Comparison of Support Vector Machine and Random Forest Algorithms in Sentiment Analysis of the JMO Mobile Application

Inneke Via Mariska [1]*, Allsela Meiriza [2]*, Dinda Lestarini [3]*
* Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya
09031282227043@student.unsri.ac.id [1], allsela@unsri.ac.id [2], dinda@unsri.ac.id [3]

Article Info

Article history:

Received 2025-08-12
Revised 2025-08-31
Accepted 2025-09-04

Keyword:

JMO Mobile,
Machine Learning,
Sentiment Analysis,
Support Vector Machine,
Random Forest.

ABSTRACT

JMO Mobile is a digital service application that enables the public to access employment-related information and benefits. User reviews serve as a valuable resource for evaluating service quality, yet systematic sentiment analysis on this application remains limited. This study aims to classify the sentiment of user reviews and compare the performance of Support Vector Machine (SVM) and Random Forest (RF) algorithms. A total of 41,673 reviews were collected through web scraping, then preprocessed through text cleaning, tokenization, stopword removal, stemming, and feature extraction using TF-IDF. The reviews were categorized into positive, negative, and neutral sentiments, and divided into training and testing datasets with an 80:20 ratio. The choice of SVM and RF was based on their proven effectiveness in text classification tasks, with SVM excelling in handling high-dimensional data and RF recognized for its stability in producing reliable results. Model evaluation was conducted using accuracy as the primary metric. The findings indicate that Random Forest achieved an accuracy of 86.15 percent, slightly outperforming SVM at 86.06 percent. While SVM showed superior performance in identifying positive sentiment, Random Forest demonstrated greater consistency across classifications. Overall, Random Forest is considered more suitable for sentiment analysis of public service application reviews. This study contributes an automated approach to understanding user perceptions and offers a reference for selecting classification algorithms in similar cases.

I. INTRODUCTION

There has been a significant shift toward digital services through mobile applications in Indonesia [1]. One service that has followed this trend is the JMO Mobile application created by BPJS Ketenagakerjaan. This app is designed to provide easy access for participants by offering various features such as checking JHT balances, submitting online claims, checking claim status, and updating participant data. With these services, participants no longer need to visit branch offices in person but can access various employment-related needs conveniently through their mobile devices.

So far, the JMO app has shown a positive trend with over five million downloads on the Google Play Store platform, a figure that clearly illustrates the high level of interest and enthusiasm among the public in utilizing digital-based employment services. This phenomenon indicates a shift in public behavior, with people increasingly accustomed to using online services to access their rights independently without the hassle of physically visiting service offices [2]. The popularity of this app not only reflects the penetration of information technology in public services but also mirrors the public's need for more practical and efficient services. However, the high number of downloads does not always correlate directly with user satisfaction, as various other factors also influence users' perceptions of the app's service quality.

Many complaints have been raised by JMO Mobile users, particularly regarding difficulties in accessing the application, technical issues during login, and challenges in submitting JHT claims online [3]. Users also frequently criticize the app's technical performance, which is perceived as slow to respond, as well as the instability of service access, which often experiences disruptions at certain times. These conditions indicate gaps in the app's performance that require further attention, necessitating research that can evaluate public perception through sentiment analysis of user reviews, to better comprehend the overall patterns of public sentiment toward the quality of the JMO Mobile application's services.

Previous studies have used sentiment analysis to categorize user opinions into positive and negative categories. One study using the Naive Bayes method showed fairly good sentiment classification results with an accuracy rate of 86%, but it had weaknesses in handling the complexity of application review text data, especially reviews with non-standard sentence structures and informal language use [4]. Therefore, further research is required by employing a classification method that is more adaptive to the characteristics of text data in order to provide more accurate classification results.

Support Vector Machine (SVM) algorithm is recognized for its effectiveness in classifying high-dimensional text data. Research on TikTok reviews shows that SVM is capable of producing a high accuracy rate of 89%, making it suitable for handling review data, especially in grouping sentiments into several categories with consistent results [5]. On the other hand, Random Forest is also frequently used due to its advantages in handling imbalanced data and its ability to provide stable multi-class classification. Research findings indicate that Random Forest achieved an F1-score of 86% in testing on review data [6].

Other studies examining the performance of classification algorithms show varying results depending on the type of data used. A study on the PLN Mobile application found that Random Forest was able to achieve an F1-score of 93.14%, proving its effectiveness in classifying public service sentiment [7]. Another study also showed that Random Forest outperformed in classifying sentiment on fuel shortages on social media with an accuracy rate of nearly 90% [8]. On the other hand, sentiment analysis of YouTube comments showed that SVM outperformed with an accuracy rate of 85% on test data, compared to Random Forest, which only achieved 80%. These results indicate that Random Forest is suitable for public service data, while SVM is more effective for unstructured textual information like user reviews or comments on social media, which tend to be complex [9].

However, research comparing how well the Support Vector Machine and Random Forest algorithms perform, particularly in the context of sentiment classification of user reviews of the JMO Mobile application, is still very limited. Most existing studies tend to use only one classification method and limit sentiment classification to two categories, namely positive and negative, without considering neutral sentiment, which also has important informational value in evaluating service satisfaction. Based on this, this study aims to classify sentiment into three categories positive, neutral, and negative while comparing the performance of two classification algorithms, SVM and Random Forest. Therefore this study aims to offer a more comprehensive understanding of public opinion regarding the JMO Mobile app service, thereby contributing to efforts to improve the quality of public services based on mobile applications.

II. METHODS

This study consists of several main stages, ranging from collecting data on user reviews of the JMO Mobile application to evaluating the classification results. Each stage is carried out sequentially so that the analysis process runs systematically. The complete flow of the research stages is shown in Figure 1.

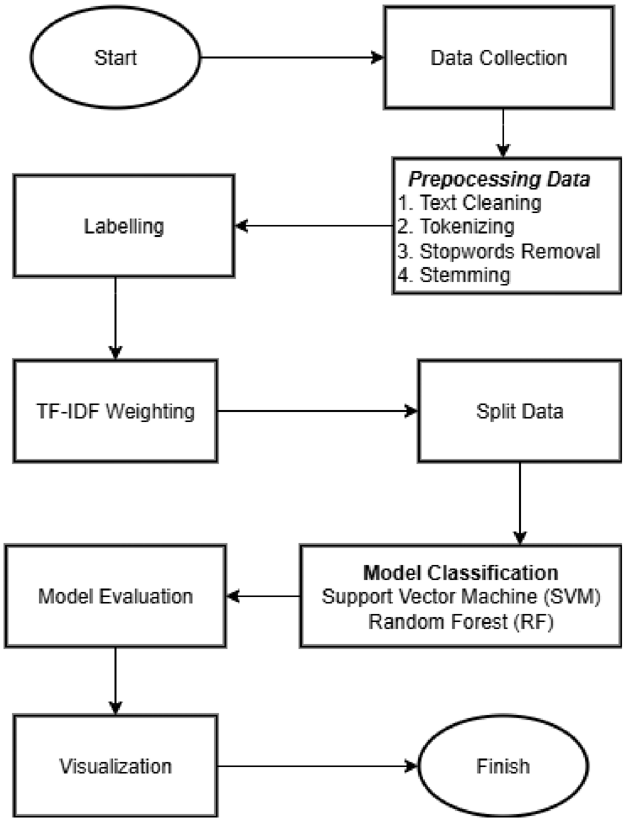

Figure 1. Research Methodology

A. *Data Collection*

Data collection in this study was conducted using the web scraping method, which is widely used for automatically extracting text data from online sources [10]. The scraping process was carried out through Google Colab with the help of the Python library google-play-scraper to retrieve user reviews of the JMO Mobile application from the Google Play Store. The use of scraping enables the efficient collection of large amounts of data, including structured text reviews and star ratings [11].

B. *Prepocessing Data*

Data preprocessing is performed to convert review text into a more structured form before it is used in the classification process. The preprocessing stage aims to clean the text of various irrelevant elements and normalize the data so that it is easier for machine learning algorithms to process [8].

C. Labelling Data

The data labeling process was carried out by dividing review scores into three sentiment categories. Review scores with values above 3 are classified as positive sentiment, scores below 3 as negative sentiment, and scores exactly 3 are placed in the neutral sentiment category. This categorization helps transform the score data into well-defined sentiment labels, thereby facilitating the analysis process of users' views on the application being studied [12].

D. TF-IDF Weighting

TF-IDF is a weighting approach that represents textual data as numerical values by assessing how frequently a term appears in a single document compared to its occurrence throughout the whole collection of documents. Words that appear frequently in one review but are rarely found in the entire data set will receive a high weight, while common words will receive a low weight. This method helps highlight important words in the text classification process.

The use of TF-IDF can minimize the influence of common words and make the data more focused on relevant words [13]. With this weighting, text data becomes more structured, thereby improving the effectiveness of the classification process using machine learning.

E. Split Data

Data splitting refers to separating a dataset into two subsets which are training data and testing data. The training subset is utilized to teach the model to identify patterns, whereas the testing subset is employed to evaluate how well the model predicts unseen information. Generally, the dataset is split in an 80% to 20% ratio between training and testing data to ensure the model's ability to generalize effectively. This data splitting approach is also applied in sentiment analysis studies of application reviews, as it helps minimize the likelihood of overfitting and enhances the reliability of model evaluation [10].

F. Model Classification

The classification stage in this study was conducted using two machine learning algorithms, namely Support Vector Machine (SVM) and Random Forest (RF). The selection of these two algorithms was based on their proven track record in sentiment analysis tasks and their ability to handle large amounts of text data [14], [15].

1) Support Vector Machine

Support Vector Machine is a classification algorithm that works by finding the best separating line or plane (hyperplane) to separate data into different classes. SVM is known to be effective in handling high-dimensional data such as TF-IDF extracted text data, because it can form an optimal separating boundary between classes [14]. In this study, SVM was implemented using LinearSVC with a linear kernel, which is efficient for TF-IDF text data. No additional hyperparameter settings were made so that the configuration reflects the basic performance of the model.

2) Random Forest

Random Forest is an ensemble learning-based classification algorithm that builds a number of decision trees randomly from subsets of data and features. Each tree produces a prediction, and the final result is determined through a voting mechanism. This approach makes Random Forest more stable against data variation and more resistant to overfitting, especially when used on data with many features such as text [15]. In this study, Random Forest was used with the default scikit-learn parameters, namely n_estimators = 100, without additional tuning to maintain evaluation consistency.

G. Model Evaluation

Model evaluation was conducted to measure the performance of classification algorithms in grouping sentiment data. In this study, the model was tested using test data obtained from an 80:20 hold-out split, where 80% of the data was used for training and 20% for testing. The following performance assessments were conducted using several evaluation metrics commonly used in text classification, which are accuracy, precision, recall, and f1-score.

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

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

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

$$F1 - Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (4)$$

True Positive (TP) represents instances that are genuinely positive and correctly identified as such. True Negative (TN) denotes cases that are genuinely negative and correctly recognized as negative. False Positive (FP) is a condition where data that should be negative is instead predicted as positive. Meanwhile, False Negative (FN) occurs when data that should be positive is instead classified as negative [8]. The results of the Confusion Matrix test are shown in Table I.

TABLE I
CONFUSION MATRIX

	Prediction		
	Class	**Negative**	**Positive**
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

H. Visualization

After applying both algorithms, data visualization was performed using WordCloud and graphs for every sentiment category, which are positive, neutral, and negative derived from the labeling results in the training data. WordCloud displays the most frequently occurring words, where the higher the frequency, the larger the text size of the word will be displayed [16].

III. Result and Discussion

A. Data Collection Results

Data was collected using web scraping methods from the JMO Mobile application available on the Google Play Store platform. The collection focused on the latest reviews in the period from January 1, 2025, to June 25, 2025. A total of 41,673 reviews were obtained. The collected data includes review content, ratings, and review dates. The results of the data collection are shown in Table II.

TABLE II
DATA COLLECTION RESULTS

	Content	Score	Review At
0	Aplikasi bagus bangeeet Sangat Membantu Bagi Kalangan Bagi ga Mampu Sukses Terus Semoga Semakin Jaya Is the best	5	6/25/2025 7:26
1	sangat mempermudah dijaman yg serba aplikasi ini. semoga tetap menjadi penenang seluruh peserta JMO. semoga dapat melayani semua keperluan peserta dgn baik. salam sehat selalu	5	6/18/2025 2:47
2	kenapa setiap mau login harus update2 terus. ini sangat mengganggu. mohon di review lagi mengenai hal ini. karena bikin tidak nyaman. tolong di perbaiki. jangan tiap masuk harus update	1	6/15/2025 7:40

B. Prepocessing Results

This process consists of the following steps:

1) Text Cleaning

Removing irrelevant characters such as numbers, punctuation marks, special symbols, and links or URLs that do not contribute to sentiment analysis. The results of text cleaning are shown in Table III.

TABLE III
CLEANING TEXT

Before Text Cleaning	After Text Cleaning
Aplikasi bagus bangeeet Sangat Membantu Bagi Kalangan Bagi ga Mampu Sukses Terus Semoga Semakin Jaya Is the best	aplikasi bagus bangeeet sangat membantu bagi kalangan bagi ga mampu sukses terus semoga semakin jaya is the best
sangat mempermudah dijaman yg serba aplikasi ini. semoga tetap menjadi penenang seluruh peserta JMO. semoga dapat melayani semua keperluan peserta dgn baik. salam sehat selalu	sangat mempermudah dijaman yg serba aplikasi ini semoga tetap menjadi penenang seluruh peserta jmo semoga dapat melayani semua keperluan peserta dgn baik salam sehat selalu
kenapa setiap mau login harus update2 terus. ini sangat mengganggu. mohon di review lagi mengenai hal ini. karena bikin tidak nyaman. tolong di perbaiki. jangan tiap masuk harus update	kenapa setiap mau login harus update terus ini sangat mengganggu mohon di review lagi mengenai hal ini karena bikin tidak nyaman tolong di perbaiki jangan tiap masuk harus update

2) Tokenizing

Breaking down review sentences into word segments or tokens so that text data can be processed in word units. The results of tokenizing are shown in Table IV.

TABLE IV
TOKENIZING

Text Cleaning Data	After Tokenizing
aplikasi bagus bangeeet sangat membantu bagi kalangan bagi ga mampu sukses terus semoga semakin jaya is the best	['aplikasi', 'bagus', 'bangeeet', 'sangat', 'membantu', 'bagi', 'kalangan', 'bagi', 'ga', 'mampu', 'sukses', 'terus', 'semoga', 'semakin', 'jaya', 'is', 'the', 'best']
sangat mempermudah dijaman yg serba aplikasi ini semoga tetap menjadi penenang seluruh peserta jmo semoga dapat melayani semua keperluan peserta dgn baik salam sehat selalu	['sangat', 'mempermudah', 'dijaman', 'yg', 'serba', 'aplikasi', 'ini', 'semoga', 'tetap', 'menjadi', 'penenang', 'seluruh', 'peserta', 'jmo', 'semoga', 'dapat', 'melayani', 'semua', 'keperluan', 'peserta', 'dgn', 'baik', 'salam', 'sehat', 'selalu']
kenapa setiap mau login harus update terus ini sangat mengganggu mohon di review lagi mengenai hal ini karena bikin tidak nyaman tolong di perbaiki jangan tiap masuk harus update	['kenapa', 'setiap', 'mau', 'login', 'harus', 'update', 'terus', 'ini', 'sangat', 'mengganggu', 'mohon', 'di', 'review', 'lagi', 'mengenai', 'hal', 'ini', 'karena', 'bikin', 'tidak', 'nyaman', 'tolong', 'di', 'perbaiki', 'jangan', 'tiap', 'masuk', 'harus', 'update']

3) Stopword Removal

Removing common words such as “dan”, “di”, and “yang” which are considered to have no significant meaning in sentiment analysis. The results of stopword removal are shown in Table V.

TABLE V
STOPWORD REMOVAL

Tokenizing Data	After Stopword Removal
['aplikasi', 'bagus', 'bangeeet', 'sangat', 'membantu', 'bagi', 'kalangan', 'bagi', 'ga', 'mampu', 'sukses', 'terus', 'semoga', 'semakin', 'jaya', 'is', 'the', 'best']	['aplikasi', 'bagus', 'bangeeet', 'membantu', 'kalangan', 'ga', 'sukses', 'semoga', 'jaya', 'is', 'the', 'best']
['sangat', 'mempermudah', 'dijaman', 'yg', 'serba', 'aplikasi', 'ini', 'semoga', 'tetap', 'menjadi', 'penenang', 'seluruh', 'peserta', 'jmo', 'semoga', 'dapat', 'melayani', 'semua', 'keperluan', 'peserta',	['mempermudah', 'dijaman', 'yg', 'serba', 'aplikasi', 'semoga', 'penenang', 'peserta', 'jmo', 'semoga', 'melayani', 'keperluan', 'peserta', 'dgn', 'salam', 'sehat']

'dgn', 'baik', 'salam', 'sehat', 'selalu']	
['kenapa', 'setiap', 'mau', 'login', 'harus', 'update', 'terus', 'ini', 'sangat', 'mengganggu', 'mohon', 'di', 'review', 'lagi', 'mengenai', 'hal', 'ini', 'karena', 'bikin', 'tidak', 'nyaman', 'tolong', 'di', 'perbaiki', 'jangan', 'tiap', 'masuk', 'harus', 'update']	['login', 'update', 'mengganggu', 'mohon', 'review', 'bikin', 'nyaman', 'tolong', 'perbaiki', 'masuk', 'update']

4) Stemming

Change each word to its root form, so that derivative words such as "mempermudah" and "dipermudah" are returned to their root form "mudah." The results of stemming are shown in Table VI.

TABLE VI
STEMMING

Stopword Removal Data	After Stemming
['aplikasi', 'bagus', 'bangeeet', 'membantu', 'kalangan', 'ga', 'sukses', 'semoga', 'jaya', 'is', 'the', 'best']	['aplikasi', 'bagus', 'bangeeet', 'bantu', 'kalang', 'ga', 'sukses', 'moga', 'jaya', 'is', 'the', 'best']
['mempermudah', 'dijaman', 'yg', 'serba', 'aplikasi', 'semoga', 'penenang', 'peserta', 'jmo', 'semoga', 'melayani', 'keperluan', 'peserta', 'dgn', 'salam', 'sehat']	['mudah', 'jam', 'yg', 'serba', 'aplikasi', 'moga', 'tenang', 'serta', 'jmo', 'moga', 'layan', 'perlu', 'serta', 'dgn', 'salam', 'sehat']
['login', 'update', 'mengganggu', 'mohon', 'review', 'bikin', 'nyaman', 'tolong', 'perbaiki', 'masuk', 'update']	['login', 'update', 'ganggu', 'mohon', 'review', 'bikin', 'nyaman', 'tolong', 'baik', 'masuk', 'update']

C. Labelling Data Results

The data labeling results indicate that most reviews fall into the positive sentiment category, with 32,345 data points. Meanwhile, negative sentiment was recorded in 7,355 data points, and the remaining 1,973 data points fell into the neutral category. The following is a visualization of the sentiment labeling results based on the data collected, as shown in Figure 2.

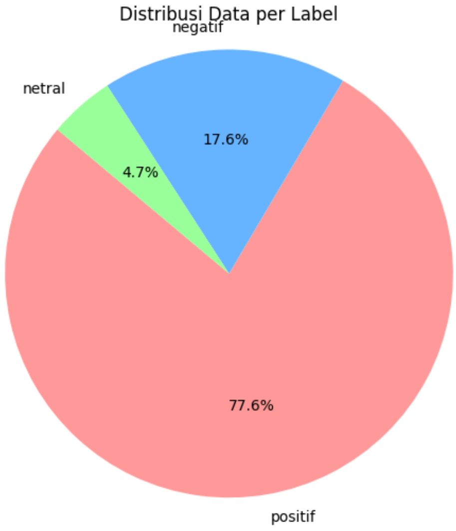

Figure 2. Labelling Data Results

D. TF-IDF Weighting Results

The TF-IDF weighting results show the five words with the highest values that represent the most prominent terms in user reviews, as shown in Table VII.

TABLE VII
PEMBOBOTAN TF-IDF

Kata	**TF-IDF**
bagus	351.992972
update	300.958222
aplikasi	284.748440
bantu	266.051100
buka	235.285976

E. Split Data

The dataset is split into two parts, namely training data and testing data, with a ratio of 80:20. Of the total 41,673 data, 33,338 are used as training data and the remaining 8,335 as testing data. This division allows the model to be trained with the majority of available data while testing its performance on unrecognized data.

F. Model Classification

The classification models used in this study consist of Support Vector Machine (SVM) and Random Forest (RF), each of which was applied to user review data after undergoing preprocessing and TF-IDF weighting. Based on testing results on a test dataset of 8,335 samples, both models demonstrated nearly identical accuracy, exceeding 86%. SVM achieved high performance in the positive sentiment category, while Random Forest provided a more balanced classification distribution, particularly in the negative category. These results indicate that both models are capable of classifying sentiment effectively, albeit with differing strengths across each class.

G. Model Evaluation

The SVM model evaluation has an accuracy value of 86.06%, which shows that overall the model is quite reliable. The positive class dominates performance with precision of 0.91, recall of 0.95, and f1-score of 0.93. For the negative class, precision of 0.67 and f1-score of 0.68 show moderate results. Meanwhile, the neutral class obtained low performance with an F1-score of only 0.05 due to very low recall, namely 0.03. This condition causes the macro average F1-score to be at 0.55, while the weighted average F1-score remains high at 0.84 because the largest contribution comes from the positive class. The SVM evaluation results are shown in Figure 3.

```
📊 Evaluasi SVM
Akurasi : 0.8605878824235152

Classification Report:
              precision    recall  f1-score   support

     negatif       0.67      0.70      0.68      1471
      netral       0.28      0.03      0.05       395
     positif       0.91      0.95      0.93      6469

    accuracy                           0.86      8335
   macro avg       0.62      0.56      0.55      8335
weighted avg       0.84      0.86      0.84      8335
```

Figure 3. SVM Evaluation Results

The confusion matrix in the SVM model shows that the model performs very well in classifying data in the positive class, with 6,136 out of 6,469 data points correctly classified. Conversely, in the negative class, only 1,026 out of 1,471 data points were correctly predicted, while the remainder were mostly misclassified as the positive class. The lowest performance is observed in the neutral class, with only 11 out of 395 data points correctly identified, indicating that the SVM model struggles to distinguish the neutral class from the others. The confusion matrix results are shown in Figure 4.

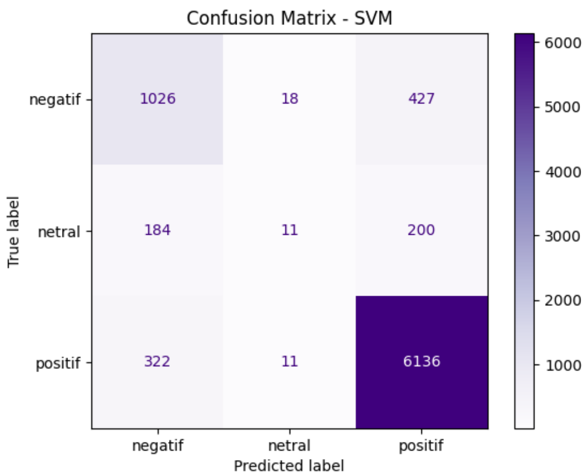

Figure 4. Confusion Matrix SVM

On the other side, the Random Forest model evaluation results show an accuracy of 86.15%, which indicates that most predictions match the actual labels. Positive sentiment is recognized well, as shown by precision, recall, and f1-score of 0.93. For negative sentiment, precision of 0.63 and recall of 0.79 show fairly good performance. However, the model had difficulty classifying neutral sentiment, as seen from the very low recall and f1-score, which were only 0.01 each. The weighted average precision and recall were 0.85 and 0.86, respectively, indicating that the overall performance of the model was still consistent on the test data. The Random Forest evaluation results are shown in Figure 5.

```
📊 Evaluasi Random Forest
Akurasi : 0.8615476904619076

Classification Report:
              precision    recall  f1-score   support

     negatif       0.63      0.79      0.70      1471
      netral       0.43      0.01      0.01       395
     positif       0.93      0.93      0.93      6469

    accuracy                           0.86      8335
   macro avg       0.66      0.57      0.55      8335
weighted avg       0.85      0.86      0.85      8335
```

Figure 5. Random Forest Evaluation Results

The confusion matrix shows that the model is very dominant in recognizing positive reviews, with 6,023 out of 6,469 data classified correctly. For the negative class, 1,155 out of 1,471 data were classified correctly, while the rest were mostly misclassified as positive. The neutral class has the lowest performance, with only 3 out of 395 data points correctly identified. Most neutral data points were instead classified as negative (228) and positive (164), reinforcing the finding that the model struggles to effectively distinguish neutral sentiment. The confusion matrix results are shown in Figure 6.

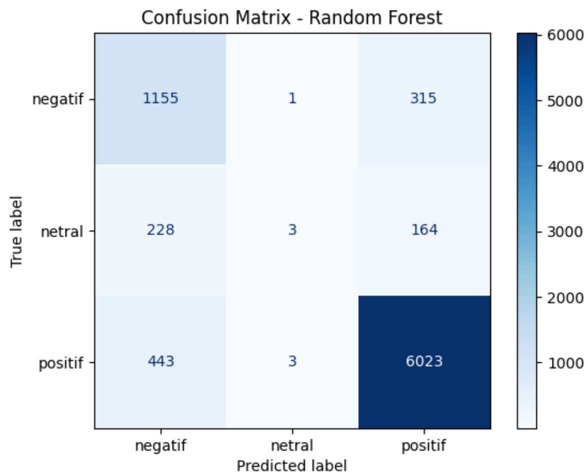

Figure 6. Confusion Matrix Random Forest

The following are the results of the comparison between the Support Vector Machine and Random Forest models, as shown in Table VIII.

TABLE VIII
COMPARISON OF MODEL

Sentiment Class	Metric	SVM	Random Forest
Positive	Precision	0.91	0.93
	Recall	0.95	0.93
	F-1 Score	0.93	0.93
Negative	Precision	0.67	0.63
	Recall	0.70	0.79
	F-1 Score	0.68	0.70
Neutral	Precision	0.28	0.43
	Recall	0.03	0.01
	F-1 Score	0.05	0.01
Accuracy	-	86.06%	86.15%

Based on Table VIII, shows the results of comparing the Support Vector Machine (SVM) and Random Forest (RF) models. The evaluation results show that the accuracy of both models is almost the same, namely 86.06% for SVM and 86.15% for RF, with a very small difference. However, other metrics show differences in performance trends.

In the positive class, SVM shows the highest recall of 0.95 with an f1-score of 0.93, while Random Forest obtains a recall of 0.93 with the same f1-score. This shows that SVM is more reliable in recognizing positive reviews. Conversely, in the negative class, Random Forest had a higher recall (0.79 compared to 0.70 for SVM) and a better f1-score (0.70 compared to 0.68). This difference indicates that Random Forest is more stable in detecting negative reviews.

Both models are equally weak in classifying neutral classes, with very low recall, namely 0.03 for SVM and 0.01 for Random Forest, which results in a low f1-score value below 0.05. This shows that unbalanced data distribution affects the model's ability to recognize minority classes.

Overall, although the accuracy of both models is almost the same, the results of other metric evaluations indicate that Random Forest is more consistent in handling negative reviews, while SVM is superior in detecting positive reviews. However, claims of differences in effectiveness between algorithms still need to be further tested using statistical significance analysis in future research.

Unlike previous studies that discussed JMO Mobile app reviews using a single algorithm and binary classification, this study adopted a more comprehensive approach. Other studies have used the Naïve Bayes algorithm to classify review sentiment into two categories, positive and negative, without considering the existence of neutral sentiment [3]. Although this model produced very high accuracy, namely 96%, as well as 100% recall and 96% precision for negative sentiment, this approach did not describe the nuances of user sentiment comprehensively.

In contrast, this study applied two classification algorithms, namely Support Vector Machine and Random Forest, and grouped sentiment into three categories, namely positive, negative, and neutral. Although the accuracy values are lower than in previous studies, namely 86.15% for Random Forest and 86.06% for SVM, this multi-class approach provides a more detailed picture of user perceptions. The advantage of this study lies in its more comprehensive classification coverage and evaluation of two algorithms at once, which has not been widely done in previous studies related to the JMO Mobile application.

H. Visualization

The final stage of this process is data visualization. The word cloud visualization in the image shows the distribution of the most frequently occurring words in each sentiment category, namely positive, negative, and neutral, related to user reviews of the JMO application.

In the positive sentiment category, words such as “bagus”, “bantu” “mantap” “mudah” and “aplikasi” dominate. The presence of these words reflects users' appreciation for the quality of service provided by the JMO application, including ease of use, perceived benefits, and overall positive experiences. The word cloud results for positive sentiment are shown in Figure 7.

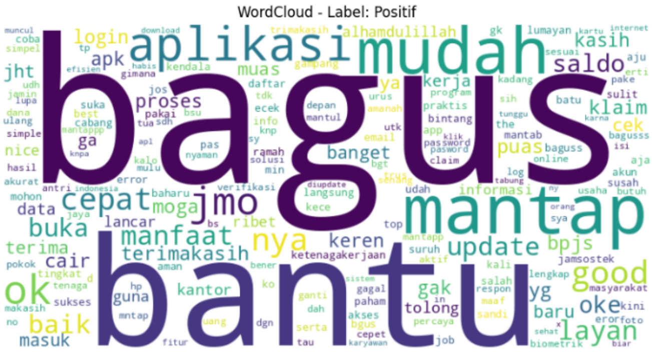

Figure 7. WordCloud Positive Sentiment

On the neutral sentiment label, words such as “update” “aplikasi” “buka” “login,” and “masuk” appear most frequently. These words reflect statements that are informative or descriptive in nature without containing strong emotional content, such as notifications about application updates or user activities when using JMO. The word cloud results for neutral sentiment are shown in Figure 8.

Figure 8. WordCloud Neutral Sentiment

On the negative sentiment label, words such as "aplikasi," "update," "login," "buka," and "gak" often appear and represent user complaints or dissatisfaction with their experience using the JMO application. The presence of these words indicates obstacles or inconveniences in the usage process, such as difficulties logging in, the app not opening, or updates that are perceived as disruptive. The word cloud results for negative sentiment are shown in Figure 9.

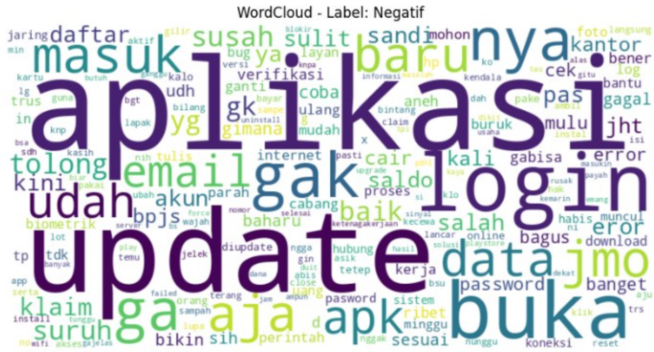

Figure 9. WordCloud Negative Sentiment

IV. Conclusion

This study compares the performance of Support Vector Machine (SVM) and Random Forest (RF) in classifying user sentiment reviews of the JMO Mobile application with a total of 41,673 data processed through text preprocessing and TF-IDF feature extraction, then divided into training and test data with a ratio of 80:20. The evaluation results show that Random Forest achieved an accuracy of 86.15%, slightly higher than SVM with 86.06%. SVM is superior in recognizing positive sentiment, while Random Forest shows better consistency in negative sentiment. Both algorithms still face challenges in classifying neutral sentiment, which is influenced by the unbalanced data distribution. Overall, Random Forest can be considered more stable in producing classifications, while SVM remains relevant for handling high-dimensional text data. This study provides a comparative overview of two popular algorithms in sentiment analysis and can be used as a reference for selecting classification methods in public service application reviews such as JMO Mobile.

Daftar Pustaka

[1] K. Kusuma Dewi, I. Kaniawulan, and C. Dewi Lestari, "Analisis Sentimen Pengguna Aplikasi Jamsostek Mobile (Jmo) Pada Appstore Menggunakan Metode Naive Bayes," *Simtek J. Sist. Inf. dan Tek. Komput.*, vol. 8, no. 2, pp. 333–338, 2023, doi: 10.51876/simtek.v8i2.286.

[2] S. A. R. Rizaldi, S. Alam, and I. Kurniawan, "Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 109–117, 2023, doi: 10.55123/storage.v2i3.2334.

[3] A. Noviriandini, H. Hermanto, and D. Ayu Ambarsari, "Analisis Tingkat Kepuasan Pengguna Aplikasi JMO (Jamsostek Mobile) Menggunakan Algoritma Naive Bayes Classifier," *Reputasi J. Rekayasa Perangkat Lunak*, vol. 4, no. 1, pp. 33–37, 2023, doi: 10.31294/reputasi.v4i1.1986.

[4] D. Azarine, N. Rahaningsih, and R. D. Dana, "Analisis Sentimen Terhadap Pengguna Aplikasi Jamsostek Mobile pada Google Play Store Menggunakan Metode Naïve Bayes," *Media Inform.*, vol. 24, no. 1, pp. 13–21, 2025, doi: 10.37595/mediainfo.v24i1.316.

[5] V. Fitriyana, L. Hakim, D. Novitasari, and A. Asyhar, "Analisis Sentimen Ulasan Aplikasi Jamsostek Mobile Menggunakan Menggunakan Metode Support Vector Machine," *J. Inf. Syst. Res.*, vol. 4, no. 3, pp. 865–873, 2023, doi: 10.47065/josh.v4i3.3353.

[6] M. R. Adrian, M. P. Putra, M. H. Rafialdy, and N. A. Rakhmawati, "Perbandingan Metode Klasifikasi Random Forest dan SVM Pada Analisis Sentimen PSBB," *J. Inform. Upgris*, vol. 7, no. 1, pp. 6–11, 2021, doi: 10.26877/jiu.v7i1.7099.

[7] M. D. Rizkiyanto, M. D. Purbolaksono, and W. Astuti, "Sentiment Analysis Classification on PLN Mobile Application Reviews using Random Forest Method and TF-IDF Feature Extraction," *INTEK J. Penelit.*, vol. 11, no. 1, pp. 37–43, 2024, doi: 10.31963/intek.v11i1.4774.

[8] T. Ariaji, S. N. Wahyuni, and M. Ikhsan, "Perbandingan Algoritma SVM dan RF pada Analisis Sentimen menggunakan Pendekatan Machine Learning," *J. Autom. Comput. Inf. Syst.*, vol. 5, no. 1, pp. 80–92, 2025, doi: 10.47134/jacis.v5i1.107.

[9] A. N. Syafia, M. F. Hidayattullah, and W. Suteddy, "Studi Komparasi Algoritma SVM Dan Random Forest Pada Analisis Sentimen Komentar Youtube BTS," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 207–212, 2023, doi: 10.30591/jpit.v8i3.5064.

[10] M. Iqbal, M. Afdal, and R. Novita, "Implementasi Algoritma Support Vector Machine Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online di Google Play Store," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1244–1252, 2024, doi: 10.57152/malcom.v4i4.1435.

[11] T. A. U. Azmi, L. Hakim, D. C. R. Novitasari, and W. D. U. D. Utami, "Application Random Forest Method for Sentiment Analysis in Jamsostek Mobile Review," *Telematika*, vol. 20, no. 1, p. 117, 2023, doi: 10.31315/telematika.v20i1.8868.

[12] N. D. Kurniawan, P. R. Ferdian, and N. Hidayati, "Analisis Sentimen Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest Pada Ulasan Aplikasi Ajaib," *J. Nas. Teknol. dan Sist. Inf.*, vol. 11, no. 1, pp. 87–97, 2025, doi: 10.25077/teknosi.v11i1.2025.87-97.

[13] K. P. Harmandini and K. M. L, "Analysis of TF-IDF and TF-RF Feature Extraction on Product Review Sentiment," *Sinkron*, vol. 8, no. 2, pp. 929–937, 2024, doi: 10.33395/sinkron.v8i2.13376.

[14] A. R. Makhtum and M. Muhajir, "Sentiment Analysis of Omnibus Law Using Support Vector Machine (Svm) With Linear Kernel," *Barekeng*, vol. 17, no. 4, pp. 2197–2206, 2023, doi: 10.30598/barekengvol17iss4pp2197-2206.

[15] M. A. Tsaqif and W. Maharani, "Comparison of Random Forest and Decision Tree Methods for Emotion Classification based on Social Media Posts," *Technol. Sci.*, vol. 6, no. 4, pp. 2240–2248, 2025, doi: 10.47065/bits.v6i4.6677.

[16] L. Enjelia, Y. Cahyana, and D. Wahiddin, "Comparison of K-Nearest Neighbors and Naive Bayes Classifier Algorithms in Sentiment Analysis of 2024 Election in Twitter (X)," *J. Appl. Informatics Comput.*, vol. 9, no. 3, p. 946, 2025, [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC