JITE, 7 (1) July 2023 ISSN 2549-6247 (Print) ISSN 2549-6255 (Online)

JITE (Journal of Informatics and Telecommunication Engineering)

Available online http://ojs.uma.ac.id/index.php/jite **DOI : 10.31289/jite.v7i1.9768**

Received: 12 June 2023	Accepted: 11 July 2023	Published: 28 July 2023

Aspect-Based Sentiment Analysis On FLIP Application Reviews (Play Store) Using Support Vector Machine (SVM) Algorithm

Nurul Hidayati1), Faqih Hamami2) & Riska Yanu Fa'rifah3)

1,2,3)Sistem Informasi, Fakultas Rekayasa Industri, Universitas Telkom, Indonesia

*Corresponding Email: nurulhidayatinh@student.telkomuniversity.ac.id

Abstrak

Perkembangan fintech telah mendorong pertumbuhan pesat e-wallet seperti Flip, yang menawarkan solusi praktis untuk transfer antarbank tanpa biaya administrasi. Ulasan pengguna di Play Store sangat penting untuk memahami pengalaman pengguna. Penelitian ini menggunakan analisis sentimen berbasis aspek (ABSA) dengan metode SVM untuk mendeteksi pendapat, persepsi, dan ulasan terkait aspek kecepatan, keamanan, dan biaya aplikasi Flip. Tujuannya adalah memberikan wawasan berharga kepada pengguna dan perusahaan mengenai pengalaman mereka dalam melakukan transaksi keuangan dengan Flip. Penelitian ini menggunakan dataset yang terdiri dari 13.500 data yang telah diproses dan dibersihkan, diikuti dengan vektorisasi TF-IDF. Data dibagi menjadi set pelatihan dan pengujian, menggunakan teknik seperti pemisahan latihan-ujian dan validasi silang K-Fold untuk menilai kinerja model. Analisis GridSearch menunjukkan bahwa kombinasi parameter tertentu, khususnya C=1.0 dan test_size=0.1, menghasilkan akurasi tinggi untuk semua aspek, dengan kernel linear memberikan akurasi tertinggi secara keseluruhan. Evaluasi model dilakukan menggunakan confusion matrix dan classification report, yang menyajikan akurasi, presisi, recall, dan skor F1 untuk setiap aspek. Secara mencolok, model Mesin Vector Pendukung (SVM) menunjukkan kinerja baik, terutama dalam aspek kecepatan, keamanan, dan biaya, di mana aspek biaya menunjukkan hasil yang sangat kuat. Secara ringkas, penelitian ini menggunakan ABSA untuk menganalisis ulasan aplikasi Flip, dengan model Mesin Vector Pendukung menunjukkan kinerja mengesankan dalam berbagai aspek, memberikan wawasan berharga bagi pengguna dan perusahaan yang menggunakan layanan transaksi keuangan Flip.

Kata Kunci: analisis sentimen berbasis aspek, support vector machine, ulasan, flip

Abstract

The development of fintech has driven the rapid growth of e-wallets like Flip, offering a convenient solution for interbank transfers without administrative fees. User reviews on the Play Store serve as crucial feedback for understanding the user experience. This research utilizes aspect-based sentiment analysis (ABSA) in combination with the SVM method to detect opinions, perceptions, and reviews pertaining to Flip's speed, security, and cost aspects. The objective is to provide valuable insights to both users and companies regarding their experiences with Flip in conducting financial transactions. The study employs a dataset comprising 13,500 preprocessed and cleansed data points, followed by TF-IDF vectorization. The data is divided into training and testing sets, utilizing techniques such as the train-test split and K-Fold Cross Validation to assess model performance. GridSearch analysis reveals that specific parameter combinations, notably C=1.0 and test_size=0.1, yield high accuracy across all aspects, with the linear kernel displaying the highest overall accuracy. Model evaluation is conducted using the confusion matrix and classification report, presenting accuracy, precision, recall, and F1-scores for each aspect. Notably, the Support Vector Machine model performs well, particularly in the speed, security, and cost aspects, where the cost aspect demonstrates exceptionally strong results. In summary, this study employs ABSA to analyze Flip application reviews, with the Support Vector Machine model showcasing impressive performance across various aspects, providing valuable insights for users and companies engaging with Flip's financial transaction services.

Keywords: aspect-based sentiment analysis, support vector machine, reviews, Flip

How to Cite: Hidayati, N., Hamami, F., & Fa'rifah, R. Y. (2023). Aspect-Based Sentiment Analysis On FLIP Application Reviews (Play Store) Using Support Vector Machine (SVM) Algorithm. JITE *(Journal of Informatics and Telecommunication Engineering)*, 7(1), 183-197.

I. INTRODUCTION

The advent of the digital era has brought innovation in day-to-day financial transactions through the use of fintech (Arner et al., 2015). A survey by the Ministry of Communication and Information Technology shows that 65.4% of respondents frequently use digital wallets (Ameliah et al., 2021). Based on an independent survey by Wise and JAKPAT, it is known that 70% of respondents admit that high costs pose difficulties when sending money. While occasional transactions may not have a significant impact, frequent transactions can result in substantial expenses due to administrative fees (Wise, 2023). E-wallets, such as Flip, provide the convenience of interbank transfers without administrative fees for transfers below Rp5,000,000 (Flip, 2023). However, despite Flip's advantage in reducing administrative costs, there are several shortcomings that need to be considered. Many user complaints can diminish loyalty, such as frequent transaction failures, difficult and time-consuming refund processes, inefficiency, poor customer service with inadequate answers and solutions, slow responsiveness, lack of recommendations to others, and a number of users uninstalling the Flip app (Harahap et al., 2020). This indicates that the Flip e-wallet lacks strong customer loyalty. The Flip app is difficult to use and not user-friendly for both new and existing customers, and it does not provide prompt responses (Harahap et al., 2020). As a result, Flip may lose customers and face difficulties in acquiring new customers because acquiring new customers is nearly five times more expensive than retaining existing customers, emphasizing the need to maintain loyalty (Derakhshanfar & Hasanzadeh, 2016). Therefore, in-depth research is needed to understand the factors influencing customer decisions to use the Flip app and how the company can enhance the user experience with the app. Although Flip has the advantage of addressing the administrative cost issue in money transfers, the company needs to make significant improvements to address the mentioned complaints, enhance customer loyalty, and reduce the risk of customer churn in the future (Derakhshanfar & Hasanzadeh, 2016).

The Flip app has been downloaded over 10 million times on the Play Store and has a rating of 4.5 based on 422 thousand user reviews (Play, 2022). The sentiment analysis can be depicted through the rating count on the Play Store, but it does not provide a detailed understanding of the users' app experience. Therefore, Aspect-based Sentiment Analysis (ABSA) is conducted based on the summary of user reviews of Flip. Negative and positive sentiments refer to various aspects, and sometimes the given reviews are specific to certain aspects (Adminlp2m, 2022). Aspect-based Sentiment Analysis or ABSA is a subarea of opinion mining that allows obtaining more in-depth information about the aspects referred to by users in mining reviews (Anand & Naorem, 2016). The sentiment analysis using the Aspect-Based Sentiment Analysis (ABSA) method can obtain more detailed insights into the usage of Flip (Anand & Naorem, 2016). In this study, the Support Vector Machine (SVM) algorithm is used for sentiment analysis (Brandusoiu & Brand, 2020). This research focuses on user reviews of the Flip app on the Google Play Store and utilizes text mining techniques for data analysis (Manning et al., 2009). The data preprocessing process is conducted to clean, normalize, and categorize the data (Adiwijaya, 2006). The results of sentiment analysis using the SVM algorithm can provide indications of the usage of the Flip app by users (Firdausi et al., 2022).

II. RESEARCH METHOD

A. Conceptual Model

A conceptual model is a design model that depicts the relationships among relevant factors in a research study and demonstrates their relevance to the research objectives. In this study, three main cycles are used: the Relevance Cycle, Design Cycle, and Rigor Cycle (Hevner, 2007). The model used in this research can be illustrated by Figure 1.

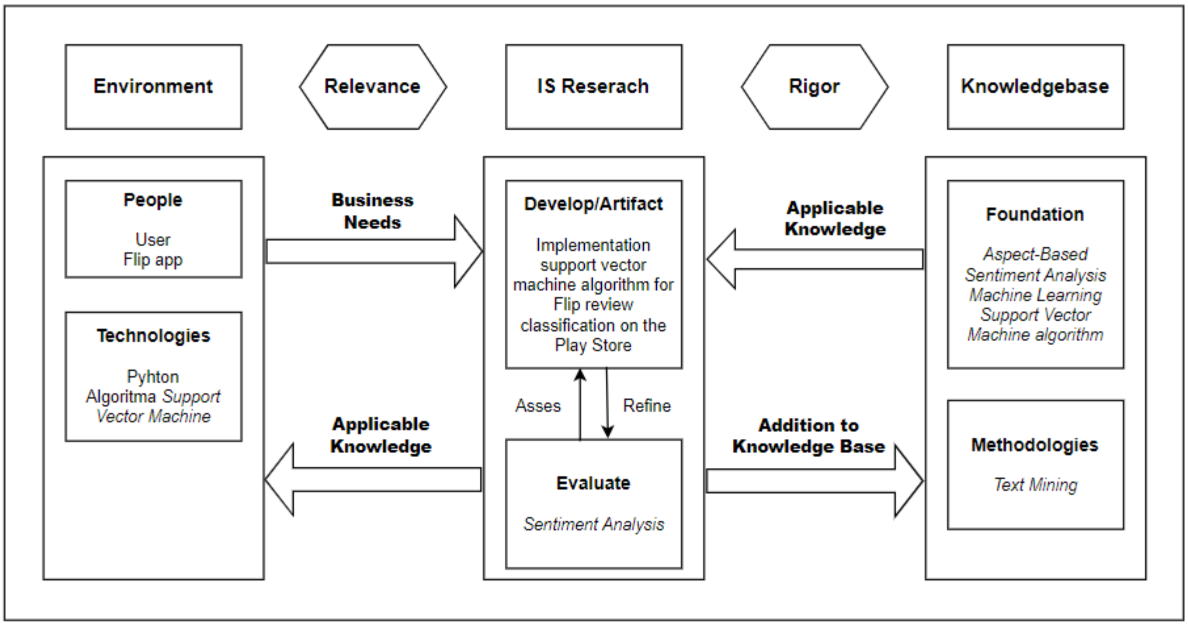

Figure 1 Conceptual Model

The environment of the Flip app consists of users who utilize the application for their financial transactions. The relevance of users in the context of Flip lies in the business relationship established with the app, which prompts the need for research and development to enhance user experience. As a result, aspect-based sentiment analysis using the Support Vector Machine (SVM) algorithm is implemented to gain insights into the Flip app. This analysis is conducted through text mining methodology, where user reviews are collected and evaluated using SVM and the Python programming language. Thus, the Flip app users, as part of the environment, play a crucial role in driving improvements and advancements in the application. Through aspect-based sentiment analysis with SVM, developers can acquire deeper insights into user experiences. The text mining methodology, utilizing the SVM algorithm and the Python programming language, is employed to efficiently gather and evaluate user reviews. With this approach, the Flip app can continuously be enhanced to better meet users' needs and expectations.

B. Systematic Research

The research framework is divided into three stages: initiation, implementation, and evaluation. Figure III.2 presents a diagram illustrating the research framework that will be developed. This diagram demonstrates how the research will be conducted to perform aspect-based sentiment analysis of Flip application reviews on the Play Store. Additionally, the research involves data training and data testing, as well as the pre-processing process and the utilization of the support vector machine method. These steps have been described to outline the research methodology.

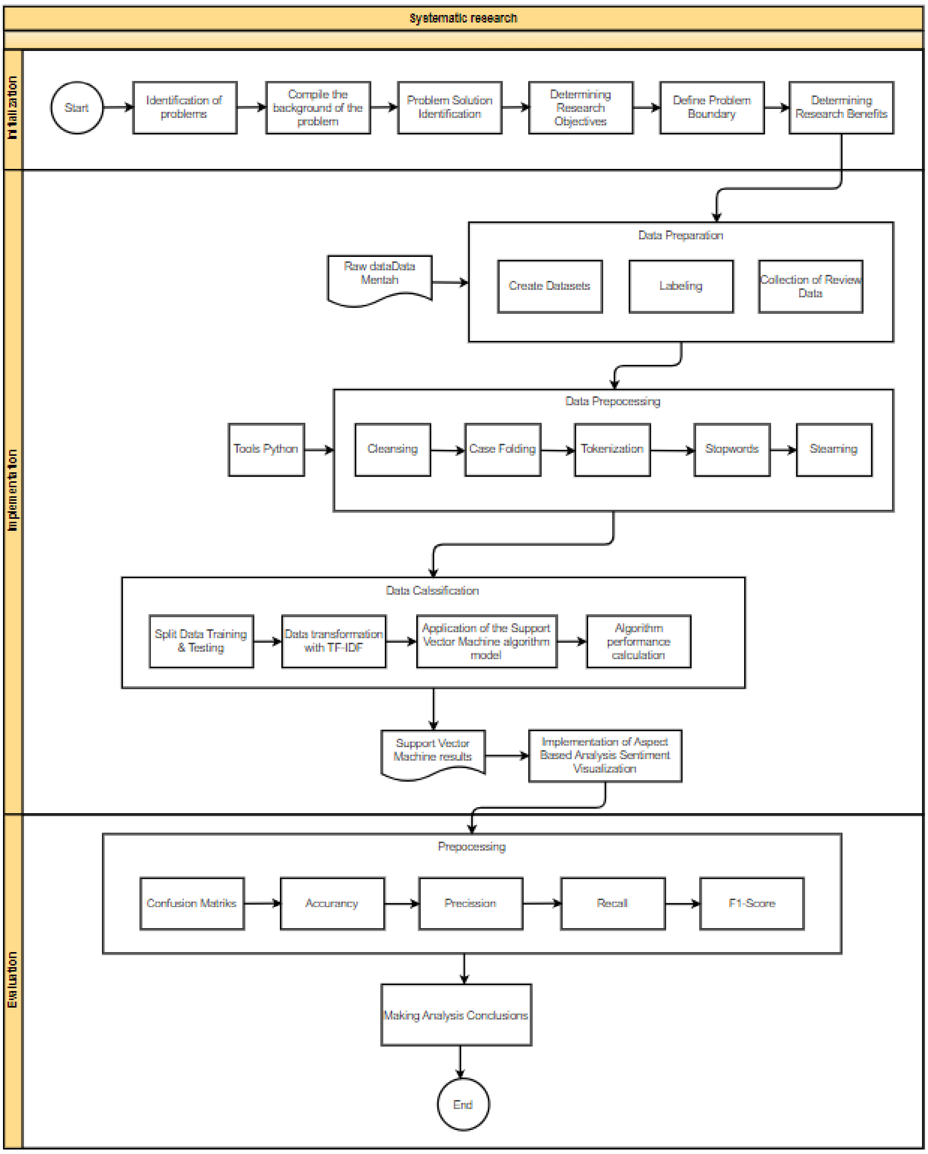

Figure 2 Systematic Research

The research methodology consists of three key phases: initiation, implementation, and evaluation. In the initiation phase, the researchers identify the problem by examining user reviews of the Flip app, understanding the challenges and issues faced by users. They then develop the research objectives to address these problems and determine the scope and limitations of the study. Additionally, potential solutions and their benefits are identified to highlight the significance of the research. In the implementation phase, the researchers collect raw data from user reviews and prepare a dataset for analysis. The collected data is labeled to classify sentiments expressed in the reviews. Preprocessing techniques, such as text cleansing, case folding, tokenization, removal of stopwords, and stemming, are applied to enhance the data quality. The dataset is divided into training and test sets. The training set is used to train a Support Vector Machine (SVM) model, which is a popular algorithm for text classification. The TF-IDF technique is employed to transform the dataset, assigning weights to each term based on its importance. The aspect-based sentiment analysis is implemented using the trained SVM model. In the evaluation phase, the performance of the aspect-based sentiment analysis model is assessed using evaluation metrics such as accuracy, precision, recall, and F1-score. These metrics provide insights into the model's effectiveness in classifying sentiments accurately. The results are analyzed to draw conclusions and implications for the Flip app. Overall, this research methodology ensures a systematic approach to understand user sentiments, improve the user experience, and provide valuable insights for the enhancement of the Flip app.

III. RESULT AND DISCUSSION

A. Data Understanding

The data collection process involves scraping the reviews or comments from the Flip application on the Play Store. The data is obtained using the scraping technique, which refers to the automated process of retrieving data from the Play Store using the Python programming language, specifically Google Colaboratory, integrated with Google Drive as the data storage. The application ID for Flip, "id.flip," is identified on the Google Play Store. Additionally, parameters such as Lang and country are used to specify the desired data, such as the Indonesian language ("id") and country. The sorting parameter can be set to sort.MOST_RELEVANT to arrange the search results or reviews based on their relevance to the Flip application. The scraping process results in data similar to that shown in Figure XX, consisting of 121,627 entries with 11 columns, including reviewId, userName, userImage, content, score, thumbsUpCount, reviewCreatedVersion, at, replyContent, repliedAt, and appVersion. The data scraping process focuses on extracting the "content" column from the Flip application reviews. Once the content data is collected, it can be exported as a .CSV file format.

B. Labeling Data

The dataset is in the .CSV file format and consists of 13,500 rows of data and attribute columns. The dataset is obtained after combining the entire dataset from the reviews of the Flip application on the Play Store with labels based on their aspect and sentiment. The labels indicate whether the reviews fall into the categories of security, speed, or cost aspect, and whether they have a positive, negative, or none sentiment. The scraped data is ready to be processed and classified based on the aspects of speed, security, and cost, as well as the sentiments of negative, positive, or none. A numerical scale is used to classify the sentiments into several categories. In sentiment classification using a scale of values 0, 1, and 2, there are three sentiment categories that can be identified. A sentiment value of 0 indicates that there is no significant aspect related to the application review, while a sentiment value of 1 signifies a positive perspective with expressions of satisfaction and praise for the application. On the other hand, a sentiment value of 2 reflects negative opinions or feelings, including criticism or dissatisfaction towards the application. By utilizing this value scale, we can classify sentiments into relevant categories based on positive, negative, or none views towards the application.

Table 1. Example for labeling data

content	speed	security	cost
So far, it has been very helpful to have the feature of free interbank transfers in the Flip app. It may be necessary to add a security menu to enhance the safety of the application. This would ensure a safer user experience by incorporating additional security measures.	0	2	1
The money transfer application with no administrative fees is indeed convenient, but there are occasional issues with the validation process taking too long. Moreover, if there is an error with the unique refund code, it can be challenging to obtain a refund, even after lodging a complaint.	2	2	1

C. Preprocessing Data

In the previously presented theoretical and methodological framework, the next step is data preprocessing. In this stage, data cleansing and refinement are performed to ensure that the data used have good quality by transforming the raw data into a more structured form that is ready for further analysis.

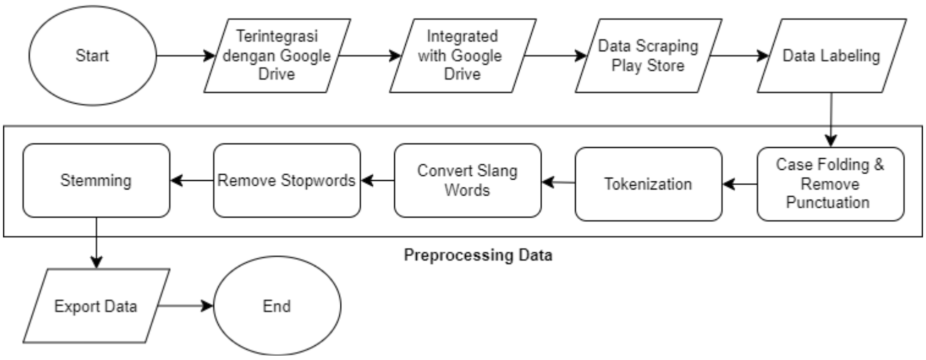

Figure 3 Preprocessing Data Flow

1) The preprocessing phase of text data in this study involves several steps to clean and prepare the text for further analysis. The first step is to convert all text to lowercase and remove punctuation marks using the `lower()` method and the `replace()` method, respectively. Next, numbers and emojis are removed using regular expressions with specific patterns. Excessive whitespace is eliminated by matching and replacing one or more spaces with a single space. Special characters, other than letters, are removed by replacing them with spaces using regular expressions.
2) After preprocessing, the next step is tokenization, which involves dividing the text into tokens or smaller parts. This is done to facilitate further text analysis. The Natural Language Toolkit (NLTK) library is imported to utilize the tokenizer provided by NLTK, and the `RegexpTokenizer` object is created with the pattern `\w+` to extract each word from the text, ignoring punctuation and spaces.
3) In text processing, slang words are converted into standard or formal words with the same or equivalent meaning. This process requires creating a dictionary or list of slang words along with their corresponding replacements. The purpose is to transform slang text into more formal and understandable text.
4) The subsequent step is removing stopwords, which aims to identify and filter out unnecessary words in the text. The NLTK library is used to import the stopwords module, which contains a collection of commonly used words that do not contribute significantly to text processing. In this study, a list of Indonesian stopwords is downloaded and utilized.
5) The final preprocessing step is stemming, which involves transforming words into their base form by removing prefixes, suffixes, and inflections. The Sastrawi library in Python is used for stemming based on the rules of the Indonesian language. The StemmerFactory object is created to generate a stemmer that is employed for stemming. The swifter package is also utilized to enhance the performance of the stemming process on the data.

In summary, the preprocessing phase consists of removing punctuation and converting to lowercase, tokenization, converting slang words, removing stopwords, and stemming. These steps are essential to prepare the text data for subsequent analysis and to improve the accuracy and effectiveness of the sentiment analysis conducted in this study.

D. TF-IDF Victorizer

After completing the data preprocessing process, the next step is the TF-IDF Vectorizer process. In TF-IDF, it is used to highlight words that contribute to the reviews by considering the frequency or occurrence of words across all the reviews. TF-IDF is capable of identifying relevant and important words in the reviews, and the result is a weight that indicates the importance of each word in the overall context of the reviews.

E. GridSearch

In the GridSearch process, the data has been divided into a training set and a test set with varying test ratios. The parameter values to be tested include the "C" value and the kernel function of the Support Vector Machine (SVM). The process is performed for each test ratio, and the best parameters are saved while displaying all the parameter combinations along with their corresponding accuracies. The GridSearch method is conducted with the following combinations of parameters: "C" values of 0.01, 0.1, 1, 10, and 100, and test ratios of 0.1, 0.2, and 0.3. The results of the parameter combinations obtained using the GridSearch method, grouped by their kernel types, can be seen in the table provided below:

Table 2. GridSearch Kernel Linear

Kernel Linear			
Test Size : 0.1			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.72	0.81	0.89
0.1	0.87	0.89	0.92
1.0	0.89	0.90	0.93
10	0.87	0.89	0.92
100	0.84	0.88	0.90
Kernel Linear			
Test Size : 0.2			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.69	0.82	0.88
0.1	0.86	0.90	0.92
1.0	0.87	0.91	0.93
10	0.85	0.90	0.92
100	0.83	0.89	0.90
Kernel Linear			
Test Size : 0.3			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.65	0.82	0.89
0.1	0.85	0.90	0.92
1.0	0.86	0.91	0.93
10	0.85	0.90	0.91
100	0.83	0.89	0.90

Table 3. GridSearch Kernel RBF

Kernel RBF			
Test Size : 0.1			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.59	0.81	0.86
0.1	0.84	0.86	0.91
1.0	0.88	0.90	0.94
10	0.89	0.89	0.93
100	0.89	0.89	0.93
Kernel RBF			
Test Size : 0.2			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.57	0.82	0.80
0.1	0.82	0.86	0.91

1.0	0.87	0.91	0.93
10	0.87	0.91	0.93
100	0.87	0.91	0.93

Kernel RBF			
Test Size : 0.3			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.56	0.82	0.55
0.1	0.81	0.86	0.91
1.0	0.86	0.91	0.93
10	0.86	0.91	0.93
100	0.86	0.91	0.93

Table 4. GridSearch Kernel Polynomial

Kernel Polynomial			
Test Size : 0.1			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.58	0.81	0.55
0.1	0.66	0.82	0.88
1.0	0.84	0.87	0.91
10	0.84	0.87	0.92
100	0.84	0.86	0.91
Kernel Polynomial			
Test Size : 0.2			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.57	0.82	0.56
0.1	0.64	0.83	0.87
1.0	0.83	0.89	0.90
10	0.83	0.88	0.90
100	0.83	0.88	0.90
Kernel Polynomial			
Test Size : 0.3			
C	Accuracy		
	Kecepatan	Keamanan	Biaya
0.01	0.56	0.82	0.55
0.1	0.62	0.82	0.88
1.0	0.81	0.88	0.90
10	0.82	0.88	0.90
100	0.82	0.87	0.89

Based on the results of the GridSearch, the following analysis can be made:

- The Linear kernel performs well across all test sizes. Setting the parameter C to either 1 or 10 yields good results with high accuracy for the aspects of speed, security, and cost. Choosing the Linear kernel in SVM can be expected to provide good results in terms of classification accuracy and modeling.
- The RBF kernel also produces good results, particularly for the speed and security aspects. Setting C to 1 or 10 leads to high accuracy. However, the RBF kernel may not perform as well for the cost aspect.
- The Polynomial kernel shows relatively lower performance compared to the Linear and RBF kernels in the dataset. Although there is some improvement in performance with increasing values of C and test size, the accuracy remains lower compared to other kernels. Therefore, in the context of using SVM for this dataset, the Polynomial kernel may not be the optimal choice.
- Although there is no direct impact observed from the test_size values on SVM performance in the data, it is important to select an appropriate test_size for comprehensive evaluation and validation of the SVM model.

In conclusion, it can be inferred that when C = 1, the Linear kernel exhibits high accuracy for the speed, security, and cost aspects. This indicates that SVM with C = 1 has good capabilities in classifying data with high accuracy for various aspects.

F. Splitting Data

In the simple data splitting method using train-test split in sklearn, based on the results of the GridSearch, the optimal combination of parameters is found to be C=1.0 and test_size=0.1. Therefore, in this simple splitting, we will use similar parameter values as obtained from Table V.

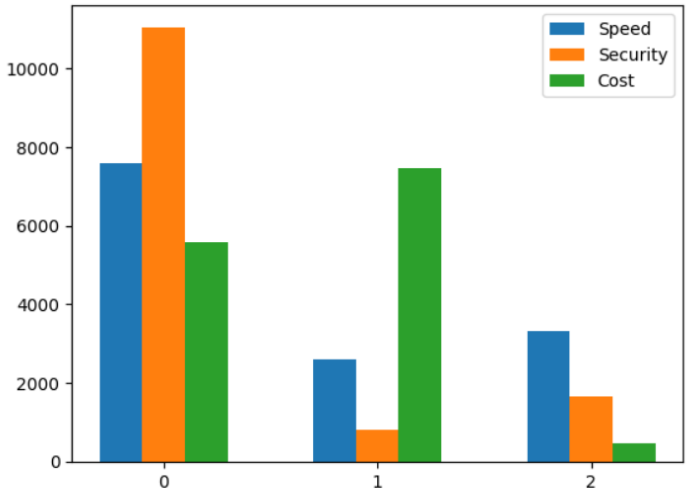

Figure 4 Splitting Data

Based on the results of data splitting, the number of test data for each category is as follows: for the speed category, there are 781 data; for the security category, there are 1,100 data; and for the cost category, there are 566 data. Furthermore, for the positive sentiment category (value 1), there are 245 data for speed, 76 data for security, and 748 data for cost. On the other hand, for the negative sentiment category (value 2), there are 324 data for speed, 174 data for security, and 36 data for cost. In conclusion, the distribution of test data in the speed, security, and cost categories is quite diverse, and there is variation in the number of data in each sentiment category.

Table 5. Splitting Data

Splitting Data (train_test split)				
Split Data	Accuracy			Average
	Speed	Securitty	Cost	
Train	0.90	0.93	0.95	0.93
Test	0.89	0.90	0.93	0.91

Average	0.90	0.92	0.94	

The SVM (Support Vector Machine) model used to predict the aspects of speed, security, and cost in the Flip application demonstrates good accuracy. The training data shows that the SVM model achieves an accuracy of 0.90 for the speed aspect, 0.93 for the security aspect, and 0.85 for the cost aspect, with an overall average accuracy of 0.93. In the test data, the SVM model also delivers satisfactory results with an accuracy of 0.89 for the speed aspect, 0.90 for the security aspect, and 0.93 for the cost aspect, with an overall average accuracy of 0.91. The analysis of the data reveals that the SVM model is effective in predicting the aspects of the Flip application with a good level of accuracy. There are variations in the average accuracy among the aspects of speed, security, and cost, where the cost aspect shows a higher average accuracy compared to security and speed. This indicates that the SVM model is capable of recognizing relevant patterns in the data to produce accurate predictions regarding the cost aspect. However, it is important to note that the evaluation of the SVM model should not solely rely on accuracy but also consider other factors such as precision, recall, and F1-score, which can provide a more comprehensive understanding of the model's performance.

G. K-Fold Cross Validation

In the K-Fold Cross Validation method, the data is divided into k subsets or folds. The cross-validation process with k=10 involves dividing the data into 10 equal-sized folds. In each iteration, one of the 10 folds is used as the test data, while the other 9 folds are used as the training data. This process is repeated 10 times, so that each fold is used as the test data once. The results of model evaluation using the K-Fold Cross Validation method can be seen in Table VI.

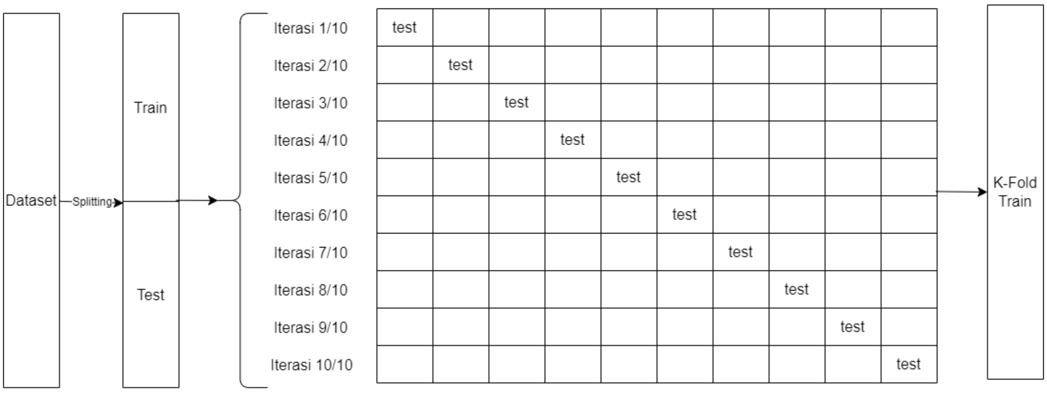

Figure 5 K-Fold Cross Validation

Table 6. K-Fold Cross Validation

K-Fold Cross Validation				
Split Data	Accuracy			Average
	Kecepatan	Keamanan	Biaya	
Fold 1	0.88	0.90	0.93	0.90
Fold 2	0.87	0.92	0.93	0.91
Fold 3	0.85	0.91	0.93	0.90
Fold 4	0.87	0.91	0.93	0.90
Fold 5	0.87	0.92	0.94	0.91
Fold 6	0.87	0.91	0.93	0.90
Fold 7	0.87	0.90	0.93	0.90

Fold 8	0.88	0.91	0.93	0.91
Fold 9	0.85	0.91	0.92	0.89
Fold 10	0.88	0.92	0.93	0.91
Average	0.87	0.91	0.93	

It can be analyzed that k-fold cross-validation was performed on the SVM (Support Vector Machine) model using 10 folds. The accuracy results for each fold show relatively high consistency, with a range of values between 0.85 to 0.88 for the speed aspect, 0.90 to 0.92 for the security aspect, and 0.92 to 0.94 for the cost aspect. By using k-fold cross-validation, the performance of the SVM model can be measured more comprehensively by dividing the data into multiple subsets and repeatedly training and testing the model. The average accuracy results from all the folds indicate that the SVM model has consistent and high accuracy, with average values of 0.87 for the speed aspect, 0.91 for the security aspect, and 0.93 for the cost aspect. k-fold cross-validation also helps evaluate the overall performance of the model, providing a more accurate understanding of the SVM model's ability to predict the aspects of speed, security, and cost in the application.

H. *Evaluation Confusion Matrix & Classifiation Report*

Confusion matrix is used to evaluate the performance of a classification model by providing information about the number of correct and incorrect predictions, as well as how well the model predicts each target class. The confusion matrix for all aspects is shown in Figure 6, Figure 7, and Figure 8.

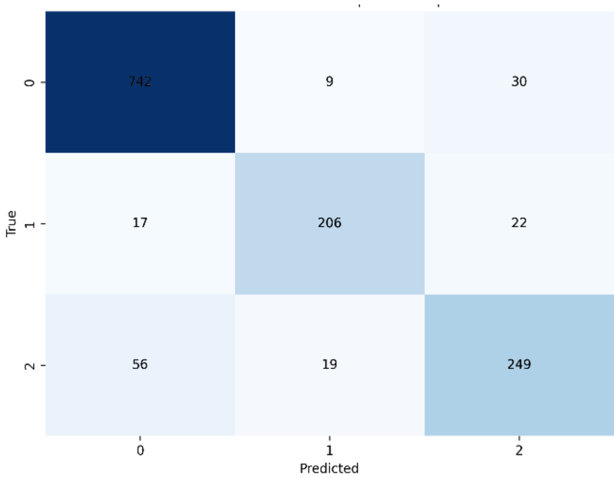

Figure 6 Confusion Matrix Aspect Speed

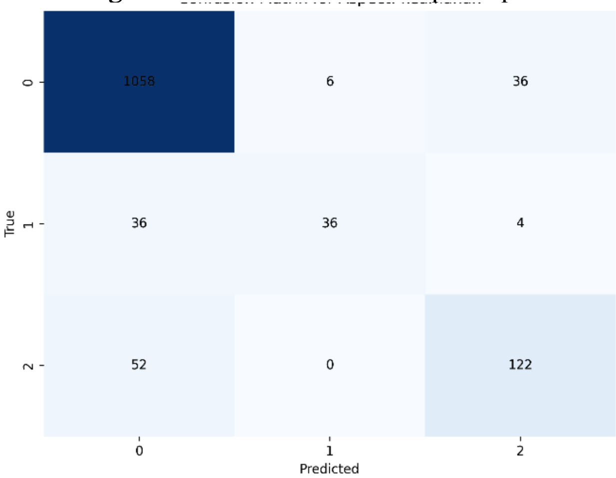

Figure 7 Confusion Matrix Aspect Security

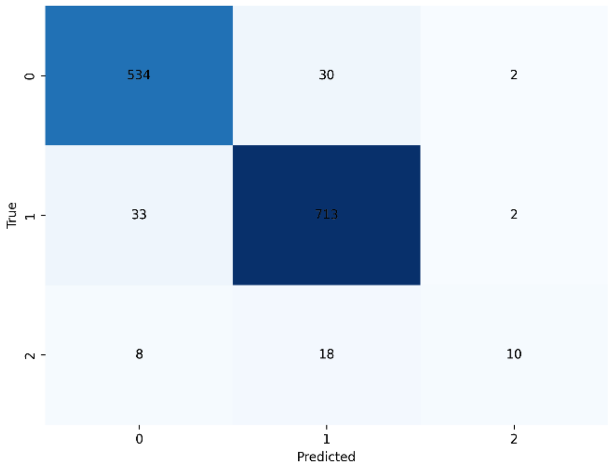

Figure 8 Confusion Matrix Aspect Cost

Overall, SVM performs well in predicting class 0 for all three aspects. However, improvements are needed in distinguishing between closely related classes and reducing the number of false positives, particularly in class 1. Further evaluation and parameter tuning can help enhance the performance of SVM in predicting the aspects of speed, security, and cost.

The purpose of a classification report is to present a comprehensive evaluation of the performance metrics of a classification model, providing information on how well the model can make predictions and classify data into the correct classes. Some commonly used evaluation metrics in a classification report include accuracy, precision, recall, and F1-score.

Table 7. Classification Report for Aspect

Classification report for speed aspect				
Class	Precision	Recall	F1-Score	Support
0	0.91	0.95	0.93	781
1	0.88	0.84	0.86	245
2	0.83	0.77	0.80	324
Accuracy			0.89	1350
Macro avg	0.87	0.85	0.86	1350
Weighted avg	0.89	0.89	0.89	1350
Classification report for security aspect				
Class	Precision	Recall	F1-Score	Support
0	0.92	0.96	0.94	1100
1	0.86	0.47	0.61	76
2	0.75	0.70	0.73	174
Accuracy			0.90	1350
Macro avg	0.84	0.71	0.76	1350
Weighted avg	0.90	0.90	0.90	1350
Classification report for cost aspect				
Class	Precision	Recall	F1-Score	Support

0	0.93	0.94	0.94	566
1	0.94	0.95	0.94	748
2	0.71	0.28	0.40	36
Accuracy			0.93	1350
Macro avg	0.86	0.72	0.76	1350
Weighted avg	0.93	0.93	0.93	1350

Overall, the SVM model demonstrated good performance in predicting the three evaluated aspects in this study. The consistent accuracy rate above 0.89 indicates the reliability of the model in making accurate predictions. However, there were variations in performance across classes, particularly in the security and cost aspects. The SVM model showed lower performance in predicting class 2 in the cost aspect, indicating challenges in classifying reviews related to FLIP application costs. The SVM model performed well in predicting the majority class across all aspects, but improvement is needed in predicting the minority classes (class 1 and class 2) in some aspects, particularly in the security and cost aspects. In conclusion, this research highlights the need for adjustments and improvements in the SVM model's performance to enhance the prediction quality, especially for classes that exhibit lower performance.

IV. CONCLUSION

In this research, the Aspect-Based Sentiment Analysis (ABSA) method was used to analyze reviews of the Flip application on the Play Store. The goal was to detect opinions, perceptions, and reviews related to the aspects of security, speed, and cost. The dataset consisted of 13,500 rows of data, which were cleaned and preprocessed to ensure data quality. The TF-IDF vectorizer was then applied to the data. The data was split into a training set and a test set for modeling purposes. Both the train-test split method and K-Fold Cross Validation were used to measure the performance of the model. The results obtained from GridSearch indicated that the combination of parameter values, particularly C=1.0 and test_size=0.1, provided high accuracy for all aspects, with the linear kernel yielding the highest accuracy overall. The evaluation was conducted using a confusion matrix and a classification report, which demonstrated the model's accuracy, precision, recall, and F1-score for each aspect. The Support Vector Machine model performed well, especially in the aspects of speed, security, and cost, with the aspect of cost showing particularly strong results.

In summary, the research employed the Aspect-Based Sentiment Analysis method to analyze reviews of the Flip application. The dataset was preprocessed and split into training and test sets. The modeling phase involved using both the train-test split method and K-Fold Cross Validation to evaluate the model's performance. The results obtained from GridSearch indicated that the combination of parameter values with C=1.0 and test_size=0.1 yielded the highest accuracy. The evaluation using a confusion matrix and classification report showed that the Support Vector Machine model performed well across all aspects, with the aspect of cost exhibiting particularly strong results.

BIBLIOGRAPHY

Adiwijaya, I. (2006). Texi Mining dan Knowledge Discovery. Komunitas Data Mining Indonesia & Soft-Omputing Indonesia.

Adminlp2m. (2022). Analisis Sentimen (Sentiment Analysis) : Definisi, Tipe dan Cara Kerjanya. PDAI - Universitas Medan Area.

Ameliah, R., Negara, R., & Rahmawati, I. (2021). No Title. In Status Literasi Digital di Indonesia 2021.

Anand, D., & Naorem, D. (2016). Semi-supervised Aspect Based Sentiment Analysis for Movies Using Review Filtering. Procedia Computer Science, 84, 86–93.

Anggi. (2023). Arti Google Play, Produk Layanan Dan Fungsinya. TEKNOVIDIA.

Arner, D., Barberis, J., & Buckley, R. (2015). The Evolution of Fintech: A New Post-Crisis Paradigm? Georgetown Journal of Internaational Law.

Babu, M. Y., Reddy, P. V. P., & Bindu, C. S. (2020). Aspect Category Polarity Detection Using Multi Class Support Vector Machine With Lexicons Based Features And Vector Based Features. Journal of Critical Reviews, 7(12).
Belete, D. M., & Huchaiah, M. D. (2022). Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. International Journal of Computers and Applications, 44(9), 875–886.
Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of Machine Learning Research, 13, 281–305.
Brandusoiu, I., & Brand, J. (2020). How To Fine-Tune Support Vector Machines For Classification.
Cahyaningtyas, C., Natalian, Y., & Widiasar, I. R. (2021). Analisis sentimen pada rating aplikasi Shopee menggunakan metode Decision Tree berbasis SMOTE. AITI: Jurnal Teknologi Informasi, 18, 173–184.
Carrasco, O. C. (2019). Support Vector Machines for Classification. Towards Data Science.
Center, N. D. R. (2021). Financial Technology (Fintech).
Chauhan, V. K., Dahiya, K., & Sharma, A. (2019). Problem formulations and solvers in linear SVM: a review. Artificial Intelligence Review, 52(2), 803–855.
Dang, S., & Ahmad, P. H. (2014). Text Mining: Techniques and its Application. IJETI International Journal of Engineering & Technology Innovations, 1(4).
Darapureddy, N., Karatapu, N., & Battula, T. K. (2019). Research of machine learning algorithms using k-fold cross validation. International Journal of Engineering and Advanced Technology, 8(6 Special issue), 215–218.
Derakhshanfar, R., & Hasanzadeh, M. (2016). The impact of customer experience and satisfaction on loyalty and word of mouth advertising in a Samen credit institution of Rasht. International Journal of Humanities and Social Sciences, 1, 1398–1406.
Dharmadhikari, S., Ingle, M., & Kulkarni, P. (2011). Empirical Studies on Machine Learning Based Text Classification Algorithms. Advanced Computing: An International Journal (ACIJ), 2.
Feldman, R., & Sanger, J. (2007). The Text Mining Handbook. Cambridge University Press.
Firdausi, A. J., Astuti, W., & Adiwijaya. (2022). Perbandingan Algoritma Klasifikasi SVM dan Naive Bayes Dalam Analisis Sentimen Pembelajaran Daring di Masa Pandemi COVID-19 di Twitter. E-Proceeding of Engineering, 9, 2065.
Flip. (2022). Tentang Flip. PT Fliptech Lentera Inspirasi Pertiwi.
Flip. (2023). Apakah ada limit transaksi di Flip untuk mendapatkan free biaya transfer?
Gaikwad, S. V., Chaugule, A., & Patil, P. (2014). Text Mining Methods and Techniques. International Journal of Computer Applications, 85(17).
Gorunescu, F. (2011). Data Mining Concepts, Models and Techniques. In Springer-Verlag Berlin Heidelberg, Intelligent Systems References Library (Vol. 12). Intelligent Systems Reference Library.
Harahap, F. Y., Fauzi, A., & Situmorang, S. H. (2020). THE INFLUENCE OF DIGITAL CUSTOMER EXPERIENCE AND ENJOYMENT ON FLIP E-WALLET E-LOYALTY THROUGH E-TRUST IN MEDAN CITY MILLENIAL GENERATIONS. 488–505.
Hevner, A. R. (2007). A Three Cycle View of Design Science Research. Scandinavian Journal of Information System, 19(2).
Johansson, R. (2015). Numerical Python A Practical Techniques Approach for Industry (S. Anglin (ed.)). Apress.
Jumba, J., & Wephukhulu, J. . (2019). Effect of Cashless Payments on the Financial Performance of Supermarkets in Nairobi County. International Journal of Academic Research in Business and Social Sciences, 9(3), 1372–1397.
Jumeilah, F. S. (2017). Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 1(1), 19–25.
Kanan, G., Al-Shalabi, R., AL_Khaleel, A., Muneera, A., Kanaan, T., & Salameh, D. bani. (2015). On The Scalability of Text Categorization Classifier Based On Arabic Text. International Symposium on INnovations in Intelligent SysTems and Applications.
Komunikasi, D. (2018, December). MENGENAL FINANCIAL TEKNOLOGI. Bank Indonesia.
Kumar, R. (2020). Machine Learning: Support Vector Regression (SVR). Analytics Vidhya.
Linkaja. (2022). Transfer Uang Lebih Mudah ke Bank Apa Pun Pakai LinkAja. Fintek Karya Nusantara.
Liu, B. (2010). Sentiment Analysis: A Multi-Faceted Problem. In IEEE Intelligent Systems.
Liu, B. (2011). Web Data Mining Exploring Hyperlinks, Contents, and Usage Data. In Springer (Second Edi). Springer-Verlag Berlin Heidelberg.
Liu, B. (2012). Sentiment Analysis and Opinion Mining. Springer Nature Switzerland AG 2022.
Liu, B. (2015). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press.

Lutz, M. (2011). Programming Python (J. Steele (ed.); 4th Editio). O'Reilly Media, Inc.
Mahendro, I. (2022). Analisa Kepuasan Mahasiswa Terhadap E-Learning Menggunakan Algoritma Support Vector Machine. Jurnal Saintek Maritim, 23(1), 1412–6826.
Manning, C., Raghavan, P., & Schütze, H. (2009). Introduction to Information Retrieval. Cambridge University Press.
Moraes, R., Valiati, J., & Gavião Ne, W. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. Expert Systems with Applications 40(2):621–633.
Murfi, H., Siagian, F., & Satria, Y. (2019). Topic features for machine learning-based sentiment analysis in Indonesian tweets. International Journal of Intelligent Computing and Cybernetics, 12(1), 70–81.
Play, G. (2022). Google Play. Https://Play.Google.Com/Store/.
Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., & Manandhar, S. (2014). SemEval-2014 Task 4: Aspect Based Sentiment Analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 27–35.
Ramdhani, G. (2020). Sering Lakukan Transfer Uang, Ini Cara Jitu Agar Bebas Biaya Administrasi. Liputan6.Com.
Ramsingh, J., & Bhuvaneswari, V. (2021). An efficient Map Reduce-Based Hybrid NBC-TFIDF algorithm to mine the public sentiment on diabetes mellitus – A big data approach. Journal of King Saud University - Computer and Information Sciences, 33(8), 1018–1029.
Saddam, M. A., Dewantara, E. K., & Solichi, A. (2023). Sentiment Analysis of Flood Disaster Management in Jakarta on TwitterUsing Support Vector Machines. Jurnal Dan Penelitian Teknik Informatika, 8(1), 470–479.
Syahputra, E. (2022). VP Flip Ungkap Komitmen Dorong Digitalisasi Solusi Keuangan. CNBC Indonesia.
Thaker, N., & Shukla, A. (2020). Python as Multi Paradigm Programming Language. International Journal of Computer Applications, 177(31), 38–42.
Ubaidillah, D, S., T, S., R, M., Fepiliana, Heriyanto, A., & Budiarto, R. (2019). Analysis WhatsApp Forensic and Visualization in Android Smartphone with Support Vector Machine (SVM) Method. Journal of Physics Conference Series 1196(1):012064.
Žižka, J., Dařena, F., & Svoboda, A. (2019). Text Mining with Machine Learning.