Jurnal Teknik Informatika (JUTIF)
Vol. 5, No. 1, February 2024, pp. 233-241

DOI: https://doi.org/10.52436/1.jutif.2024.5.1.1680
p-ISSN: 2723-3863
e-ISSN: 2723-3871

SENTIMENT ANALYSIS OF CUSTOMER SATISFACTION IN GOJEK AND GRAB APPLICATION REVIEWS USING THE NAIVE BAYES ALGORITHM

Ridha Faiz Ananda[1], Alfi Syahri[2], Firman Noor Hasan[*3]

[1,3]Information Technology, Faculty of Industrial Technology and Informatics, Universitas Prof. Dr. HAMKA Muhammadiyah, Indonesia
[2]Arabic Literature, Faculty of Arabic Language, Al-Azhar University, Egypt.
Email: [1]rfananda19@gmail.com, [2]alfi.sy18@gmail.com, [3]firman.noorhasan@uhamka.ac.id

(Article received: January 06, 2024; Revision: February 02, 2024; published: February 16, 2024)

Abstract

Online motorcycle taxis are a widely favored mode of public transportation in Indonesia. There are several companies providing online motorcycle taxi services in Indonesia, with Gojek and Grab dominating the market. In this rapidly digitizing era, social media has become a platform for Indonesian citizens to express their evaluations and opinions. One common platform used by users to express their evaluations is the Google Play Store, where users can provide ratings and opinions on the applications they use, including users of Gojek and Grab applications.This research aims to understand and analyze the sentiments of the public towards the two dominant giants in the online motorcycle taxi market in Indonesia based on review data from the Google Play Store using the Naive Bayes algorithm. The data used consists of user reviews from May 14, 2023, to July 26, 2023, totaling 300 data points for each application. This data will undergo pre-processing to remove irrelevant elements. The Naive Bayes algorithm is used to classify the existing sentiments into two classes: positive and negative.The results of this research conclude that Gojek users give positive reviews at 49% and negative reviews at 51%, which include praises for the drivers and services provided by the company, complaints about the heaviness of the application, and some disruptions in the Gopay payment method. Meanwhile, Grab users give positive reviews at 67% and negative reviews at 33%, which include customer satisfaction with attractive promos, complaints about the heaviness of the application after the latest update, and the high cost of Grabexpress and Grabfood services.

Keywords: *Customer satisfaction, Gojek, Grab, Naive Bayes, Online motorcycle taxi, Sentiment analysis.*

1. INTRODUCTION

The rapid development of technology has had a significant impact on our daily lives. One such example is in transportation, where, as we know, over the past eight years, particularly in 2015, the Indonesian public was introduced to a new phenomenon: online motorcycle taxis [1]. This year marked the emergence of online motorcycle taxis as a sensational development in the nation's transportation sector. This mode of transport became increasingly popular because it offers advantages not found in traditional motorcycle taxis, including practicality. Passengers simply wait at a designated pick-up point, and the driver will arrive to collect them. This convenience has become a selling point for online motorcycle taxis, continually attracting interest from the Indonesian populace. Moreover, these online taxi applications are usually hybrid, accessible to both Android and iOS users [2].

Gojek was the first company to offer online motorcycle taxi services in Indonesia, followed by its competitor, Grab, which began expanding its market to Indonesia [3]. This has made Gojek and Grab the two dominant players in the online motorcycle taxi market to date [4]. These two companies are continually competing by introducing cutting-edge innovations to appeal to online motorcycle taxi service enthusiasts in the country. The competition between these two giants certainly does not escape user evaluations. These user assessments are crucial for the companies to introspect and further develop their respective applications.

Google Play, commonly known as the Play Store, is a platform where users can share their opinions and ratings of the applications available. It is where the majority of users download Gojek and Grab, allowing them to provide feedback and ratings for these applications. Researchers utilize this user feedback from Google Play as a data source for this study. The data obtained from Google Play will undergo an analysis process to extract sentiment information contained within.

Sentiment analysis is a part of data mining research that focuses on text processing with the aim of understanding the meanings and emotions conveyed in a text [5]. There are many algorithms available for conducting sentiment analysis, with one of the most popular being the Naïve Bayes algorithm. This popularity is due to its high accuracy compared to other algorithms. The Naïve Bayes algorithm will be used throughout this research.

According to previous research by Dwijayanti et al., which discussed user satisfaction with Grab, the accuracy achieved by the Naïve Bayes algorithm was 92.5% [6]. This demonstrates the effectiveness of the Naïve Bayes algorithm in this study. Their research also indicated that Grab users felt satisfied with the services provided by the company. Another study by Wibowo et al., which focused on sentiment analysis of online learning during COVID-19, showed that the Naïve Bayes algorithm's performance was very satisfactory [7]. The accuracy achieved in this study was 88.5%, indicating that the Naïve Bayes algorithm is an effective tool for sentiment analysis. Another study conducted by Alfandi Safira and Firman Noor Hasan, which discusses the sentiment analysis of the public towards the paylater payment method, shows that the Naive Bayes algorithm is capable of classifying public sentiment well. The research resulted in an accuracy score of 95%[8].

2. RESEARCH METHODOLOGY

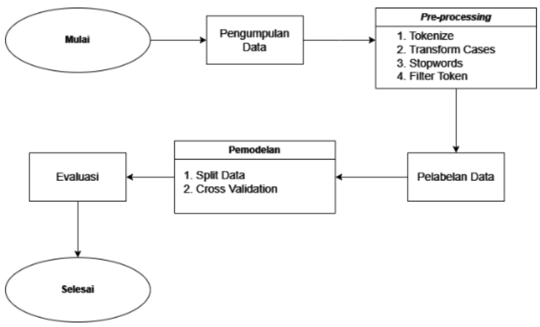

Figure 1. Research Methodology

Figure 1 illustrates the stages followed by the researcher in conducting this study. In this research, the researcher utilizes the Naïve Bayes algorithm as the main algorithm. The following is an explanation of each stage:

1. Data Collection

The researcher utilizes Google Colab as the primary tool for data extraction in this stage. The data extracted consists of reviews for the Gojek and Grab applications from Google Play. The review data is collected for the period from May 14, 2023, to July 26, 2023.

2. Pre-processing

Data Preparation or Pre-processing is the stage of transforming data into a form that is easier, simpler, and more suitable for the user's needs [9]. In this study, there are several processes involved in this data preparation stage, which are as follows:

- Tokenize

Tokenization is a process aimed at grouping words that are present in each sentence [10],

- Transform cases

Case transformation is a process that functions to convert all letters in the data into lowercase [11].

- Stopwords

Stopword removal is the process of eliminating frequently used words in a language. This process is necessary to allow users to focus more on important words [12]. Table 1 displays examples of stop words:

Table 1. Examples of Stopwords

Ada	Aku	Di
Hal	Ia	Ini
Itu	Jika	Kami
Kata	Ke	Kini

- Filter token

This stage is carried out with the aim of filtering tokens based on the number of characters contained in each token [13].

3. Data Labeling

Data labeling here refers to the process of assigning sentiments to data. In this process, the researcher labels each piece of data with either a positive or negative sentiment.

4. Modeling

The method used by the researcher in this modeling stage includes split data and cross-validation, where in each model, the Naïve Bayes algorithm is employed.

- Split Data

Split data is a model used by dividing the data into a ratio desired by the user. This is because there are no definitive rules for determining the ratio [14].

- Cross validation

Cross validation is a technique used to validate a model and assess how high the accuracy is in the results of the analysis [15].

5. Evaluation

he researcher employs a confusion matrix as a form of evaluation of the model used in the modeling stage. A confusion matrix is a table that illustrates the correct or incorrect classification outcomes of a value from the modeling used in the previous stage [16]. Figure 2 illustrates the confusion matrix table.

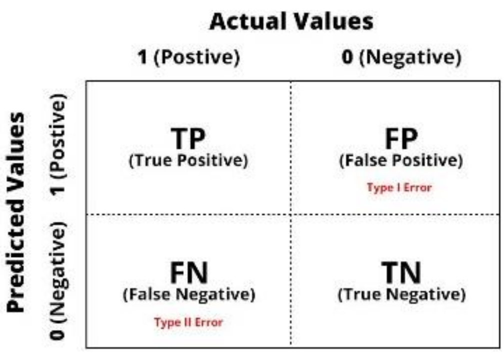

Figure 2. Confusion matrix

After obtaining the values from the confusion matrix, the process continues to the calculation stage, including Accuracy, Precision, and Recall. Accuracy is useful for determining the accuracy of the model used in the previous stage. Accuracy can be obtained using the following equation.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision is the result of the classification of true positives and all classes that potentially have a positive value [17]. Precision can be determined using the following equation.

$$Precission = \frac{TP}{TP+FP} \quad (2)$$

Recall is a value that indicates the effectiveness or success of the model in identifying information within the data. Below is the equation used to obtain the Recall value.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

3. RESULTS AND DISCUSSION

3.1. Data Collection

The researcher conducted data crawling using Google Colab, where the data extracted consisted of reviews from two applications, namely Gojek and Grab. These reviews were retrieved from the Google Play Store, which serves as a platform for Android-based applications. The researcher gathered 3000 data points for each application as the research objects. However, in the end, the researcher only utilized 300 data points for each application in CSV format. This was due to the age of the reviews being too old, rendering them irrelevant to current conditions. After obtaining the desired data, the researcher used RapidMiner as the main tool to process the available data. Figure 3 provides an overview of the researcher's data collection process.

Firgure 3. Data Crawling

Before moving on to the pre-processing/data preparation stage, the researcher created a word cloud to identify the most frequently appearing words. In this study, the researcher selected the top 25 words with the highest frequency in the data used for each application. The following are the word clouds created by the researcher to understand which words commonly appear in the reviews of both applications.

Figure 4. Gojek Word Cloud

Figure 4 shows that in the Gojek application, the most frequently occurring word in the reviews is "application," which appears 88 times in 76 reviews, followed by the word "driver," which appears 79 times in 48 reviews, and the word "Gojek," which appears 73 times in 62 reviews.

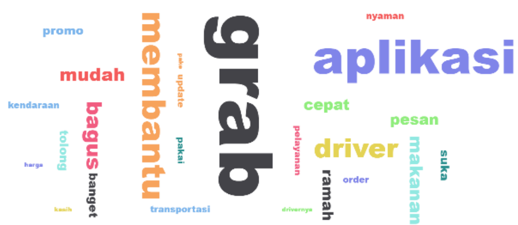

Figure 5. Grab Word Cloud

Meanwhile, in Figure 5, it is shown that in the Grab application reviews, the most frequently occurring word in the data used by the researcher is "Grab," which appears 111 times in 85 reviews, followed by the word "application," which appears 73 times in 63 reviews, and the word "helps," which appears 52 times in 52 reviews.

3.2. Pre-processing

The researcher utilized several operators in the data preparation (data pre-processing) process. Below is an image depicting the stages of data preparation that the researcher underwent. Figure 6 displays the data preparation process undertaken by the researcher.

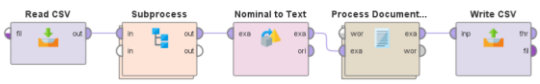

Figure 6. Pre-processing Stage

a) Read CSV

The operator is designed to read the data used by the researcher for this study. As explained earlier, the researcher utilized 300 sample data for each application in CSV format.

b) Subprocess

It serves as a terminal containing the operators needed for the data pre-processing stage. Figure 7 displays the operators within the subprocess.

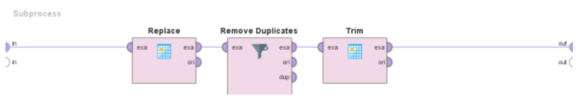

Figure 7. Operators in the Subprocess

- Replace

Replace is an operator that functions to substitute or remove unwanted characters [18]. In this stage, the researcher uses the replace operator to eliminate punctuation marks ([-!"#$%&'()*+,./:;<=>?@\[\\\]_`{|}~]) and emoticons. This is done to ensure that the data to be used by the researcher is free from punctuation marks, making it easier to determine the sentiments contained in the data. Table 2 displays the difference in data before and after the replacement process

Table 2. Results of Replace

Before	After	Sentimen
Bugs nya parah, ngefreeze lama banget.. Dulu ga kaya gini... Jadi pindah ke lain hati deh	Bugs nya parah ngefreeze lama banget Dulu ga kaya gini Jadi pindah ke lain hati deh	Negatif
aplikasi sangat membantu, lebih banyak lagi bikin promo...	aplikasi sangat membantu lebih banyak lagi bikin promo	Positif
Aplikasi transportasi yang bermanfaat bagi kalangan masyarakat bawah	Aplikasi transportasi yang bermanfaat bagi kalangan masyarakat bawah	Positif

- Remove Duplicate

This operator functions as a filter for sentences that have the same value, ensuring that the data used does not have duplicates or identical values. Table 3 displays the difference in data before and after the process of removing duplicates.

Table 3. Results of Remove Duplicate

Before	After	Sentimen
nyaman dan mengutamakan keselamatan	nyaman dan mengutamakan keselamatan	Positif
nyaman dan mengutamakan keselamatan	sangat puas tepat wkt sampai tujuan	Positif
sangat puas tepat wkt sampai tujuan	Tiap kali pesan makanan selalu terkendala dengan pembayaran Mendingan saya menggunakan Gojek atau shopeefood	Negatif

- Trim

This operator serves as a remover of spaces at the beginning and end of nominal attribute values that have been determined [19]. Table 4 displays the difference in data before and after the trimming process.

Table 4. Results of Trim

Before	After	Sentimen
Baik membantu dan aman	Baik membantu dan aman	Positif
saya tidak suka karna penulisan alamat tidak sesuai	saya tidak suka karna penulisan alamat tidak sesuai	Negatif
Suka banget sama grep tolongin ane gan maaf mas bro biar bisa aplikasi Android bisa berjalan dengan baik terimakasih atas kunjungan Saya	Suka banget sama grep tolongin ane gan maaf mas bro biar bisa aplikasi Android bisa berjalan dengan baik terimakasih atas kunjungan Saya	Positif

c) Nominal To Text

This operator functions to convert data from nominal format to text format. This step is necessary because the researcher uses the "process document from data" operator, which requires that the data used must be in text format

d) Process Document From Data

In this operator, there are several sub-operators that play a crucial role in the data pre-processing stage. Figure 8 displays the operators used in the "process document from data" process.

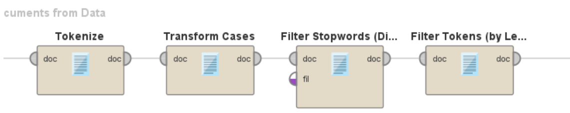

Figure 8. Operators in the Subprocess 'Process Document From Data'

- Tokenize

It functions to group the words present in the data being used. Table 5 displays the difference in data before and after the Tokenize process.

Table 5. Tokenize Results

Before	After	Sentimen
aplikasinya bagus bangt cuma kalau lupa pasword nunggu nya terlalu lama baru bisa di resep tapi aplikasinya udh nyaman banget	“aplikasinya”,”bagus”,”bangt”,”Cuma”,”kalau”,”lupa”,”pasword”,”nunggunya”,”terlalu”,”lama”,”baru”,”bisa”,”di”,”resep”,”tapi”,”aplikasinya”,”udh”,”nyaman”,”banget”	Positif
Mudah aplikasi nya dan lengkap	“Mudah”,”aplikasi”,”nya”,”dan”,”lengkap”	Negatif
Mudah cepat dan banyak promo	“Mudah”,”cepat”,”dan”,”banyak”,”promo”	Positif

- Transform cases

This operator functions to convert all uppercase characters in the data to lowercase. Table 6 displays the difference in data before and after the Transform Cases process.

Table 6. Results of Transform Cases

Before	After	Sentimen
“Mahal”, ”kali”, ”ongkos”, ”klen”, ”Beda”, ”13k”, ”dari”, ”Grab”, ”Begitu”, ”dicancel”, ”sekali”, ”dan”, ”mau”, ”order”, ”lagi”, ”naik”, ”dari”, ”73k”, ”ke”, ”90k”	“mahal”, ”kali”, ”ongkos”, ”klen”, ”beda”, ”13k”, ”dari”, ”Grab”, ”begitu”, ”dicancel”, ”sekali”, ”dan”, ”mau”, ”order”, ”lagi”, ”naik”, ”dari”, ”73k”, ”ke”, ”90k”	Positif
“Cukup”, ”membantu”, ”untuk”, ”kebutuhan”, ”cepat”	“cukup”, ”membantu”, ”untuk”, ”kebutuhan”, ”cepat”	Negatif
“Dah”, ”login”, ”tpi”, ”kalo”, ”gk”, ”dipake”, ”beberapa”, ”hari”, ”disuruh”, ”login”, ”ulang”, ”tapi”, ”gak”, ”dikirm”, ”SMS”, ”nya”	“dah”, ”login”, ”tpi”, ”kalo”, ”gk”, ”dipake”, ”beberapa”, ”hari”, ”disuruh”, ”login”, ”ulang”, ”tapi”, ”gak”, ”dikirm”, ”sms”, ”nya”	Positif

- Filter Stopwords

It is the process of filtering out commonly used words in a language, with the aim of allowing users to focus on words that carry specific sentiment values. Table 7 displays the difference in data before and after the stop words filtering process.

Table 7. Results of Filtering Stopwords

Before	After	Sentimen
“mahal”, ”kali”, ”ongkos”, ”klen”, ”beda”, ”13k”, ”dari”, ”Grab”, ”begitu”, ”dicancel”, ”sekali”, ”dan”, ”mau”, ”order”, ”lagi”, ”naik”, ”dari”, ”73k”, ”ke”, ”90k”	“mahal”, ”kali”, ”ongkos”, ”klen”, ”beda”, ”13k”, ”Grab”, ”dicancel”, ”order”, ”73k”, ”90k”	Positif
“cukup”, ”membantu”, ”untuk”, ”kebutuhan”, ”cepat”	”membantu”, ”kebutuhan”, ”cepat”	Negatif
“dah”, ”login”, ”tpi”, ”kalo”, ”gk”, ”dipake”, ”beberapa”, ”hari”, ”disuruh”, ”login”, ”ulang”, ”tapi”, ”gak”, ”dikirm”, ”sms”, ”nya”	“dah”, ”login”, ”kalo”, ”gk”, ”dipake”, ”disuruh”, ”login”, ”ulang”, ”gak”, ”dikirm”, ”sms”	Positif

- Filter tokens

This operator functions as a filter for tokens in the data based on the number of characters contained within a token. In this operator, the author used parameter values of a minimum of 4 characters and a maximum of 25 characters, meaning that words with fewer than 4 characters or more than 25 characters will be removed. Table 8 displays the difference in data before and after the token filtering process.

Table 8. Results of Filter Tokens

Before	After	Sentimen
“mahal”, ”kali”, ”ongkos”, ”klen”, ”beda”, ”13k”, ”Grab”, ”dicancel”, ”order”, ”73k”, ”90k”	“mahal”, ”kali”, ”ongkos”, ”klen”, ”beda”, ”Grab”, ”dicancel”, ”order”	Positif
”membantu”, ”kebutuhan”, ”cepat”	”membantu”, ”kebutuhan”, ”cepat”	Negatif
“dah”, ”login”, ”kalo”, ”gk”, ”dipake”, ”disuruh”, ”login”, ”ulang”, ”gak”, ”dikirm”, ”sms”	”login”, ”kalo”, ”dipake”, ”disuruh”, ”login”, ”ulang”, ”dikirm”	Positif

e) write CSV

Functions to save data that has gone through the pre-processing stage using the CSV format (Yarist Kusnaedi & Hisyam, n.d.). In this operator, users can specify where they want to save the file that has undergone pre-processing according to their preferences.

3.3. Data Labeling

In this stage, each data point is assigned a sentiment value based on the nature of the sentiment it contains. The criterion used by the researcher for assigning sentiment values is as follows: if a data point contains positive words that are supportive or constructive, the researcher will assign a positive sentiment value to that data. Conversely, a negative sentiment value is assigned if a data point contains negative words that are hateful or derogatory.

The researcher labeled all the data used, meaning the researcher labeled 100% of the data used. Figure 8 and Figure 9 display the data labeled by the researcher.

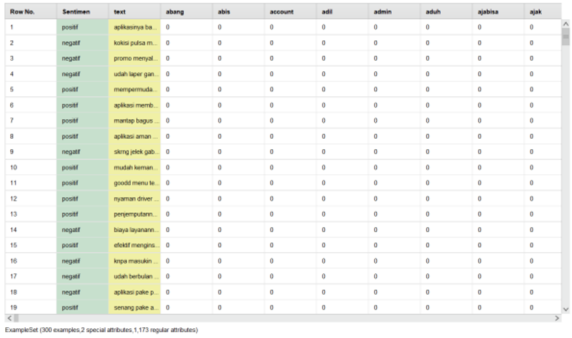

Figure 8. Gojek Data After Labeling Process

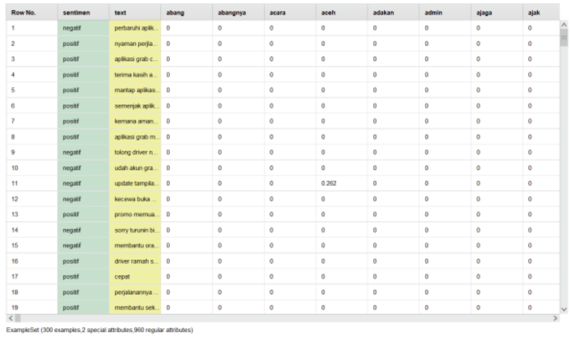

Figure 9. Grab Data After Labeling Process

3.4. Modeling

The researcher uses two methods for modeling: split data and cross-validation. In each of these methods, the researcher employs the Naïve Bayes algorithm for each model.

a) Split data

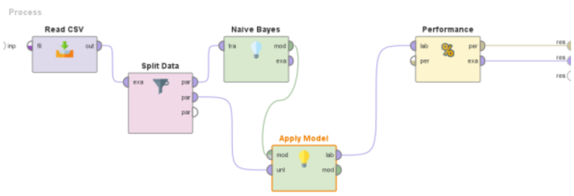

Figure 10. Data Split Using Naïve Bayes

Figure 10 displays the operators used in the modeling process. Here are the explanations for each operator used.

- Read CSV

To read data that has undergone the pre-processing and labeling stages.

- Split Data

Is an operator that functions to divide data into a certain ratio according to the user's desire.

- Algoritma naïve bayes

Naïve Bayes is a probability prediction method that uses Bayes' Theorem, with the assumption that features in the data are not interrelated [20].

- Apply Model

This operator functions as a processor for comparing data used as training data with data used as testing data [21].

- Performance

This operator is useful as an indicator of the accuracy results from the model that has been used [22].

b) Cross validation

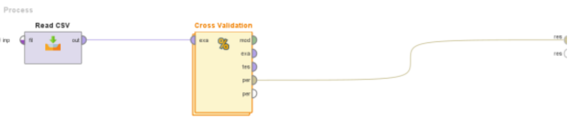

Figure 11. Cross-Validation Process

The Cross-validation method is divided into two parts: the process and subprocess. Figure 11 illustrates the process part, where the Cross-validation operator contains a subprocess that serves as the placement of the algorithm to be used and also as the divider between training and testing. Figure 12 shows the contents of the subprocess of cross-validation.

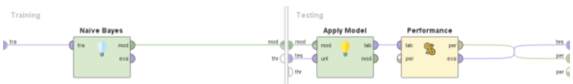

Figure 12. Subprocess of Cross Validation

The parameter used in this Cross Validation process is the number of folds. In this process, the researcher uses the default number of folds, which is 10. The number of folds serves as a basis for assessing the best and worst combinations [23].

3.5. Evaluation

After conducting the modeling, the researcher proceeds to the next process, which is the evaluation stage. The following are the results of the above modeling.

a) Split data

Table 9 presents the accuracy results of the model for the Gojek dataset and the Grab dataset using the split data method.

Table 9. Split Data Accuracy Values

	Hasil Nilai Akurasi	
Rasio	**Gojek**	**Grab**
80:20	83.33%	80.00%
70:30	**85.56%**	**84.44%**
60:40	84.44%	83.33%

Based on the table above, it shows that the highest accuracy is achieved using a 70:30 comparison ratio with an accuracy value of 85.56% for the Gojek application and 84.44% for the Grab application.

b) Cross validation.

Table 10 presents the accuracy results of the model for the Gojek dataset and the Grab dataset using the Cross-validation method.

Table 10. Accuracy Value Cross-validation

	Hasil Nilai Akurasi	
Naïve Bayes	**Gojek**	**Grab**
	83.33%	85.67%

The table above shows that the highest accuracy achieved is 88.33% for the Gojek app and 85.67% for the Grab app.

From the two methods used to measure the performance level of the classification, namely the Split data method and Cross-validation, it can be concluded that the highest result achieved from both classification performance measurement methods is the model using the Cross-validation method, with 88.33% for the Gojek app and 85.67% for the Grab app.

From the modeling results above, the researcher uses the highest value for the confusion matrix, where the data used is obtained from the modeling results using the Cross-validation method. Here are the confusion matrix results from the two datasets used.

a) Confusion matrix Aplikasi Gojek

Table 11. Confusion Matrix for the Gojek Application

	True Positive	***True Negative***	***Class Precission***
Pred. Positive	118	18	86.76%
Pred. Negative	32	132	80.49%
Class Recal	78.67%	88.00%	

The following is the calculation from table 11:

- *Precission* $= \frac{TP}{TP+FP} = \frac{118}{118+18} = 0.8676 = 86.76\%$
- *Recall* $= \frac{TP}{TP+FN} = \frac{118}{118+32} = 0.7867 = 78.67\%$
- *Accuracy* $= \frac{TP+TN}{TP+TN+FP+FN} = \frac{118+132}{118+132+18+32} = 0.8333 = 83.33\%$

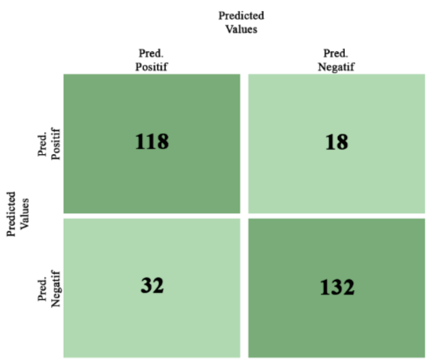

Figure 13. Gojek Confusion Matrix

Based on the confusion matrix results above, it shows that the Naive Bayes algorithm is capable of classifying public reviews on the Gojek application.

This is because the accuracy value generated by the Naive Bayes algorithm is 83.33%.

b) Confusion matrix Aplikasi Grab

Table 12. Confusion Matrix for the Grab Application

	True Positive	***True Negative***	***Class Precission***
Pred. Positive	81	27	75.00%
Pred. Negative	16	176	91.67%
Class Recal	83.51%	86.70%	

The following is the calculation from table 12:

- *Precission* = $\frac{TP}{TP+FP} = \frac{81}{81+27}$ = 0.75 = 75.00%
- *Recall* =$\frac{TP}{TP+FN} = \frac{81}{81+16}$ = 0.8351= 83.51%
- *Accuracy* = $\frac{TP+TN}{TP+TN+FP+FN} = \frac{81+176}{81+176+27+16}$ = 0.8567 = 85.67%

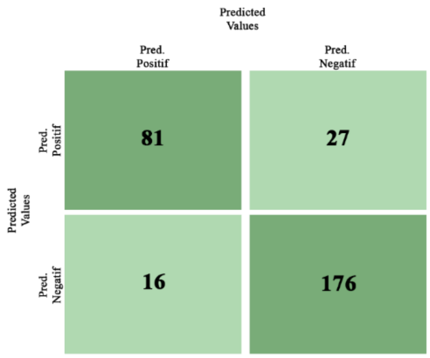

Figure 14. Grab Confusion Matrix

Based on the confusion matrix results above, it shows that the Naive Bayes algorithm is more accurate in classifying the Grab application compared to the Gojek application. This is because the accuracy value generated by the Naive Bayes algorithm is higher than the accuracy value for the Gojek application, which is 85.67%.

4. DISCUSSION

The following are some previous studies that discuss satisfaction using the Naive Bayes classification method. The first study was conducted by Anisa Halifa and colleagues, focusing on customer satisfaction with postal services in the Rumbai sub-district. This research utilized the Naive Bayes algorithm to classify sentiments in the dataset, resulting in an accuracy score of 94.74%, a recall of 94.12%, and precision of 100% [24]. The second study, conducted by Muhammad Sidik and colleagues, explored student satisfaction with university services using the Naive Bayes algorithm, achieving an accuracy of 96.24%, a recall of 98.96%, and precision of 93.14% [25]. Both of these studies affirm that the Naive Bayes algorithm is effective for sentiment analysis classification. This forms the basis for the researcher's choice to use the Naive Bayes algorithm in the current study. Moreover, when compared to other algorithms, Naive Bayes demonstrates the highest level of accuracy.

5. CONCLUSION

Based on the analysis results of the datasets for the two online motorcycle taxi applications, Gojek and Grab, it is found that the best accuracy value is obtained through the model that uses the Cross-validation method. For the Gojek application, the best accuracy value is 83.33%, with 188 true positives, 18 false positives, 32 false negatives, and 132 true negatives. Meanwhile, for the Grab application, the best accuracy value is 85.67%, with 81 true positives, 27 false positives, 16 false negatives, and 176 true negatives. In each application, the sentiments obtained from each opinion are divided into negative and positive classes.

For the Gojek application, negative sentiments include user complaints about the numerous disruptions when using the application, especially with Gojek's exclusive payment method, Gopay, which causes inconvenience for users when using the Gojek application. On the other hand, positive sentiments mostly consist of user satisfaction with the services provided by Gojek and praise for the skilled and friendly drivers.

As for the Grab application, negative sentiments include user complaints about the heaviness of the Grab application after the latest update, as well as the high delivery fees for Grabfood and Grabexpress features compared to its competitors. On the positive side, most sentiments express user satisfaction with the satisfying driver services and attractive promotions that encourage users to subscribe to the Grab application.

DAFTAR PUSTAKA

[1] O. Halilintarsyah, "Ojek Online, Pekerja atau Mitra?," *Jurnal Persaingan Usaha*, vol. 02, pp. 64–74, Dec. 2021, doi: https://doi.org/10.55869/kppu.v2i.24.

[2] Yusuf Wahyu Setiya Putra *et al.*, *Pengantar Aplikasi Mobile*. 2023.

[3] K. Wongso and W. Purnama Sari, "Analisa UX Writing terhadap User Experience pada Pengguna Aplikasi Grab," *Prologia*, vol. 4, pp. 1–9, Feb. 2020, doi: https://doi.org/10.24912/pr.v4i1.6415.

[4] J. C. Jansela Situmorang, F. Akbar Maulana Hutabarat, W. Pong Lan, N. Nugroho, and Lisa, "Kualitas Pelayanan Mitra Gojek Terhadap Kepuasan Pengguna Layanan Go-Ride," *ARBITRASE*, vol. 2, pp. 51–54, Nov. 2021, doi: https://doi.org/10.47065/arbitrase.v2i2.285.

[5] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter

BMKG Nasional," *TEKNO KOMPAK*, vol. 15, no. 1, pp. 131–145, 2021, doi: https://doi.org/10.33365/jtk.v15i1.744.

[6] M. Dwijayanti, F. Noor Hasan, and R. Zein Adam, "Analisis Sentimen Pada Ulasan Pelanggan Menggunakan Metode Naïve Bayes Classifier (Studi Kasus: Grab Indonesia)," *Prosiding Seminar Nasional Teknoka*, vol. 6, pp. 93–99, Jan. 2022, doi: 10.22236/teknoka.v6i1.441.

[7] A. Wibowo, F. Noor Hasan, L. Akbar Ramadhan, R. Nurhayati, and dan Arief Wibowo, "Analisis Sentimen Opini Masyarakat Terhadap Keefektifan Pembelajaran Daring Selama Pandemi COVID-19 Menggunakan Naïve Bayes Classifier," *Asiimetrik:*, vol. 4, pp. 239–248, 2022, doi: https://doi.org/10.35814/asiimetrik.v4i1.3577.

[8] A. Safira and F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," *Jurnal Sistem Informasi*, vol. 5, no. 1, 2023, doi: https://doi.org/10.31849/zn.v5i1.12856.

[9] F. Setya Ananto and F. N. Hasan, "Implementasi Algoritma Naïve Bayes Terhadap Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store," *Jurnal ICT : Information Communication & Technology*, vol. 23, no. 1, pp. 75–80, 2023, [Online]. Available: https://ejournal.ikmi.ac.id/index.php/jict-ikmi

[10] N. Tri Romadloni, I. Santoso, S. Budilaksono, and M. Ilmu Komputer STMIK Nusa Mandiri Jakarta, "Perbandingan Metode Naive Bayes, Knn Dan Decision Tree Terhadap Analisis Sentimen Transportasi Krl Commuter Line," *Jurnal IKRA-ITH Informatika*, vol. 3, no. 2, pp. 1–9, 2019.

[11] M. I. Ahmadi, D. Gustian, and F. Sembiring, "Analisis Sentiment Masyarakat terhadap Kasus Covid-19 pada Media Sosial Youtube dengan Metode Naive bayes," *Jurnal Sains Komputer & Informatika (J-SAKTI*, vol. 5, no. 2, pp. 807–814, 2021, doi: http://dx.doi.org/10.30645/j-sakti.v5i2.378.

[12] W. Yulita *et al.*, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier," *JDMSI*, vol. 2, no. 2, pp. 1–9, 2021, doi: https://doi.org/10.33365/jdmsi.v2i2.1344.

[13] D. Setian and I. Seprina, "Analisis Sentimen Masyarakat Terhadap Data Tweet Lazada Menggunakan Text Mining Dan Algoritma Naive Bayes Classifier," *Bina Darma Conference on Computer Science*, vol. 1, no. 4, pp. 998–1004, 2020.

[14] Rahmadhan Adinugroho, "Perbandingan Rasio Split Data Training Dan Data Testing Menggunakan Metode Lstm Dalam Memprediksi Harga Indeks Saham Asia," 2022.

[15] D. Azzahra Nasution, H. H. Khotimah, and N. Chamidah, "Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *CESS (Journal of Computer Engineering System and Science)*, vol. 4, no. 1, pp. 2502–7131, 2019.

[16] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *Jurnal Sains Komputer & Informatika (J-SAKTI*, vol. 5, no. 2, pp. 697–711, 2021, doi: http://dx.doi.org/10.30645/j-sakti.v5i2.369.

[17] A. Rahim, K. Kusrini, and E. T. Luthfi, "Convolutional Neural Network untuk Kalasifikasi Penggunaan Masker," *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 10, no. 2, p. 109, Dec. 2020, doi: 10.35585/inspir.v10i2.2569.

[18] K. Nugroho, F. N. Hasan, P. Korespondensi, : Firman, N. Hasan, and R. Artikel, "Analisis Sentimen Masyarakat Mengenai RUU Perampasan Aset Di Twitter Menggunakan Metode Naïve Bayes Analysis of Public Sentiment Regarding RUU Perampasan Aset on Twitter Using Naïve Bayes Method," *SMATIKA : STIKI Informatika Jurnal*, vol. 13, no. 2, p. 13, 2023, doi: 10.32664/smatika.v13i02.899.

[19] A. Saputra and F. Noor Hasan, "Analisis Sentimen Terhadap Aplikasi Coffee Meets Bagel Dengan Algoritma Naïve Bayes Classifier," *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 2, no. 2, pp. 465–474, Jan. 2023, doi: 10.54443/sibatik.v2i2.579.

[20] M. E. Lasulika, "Komparasi Naïve Bayes, Support Vector Machine Dan K-Nearest Neighbor Untuk Mengetahui Akurasi Tertinggi Pada Prediksi Kelancaran Pembayaran Tv Kabel," *ILKOM Jurnal Ilmiah*, vol. 11, no. 1, pp. 11–16, May 2019, doi: https://doi.org/10.33096/ilkom.v11i1.408.11-16.

[21] M. A. Rosari,) Wasino, and) Tony, "Analisis Sentimen Tanggapan Masyarakat Terhadap Bantuan Sosial Pemerintah Di Masa Pandemi Covid-19 Pada Platform Twitter," *Jurnal*

Ilmu Komputer dan Sistem Informasi, vol. 10, no. 1, 2022, doi: https://doi.org/10.24912/jiksi.v10i1.17867.

[22] N. L. W. S. R. Ginantra, C. P. Yanti, G. D. Prasetya, I. B. G. Sarasvananda, and I. K. A. G. Wiguna, "Analisis Sentimen Ulasan Villa di Ubud Menggunakan Metode Naive Bayes, Decision Tree, dan K-NN," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 11, no. 3, pp. 205–215, Dec. 2022, doi: 10.23887/janapati.v11i3.49450.

[23] D. Pardede, I. Firmansyah, M. Handayani, M. Riandini, and R. Rosnelly, "Comparison Of Multilayer Perceptron's Activation And Optimization Functions In Classification Of Covid-19 Patients," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 8, no. 3, pp. 271–278, Aug. 2022, doi: 10.33330/jurteksi.v8i3.1482.

[24] A. Halifa and R. Novita, "Application Of Naïve Bayes Classifier Algorithm In Determining The Level Of Customer Satisfaction With Rumbai Post Office Services," *Jurnal Teknik Informatika (JUTIF)*, vol. 4, no. 6, pp. 1295–1304, 2023, doi: 10.52436/1.jutif.2023.4.6.1054.

[25] M. Siddik, R. Noratama Putri, and Y. Desnelita, "Classification Of Student Satisfaction On Higher Education Services Using Naïve Bayes Algorithm," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 3, no. 2, 2020, doi: https://doi.org/10.31539/intecoms.v3i2.1654