Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

Sentiment Analysis of Maxim Online Transportation App Reviews using Support Vector Machine (SVM) Algorithm

Putri Kurniawati*, Riska Yanu Fa'rifah, Deden Witarsyah

School of Industrial Engineering, Information System, Telkom University, Bandung, Indonesia
Email: [1,*]putrikurniaa@student.telkomuniversity.ac.id, [2]riskayanu@telkomuniversity.ac.id, [3]dedenw@telkomuniversity.ac.id
Correspondence Author Email: putrikurniaa@student.telkomuniversity.ac.id
Submitted: **10/09/2023**; Accepted: **25/09/2023**; Published**: 27/09/2023**

Abstract−The continuous emergence of online transportation service platforms is one of the effects of the ever-increasing technological advancements. One such online transportation service application, Maxim, has recently been slowly gaining ground in the ride-hailing market in Indonesia. According to data collected by one media outlet in 2022, Maxim ranks third as the most preferred online transportation platform by the public, following Gojek and Grab. This suggests that there are factors causing users to lack interest in or hesitate to use the Maxim application. On the Google Play Store, user ratings (in numerical values) and written reviews serve as reasons for the potential users lack of interest. Analyzing ratings alone is less accurate and does not provide in-depth information and meaning regarding users experiences. To understand user opinions about Maxim's service and functionality, an analysis of user reviews is crucial. Therefore, this research conducts sentiment analysis on Maxim user reviews using the Support Vector Machine (SVM) algorithm to classify reviews quickly. The reviews are categorized into two classes: positive and negative sentiment. The classification process is carried out in three scenarios with different data training and testing ratios: 60:40, 70:30, and 80:20, using a Linear kernel and hyperparameter optimization with GridSearch. The best accuracy is achieved with a 70:30 ratio, which is 89.82%. Evaluation using the confusion matrix also yields a precision of 92.66%, recall of 94.09%, and an F1 score of 93.38%. The ROC-AUC curve evaluation results in an AUC value of 0.8505. The sentiment analysis results tend to lean towards positive sentiment, indicating a high level of user satisfaction with the Maxim application. Based on these sentiment results, developers can identify what aspects of the Maxim application need to be maintained and improved.

Keywords: Maxim; Online Transportation; Sentiment Analysis; Classification; Support Vector Machine

1. INTRODUCTION

The continuous advancement of technology has had a significant impact, leading to more and more young generations continually innovating to simplify daily activities. One such innovation is the emergence of online transportation service providers. Online transportation has become an integral part of daily life in Indonesia. According to a survey conducted by the Association of Indonesian Internet Service Providers (Asosiasi Penyelenggara Jasa Internet Indonesia or APJII), online transportation ranks 16th among the 22 reasons for using the internet in Indonesia. This indicates that using online transportation services has become a habit for Indonesian society [1]. Online transportation service providers in Indonesia are currently in fierce competition. The Chairman of the Online Driver Association (Asosiasi Driver Online or ADO) revealed that in recent years, more than 10 online transportation companies have been forced to shut down because they couldn't capture and dominate the ride-hailing market in Indonesia. Currently, Gojek, Grab, and Maxim are the top online transportation services based on the number of users [2][3].

Maxim is one of the online transportation service providers operating in Indonesia since 2018. In addition to providing pick-up and delivery services, Maxim also offers food delivery and freight transport services. Maxim has consistently improved its technology and business processes with the main mission of enhancing user interaction and assisting many people in reaching their destinations [4].

According to a survey conducted by the Ministry of Transportation's Research and Development Agency (Balitbang Kementerian Perhubungan), the most widely used online transportation service in the Jabodetabek area is Gojek, with 59.13% of the total respondents using it. In second place is Grab, with 32.24% of users, while Maxim ranks third with 6.93% of users [5]. Therefore, it can be concluded that there may still be areas for improvement. One way to assess Maxim's performance and quality is by understanding user experiences, which can be obtained from reviews and comments on the Google Play Store. User reviews can reflect their views on the application experience, including service quality, user-friendliness, and more.

As Maxim's popularity grows, the number of users and feedback on Google Play Store has also increased. Based on data recorded on Google Play Store, the Maxim application has an overall rating of 4.7 out of 5 and has received 2.85 million product reviews to date. Achieving high ratings and reviews is crucial for assessing an application. New users often consider ratings and reviews as references and considerations before using an app. However, rating acquisition is not always accurate and may not provide in-depth information about user experiences. Rating analysis is easy to understand and aids in decision-making by looking at the average ratings provided by users. However, rating analysis has some drawbacks, such as the inability to capture complex emotions that may be present in textual reviews and susceptibility to bias and manipulation, such as ratings without explanatory reviews and mismatched reviews and ratings.

Another way to assess Maxim's performance and quality is by understanding user experiences through reviews and comments on the Google Play Store. User reviews can reflect their experiences with the app, including service quality, ease of use, and more [6]. Therefore, sentiment analysis of user reviews is necessary. Sentiment analysis

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

provides deeper insights compared to rating analysis alone. The advantage of sentiment analysis over rating analysis is its ability to identify the emotions and context conveyed in user-written text [7]. Sentiment analysis categorizes data into two classes: positive sentiment and negative sentiment, providing valuable insights from the text. With numerous user reviews, sentiment analysis offers a faster and more efficient way to process review data.

In this study, the Support Vector Machine (SVM) algorithm is used as the research method. Previous research has shown that the SVM algorithm outperforms other algorithms in review data processing. In previous studies, SVM achieved better accuracy than Naïve Bayes, with SVM accuracy at 77.00% and Naïve Bayes at 70.40%. The study also suggested that SVM could be used effectively for further research with the same data characteristics [8]. Furthermore, a study comparing sentiment analysis of review data between SVM and Decision Tree found that SVM outperformed with an accuracy of 90.20% compared to Decision Tree's 89.80% [9]. Another study comparing Naïve Bayes and SVM for sentiment analysis found that SVM achieved the best accuracy at 82.48%, while Naïve Bayes achieved 76.56% [10]. Additionally, a study conducting sentiment analysis with Naïve Bayes reported an accuracy of 74.37% and an AUC of 0.659, while SVM performed better with an accuracy of 81.22% and an AUC of 0.886 [11].

This research uses the Support Vector Machine (SVM) algorithm for text classification into positive and negative sentiments. Proper preprocessing, processing, and evaluation steps are followed to obtain more accurate results. The goal of this research is to implement the Support Vector Machine (SVM) algorithm for sentiment analysis and to compare sentiment results for the Maxim application. The results of this research, which include positive and negative sentiment results, can serve as an evaluation and reference for improving Maxim's performance and services.

2. RESEARCH METHODOLOGY

2.1 Research Stages

The research conducted focuses on the user environment of the Maxim online transportation service application, where users provide opinions in the form of comments or reviews about the Maxim application on Google Play Store. The method used in this research is text data classification using Support Vector Machine (SVM). The problem-solving process in this research consists of four stages: Problem and Solution Identification Stage, Preprocessing Stage, Modeling Stage, and Conclusion Stage.

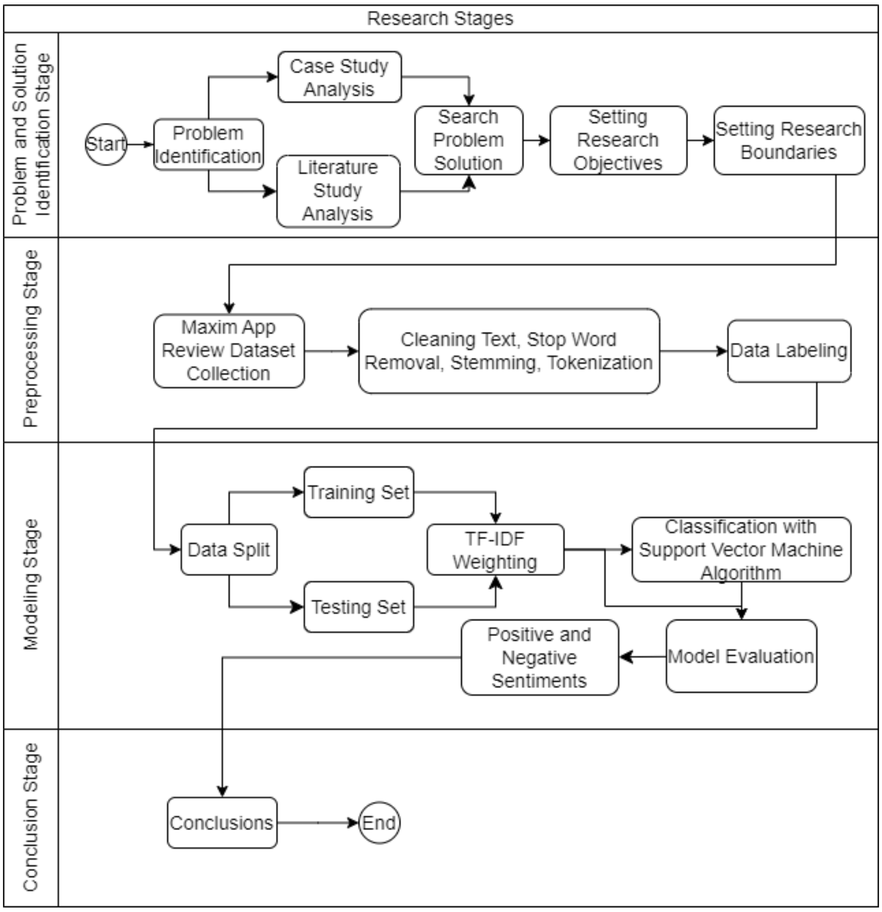

Figure 1. Research Stages

The Figure 1 above represents the problem-solving framework consisting of four stages as follows.

Copyright © 2023 **Putri Kurniawati**, Page 467
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

2.1.1 Problem Identification and Solution Stage

In this stage, problem identification is carried out based on the analysis of the selected case study and a review of related literature relevant to the research case. This serves as a basis for determining problem solutions. Once the problem solution is determined, the next step is to establish the research objectives and define the scope of the study.

2.1.2 Preprocessing Stage

After data collection, several data cleaning steps are performed to remove unnecessary data. This includes text cleaning, tokenization, stop-word removal, and stemming. Following data cleaning, the next step involves manual data labeling, dividing the data into two labels: Positive and Negative.

2.1.3 Modeling Stage

The next stage is the Modeling Stage. After going through the Preprocessing Stage and data labeling, the data is split into training and testing sets. Text is then numerically encoded using TF-IDF (Term Frequency-Inverse Document Frequency) for both training and testing data. This research also implements hyperparameter optimization techniques to find the best combination of kernel and parameters that enhance model performance. Subsequently, the classification process is performed using the Support Vector Machine algorithm, and model performance is evaluated using the confusion matrix method and AUC (Area Under the Curve) values.

2.1.4 Conclusion Stage

After obtaining the results from sentiment classification, conclusions are drawn from the research findings, and recommendations are provided based on the outcomes of the study.

2.2 Text Mining

Text mining, also known as text data mining, is a technique used for extracting information and knowledge from textual data. This technique involves the analysis of text data that contains specific features or can be considered unstructured data, requiring specific preprocessing steps. Text mining is one of the information extraction techniques within the field of data mining [11]. Data mining, or data discovery, is a broader field of study or research that focuses on the collection, cleaning, processing, analysis, and discovery of insights from data. Data mining is applied in various problem-solving processes that involve structured data [12][11]. In summary, text mining is a subset of data mining that specifically deals with unstructured textual data, aiming to extract valuable information and knowledge from such data sources.

2.3 Sentiment Analysis

Sentiment analysis, also known as sentiment mining, is a field of study that involves the analysis of sentiments, opinions, evaluations, and emotions expressed by individuals when assessing products, services, topics, or other people. Sentiment analysis is often referred to as Opinion Mining [12]. Sentiment classification is typically formulated as a binary classification task, where sentiments are categorized into two classes: positive and negative [12]. In sentiment classification, several popular algorithms are used in research, such as Decision Trees, Naïve Bayes Classifiers, Support Vector Machines, K-Nearest Neighbor Classifiers, and more [13]. The benefits of sentiment analysis include the ability to quickly analyze a product and use it as a tool to understand how consumers or users respond to and perceive that product. Sentiment analysis is valuable for businesses and organizations as it allows them to gain insights into customer feedback, public opinion, and market trends, helping them make informed decisions and improve their products or services based on customer sentiment and feedback.

2.4 Support Vector Machines

Support Vector Machine (SVM) is a classification or categorization algorithm used in data mining for processing both linear and non-linear data [13]. SVM is classified as one of the supervised learning methods. The main goal of SVM is to find a separator, often referred to as a hyperplane, that can effectively separate two classes of data points [14][15]. Initially, SVM algorithms were designed for handling two-dimensional data (linear classifiers). However, they have been extended to handle high-dimensional data by employing a technique known as the kernel trick, which is commonly used to address non-linear data [14][16]. Commonly used kernel functions in SVM include the Linear kernel, Radial Basis Function (RBF) kernel, and Polynomial kernel. These kernels allow SVM to work with complex, non-linear data by mapping the data into higher-dimensional spaces where it can be linearly separated . SVM is a powerful and versatile algorithm widely used in various machine learning and data mining tasks, including classification and regression problems.

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

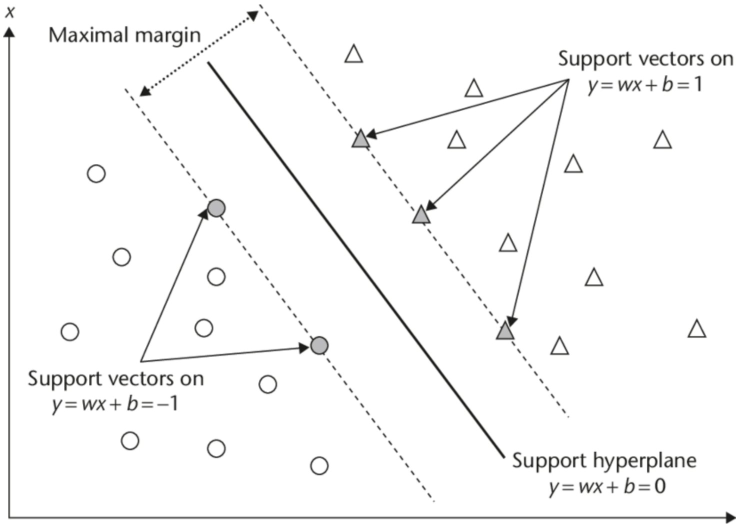

Figure 2. Optimal Separating Hyperplane

The kernel used in this research is the Linear kernel. The Linear kernel was chosen because it has several advantages, including computational efficiency, interpretability, low risk of overfitting, and high stability [17]. The selection of the Linear kernel is also based on previous research where a comparison was made between the Linear, Polynomial, and Sigmoid kernels. That research showed that the model with the Linear kernel achieved the highest accuracy [18].

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

TF-IDF, which stands for Term Frequency-Inverse Document Frequency, is a method used to measure the importance of a word in a document. In this research, TF-IDF calculations are performed to determine the weight value of each word in the dataset documents. TF-IDF weighting in this study utilizes one of the open-source Python programming language libraries, which is Scikit-learn or Sklearn. Sklearn has a text feature extraction module to transform or represent text data into numeric form, and one of its methods is TF-IDF Term Weighting. Here is the formula for calculating TF-IDF.

$$tf_{t,d} = frequency_{t,d} \quad (1)$$

$$idf_{t,d} = ln\left(\frac{1+N}{1+df_t}\right) + 1 \quad (2)$$

$$tf - idf_{t,d} = tf_{t,d} * idf_{t,d} \quad (3)$$

2.6 Confusion Matrix

Confusion matrix is an evaluation method used to obtain accuracy, recall, precision, and F1 score results from a classification process [19]. This method measures the performance of a classification model by displaying the number of data points that were classified correctly and incorrectly, along with their true labels. Confusion matrix with 4 combinations as a representation of different predicted and actual values as shown in Table 1 below.

Table 1. Confusion Matrix

Confusion Matrix		***Prediction***	
		Negative	*Positive*
Actual	*Negative*	TN	FP
	Positive	FN	TP

TP (True Positive) is the correctly predicted positive data; a high TP value indicates that the model has high sensitivity in predicting positive data. FN (False Negative) is positive data predicted as negative; a high FN value indicates that the model is less capable of correctly predicting actual positive data as negative. FP (False Positive) is negative data predicted as positive; a high FP value indicates that the model is less capable of correctly predicting actual negative data as positive. TN (True Negative) is the correctly predicted negative data; a high TN value indicates that the model can accurately identify and classify actual negative data. Based on the values obtained in the confusion matrix, model performance calculations can be made, including accuracy, recall, precision, and F1 score, using the following equations.

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

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

$$F1\ Score = \frac{2(Recall\ x\ Precision)}{Recall+Precision} \quad (7)$$

2.7 ROC-AUC (Receiver Operating Characteristic-Area Under the Curve)

Performance measurement or evaluation of a classification model can also be done using a ROC (Receiver Operating Characteristic) curve. The ROC curve depicts the relationship between TPR (True Positive Rate) and FPR (False Positive Rate). On the ROC curve, the area under the curve, known as AUC (Area Under Curve), can be determined. A high AUC value indicates that the model performs well in classification. Table 2 below represents the AUC criteria that indicate how well the classification model performs [13].

Table 2. AUC Score Criteria

AUC Score	Description
0.90 - 1.00	Sangat baik (*Excellent Classification*)
0.80 - 0.90	Baik (*Good Classification*)
0.70 - 0.80	Adil (*Fair Classification*)
0.60 - 0.70	Buruk (*Poor Classification*)
0.50 - 0.60	Gagal (*Failure*)

3. RESULT AND DISCUSSION

3.1 Data Collection

The data used in this research consists of reviews from users of the Maxim application, which were obtained from the Google Play Store through web scraping. The data collected covers reviews from January 12, 2023, to March 10, 2023. A total of 20,000 data were collected, and after the preprocessing stage, 18,242 data were used for classification. The following is an example of Maxim application reviews data as shown in Table 3.

Table 3. Review Dataset

No	Review Dataset
1.	harga bersahabat & pelayanan ok. (friendly price & ok service.)
2.	maps nya mohon diperbaharui donk pemetaan lokasi nya masih maps yg lama belum yg terbaru., trimakasih maxim #dpt_driver_yg_ramah ðŸ‘• (Please update the location mapping, it's still the old maps, not the latest. Thanks maxim #dpt_driver_yg_ramah ðŸ')
3.	akurasi GPS nya sangat kacau (GPS accuracy is very messed up)

Based on the Table 3 above, the review data obtained from the Google Play Store consists of various characters such as letters, punctuation marks, and specific symbols. Preprocessing will be performed on the review data to remove or clean the data of unused and meaningless attributes.

3.2 Data Preprocessing

The data collected through the web scraping process is raw data that includes not only letters and words but also characters, numbers, and specific symbols. The classification process will not run optimally if the data is not readable or identifiable by the system. Therefore, preprocessing is performed to ensure that the data can be processed effectively. Figure 3 below illustrates the data preprocessing flow in this research.

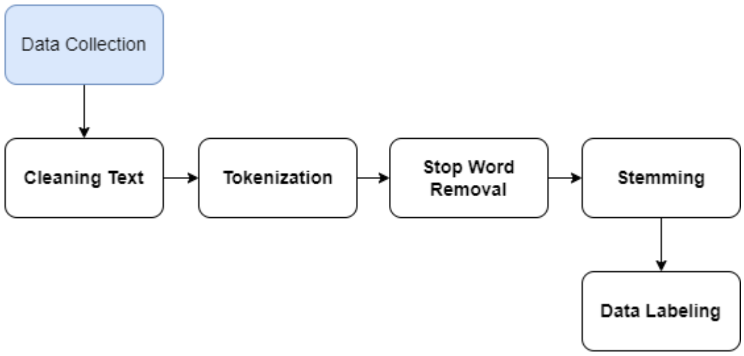

Figure 3. Data Preprocessing Stages

Copyright © 2023 **Putri Kurniawati**, Page 470
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

The preprocessing stages carried out include Cleaning Text to remove meaningless characters such as punctuation marks, numbers, specific symbols, and convert the text to lowercase, Tokenization to separate each word from a review sentence, Stop Word Removal to eliminate words that have no meaning, and Stemming to transform words with affixes into their base form. The following is an example of the preprocessing results as shown in Table 4.

Table 4. Preprocessing Stages Results

Preprocessing Stages	Results
Data	Layanan mantap terima kasih sudah mengantar saya sampai tujuan. (Great service, thank you for taking me to my destination.)
Cleaning Text	layanan mantap terima kasih sudah mengantar saya sampai tujuan (Great service, thank you for taking me to my destination)
Tokenization	['layanan', 'mantap', 'terima', 'kasih', 'sudah', 'mengantar', 'saya', 'sampai', 'tujuan'] (['service', 'steady', 'accept', 'love', 'already', 'delivered', 'me', 'arrived', 'destination'])
Stop Word Removal	['layanan', 'mantap', 'terima', 'kasih', 'mengantar', 'tujuan'] (['service', 'steady', 'accept', 'love', 'deliver', 'destination'])
Stemming	['layan', 'mantap', 'terima', 'kasih', 'antar', 'tuju'] (['serve', 'steady', 'receive', 'love', 'deliver', 'go'])

Table 4 shows that the initial data consists of reviews with punctuation marks and initial capitalization in sentences. During the Text Cleaning, punctuation marks are removed, and sentences are converted to lowercase. Then, in the Tokenization, words within sentences are separated for ease of processing. In the Stop Word Removal, meaningless words are eliminated as they do not play a role in the processing. Finally, in the Stemming, words containing affixes are converted to their base forms.

3.3 Data Labeling

In this stage, labels are assigned to the data based on the content and meaning of the review words. There are two label groups: positive and negative labels. In this context, a positive label represents a review or comment that contains indicators of satisfaction from Maxim's online transportation users expressed in various words or sentences. Examples of reviews falling into the positive label category include words of praise and expressions of gratitude. On the other hand, negative labels represent reviews that contain complaints or user dissatisfaction, and there are many reviews mentioning shortcomings in the Maxim application or services. An example of data labeling results, as found in Table 5, is as follows.

Table 5. Data Labelling Results

Review Dataset	Label
['maxim', 'jalan', 'mudah', 'terima', 'kasih', 'maxim', 'sukses'] (['maxim', 'way', 'easy', 'accept', 'love', 'maxim', 'success'])	POSITIF
['driver', 'nya', 'nakal', 'apk', 'ribu', 'suruh', 'bayar', 'ribu', 'alas', 'sesuai', 'aplikasi', 'pake', 'maxim', 'drivernya', 'nakal', 'curang'] (['driver', 'its', 'naughty', 'apk', 'thousand', 'tell', 'pay', 'thousand', 'base', 'appropriate', 'app', 'use', 'maxim', 'driver', 'naughty', 'cheat'])	NEGATIF
['layan', 'ramah', 'nyaman', 'kasih', 'bintang'] (['serve', 'friendly', 'cozy', 'love', 'star'])	POSITIF
['tolong', 'lokasi', 'kurang', 'antar', 'susah', 'akurat', 'tuju'] (['please', 'location', 'less', 'deliver', 'difficult', 'accurate', 'go'])	NEGATIF

In Table 5 above, there are examples of review data that have undergone preprocessing and manual labeling stages to prevent potential sentiment misinterpretation. For instance, the first data entry is labeled as positive because it contains positive sentiments, such as appreciation and expressions of gratitude from a user for the user-friendly Maxim service. The second data indicates that the user expressed disappointment in their experience with the Maxim application, as evidenced by words like 'nakal' (naughty) and 'curang' (cheat).

3.3 Modeling

The modeling stage, also known as the modeling phase, is the stage where a model is created to represent a system or problem that needs to be solved. The steps in this stage consist of data preparation, including dividing the dataset into two parts: testing data and training data, assigning weights to all data using TF-IDF, forming a model by selecting a classification algorithm and the parameters to be used, and finally evaluating the trained model.

3.3.1 Splitting the Data into Training Data and Testing Data.

The next step is data splitting, where the data will be divided into two categories: training data and testing data. Training data is the portion of data needed by the machine to learn the characteristics of the data in order to generate a model for making predictions on new data. Meanwhile, testing data is the new data that will be predicted using the

Copyright © 2023 **Putri Kurniawati**, Page 471
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

trained model [20]. There are three data splitting ratios used in this research: a 60:40 ratio, which means dividing the data into 60% training data and 40% testing data; a 70:30 ratio, which means dividing the data into 70% training data and 30% testing data; and an 80:20 ratio, which means dividing the data into 80% training data and 20% testing data. Each of these ratios will undergo the classification process to test accuracy and determine the best accuracy value. The details of the data distribution or split, as shown in Table 6 below, are as follows.

Table 6. Split Data Training and Testing

Data Split Ratio	Data Split		Total
	Training	*Testing*	
60:40	10.945	7.297	
70:30	12.769	5.473	18.242
80:20	14.593	3.649	

As seen in the Table 6 above, the total overall data before the split amounted to 18,242 review data. In the 60:40 data split ratio, there were 10,945 data for training and 7,297 data for testing. In the 70:30 ratio, there were 12,769 data for training and 5,473 data for testing. Finally, in the 80:20 ratio, there were 14,593 data for training and 3,649 data for testing.

3.3.2 TF-IDF Weighting

Before being used in classification, text data needs to be transformed into a numeric form that represents the weight of each word. In Table 7 below, examples of word weighting using TF-IDF are presented for three sample documents.

Table 7. TF-IDF Weighting Examples

Term	TFIDF		
	D1	**D2**	**D3**
bagus	0	0	0,46735
cepat	0,40204	0,51785	0
jalan	0	0	0,46735
jemput	0	0,68091	0
motor	0	0	0,46735
mudah	0,52863	0	0
murah	0,52863	0	0
nyaman	0	0	0,46735
ramah	0	0,51785	0,35543
terimakasih	0,52863	0	0

In the classification process, during model training, SVM uses training model vectors (training data) in the form of TF-IDF weight values as input in the equation to calculate the best hyperplane that can effectively separate the two classes. After the model is trained, the TF-IDF weight value vectors from the testing data are used as input to predict the correct class for each document. The testing data is projected into the same feature space used in the training data, and then the predicted class is determined based on the position or coordinates of the vector relative to the hyperplane that was defined during the training.

3.3.3 Model Formation

In determining the parameters of the classification model, this research uses the GridSearch method to find the best combination of kernel and parameter C that can optimize the performance of the classification model. The kernel used in this research is the Linear kernel. The parameter C was tested with values of [0.01, 0.1, 1, 100]. After parameter optimization, the best combination was found to be the Linear kernel with a parameter C value of 1. These parameters can be used to create a classification model with SVM and perform training on the training data before testing the model on the testing data. The training and testing of the best model are conducted using the syntax in the following Table 8.

Table 8. Classification Model

```
svm_model = svm.SVC(kernel='linear', C=1)
svm_model.fit(tf_train, y_train)
predicted = svm_model.predict(tf_test)
```

The syntax above indicates that a linear kernel SVM is used with a parameter C = 1. Next, the model is 'fit' or applied to the pre-weighted training data. After the training process with the training data, prediction is carried out using new data, which is the testing data, to predict its labels or sentiments. Then, to measure how accurate or good the classification model's performance is, evaluation is performed in the next stage.

3.3.4 Model Formation

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

In the SVM classification process, accuracy and evaluation results were obtained for three data splitting scenarios: 60:40, 70:30, and 80:20, as shown in Table 9 below.

Table 9. Model Performance Evaluation Results

Rasio	Akurasi	*Precision*	*Recall*	*F1 Score*	AUC
60:40	89.02%	92.17%	93.48%	92.82%	0.8419
70:30	89.82%	92.69%	94.09%	93.38%	0.8505
80:20	88.98%	92.26%	93.39%	92.82%	0.8408

Based on the model evaluation conducted, the model with a data ratio of 60:40 achieved an accuracy of 92.17%, the 70:30 ratio had an accuracy of 89.82%, and the 80:20 ratio had an accuracy of 88.98%. The model with the best accuracy was the 70:30 ratio, which resulted in a precision of 92.69%, recall of 94.09%, and an F1 score of 93.38% with a confusion matrix as shown in Table 10 below.

Table 10. Confusion Matrix Result

Confusion Matrix		***Prediction***	
		Negative	***Positive***
Actual	***Negative***	982	310
	Positive	247	3934

The precision value of 92.69% indicates that the model can correctly identify 92.69% of all data declared as positive by the model. The recall value of 94.09% shows that the model can correctly identify 94.09% of all positive data. The F1 score value of 93.38% indicates that the model has a good balance between recall and precision values. In addition to evaluation with the confusion matrix, another technique that can be used to evaluate model performance is by considering the area under the ROC curve or AUC (Area Under Curve). In Table 9 above, the best AUC value was obtained by the model with a 70:30 ratio, with an AUC value of 0.8505. This value indicates that the model has a fairly good ability to distinguish between positive and negative label classes and falls into the criteria of good classification based on the criteria in Table 2 and the curve visualization as shown in the Figure 4 below.

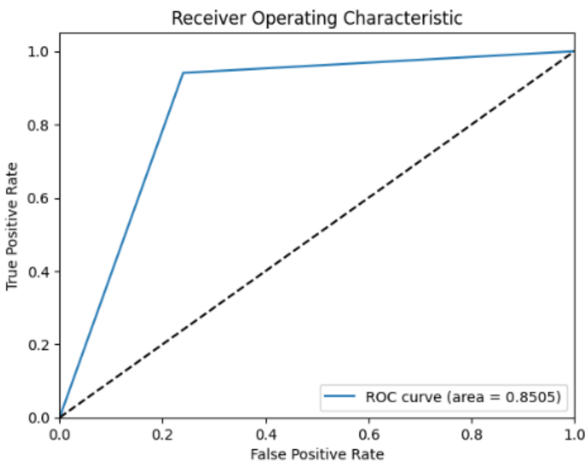

Figure 4. ROC-AUC Curve

Based on the results of the classification process, it was found that 76% of the sentiment was positive, while 24% was negative. The following is a visualization of word distribution in the reviews in the form of a word cloud, which visually represents words, where the frequency of occurrence of a word is indicated by the size of the word in the image.

Figure 5. Wordcloud Visualization

Copyright © 2023 **Putri Kurniawati**, Page 473
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

The expression of positive sentiment can be depicted using various words. In the image above, there is a visualization of frequently occurring words in the reviews. To get a clearer view of the sequence of words that appear most dominantly in positive reviews, this research also includes a visualization with a bar chart to display the words and their frequency of occurrence.

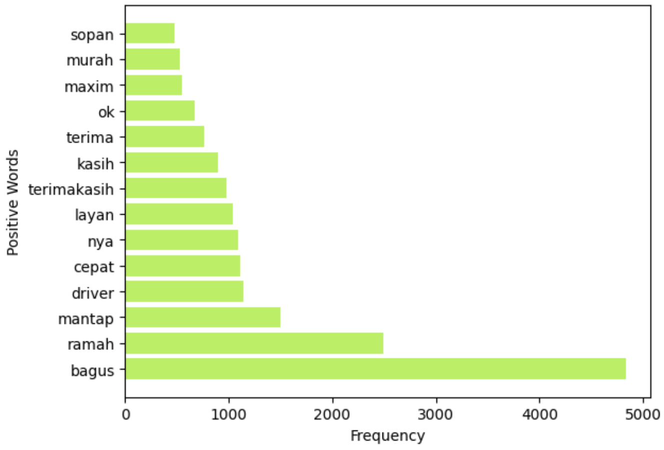

Figure 6. Positive Sentiment Word Distribution

The image above displays several words that dominantly appear in positive reviews, meaning that these words are frequently written and are highly present in the reviews provided by users. The most frequently appearing word is 'bagus,' which appears 4,000 times. Other dominant words that also appear in positive reviews include 'ramah,' 'mantap,' 'driver,' 'cepat,' 'layan,' 'terimakasih,' and others. Examples of positive review sentences include "mantap dan yang terbagus di Maxim, sangat cepat dan aman" (great and the best in Maxim, very fast and safe) and "Layanan Cepat , Driver Ramah" (Fast Service, Friendly Driver). These example sentences indicate that users are satisfied with Maxim's fast service, safety, and friendly drivers. This highlights the strengths of Maxim that should be maintained in providing services to users.

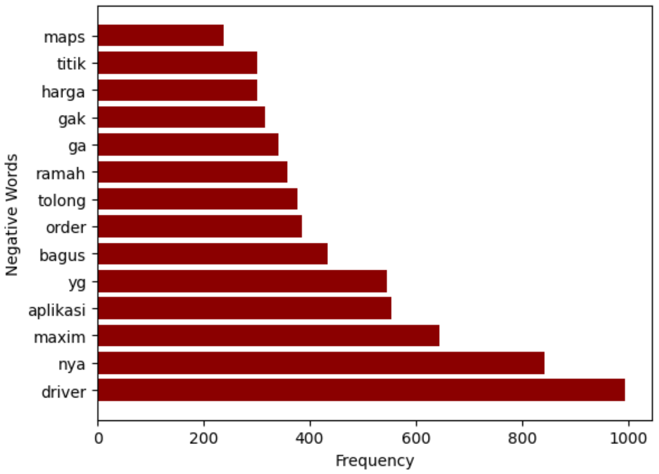

Figure 7. Negative Sentiment Word Distribution

The figure above displays several words that dominantly appear in negative reviews. The most frequently appearing word is 'driver,' which appears 1,000 times. Other dominant words that also appear in negative reviews include 'maxim,' 'aplikasi,' 'bagus,' 'order,' 'tolong,' 'ramah,' 'harga,' 'titik,' and others. Examples of negative review sentences include "ada kekurangan di titik penjemputan yg tidak tepat" (there are inaccuracies in the pick-up point), "kadang drivernya kejauhan dari titik penjemputan, tolong di perbagusi lagi" (sometimes the driver is far from the pick-up point, please improve it), and "Drivernya kurang ramah dan banyak maunya, selalu minta ganti rugi dan biaya lebih dr aplikasi, dan kalau pesanan tdk bisa dibatalkan" (The driver is not friendly and demanding, always asking for compensation and more fees than the app, and if the order cannot be canceled). These example sentences indicate that users are expressing complaints about the application. This also highlights the shortcomings of Maxim that need to be addressed, such as inaccurate map locations, drivers sometimes being too far, and drivers requesting additional fees and not being friendly to users.

4. CONCLUSION

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 466−475
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4265

The conclusions drawn from the conducted research include the following, The implementation of the Support Vector Machine algorithm in sentiment analysis of online transportation application reviews, specifically Maxim, resulted in a reasonably good model performance. Performance evaluation using the confusion matrix revealed the highest accuracy achieved by the classification model with a 70:30 data splitting ratio, which reached 89.82%. It also achieved precision of 92.17%, recall of 94.09%, and an F1 score of 93.38%. Additionally, this model obtained the best AUC (Area Under Curve) value of 0.8505, indicating good classification performance. Based on sentiment analysis results, the Maxim application tends to receive more positive comments from users, accounting for 76% of the feedback. Users expressed satisfaction with Maxim's services, such as the friendly and prompt behavior of drivers. This is one of Maxim's strengths that should be maintained. However, there were negative comments accounting for 24% of the feedback, which also require attention from developers to improve the application's performance and functionality. Some users reported complaints while using the application, such as malfunctioning features. This feedback can serve as an evaluation for developers to enhance performance, including improvements to crucial features like location accuracy. This research can also serve as a basis for future studies in sentiment analysis of application reviews, especially for online transportation applications. Overall, the research findings provide valuable insights into user sentiments and feedback regarding the Maxim application, which can guide developers in making improvements and maintaining user satisfaction.

REFERENCES

[1] M. Arif and K. U. Apjii, “Profil Internet Indonesia 2022,” no. June, 2022.

[2] R. K. Hastuti, “Duh, 2 Tahun Terakhir Ada 10 Ojol yang Tergilas Grab & Gojek,” CNBC Indonesia, 2019. https://www.cnbcindonesia.com/tech/20190813122355-37-91630/duh-2-tahun-terakhir-ada-10-ojol-yang-tergilas-grab-gojek

[3] Shopback, “Sering Membandingkan Harga Transportasi Online? Aplikasi Ini Akan Memudahkan Penggunanya,” Shopback.co.id, 2018. https://www.shopback.co.id/katashopback/transportasi-online-makin-digemari

[4] Taximaxim.com, “Tentang Perusahaan.” https://id.taximaxim.com/id/2093-jakarta/about/ (tanggal akses 10 Juli 2023)

[5] A. Mutia, “Survei: Publik Jabodetabek Paling Sering Pakai Gojek, Bagaimana Grab, Maxim, dan InDriver?,” Databoks.katadata.co.id, 2022. https://databoks.katadata.co.id/datapublish/2022/10/11/survei-publik-jabodetabek-paling-sering-pakai-gojek-bagaimana-grab-maxim-dan-indriver

[6] P. A. Permatasari, L. Linawati, and L. Jasa, “Survei Tentang Analisis Sentimen Pada Media Sosial,” Maj. Ilm. Teknol. Elektro, vol. 20, no. 2, p. 177, 2021, doi: 10.24843/mite.2021.v20i02.p01.

[7] S. W. Iriananda et al., “ANALISIS SENTIMEN DAN ANALISIS DATA EKSPLORATIF ULASAN,” no. Ciastech, pp. 473–482, 2021.

[8] A. Saepulrohman, S. Saepudin, and D. Gustian, “Analisis Sentimen Kepuasan Pengguna Aplikasi Whatsapp Menggunakan Algoritma NaÃ¯ve Bayes Dan Support Vector Machine,” is Best Account. Inf. Syst. Inf. Technol. Bus. Enterp. this is link OJS usf@, vol. 6, no. 2, pp. 91–105, 2021, doi: 10.34010/aisthebest.v6i2.4919.

[9] K. A. Rokhman, B. Berlilana, and P. Arsi, “Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online,” J. Inf. Syst. Manag., vol. 3, no. 1, pp. 1–7, 2021, doi: 10.24076/joism.2021v3i1.341.

[10] A. M. Rahat, A. Kahir, and A. K. M. Masum, “Comparison of Naive Bayes and SVM Algorithm based on Sentiment Analysis Using Review Dataset,” Proc. 2019 8th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2019, pp. 266–270, 2020, doi: 10.1109/SMART46866.2019.9117512.

[11] N. Herlinawati, Y. Yuliani, S. Faizah, W. Gata, and S. Samudi, “Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine,” CESS (Journal Comput. Eng. Syst. Sci., vol. 5, no. 2, p. 293, 2020, doi: 10.24114/cess.v5i2.18186.

[12] S. A. Salloum, M. Al-emran, and A. A. Monem, “Using Text Mining Techniques for Extracting Information from Research Using Text Mining Techniques for Extracting Information from Research Articles,” no. January, 2018, doi: 10.1007/978-3-319-67056-0.

[13] S. Bhatia, M. Sharma, and K. K. Bhatia, “Sentiment Analysis and Mining of Opinions,” Stud. Big Data, vol. 30, no. May, pp. 503–523, 2018, doi: 10.1007/978-3-319-60435-0_20.

[14] J. Han, J. Pei, and H. Tong, Data Mining : Concepts and Techniques. Morgan Kaufmann, 2022.

[15] K. X. Han, W. Chien, C. C. Chiu, and Y. T. Cheng, “Application of support vector machine (SVM) in the sentiment analysis of twitter dataset,” Appl. Sci., vol. 10, no. 3, 2020, doi: 10.3390/app10031125.

[16] P. S. Saragih, D. Witarsyah, F. Hamami, and J. M. MacHado, “Sentiment Analysis of Social Media Twitter with Case of Large Scale Social Restriction in Jakarta using Support Vector Machine Algorithm,” 2021 Int. Conf. Adv. Data Sci. E-Learning Inf. Syst. ICADEIS 2021, vol. 19, no. January 2020, pp. 1–6, 2021, doi: 10.1109/ICADEIS52521.2021.9701961.

[17] S. S. Chaeikar, A. A. Manaf, A. A. Alarood, and M. Zamani, “PFW: Polygonal fuzzy weighted—an SVM kernel for the classification of overlapping data groups,” Electron., vol. 9, no. 4, 2020, doi: 10.3390/electronics9040615.

[18] A. Zaiem and N. Charibaldi, “Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi),” vol. 9, no. 2, pp. 33–42, 2021.

[19] A. Pranata, E. Budianita, Yusra, and E. P. Cynthia, “Klasifikasi Sentimen Terhadap Maxim Menggunakan Algoritma SVM Pada Media Sosial TwittAnggi Pranata, N. (2022). Klasifikasi Sentimen Terhadap Maxim Menggunakan Algoritma SVM Pada Media Sosial Twitter. Klasifikasi Sentimen Terhadap Maxim Menggunakan Algorit,” Klasifikasi Sentimen Terhadap Maxim Menggunakan Algoritm. SVM Pada Media Sos. Twitter, vol. 5, no. 3, pp. 332–341, 2022.

[20] W. Musu, A. Ibrahim, and Heriadi, “Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4 . 5,” Pros. Semin. Ilm. Sist. Inf. Dan Teknol. Inf., vol. X, no. 1, pp. 186–195, 2021.

Copyright © 2023 **Putri Kurniawati**, Page 475
This Journal is licensed under a Creative Commons Attribution 4.0 International License