Acceleration, Quantum, Information Technology and Algorithm Journal (AQILA)
Vol. 3, No. 1, June 2026, pp 9-17
ISSN: 3062-8555, DOI: 10.62123/aqila.v3i1.171

Sentiment Analysis of Fintech Application Users in Indonesia Using Machine Learning Algorithms

Made Marshall Vira Deva[1*], Lukman Abdurrahman[1], Hanif Fakhrurroja[1]

[1]Telkom University, Bandung, Indonesia

*Corresponding Email: mademarshall97@gmail.com

DOI : 10.6213/aqila.v3i1.171

Received : May 17, 2026
Revised : May 19, 2026
Accepted : May 25, 2026

Keywords:
Sentiment Analysis
Machine Learning
Fintech
Natural Language Processing
Indonesian Text
Support Vector Machine

ABSTRACT

This study focuses on Indonesian users' sentiments regarding 9 fintech apps based on their Google Play Store reviews. The rapid growth of the fintech industry in Indonesia makes it crucial to understand user perceptions and satisfaction. Around 2,554 reviews from users of Kredivo, ShopeePay, Dana, GoPay, LinkAja, Bareksa, Flip, Jenius, and OVO were analyzed. The user review text and data were preprocessed using text cleaning, slang normalization, stopword removal, stemming, and the Sastrawi library and were moved through the TF-IDF vectorizer (term frequency-inverse document frequency). The four algorithms were Naive Bayes, Logistic Regression, Support Vector Machine (SVM), and Random Forest. The results showed that SVM (Linear) achieved the best overall balanced performance with an accuracy of 80.23%, precision of 77.79%, recall of 80.23%, and the highest F1-score of 78.53%, outperforming Naive Bayes (accuracy 81.21%, F1-score 78.32%), Logistic Regression (accuracy 80.43%, F1-score 77.81%), and Random Forest (accuracy 78.08%, F1-score 75.81%). While Naive Bayes recorded the highest raw accuracy, SVM was selected as the best model due to its superior F1-score, which provides a more balanced evaluation across all sentiment classes. Machine learning provided a snapshot of the reviews' sentiments, with 42.4% positive, 51.4% negative, and 6.1% neutral. Kredivo and ShopeePay had the most favorable sentiments of 72.4% and 70.9%. The most salient sentiment indicators include 'bagus' (good) and 'bantu' (help) as top positive classifiers, while 'buruk' (bad) and 'kecewa' (disappointed) emerged as the most prominent negative classifiers, with 'mudah' (easy) and 'cepat' (fast) also strongly associated with positive sentiment. The results of this study give fintech firms a better grasp of user satisfaction, and fintech user positive sentiments.

1. INTRODUCTION

In Indonesia's financial sector, digital transformation has changed the way transactions are done. Traditional financial services have been transformed by financial technology (fintech) to become more accessible, easier, and more efficient. The growth of Indonesia's fintech sector has been unprecedented, with millions of people using its various services, according to the Financial Services Authority (OJK) [1][2].

Owing to its geographic landscape and population size, Indonesia has a fast-growing technology scene that is primarily driven by a wide range of financial technology services encompassing digital payments, virtual banking, peer-to-peer lending, investment technology, and digital insurance. Fintech companies have already become a part of everyday payment transactions, including GoPay, Dana, OVO, and ShopeePay. The rapid adoption of these platforms has resulted in extensive user feedback which has been analyzed to understand user satisfaction and pain points [3][4][5].

Understanding user sentiment is essential for a fintech company to outperform rivals and enhance the quality of the services provided. Traditional surveys take a long time and do not capture user feelings and sentiments in real time. Sentiment analysis through machine learning offers a quicker and scalable alternative to determining the sentiments of customers in bulk user reviews [3][6][7]. Companies can leverage action-oriented insights through document processing and machine learning applied to unstructured data. Previous research has demonstrated the effectiveness of machine learning algorithms in sentiment analysis. It has been established in the literature that algorithms in Support Vector Machine, Naive Bayes, Logistic Regression, and Random Forest can accurately analyze sentiments in Indonesian texts [8][9][10]. However, very few studies have attempted to conduct sentiment analysis of multiple fintech applications in Indonesia, and if they have, none have employed a comparative machine learning approach.

This research aims to fill this gap by conducting a comprehensive sentiment analysis of 9 major fintech applications in Indonesia. We employ multiple machine learning algorithms to compare their performance and identify the most effective approach for fintech sentiment classification. Specifically, this study has several objectives: (1) to collect and preprocess user reviews from Indonesian fintech applications, (2) to implement and compare four machine learning algorithms for sentiment classification, (3) to evaluate model performance using multiple metrics, (4) to identify key features that influence sentiment, and (5) to provide insights into user satisfaction across different fintech platforms.

This research is important because it has real-world implications for the fintech sector. Our research can help fintech companies prioritize their development efforts, better understand user perception, spot service flaws, and create strategies that will increase customer satisfaction. In addition to its usefulness, this study advances academia by offering a thorough methodology for sentiment analysis of Indonesian text in the fintech industry.

2. LITERATURE REVIEW

This section examines relevant literature on fintech, sentiment analysis techniques, and the machine learning algorithms employed in this research. The review establishes the theoretical foundation and positions this research within the broader academic context.

2.1. Financial Technology (Fintech)

The application of technology to enhance the availability and provision of financial services is known as fintech, or financial technology. The fintech landscape encompasses a wide range of industries, including digital payments, peer-to-peer lending, crowdfunding, robo-advisors, insurtech, and blockchain-based services [11]. Due to the widespread use of smartphones and internet connectivity, fintech has grown globally, with emerging markets like Indonesia experiencing particularly rapid growth.

Fintech, which offers banking services to those previously underserved or excluded from traditional financial systems, has emerged as a significant force behind financial inclusion in Indonesia. The Indonesian fintech ecosystem consists of lending platforms, investment apps, e-wallets, and payment gateways [1]. A number of notable companies have attracted substantial investment and achieved unicorn status. In this highly competitive environment, understanding user sentiment has become essential for fintech companies hoping to maintain customer trust and enhance service quality.

2.2. Sentiment Analysis

The computational analysis of people's opinions, sentiments, assessments, attitudes, and emotions as they are expressed in text is known as sentiment analysis, or opinion mining [3][12]. With useful applications in business intelligence, political analysis, social media monitoring, and customer feedback evaluation, this field has grown significantly in prominence within natural language processing.

Sentiment analysis functions at various granularity levels: sentence-level analysis establishes sentiment for individual sentences, document-level analysis categorizes the general sentiment of entire documents, and aspect-level analysis identifies sentiment toward particular features or aspects [13][8][14]. Using a document-level methodology, this study treats every review as a distinct classification unit. The three-class scheme (positive, neutral, and negative) provides sufficient granularity for obtaining actionable insights.

2.3. Support Vector Machine (SVM)

Support Vector Machine is a supervised learning algorithm that finds the best hyperplane in the feature space to maximize the margin between various classes in order to perform classification [15]. SVM uses kernel functions to map data into higher-dimensional spaces where linear separation is possible when working with linearly non-separable data.

Linear SVM proves particularly effective for text classification because TF-IDF transformed text data often exhibits linear separability in high-dimensional space [16] [15]. The regularization parameter C controls the trade-off between reducing classification errors and increasing the margin. Due to its strong performance in a variety of languages and application domains, SVM has become widely used in sentiment analysis research.

2.4. Naive Bayes

Naive Bayes is a probabilistic classifier that applies Bayes' theorem while making strong independence assumptions about features [17]. This algorithm performs surprisingly well in practice despite its "naive" assumption that features are conditionally independent given the class label, especially for text classification.

Multinomial Naive Bayes is ideal for text classification using TF-IDF features because it is specifically designed for discrete features like word counts [17]. It is particularly attractive for large-scale text classification applications due to its computational efficiency and short training time. The primary drawback is the independence assumption, which can limit performance when feature correlations are essential for precise classification.

2.5. Logistic Regression

Logistic Regression is a linear model designed for binary and multi-class classification that estimates the probability of class membership through the logistic function [18]. Using optimization techniques like gradient descent or L-BFGS, the algorithm learns feature weights that maximize the log-likelihood of accurate classifications.

Logistic regression usually uses either one-vs-rest or multinomial approaches when dealing with multi-class problems. By penalizing overly large weights, the regularization parameter helps prevent overfitting and enhances the model's capacity for generalization [18]. When the relationship between features and class log-odds is roughly linear, logistic regression performs well and yields results that are easy to understand.

2.6. Random Forest

Random Forest is an ensemble learning technique that builds multiple decision trees and combines their predictions through majority voting [19][20]. Each tree trains on a bootstrap sample of the data, and at each split point, only a random subset of features is evaluated. This randomization technique lessens tree correlation and aids in avoiding overfitting.

Random Forest is flexible in a variety of classification scenarios because it is excellent at capturing non-linear relationships and feature interactions that linear models are unable to handle [19]. The number of trees and maximum depth are important hyperparameters that have a significant impact on model performance. Although Random Forest usually yields high accuracy, it requires more processing power for both training and prediction and compromises some interpretability when compared to linear models.

3. RESEARCH METHODS

This study compares the effectiveness of machine learning algorithms in sentiment classification using a quantitative method and an experimental design. From data collection to model evaluation and deployment, our research framework adheres to a methodical procedure. The entire research workflow used in this study is shown in Figure 1.

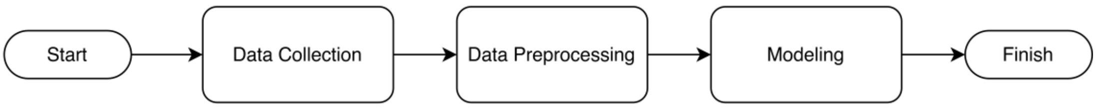

Figure 1. Research Flow Chart

The four primary phases of the research methodology are (1) data collection, (2) data preprocessing, (3) modeling, and (4) evaluation. To guarantee data quality, model dependability, and the validity of our findings, we meticulously planned each step.

3.1. Data Collection

Using the google-play-scraper library, web scraping of Google Play Store reviews was used to gather data. Kredivo, ShopeePay, Dana, GoPay, LinkAja, Bareksa, Flip, Jenius, and OVO were the nine main fintech apps in Indonesia that were the focus of the process. Three main factors were taken into consideration when choosing these apps: their widespread market appeal, sizeable user base, and representation of various fintech service categories.

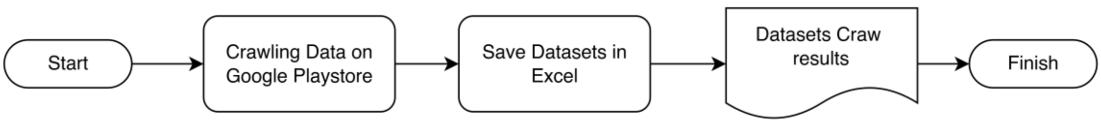

Figure 2. Data Collection Process

For every review, the following information was gathered by the scraping process: (1) review text; (2) rating score (1–5 stars); (3) reviewer name; (4) review date; (5) application name; and (6) developer response, if available. Between January 31, 2025, and November 18, 2025, data collection was carried out to gather recent user feedback that represents the state of service quality and user experience.

3.2. Data Preprocessing

A crucial step in turning unprocessed review text into clean, organized data that machine learning algorithms can use is data preprocessing [21]. This stage is made up of a number of sequential steps intended to address various aspects of the Indonesian language, such as slang terms, informal language, and inconsistent spelling.

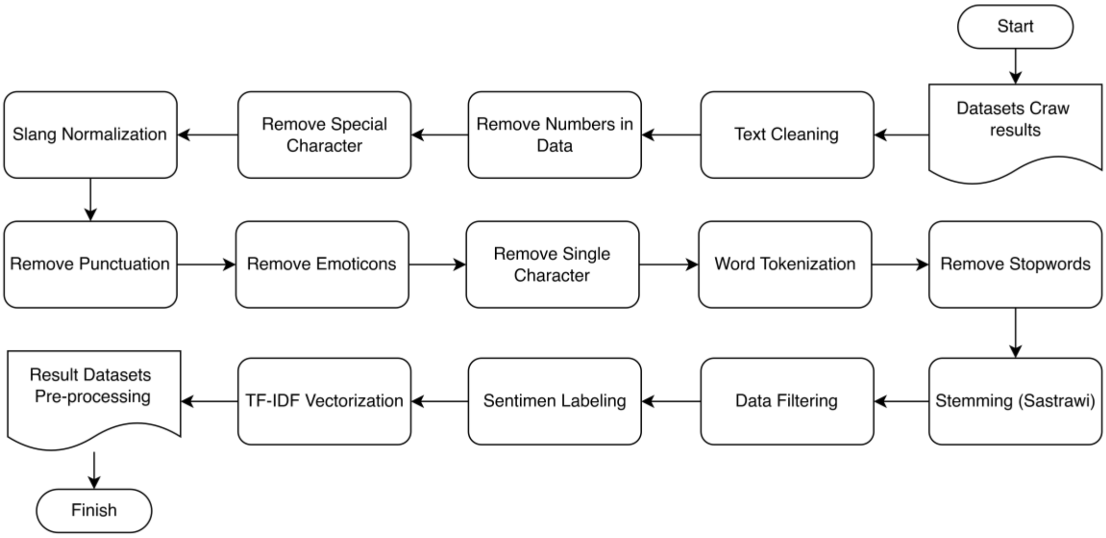

Figure 3. Data Preprocessing

The data preprocessing pipeline starts with the collection of raw datasets and moves through methodical cleaning and transformation steps. Steps include text cleaning, special character and number removal, slang normalization, punctuation and emoticon removal, single character removal, word tokenization, and stopword removal. Sentiment labeling assigns categories based on user ratings (1–2 stars for negative, 3 stars for neutral, and 4–5 stars for positive) after stemming using Sastrawi reduces words to their root forms. The processed text is then transformed into numerical representations using TF-IDF vectorization [22] [23].

3.3. Modeling

The modeling stage implements and compares four machine learning algorithms for sentiment classification: Naive Bayes, Logistic Regression, Support Vector Machine (SVM), and Random Forest. This stage starts by splitting the preprocessed dataset into training and testing sets using an 80:20 ratio for proper model validation. Each algorithm was trained and evaluated using the following standardized procedure. The preprocessed dataset was split into training (80%) and testing (20%) sets using stratified random sampling with a fixed random seed (random_state = 42) to ensure reproducibility and preserve the class distribution in both subsets [24].

Hyperparameter configurations for each algorithm were as follows: (1) SVM (Linear) used a linear kernel with regularization parameter C = 1.0; (2) Naive Bayes used Multinomial Naive Bayes with Laplace smoothing parameter alpha = 1.0; (3) Logistic Regression used the L2 regularization with C = 1.0 and the lbfgs solver with a maximum of 1,000 iterations; (4) Random Forest used 100 estimators with a maximum depth of None and minimum samples per split set to 2. These configurations were selected based on empirical performance on a held-out validation subset during preliminary experiments.

To ensure robust evaluation and mitigate the effect of data partitioning variance, a 5-fold stratified cross-validation was additionally performed on the training set for each algorithm. The final performance metrics reported in Table 2 including accuracy, precision (weighted), recall (weighted), and F1-score (weighted) — were computed on the held-out test set (20%) after training on the full training set, ensuring that the test set was never used during model selection or hyperparameter tuning.

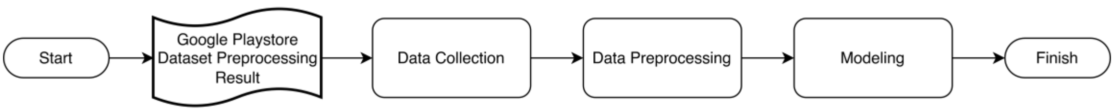

Figure 4. Modeling

4. DISCUSSION AND RESULTS

The complete results of the sentiment analysis are presented in this section, along with the dataset's features, preprocessing outcomes, comparative model performance, and a detailed examination of the top-performing model.

4.1 Raw Data Characteristics

The collected dataset consisted of 2,554 reviews from 9 fintech applications. Figure 5 presents a comprehensive visualization of the raw data distribution across multiple dimensions.

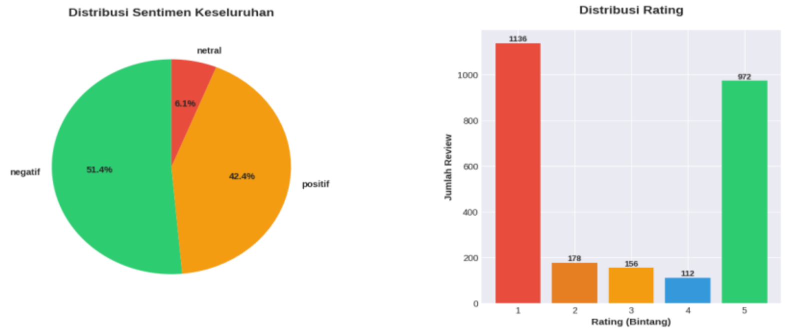

Figure 5. Overall Sentiment and Rating Distribution

There is a significant amount of space for service improvement throughout the fintech sector, as evidenced by the sentiment distribution, which is skewed toward negative sentiment (51.4%) as opposed to positive sentiment (42.4%) and neutral sentiment (6.1%). With peaks at 1-star (1,139 reviews) and 5-star (972 reviews), the rating distribution exhibits a prominent bimodal pattern that suggests polarized user experiences where users are more likely to express extreme feedback than moderate opinions.

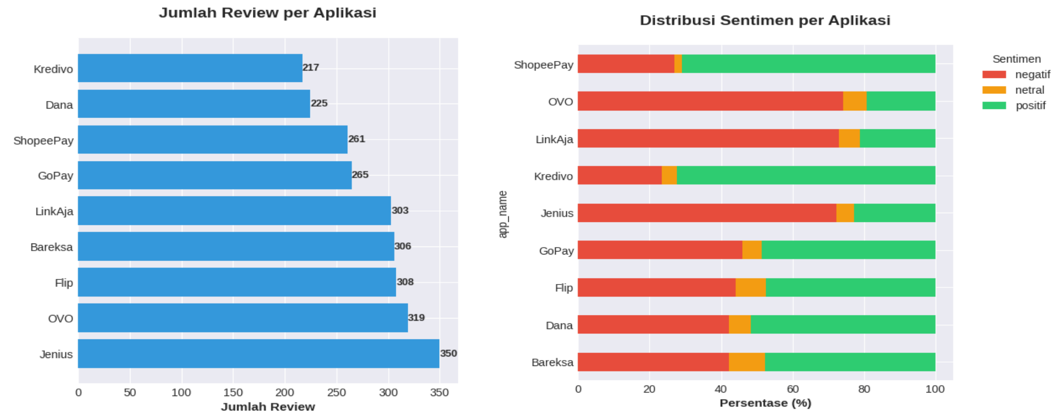

Figure 6. Application-wise Review Volume and Sentiment Distribution

The number of reviews varies greatly between applications; Jenius has the most (350), followed by LinkAja (303), OVO (313), and Flip (308), while Bareksa has the fewest (264). Over 70% of reviews on Kredivo and ShopeePay are positive, indicating high user satisfaction. OVO, LinkAja, and Jenius, on the other hand, have higher negative sentiment proportions—more than 50% of all reviews are negative—indicating possible problems with service quality.

4.2 Data Preprocessing

Raw review text was successfully converted by the preprocessing pipeline into standardized, clean data that could be used for machine learning. Examples of the preprocessing transformation are shown in Table 1.

Table 1. Preprocessing Transformation Examples

Original Review	After Cleaning	After Normalization	After Stopword Removal	After Stemming
"App nya lemot bgt!!!"	"app nya lemot bgt"	"app nya lambat banget"	"lambat"	"lambat"
"Gak bisa masuk terus nih"	"gak bisa masuk terus nih"	"tidak bisa masuk terus nih"	"masuk terus"	"masuk terus"
"Bagus banget membantu sekali"	"bagus banget membantu sekali"	"bagus banget membantu sekali"	"bagus bantu"	"bagus bantu"

While stopword removal removed 12,847 common words that have little discriminative value for sentiment classification, slang normalization successfully standardized 847 instances of informal terms. By reducing the vocabulary from 5,421 distinct word forms to 3,138 root forms, the stemming process improved model generalization by 42%. The vocabulary consolidation shows how effective the preprocessing was. For example, variations like "membantu," "bantuan," "dibantu," and "pembantu" are all reduced to the root "bantu" so that the model can identify their semantic similarity. In a similar vein, semantically related terms "pelayanan," "melayani," and "dilayani" converge to "layan.".

4.3 Feature Extraction Results

A feature matrix with dimensions of 2,554 × 3,138 and 99.22% sparsity was generated by the TF-IDF vectorization. The most discriminative terms for sentiment classification revealed by top TF-IDF scores include. Positive indicators: bagus (good), mudah (easy), cepat (fast), lancar (smooth), bantu (assistance), praktis (practical), aman (safe), and rekomendasi (recommend). Negative indicators: buruk (bad), lambat (slow), susah (difficult), error, masalah (problem), kecewa (disappointed), gagal (failed), and tolak (reject). Neutral indicators, biasa (ordinary), standar (standard), cukup (sufficient), and normal. Additional bigram features such as "sangat bagus" (very good), "tidak bisa" (cannot), "gagal terus" (keep failing), "mudah digunakan" (easy to use), and "buruk sekali" (very bad) capture sentiment intensity and negation.

4.4 Feature Extraction Results

Four machine learning algorithms were trained and evaluated on the preprocessed dataset. Table 2 presents a comprehensive performance comparison.

Table 2. Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	Training Time (s)
SVM (Linear)	0.8023	0.7779	0.8023	0.7853	0.34
Naive Bayes	0.8121	0.7699	0.8121	0.7832	0.12
Logistic Regression	0.8043	0.7603	0.8043	0.7781	0.28
Random Forest	0.7808	0.7450	0.7808	0.7581	2.47

The results show that all four algorithms achieved accuracy above 78%, demonstrating the effectiveness of the preprocessing and feature extraction pipeline. With 80.23% accuracy and a 78.53% F1-score, SVM (Linear) performed the best overall in terms of balanced metrics [25]. Naive Bayes came in second with 81.21% accuracy but slightly lower F1-score. The comparatively small performance differences imply that the TF-IDF feature space has linear separability and that the problem is clearly defined.

Significant variations in computational efficiency are revealed by training time analysis. The quickest algorithms were Naive Bayes (0.12 seconds), Logistic Regression (0.28 seconds), and SVM (0.34 seconds). Because Random Forest had to build 100 decision trees, it took a lot longer (2.47 seconds). SVM and Naive Bayes are especially appealing options for real-time applications due to their high accuracy and quick training.

4.5 SVM Detailed Analysis

A thorough analysis was carried out to comprehend SVM's classification behavior since it produced the best balanced performance. The SVM model's confusion matrix is shown in Figure 7.

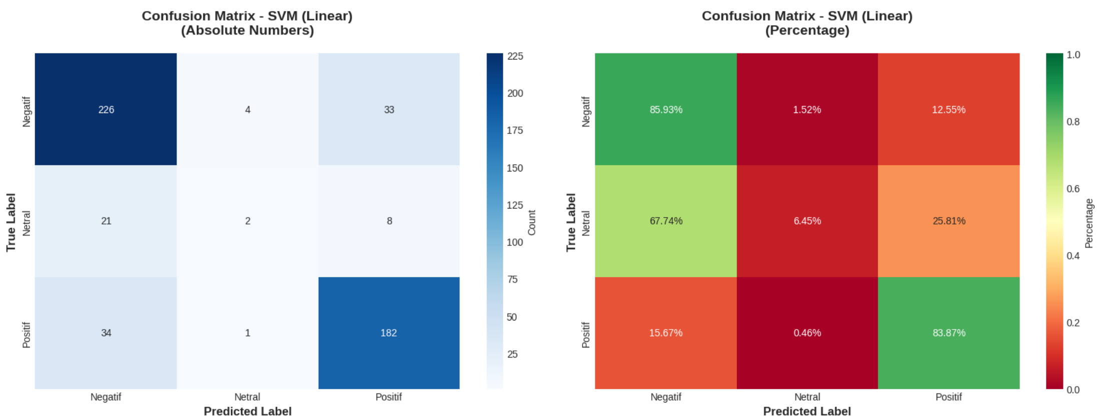

Figure 7. SVM Confusion Matrix (Absolute Numbers and Percentage)

SVM performs remarkably well on both positive and negative sentiment classification. The model correctly identified 182 out of 217 instances (83.87% recall) for positive reviews. The model correctly classified 226 out of 263 negative reviews (85.93% recall). With only 2 out of 31 cases correctly classified (6.45% recall), neutral sentiment proved to be the most difficult, primarily due to the scarcity of neutral samples (6.1% of the dataset) and the inherent ambiguity of 3-star ratings.

The neutral class exhibited a critically low recall of 6.45%, with only 2 out of 31 instances correctly classified. This constitutes a significant limitation of the model and warrants careful discussion. Three interconnected factors contribute to this weakness. First, severe class imbalance: neutral reviews represent only 6.1% of the dataset (157 out of 2,554 reviews), providing an insufficient number of training examples for the model to learn discriminative patterns for this class. Second, inherent label ambiguity: the neutral class was defined using 3-star ratings, which represent a gray zone where users may express mixed sentiments partially positive and partially negative within the same review. This ambiguity makes it difficult for any classifier to consistently distinguish neutral from positive or negative reviews. Third, linguistic overlap: the vocabulary used in neutral reviews shares substantial overlap with both positive and negative vocabulary, reducing the discriminative power of TF-IDF features for this class.

The misclassification of neutral reviews is predominantly toward the negative class (21 out of 31 instances, or 67.74%), suggesting that users writing 3-star reviews tend to include more negative language despite an overall neutral intent. Future work should address this limitation by employing oversampling techniques such as SMOTE for the neutral class, applying aspect-level sentiment analysis to capture mixed sentiments more accurately, or reconfiguring the task as a binary classification (positive vs. negative) to eliminate the ambiguous neutral class. Additionally, richer labeling strategies such as manual annotation by human raters rather than purely star-rating-based labeling could improve the reliability of neutral class boundaries.

4.6 Word Cloud Analysis

Word cloud visualizations provide intuitive understanding of vocabulary patterns across sentiment categories. Three word clouds generated from the analysis reveal distinct linguistic characteristics.

Figure 8. Positive Sentiment Word Clouds

Positive Sentiment Word Cloud: Terms like “bagus” (good), “mudah” (easy), “cepat” (fast), “bantu” (help), “lancar” (smooth), “praktis” (practical), and “aman” (safe) predominate. Users value dependability and practical advantages, as evidenced by the prevalence of functionality-related terms. Service-oriented words such as “bantu” and “terima kasih” (thank you) imply gratitude for assistance.

Figure 9. Neutral Sentiment Word Clouds

With words like “biasa” (ordinary), “cukup” (sufficient), “standar” (standard), “layan” (serve), and “proses” (process), the Neutral Sentiment Word Cloud has a more balanced vocabulary reflecting neither extreme satisfaction nor discontent.

Figure 10. Negative Sentiment Word Clouds

Negative Sentiment Word Cloud: Terms “buruk” (bad), “masuk” (enter/login), “dana” (funds), “transaksi” (transaction), “saldo” (balance), “gagal” (failed), “tolak” (reject), and “susah” (difficult) indicate operational problems. Words expressing emotional frustration include “kecewa” (disappointed) and “parah” (terrible). The word cloud analysis demonstrates that while negative sentiment concentrates on technical issues and service failures, positive sentiment is centered on functionality and ease of use. Enhancing transaction reliability, streamlining user interfaces, and fortifying customer support infrastructure are the improvement priorities guided by this insight.

5 CONCULUSSION

Using machine learning algorithms, this study successfully performed a thorough sentiment analysis of nine significant fintech applications in Indonesia. Using a strong preprocessing pipeline created especially for Indonesian text characteristics, the study gathered and examined 2,554 user reviews from the Google Play Store. Among the four machine learning algorithms evaluated, SVM (Linear) achieved the best overall balanced classification performance with an accuracy of 80.23%, precision of 77.79%, recall of 80.23%, and the highest F1-score of 78.53%. Although Naive Bayes recorded a marginally higher raw accuracy of 81.21%, its F1-score of 78.32% was slightly lower than SVM. The F1-score was prioritized as the primary evaluation metric because it accounts for class imbalance particularly the underrepresented neutral class and provides a more balanced measure of model performance across all sentiment categories. These results demonstrate the effectiveness of linear models for TF-IDF-transformed text data in Indonesian fintech sentiment classification.

The overall sentiment distribution shows 42.4% positive, 51.4% negative, and 6.1% neutral reviews, indicating significant room for service quality improvement across the fintech industry. Kredivo (72.4% positive sentiment) and ShopeePay (70.9% positive sentiment) have the best performance in terms of user satisfaction, while OVO (19.1%), LinkAja (21.1%), and Jenius (22.9%) have serious issues that need to be addressed immediately.

The most significant positive sentiment indicators include "bagus" (good), "bantu" (help), "mudah" (easy), and "cepat" (fast), while "buruk" (bad) and "kecewa" (disappointed) emerged as the most prominent negative indicators, along with transaction-related terms such as "gagal" (failed) and "tolak" (reject). These findings indicate that users value dependability, functionality, and ease of use. However, due to severe class imbalance and intrinsic label ambiguity, the model has the greatest difficulty with neutral sentiment classification (6.45% recall). Sarcasm and mixed-sentiment reviews also pose additional classification difficulties.

Based on the findings of this study, several directions for future research are recommended. First, the severe class imbalance for neutral sentiment (6.1%) should be addressed through oversampling techniques or manual annotation to improve neutral class classification. Second, deep learning approaches such as BERT-based models pre-trained on Indonesian corpora (e.g., IndoBERT) could be explored to capture contextual semantics that TF-IDF cannot represent. Third, the scope of this study could be extended to include a temporal analysis of sentiment trends across different time periods, as well as cross-platform comparison between Google Play Store and App Store reviews. Finally, aspect-based sentiment analysis could provide more granular insights into specific service attributes such as transaction speed, customer support, and interface usability.

ACKNOWLEDGMENT

The authors would like to thank Telkom University and all parties who supported this research.

REFRENCES

[1] D. P. Nugraha, B. Setiawan, R. J. Nathan, and M. Fekete-Farkas, "Fintech Adoption Drivers for Innovation for SMEs in Indonesia," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 8, no. 4, Dec. 2022, doi: 10.3390/joitmc8040208.

[2] B. Setiawan, T. D. Phan, J. Medina, M. Wieriks, R. J. Nathan, and M. Fekete-Farkas, "Quest for financial inclusion via digital financial services (Fintech) during COVID-19 pandemic: case study of women in Indonesia," *Journal of Financial Services Marketing*, vol. 29, no. 2, pp. 459–473, Jun. 2024, doi: 10.1057/s41264-023-00217-9.

[3] A. D. Widiantoro, A. Wibowo, and B. Harnadi, "User Sentiment Analysis in the Fintech OVO Review Based on the Lexicon Method," in *2021 6th International Conference on Informatics and Computing, ICIC 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICIC54025.2021.9632909.

[4] R. Nurcahyo, A. S. Prabuwono, A. F. Fainusa, N. Wibowo, M. Habiburrahman, and K. Hindriyandhito, "Enhancing User Satisfaction in Indonesia's e-wallet Market: A Comprehensive Analysis of Factors and Priorities," *Hum. Behav. Emerg. Technol.*, vol. 2023, 2023, doi: 10.1155/2023/8864865.

[5] W. A. Deviani, K. Kusumahadi, and E. Nurhazizah, "Service Quality For Digital Wallet In Indonesia Using Sentiment Analysis And Topic Modelling," *International Journal of Business and Technology Management*, Apr. 2022, doi: 10.55057/ijbtm.2022.4.1.6.

[6] B. Setiawan, "A Review of Sentiment Analysis Applications in Indonesia Between 2023-2024," 2024.

[7] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," Dec. 01, 2021, *Springer*. doi: 10.1007/s13278-021-00776-6.

[8] J. Singh, G. Singh, and R. Singh, "Optimization of sentiment analysis using machine learning classifiers," *Human-centric Computing and Information Sciences*, vol. 7, no. 1, Dec. 2017, doi: 10.1186/s13673-017-0116-3.

[9] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, "More than a Feeling: Accuracy and Application of Sentiment Analysis," *International Journal of Research in Marketing*, vol. 40, no. 1, pp. 75–87, Mar. 2023, doi: 10.1016/j.ijresmar.2022.05.005.

[10] C. H. Lin and U. Nuha, "Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy," *J. Big Data*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-023-00782-9.

[11] S. Liang, "The Future of Finance: Fintech and Digital Transformation," 2023.

[12] D. Tiwari, B. Nagpal, B. S. Bhati, A. Mishra, and M. Kumar, "A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 13407–13461, Nov. 2023, doi: 10.1007/s10462-023-10472-w.

[13] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, "A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges," Nov. 2022, [Online]. Available: http://arxiv.org/abs/2203.01054

[14] Y. C. Hua, P. Denny, J. Wicker, and K. Taskova, "A systematic review of aspect-based sentiment analysis: domains, methods, and trends," *Artif. Intell. Rev.*, vol. 57, no. 11, Nov. 2024, doi: 10.1007/s10462-024-10906-z.

[15] J. Ipmawati, S. Saifulloh, and K. Kusnawi, "Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 247–256, Jan. 2024, doi: 10.57152/malcom.v4i1.1066.

[16] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, "Enhancing machine learning-based sentiment analysis through feature extraction techniques," *PLoS One*, vol. 19, no. 2 February, Feb. 2024, doi: 10.1371/journal.pone.0294968.

[17] A. Bhavani and B. Santhosh Kumar, "A Review of State Art of Text Classification Algorithms," in *Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*, Institute of Electrical and Electronics Engineers Inc., Apr. 2021, pp. 1484–1490. doi: 10.1109/ICCMC51019.2021.9418262.

[18] A. Ichsan, S. Riyadi, and D. Pardede, "Analysis of Logistic Regression Regularization in Wild Elephant Classification with VGG-16 Feature Extraction," *Journal of Computer Networks, Architecture and High Performance Computing*, vol. 6, no. 2, pp. 783–793, Apr. 2024, doi: 10.47709/cnahpc.v6i2.3789.

[19] R. A. Disha and S. Waheed, "Performance analysis of machine learning models for intrusion detection system using Gini Impurity-based Weighted Random Forest (GIWRF) feature selection technique," *Cybersecurity*, vol. 5, no. 1, Dec. 2022, doi: 10.1186/s42400-021-00103-8.

[20] G. Biau and E. Scornet, "Rejoinder on: A random forest guided tour," Jun. 01, 2016, *Springer New York LLC*. doi: 10.1007/s11749-016-0488-0.

[21] A. W. Pradana and M. Hayaty, "The Effect of Stemming and Removal of Stopwords on the Accuracy of Sentiment Analysis on Indonesian-language Texts," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, pp. 375–380, Oct. 2019, doi: 10.22219/kinetik.v4i4.912.

[22] M. Das, S. Kamalanathan, and P. Alphonse, "A Comparative Study on TF-IDF feature Weighting Method and its Analysis using Unstructured Dataset," 2020.

[23] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning-Based Text Classification," Apr. 30, 2022, *Association for Computing Machinery*. doi: 10.1145/3439726.

[24] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, Jul. 2009, doi: 10.1016/j.ipm.2009.03.002.

[25] P. H. C. Samanmali and R. A. H. M. Rupasingha, "Sentiment analysis on google play store app users' reviews based on deep learning approach," *Multimed. Tools Appl.*, vol. 83, no. 36, pp. 84425–84453, Nov. 2024, doi: 10.1007/s11042-024-19185-w.