment analysis will provide valuable insights and opinions regarding the users' experiences with the health services offered through the application [4].

Sentiment analysis is an approach employed to extract opinion-based information, automatically comprehend and process textual data, aiming to identify and determine the sentiment expressed within an opinion [5]. The results of sentiment analysis on Practo application reviews will be divided into two parts, namely positive and negative reviews. Positive reviews reflect the valuable feedback provided by application users, whereas negative reviews suggest that the comments given have lesser value, implying that the reviewed Practo application is not satisfactory. The importance of doing this sentiment analysis is because the sentiment obtained can be sampled according to the comments, both positive and negative. In addition, it can also be seen that there is an inconsistency between the rating and comments given by the user, which in turn, the data containing this information is useful for developing the product quality of the Practo application in the future.

In sentiment analysis, several types of features can be used that will make it easier to get sentiment analysis Feature extraction is necessary because certain features can have a negative impact on classification performance [6]. In this research problem, the feature extraction technique employed is TF-IDF, which quantifies the importance of words by assigning them weights based on their frequency and rarity in the dataset [7]. The reason for selecting this feature is its ability to assign a numerical value to each word in a document, reflecting its importance or relevance within that document [8].

The approach employed for sentiment analysis in this study utilizes the Naïve Bayes method. Naïve Bayes is a classification method by calculating probabilities based on the Bayes rule formula used to solve classification problems on numeric or nominal data features [4]. Naïve Bayes is used because it is known to have an efficient classification algorithm but has high accuracy. In research [5] After conducting the necessary analysis, it has been established that utilizing the Naïve Bayes technique for categorizing opinions on the Halodoc application achieves an accuracy level of 81.68%.

In research [9] discusses the Implementation of Particle Swarm Optimization (PSO) on Sentiment Analysis of Halodoc Application Reviews Using the Naïve Bayes Algorithm. In this research using the PSO feature selection method to increase accuracy with the Naïve Bayes algorithm. Naïve Bayes accuracy with n-gram features before combining feature selection gets 88.50% results with AUC 0.535. then accuracy increases after being optimized with PSO the accuracy value is 90.50% with AUC 0.525. The increase in accuracy reached 2%. The

Submitted: **09/06/2023**; Accepted: **14/07/2023**; Published: **23/07/2023**

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1070-1078
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6311

utilization of the PSO-based Naïve Bayes approach for categorizing reviews in the halodoc application has been demonstrated to yield significantly improved accuracy outcomes.

In the journal [10] discusses the application of sentiment classification on smartphone reviews on web-based online buying and selling sites using Naïve Bayes with TF-IDF. In this journal, the application strength test results are determined by precision, accuracy and recall tests. Based on these three tests, the success rate of the application can be measured in relation to the algorithm used. The application's performance yielded outstanding outcomes, boasting an impressive accuracy rate of 93%. The accuracy measurement outcome of the neutral class get the highest score of 100%, positive class 86.6%, negative class 93.3%. While the recall results of the neutral category reached 93.7%, the positive category 92.8% and the negative category 93.3%. The results, where precision and recall exceed 85%, provide strong evidence that the Naive Bayes algorithm is highly effective in this specific application.

Research [11] discusses Sentiment Analysis of the Effect of the Combination of TF-IDF and Lexicon Feature Extraction on Movie Reviews Using the KNN Method. In this study, when testing a combination using TF-IDF feature extraction with SentiWordnet Lexicon, it was found that the accuracy value of the combination of the two selection features was not better than TF-IDF feature extraction, which was 73.31%, while using TF-IDF alone had 81.04%. This test shows that using TF-IDF feature alone has better accuracy.

Journal [12] discusses Sentiment Analysis of Independent Campus Policies Using Naïve Bayes and TF-IDF Weighting Based on comments on Youtube. In this study, there are five main processes including manual labeling, text preprocessing, TF-IDF weighting, data validation using k-fold cross validation, and classification. Manual labeling is done by three processes, namely using text preprocessing, TF-IDF weighting, and classification itself. This test shows that utilizing the characteristics derived from tf-idf affects the classification outcome. In the process without being treated with TF-IDF, the accuracy is 91% while the process using TF-IDF produces an accuracy of 96%. Based on these findings, it can be inferred that employing TF-IDF weighting has a positive impact on enhancing the accuracy value.

Research [13] focuses on analyzing the sentiment of Mola Application Reviews obtained from the Google Play Store. The study utilizes the Support Vector Machine algorithm for this objective. The aim of this study is to employ the Support Vector Machine technique to categorize user feedback on the mola application. After conducting research using three different scenarios, it was observed that the most effective data split involved 90% for training and 10% for testing. This split yielded impressive performance metrics, including an accuracy rate of 92.31%, precision of 96.3%, recall of 89.66%, and an f1-score of 92.86%.

2. RESEARCH METHODOLOGY

The focus of this research is to develop a sentiment analysis model for analyzing user reviews of the Practo app. The chosen approach involves utilizing the Naïve Bayes algorithm in conjunction with the TF-IDF feature extraction. The diagram outlining the process of this model is depicted in Figure 1:

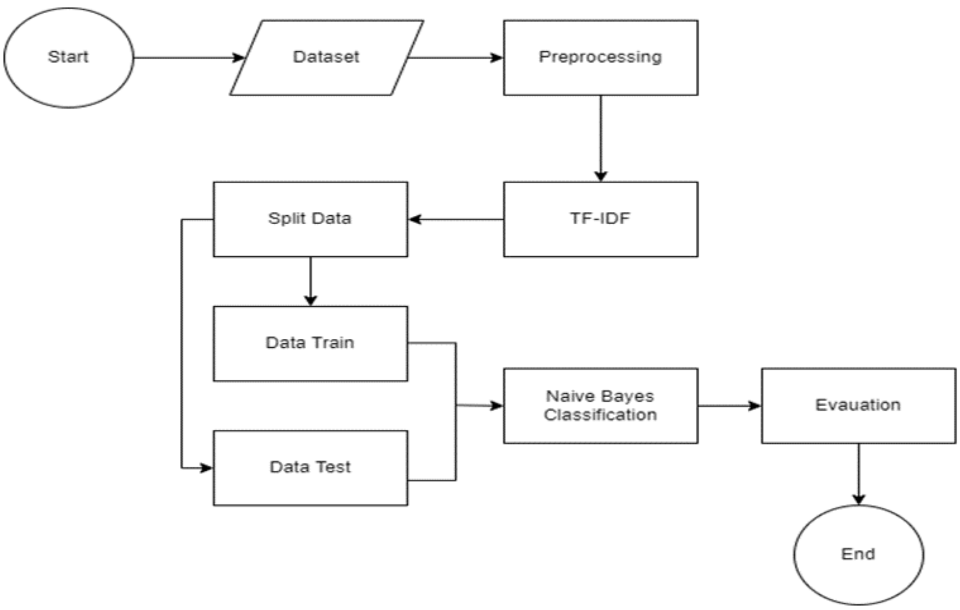

Figure 1. System Design

2.1 Dataset

This research utilizes a dataset obtained from the analysis of user reviews related to the Practo application, which were: (https://www.kaggle.com/datasets/manojgowda65/practo-android-appstore-reviews?resource=download).

Submitted: 09/06/2023; Accepted: 14/07/2023; Published: 23/07/2023

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1070-1078
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6311

This study analyzes a collection of 7,156 English-language reviews from the Practo application reviews. The data will pass the classification stage which previously performed positive and negative class labeling.

Table 1. Research Dataset

Label	Sentence
Fresh (Positive)	"Amazing app if doctor does not replies to your massges during consultation period of five days you get refunded.. superb!! Great respect."
Rotten (Negative)	"Atomic blonde was falls between james bond and john wick, carrying the violent spy chase of the former while increading the violence level, but not enough reaching that of the latter."

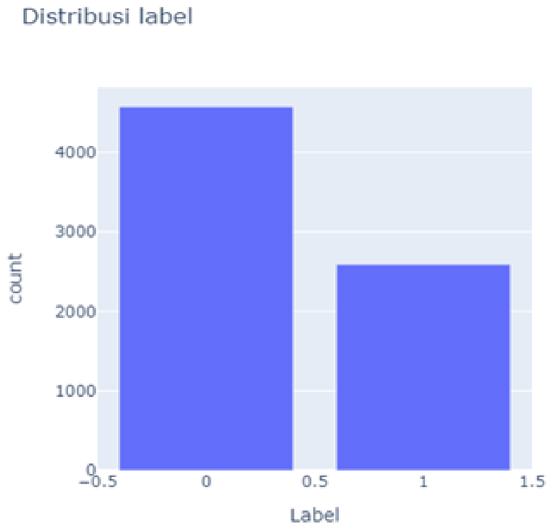

Figure 2. Label Distribution

2.2 Preprocessing

After all the data has been collected and prepared, the next step is preprocessing. The preprocessing process is carried out to fix problems that arise in data processing. In this research, the preprocessing process used is Cleansing, Case Folding, Tokenization, Stopword Removal, and Stemming. During the preprocessing process, data type conversion is also carried out on the "review_type" column. This data type change includes converting review values from 1 to 3 into 0, while review values from 4 to 5 are converted into 1.

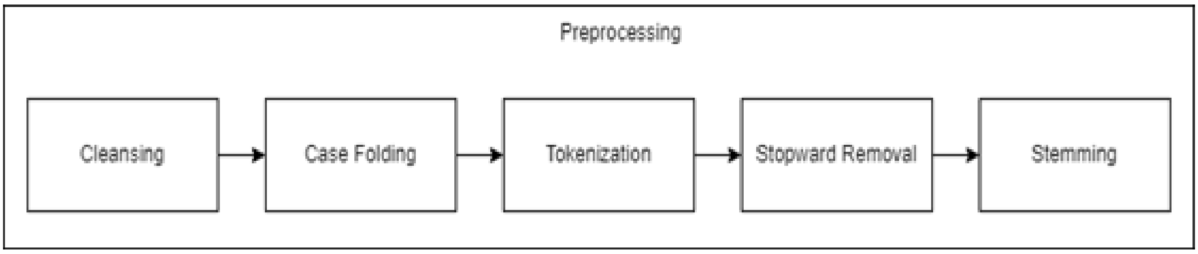

Figure 3. Preprocessing Stage

2.2.1 Cleansing

Cleansing is a process to remove or clean elements that have no influence on the classification process, such as punctuation and irrelevant words [11].

Table 2. Cleansing Data

Sentence	Cleansing Result
"It is a very useful app, but never loads the existing appointments page. This blocks me from rescheduling appointments. The app developers should fix this. It	"It is a very useful app but never loads existing appointments page This blocks me from rescheduling appointments The app developers should fix this It has

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1070-1078
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6311

Sentence	Cleansing Result
has been an issue for ages now. Edit: Prompt response from the support team and issue now resolved. 5 stars"	been an issue for ages now Edit Prompt response from the support team and issue now resolved stars".

2.2.2 Case Folding

Case folding is a data processing process that aims to convert or remove all capital letters in a document to lowercase [14]. This is done to facilitate searching since not all text documents are consistent in their use of uppercase letters. It promotes consistency and ease of analysis by standardizing the letter casing throughout the document.

Table 3. Case Folding Data

Cleansing Result	Case folding Result
"It is a very useful app but never loads existing appointments page This blocks me from rescheduling appointments The app developers should fix this It has been an issue for ages now edit Prompt response from the support team and issue now resolved stars."	""it is a very useful app but never loads existing appointments page this blocks me from rescheduling appointments the app developers should fix this it has been an issue for ages now edit prompt response from the support team and issue now resolved stars".

2.2.3 Tokenization

Tokenization is the collection of all words and removal of punctuation and non-letter symbols that do not affect the classification process [15].

Table 4. Tokenization Data

Case folding Result	Tokenization Result
"it is a very useful app but never loads existing appointments page this blocks me from rescheduling appointments the app developers should fix this it has been an issue for ages now edit prompt response from the support team and issue now resolved stars"	"[useful], [app], [never], [loads], [existing], [appointments], [page], [blocks], [rescheduling], [appointments], [app], [developers], [fix], [issue], [ages], [edit], [prompt], [response], [support], [team], [issue], [resolved], [stars]" to [use], [app], [never], [load], [exist], [appoint], [page], [block], [reschedule], [appoint], [app], [develop], [fix], [issue], [age], [edit], [prompt], [response], [support], [team], [issue], [resolv], [star]".

2.2.4 Stopword Removal

Stopword removal involves the elimination of insignificant words, while preserving the important ones, in order To minimize the amount of words stored in the token list. This process is performed as a preliminary step before further processing of the text [16].

Table 5. Stopword Removal Data

Tokenization Result	Stopword Removal Result
"[it], [is], [a], [very], [useful], [app], [but], [never], [loads], [existing], [appointments], [page], [this], [blocks], [me], [from], [rescheduling], [appointments], [the], [app], [developers], [should], [fix], [this], [it], [has], [been], [an], [issue], [for], [ages], [now], [edit], [prompt], [response], [from], [the], [support], [team], [and], [issue], [is], [now], [resolved], [stars]"	"[useful], [app], [never], [loads], [existing], [appointments], [page], [blocks], [rescheduling], [appointments], [app], [developers], [fix], [issue], [ages], [edit], [prompt], [response], [support], [team], [issue], [resolved], [stars]".

2.2.5 Stemming

Stemming is a method of transforming words into their base form by removing affixes in the words of a document or converting verbs into nouns, resulting in these words appearing in their base form. The stem (root word) is the core word after removing prefixes and suffixes [17].

Table 6. Stemming Data

Stopword Removal Result	Stemming Removal Result
"[useful], [app], [never], [loads], [existing], [appointments], [page], [blocks], [rescheduling], [appointments], [app], [developers], [fix], [issue],	[use], [app], [never], [load], [exist], [appoint], [page], [block], [reschedule], [appoint], [app], [develop], [fix], [issue], [age], [edit], [prompt], [response], [support], [team], [issue], [resolv], [star]".

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1070-1078
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6311

Stopword Removal Result	Stemming Removal Result
[ages], [edit], [prompt], [response], [support], [team], [issue], [resolved], [stars]"	

2.4 TF-IDF

TF-IDF (Term Frequency - Inverse Document Frequency) feature extraction is a commonly employed method for examining the association between sentences and a collection of documents. It is primarily used to determine the significance of individual words by calculating their respective weights [18]. In this study, the TF-IDF Feature Extraction method is utilized to transform words into numerical data. There are several variants of TF-IDF, namely Unigram and Bigram. The TF-IDF calculation in this study is expressed in the following equation:

$$Wt = TF_{t,d} \times IDF_t = TF_{t,d} \times log\frac{n}{DF_t} \quad (1)$$

Description:
Wt : The value representing the weight of a term (t) in a document (d).
TF_(t,d) : The frequency of the term (t) within the document (d).
IDF_t : The inverse frequency of the term (t) across the entire document corpus.
n : The overall count of documents present in the corpus
DF_t : Count of documents within the corpus that contain the term (t).

In this research, the TF-IDF Unigram model is one of the weighting methods included in the n-gram concept, where the focus is on using one word as a unit of analysis [19]. Bigram is one of the commonly used n-gram concepts in natural language processing (NLP) for weighting methods. The concept used by bigrams in this research is to sort pairs of contiguous words of length n from two words [19]. The flowchart of TF-IDF feature extraction in this research is expressed in the following figure:

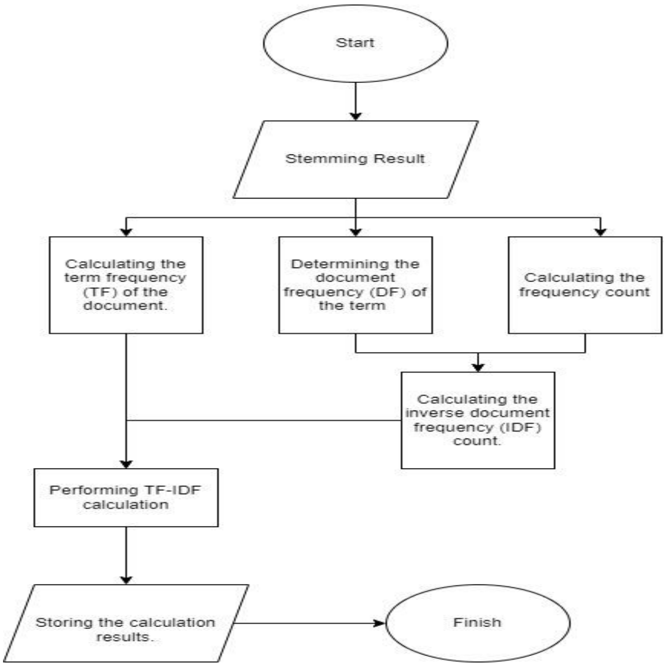

Figure 5. TF-IDF Flowchart

2.3 Split Data

After doing the preprocessing stage, the next stage is split data. The data in this study is divided into two parts, where 80% is allocated for training purposes, and the remaining 20% for testing purposes. The outcomes of the data split are presented in Table 2 provided below:

Table 2. Split Data Result

Split Data	Data Train	Data test
Total Data	5724	1432
Sentiment (Positive Label)	2062	523
Sentiment (Negative Label)	3662	909

2.5 Naïve Bayes Classification

After completing data processing and TF-IDF feature extraction, the subsequent task involves classification. The chosen approach for classification is Naïve Bayes, which is a statistical analysis algorithm utilizing Bayesian

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1070-1078
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6311

probabilities to process numerical data [20]. various variations of Naïve Bayes are available, such as Gaussian and Multinomial. Gaussian Naïve Bayes is a commonly employed variation of the Naïve Bayes method. The Gaussian method is employed to derive outcomes from training data or evaluate model performance by considering the probability value of the training data [21]. Multinomial Naïve Bayes is a specific variation of the Naïve Bayes method. In the context of text analysis, In Multinomial Naïve Bayes, it is assumed that the existence or occurrence presence of words in documents is independent of their order or contextual information [22]. This method enables the counting of word occurrences within each document, allowing the computation of document class frequencies based not only on the presence of words but also on the frequency of word occurrences in every individual document. In a general overview, the Naïve Bayes classification process can be visualized through the following equation:

$$P(c|x) = P(x|c) \times \frac{P(c)}{P(x)} \tag{2}$$

Description:

P(c|x) : The probability of class (c) given an observation (x). this is the posterior probability that we want to estimate

P(x|c) : The probability of observation (x) given class (c). It is a likelihood probability that represents the probability of observing a particular data point (x) assuming the class (c) is true.

P(c) : The probability of class (c). This is a prior probability that represents the overall probability of a random observation belonging to class (c) before considering specific features.

P(x) : Probability of observation (x). It is the probability of observing a particular data point (x) regardless of class.

The flowchart of the Naive Bayes classifier in this research is expressed in the following figure:

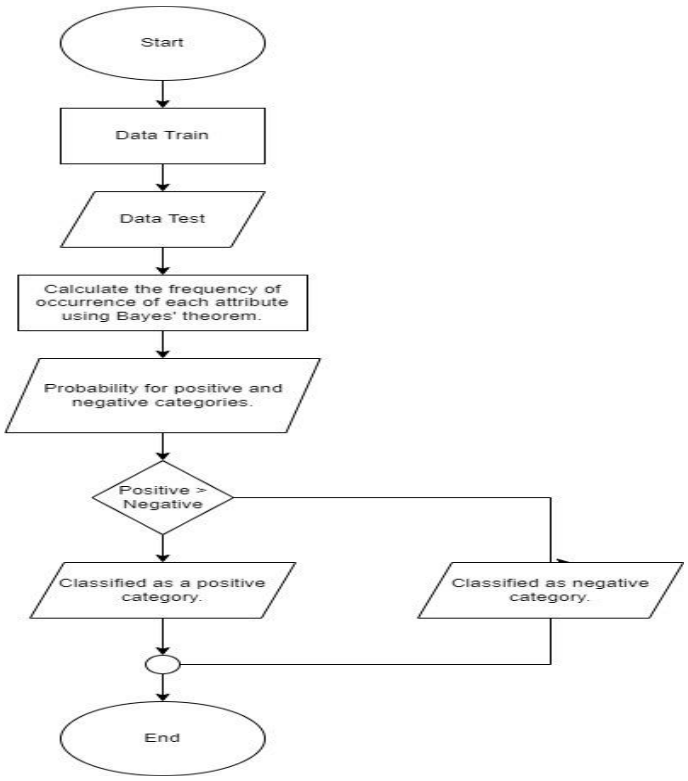

Figure 6. Naïve Bayes Flowchart

2.6 Evaluation

The evaluation stage in this research is carried out using a confusion matrix. The confusion matrix is utilized to evaluate the classification outcomes by comparing the actual classes and predicted classes [23]. The evaluation of the classification system's performance is typically conducted using the data matrix. Presented below is a tabular representation of the confusion matrix:

Table 3. Confusion Matrix

Confusion Matrix		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP	FP
	Negative	FN	TN

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1070-1078
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6311

Description:
TP = The count of instances categorized as positive data records that have been classified as positive values
FN = The count of instances categorized as positive data records that have been classified as negative values
FP = The count of instances categorized as negative data records that have been classified as positive values
TN = The count of instances categorized as negative data records that have been classified as negative values

The value produced by the Confusion Matrix method is in the form of an evaluation as follows:

The F1-Score represents the similarity or proximity between the predicted values generated by the application and the actual values of the real data [10]. The formula of f1-score, as follows:

$$F1 - Score = \frac{TP + TN}{TP + TN + FN + FP} \tag{3}$$

Precision refers to the degree of accuracy in matching the information requested by the user with the answers provided by the system [10]. The formula of precision, as follows:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Recall measures the effectiveness of the system in retrieving information successfully [10]. The formula of recall, as follows:

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

3. RESULT AND DISCUSSION

In this study, The system that has been built is tested for success by using F1-score as a reference. This system goes through several stages, starting with the Preprocessing stage which includes Cleansing, Case folding, Tokenization, Stopword Removal, and Stemming. The dataset obtained of 7156 data, which were subsequently split into 5724 training data and 1432 test data using an 80:20 division ratio. Furthermore, this data is fed into the TFI-IDF feature extraction. At this stage, a comparison is made between the use of unigram variants and bigram variants in TF-IDF extraction features. For this research, the classification model employed is a comparison between Gaussian Naïve Bayes and Multinomial Naïve Bayes. Three different test scenarios were carried out. In the first scenario, the comparison is made between using stemming and not using stemming in the Preprocessing stage. In the second scenario, the aim is to evaluate the impact of the number of N in TF-IDF feature selection on the performance of Naïve Bayes. This is accomplished by comparing the use of unigram and bigram as the parameter for the number of N in TF-IDF. The third scenario involves comparing different variants of the Naïve Bayes method, specifically Gaussian Naïve Bayes and Multinomial Naïve Bayes.

3.1 The Effect of Stemming

In the first scenario of the experiment, an assessment is conducted to examine how the application of stemming impacts the effectiveness of the system model that has been developed. In this test, the effectiveness of stemming is assessed by comparing its usage during the preprocessing stage with and without stemming. The evaluation is carried out using data extracted through TF-IDF with unigram variations and employing the Multinomial Naïve Bayes method. The outcomes of the evaluation conducted in scenario 1 are presented in the table below:

Table 4. The Effect of Stemming

Preprocessing	Precision	Recall	F1-Score
With Stemming	**96.62%**	**76.67%**	**85.50%**
Without Stemming	90.43%	85.31%	**84.15%**

According to the test results presented in Table 4, it is evident that employing stemming during the preprocessing stage leads to superior performance in terms of F1-score, Precision, and Recall compared to not utilizing stemming in the preprocessing stage. The test resulted in an F1-score of 85.50%, precision is 96.62% and recall is 76.67%. The increase in performance by using stemming is because stemming helps eliminate word variants that have the same root word. By combining these words into their basic form, it can reduce duplication in the data. As a result, the number of features generated by TF-IDF feature extraction can be reduced, reducing the complexity and dimension of the feature space. This can improve model performance by preventing overfitting and improving generalization ability.

3.2 The Effect of of the number N in TF-IDF feature selection on the performance of Naïve Bayes

In scenario 2, testing was carried out to assess the impact of the number of N in TF-IDF feature selection on Naïve Bayes performance. In this scenario, the comparison is made between unigram and bigram using data that has

Volume 7, Nomor 3, Juli 2023, Page 1070-1078
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6311

gone through stemming process and Multinomial Naïve Bayes method as a parameter of the number of N in TF-IDF. The outcomes of the evaluation conducted in scenario 2 are presented in the table presented beneath:

Table 5. The Effect of the number N in TF-IDF feature selection on the performance of Naïve Bayes

Word2Vec Dimension	Precision	Recall	F1-Score
TF-IDF Unigram	**96.50%**	**76.67%**	**85.50%**
TF-IDF Bigram	90.15%	76.81%	85.04%

According to the test results presented in Table 5, it is evident that employing the unigram variant of Naive Bayes yields superior and more consistent performance compared to using the bigram variant.. there is a significant difference in the F1-score value between the two methods, with a difference of 0.46%. The result of the Unigram variant is greater because unigram focuses on the frequency of occurrence of individual words, while the stemming process reduces the variation of words thus increasing the relevance of words in the dataset. In addition, unigrams have a lower number of feature space dimensions than bigrams. So it is proven that the number of N affects the selection of TF-IDF features on Naïve Bayes performance.

3.3 Naïve Bayes Variant comparison

In scenario 3, tests were conducted to evaluate the performance of sentiment analysis of Practo application reviews using the Naïve Bayes method. This test was conducted to test the comparison between the use of Naïve Bayes method variants with data that had been processed using stemming, and the use of TF-IDF method with unigram variations. In this test, two variants of Naïve Bayes method were used, namely Gaussian Naïve Bayes and Multinomial Naïve Bayes. The test results for scenario 3 are presented in the table below:

Table 6. Naïve Bayes Comparison

Classifier	Precision	Recall	F1-Score
Gaussian Naïve Bayes	50.36%	80.22%	61.84%
Multinomial Naïve Bayes	**96.62%**	**76.67%**	**85.50%**

Based on the analysis in table 6 above using the Naïve Bayes method, it can be concluded that the Multinomial Naïve Bayes method has better performance than the Gaussian Naïve Bayes method. There is a significant difference in the F1-score value between the two methods, with a difference of 19.66%. This difference is due to the different basic principles used by each method. Naïve Bayes Multinomial treats each attribute (word) within a class as independent of one another, without taking into account the interdependencies among the attributes. This method discretely computes the occurrence or frequency of each word's occurrence in individual documents, making it well-suited for modeling text data based on word frequencies. By analyzing the frequency of occurrence of words relevant to the class being analyzed, this gives Multinomial Naïve Bayes better performance. The F1-score resulting from using the Multinomial Naïve Bayes Method is 85.50%.

4. CONCLUSION

Based on the performed testing and analysis in the research regarding Sentiment Analysis of Practo Application Reviews using Naïve Bayes with TF-IDF, the following conclusions were derived. The stemming process at the preprocessing stage has an influence on the performance results obtained. Comparing the performance results of testing scenario 1, it is observed that utilizing stemming on the data yields superior outcomes compared to data that has not undergone the stemming process. The performance outcomes are influenced by various variations in the construction of TF-IDF extraction features. In scenario 2, TF-IDF constructed with unigram variant paramater gives better accuracy than the model built with bigram variant and it is proven that the number of N has an effect in TF-IDF feature selection on Naïve Bayes performance. The selection of Naïve Bayes method variants has an impact on the results, with Multinomial Naïve Bayes being demonstrated to offer the most optimal performance. Among the three scenarios, the most successful model achieved an F1-Score of 85.50%. This model utilized stemming during the preprocessing stage, TF-IDF unigram variants, and the Multinomial Naïve Bayes classification method. Suggestions for future research are to consider the use of other parameters in the construction of TF-IDF extraction features apart from unigrams and bigrams. For example, exploring the use of trigrams or more complex word clustering schemes can be done to enhance the model's performance. In addition, it is also recommended to perform data normalization before the analysis process to ensure the uniformity and consistency of the data used.

REFERENCES

[1] J. Nurvania, Jondri, and K. Muslim Lhaksamana, “Analisis Sentimen Pada Ulasan di TripAdvisor Menggunakan Metode Long Short-Term Memory (LSTM),” e-Proceeding Eng., vol. 8, no. 4, pp. 4124–4135, 2021.

[2] N. R. Wardani and A. Erfina, “Konsultasi Dokter Menggunakan Algoritma Naive,” SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform., pp. 11–18, 2021.

Submitted: 09/06/2023; Accepted: 14/07/2023; Published: 23/07/2023

[3] M. F. Maulana, L. Ramadani, and F. M. Al-Anshary, "Pengembangan Sistem Telemedicine Berbasis Aplikasi Mobile Menggunakan Metode Iterative Dan Incremental Development of a Telemedicine System Based on Mobile Applications Using Itrative and Incremental Methods," e-Proceeding Eng., vol. 8, no. 5, pp. 9475–9487, 2021.
[4] R. N. CIKANIA, "Implementasi Algoritma Naïve Bayes Classifier Dan Support Vector Machine Pada Klasifikasi Sentimen Review Layanan Telemedicine Halodoc," Jambura J. Probab. Stat., vol. 2, no. 2, pp. 96–104, 2021, doi: 10.34312/jjps.v2i2.11364.
[5] A. Hendra and F. Fitriyani, "Analisis Sentimen Review Halodoc Menggunakan Nai ̈ve Bayes Classifier," JISKA (Jurnal Inform. Sunan Kalijaga), vol. 6, no. 2, pp. 78–89, 2021, doi: 10.14421/jiska.2021.6.2.78-89.
[6] M. B. Hamzah, "Classification of Movie Review Sentiment Analysis Using Chi-Square and Multinomial Naïve Bayes with Adaptive Boosting," J. Adv. Inf. Syst. Technol., vol. 3, no. 1, pp. 67–74, 2021, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/jaist
[7] M. T. Razaq, D. Nurjanah, and H. Nurrahmi, "Analisis Sentimen Review Film menggunakan Naive Bayes Classifier dengan fitur Tf-Idf," vol. 9, no. 5, pp. 6053–6071, 2022.
[8] S. Adji Pratomo, S. Al Faraby, and M. Dwifebri Purbolaksono, "Analisis Sentimen pada Ulasan Film dengan Kombinasi Seleksi Fitur Chi-Square dan TF-IDF menggunakan Metode KNN," e-Proceeding Eng., vol. 8, no. 5, pp. 10116–10126, 2021.
[9] N. Nuris, E. R. Yulia, and K. Solecha, "Implementasi Particle Swarm Optimization (PSO) Pada Analysis Sentiment Review Aplikasi Halodoc Menggunakan Algoritma Naïve Bayes," J. Teknol. Inf., vol. 7, no. 1, pp. 17–23, 2021, doi: 10.52643/jti.v7i1.1330.
[10] R. Karisma, S. Lestanti, and M. T. Chulkamdi, "Aplikasi Klasifikasi Sentimen Pada Ulasan Smartphone Di Situs Jual Beli Online Berbasis Web Menggunakan Naive Bayes Dengan Tf-Idf," JATI (Jurnal Mhs. Tek. Inform., vol. 6, no. 1, pp. 31–37, 2021, doi: 10.36040/jati.v6i1.4365.
[11] S. A. Pratomo, S. Al Faraby, and M. D. Purbolaksono, "Analisis Sentimen Pengaruh Kombinasi Ekstraksi Fitur TF-IDF dan Lexicon Pada Ulasan Film Menggunakan Metode KNN," eProceedings Eng., vol. 8, no. 5, pp. 10116–10126, 2021.
[12] D. F. Zhafira, B. Rahayudi, and I. Indriati, "Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube," J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf., vol. 2, no. 1, pp. 55–63, 2021, doi: 10.25126/justsi.v2i1.24.
[13] M. D. Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine," INTECOMS J. Inf. Technol. Comput. Sci., vol. 5, no. 1, pp. 1–7, 2022, doi: 10.31539/intecoms.v5i1.3708.
[14] Merinda Lestandy, Abdurrahim Abdurrahim, and Lailis Syafa'ah, "Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes," J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 5, no. 4, pp. 802–808, 2021, doi: 10.29207/resti.v5i4.3308.
[15] S. Ernawati and R. Wati, "Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Review Agen Travel," J. Khatulistiwa Inform., vol. 6, no. 1, pp. 64–69, 2018.
[16] J. A. Septian, T. M. Fachrudin, and A. Nugroho, "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor," J. Intell. Syst. Comput., vol. 1, no. 1, pp. 43–49, 2019, doi: 10.52985/insyst.v1i1.36.
[17] A. V. Sudiantoro et al., "Analisis Sentimen Twitter Menggunakan Text Mining Dengan," vol. 10, no. 2, pp. 398–401, 2018.
[18] H. Nurrun Muchammad Shiddieqy, S. Paulus Insap, and W. Wing Wahyu, "Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter," Semin. Nas. Teknol. Inf. dan Komun., vol. 2016, no. March, pp. 57–64, 2016.
[19] A. Munir, E. P. Atika, and A. D. Indraswari, "Analisis Sentimen pada review hotel menggunakan metode pembobotan dan klasifikasi," Jnanaloka, vol. 3, no. 1, pp. 33–38, 2022, doi: 10.36802/jnanaloka.2022.v3-no1-33-38.
[20] D. A. Nugroho et al., "Analisis Sentimen Data Presiden Jokowi Dengan Preprocessing Normalisasi Dan Stemming Menggunakan Metode Naive Bayes Dan SVM," Semin. Nas. Teknol. Fak. Tek. Univ. Krisnadwipayana, vol. 3, no. 1, pp. 1–11, 2021, [Online]. Available: http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/4793%0Ahttps://jurnal.teknikunkris.ac.id/index.php/semnastek2019/article/view/343/342
[21] K. S. Sara, H. Dan, R. Pada, F. E. Purwiantono, and A. Aditya, "POSTINGAN MEDIA SOSIAL MENGGUNAKAN ALGORITMA NAIVE," vol. 14, no. 2, pp. 68–73, 2020.
[22] D. E. Kurniawan, D. Swanjaya, and ..., "Penerapan Metode Naïve Bayes Pada Aplikasi Ayo Playon," Pros. SEMNAS …, pp. 167–171, 2022.
[23] I. Prayoga and M. D. P, "Sentiment Analysis on Indonesian Movie Review Using KNN Method With the Implementation of Chi-Square Feature Selection," J. Media Inform. Budidarma, vol. 7, pp. 369–375, 2023, doi: 10.30865/mib.v7i1.5522.