Implementation of Classification Algorithm for Sentiment Analysis: Measuring App User Satisfaction

Rizki Aulia Putra[1*], Rice Novita[2], Tengku Khairil Ahsyar[3], Zarnelly[4]

[1,2,3,4]Information System Study Program, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia
Email: [1*]12050313356@students.uin-suska.ac.id, [2]rice.novita@uin-suska.ac.id, [3]tengkukhairil@uin-suska.ac.id, [4]zarnelly@uin-suska.ac.id

(Received: 16 May 2024, revised: 1 Jun 2024, accepted: 3 Jun 2024)

Abstract

Google Play Store is the official app store for Android devices from Google that offers rating and review features. This feature on the platform is a source of data for sentiment analysis in research on app user satisfaction. The purpose of this study is to provide an overview of app user satisfaction and evaluate the accuracy of the algorithms used. The algorithms compared include Support Vector Machine (SVM), namely linear, rbf, sigmoid, and polynomial kernels with Naïve Bayes Classifier (NBC). The key variables analyzed include perceived usefulness, perceived ease of use, relia-bility, responsiveness, and website design. The results showed that the SVM algorithm with a linear kernel achieved the highest accuracy of 95.23% compared to the NBC algorithm of 91.43%. For other accuracy results, rbf kernel 94.35%, sigmoid kernel 95.19% and polynomial kernel 93.31%. In addition, the results of sentiment analysis on application user satis-faction revealed that 75% of users were dissatisfied, with the service indicator having the highest number of negative sen-timents. These findings suggest that sentiment analysis can be an effective tool for companies to measure and improve user satisfaction. In addition, these results can also be a useful reference for new users in assessing apps before using them.

Keywords: Google Play Store, Naïve Bayes, Sentiment Analysis, Support Vector Machine, User Satisfaction.

I. INTRODUCTION

Improvements in technology have led to a notable surge in Internet usage, resulting in the emergence of novel routines like social networking and e-commerce-based online purchasing [1]. Transactions now feel more customized, flexible, and convenient because to the growth of e-commerce [2]. Every year, the value of e-commerce transactions in Indonesia rises. Delivery services, acting as a middleman between buyers and sellers, are crucial to getting items to final customers. Because of this, there is a greater demand for dependable delivery services as well as an increase in public interest in online purchasing [3].

J&T Express, a leading company in the expedition sector operating in Indonesia [4], recorded a 40% increase in shipment volumes over 2023, with an average of 2.5 million packages shipped daily [5]. According to the Top Brand Award 2023, J&T Express was named the best brand in the courier service category with 33.3%, marking a significant increase from the previous year. According to one of Indonesia's online survey platforms Populix, J&T Express is the favorite courier service among both Generation Z and Millennials for their online shopping needs.

To stay competitive and retain customers, innovations are constantly being made. PT Global Jet Express launched a smartphone app in 2015, allowing customers to track shipments, find out rates, and get package information [4]. The J&T Express app continues to be popular with ever-increasing downloads, according to data from app.sensortower.com, the number of downloads in Indonesia for all platforms from December 2015 to December 2023 reached 50 million to 100 million downloads, with more than one million active users according to similar web [6].

The J&T Express app received a low rating of 2.2 on the Play Store and 2.0 on the App Store, with many 1-star reviews continuing to drop. Comments indicate disappointment with delivery accuracy, poor service, including delays, breakdowns, and a tracking feature that doesn't work properly despite being updated [7]. J&T Express app user satisfaction is key to maintaining and increasing the company's market share. In order to ensure customer satisfaction, J&T Express needs to deeply understand the level of user satisfaction [8]. Sentiment analysis of app comments is becoming an important tool in understanding user views and experiences, allowing companies to identify issues, improve services, and maximize user experience [9].

Sentiment analysis has been a major research topic since 2008, addressing methods such as machine learning or lexicon-based techniques as well as applications in various contexts [10]. This research applies data mining methods with the Support Vector Machine (SVM) and Naive Bayes Classifier (NBC) algorithms to analyze user comments on the J&T Express application on the Google Play Store [11]. The goal is to measure user satisfaction with application services, by considering variables from the Technology Acceptance Model (TAM), Service Quality (Servqual), and from previous research [12].

II. LITERATURE REVIEW

In recent years, there have been many sentiment analysis studies that discuss the comparison of SVM and NBC classification algorithms. Support Vector Machine was developed by Boser, Guyon, Vapnik, and first published in 1992 at the Annual Workshop on Computational Learning Theory [13]. The SVM algorithm has several kernels including Linear, Radial Basis Function (RBF), Polynomial, and Sigmoid [14]. Meanwhile, Naive Bayes Classifier is a simple classification method with the application of Bayes Theorem (Bayesian Statistics) proposed by an English scientist named Thomas Bayes [15] [16]. Finds the highest probability of the most appropriate category and is one of the classification techniques used in sentiment analysis, and has good potential for computational classification of data and precision [1]. The following are four studies that are considered relevant because they use SVM and NBC methods. These studies are presented in Table 1.

Table 1. Relevant Research

Information	Algorithm	Result
Research I [14]	SVM Kernel Lienar and RBF	1. Scenario (70:30) Kernel Linear 87% Acuracy 2. Scenario (90:10) Kernel RBF 87% Acuracy
Research II [16]	NBC and PSO	Using 10-Fold Cross Validation, showing an increase in accuracy from 77.96% to 79.85%
Research III [17]	SVM Kernel Linear, RBF, Sigmoid, Polynomial and Logistic Regression	SVM Acuracy 1. Linear 87,02% 2. RBF 84,59% 3. Sigmoid 86,63% 4. Polynomial 71,12% 5. Logistic Regression 85,17% Acuracy
Research IV [18]	NBC and KNN	1. NBC 75,53% Acuracy 2. KNN 48,66% Acuracy

Table 1 shows the four most relevant different studies sourced from scientific articles. The four studies discuss the comparison of classification algorithms for sentiment analysis with the main problem of providing public sentiment on an application or a particular topic. The similarity of these studies discusses sentiment analysis with several classification algorithms, several kernel parameters in the SVM algorithm in the first and third studies and algorithm comparisons in the second and fourth studies. Based on these findings, it can be concluded that the SVM and NBC algorithms provide good accuracy results with some additional parameters making them suitable for sentiment analysis. In addition to the use of kernal parameters in the SVM algorithm, there are other studies of derivative forms of the NBC algorithm such as Multinomial also provide good results in sentiment analysis [19]. In addition, sentiment analysis can be further developed to be better so as to provide novelty research by utilizing sentiment results. By adding several variables to the final sentiment results that were previously only completed on positive and negative sentiments.

In this study, variables will be used to measure user satisfaction from applications with several variables such as perceived usefulness, perceived ease of use [20], reliability, responsiveness [21] and website design [11]. The reason for using these variables has a match with the data and research topics, while the algorithm used is SVM with four kernels, namely, Linear, RBF, Sigmoid, Polynomial and Multinomial, Gaussian and Bernoulli Naïve Bayes.

III. METHODOLOGY

This research aims to apply classification algorithms in performing sentiment analysis on Google Play Store reviews of the J&T Express application. Processing is done with SVM kernel Linear, RBF, Sigmoid, Polynomial algorithms as well as Multinomial, Gaussian, Bernoulli Naïve Bayes algorithms. The use of variables in this study is expected to measure the satisfaction of application users in sentiment analysis. The entire research process is summarized in the following method (Figure 1).

A. Research Methods

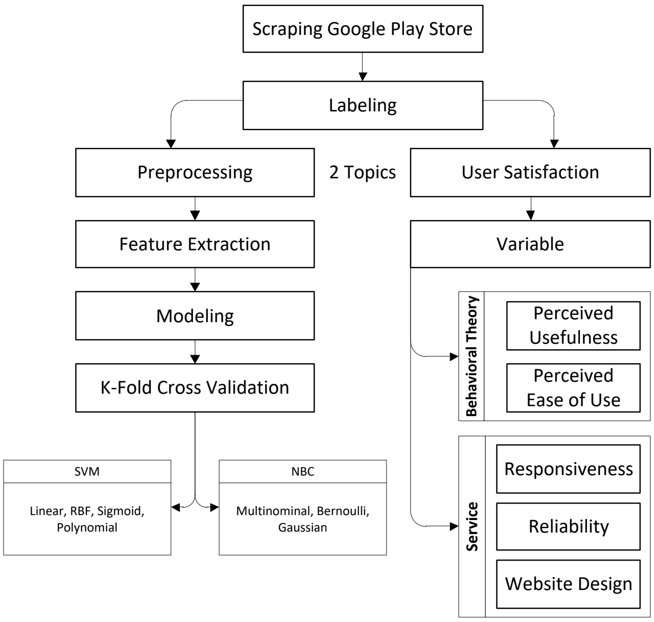

Figure 1. Research Procedure

1. Data Labeling

In this study, data labeling (annotation) of J&T Express application user reviews was carried out manually to facilitate the assessment of positive and negative comments. Annotations were carried out by annotators with experience as Indonesian language teachers. User reviews were marked as positive or negative and stored in a Comma Separated Values (CSV) document.

2. Preprocessing

The initial stages of text mining in changing the structure of irregular text data to become more structured [22]. According to the example in the book Tala, 2003 [23] preprocessing there are 5 processes consisting of Cleaning, Case Folding, Tokenizing, Filtering (Stop-word), Stemming.

a) Cleaning
The data obtained from scraping results will be cleaned and selected first to avoid repetitive data and empty data, such as neutral comment data that is not needed for this research.

b) Case Folding
The processed data is text from the results of data cleaning to convert all uppercase text to lowercase, remove urls, numbers, and symbols (including emojis).

c) Stopword Removal
A common word that is usually used in large quantities and is considered meaningless. Examples of stopwords in Indonesian are “yang”, “dan”, “di”, “dari”, etc.

d) Tokenizing
The process of dividing text into parts called tokens for later analysis. Words, numbers, symbols, punctuation, and other significant units can be considered as tokens.

e) Stemming
The process of mapping and decomposing word forms into their basic words. To perform Indonesian language stemming, we use the Sastrawi library which applies the Nazief and Adriani Algorithm to perform Indonesian language stemming.

3. Feature Extraction

Using the Term Frequency-Iverse Document Frequency (TF-IDF) weighting method, a weighting system that gives weight to each word in a document based on Term Frequency (TF) and Inverse Document Frequency (IDF).

$$W_{t,d} = W_{tf_{t,d}} \times idft \quad (1)$$

Wt ft,d : Term Frequency
id ft : Invers Dokument Frequency

4. Modeling

a) K-Fold Cross Validation
It is one of the techniques used to sort data into train and test data. In addition, K-Fold Cross Validation implements data splitting in equal numbers into k sub data sets. Division by repeating multiples of K-3, 6, 9 and 10 into sub-sets of data.

b) Support Vector Machine (SVM)
SVM is a machine learning method that aims to find the best hyperplane that separates two classes in the input space [8]. As in the previous explanation, the kernel parameters used in this study are Linear, RBF, Sigmoid and Poly-nomial. The equations of Linear kernel, RBF kernel, Polynomial kernel and Sigmoid kernel [14].

$$K\left(\vec{x_I}, \vec{x_J}\right) = \emptyset(\vec{x_I}) \times \emptyset\left(\vec{x_J}\right) \quad (2)$$

$$K\left(\vec{x_I}, \vec{x_J}\right) = \exp\left(-\frac{\left\|\vec{x_I} \times \vec{x_J}\right\|^2}{2\sigma^2}\right) \quad (3)$$

$$K\left(\vec{x_I}, \vec{x_J}\right) = \left(1 + x_I \times x_J\right)^d \quad (4)$$

$$K\left(\vec{x_I}, \vec{x_J}\right) = \tanh(\alpha x^T y + c) \quad (5)$$

x: expanded input vector
α: alpha
β: beta
σ: sigmoid

c) Naïve Bayes Classifier (NBC)
The NBC method is also one of the classification methods in text mining used in sentiment analysis and has good potential in terms of data computation classification and precision [1].

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (6)$$

X: input vector
C: a specific class
P(C|X): probability class based on known input vectors (posteriori probability)
P(c): probability class searched (prior probability) of the entire data
P(x|c): the probability of each input based on the condition of the class
P(x): probability of an input of the entire data

Multinomial, Bernoulli and Gaussian will be implemented in this study.

$$P(D|c) = \frac{Tc!}{\prod_{i=1}^{V}(x_i!)} \quad (7)$$

$$P(x_i|y) = p(x_i = 1|y)xi + \big(1 - p(x_i = 1|y)\big)(1 - x_i) \quad (8)$$

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}\right) \quad (9)$$

Tc: total number of words in documents of class c
i: the event
xi: holds binary value either 0 or 1 (Bernoulli)
p(xi=1|y): probability of feature xi appearing in class
σy2: variance of feature i in class
μy: mean of feature I in class

5. User Satisfaction Variables

User satisfaction is a measurement of an application that helps developers evaluate, later improve the quality of the application [24]. The TAM and Serv-qual methods are two methods that can understand user behavior and service quality in the context of technology and services [25]. Therefore, the main variables of the two methods will be taken to be used as a measurement of user satisfaction of the J&T Express application based on the results of sentiment analysis containing user comments or responses that will refer to each keyword from the comment search results. These variables include perceived usefulness, perceived ease of use, reliability, responsiveness and website design variables obtained from previous research [11].

B. Datasets

In this research, data is obtained from comments or reviews of J&T Express application users on the Google Play Store Indonesia with scraping techniques on python. The data used in a period of less than one year, namely, July 2023-December 2023, adjusts the final results of the survey (Top Brand Award and Populix) with a total data of 4,752 comments.

C. Tools

Google Colab was used as a tool for processing data in this study. It is a cloud-based Jupyter notebook that allows machine learning programming with the python language, which is known to be easy to learn and use [26]. In this process, several python libraries were used, such as Pandas, Numpy, Scikit-learn, and Google-play-scraper, to help with data retrieval and processing. In addition, Microsoft Excel was used to categorize words, which helped in measuring user satisfaction.

IV. RESULT AND DISCUSSION

A. Data Collection

The initial data collection stage for this research was carried out with google-play-scraper python on the Google Play Store of the J&T Express Indonesia application. Data was taken in Indonesian with a period of July 2023-December 2023, with a total of 4,752 comment data (Table 2). The comment data taken are reviews that have a star rating of 1 to 5 as in Figure 2. To determine whether a comment is positive or negative, it is done by adjusting the rating from the Google Play Store, where: Classifying ratings 4 and 5 as positive sentiment, ratings 1 and 2 as negative sentiment and rating 3 as neutral sentiment [14] [27]. Then the results of sentiment on each comment are obtained, namely the positive class totaling 434 comments and the negative class totaling 4,204 comments while the neutral class totals 114 comments as in Figure 2.

Table 2. Initial Data

No	Content (in Indonesian)	Score
1	j&T petir serang ngga jelas paket malah d retrun retrun terus kemana kurir nya ngapain d Adain pengiriman kalo ngga mau ngirim ke kitaa aneh	1
2	Jnt Kedung Tuban tolong donk ditegor itu adminya males banget dah paket sudah di kantor kagak di input.. lemot banget sejak ganti admin baru .	2
3	Mohon jnt melayani COD untuk usaha perorangan klu melalui pihak ketiga memberatkan pada pelanggan kami,karena ada biaya tambahan klu memakai pihak ketiga	5
…	…	…
4752	Lambat kecepatan pengirimanya.	1

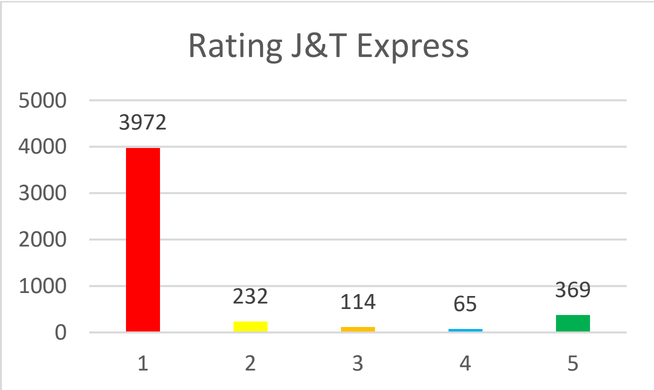

Figure 2. Aplication Rating of J&T Express

1. Comparison of Sentiment Results with Experts

The sentiment labeling annotation done manually by an expert has different results. By comparison, the rating generated from the application is not the same as the sentiment generated by the expert, the annotator presents different labeling results as in Table 3.

Table 3. Labeling Annotations by Experts

Sentiment	Google Play Store Rating Results	Changes	Expert Labeling Results
Positive	434	52 comments become neutral	387

		or negative	
Neutral	114	70 comments become negative or positive	56
Negative	4,204	11 comments become neutral	4,309

This gap occurs because users give views that do not match the reviews. Some users give a rating of 4 or 5, which should be positive, but the reviews written are mostly neutral in sentiment. Likewise, reviews with a score of 3, which should be neutral, are written with negative sentiment. For scores of 1 or 3, at least there are mistakes in giving negative sentiment reviews. The results of labeling annotations by this expert will be used as new data for future processes in this study, the data will be stored in CSV format and will be processed to the next stage.

B. Preprocessing

Previously labeled data will be processed for processing in classification algorithm modeling, which starts from cleaning, case folding, stopwords, tokenizing and stemming as well as word weighting.

1. Cleaning

The data will be cleaned and selected to avoid repetitive data and empty data, one of which is neutral sentiment data which will not be used because it does not have a strong argument or does not provide useful information for this research. Based on this, class reduction is done by eliminating neutral sentiment, from the initial data used 4,752 to 4,696 review data as shown in Table 4.

Table 4. Cleaning Data

No	Comments (in Indonesian)	Sentiment
1	j&T petir serang ngga jelas paket malah d retrun retrun terus kemana kurir nya ngapain d Adain pengiriman kalo ngga mau ngirim ke kitaa aneh	Negative
2	Jnt Kedung Tuban tolong donk ditegor itu adminya males banget dah paket sudah di kantor kagak di input.. lemot banget sejak ganti admin baru .	Negative
3	Mohon jnt melayani COD untuk usaha perorangan klu melalui pihak ketiga memberatkan pada pelanggan kami,karena ada biaya tambahan klu memakai pihak ketiga	Positive
…	…	…
4696	Lambat kecepatan pengirimanya .	Negative

2. Case Folding

The data processed as text from the data cleaning results will be first cleared, to convert the entire text of large letters into small letters, remove the url, numbers, and symbols (including emoji) as shown in Table 5.

Table 5. Case Folding Data

No	Comments (in Indonesian)	Sentiment
1	jt petir serang ngga jelas paket malah d retrun retrun terus kemana kurir nya ngapain d adain pengiriman kalo ngga mau ngirim ke kitaa aneh	Negative
2	jnt kedung tuban tolong donk ditegor itu adminya males banget dah paket sudah di kantor kagak di input lemot banget sejak ganti admin baru	Negative
3	mohon jnt melayani cod untuk usaha perorangan klu melalui pihak ketiga memberatkan pada pelanggan kamikarena ada biaya tambahan klu memakai pihak ketiga	Positive
…	…	…
4696	lambat kecepatan pengirimanya	Negative

3. Stopword Removal

Examples of stopwords in Indonesian are “yang”, “dan”, “di”, “dari”, etc. The purpose of using stopwords is to remove less informative words from the text, thus focusing on important words only as shown in Table 6.

Table 6. Stopword Data

No	Comments (in Indonesian)	Sentiment
1	jt petir serang ngga paket retrun retrun kemana kurir nya ngapain adain pengiriman kalo ngga ngirim kitaa aneh	Negative
2	jnt kedung tuban tolong donk ditegor adminya males banget dah paket kantor kagak input lemot banget ganti admin	Negative
3	mohon jnt melayani cod usaha perorangan klu ketiga memberatkan pelanggan kamikarena biaya tambahan klu memakai ketiga	Positive
…	…	…
4696	lambat kecepatan pengirimanya	Negative

4. Tokenizing

Words, numbers, symbols, punctuation marks and other important entities can be considered as tokens. In NLP, tokens are defined as “words” although tokenization can also be done on paragraphs or sentences as shown in Table 7.

Table 7. Tokenize Data

No	Comments (in Indonesian)	Sentiment
1	['jt', 'petir', 'serang', 'ngga', 'paket', 'retrun', 'retrun', 'kemana', 'kurir', 'nya', 'ngapain', 'adain', 'pengiriman', 'kalo', 'ngga', 'ngirim', 'kitaa', 'aneh']	Negative

No	Comments (in Indonesian)	Sentiment
2	['jnt', 'kedung', 'tuban', 'tolong', 'donk', 'ditegor', 'adminya', 'males', 'banget', 'dah', 'paket', 'kantor', 'kagak', 'input', 'lemot', 'banget', 'ganti', 'admin']	Negative
3	['mohon', 'jnt', 'melayani', 'cod', 'usaha', 'perorangan', 'klu', 'ketiga', 'memberatkan', 'pelanggan', 'kamikarena', 'biaya', 'tambahan', 'klu', 'memakai', 'ketiga']	Positive
…	…	…
4696	['lambat', 'kecepatan', 'pengirimanya']	Negative

5. Stemming

Stemming is the process of mapping and decomposing word forms into their basic words. To do Indonesian stemming, we can use the Python Sastrawi library that we have prepared at the beginning as shown in Table 8.

Table 8. Stemming Data

No	Comments (in Indonesian)	Sentiment
1	jt petir serang ngga paket retrun retrun mana kurir nya ngapain adain kirim kalo ngga ngirim kitaa aneh	Negative
2	jnt kedung tuban tolong donk ditegor adminya males banget dah paket kantor kagak input lot banget ganti admin	Negative
3	mohon jnt layan cod usaha orang klu tiga berat langgan kamikarena biaya tambah klu pakai tiga	Positive
…	…	…
4696	lambat cepat pengirimanya	Negative

6. TF-IDF

Data after the preprocessing stage is processed into numerical form using the TF-IDF weighting method. TF-IDF is a weighting system that gives weight to each word in a document based on term frequency (tf) and inverse document frequency (tf) (idf). In the TF-IDF stage, the calculation is done using python using the skicit-learn model.

C. Classification Modeling

Based on the research methodology described, the division of training data and testing data is taken from a dataset with a total of 4696 data after going through several pre-processing stages and the TF-IDF weighting stage. In this study, 4 comparison values will be used in the division of training data and testing data, namely K 3, 6, 9 and 10. Then the data will be used for testing using the SVM and NBC algorithms.

1. SVM Classification

The SVM algorithm with its kernel parameters has an influence on the accuracy performance results obtained. In this algorithm itself has four kernels namely Linear, Polynomial, Radial Basis Function (RBF), and Sigmoid. From each kernel, the highest accuracy performance will be sought to continue in the comparison process.

Table 9. SVM Kernel Accuracy Comparison

Kernel	K-Fold			
	K-3	K-6	K-9	K-10
Linear	95.06%	95.16%	95.23%	95.21%
RBF	93.96%	94.32%	94.28%	94.35%
Sigmoid	94.99%	95.01%	95.19%	95.08%
Polynomial	93.27%	93.31%	93.27%	93.31%

Based on repeated experiments carried out on each previous kernel with each K-repetition (Table 9), the best accuracy result is obtained for the Support Vector Machine (SVM) algorithm present in the Linear kernel, with an accurate value of 95.23%. Therefore, the support vector machine (SVM) with the Lineary kernel will be compared to the Naive Bayes Classifier (NBC) algoritm to determine the best.

2. NBC Classification

The naive bayes algorithm is divided into several types that are distinguished based on their functions. This algorithm itself has several derivatives including those used in this research are Multinomial Naive Bayes, Bernoulli Naive Bayes and Gaussian Naive Bayes. From each type of naive bayes, the highest accuracy performance will be sought to continue in the comparison process. For this algorithm, it will be compared to determine the best type of naïve bayes algorithm, then will continue in the same way on the SVM algorithm with K-Fold Cross Validation. As shown in Table 10, the comparison results where the Multinomial Naïve Bayes type is the best.

Table 10. Naïve Bayes Type Accuracy Comparison

Naive Bayes	Multinomial NB	Bernoulli NB	Gaussian NB
Accuracy	91.45 %	90.01 %	81.77 %

The Multinomial Naïve Bayes algorithm will be tested with the help of the same K-Fold, namely 3, 6, 9 and 10 repetitions of K that have been determined. This will be repeated until we get the best results in Table 11, which will then be compared with the SVM algorithm that has been trained before.

Table 11. K-Fold Accuracy Multinomial Naïve Bayes

K-Fold	Acuracy
K-3	91.05%
K-6	91.39%
K-9	91.43%
K-10	91.43%

From the results of K-Fold split data, K-9 is the highest with 91.43% accuracy. So that Multinomial Naïve Bayes K-9 is the best algorithm and will represent the NBC algorithm for comparison with the SVM algorithm.

3. Comparison of Result

The classification analysis process is first carried out by dividing training data and testing data by using K-Fold Cross Validation on 4696 comment data. Where in this study, previously the best parameters in each SVM and NBC algorithm will be determined so that these two algorithms will be compared to determine the best algorithm for data processing.

Table 12. Comparison of SVM and NBC Classification Algorithms

Model	K-*Fold*	Acuracy
Linear Support Vector Machine	K-9	95.23 %
Multinomial Naïve Bayes	K-9	91.43 %

In Table 12, it shows that the performance of the algorithm suitable for the J&T Express application comment review dataset on the Google Play Store is the Support Vector Machine Algorithm with a Linear kernel. With the same dataset and division of training and test data, SVM is slightly superior to NBC represented by Multinomial Naïve Bayes.

D. User Satisfaction

In this study, to measure user satisfaction from the J&T Express Indonesia application is carried out using variables of perceived usefulness, perceived ease of use, reliability, responsiveness and website design. Where the results of the sentiment analysis that has been done before, will be reviewed again to find words related to user satisfaction from the J&T Express application which will be calculated whether the user is satisfied or not fasting on the J&T Express application.

The first variable to be measured for user satisfaction is Perceived Usefulness, which contains user comments on how useful the J&T application is (Figure 3).

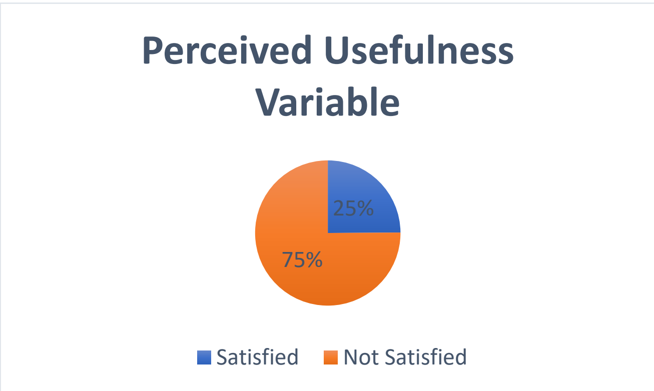

Figure 3. Perceived Usefulness Variable

The second variable to be measured for user satisfaction is Perceived Ease of Use, which contains user comments on how easy the J&T application is to use and run (Figure 4).

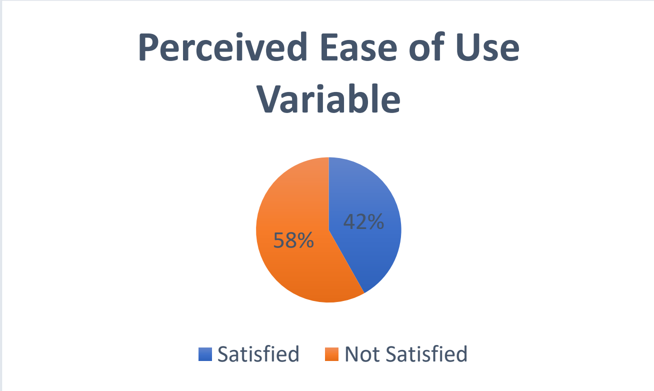

Figure 4. Perceived Ease of Use Variable

The third variable to be measured for user satisfaction is Reliability, which contains user comments on the quality of service provided in terms of reliability (Figure 5).

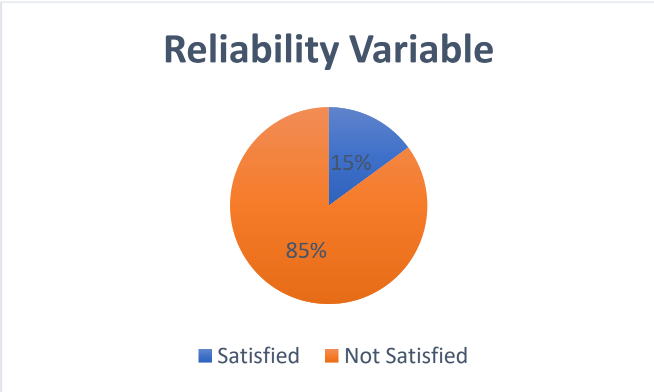

Figure 5. Reliability Variable

The fourth variable to be measured for user satisfaction is Responsiveness, which contains user comments on the quality of service provided in terms of responsiveness (Figure 6).

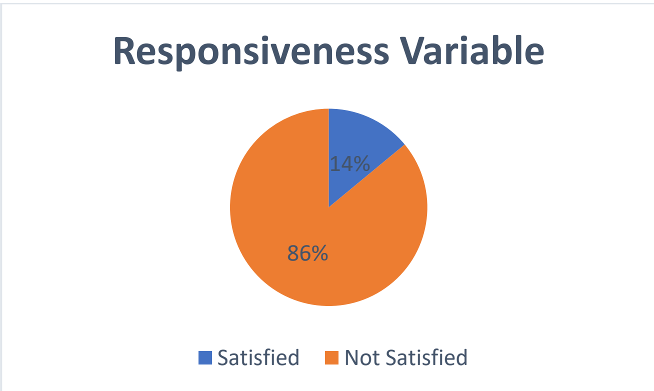

Figure 6. Responsiveness Variable

The fifth variable that will be measured for user satisfaction is Website Design, which contains user comments regarding the user experience of using the application (Figure 7).

DOI: 10.34148/teknika.v13i2.827

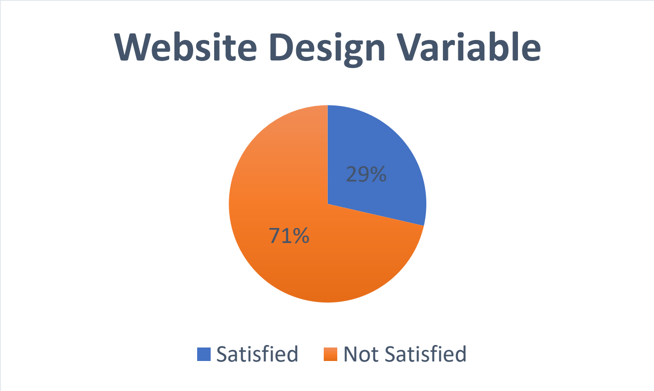

Figure 7. Website Design Variable

The results of the calculation of sentiment words on variables to measure user satisfaction beforehand, will be implemented in two classes, namely satisfied and dissatisfied to see the responses of the J&T Express Indonesia application users.

From the measurement results of each variable above, it can be concluded that users of the J&T Express Indonesia application gave unsatisfied responses in each variable measured, with an average response of 75% unsatisfied and 25% satisfied. So this is why the rating of the J&T Express Indonesia application given is very low even though the company has increased shipping volume and received many awards but is not in line with the services provided. Although from 4696 measured comment data only get 1154 comments related to user satisfaction based on variables. The remaining 3542 data from the comments contain other topics that are not related to user satisfaction, such as couriers, shipping and others. Can be seen in Figure 8.

Figure 8. Topic Comments J&T Express Application

Based on the statistics above, most positive comments are on the topic of user satisfaction with 267 comments out of 387 positive sentiment comments, and most negative comments are on the topic on delivery with 1,401 comments out of 4,309 negative sentiment comments. With the overall conclusion, the bad comments obtained by the J&T Express Indonesia application are in the section of bad and long delivery problems, even though user satisfaction only fills 25% of the comments but still produces a very high negative sentiment compared to the positive.

V. CONCLUSION

Sentiment analysis on app reviews is useful for gathering valuable feedback for app developers and can be a tool for assessing user satisfaction. The results of this analysis can help in understanding user needs and identifying areas of improvement, as well as giving new users an idea of what they can expect from the app. The application of classification algorithms, especially support vector machine with linear kernel, showed better accuracy than multinomial naive bayes algorithm. Testing using the K-Fold method with K-9 achieved an accuracy of 95.23%. This shows that the right parameters and algorithms can significantly improve the reliability of sentiment analysis. On the other hand, the review results show a low level of user satisfaction, with the service indicators having many negative sentiments and dissatisfaction levels reaching more than 80% while the behavioral theory indicators have dissatisfaction above 60%. This emphasizes the need for improvements in certain aspects of the app to enhance user experience. In future research, it would be better to expand the range of data and try other classification algorithms to increase the accuracy and improve the quality of sentiment analysis.

REFERENCE

[1] J. Khitamasari, I. F. Hanif, P. Studi, and T. Informatika, "Analisis Sentimen Ulasan User Terhadap Aplikasi Tix-Id Menggunakan Metode Naïve Baye s Pada Google Play Store," vol. 23, no. 6, pp. 81–83, 2023.

[2] F. Z. Shaumi and E. R. Cahyadi, "Perbandingan Kepuasan Pelanggan Terhadap Kualitas Layanan Logistik Pada JNE dan J&T Express di Jabodetabek," *J. Apl. Bisnis dan Manaj.*, vol. 8, no. 3, pp. 939–950, 2022, doi: 10.17358/jabm.8.3.939.

[3] N. Aula, M. Ula, and L. Rosnita, "Analisis Sentimen Review Customer Terhadap Perusahaan Ekspedisi Jne, J&T Express Dan Pos Indonesia Menggunakan Metode Support Vector Machine (Svm) Analysis of Customer Review Sentiment To Jne, J&T Express and Pos Indonesia Expedition Companies Using Svm M," *J. Informatics Comput. Sci.*, vol. 9, no. 1, pp. 81–86, 2023.

[4] A. Nurjanah, N. Aulia, E. R. Rasywir, and A. Feranika, "Analisis Tingkat Kepuasan Konsumen Terhadap Layanan Ekspedisi J&T Express Dengan Metode Algoritma Naive Bayes," *J. Manaj. Teknol. Dan Sist. Inf.*, vol. 2, no. 1, pp. 78–88, 2022, doi: 10.33998/jms.2022.2.1.39.

[5] V. Elvira and H. K. Dewi, "No Title," *Kontan.co.id*, 2023. https://industri.kontan.co.id/news/volume-pengiriman-jt-express-tumbuh-40-sepanjang-2022

[6] M. A. Syainal *et al.*, "Analisis User Experience Pada

Aplikasi J & T EXPRESS Menggunakan Metode Heart Matrics .," *J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. Vol. 12, no. 01, pp. 367–375, 2023.

[7] O. D. Mawarni and S. W. Adi, "Pengaruh Fasilitas Sistem Pelacakan Online, Ketepatan Waktu Pengiriman, Harga Dan Kualitas Pelayanan Terhadap Kepuasan Pelanggan J&T Express (Studi Pada J&T Express Cabang Kota Karanganyar)," *Eqien - J. Ekon. dan Bisnis*, vol. 10, no. 1, pp. 445–450, 2022, [Online]. Available: https://stiemuttaqien.ac.id/ojs/index.php/OJS/article/view/556/372

[8] E. Suryati, Styawati, and A. A. Aldino, "Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM)," vol. 4, no. 1, pp. 96–106, 2023.

[9] P. Arsi, R. Wahyudi, and R. Waluyo, "Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 231–237, 2021, doi: 10.29207/resti.v5i2.2698.

[10] M. Rodríguez-Ibánez, A. Casánez-Ventura, F. Castejón-Mateos, and P.-M. Cuenca-Jiménez, "A review on sentiment analysis from social media platforms," *Expert Syst. Appl.*, vol. 223, no. March, p. 119862, 2023, doi: 10.1016/j.eswa.2023.119862.

[11] A. B. Elciano, D. R. Atmaja, R. Rojuaniah, I. L. H. Hafidz, and F. Trianda, "Pengaruh Faktor Kualitas Layanan Aplikasi J&T Express terhadap Kepuasan Pelanggan, Niat Beli, dan Pembelian Aktual," *J. Inform. Ekon. Bisnis*, vol. 5, no. No. 3, pp. 844–851, 2023, doi: https://doi.org/10.37034/infeb.v5i4.673.

[12] U. Rastryana, U. Masahere, and S. Rusmayanti, "Analisis Ipa Dan Csi Terhadap Kepuasan Pelanggan Market Place Pada Pelayanan J&T Express Indonesia," *Akrab Juara J. Ilmu-ilmu Sos.*, vol. 7, no. 2, p. 41, 2022, doi: 10.58487/akrabjuara.v7i2.1814.

[13] U. Amelia *et al.*, "IMPLEMENTASI ALGORITMA SUPPORT VECTOR MACHINE (SVM) UNTUK PREDIKSI PENYAKIT STROKE DENGAN ATRIBUT BERPENGARUH," vol. III, pp. 254–259, 2022.

[14] P. Aditiya, U. Enri, and I. Maulana, "Analisis Sentimen Ulasan Pengguna Aplikasi Myim3 Pada Situs Google Play Menggunakan Support Vector Machine," vol. 9, no. 4, pp. 1020–1028, 2022, doi: 10.30865/jurikom.v9i4.4673.

[15] R. N. Sari and R. S. Hayati, "Penerapan Metode Multi Attribute Utility Theory (MAUT) Dalam Pemilihan Rumah Kost," *J-SAKTI (Jurnal Sains Komput. dan Inform.*, vol. 3, no. 2, p. 243, 2019, doi: 10.30645/j-sakti.v3i2.144.

[16] T. Astuti and Y. Astuti, "Analisis Sentimen Review Produk Skincare Dengan Naïve Bayes Classifier Berbasis Particle Swarm Optimization (PSO)," vol. 6, pp. 1806–1815, 2022, doi: 10.30865/mib.v6i4.4119.

[17] A. Novantika and Sugiman, "Analisis Sentimen Ulasan Pengguna Aplikasi Video Conference Google Meet menggunakan Metode SVM dan Logistic Regression," *Prism. Pros. Semin. Nas. Mat.*, vol. 5, pp. 808–813, 2022, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/prisma/

[18] D. Era, S. Andryana, and A. Rubhasy, "Perbandingan Algoritma Naïve Bayes Dan K-Nearest Neighbor pada Analisis Sentimen Pembukaan Pariwisata Di Masa Pandemi Covid 19," *J. Sains Komput. Inform. (J-SAKTI*, vol. 7, no. 1, pp. 263–272, 2023.

[19] M. Raffi, A. Suharso, I. Maulana, and U. S. Karawang, "Analisis sentimen ulasan aplikasi binar pada google play store menggunakan algoritma naïve bayes sentiment analysis of binar application reviews on google play store using naïve bayes algorithm," vol. 6, 2023.

[20] R. Adityawan, A. Gunawan, and G. L. Ginting, "Analisis Kepuasan Penggunaan Aplikasi Pospay Menerapkan Metode TAM," vol. 1, no. 2, pp. 57–62, 2023.

[21] N. Charoonsri *et al.*, "Analisis Kepuasan Pelanggan Terhadap Pelayanan J & T Express Saribudolok Dengan Metode Servqual," vol. 26, no. 1, pp. 44–49, 2024.

[22] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government Pada Google Play Menggunakan Algoritma Naïve Bayes," vol. 9, no. 2, pp. 785–795, 2022.

[23] F. Z. Tala, "A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia," 2003.

[24] R. Sholihah and A. D. Indriyanti, "Analisis Kepuasan Pengguna Aplikasi Camscanner Menggunakan Metode Technology Acceptance Model (TAM) dan End-User Computing Satisfaction (EUCS)," vol. 03, no. 03, pp. 102–109, 2022.

[25] P. Hadroh, M. Maheis, and I. Farida, "Jurnal Mirai Management Pengaruh Technology Acceptance Model (TAM) dan Kualitas Layanan Terhadap Minat Menggunakan Aplikasi Flip di Jawa Tengah," *J. Mirai Manag.*, vol. 8, no. 2, pp. 408–421, 2023, [Online]. Available: https://databoks.katadata.co.id/

[26] A. Rianti, N. Wachid, A. Majid, and A. Fauzi, "Machine Learning Journal Article Recommendation System using Content based Filtering," *JUTI J. Ilm. Teknol. Inf.*, vol. 22, no. 1, pp. 1–10, 2024.

[27] R. Maulana, A. Voutama, and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC," *J. Teknol. Terpadu*, vol. 9, no. 1, pp. 42–48, 2023, doi: 10.54914/jtt.v9i1.609.