IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.20, No.1, January 2026, pp. 83~94
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: https://doi.org/10.22146/ijccs.110409

Sentiment Classification of MyTelkomsel Reviews Using SVM and Logistic Regression

Rijal Bagus Adinata*[1], Supriyono[2], Diana Laily Fithri [3]
[1,2,3]Department of Information System, Faculty of Engineering, Muria Kudus University, Kudus, Indonesia
e-mail: ***[1]202153112@std.umk.ac.id**, [2]supriyono.si@umk.ac.id, [3]diana.laily@umk.ac.id

Abstract

Dalam memungkinkan pengguna mengakses informasi dan fitur layanan pada satu platform dibutuhkan penyedia layanan telekomunikasi. Aplikasi MyTelkomsel sebagai layanan digital dari penyedia telekomunikasi yang memiliki banyak pengguna di Indonesia. Namun, ulasan pengguna memiliki berbagai tanggapan, mulai dari apresiasi hingga keluhan terkait performa aplikasi dan layanan pelanggan. Studi ini dimaksudkan untuk mengevaluasi peforma Support Vector Machine (SVM) dan Logistic Regression dalam menganalisis sentimen. Data dikumpulkan dari Google Play Store dan melalui serangkaian tahap prapemrosesan, termasuk pembersihan data, case folding, normalisasi, tokenisasi, stopword removal, serta stemming. Proses ekstraksi fitur menggunakan pendekatan TF-IDF, sementara penilaian performa model didasarkan pada metrik akurasi, presisi, recall, F1-score, dan Area Under Curve (AUC). Hasil menunjukkan bahwa SVM dan Regresi Logistik memiliki nilai presisi, recall, dan skor F1 yang hampir sama bedasarkan weighted average. Namun dibanding dengan Logistic Regression, SVM lebih unggul karena memiliki nilai dari akurasi sebesar 93,36%, Nilai dari AUC 0,9680, dan recall untuk kelas positif sebesar 82%. Oleh karena itu, model SVM dapat dianggap sebagai algoritma yang lebih optimal untuk mengklasifikasikan sentimen

Kata kunci*—Analisis Sentimen, MyTelkomsel, Logistic Regression, Support Vector Machine, TF-IDF*

Abstract

Telecommunication service providers are needed to enable users to access information and service features on a single platform. The MyTelkomsel application is a digital service from a telecommunications provider that has many users in Indonesia. However, user reviews have various responses, ranging from appreciation to complaints regarding application performance and customer service. This study aims to evaluate the performance of Support Vector Machine (SVM) and Logistic Regression in analyzing sentiment. Data was collected from the Google Play Store and went through a series of preprocessing stages, including data cleaning, case folding, normalization, tokenization, stopword removal, and stemming. The feature extraction process uses the TF-IDF approach, while the model performance assessment is based on accuracy, precision, recall, F1-score, and Area Under Curve (AUC) metrics. The results show that SVM and Logistic Regression have almost the same precision, recall, and F1-score values based on the weighted average. However, compared to Logistic Regression, SVM is superior because it has an accuracy value of 93.36%, an AUC value of 0.9680, and a recall for the positive class of 82%. Therefore, the SVM model can be considered as a more optimal algorithm for classifying sentiment.

Keywords— *Sentiment Analysis, MyTelkomsel, Logistic Regression, Support Vector Machine, TF-IDF*

Received December 19th,2025; Revised January 21th, 2026; Accepted January 28th, 2026

1. INTRODUCTION

In allowing users to access information and service features on a single platform, a telecommunication service provider is required [1]. In the development of digital technology, service provider applications play a crucial role as the primary medium for accessing services effectively and efficiently. One of the most popular telecommunications service providers in Indonesia is MyTelkomsel, with downloads reaching 50 million users[2].

With this application, users can display their mobile phone number, remaining internet quota, credit, and other information services on the main page [3]. In addition, MyTelkomsel app users can purchase credit, internet data, and redeem points. This leads to several issues that need to be addressed regarding slow quota updates, network performance, disrupted transaction processes, and a less responsive interface.

In this study, the MyTelkomsel application was chosen to conduct sentiment analysis on user reviews on the Google Play Store. These reviews are used as an important source of information so that they can assess the quality and performance of an application. Sentiment analysis can identify emotions in written content with the aim of categorizing opinions based on their polarity, specifically distinguishing between positive and negative sentiments. In sentiment analysis, a technique that is often used is Text Mining to extract valuable information from unstructured text data [4].

In a previous study, Ningsih et al. found that sentiment analysis in three popular mobile applications in Indonesia using Support Vector Machine (SVM) was superior to Naïve Bayes Classification (NBC) in classifying sentiment [5]. In another study, Hardiansyah et al. and Fauzan et al. showed that in classifying MyTelkomsel reviews with SVM, there was a high level of accuracy [6] [7].

Additional research that uses machine learning in conducting sentiment analysis. Sentiment analysis in research conducted by Maulana et al. related to retail applications shows that Logistic Regression excels in positive sentiment classification with the highest F1-score value, but the accuracy level of Logistic Regression is the same as Multinomial Naïve Bayes, SVM, and K-Nearest Neighbor (KNN) [8]. The research conducted by Budianto et al. also found that Logistic Regression and SVM, with the number of misclassification errors and the accuracy of sentiment classification [9]. Meanwhile, research conducted by Husein et al. found that sentiment analysis on Google Play reviews using the Naïve Bayes and SVM methods showed that SVM provides better precision and recall values in detecting negative sentiment than Naïve Bayes [10] . In the research conducted by Putri et al., it was revealed that the sentiment analysis of public opinion towards Nadiem Makarim showed that SVM with an optimal kernel in analyzing public opinion [11]. Another finding, in the public tweets of public figures in Sinaga and Aji's research, showed that Logistic Regression had stable performance [12]. Meanwhile, research conducted by Aditiya et al. found that the classification of BMKG Info application reviews with Naïve Bayes is superior to SVM [13]. In a study conducted by As'ari et al., the classification of Spotify application reviews with Naïve Bayes has high accuracy [14].

Based on previous studies, sentiment analysis on MyTelkomsel app user reviews using SVM effectively outperforms reviews from Naïve Bayes (NB), Random Forest (RF), and Gradient Boosting [5][6][7]. However, another study, sentiment analysis using Logistic Regression on other objects, outperformed SVM [9]. Therefore, this study is focused on conducting sentiment analysis on MyTelkomsel application user reviews by directly comparing the performance of the two commonly used machine learning algorithms, namely Logistic Regression and Support Vector Machine (SVM). This study uses a medium-sized dataset taken from user reviews on the Google Play Store in April 2025. This study uses a medium-sized dataset taken from user reviews on the Google Play Store in April 2025. This study is intended to compare SVM and Logistic Regression on the MyTelkomsel application by evaluating performance using various metrics. The metrics used in the comparison between SVM and Logistic regression were F1-score, accuracy, precision, recall, and Area Under the Curve (AUC). This was done to assess the effective performance between the SVM model and Logistic Regression in classifying positive and negative sentiments.

This research was carried out in several stages. In the initial stage, data collection was carried out. The next stage is pre-processing of the text. In pre-processing, it is divided into several stages. At first, pre-processing is carried out by cleaning. The next stage is tokenization. After tokenization, normalization is carried out. After normalization, the stopword was removed. The last stage of the pre-processing process is the stemming. After pre-processing, TF-IDF (Term Frequency-Inverse Document Frequency) vectorization is carried out. The next stage is sentiment labeling. After labeling, model training is carried out. In the last stage, performance evaluation is carried out on the model. In the evaluation of model performance, the metrics used in the comparison between SVM and Logistic regression were F1-score, accuracy, precision, recall, and Area Under the Curve (AUC). This aims to provide insight into the most effective algorithm to calculate MyTelkomsel app reviews between SVM and Logistic Regression, so that it is expected to contribute to the development of a more efficient and accurate sentiment analysis model in the future. In addition, it is hoped that developers will find the research results to provide practical benefits. It is also hoped that service providers can understand user sentiment automatically so that it can be the basis for decision-making in an effort to improve service quality, user experience, and customer satisfaction.

2. METHODS

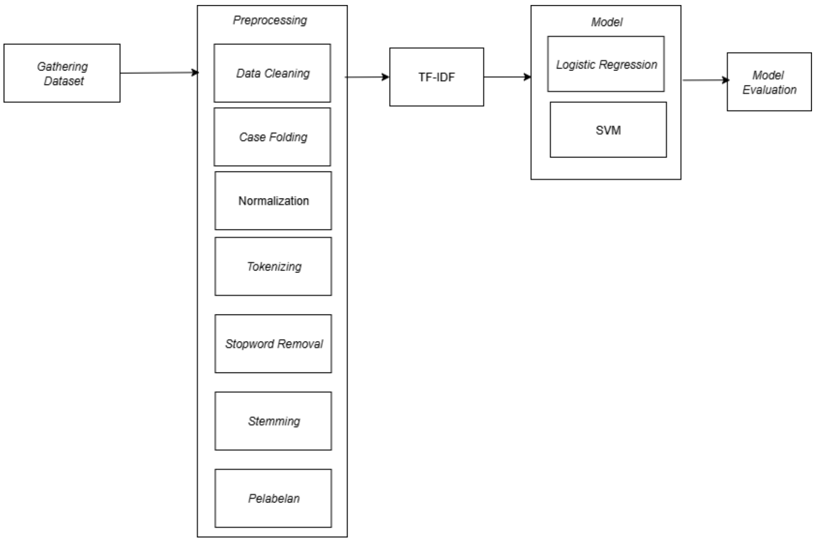

Figure 1 Sentiment analysis research technique, SVM and Logistic Regression in MyTelkomsel

An illustration of the methodological framework shown in Figure 1, the process of analyzing the sentiment of the Mytelkomsel application using SVM and Logistic Regression will be carried out at Google Colaboratory. The discussion of the stages will be explained in detail in the next section.

2.1 Data Gathering

In this study, the initial stage was data collection. This data collection was used to support the entire analysis process. A total of 7,000 reviews were collected in April from the Google Play Store using the Google Play Scraper library. The library was run in the Google Colaboratory development environment.

Table 1 Shows a example of MyTelkomsel application user review data

Username	Rating	Comment	Review Date
Pengguna Google	5	semua di My Telkomsel sat set sat set #Luarbiasa @A213	27/04/2025
Pengguna Google	3	kenapa udah 2 bulan ini jaringannya buruk banget saya di kota pekanbaru susah dapat internet laju, ada aja dengan telkomsel	18/04/2025
Pengguna Google	2	lag mulu padahal jaringan full apk aneh	18/04/2025

2. 2 Data Preprocessing

Preprocessing there are several stages that are used.The following are the stages in prepocessing.

2. 2.1 Data Cleaning

This data cleaning process aims to reduce noise and produce cleaner text. Data cleaning removes mentions, usernames, numbers, non-alphanumeric characters, and various emojis and symbols. An illustration of this process is presented in Table 2.

Table 2 Shows a sample of user review data following the data cleaning process

Comment	Cleaning
semua di My Telkomsel sat set sat set #Luarbiasa @A213	semua di My Telkomsel sat set Luarbiasa
kenapa udah 2 bulan ini jaringannya buruk banget saya di kota pekanbaru susah dapat internet laju, ada aja dengan telkomsel	kenapa udah bulan ni jaringannya buruk banget saya di kota pekanbaru susah dapat internet laju ada apa dengan telkomsel
lag mulu padahal jaringan full apk aneh	lag mulu padahal jaringan full apk aneh

2. 2.2 Case Folding

The case folding stage involves converting all letters in the review text to lowercase. This step aims to ensure consistency in word representation, as in text analysis, the difference between capital and lowercase letters can affect the results of data processing. This process is applied to the text cleanup results and stored in a new text form, which is entirely lowercase, as the basis for the next preprocessing stage. The series of processes is shown in Table 3.

Table 3 Shows the results of data normalization through case folding

Cleaning	Case_Folding
semua di My Telkomsel sat set Luarbiasa	semua di my telkomsel sat set luarbiasa
kenapa udah bulan ni jaringannya buruk banget saya di kota pekanbaru susah dapat internet laju ada apa dengan telkomsel	kenapa udah bulan ni jaringannya buruk banget saya di kota pekanbaru susah dapat internet laju ada apa dengan telkomsel
lag mulu padahal jaringan full apk aneh	lag mulu padahal jaringan full apk aneh

2. 2.3 Normalization

At the normalization stage, non-standard words or slang terms in the text are changed into standard forms. This process is carried out by matching each word in the case folding result to the standard word dictionary reference. If a suitable form is found, the word is replaced with its official form. This stage aims to improve the uniformity of language use in the data, resulting in cleaner text that is ready for analysis at a later stage. The sequence of processes is shown in Table 4.

Table 4 Example of case folding applied post-normalization

Case_Folding	Normalization
semua di my telkomsel sat set luarbiasa	semua di my telkomsel sat set Luarbiasa
kenapa udah bulan ni jaringannya buruk banget saya di kota pekanbaru susah dapat internet laju ada apa dengan telkomsel	kenapa sudah bulan nih jaringannya buruk sangat saya di kota pekanbaru susah dapat internet laju ada apa dengan telkomsel
lag mulu padahal jaringan full apk aneh	lag mulu padahal jaringan full aplikasi aneh

2. 2.4 Tokenizing

In this study, tokenization was performed by separating words with spaces, allowing each word in the text to be identified individually. In the next analysis stage, the tokenization results are saved in a new format. The sequence of these processes is shown in Table 5.

Table 5 Example of tokenization applied post-normalization

Normalization	Tokenizing
semua di my telkomsel sat set Luarbiasa	['semua', 'di', 'my', 'telkomsel', 'sat', 'set', 'luarbiasa']
kenapa sudah bulan nih jaringannya buruk sangat saya di kota pekanbaru susah dapat internet laju ada apa dengan telkomsel	['kenapa', 'sudah', 'bulan', 'nih', 'jaringannya', 'buruk', 'sangat', 'saya', 'di', 'kota', 'pekanbaru', 'susah', 'dapat', 'internet', 'laju', 'ada', 'apa', 'dengan', 'telkomsel']
lag mulu padahal jaringan full aplikasi aneh	['lag', 'mulu', 'padahal', 'jaringan', 'full', 'aplikasi', 'aneh']

2. 2.5 Stopword Removal

In stopword removal, only words with significant information weight are retained. Therefore, common words that do not have a significant contribution to the meaning of the text, such as conjunctions and pronouns, are removed. In addition, the stopword list is adjusted to the Indonesian context. The process sequence is shown in Table 6.

Table 6 Tokenized text sample after removing stopwords

Tokenizing	Stopword Removal
['semua', 'di', 'my', 'telkomsel', 'sat', 'set', 'luarbiasa']	['my', 'telkomsel', 'sat', 'set', 'luarbiasa']
['kenapa', 'sudah', 'bulan', 'nih', 'jaringannya', 'buruk', 'sangat', 'saya', 'di', 'kota', 'pekanbaru', 'susah', 'dapat', 'internet', 'laju', 'ada', 'apa', 'dengan', 'telkomsel']	['nih', 'jaringannya', 'buruk', 'kota', 'pekanbaru', 'susah', 'internet', 'laju', 'telkomsel']
['lag', 'mulu', 'padahal', 'jaringan', 'full', 'aplikasi', 'aneh']	['lag', 'mulu', 'jaringan', 'full', 'aplikasi', 'aneh']

2. 2.6 Stemming

At this stage, the Sastrawi Stemmer library is applied for the stemming process, which is useful for generating basic forms for transforming words in text. This process can result in some empty data, which needs to be removed. This process is illustrated in Table 7.

Table 7 Result of stopword elimination and stemming

Stopword Removal	Stemming
['my', 'telkomsel', 'sat', 'set', 'luarbiasa']	my telkomsel sat set luarbiasa
['nih', 'jaringannya', 'buruk', 'kota', 'pekanbaru', 'susah', 'internet', 'laju', 'telkomsel']	nih jaring buruk kota pekanbaru susah internet laju telkomsel
['lag', 'mulu', 'jaringan', 'full', 'aplikasi', 'aneh']	lag mulu jaring full aplikasi aneh

2. 2.6 Labeling

At this stage, the initial total of reviews stemming to 6,564. In labeling, reviews were categorized as positive or negative sentiment based on the rating value, with ratings 1–2 considered negative, 4–5 considered positive, and rating 3 excluded because it was considered ambiguous and could improve classification accuracy [5]. This approach was applied to simplify the data, resulting in 6,324 labeled review data. The results of sentiment grouping based on rating values are presented in Table 8.

Table 8 Example of labeled data based on rating values

Stemming	Rating	Label
my telkomsel sat set luarbiasa	5	positive
nih jaring buruk kota pekanbaru susah internet laju telkomsel	3	negative
lag mulu jaring full aplikasi aneh	2	negative

2. 3 TF-IDF

The TF-IDF method is applied to assign weights to words based on their frequency in a document, which is useful for converting text data into numeric vectors[15]. Not only from its frequency in a single document, but also from how often the word appears in the corpus. Therefore, TF-IDF in sentiment analysis can help the model focus on the most important words. The TF calculation formula can be seen in Equation (1):

$$tf_{t,d} = \frac{n_{t,d}}{(Total\ number\ of\ term\ in\ document)} \tag{1}$$

The next step is to determine the IDF (Inverse Document Frequency) to determine the significance of a word. The IDF formula is shown in Equation (2):

$$Idf = log\left(\frac{number\ of\ document}{Total\ number\ of\ term\ in\ document}\right) \tag{2}$$

The obtained TF and IDF values can be used to calculate the TF-IDF score by multiplying TF by IDF. This score indicates the relevance of a term in a document to the entire corpus. The TF-IDF formula is expressed in Equation (3):

$$tfidf_{t,d} = tf_{t,d} \times idf_d \tag{3}$$

2. 4 Models

In this study, the machine learning techniques used were Support Vector Machine (SVM) and Logistic Regression.

2. 4.1 SVM

This algorithm works to find a hyperplane that separates classes with the most significant margin by transforming the data into a higher-dimensional space [16]. Therefore, SVM is very effective in handling TF-IDF results. This process aims to maximize the distance margin between the two classes, thereby increasing the accuracy and generalization of the model. The SVM calculation formula is given in Equation (4).

$$f(x) = sign\ (w \cdot x + b) \tag{4}$$

2. 4.2 Logistic Regression

This model predicts the probability of the target variable being in a certain class[17]. The prediction process is carried out by calculating the weights of the text extraction features, then adding them up in the form of logit values, and then entering them into the sigmoid function to make the probability value range from 0 to 1. The P(Y) formula can be seen in equation (5).

$$P(Y) = \frac{1}{1 + e^{-(\beta 0+\beta 1\ X1\ +\ \beta 2\ X2+\cdots+\beta 8X8)}} \tag{5}$$

2. 5 Model Evaluation

In model performance evaluation, the metrics used in the comparison between SVM and logistic regression are F1 score, accuracy, precision, recall, and Area Under the Curve (AUC).

2. 5.1 Accuracy

Accuracy is used to evaluate how often a classification algorithm produces correct results. Accuracy measurement can be obtained by calculating the proportion of correct results relative to the total number of predictions, as described in Equation (6).

$$Accuracy = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \tag{6}$$

2. 5.2 Precision

Precision is useful for accurately identifying the true positive class from all positive predictions. Thus, if the precision of the model is high, the error in predicting data as positive will be less frequent. The calculation of precision is presented in Equation (7).

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{7}$$

2. 5.3 Recall

The recall metric measures the model's ability to detect all true positive data. Equation (8) describes the calculation for recall.

$$Recall = \frac{\text{TP}}{\text{TP+FN}} \quad (8)$$

2. 5.1 F1 Score

F1-Score is used as a measure of classification model performance by considering precision and recall values in a balanced manner. Equation (9) describes the calculation for the F1 score.

$$F1 - Score = 2 - \frac{\text{Precission} \times \text{Recall}}{\text{Precission} + \text{Recall}} \quad (9)$$

2. 5.1 Area Under the Curve (AUC)

The function of the AUC matrix assesses the binary classification performance, which is derived from the area under the ROC (Receiver Operating Characteristic) curve. Therefore, the AUC value can distinguish between positive and negative categories, with AUC values close to 1 reflecting strong ability. On the other hand, a value of 0.5 means that the model performs similarly to random selection.

3. RESULTS AND DISCUSSION

This section discusses the results of performance testing between SVM and Logistic Regression models, which will be discussed in several stages. In the first stage, discuss the extraction results of TF-IDF. In the second stage, it discusses the comparison of models in analyzing sentiment on MyTelkomsel, measured through the results of model performance and analysis errors. The final stage discusses the visualization of the sentiment distribution, which can explain the revealed patterns.

3. 1 Feature Extraction

The SVM and Logistic Regression classification models are used to predict the sentiment of MyTelkomsel service reviews from the Play Store. For both models to process review data properly, a numerical representation of the text is required. Therefore, the TF-IDF feature extraction technique is used because it can emphasize words that have significant meaning in a particular context, namely. These words often appear in one review but rarely appear in another. In this implementation, TF-IDF is used with default parameters, allowing only a single word (unigram) to be considered, and all words that appear at least once in the document are included as features, with no limitation on maximum frequency. With this representation, the model can identify word patterns that contribute to sentiment polarity. Figure 2 displays the most relevant terms based on TF-IDF scores.

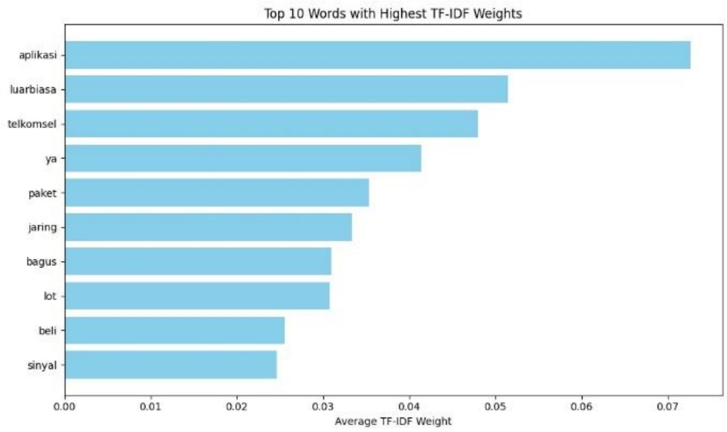

Figure 2 Ten words with the highest TF-IDF weight

3. 2 Model Comparison And Error Analysis

In this section, the results of training SVM and Logistic Regression models with a ratio in the training data of 80% and a ratio in the test data of 20% will be compared to the performance of the model through the measurement of several matrices. In addition, the SVM and Logistic Regression model classification errors are analyzed. This stage aims to find out which model is superior in classifying reviews.

3. 2.1Performance Models

In the model performance assessment stage, measurements are made using several evaluation metrics such as F1-score, accuracy, precision, recall, and AUC. This aims to identify the advantages of each model in overcoming class distribution imbalance. The results of the model performance comparison are shown in Table 9 below.

Table 9 Evaluation of model performance based on weighted average

Metric	Logistic Regression	SVM
Accuracy	93,12%	93,37%
Precision	94%	94%
Recall	93%	94%
F1 Score	93%	93%

It is known from Table 9's assessment findings that the two models, SVM and Logistic Regression, have the same precision value, which is 94%, as well as identical recall and F1-score values, each of 93%, based on the weighted average. This similarity suggests that both models have a balanced classification performance in handling the distribution of data between classes. However, when reviewed from the results of the classification per class, there is a significant difference in performance. A more detailed result of the classification per class can be seen in Table 11, which presents the precision, recall, and F1-score values of each class, as well as opening up opportunities for a more in-depth analysis of the strengths and weaknesses of each model.

Table 10 Evaluate Model Performance Per Class

Type	Class	Precision	Recall	F1 Score
SVM	Positive	96%	82%	89%
	Negative	92%	98%	95%
Logistic Regression	Positive	99%	79%	88%
	Negative	91%	100%	95%

In Table 10, Evaluation of Model Performance per Class, above, the total accuracy value obtained by SVM is 93.36%, slightly higher than that of Logistic Regression, which reaches 93.12%. The recall value for SVM shows 82%. This value is higher than the Logistic Regression value, which is 79%. This shows that SVM is better able to recognize positively labeled data. However, in the positive class, the precision value owned by Logistic Regression, with a value of 99% is superior to SVM, which is worth 96% in the positive class. This shows that Logistic Regression has a higher level of accuracy in predicting positive classes.

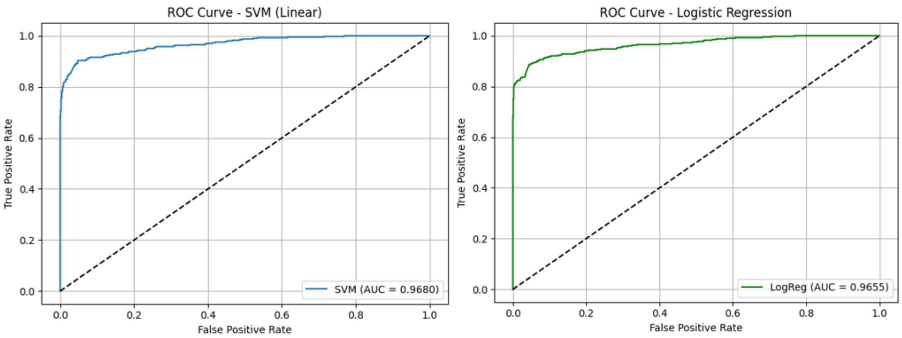

Figure 3 AUC ROC curve SVM and Logistic Regression

The ROC curve in Figure 3 above illustrates the relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) at various threshold values. The AUC value is obtained by calculating the ROC curve. In the SVM model, the AUC value is 0.9680. Meanwhile, in the Logistic Regression model, the value is 0.9655. With AUC as a statistic that can measure how well a model distinguishes between positive and negative results, the small difference between SVM and Logistic Regression indicates the superior generalization ability of the SVM in accurately distinguishing the two groups.

3. 2.2 Confusion Matrix

The Confusion Matrix is crucial for gaining a deeper understanding of a model's classification performance. It can also provide a detailed overview of the model's ability to correctly classify each sentiment class (negative and positive). Furthermore, the Confusion Matrix can be used to identify sentiment misclassifications. This matrix uses the predicted distribution of true positives, true negatives, false positives, and false negatives to evaluate the results. In this study, the SVM model correctly identified 848 reviews with negative sentiment (true negatives), while correctly classifying 333 positive reviews as such (true positives). This confusion matrix indicates that the SVM model is generally effective in detecting negative reviews, but still struggles to consistently recognize all positive responses. A detailed summary of these classification results is illustrated in the confusion matrix in Figure 4.

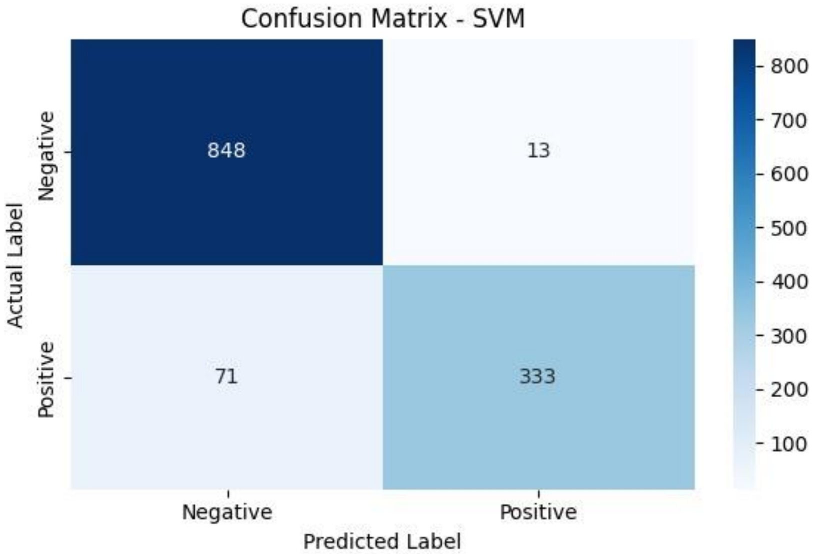

Figure 4 SVM matrix confusion

The Logistic Regression model can classify most of the data correctly. A total of 859 negative data points were correctly predicted as true negatives, and 319 positive data points were correctly recognized as true positives. The misclassification rate is minimal in the negative class, specifically, there are only 2 data points that are incorrectly predicted as positive (false positives). However, there are still 85 positive data points that are indicated as negative (false negative). This shows that the model is highly accurate in recognizing negative reviews with a high level of precision; however, its sensitivity to positive reviewsreviews still needs to be improved. Nevertheless, the performance of Logistic Regression in general remains competitive, especially in the context of text-based binary classification, as it can provide stable and easy-to-interpret prediction results. A detailed summary of these classification outcomes is illustrated through the confusion matrix in Figure 5.

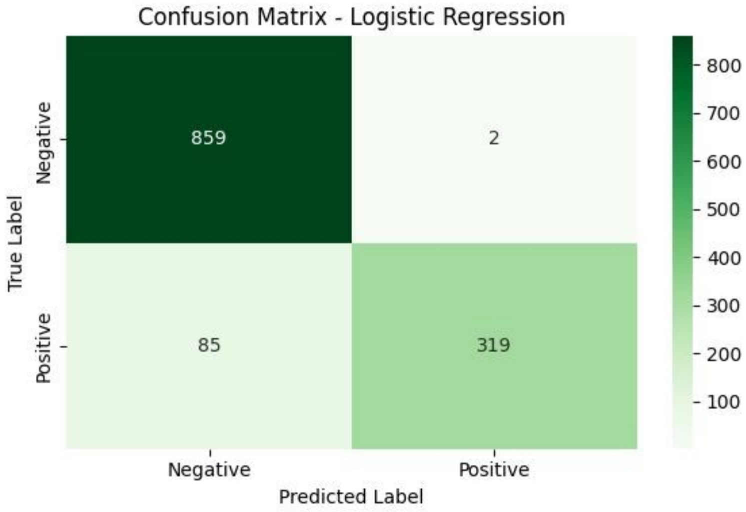

Figure 5 Logistic Regression Matrix Confusion

3. 2.3 Error Analysis in Sentiment Classification

Apart from the general results shown in the confusion matrices (Figures 4 and 5), a close look at misclassified instances reveals the constraints of Logistic Regression and SVM models. Most misclassifications occur when positive reviews are projected as negative. Examples include words like "lelet", "aplikasi my telkomsel loading buka", and "iya nih my telkomsel buruk banget sih kemarin masuk app my telkomsel hadeuhhh". Users rated tolng jaring ya baik as good; however, both models mistakenly projected it as bad. On the other hand, some unfavorable reviews, including somewhat positive comments such as "lapor tindak informasi overload jaring internet tsel lalod" and "terima kasih telkomsel rubah paket serba", were incorrectly categorized as positive. Particularly by Logistic Regression, ribu telkomsel telkomsel terima kasih lho were wrongly classified as positive.

Table 11 shows that SVM misclassified 71 positive and 13 negative cases, whereas Logistic Regression misclassified 2 negative and 85 positive cases. The results show that, unlike SVM, which has a somewhat more balanced error distribution, Logistic Regression tends to misclassify positive sentiments more often. Uncertain, brief, or informal phrases, as well as texts with varied emotion, which both models struggle to understand correctly, primarily drive the misclassifications. Generally speaking, examining mislabeled instances reveals the difficulties of sentiment categorization in user-generated content and points to possible fixes, including better feature extraction, context-aware embeddings, and more effective approaches to capture subtle and mixed emotions in future research.

Table 11Misclassified Cases per Model and Class

Models	Positif Misclassified	Negatif Misclassified	Total Error
SVM	71	13	84
Logistic Regression	85	2	87

3. 4 Visualization Feeling

To obtain a visual representation of the sentiment predicted by the model, visualizations were carried out using WordCloud. This visualization presents the most frequently occurring sets of words for each sentiment label (positive and negative) based on the results of classification using the SVM and Logistic Regression models. The larger the size of a word in a WordCloud, the higher its frequency of occurrence in user reviews. This visualization aims to identify general patterns in users' perceptions of the MyTelkomsel App.

Figure 6 Word cloud sentiment analysis for each class

WordCloud visualization of the results of sentiment classification by the SVM and Logistic Regression models shown in Figure 6, shows that there is real consistency in the dominant words in each sentiment class. In the figure, words frequently included in positive reviews include luar biasa", "telkomsel", "mudah", "bagus", "mantap", and "paket". On the other hand, words frequently included in negative reviews include "aplikasi", "lot", "ya", "jelek", "kuota", "lelet", "loading", and "parah". This shows that the SVM and Logistic Regression models in classifying sentiment are able to capture general patterns in user opinions, whether reflecting user satisfaction or complaints.

4. CONCLUSIONS

The conclusion of this study shows that SVM and Logistic Regression have nearly identical precision, recall, and F1 scores based on a weighted average. However, compared to Logistic Regression, SVM is superior, with an accuracy of 93.36%, an AUC of 0.9680, and a recall of 82% for the positive class. The results of the study showed that SVM had fewer Classification errors per Model in the class with a total error of 84, while in Logistic regression, the total error was 87. The results also show that Logistic Regression outperforms SVM in predictive accuracy for the positive class, with a precision of 99%. Therefore, the SVM model can be considered a more optimal algorithm for classifying the sentiment of MyTelkomsel app reviews, especially in the context of effectively recognizing positive user opinions.

This study, as a recommendation, due to limitations, suggests that further research should consider including a neutral sentiment category. Furthermore, this study still uses ratings for labeling, so future research could utilize Lexicon-Based Labeling for data labeling. In this case, it can be focused on the content of the text, not just the rating.

REFERENCES

[1] D. Patmalasari and A. D. Indriyanti, “Analisis Kepuasan Pengguna Layanan Aplikasi MyTelkomsel dengan Menggunakan Model UTAUT,” *J. Emerg. Inf. Syst. Bus. Intell.*, vol. 02, no. 02, pp. 37–45, 2021, doi: https://doi.org/10.26740/jeisbi.v2i2.39559.

[2] Novenya Rahmadani and Siti Ning Fardida, “Peran UI/UX Pada Layanan Aplikasi Mytelkomsel Terhadap Keputusan Pembelian dan Loyalitas Pelanggan,” *SAFARI J. Pengabdi. Masy. Indones.*, vol. 4, no. 3, pp. 220–227, 2024, doi: 10.56910/safari.v4i3.1655.

[3] F. Nugraha, D. A. Rahayu, and I. Bastian, “Analisis Tingkat Kepuasan Pengguna Aplikasi Mytelkomsel Menggunakan Evaluasi Heuristik dan Metode Pieces (Studi Kasus :

Mahasiswa Kampus Karawaci Universitas Gunadarma),” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 3, pp. 463–468, 2022, doi: 10.25126/jtiik.2022924403.

[4] C. Apriyadi and S. Styawati, “Sentiment Analysis of Cyber Attacks in Bank Syariah Indonesia Using SVM and Indobert Method,” *J. Tek. Inform.*, vol. 6, no. 2, pp. 819–838, 2025, doi: 10.52436/1.jutif.2025.6.2.2636.

[5] T. S. Ningsih, T. I. Hermanto, and I. M. Nugroho, “Sentiment Analysis of Mobile Provider Application Reviews Using Naive Bayes Algorithm and Support Vector Machine,” *Sinkron*, vol. 8, no. 2, pp. 824–835, 2024, doi: 10.33395/sinkron.v8i2.13469.

[6] F. J. Fauzan, M Afdal, Rice Novita, and Mustakim, “Penerapan Machine Learning Pada Analisis Sentimen Aplikasi Mytelkomsel Mengunakan Data Ulasan Google Playstore,” *Indones. J. Comput. Sci.*, vol. 13, no. 3, pp. 4747–4761, 2024, doi: 10.33022/ijcs.v13i3.4024.

[7] D. Hardiansyah, R. Z. Abdul, and M. Said, “The Classification Method is Used for Sentiment Analysis in My Telkomsel,” vol. 8, no. 2, 2024, doi: https://doi.org/10.29099/ijair.v8i2.1229.

[8] A. Maulana, I. K. Afifah, A. Mubarrak, K. R. Fauzan, A. Dwintara, and B. P. Zen, “Comparison of Logistic Regression, Multinomialnb, Svm, and K-Nn Methods on Sentiment Analysis of Gojek App Reviews on the Google Play Store,” *J. Tek. Inform.*, vol. 4, no. 6, pp. 1487–1494, 2023, doi: https://doi.org/10.52436/1.jutif.2023.4.6.863.

[9] A. G. Budianto, A. Trisno, E. Suryo, and G. Rudi, “Perbandingan Performa Algoritma Support Vector Machine (SVM) dan Logistic Regression untuk Analisis Sentimen Pengguna Aplikasi Retail di Android,” *J. Sains dan Indormatika*, vol. 10, no. November, pp. 1–10, 2024, doi: 10.34128/jsi.v10i2.911.

[10] R. A. Husen, R. Astuti, L. Marlia, R. Rahmaddeni, and L. Efrizoni, “Analisis Sentimen Opini Publik pada Twitter Terhadap Bank BSI Menggunakan Algoritma Machine Learning,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 211–218, 2023, doi: 10.57152/malcom.v3i2.901.

[11] S. R. Putri, M. Arifin, and Supriyono, “Public Sentiment Analysis of Nadiem Makarim as Minister of Education, Culture, Research, and Technology using Support Vector Machine (SVM),” *Sist. J. Sist. Inf.*, vol. 14, no. 2, pp. 826–834, 2025, doi: https://doi.org/10.32520/stmsi.v14i2.5067.

[12] A. S. D. P. Sinaga and A. S. Aji, “Analisis Sentimen Publik Terhadap Mayor Teddy Indra Wijaya dengan Pendekatan Logistic Regression: Public Sentiment Analysis Towards Mayor Teddy Indra Wijaya with Logistic Regression Approach,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 1, pp. 222–231, 2025, doi: https://doi.org/10.57152/malcom.v5i1.1752.

[13] N. Aditiya, P. Setiaji, and Supriyono, “Sentiment Analysis of Public Satisfaction with the’INFO BMKG’Application using Naive Bayes, SVM, and KNN,” *Sist. J. Sist. Inf.*, vol. 14, no. 3, pp. 1418–1432, 2025, doi: https://doi.org/10.32520/stmsi.v14i3.5223.

[14] W. Ramadlan As’ari, M. Arifin, D. Laily Fithri, and P. Setiaji, “Analisis Sentimen Ulasan Pengguna Aplikasi Spotify Di Google Play Store Menggunakan Algoritma Naive Bayes,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 2, pp. 3600–3607, 2025, doi: 10.36040/jati.v9i2.13543.

[15] M. T. Hidayat, M. Arifin, and S. Muzid, “Prediction Sentiment Analysis Grab Reviews Using SVM Linear Based Streamlit,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 9, No.12, no. 2, pp. 189–200, 2025, doi: 10.22146/ijccs.104924.

[16] F. R. Razak, M. Kunta Biddinika, H. Yuliansyah, and D. Soyusiawaty, “SMOTE-SVM for Handling Imbalanced Data in Obesity Classification,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 19, No.2, no. 2, pp. 142–152, 2024, doi: 10.22146/ijccs.103994.

[17] R. Y. Lewu, K. Kusrini, and A. Yaqin, “Comparing text classification algorithms with n-grams for mediation prediction,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 18, no. 2, 2024, doi: 10.22146/ijccs.93929.