Jurnal Media Teknik Elektro dan Komputer
Vol. 02, No.02 (July - December, 2025)
E-ISSN: 3063-2927

Analysis of MyASN Application User Sentiment on Twitter Using the Support Vector Machine Algorithm

Amirah Nafiah Zalfa[1], Anisa Ninda Cahyani[2]

[1]Universitas Islam Negeri Sumatera Utara; amirah0701232082@uinsu.ac.id
[2]Universitas Islam Negeri Sumatera Utara; anisa0701232074@uinsu.ac.id

ABSTRACT

This study aims to analyze user sentiment toward the MyASN application based on public discussions on X. A total of 1059 tweets were collected through a web scraping process using the Tweet-Harvest tool on Google Colab, with data obtained from tweets containing keywords and hashtags related to MyASN and BKN. The collected data were preprocessed through case folding, tokenizing, normalization, stopword removal, and stemming, then represented using the TF-IDF weighting scheme. Sentiment labels were assigned into three categories: positive, negative, and neutral. The classification process employed the Support Vector Machine (SVM) algorithm, with data divided into 80% training data and 20% testing data. The experimental results show that the Support Vector Machine (SVM) algorithm achieved an accuracy rate of 98.1% with a precision value of 0.983, a recall of 0.981, and an F1-score of 0.982. Evaluation based on sentiment class shows that in the negative class, SVM produced a precision of 1.000, a recall of 0.977, and an F1-score of 0.989. In the neutral class, it achieved a precision of 0.929, a recall of 1.000, and an F1-score of 0.963, while in the positive class, SVM achieved a precision of 0.885, a recall of 1.000, and an F1-score of 0.939. These results show that the indicates that the implemented SVM model demonstrates strong reliability in handling text-based sentiment classification, particularly in datasets with imbalanced sentiment distributions. Overall, the results demonstrate that SVM is effective in capturing user sentiment patterns and can provide meaningful insights for evaluating and improving the MyASN application service.

Keywords: Sentiment Analysis, MyASN Application, Social Media X, SVM

Corresponding Author:
Amirah Nafiah Zalfa
Universitas Islam Negeri Sumatera Utara; amirah0701232082@uinsu.ac.id

1. INTRODUCTION

The development of information technology has had a significant impact on the way government agencies provide public services. In Indonesia, the State Civil Service Agency (BKN) has responded to this by launching the MyASN application, which is a web-based and mobile-based civil service portal aimed at making it easier for State Civil Servants (ASN) both PNS and PPPK to access data and civil service administration services. With MyASN, every civil servant can independently access their personnel data and profile information. The available features are also increasingly comprehensive and convenient, ranging from data updates and service requests to monitoring the status of submitted services (Septiandika & Nurrahmana, 2022). The existence of MyASN is a step forward towards efficient e-Government. However, the success of a digital application is not only measured by its availability, but also by the user experience and the level of user satisfaction with the application.

The information needs of a unit or agency vary greatly, so a platform or medium is needed to share the required information. One way to disseminate information about public services is through e-government or digital tools, such as social media. X is one of the most popular social media platforms adopted by the government in its efforts to help disseminate government information, partly because the number of X users continues to increase every year (Nurjanna & Ivanna, 2025). A report by We Are Social and Meltwater shows that X has reached 561 million users globally, with Indonesia ranking fourth with 24.69 million users in 2024 (Rizqika et al., 2025). This large number of users makes X a relevant source of data for observing public perception of a service.

The MyASN application owned by the State Civil Service Agency (BKN) has become one of the most discussed public services on X, through criticism, complaints, and user appreciation. These posts generate large amounts of unstructured text data that is constantly changing. Therefore, to systematically collect this information, the study employs an API-based tweet crawling approach using the Tweet-Harvest tool, which retrieves public tweets containing relevant keywords and hashtags related to MyASN. The collected data are then analyzed using sentiment analysis techniques to automatically classify user opinions into positive, negative, and neutral categories. Sentiment analysis is a measurement of people's opinions regarding the level of agreement on a particular topic, product, or service, or even an event (Ramlan et al., 2023). In this sentiment analysis, the study uses the Support Vector Machine (SVM) algorithm. SVM works by finding the best hyperplane that can separate data based on its class, making it effective for use in cases of classification into two or more sentiment categories.

Previous research shows that SVM is a highly effective algorithm for text classification. Research conducted by (Nurkholis et al., 2025) discusses the analysis of public sentiment towards the COVID-19 booster vaccine policy using the Support Vector Machine (SVM) algorithm. A total of 5,000 X data were collected during 2022. The data was cleaned and processed first, then converted into numbers using the TF-IDF method so that it could be processed by a computer. The test results showed that the OAO method with the RBF kernel provided good results with an accuracy of 96%. This was followed by research by Arfian et al., (2025), which analyzed sentiment regarding the KIP-K program using 941 X data points collected from X. Classification using TF-IDF and SVM produced an accuracy of 59%, with precision, recall, and F1-score values of 0.69, 0.75, and 0.72 for the negative sentiment class. Another study conducted by Ramlan et al., (2023) examined public sentiment related to fuel price increases using 258 X data samples. Sentiment classification using the SVM algorithm achieved an accuracy of 82.69%, with a recall value of 100% and acceptable precision and specificity values. Research conducted by Lestari & Berliani, (2024) analyzed X user sentiment regarding the dismissal of Sri Mulyani using the Naive Bayes and Support Vector Machine (SVM) methods with 1958 X data samples. Evaluation using a confusion matrix showed that the SVM model achieved precision, recall, and F1-score values above 95%, indicating that SVM performed well in classifying public sentiment towards the Sri Mulyani dismissal issue. Further research by Ma’rufudin & Yudhistira (2025), compared SVM with Naïve Bayes in non-linear text classification and found that SVM achieved higher accuracy and more stable F1-score values, particularly when handling high-dimensional text data.

In general, based on five previous studies, the average accuracy of SVM in sentiment classification ranges from approximately 59% to above 82%. Thus, SVM can be used in public sentiment analysis on various social media platforms. In addition to accuracy, SVM offers operational advantages such as effectiveness in high-dimensional feature spaces, robustness against overfitting, and flexibility in handling non-linear data through kernel selection.

The MyASN application has become an important part of the lives of civil servants (ASN) in Indonesia. The MyASN application has been downloaded more than 5 million times on Google Play Store to date. This shows the high level of significant application utilization among ASN. With a growing number of users, this application plays a crucial role in managing personnel data and information (Aditya et al., 2024). Given its broad user base and the sensitive nature of personnel data management, the security and service quality of the MyASN app are of utmost importance. Therefore, this study aims to provide a more complete picture of public opinion regarding MyASN, which is classified into positive, negative, and neutral sentiments using the latest X data and SVM algorithms. These findings are expected to serve as evaluation material for the State Civil Service Agency (BKN) in improving the quality of MyASN services in the future.

2. LITERATURE REVIEW (10 Pt)

2.1. Sentiment Analysis

Sentiment analysis is a computational process that analyzes digital text to determine whether the words or sentences conveyed have a meaningful or emotional message. Sentiment analysis understands and categorizes emotions into positive, negative, and neutral categories. This technique aims to identify public opinion into specific sentiment categories—positive and negative—so that it can provide an overview of public attitudes toward the service being studied (Iqbal et al., 2024). In conducting sentiment analysis, several important aspects must be considered, such as the topic being analyzed, relevant data sources, and the algorithms used for analysis. One algorithm that can be used is Support Vector Machine (SVM), which currently shows quite good performance in sentence classification (Arfian et al., 2025).

2.2. Evaluation of App Usage

Evaluation has related meanings, each referring to the application of several value scales to the results of policies and programs. In general, evaluation can be equated with appraisal, rating and assessment words that express an effort to analyze policy outcomes in terms of their value. Application usage evaluation is an activity to measure or explore all attributes of the system (in planning, development, implementation, or operation) and evaluations that will be carried out in relation to the acceptance of the system by end users. End users are one indicator for assessing the extent to which the use of health applications can provide convenience and benefits to the users of these applications (Sudirman et al., 2026). Therefore, the research results are expected to identify potential problems experienced by system users. Evaluation in the use of applications for public services is a concrete effort to determine the actual conditions of the implementation of an information system.

2.3. Implementation of SVM in Sentiment Classification

SVM is a technique used to perform sentiment classification. In the context of classification modeling, SVM offers a better and more mathematically defined concept than other classification approaches (Iqbal et al., 2024). Support Vector Machine (SVM) is a supervised learning-based machine learning algorithm that is effective for text classification by forming a hyperplane to separate data into different classes. This algorithm works by finding the optimal hyperplane that maximizes the margin between classes, thereby improving classification accuracy. SVM is very efficient in handling high-dimensional data and remains reliable even when the number of samples is smaller than the data dimension (Adrian et al., 2021). Because it can handle various data patterns, both linear and non-linear, SVM is often chosen to classify sentiment from social media, which has diverse data forms. Support Vector Machine (SVM) is a method that learns training data to assign classes or labels to objects. Support Vector Machine can learn to recognize handwritten numbers by examining a large collection of scanned images of handwritten zeros, ones, and so on (Siregar, 2023).

3. METHODS

The research process was carried out in a structured manner through several stages, starting from data collection using a scraper to evaluating the performance of the Support Vector Machine (SVM) algorithm. In general, the research stages can be seen in Figure 1.

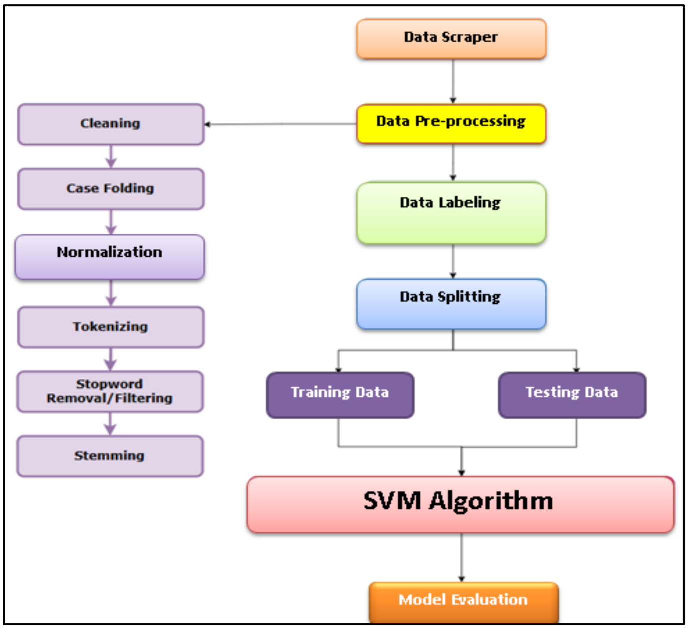

Figure 1. Diagram of the SVM Computation Process

3.1. Data Scraper

The dataset was collected through web scraping using Tweet-Harvest, which was run through Node.js commands using a data crawling method in Python programming language. The goal was to collect tweets containing the keywords “myASN OR myASN application OR BKN.” The data collection process was carried out using Google Colab, and the data was taken from tweets with hashtags related to the application. Google Colab was used as a platform to run machine learning and Python codes. Due to its ease of use, namely without the need for installation, Google Colab is the best place to learn Python (Nugrahanti et al., 2025). A total of 1060 tweets were collected between 2023 and October 2025. This dataset contains Indonesian text that will be used to perform sentiment analysis on the myASN application on X using the Support Vector Machine (SVM) method. The evaluation was carried out by dividing the dataset into 80% for training and 20% for testing, which ultimately showed a significant increase in the accuracy of the classification model. After data collection was completed, the data was stored in CSV format for further analysis.

3.2. *Data Pre-Processing*

This stage aims to clean the text of unnecessary elements and standardize the sentence structure so that it is ready for processing. Preprocessing in text mining refers to the words that appear most frequently in a document, which have little value and are of little help when selecting the documents needed by users (Mulyana & Lutfianti, 2023). The process follows the sequence shown in the diagram:

1. Cleaning: At this stage, data is cleaned of elements such as URLs, emoticons, punctuation marks, mentions, numbers, and characters that have no meaning for the analysis. The goal is to make the text content more focused on relevant words.
2. Case Folding: All text is converted to lowercase letters. This is to standardize the format so that words such as "myASN" and "*aplikasi myASN*" or "BKN" are considered the same.
3. Word Normalization: Non-standard words or abbreviations that often appear on social media are replaced with standard forms. For example, *"gk" becomes "tidak", 'mybknny' becomes "mybkn"*, and so on.
4. Tokenizing, Sentences are broken down into individual words. The results of tokenization are needed for filtering, stopword removal, and subsequent stages.
5. Stopword Removal/Filtering, Common words that do not affect meaning, such as *"yang", 'dan', "di",* and other conjunctions, are removed from the text.
6. Stemming: Each word is returned to its root form using the Sastrawi library. For example, *"berjalan"* becomes *"jalan"* and '*menggunakan*' becomes "*guna*".

The final result of this pre-processing is then visualized in the form of a word cloud to see the changes in word density before and after cleaning.

3.3. *Data Labeling*

After cleaning the text, the next step is to label the sentiment. In this study, the labeling process was carried out using a lexicon-based approach. Lexicon-based is used to identify and analyze sentiment or emotion in text by matching words in the text with a dictionary of words that have been categorized based on specific sentiments or emotions (Ratnaswari et al., 2025). In this lexicon-based approach, words in tweets are matched with a list of words with positive or negative polarity. If neither applies, the text is labeled as neutral. This labeling results in three sentiment categories: positive, negative, and neutral.

3.4. *Data Splitting*

The labeled data is then divided into two parts: training data and testing data. The division is done with a ratio of 80% training data and 20% testing data. The training data is used to train the SVM model, while the testing data is used to test how well the model predicts sentiment.

3.5. *SVM Algorithm*

The next step is the application of Support Vector Machine (SVM). Before entering SVM, the text is first converted into a numerical representation using TF-IDF, so that each word has a weight. In the TF-IDF method, the TF and IDF values are multiplied together to produce a term weight for each word in a document (Setiawan & Suryono, 2024). This method works by calculating how often a word appears in a document, then reducing the value of words that are too common through the inverse document frequency process. In this way, words that are more important or unique in the data set will stand out more. In sentiment analysis, TF-IDF serves to convert text in tweets into a numerical form that can be processed by algorithms such as Support Vector Machine. This approach helps the model recognize sentiment patterns better because the features that have the most influence on the text content are given higher weights. In Support Vector Machine, the formula for predicting the class of a sample x uses the following equation.

$$f(x) = sign(w \cdot x + b) \quad (1)$$

Description :
$f(x)$ is the decision function that classifies sample x
w is the weight vector
x is the input feature vector
b is the bias
$sign\ (\cdot)$ is the sign function that produces class labels

The SVM model is then trained using training data to form a separation between sentiment classes. Once the model is formed, testing data is used to see the model's performance in classifying tweets that have never been seen before.

3.6. Model Evaluation

The final stage is the evaluation of the classification results. The evaluation is carried out using several measurements, namely accuracy, precision, recall, and F1-score. In addition, a confusion matrix is also displayed to see the distribution of predictions between positive, negative, and neutral classes. With the confusion matrix, we can see how well the model predicts positive and negative classes, as well as how often the model makes prediction errors.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$f1 - score = \frac{2 \times Precision \times Recall}{Precision \times Recall} \quad (5)$$

Description :

True Positive (TP)	: The number of data successfully identified correctly as positive by the model.
False Positive (FP)	: The number of data that the model misclassified as positive.
True Negative (TN)	: The number of data correctly identified as negative by the model.
False Negative (FN)	: The number of data incorrectly classified as negative by the model.

This evaluation provides an overview of the success rate of the SVM model in analyzing sentiment from data taken from hashtags related to the MyASN application.

4. RESULTS AND DISCUSSION

4.1. *Data Scraper Results*

Data collection was carried out using Google Colab through a scraping process with the help of the tweet-harvest library, which was run using Node.js commands. The search keywords used were the hashtags #myasn, #aplikasimyasn, #bkn, and #id.go.bkn.mybkn. This process produced a CSV file containing public tweets related to the MyASN application. The file contents include tweet text, dates, and other information necessary for analysis. From this stage, hundreds to thousands of tweets were obtained, which formed the basis for the subsequent analysis process.

Table 1. Data Scraper Results

No	user_id_str	created_at	full_text
3	1.33694E+18	Tue Oct 28 13:28:28 +0000 2025	*@BKNgoid Min temen saya PPPK email untuk masuk ke MyAsn gak bisa masuk. Email pas PPPKnya mengalami error. Bagaimana yah??*
32	1.64845E+18	Wed Oct 08 03:10:04 +0000 2025	*@BKNgoid min bantuannya dong min layanan helpdesk myasn atau ada kanal alternatif? https://t.co/B81TEKaYjy*
135	2724964376	Tue Aug 05 13:15:57 +0000 2025	*..praktis aman dan transparan. Mulai dari updating data mandiri riwayat karier pengembangan kompetensi hingga proses mutasi dan pensiun-semuanya bisa dilakukan secara online tanpa harus datang ke kantor. Dengan MyASN layanan kepegawaian kini ada di genggaman tangan. Yuk ..*

Based on the data in Table 1, the process of collecting tweets through scraping resulted in a total of 1059 tweets discussing the MyASN application from various hashtags used. The data includes important information such as user names, tweet content, and upload times. The tweets in the first row show the most recent data entered during the collection process, while the bottom row shows older tweets within the scraping time frame.

The content of the collected tweets varies considerably, ranging from user complaints about technical issues, comments related to BKN services, to tweets that do not directly discuss the application but are still included because they contain relevant hashtags. This shows that public opinion about the MyASN application comes from a variety of conversational contexts. This raw data will then undergo cleaning and preprocessing before being analyzed in more depth to identify user sentiment trends towards the MyASN application.

In the initial stage, a wordcloud of the raw data is displayed to see the most frequently occurring word patterns before cleaning. The initial wordcloud is still dominated by irrelevant words such as links, mentions, or repetitive slang words. This initial form shows that the data is still messy and needs to be further processed for more accurate analysis results.

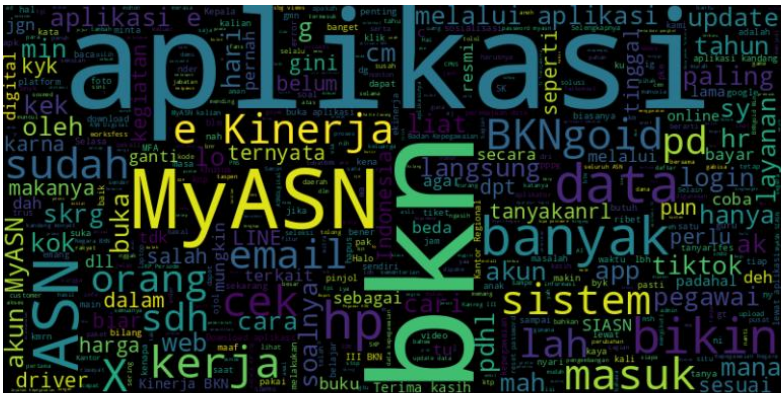

Figure 2. Wordcloud Before Pre-processing

4.2. *Data Pre-processing*

After the data collection stage, 1059 tweets were obtained. Next, the pre-processing stage was carried out to clean and prepare the data before further analysis. Cleaning was done to remove unnecessary characters or symbols, such as punctuation marks, numbers, and links. Case folding converted all text to lowercase to maintain data consistency. Word normalization aimed to replace non-standard words or slang with standard forms. Tokenization broke the text down into words or smaller linguistic units. Stopword removal eliminates common words that have no significant meaning in the analysis, such as "dan," "di," or "yang." Finally, stemming converts words to their basic form to simplify text analysis. The results of the preprocessing stage can be seen in Table 2.

Table 2. Pre-processing Results

Stage	Tweet
Tweet Data	*@BKNgoid Min temen saya PPPK email untuk masuk ke MyAsn gak bisa masuk. Email pas PPPKnya mengalami error. Bagaimana yah??*
Data Cleaning	*Min temen saya PPPK email untuk masuk ke MyAsn gak bisa masuk Email pas PPPKnya mengalami error Bagaimana yah*
Case Folding	*min temen saya pppk email untuk masuk ke myasn gak bisa masuk email pas pppknya mengalami error bagaimana yah*
Normalization	*min teman saya pppk email untuk masuk ke myasn tidak bisa masuk email pppk mengalami error bagaimana ya*
Tokenizing	*['min', 'teman', 'saya', 'pppk', 'email', 'untuk', 'masuk' ,'ke' ,'myasn' ,'tidak' ,'bisa' ,'masuk' ,'email' ,'pppk' ,'mengalami' ,'error' ,'bagaimana' ,'ya']*
Stopword/Filtering	*['min', 'teman', 'pppk', 'email', 'masuk', 'myasn', 'tidak', 'bisa', 'masuk', 'email', 'pppk', 'mengalami', 'error']*
Stemming	*min teman pppk email masuk myasn tidak bisa masuk email pppk alami error*

Based on Table 2, the preprocessing stage begins with data cleaning, which involves removing irrelevant elements such as the "@" symbol in the original tweet @BKNgoid. Next, the case folding process converts all letters to lowercase to maintain text consistency, for example PPPK becomes pppk.

In the tokenizing stage, the sentence is broken down into a list of separate words, for example "teman saya" becomes ['teman, 'saya]. Then, the stopword removal process removes words that are considered meaningless or do not contribute significantly, such as "ke", leaving only important words. Next, normalization corrects non-standard words to the standard form according to the data context, so that inconsistent words can be changed to the correct form. Finally, the stemming process simplifies words to their basic form, so that, for example, "mengalami" becomes "alami". To see the frequency of words that often appear in tweets at the filtering and stemming stages, see Figure 3 and Figure 4. This entire process aims to simplify and standardize the text so that it is ready for analysis.

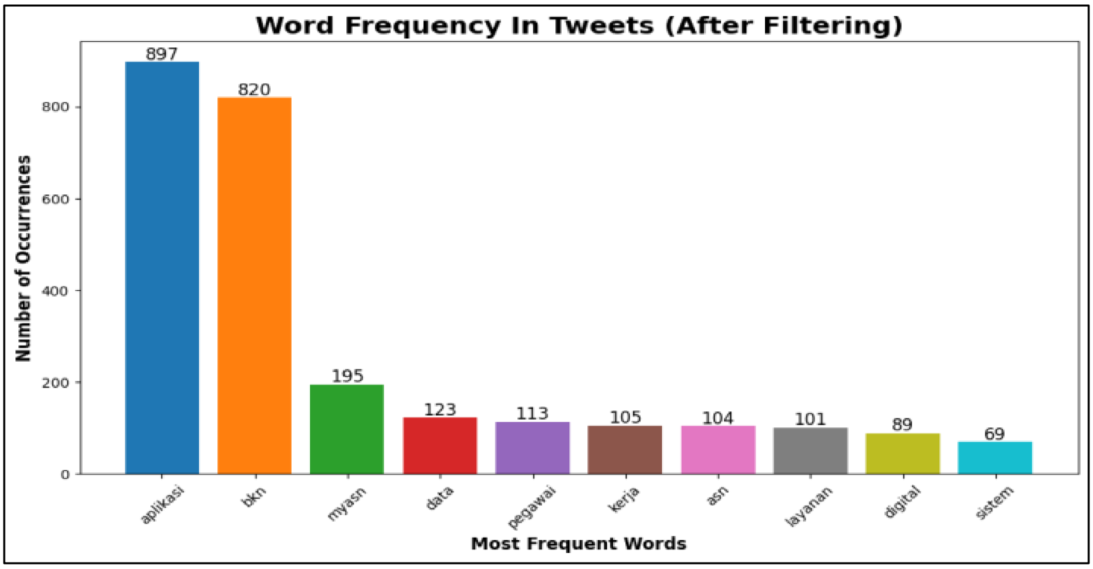

Figure 3. Word Frequency in Tweets After Filtering

After going through the cleaning and filtering process, Figure 3 shows the most frequently occurring words in tweets related to the MyASN application after filtering. The word "aplikasi" appears most often with 897 occurrences, followed by "bkn" with 820 occurrences, indicating that user discussions mainly focus on the application and the institution managing it. In addition, words such as "myasn", "data", "pegawai", and "asn" also appear frequently, showing that many tweets discuss employee-related data and services within the MyASN system. Other words like "layanan", "digital", and "sistem" suggest that users are also paying attention to the digital service aspects and system performance of the application.

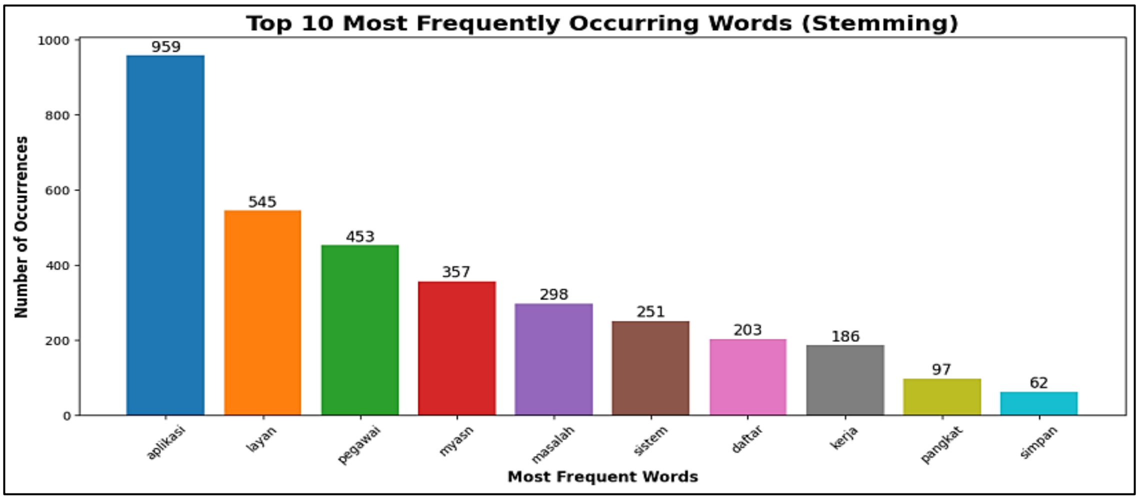

Figure 4. Word Frequency in Tweets at the Stemming Stage

Figure 4 above shows the ten most frequently occurring words after the stemming process. The word "aplikasi" remains the most dominant term with 959 occurrences, indicating that user discussions are primarily centered on the MyASN application. This is followed by "layan", "pegawai", and "myasn", which reflect frequent discussions related to services and civil servants as the main users of the application. Other commonly appearing words such as "masalah", "sistem", and "daftar" suggest that tweets often discuss service processes, usage experiences, and data-related matters. In addition, words like "kerja", "pangkat", and "simpan" indicate that users also talk about work-related aspects and data storage within the MyASN system.

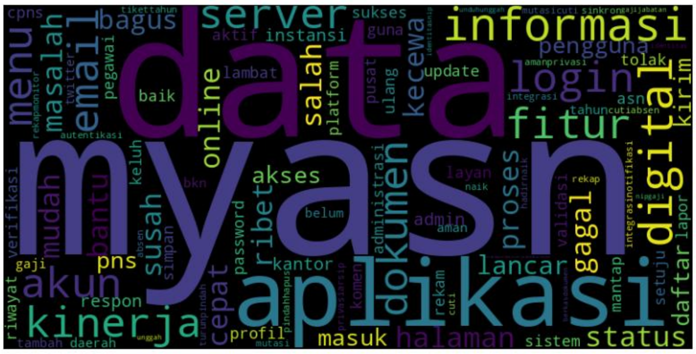

Figure 5. Wordcloud After Pre-processing

4.3. Data Labeling

After the preprocessing stage, 1,059 tweets were labeled using a lexicon-based approach with the Indonesian Sentiment Lexicon (INSET) dictionary. Each word in the tweet was compared to a list of positive and negative words in the dictionary. If a word was found in the positive list, it was given a score of +1, and if it was in the negative list, it was given a score of -1. The sentiment score was calculated by subtracting the number of negative words from the positive words. Tweets were categorized as positive if the score was greater than 0, negative if less than 0, and neutral if the score was equal to 0. The results are presented in Table 3.

Table 3. Data Labeling Results

No	user_id_str	Tweet	Score	Sentiment
3	1.33694E+18	*bkngoid min teman saya pppk email untuk masuk ke myasn tidak bisa masuk email pas pppknya alami error*	-4	Negative
32	1.64845E+18	*bkngoid min bantu dong min layan helpdesk myasn atau ada kanal alternatif*	0	Neutral
135	2724964376	*praktis aman dan transparan mulai dari updating data mandiri riwayat karier kembang kompetensi hingga proses mutasi dan pensiunsemuanya bisa laku cara online tanpa harus datang ke kantor dengan myasn layan pegawai kini ada di genggam tangan yuk*	5	Positive

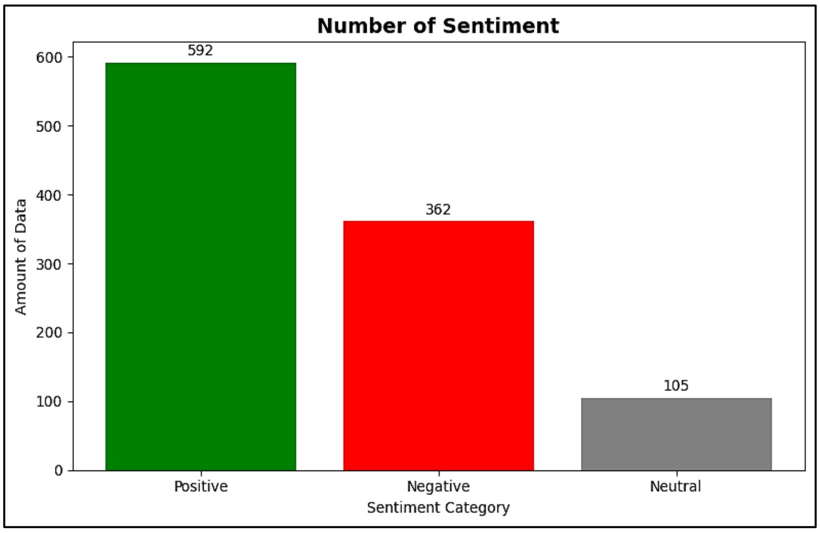

Figure 6. Sentiment Classification Results

Figure 6 shows the classification results with the number of tweets in each sentiment category, namely 592 positive tweets, 362 negative tweets, and 105 neutral tweets. This imbalance indicates a data imbalance between classes, with positive sentiment being the most dominant category. This condition can affect the performance of the SVM model because the model tends to more easily recognize patterns in classes with larger amounts of data compared to classes with less data, such as neutral and negative sentiment.

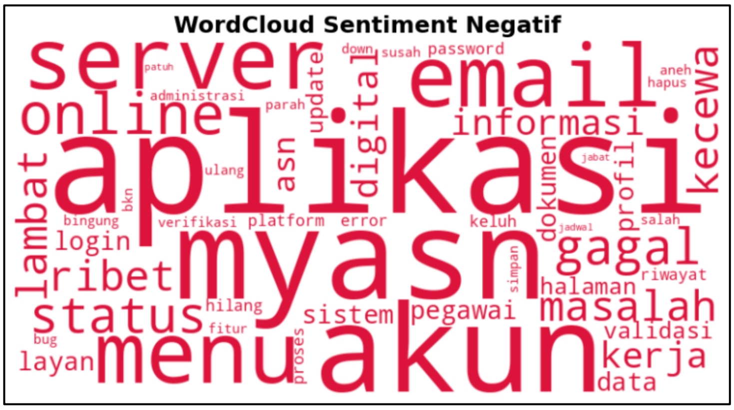

Figure 7. Negative Sentiment WordCloud

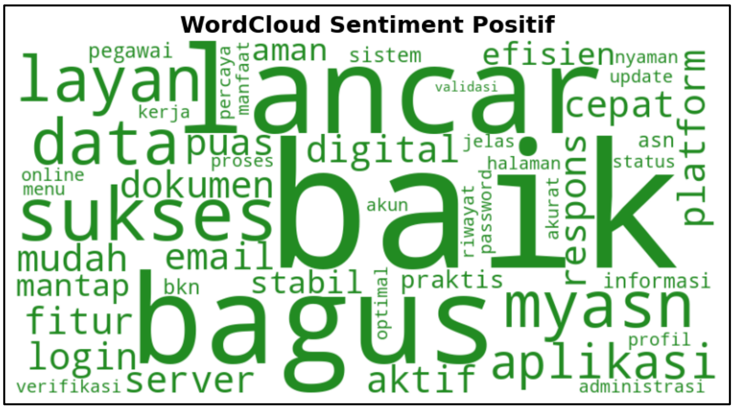

Figure 8. Positive Sentiment WordCloud

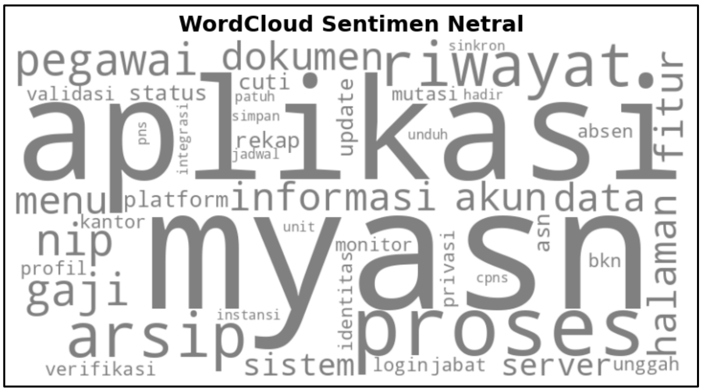

Figure 9. Neutral Sentiment WordCloud

4.4. *Data Splitting*

The dataset has been divided into two parts: training data and test data. The division was made with a ratio of 80% for training data and 20% for testing data. With a total dataset of 1059, this means that 80% of the total data, or 847 training data, will be used to train the model, while the remaining 20%, or 212 test data, will be used to test the model's performance. This division process is important for testing the reliability and consistency of the model in handling new data. A visualization of the distribution of training data and testing data can be seen in Figure 10.

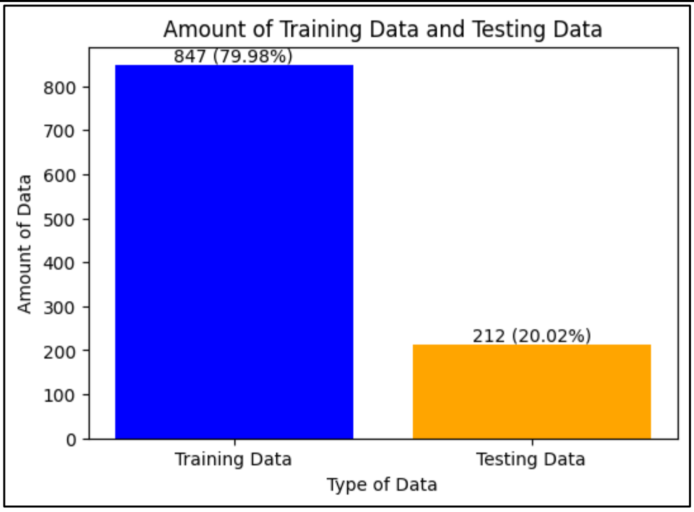

Figure 10. Visualization of Training Data and Testing Data

4.5. *Support Vector Machine (SVM)*

After the data division process, the next step is to apply the Support Vector Machine (SVM) algorithm to classify sentiment based on tweet data that has undergone preprocessing and labeling. The model was tested using test data consisting of 212 tweets and measured using several metrics, namely accuracy, precision, recall, and F1-score, to determine how well it performed in predicting tweet sentiment. The classification results of the model can be seen in Figure 11.

Classification Report for SVM:

	precision	recall	f1-score	support
Negatif	1.000	0.977	0.989	176.000
Netral	0.929	1.000	0.963	13.000
Positif	0.885	1.000	0.939	23.000
accuracy	0.981	0.981	0.981	0.981
macro avg	0.938	0.992	0.963	212.000
weighted avg	0.983	0.981	0.982	212.000

Figure 11. SVM Model Classification Report

Based on Figure 11, the SVM model's performance shows very strong results in all three sentiment categories, although there are differences in contribution between classes. In the negative category, the model recorded a precision value of 1.000, which means that all tweets predicted as negative actually came from the negative class without any prediction errors. The recall value of 0.977 indicates that almost all negative tweets can be accurately recognized by the model. The combination of these two values results in an f1-score of 0.989, with support of 176 tweets, making the negative class the most stable and dominant category in the classification process.

For the neutral category, the model's performance is also quite solid, even though the amount of data is not as much as the negative class. In this category, a precision value of 0.929 shows that most neutral predictions are on target. A recall value of 1.000 indicates that all neutral tweets were correctly identified by the model. Out of a total of 13 tweets, this category obtained an F1-score of 0.963, which shows that the model is able to capture neutral language patterns well despite limited data.

Meanwhile, the positive category also recorded excellent performance. With a precision of 0.885, most positive predictions matched their original labels. A recall value of 1.000 shows that all positive tweets in the test data were successfully recognized without any being missed. With a total of 23 tweets, this category produced an F1-score of 0.939, indicating that the model worked consistently in detecting positive sentiment.

Overall, the SVM model shows a very high accuracy of 0.981, which means that around 98.1% of sentiment predictions match the correct class. In the macro average metric, the precision is 0.938, the recall is 0.992, and the f1-score is 0.963, which describes the average performance between classes without taking into account the amount of data. On the other hand, the weighted average produces precision of 0.983, recall of 0.981, and f1-score of 0.982, indicating that the model's overall performance is greatly influenced by the negative class, which has the most data. These results show that the SVM model works very well in all three sentiment categories. With near-perfect accuracy, the model is considered highly effective in recognizing all three types of sentiment and is able to provide a clear picture of user opinions about the MyASN application.

4.6. Model Evaluation

This study also evaluated the model using a Confusion Matrix to see the extent to which the algorithm was able to group each tweet into the correct sentiment category. The Confusion Matrix helped clarify which parts of the model worked well and which parts were still often incorrect in making predictions. This image shows how the model recognizes negative, positive, and neutral data, while also revealing patterns of error that arise when the model misclassifies certain sentiments. The confusion matrix results can be seen in Figure 12.

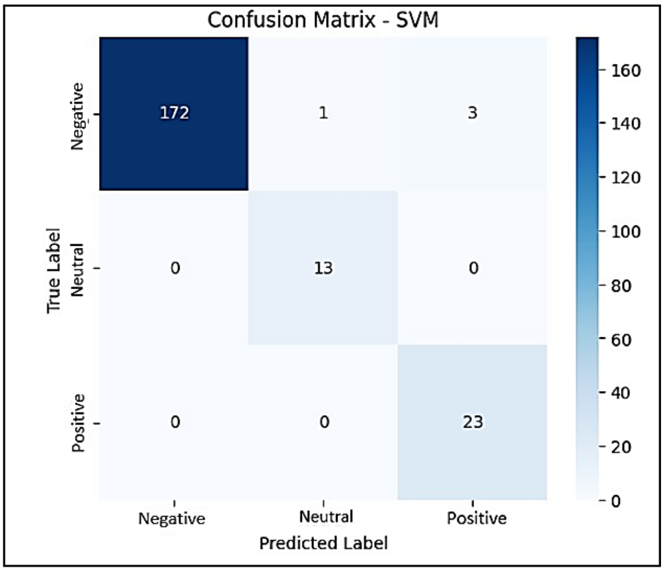

Figure 12. Confusion Matrix Results

Based on the confusion matrix visualization in Figure 12, the SVM model shows excellent classification capabilities across all sentiment categories. In the negative class, the model was able to accurately recognize most of the data, namely 172 tweets, and made only a few errors by misclassifying three tweets as positive and one tweet as neutral. This shows that the model is quite sensitive to patterns of language that are critical or complaint-oriented. For the neutral category, the model's performance is also very strong. All 13 neutral tweets were predicted correctly without any classification errors. Despite the small amount of data, the model is able to consistently understand the characteristics of neutral sentences, as reflected in the perfect recall value. In the positive category, the model again showed full accuracy. All 23 positive tweets were correctly identified and none were classified into other classes. This success indicates that the model is able to clearly distinguish appreciative or supportive language.

Overall, the pattern in the confusion matrix illustrates the SVM's very strong performance in distinguishing between the three types of sentiment. Errors only appear in a small portion of negative data that is similar to other classes. With the dominance of correct predictions in each category, these results indicate that the preprocessing, feature weighting, and model parameters have worked optimally in mapping the language variations in tweets related to the MyASN application.

5. CONCLUSION

This study analyzed public sentiment toward the MyASN application using X data and the Support Vector Machine (SVM) classification method. A total of 1,059 tweets were collected through a web scraping process using the Tweet-Harvest tool based on keywords and hashtags related to MyASN and BKN. The collected data were processed through several preprocessing stages, including cleaning, case folding, normalization, tokenizing, stopword removal, and stemming, to standardize the text before analysis. Sentiment labeling was performed using a lexicon-based approach, resulting in three sentiment classes: positive, negative, and neutral, and the data were divided into 80% training data and 20% testing data.

The results show that positive sentiment dominates user discussions about the MyASN application, followed by negative sentiment, while neutral sentiment appears in the smallest proportion. This indicates that although users still express technical issues and service complaints, positive perceptions of the application are more prominent. The SVM model demonstrated excellent performance, achieving an accuracy of 98.1%, with high precision, recall, and F1-score values across all sentiment categories. Negative sentiment achieved a precision of 1.000 and an F1-score of 0.989, while neutral and positive sentiments recorded recall values of 1.000, indicating that the model was able to identify sentiment classes accurately.

The confusion matrix confirms that the SVM model performed consistently with minimal misclassification, despite the imbalance in sentiment data. These findings indicate that the combination of preprocessing techniques, TF-IDF feature representation, and the SVM algorithm is highly effective in capturing sentiment patterns in Twitter data related to government digital services. Overall, this study demonstrates that SVM is a reliable method for sentiment analysis and can provide meaningful insights for evaluating and improving the MyASN application.

REFERENCES

Aditya, S. K., Alfiani, H., Sensuse, D. I., Lusa, S., Putro, P. A. W., & Indriasari, S. (2024). Investigating E-Government Adoption: The Intention to Use MyASN Application by Indonesian National Civil Service Agency (NCSA). *Journal of Government and Civil Society*, *8*(1), 43. https://doi.org/https://doi.org/10.31000/jgcs.v8i1.10890

Adrian, M. R. , Putra, M. P. , Rafialdy, M. H. , & Rakhmawati, N. A. (2021). Perbandingan Metode Klasifikasi Random Forest dan SVM Pada Analisis Sentimen PSBB. *JURNAL INFORMATIKA UPGRI*, *7*(1), 36–40. https://doi.org/https://doi.org/10.26877/jiu.v7i1.7099

Arfian, M. H., Maruhawa, H. S., Putri, A. A., Ratulugina, D., Khoirunnisa, E. R., Hermawan, V., & Syahr, A. J. (2025). Analisis Sentimen Pada Media Sosial Menggunakan Metode Support Vector Machine. *Jurnal Ilmu Teknik Dan Komputer (JITKOM)*, *09*(01), 1–6. https://doi.org/10.22441/jitkom.v9i1.001

Iqbal, M., Afdal, M., & Novita, R. (2024). Implementasi Algoritma Support Vector Machine Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online di Google Play Store. *Indonesian Journal of Machine Learning and Computer Science*, *4*(4), 1244–1252. https://doi.org/https://doi.org/10.57152/malcom.v4i4.1435

Lestari, S., & Berliani, S. (2024). Analisis Sentimen Masyarakat Terhadap Isu Pecat Sri Mulyani Pada Twitter Menggunakan Metode Naive Bayes Dan Support Vector Machine. *Jurnal Sains Dan Teknologi*, *5*(3), 951–960. https://doi.org/https://doi.org/10.55338/saintek.v5i3.2746

Ma'rufudin, & Yudhistira, A. (2025). Analisis Sentimen Petani Milenial Pada Media Sosial X Menggunakan Algortitma Support Vector Machine (SVM). *Jurnal Pendidikan Dan Teknologi Indonesia (JPTI)*, *5*(3), 845–857. https://doi.org/https://doi.org/10.52436/1.jpti.717

Mulyana, I. D., & Lutfianti, N. (2023). Analisis Sentimen Dengan Algoritma SVM Dalam Tanggapan Netizen Terhadap Berita Resesi 2023 Analysis Sentiment Using the SVM Algorithm in Netizen Responses to News of the 2023 Recession. *Jurnal Sistem Informasi Dan Teknologi Informatika*, *13*(1), 53–64. https://doi.org/https://dx.doi.org/10.30700/jst.v13i1.1339

Nugrahanti, F., Sari, E. R. N., Ahsan, M. S., & Pasyawati, P. R. (2025). Edukasi dan Pelatihan Pemrograman Dasar Menggunakan Google Colab bagi Guru SDN Klagen 4, Barat, Kec. Maospati Kab. Magetan. *Jurnal Pengabdian Masyarakat*, *3*(6), 2793–2799. https://doi.org/https://doi.org/10.59837/jpmba.v3i6.2824

Nurjanna, & Ivanna, J. (2025). Implementasi E-Government dalam Meningkatkan Pelayanan Publik di Dinas Kependudukan dan Catatan Sipil Kota Pematangsiantar. *Jurnal Pendidikan Kewarganegaraan*, *9*(1), 293–298. https://doi.org/https://doi.org/10.31571/jpkn.v9i1

Nurkholis, A., Styawati, S., Alim, S., Saputra, H., & Ferriyan, A. (2025). Sentiment Analysis of COVID-19 Booster Vaccines on Twitter Using Multi-Class Support Vector Machine. *Applied Information System and Management (AISM)*, *8*(1), 29–36. https://doi.org/https://doi.org/10.15408/aism.v8i1.42911

Ramlan, R., Satyahadewi, N., & Andani, W. (2023a). Analisis Sentimen Pengguna Twitter Menggunakan Support Vector Machine Pada Kasus Kenaikan Harga BBM. *Jambura Journal of Mathematics*, *5*(2), 431–445. https://doi.org/https://doi.org/10.34312/jjom.v5i2.20860

Ramlan, R., Satyahadewi, N., & Andani, W. (2023b). Analisis Sentimen Pengguna Twitter Menggunakan Support Vector Machine Pada Kasus Kenaikan Harga BBM. *JAMBURA JOURNAL OF MATHEMATICS*, *5*(2), 431–445. https://doi.org/https://doi.org/10.34312/jjom.v5i2.20860

Ratnaswari, S., Wibowo, N. C., & Kartika, D. S. Y. (2025). ANALISIS SENTIMEN MENGGUNAKAN METODE LEXICON-BASED DAN SUPPORT VECTOR MACHINE PADA PRESIDEN DAN WAKIL PRESIDEN INDONESIA PERIODE 2024–2029. *Jurnal Informatika Dan Teknik Elektro Terapan*, *13*(1), 362–368. https://doi.org/https://doi.org/10.23960/jitet.v13i1.5604

Rizqika, V., Handoko, P., & Rochmania, A. (2025). SELF DISCLOSURE GENERASI Z MELALUI MEDIA SOSIAL X (TWITTER). *Jurnal Ilmu Komunikasi UHO*, *10*(2), 397–411. https://doi.org/https://dx.doi.org/10.52423/jikuho.v10i2.348

Septiandika, V., & Nurrahmana, A. (2022). Analisis Implementasi Keunggulan Aplikasi My Sapk Bkn Terhadap Pemberkasan Administrasi Aparatur Sipil Negara (ASN) Pemerintah Kota Probolinggo.

Jurnal Ilmu Sosial Dan Pendidikan (JISIP), *6*(4), 2598–9944. https://doi.org/https://dx.doi.org/10.58258/jisip.v6i4.3492

Setiawan, A., & Suryono, R. R. (2024). Analisis Sentimen Ibu Kota Nusantara menggunakan Algoritma Support Vector Machine dan Naïve Bayes. *Jurnal Pendidikan Informatika*, *8*(1), 183–192. https://doi.org/https://doi.org/10.29408/edumatic.v8i1.25667

Siregar, A. M. (2023). Analisis Sentimen Pindah Ibu Kota Negara (IKN) Baru pada Twitter Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM). *Faktor Exacta*, *16*(3), 170–181. https://doi.org/https://dx.doi.org/10.30998/faktorexacta.v16i3.16703

Sudirman, A. N. A., Firmawati, F., & Sedana, G. S. (2026). Evaluasi Penerapan Aplikasi Sistem Informasi Kesehatan Jiwa (SIMKESWA) di Wilayah Puskesmas Kabupaten Gorontalo. *Malahayati Health Student Journal*, *6*(1), 465–476. https://doi.org/https://doi.org/10.33024/mahesa.v6i1.17199