JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XII No 1, Desember 2025, hlm. 69 – 80 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/ 10.33330/jurteksi.v12i1.4189
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

COMPARISON OF NAÏVE BAYES, SVM, K-NN, DECISION TREE, AND RANDOM FOREST IN SENTIMENT ANALYSIS BASED ON SEABANK APPLICATION ASPECTS

Muhammad Al Fachrozi[1], Ken Ditha Tania[1*]
[1]Information Systems, Sriwijaya University
email: *muhammadalfachrozi03@gmail.com

Abstract: The increasing use of digital banking applications has led to the need for a deeper understanding of user perceptions, especially through aspect-based sentiment analysis. This study aims to classify the sentiment of SeaBank app users by focusing on four main aspects: learnability, efficiency, technical issues or errors, and satisfaction. Review data totaling 1,971 comments were collected from the Google Play Store and labeled with sentiments based on the scores (ratings) given by users. The CRISP-DM approach serves as the methodological framework for this study, which includes five classification algorithms: Naïve Bayes, Support Vector Machine (SVM), k-Nearest Neighbor (k-NN), Decision Tree, and Random Forest. The evaluation results show that the SVM algorithm provides the best performance with the highest average value of the four aspects achieving accuracy of 93.91%, Precision of 91.16%, recall of 97.96% and F1-Measure of 94.33%. According to the research findings, the Support Vector Machine (SVM) algorithm provides the best performance when performing aspect-based sentiment analysis on text data from digital banking application reviews. The findings are expected to serve as a reference for the development of automated evaluation systems that rely on user opinions as the basis for decision making.

Keywords: aspects; CRISP-DM; digital Banking; seabank; sentiment analysis

Abstrak: Peningkatan pemakaian aplikasi perbankan digital mendorong perlunya pemahaman yang lebih dalam mengenai persepsi pengguna, terutama melalui analisis sentimen berbasis aspek. Penelitian ini bertujuan untuk mengklasifikasikan sentimen pengguna aplikasi SeaBank dengan berfokus pada empat aspek utama: kemudahan dipelajari *(learnability)*, efisiensi penggunaan *(efficiency)*, kendala atau kesalahan teknis *(error)*, serta tingkat kepuasan *(satisfaction)*. Data ulasan berjumlah 1.971 komentar dikumpulkan dari Google Play Store dan diberi label sentimen berdasarkan skor (rating) yang diberikan oleh pengguna. Pendekatan CRISP-DM berfungsi sebagai kerangka metodologis untuk penelitian ini, yang mencakup lima algoritma klasifikasi: Naïve Bayes, Support Vector Machine (SVM), k-Nearest Neighbor (k-NN), Decision Tree, dan Random Forest. Hasil evaluasi menunjukkan bahwa algoritma SVM memberikan performa terbaik dengan nilai rata-rata dari ke empat aspek tertinggi yang mencapai *accuracy* sebesar 93.91%, *Precision* sebesar 91.16%, *recall* sebesar 97.96% dan *F1-Measure* sebesar 94.33%. Menurut temuan penelitian, algoritma Support Vector Machine (SVM) memberikan kinerja terbaik saat melakukan analisis sentimen berbasis aspek pada data teks dari ulasan aplikasi Seabank. Temuan ini diharapkan dapat menjadi referensi bagi pengembangan sistem evaluasi otomatis yang mengandalkan opini pengguna sebagai dasar pengambilan keputusan.

Kata kunci: Analisis Sentimen, Aspek, Bank Digital, SeaBank, CRISP-DM

INTRODUCTION

One area of human life that has been greatly affected by recent technological advances is finance and digital transactions [1]. Many banking companies are required to continue innovating in order to improve the quality of their services because digital banking applications offer a number of benefits, including ease of access, time efficiency, and high flexibility [2]. This application is part of a digital financial services ecosystem developed by a technology company that also oversees the Shopee e-commerce platform, giving it strong digital infrastructure support [3]. Through this application, users can carry out various banking transactions such as fund transfers, balance checks, bill payments, and financial management quickly and efficiently [4].

Despite offering various conveniences, the application's unstable performance and various other technical obstacles often trigger negative reviews from some users. These reviews are a valuable source of information because they reflect the level of satisfaction, complaints, and expectations of users regarding the services provided [5]. Given the volume of comments that continues to grow every day, application developers need an efficient approach to understand user opinions comprehensively and make data-driven decisions. In this context, sentiment analysis is a relevant solution. In this context, sentiment analysis is used to categorize user opinions into positive or negative categories by looking at the meaning that emerges in the review text [6] [7].

Therefore, an aspect-based sentiment analysis approach is important for exploring more specific information related to certain service dimensions. To classify sentiment based on these aspects, it is necessary to apply a machine learning algorithm-based classification method that is capable of efficiently handling large amounts of unstructured text data. Some algorithms that are often used to analyze sentiment include: Naïve Bayes, which classifies text based on the probability of word occurrence; SVM, which identifies the optimal separator (hyperplane) that maximizes the margin between classes; k-NN classifies data based on proximity or closest distance to training data, Decision Tree offers good flexibility in the classification process, and Random Forest works by combining multiple decision trees [8][9][10][11][12]. Each algorithm has its own advantages and limitations in terms of accuracy, computational speed, and resilience to noisy data.

This study specifically aims to compare the performance of five classification algorithms for sentiment analysis based on various aspects found in user reviews of the SeaBank application. Using review data from the Google Play Store, which has been separated based on aspects such as learnability, efficiency, error, and satisfaction, this study will evaluate the performance of each method based on metrics such as accuracy, precision, recall, and F1-Measure. The selection of these four aspects is based on the usability heuristic principle introduced by Nielsen and supported by the ISO 9241-11 standard, which includes the attributes of learnability, efficiency, memorability, error, and satisfaction, where memorability is treated the same as learnability because they have similar meanings [13]. The main contributions of this study include two things. First, this study presents an aspect-based review of the SeaBank application. Second, this study offers a comparative analysis of aspect-

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XII No 1, Desember 2025, hlm. 69 – 80 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/ 10.33330/jurteksi.v12i1.4189
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

based sentiment classification algorithms, which can serve as a methodological guideline for similar studies in the digital banking sector. This study is also expected to provide strategic insights for application developers, particularly in understanding user needs and experiences. By systematically and thoroughly analyzing user opinions, developers can obtain input to improve service quality in a more targeted manner.

METHOD

This study adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) model as a methodological framework, which provides systematic guidance in the implementation of the data mining process. CRISP-DM serves as a standard guide in the data mining process that can be widely applied to systematically solve various common problems [14]. This model is used to design and execute the process of analyzing user sentiment reviews of the SeaBank application based on certain aspects. However, this study is limited to the evaluation stage and does not include the deployment phase. This is because the main focus of the study is on comparing the performance of five classification algorithms in classifying sentiment based on certain aspects, rather than on developing a system that is ready for immediate implementation. Figure 1 is a diagram illustrating the stages in the CRISP-D model.

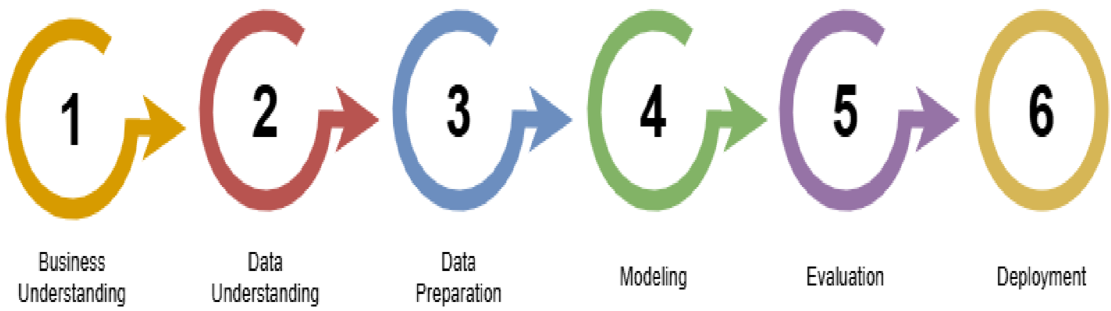

Image 1. CRISP-DM Model Flowchart [15]

Business Understanding

The purpose of this stage is to gain a comprehensive understanding of the issues and context surrounding the research. This research aims to explore the perceptions and experiences of SeaBank app users through an analysis of the reviews they have provided. This app is widely used by the public to conduct various financial transactions online. Through the review section on Google Play Store, users express their opinions, complaints, and satisfaction regarding several aspects of the app. This data has great potential to be analyzed in order to understand user sentiment towards various specific aspects of the app.

Data Understanding

This research data is sourced from user reviews found on the Google Play Store. The data collection process was carried out automatically using the Google Play Scraper Library to extract user comments. The reviews collected were filtered based on the latest content,

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XII No 1, Desember 2025, hlm. 69 – 80 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/ 10.33330/jurteksi.v12i1.4189
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

written in Indonesian, and originating from Indonesia between November 2024 and May 2025. A total of 2,000 user reviews were collected based on specific criteria to ensure that the reviews were easy to analyze and contained relevant information to support the research objectives [16]. From each data entry, the information collected includes the username, score or rating, review upload time (timestamp), and comment content. Table 1 presents some of the results of scraping the user reviews that have been collected.

Table 1. Seabank Review Scraping Results

Rating	Review
5	A very good digital banking app with excellent and easy-to use services
1	i can't log in to my account. I can't enter my PIN (NGe-BUG)...what should I do??????
3	registration is a bit difficult.. All the data is complete.. But registration still fails.. I don't know why.

After the data was collected, a cleaning stage was carried out. The purpose of this process was to produce clean data that was ready for further processing in the preprocessing and classification stages. After the data selection and cleaning process, 1,971 comments were collected that were relevant and in line with the needs and objectives of this study. All data was then labeled based on sentiment scores (ratings), and aspects contained in each review were identified through content analysis.

This labeling process uses three sentiment categories: positive, negative, and neutral [17]. The four main aspects that are the focus of analysis in this study are: learnability, which relates to the ease with which users understand and learn how to use the application; efficiency, which relates to the speed, performance, and operational efficiency of the application in conducting transactions; The error aspect covers technical complaints or problems such as failed logins, bugs, or system disruptions, and the satisfaction aspect highlights the overall satisfaction of users with their experience when using the SeaBank application.

Each aspect successfully identified from user reviews was then classified into two sentiment categories: positive, labeled with the number (1), and negative, labeled with the number (-1). Meanwhile, reviews that were not related to the aspects under study were labeled (0) as neutral or irrelevant.

Table 2. Data Labeling

Review	Learnability	Efficiency	Error	Satisfaction
A very good digital banking app with excellent and easy to use services	0	0	0	1
i can't log in to my account i can't enter my PIN NGeBUGwhat should I do	0	0	-1	0
it's a bit difficult to register the data is complete but it still fails	0	0	-1	0
it has comprehensive features and is easy to understand	1	0	0	0
any transaction is fast and hasle free	0	1	0	0

Data Preparation

This stage, also known as the preprocessing stage, is carried out to prepare the text data before it enters the classification process. The preprocessing process in this study begins with Transform Cases, which is converting all characters in the text data to lowercase letters to standardize the writing format and facilitate analysis. The next step is Tokenize, which breaks sentences into word units (tokens) based on spaces so that the text structure becomes more detailed. Next, Normalization is performed to standardize non-standard words or abbreviations so that their meanings are consistent. After that, the Stopword Removal process removes common words such as “yang”, “dan” or “di” that have little effect on sentiment analysis. Finally, Stemming is applied to return words to their basic form, so that variations of words with the same meaning are not considered different by the classification model.

Modeling

In the modeling stage, sentiment classification models based on aspects were developed using supervised learning algorithms, namely Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), Decision Tree, and Random Forest. Each model was trained to classify user sentiment towards predetermined aspects, namely learnability, efficiency, error, and satisfaction, into two sentiment classes: positive (1) and negative (-1). To overcome the problem of class imbalance in the training data, the SMOTE (Synthetic Minority Over-sampling Technique) technique was used to generate synthetic data in the minority class so that the model could learn more evenly [1].

Evaluation

The evaluation was conducted to determine how well each model performed in categorizing sentiment based on aspects. To ensure consistent results and prevent bias in data distribution, this study used the 10-fold cross-validation

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XII No 1, Desember 2025, hlm. 69 – 80 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/ 10.33330/jurteksi.v12i1.4189
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

evaluation technique. To provide a complete picture of the classification performance, evaluation metrics such as accuracy, precision, recall, and F1-Measure were used.

Accuracy
Measuring the proportion of correct predictions, using formula (1):

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

Precision
Shows the level of positive prediction accuracy, using formula (2):

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

Recall
the proportion of relevant data successfully recognized by the model from all data that should be positive, using formula (3):

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

F1-Measure
It is an evaluation metric obtained from the harmonic mean between precision and recall, with formula (4):

$$F - measure = \frac{2 \times precision \times recall}{precision + recall} \quad (4)$$

Description (1) (2) (3) (4) [18].:
TP (True Positive) = number of positive data correctly predicted by the model

TN (True Negative) = number of negative data correctly predicted by the model
FP (False Positive) = number of negative data incorrectly predicted as positive
FN (False Negative) = number of positive data incorrectly predicted as negative

RESULT AND DISCUSSION

Preprocessing Results

The reviews obtained from Google Play Store are still in the form of unstructured raw text, so a preprocessing step is required to ensure that the data can be optimally used in the classification stage. This step is carried out to produce clean, consistent text data that is suitable for processing by the classification algorithm. The final results of this preprocessing step are shown in Table 3 as a representation of the data transformation before and after cleaning.

Table 3. Preprocessing Result

Process	Preprocessing Result
Cleansing	Alhamdulillah i so can save every month to application seabank this insyaallah if i save continuously continuously so accumulate money savings i here THANKYOU VERY APPLICATION SEABANK can help life i to save application its safe guarded youknoww friends friends soo dont worry yea friends friends if want install application this guaranteed very
Case Folding	alhamdulillah i so can save every month to application seabank this insyaallah if i save continuously continuously so accumulate money savings i here thankyou very application seabank can help life i to save application its safe guarded youknoww friends friends so dont worry yea friends friends if want install application this guaranteed very

Tokenize	['alhamdulillah', 'i', 'so', 'can', 'save', 'every', 'month', 'to', 'application', 'seabank', 'this', 'insyaallah', 'if', 'i', 'save', 'continuously', 'continuously', 'so', 'accumulate', 'money', 'savings', 'i', 'here', 'thankyou', 'very', 'application', 'seabank', 'can', 'help', 'life', 'i', 'to', 'save', 'application', 'its', 'safe', 'guarded', 'youknoww', 'friends', 'friends', 'so', 'dont', 'worry', 'yea', 'friends', 'friends', 'if', 'want', 'install', 'application', 'this', 'guaranteed', 'very']
Normalization	alhamdulillah i so can save every month to application seabank this insyaallah if i save continuously continuously so accumulate money savings i here thankyou can application seabank can help life i to save application its safe guarded youknoww friends friends so dont worry yea friends friends if want install application this guaranteed very
Stopword Removal	alhamdulillah can save every month application seabank insyaallah save continuously continuously accumulate money savings thankyou can application seabank help life save application safe guarded youknow friends friends worry yea friends friends install application guaranteed very
Stemming	alhamdulillah can save each month application seabank insyaallah save continuously continuously accumulate money save thankyou can application seabank help life save application safe guard youknow friends friends worry yea friends friends install application guarantee very

Classification Model

After comments are labeled with sentiment based on scores in .CSV format, the data is loaded into RapidMiner using the Read CSV operator. Attributes with nominal types are converted to text format using the Nominal to Text operator so that they can be treated as text data and further processed in the analysis. Next, the data goes through a preprocessing stage using the Process Documents from Data operator, which includes text transformation processes such as case transformation, tokenization, normalization, stopword removal, and stemming. After generating a numerical representation in the form of TF-IDF, class balancing is performed using the SMOTE Upsampling operator to address the imbalance in data distribution between classes. The balanced dataset is then duplicated using the Multiply operator. This step allows the five classification algorithms to be tested simultaneously using the 10-fold Cross Validation method. By dividing the data into ten segments and testing each model several times, this approach produces more consistent and objective evaluation values.

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XII No 1, Desember 2025, hlm. 69 – 80 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/ 10.33330/jurteksi.v12i1.4189
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

Classification Model Evaluation

Table 4. *Accuracy* Result

Algoritma	*Learnability*	*Efficiency*	*Error*	*Satisfaction*
Naïve Bayes (NB)	92.95%	91.44%	97.72%	81.80%
Super Vector Machine (SVM)	99.09%	96.00%	90.88%	89.68%
k-Nearest Neigbors (k-NN)	61.14%	65.33%	51.31%	54.51%
Decision Tree (DT)	93.41%	90.78%	93.36%	83.77%
Random Forest (RF)	96.14%	94.78%	96.76%	88.27%

Table 5. *Precision* Result

Algoritma	*Learnability*	*Efficiency*	*Error*	*Satisfaction*
Naïve Bayes (NB)	88.24%	88.00%	95.79%	75.44%
Super Vector Machine (SVM)	99.11%	95.25%	84.86%	85.41%
k-Nearest Neigbors (k-NN)	56.31%	59.22%	50.67%	52.38%
Decision Tree (DT)	89.71%	86.69%	93.18%	97.17%
Random Forest (RF)	98.61%	96.05%	97.29%	94.34%

Table 6. *Recall* Result

Algoritma	*Learnability*	*Efficiency*	*Error*	*Satisfaction*
Naïve Bayes (NB)	100.00%	96.22%	100.00%	94.75%
Super Vector Machine (SVM)	99.09%	96.89%	100.00%	95.87%
k-Nearest Neigbors (k-NN)	100.00%	99.56%	100.00%	99.81%
Decision Tree (DT)	98.64%	96.67%	93.94%	69.59%
Random Forest (RF)	93.64%	93.56%	96.21%	81.58%

Table 7. *F-Measure* Result

Algoritma	*Learnability*	*Efficiency*	*Error*	*Satisfaction*
Naïve Bayes (NB)	93.75%	91.93%	97.85%	84.00%
Super Vector Machine (SVM)	99.10%	96.06%	91.81%	90.34%
k-Nearest Neigbors (k-NN)	72.05%	74.27%	67.26%	68.70%
Decision Tree (DT)	93.96%	91.41%	93.56%	81.10%
Random Forest (RF)	96.06%	94.79%	96.75%	87.50%

The classification model performance evaluation was conducted based on four main metrics, namely accuracy, precision, recall, and F1-Measure, against four main aspects determined in this study, namely learnability, efficiency, error, and satisfaction. Based on the test results, in general, the Support Vector Machine (SVM) algorithm showed the most superior and consistent performance in almost all aspects and evaluation metrics.

Accuracy

Based on Table 4, the Accuracy SVM results recorded the highest values in terms of learnability (99.09%) and efficiency (96.00%). This indicates the ability of the Super Vector Machine algorithm to consistently map data to the correct labels. In contrast, k-NN showed the lowest performance, especially in terms of error (51.31%) and satisfaction

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XII No 1, Desember 2025, hlm. 69 – 80 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/ 10.33330/jurteksi.v12i1.4189
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

(54.51%), indicating its limitations in handling complex text data.

Precision

In Table 5 Precision Results, Random Forest (RF) showed the best performance with an average of 96.57%, dominating in all aspects, including error (97.29%) and satisfaction (94.34%). In contrast, k-NN recorded significantly lower precision, particularly in terms of learnability (56.31%) and error (50.67%), indicating that most positive predictions by k-NN were inaccurate.

Recall

The Recall results in Table 6 show that k-NN recorded the best results with an average of 99.84% and a perfect score (100%) in terms of learnability and error. SVM came in second with an average of 97.96%, excelling in terms of satisfaction (95.87%). Conversely, Decision Tree (DT) recorded the lowest recall in almost all aspects, with the lowest value in the satisfaction aspect (69.59%).

F-Measure

The evaluation results in Table 7 show that the Support Vector Machine (SVM) algorithm obtained the highest F1 score with an average of 94.33% and recorded the highest value in terms of learnability at 99.10%. Random Forest (RF) followed with an average of 93.77% and the best performance in terms of error at 96.75%. Naïve Bayes (NB) showed stable performance across all aspects with an average of 91.88%, making it an efficient alternative choice. In contrast, k-Nearest Neighbor (k-NN) consistently recorded the lowest F1 score, with an average of only 70.57%, particularly low in the error (67.26%) and satisfaction (68.70%) aspects. This confirms that the k-NN algorithm is not suitable for sentiment classification on unstructured text data such as in this study.

After analyzing all five classification models, the average performance comparison results are displayed as a graph, as shown in Figure 2. Based on these results, the Support Vector Machine (SVM) algorithm shows the highest average value in almost all evaluation metrics compared to other algorithms, especially in the accuracy, precision, and F1-score metrics.

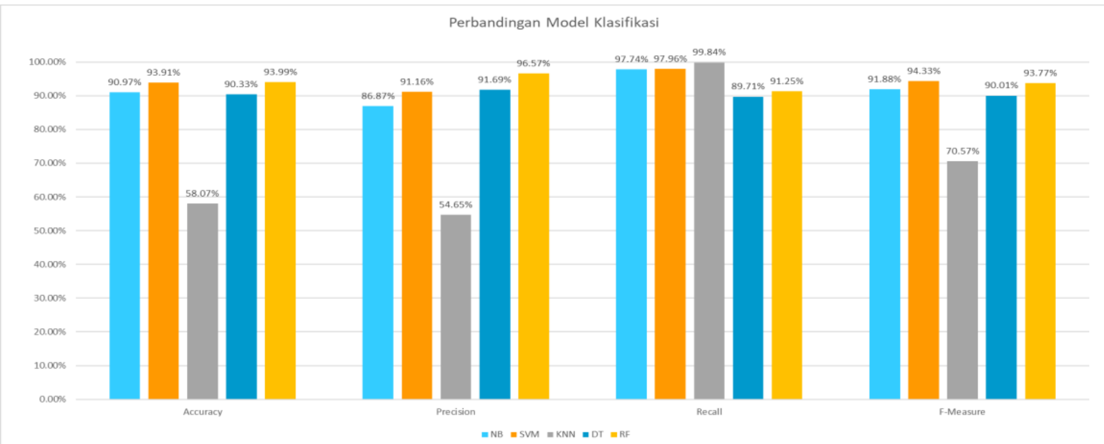

Image 2. Comparison of Classification Models

CONCLUSION

This study aims to analyze the sentiment of SeaBank application users based on four main aspects, namely learnability, efficiency, error, and satisfaction, using the Aspect-Based Sentiment Analysis approach and five classification algorithms with the CRISP-DM framework. Based on the evaluation results, the Support Vector Machine (SVM) algorithm showed the best overall performance, with the highest average F1-Measure value of 94.33%, accuracy of 93.91%, precision of 91.16%, and recall of 97.96%. Random Forest (RF) was the closest competitor with the highest accuracy of 93.99% and the highest precision of 96.57%, demonstrating its effectiveness in avoiding classification errors. On the other hand, k-NN showed the lowest performance with an F1-Measure of only 70.57%, making it less recommended in cases of aspect-based classification of unstructured text review data. According to these findings, the SVM algorithm is the best choice for aspect-based sentiment categorization in the Sea-Bank application review. These findings also reinforce that the aspect-based approach can provide deep insights into user perceptions, which can be utilized for more targeted development of digital banking application services.

BIBLIOGRAPHY

[1] D. W. Bhatara and R. R. Suryono, “Analisis Sentimen Aplikasi BCA Mobile Menggunakan Algoritma Naive Bayes dan Suport Vector Machine,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 4, pp. 1907–1917, Nov. 2024, doi: 10.29100/jipi.v9i4.5536.

[2] I. Tahyudin, A. Rafi Hananto, S. Anggun Rahayu, R. M. Anjani, and A. Nurhopipah, “Sentiment Analysis Model Development on E-Money Service Complaints,” *TEM J.*, pp. 2050–2055, Nov. 2023, doi: 10.18421/TEM124-15.

[3] N. Sakir, J. N. U. Jaya, and N. Wahyuni, “Penerapan Metode Pieces Framework Sebagai Evaluasi Tingkat Kepuasan Pengguna Aplikasi Seabank di Balikpapan,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 2, p. 344, Apr. 2022, doi: 10.30865/jurikom.v9i2.4047.

[4] H. A. Alnemer, “Determinants of digital banking adoption in the Kingdom of Saudi Arabia: A technology acceptance model approach,” *Digit. Bus.*, vol. 2, no. 2, p. 100037, 2022, doi: 10.1016/j.digbus.2022.100037.

[5] S. J and K. U, “Sentiment analysis of amazon user reviews using a hybrid approach,” *Meas. Sensors*, vol. 27, p. 100790, Jun. 2023, doi: 10.1016/j.measen.2023.100790.

[6] A. Alsokkar, M. Otair, H. E. Alfar, A. Y. Nasereddin, K. Aldiabat, and L. Abualigah, “Sentiment analysis for Arabic call center notes using machine learning techniques,” *J. Auton. Intell.*, vol. 7, no. 3, Jan. 2024, doi: 10.32629/jai.v7i3.940.

[7] V. Novalia, K. Ditha Tania, A. Meiriza, and A. Wedhasmara, “Knowledge Discovery of Application Review Using Word Embedding’s Comparison with CNN-LSTM Model on Sentiment Analysis,” in *2024 International Conference on Electrical Engineering and Computer*

Science (ICECOS), IEEE, Sep. 2024, pp. 234–238. doi: 10.1109/ICECOS63900.2024.10791113.

[8] A. Sabir, H. A. Ali, and M. A. Aljabery, "ChatGPT Tweets Sentiment Analysis Using Machine Learning and Data Classification," *Informatica*, vol. 48, no. 7, May 2024, doi: 10.31449/inf.v48i7.5535.

[9] Dhina Nur Fitriana and Yuliant Sibaroni, "Sentiment Analysis on KAI Twitter Post Using Multiclass Support Vector Machine (SVM)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 5, pp. 846–853, Oct. 2020, doi: 10.29207/resti.v4i5.2231.

[10] S. Masturoh and A. B. Pohan, "ANALISIS SENTIMEN TERHADAP E-WALLET DANA PADA ULASAN GOOGLE PLAY MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR," *J. Pilar Nusa Mandiri*, vol. 17, no. 1, pp. 53–58, Apr. 2021, doi: 10.33480/pilar.v17i1.2182.

[11] R. M. Harun and F. Fahmi, "ANALISIS SENTIMEN MENGGUNAKAN ALGORITMA NAÏVE BAYES, KNN, DAN DECISION TREE TERHADAP ULASAN APLIKASI KITALULUS," *Build. Informatics, Technol. Sci.*, vol. 6, no. 3, pp. 2033–2042, Dec. 2024, doi: 10.47065/bits.v6i3.6367.

[12] T. A.M and A. Yaqin, "Perbandingan Algoritma Naïve Bayes, K-Nearest Neighbors dan Random Forest untuk Klasifikasi Sentimen Terhadap BPJS Kesehatan pada Media Twitter," *InComTech J. Telekomun. dan Komput.*, vol. 12, no. 1, p. 01, Apr. 2022, doi: 10.22441/incomtech.v12i1.13642.

[13] G. E. Chanchí Golondrino, M. A. Ospina Alarcón, and L. M. Sierra Martínez, "DETERMINATION OF THE SATISFACTION ATTRIBUTE IN USABILITY TESTS USING SENTIMENT ANALYSIS AND FUZZY LOGIC," *Int. J. Comput. Commun. Control*, vol. 18, no. 3, May 2023, doi: 10.15837/ijccc.2023.3.4901.

[14] A. J. Syahid and D. Mahdiana, "Perbandingan Algoritma Untuk Klasifikasi Analisis Sentimen Terhadap Genose Pada Media Sosial Twitter," *semanTIK*, vol. 7, no. 1, p. 9, 2021, doi: 10.55679/semantik.v7i1.18087.

[15] J. W. Iskandar and Y. Nataliani, "Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 6, pp. 1120–1126, Dec. 2021, doi: 10.29207/resti.v5i6.3588.

[16] R. Q. Rohmansa, N. Pratiwi, and M. J. Palepa, "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI DISCORD MENGGUNAKAN METODE K-NEAREST NEIGHBOR," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 1, pp. 368–378, 2024, doi: 10.29100/jipi.v9i1.4943.

[17] R. Puspita and A. Widodo, "PERBANDINGAN METODE KNN, DECISION TREE, DAN NAÏVE BAYES TERHADAP ANALISIS SENTIMEN PENGGUNA LAYANAN BPJS," *J. Inform. Univ. Pamulang*, vol. 5,

no. 4, p. 646, Dec. 2021, doi: 10.32493/informatika.v5i4.7622.

[18] Rusydi Umar, Imam Riadi, and Purwono, "Comparison of SVM, RF and SGD Methods for Determination of Programmer's Performance Classification Model in Social Media Activities," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 2, pp. 329–335, Apr. 2020, doi: 10.29207/resti.v4i2.1770.