DOI: 10.14421/ijid.2024.4504 IJID (International Journal on Informatics for Development), *e-ISSN: 2549-7448*
Vol. 13, No. 1, June 2024, Pp. 434-447

Sentiment Analysis of TIMNAS Indonesia's Participation in the Asian Cup U23 2024 on X Using Naive Bayes and SVM

Sewin Fathurrohman
Department of Information Systems
Universitas Gunadarma
Depok, Indonesia
sewin.fathur@gmail.com

Irfan Ricky Afandi
Department of Informatics Engineering
Universitas Muhammadiyah Prof. Dr. Hamka
Jakarta, Indonesia
richky99@gmail.com

Firman Noor Hasan*
Department of Informatics Engineering
Universitas Muhammadiyah Prof. Dr. Hamka
Jakarta, Indonesia
firman.noorhasan@uhamka.ac.id

Article History
Received May 27th, 2024
Revised August 10th, 2024
Accepted August 23rd, 2024
Published August, 2024

***Abstract*—This study aims to analyze the sentiment of the Indonesian public regarding the participation of the Indonesian National Team in the 2024 U-23 Asian Cup through the social media platform X. Sentiment analysis is crucial for understanding public perception and its impact on support for the national team. The research methodology involves collecting user comments on X related to the team's performance during the tournament, followed by data cleaning. The dataset is manually labeled, with 80% used as training data for algorithmic model training and the remaining 20% as test data, classified using Naive Bayes and Support Vector Machine algorithms. The analysis results indicate that the SVM algorithm achieves a higher % accuracy rate of 95% compared to Naive Bayes, which achieves 87%. The majority of the 3367 opinions analyzed express positive or satisfactory sentiments towards the national team's participation. However, there are fewer negative sentiments, highlighting areas requiring team management's attention. This study provides valuable insights into public perception of the Indonesian National Team. Furthermore, these findings can inform policymakers and team managers' decision-making to enhance the team's quality and performance in the future.**

Keywords*— *Indonesian national team; Naive Bayes; sentiment analysis; Support Vector Machine; X

1 INTRODUCTION

The dynamic evolution in the information technology sector has had a positive impact on the way society obtains, organizes, and distributes information efficiently and quickly [1]. Social media has become an effective and popular form of digital communication in Indonesia, enabling people to obtain and share information easily [2]. The use of social media as a platform for sharing views on various topics makes it a valuable source of data [3]. Effective data processing and documentation are very important to produce information that supports decision-making [4].

Football is one of the most popular topics in Indonesia [5], with high public enthusiasm for the sport [6], especially in the participation of the Indonesian National Team (TIMNAS) in various competitions [7]. International-level football competitions attract the attention of the Indonesian people [8], such as the 2024 AFC U-23 Asian Cup, which sparked many opinions about TIMNAS' performance during the match on social media [9]. Understanding public sentiment towards TIMNAS performance is very important because it can inform decision-making and provide valuable insights that help improve the team's image, and manage relationships with fans more effectively.

Social media platform X has emerged as one of the most popular platforms in Indonesia [10], enabling users to share opinions and information via computers and smartphones [11]. The advantages of X include real-time updates [12], hashtags and trending topics features that facilitate tracking and categorizing specific discussions [13], such as those about the participation of the Indonesian national team in the AFC U-23 Asian Cup 2024. The advantages of it are that it has many data and can be processed as material for decision-making using sentiment analysis methods [14].

Sentiment analysis is a method used to determine trends in public interest and opinion on specific issues [15]. This process utilizes artificial intelligence, such as machine learning [16], to efficiently and accurately process opinion data from social media [17] by categorizing these opinions into sentiment labels such as positive, negative, and neutral [18], which can provide valuable insights for decision-making [19]. In sentiment analysis, the use of Naive Bayes and Support Vector Machine (SVM) algorithms is known to be effective due to their respective strengths in handling unstructured and complex data, such as public opinion data on social media [20]. Naive Bayes is known for its simplicity and implementation speed, as well as its effectiveness in managing small datasets [21]. However, the independence assumption among features in Naive Bayes can reduce accuracy and be less effective for data with complex feature relationships [22]. Conversely, the SVM algorithm is recognized for its ability to handle high-dimensional data and complex distributions [23]. However, this algorithm tends to be slower in training on large datasets and requires careful parameter selection as it can significantly impact accuracy [24].

Previous studies have shown the success of using Naive Bayes and SVM algorithms in classifying public opinion sentiment related to the field of football on social media platform X. A study [25] using the Naive Bayes algorithm achieved an accuracy of 96% with 1702 positive tweets and 1098 negative tweets, out of 2800 data. Another study [26] collected 946 tweets with 150 positive sentiment tweets and 796 negative sentiment tweets, showing an accuracy rate of 84%. The study [27] reported an accuracy of 78% with 642 tweet data, resulting in 287 positive sentiments, 37 negative sentiments, and 318 neutral sentiments. On the other hand, studies using SVM, such as study [28], showed an accuracy of 91% on 509 tweet data, with 445 negative sentiments, 50 positive sentiments, and 4 neutral sentiments. Another study [29] demonstrated an accuracy of 95.16% with 1507 tweet data, resulting in 676 negative labels, 87 neutral labels, and 62 positive labels. A study [30] processed 13705 text data and achieved an accuracy score of 76.40%, with 6907 neutral sentiments, 4901 negative sentiments, and 1897 positive sentiments.

This study aims to provide deeper insights into public opinion regarding the participation of the Indonesian national team in the U-23 Asian Cup 2024 through sentiment analysis from social media platform X. These opinions will be categorized into positive, neutral, and negative sentiments. These insights can assist team managers and policymakers in designing strategies to improve team quality and performance and enhance relationships with fans. Additionally, this study will compare the performance in terms of accuracy, recall, precision, and F1-score between the Naive Bayes and SVM algorithms to determine which algorithm is better for analyzing public sentiment towards the participation of the Indonesian national team in the U-23 Asian Cup 2024. The contribution of this study lies in the development of knowledge by demonstrating how sentiment analysis techniques like Naive Bayes and SVM can be applied in social and sports contexts, providing a solid foundation for future research that seeks to explore the use of sentiment analysis in various other fields.

2 METHOD

The researchers conducted a series of steps in the data processing procedure, as illustrated in the diagram provided in Fig. 1 which outlines the process that must be followed to achieve the desired results in the data analysis. Based on the research flow shown in Fig. 1, the initial phase of this study involves collecting a dataset containing public opinions regarding the Indonesian national team's participation in the U-23 Asian Cup 2024 on social media platform X. Data collection is conducted using several parameters to ensure that the data obtained is more focused and relevant to the research topic. The collected dataset will be divided into training and testing data with an 80:20 ratio.

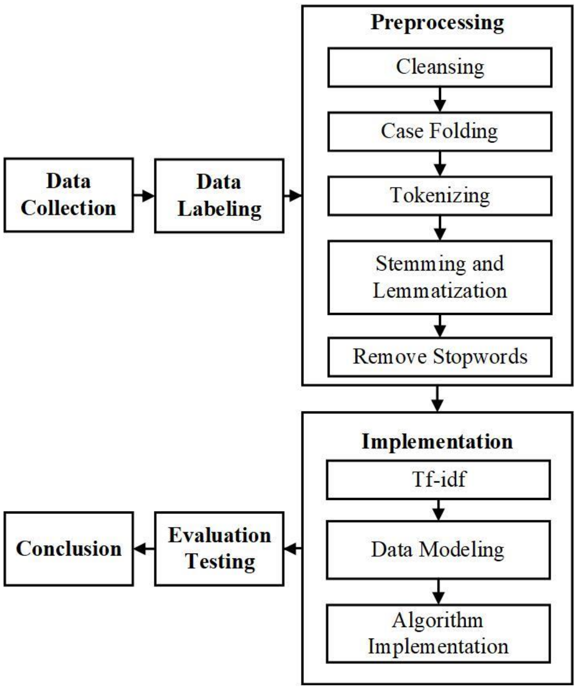

Figure 1. Stages of research

The researchers will manually label 80% of the dataset for training data, while the remaining 20%, which will not have labels, will be used as testing data. The dataset will then undergo preprocessing to transform unstructured text into a structured format. The preprocessing stages include cleansing, case folding, tokenizing, stemming and lemmatization, and removing stopwords. The cleaned dataset will then be weighted word using Term Frequency-Inverse Document Frequency (TF-IDF). A training data model will be created based on the labeled data and applied to the Naive Bayes and SVM algorithms to classify the unlabeled testing data. This study chooses the Naive Bayes and SVM algorithms because they are known for their simplicity and effectiveness in processing text data and require lower computational power. Previous studies have shown that these classical methods effectively analyse sentiment on social media platform X for Indonesian-language datasets. Although deep learning methods and pre-trained models are becoming increasingly popular, they require very large data sets and high computational resources [31], which is a consideration in choosing these algorithms due to limited computational resources. The next stage is the evaluation of testing using a confusion matrix to obtain the accuracy, recall, precision, and F1-score values of the Naive Bayes and SVM algorithms. The evaluation results will be analyzed to draw conclusions that can be used as a basis for decision-making.

2.1 *Data Collection*

In this step, the researchers accessed a dataset from the social media platform X regarding public opinions on the participation of the Indonesian National Team in the Asian Cup U-23 2024 competition using the Python programming language with the X_auth_token. The keyword phrase used was "timnas Indonesia u23". Data was collected from April 15 to May 3, 2024, as shown in Fig. 2 with texts in Indonesian, resulting in 3746 tweets then the dataset was saved in CSV file format.

	conversation_id_str	created_at	favorite_count	full_text
0	1786183300731818120	Thu May 02 23:57:39 +0000 2024	0	hasil skor 2-1 yang menandakan irak menempati ...
1	1786179851734647087	Thu May 02 23:43:56 +0000 2024	0	Timnas Indonesia U23 Ditekuk Irak 1-2 Tiket Ol...
2	1786176654093746291	Thu May 02 23:31:14 +0000 2024	1	Timnas U23 Indonesia kalah dari Irak pada laga...
3	1786176355132182957	Thu May 02 23:30:03 +0000 2024	4	Jadwal Timnas Indonesia U23 vs Guinea U23! htt...
4	1786176342608007593	Thu May 02 23:30:00 +0000 2024	0	Link siaran ulang Indonesia vs Irak dalam laga...
...	...	...	...	...
2704	1783512955788910635	Thu Apr 25 15:06:39 +0000 2024	1	Beberapat Menit Lagi Ini Link Live Streaming K...
2705	1783512416082599962	Thu Apr 25 15:04:30 +0000 2024	0	Timnas Indonesia vs Korea Selatan di Perempat ...
2706	1783512307932463433	Thu Apr 25 15:04:04	2	SC Heerenveen Ceritakan Detik-detik

Figure 2. Results of data retrieval

The next step is to remove unnecessary columns from the table and retain only the columns containing public opinions. Table 1 presents examples of tweets to provide a clearer picture of the variation in expressions and sentiment analysis used in this research.

Table 1. Several Opinions from X

Tweet
Semangat Timnas Indonesia U23 #TimNasDay
jelek sekali permainan Timnas Indonesia U23 malam ini. Sangat kecewa 👎.
Terima kasih untuk Timnas Indonesia U23 perjuangan kalian sampai sejauh ini patut diapresiasi 🔥 Tetap semangat untuk berjuang di laga selanjutnya ya 🙏 @KompasBola #PlayoffOlimpiade #OlimpiadeParis2024 #TimnasIndonesia #KitaGaruda #TimnasDay #Guinea
☑ RESMI : Nathan Tjoe telah mendapatkan izin dari Heerenveen untuk

Tweet
kembali memperkuat timnas Indonesia U23 di ajang Piala Asia U23. #timnasindonesiau23 #shintaeyong #pelatih https://www.viva.co.id/bola/liga-indonesia/1708088-alasan-heerenveen-izinkan-nathan-tjoe-a-on-kembali-ke-timnas-indonesia-u-23
Timnas Indonesia U23 buat sejarah dengan lolos ke babak 4 besar Piala Asia 2024 #timnasday #garudamuda

2.2 Data Labeling

The dataset will be manually labeled by three independent researchers who will categorize the tweets into three sentiments: positive, negative, and neutral. Each tweet will be reviewed individually to ensure consistency and accuracy of labeling. Disagreements will be resolved through discussion and consensus among the researchers to ensure data reliability. Out of a total of 3367 tweets, 80% (2693 tweets) will be manually labeled as training data, while the remaining 20% (674 tweets) will be used as unlabeled testing data. The testing data will be labeled using the Naive Bayes and SVM algorithms to assess the performance of the built model. This manual labeling process aims to ensure high-quality and representative training data and to minimize bias through researcher discussion and consensus. Table 2 presents examples of tweets along with sentiment classifications to provide a more detailed overview of the variations in expression and sentiment analysis methods that will be used in this study.

2.3 Preprocessing Data

The next step is preprocessing the CSV dataset. Before entering this stage, the dataset with similar or duplicate opinion values must be removed first, resulting in a total dataset of 3367 entries for the subsequent processes. Data preprocessing is the process of transforming unstructured text into a more structured form, which will aid in data processing [32]. Several steps in data preprocessing include cleansing, case folding, tokenizing, stemming and lemmatization, and removing stopwords.

Table 2. Several Opinions from X After Labeling

Tweet	Label
Semangat Timnas Indonesia U23 #TimNasDay	Positive
jelek sekali permainan Timnas Indonesia U23 malam ini. Sangat kecewa 👎.	Negative
Terima kasih untuk Timnas Indonesia U23 perjuangan kalian sampai sejauh ini patut diapresiasi Tetap semangat untuk berjuang di laga selanjutnya ya @KompasBola #PlayoffOlimpiade #OlimpiadeParis2024 #TimnasIndonesia #KitaGaruda #TimnasDay #Guinea	Positive
☑ RESMI : Nathan Tjoe telah mendapatkan izin dari Heerenveen untuk kembali memperkuat timnas Indonesia U23 di ajang Piala Asia U23. #timnasindonesiau23 #shintaeyong #pelatih https://www.viva.co.id/bola/liga-indonesia/1708088-alasan-heerenveen-izinkan-nathan-tjoe-a-on-kembali-ke-timnas-indonesia-u-23	Neutral
Timnas Indonesia U23 buat sejarah dengan lolos ke babak 4 besar Piala Asia 2024 #timnasday #garudamuda	Neutral

2.3.1 *Cleansing:* The cleansing process is an essential step in data preprocessing. Cleansing aims to remove all URL links, symbols, mentions, retweets, hashtags, and emoticons from the dataset [33]. This step is necessary to ensure that the data used for further analysis or processing is clean and free from irrelevant or distracting characters [34]. By performing cleansing, the dataset becomes cleaner and ready for the next preprocessing steps, where the text results have already passed through the cleansing stage, as shown in Table 3.

2.3.2 *Case Folding:* The next step is case folding, which is the process of converting all uppercase letters in the dataset to lowercase letters to standardize the text because uppercase letters can affect subsequent analysis processes [35]. The text results that have undergone the case folding stage are shown in Table 4.

Table 3. The Outcome of Data Cleansing

Before	After	Label
Semangat Timnas Indonesia U23 #TimNasDay	Semangat Timnas Indonesia	Positive
jelek sekali permainan Timnas Indonesia U23 malam ini. Sangat kecewa 👎.	jelek sekali permainan Timnas Indonesia malam ini Sangat kecewa	Negative
Terima kasih untuk Timnas Indonesia U23 perjuangan kalian sampai sejauh ini patut diapresiasi Tetap semangat untuk berjuang di laga selanjutnya ya @KompasBola #PlayoffOlimpiade #OlimpiadeParis2024 #TimnasIndonesia #KitaGaruda #TimnasDay #Guinea	Terima kasih untuk Timnas Indonesia perjuangan kalian sampai sejauh ini patut diapresiasi Tetap semangat untuk berjuang di laga selanjutnya ya	Positive
☑ RESMI : Nathan Tjoe telah mendapatkan izin dari Heerenveen untuk kembali memperkuat timnas Indonesia U23 di ajang Piala Asia U23. #timnasindonesiau23 #shintaeyong #pelatih https://www.viva.co.id/bola/liga-indonesia/1708088-alasan-heerenveen-izinkan-nathan-tjoe-a-on-kembali-ke-timnas-indonesia-u-23	RESMI Nathan Tjoe telah mendapatkan izin dari Heerenveen untuk kembali memperkuat timnas Indonesia di ajang Piala Asia	Neutral
Timnas Indonesia U23 buat sejarah dengan lolos ke babak 4 besar Piala Asia 2024 #timnasday #garudamuda	Timnas Indonesia buat sejarah dengan lolos ke babak besar Piala Asia	Neutral

Table 4. The Outcome of Case Folding

Before	After	Label
Semangat Timnas Indonesia	semangat timnas indonesia	Positive
jelek sekali permainan	jelek sekali permainan	Negative

Before	After	Label
Timnas Indonesia malam ini Sangat kecewa	timnas indonesia malam ini sangat kecewa	
Terima kasih untuk Timnas Indonesia perjuangan kalian sampai sejauh ini patut diapresiasi Tetap semangat untuk berjuang di laga selanjutnya ya	terima kasih untuk timnas indonesia perjuangan kalian sampai sejauh ini patut diapresiasi tetap semangat untuk berjuang di laga selanjutnya ya	Positive
RESMI Nathan Tjoe telah mendapatkan izin dari Heerenveen untuk kembali memperkuat timnas Indonesia di ajang Piala Asia	resmi nathan tjoe telah mendapatkan izin dari heerenveen untuk kembali memperkuat timnas indonesia di ajang piala asia	Neutral
Timnas Indonesia buat sejarah dengan lolos ke babak besar Piala Asia	timnas indonesia buat sejarah dengan lolos ke babak besar piala asia	Neutral

2.3.3 *Tokenizing:* The next stage in data preprocessing is the tokenization process, which involves splitting a sentence into several words or tokens [36] allowing for more effective processing of individual words and facilitating the subsequent preprocessing stages such as stemming and lemmatization, remove Stopwords. The text results that have undergone the tokenization stage are shown in Table 5.

2.3.4 *Stemming dan Lemmatization:* The following process is removing all affixes from words to obtain their base form in the dataset [37]. In this process, researchers use library sastrawi which can help simplify Indonesian text by converting inflected words to their basic form. It is beneficial in sentiment analysis to treat similar words consistently [38]. Sastrawi library can process related words, such as "memenangkan" and "dimenangkan" into "menang" another example is "memperkuat" and "diperkuat" into "kuat" ensuring a more consistent analysis. The text results that have undergone the stemming and lemmatization stage are shown in Table 6.

Table 5. The Outcome of Tokenizing

Before	After	Label
semangat timnas indonesia	semangat, timnas, indonesia	Positive
jelek sekali permainan timnas indonesia malam ini sangat kecewa	jelek, sekali, permainan, timnas, indonesia, malam, ini, sangat, kecewa	Negative
terima kasih untuk timnas indonesia perjuangan kalian sampai sejauh ini patut diapresiasi tetap semangat untuk berjuang di laga selanjutnya ya	terima, kasih, untuk, timnas, indonesia, perjuangan, kalian, sampai, sejauh, ini, patut, diapresiasi, tetap, semangat, untuk, berjuang, di, laga, selanjutnya, ya	Positive
resmi nathan tjoe telah mendapatkan izin dari heerenveen untuk kembali memperkuat timnas indonesia di ajang piala asia	resmi, nathan, tjoe, telah, mendapatkan, izin, dari, heerenveen, untuk, kembali, memperkuat, timnas, indonesia, di, ajang, piala, asia	Neutral
timnas indonesia buat sejarah dengan lolos ke babak besar piala asia	timnas, indonesia, buat, sejarah, dengan, lolos, ke, babak, besar, piala, asia	Neutral

Table 6. The Outcome of Stemming and Lemmatization

Before	After	Label
semangat, timnas, indonesia	semangat, timnas, indonesia	Positive
jelek, sekali, permainan, timnas, indonesia, malam, ini, sangat, kecewa	jelek, sekali, main, timnas, indonesia, malam, ini, sangat, kecewa	Negative
terima, kasih, untuk, timnas, indonesia, perjuangan, kalian, sampai, sejauh, ini, patut, diapresiasi, tetap, semangat, untuk, berjuang, di, laga, selanjutnya, ya	terima, kasih, untuk, timnas, indonesia, juang, kalian, sampai, jauh, ini, patut, apresiasi, tetap, semangat, untuk, juang, di, laga, lanjut, ya	Positive
resmi, nathan, tjoe, telah, mendapatkan, izin, dari, heerenveen, untuk, kembali, memperkuat, timnas, indonesia, di, ajang, piala, asia	resmi, nathan, tjoe, telah, dapat, izin, dari, heerenveen, untuk, kembali, kuat, timnas, indonesia, di, ajang, piala, asia	Neutral
timnas, indonesia, buat, sejarah, dengan, lolos, ke, babak, besar, piala, asia	timnas, indonesia, buat, sejarah, dengan, lolos, ke, babak, besar, piala, asia	Neutral

2.3.5 *Remove Stopwords:* The next stage is the remove stopwords process, which involves removing all conjunctions and unnecessary words from the dataset [39]. This process utilizes the NLTK library. After removing the stopwords, the remaining words are recombined into a single sentence. Removing unnecessary words makes the dataset more focused and clean, enabling more effective and efficient analysis. Stopword removal helps reduce noise in the data and ensures that only meaningful words are analyzed [40]. The text results that have undergone the remove stopwords stage are shown in Table 7.

Table 7. The Outcome of Remove Stopwords

Before	After	Label
semangat, timnas, indonesia	semangat timnas indonesia	Positive
jelek, sekali, main, timnas, indonesia, malam, ini, sangat, kecewa	jelek main timnas Indonesia malam kecewa	Negative
terima, kasih, untuk, timnas, indonesia, juang, kalian, sampai, jauh, ini, patut, apresiasi, tetap, semangat, untuk, juang, di, laga, lanjut, ya	terima kasih timnas indonesia juang patut apresiasi semangat juang laga lanjut	Positive
resmi, nathan, tjoe, telah, dapat, izin, dari, heerenveen, untuk, kembali, kuat, timnas, indonesia, di, ajang, piala, asia	resmi nathan tjoe dapat izin heerenveen kembali kuat timnas indonesia ajang piala asia	Neutral
timnas, indonesia, buat, sejarah, dengan, lolos, ke, babak, besar, piala, asia	timnas indonesia buat sejarah lolos babak besar piala asia	Neutral

3 RESULT AND DISCUSSION

In this study, several experiments were conducted to analyze public sentiment toward the participation of the Indonesian National Team (TIMNAS) in the U-23 Asian Cup 2024 through social media platform X. These experiments were carried out in several stages: collecting 3367 tweets containing opinions about TIMNAS; labeling the data where 80% of the data were manually labeled as training data while 20% was used as testing data; cleaning the dataset; weighting the words using the TF-IDF method; modeling the data and applying the Naive Bayes and SVM algorithms for sentiment classification on the testing data; evaluating the Naive Bayes and SVM algorithms using a confusion matrix to obtain accuracy, precision, recall, and F1-score values. The experiments are considered successful if the algorithms can classify sentiments with a high accuracy above 50% and the manual calculation results match the code program. A failure scenario occurs if the algorithms cannot properly distinguish between positive and negative sentiments, indicated by low precision, recall, or F1-score values below 50%, and if the manual calculation results with the code program do not match.

3.1 TF-IDF Weighting

TF-IDF weighting aims to assign values to data based on the frequency of word occurrence in a document. The frequency of each word in the document is calculated and then given a weight based on the frequency of occurrence of that word relative to the total number of documents containing that word [41]. Equation 1 is the TF-IDF formula used to calculate the total weight of a word in a document [42].

$$TF - IDF_{(t,d,D)} = TF_{(t,d)} \times IDF_{(t,D)} \qquad (1)$$

Researchers create a visualization of word frequency in the form of a bar chart as depicted in Fig. 3 which highlights the top 25 most frequently occurring words within the text and enables researchers to identify and assess the prevalence of these terms and their potential impact on sentiment determination.

The bar chart visualization presented in Fig. 3 provides a clear representation of the most frequently used words in the corpus of analyzed text. Where the dominant term is "timnas," followed by "indonesia," and then "piala." The researchers also conducted word cloud as shown in Fig. 4 to visualize text data where larger words indicate higher frequency and can identify key terms from the entire set of words in the dataset.

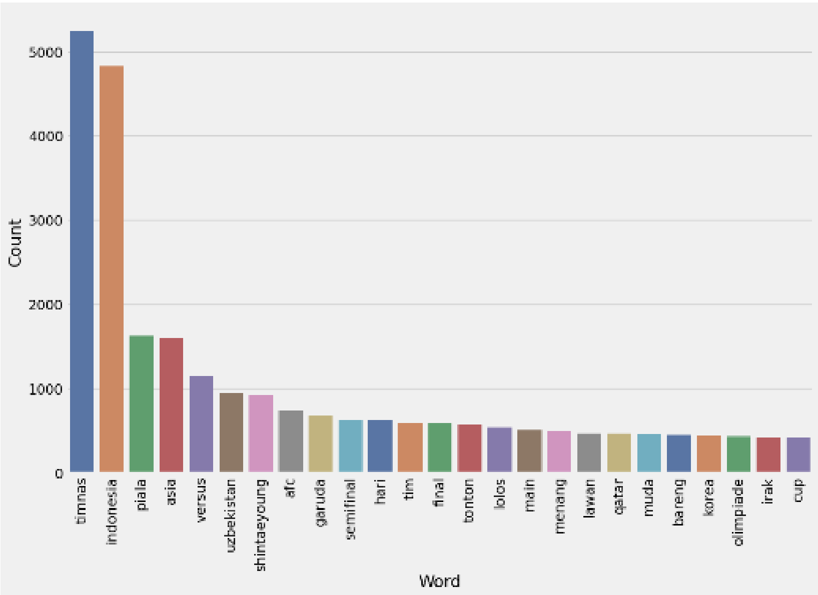

Figure 3. Visualization of word frequency in the bar chart

Figure 4. Visualization of word frequency in word cloud

3.2 *Data Modeling*

After the data weighting process using the TF-IDF method, the next step is data modeling. This process is an essential step in preparation for further analysis. The dataset, consisting of 3367 tweets, will be divided into two parts with an 80:20 ratio, where 80% will be used for training and 20% for testing. During the training phase, the Naive Bayes and SVM algorithms will be trained using 80% of the labeled data, and in the testing phase, the remaining 20% of the data will be used to evaluate the performance of the two algorithms. This division is performed to ensure that the model can be effectively trained and evaluated on data that it has not previously encountered.

3.3 *Algorithm Implementation*

After the data modeling process, the next step is the implementation of algorithms for sentiment analysis. In this study, researchers have implemented two distinct algorithms: Naive Bayes and SVM. Each algorithm offers unique and effective methods for processing data. The Naive Bayes algorithm is rooted in Bayes' theorem and provides a probability-based approach to classification [43]. On the other hand, SVM is a robust technique that operates by seeking the optimal hyperplane to separate data classes [44]. The Naive Bayes and SVM algorithms, which have been trained to learn patterns from each sentiment category, will then be applied to the 20% test data to determine the sentiment values. Subsequently, the sentiment classification results from the Naive Bayes and SVM algorithms will be combined with the previously labeled 80% data to observe the overall sentiment of the dataset. The results of the Naive Bayes and SVM algorithm implementations are shown in Fig. 5, where the classification results of these algorithms have been combined with the previously labeled data.

This visualization helps to see the sentiment distribution within the dataset graphically, making it easier to analyze and understand the overall sentiment classification results. By looking at the diagram in Fig. 5 it's easier to identify the majority sentiment as positive, negative, or neutral and gain additional insights into any patterns in the data. The difference in the number of tweets classified by the Naive Bayes and SVM algorithms is due to how they handle data and define sentiment class boundaries. Both algorithms yielded very similar results across all sentiment categories, indicating that both Naive Bayes and SVM can be effectively utilized for sentiment analysis.

The implementation results of the algorithms show that positive sentiment dominates, followed by neutral and negative sentiments. This finding indicates that most of the public opinion on social media platform X views the performance of the Indonesian National Team in the U-23 Asian Cup 2024 positively. The result of the dominant positive sentiment can be interpreted as a sign that the team's strategies and performance have met the expectations of many fans. Insights gained from this sentiment analysis can provide significant benefits to team managers and policymakers. Understanding that most of the public holds a positive view can help develop communication and promotion strategies that leverage this support to strengthen relationships with fans. Additionally, criticism found in negative sentiments can be used as constructive feedback to improve the team's performance in the future.

3.4 *Evaluation Testing*

The evaluation testing is a critical step in the process of developing and accessing models in machine learning. The researchers presented the performance analysis of the classification model using the confusion matrix as an evaluation tool. The confusion matrix maps the number of correct and incorrect predictions made by the model by comparing the predictions with the actual values [45]. This tool provides a visual representation of the algorithm's performance in classifying data into sentiment categories and is used to calculate the accuracy, recall, precision, and F1-score for each algorithm [46]. In this study, the researchers applied a confusion matrix to the 3367 data points that were split into training and testing sets as shown in Fig. 6 this analysis was conducted for the application of the Naive Bayes algorithm and in Fig. 7 for the application of the Support Vector Machine algorithm.

The confusion matrix shown in Fig. 6 illustrates the performance of the Naive Bayes algorithm in classifying into three sentiment classes there negative, neutral and positive with the following values:

1. Negative category: The model predicted 213 cases as True Negatives, but there were 54 Negative cases predicted as Neutral, and 77 Negative cases predicted as Positive.
2. Neutral category: The model correctly predicted 913 cases as True Neutrals, but there were 13 Neutral cases predicted as Negative and 124 Neutral cases predicted as Positive.
3. Positive category: The model correctly predicted 1227 cases as True Positives, but there were 21 Positive cases predicted as Negative and 51 Positive cases predicted as Neutral.

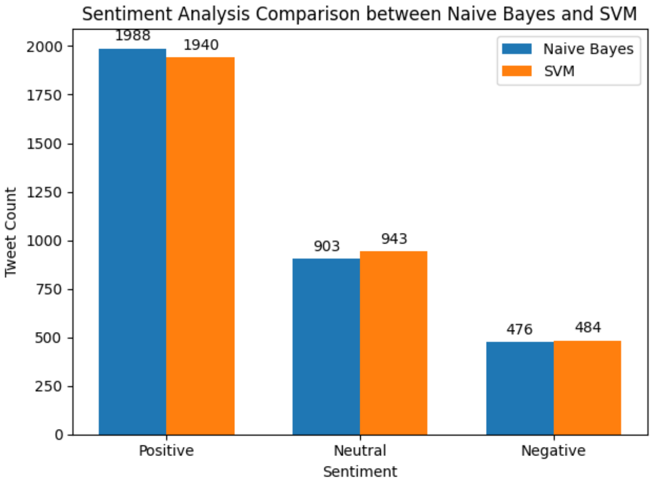

Figure 5. Comparison of classification labels from the Naive Bayes and SVM algorithms

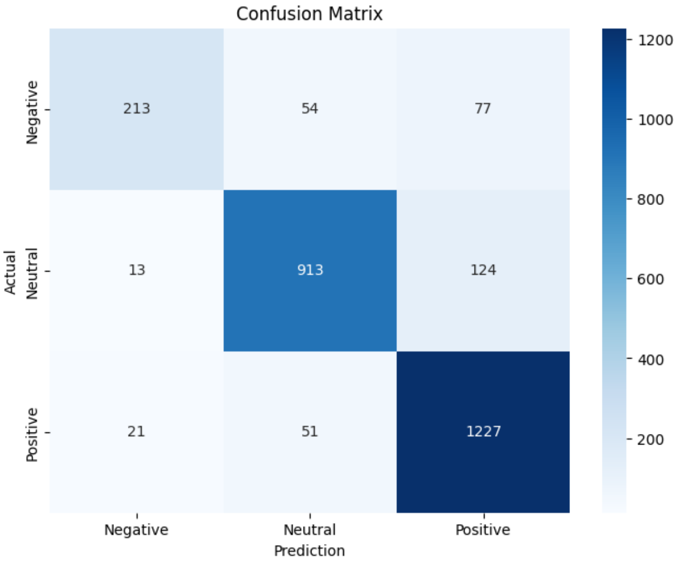

Figure 6. Confusion matrix of Naïve Bayes implementation result

The confusion matrix shown in Fig. 7 illustrates the performance of the SVM algorithm in classifying into three sentiment classes there negative, neutral and positive with the following values:

1. Negative category: The model predicted 274 cases as True Negatives, but there were 51 Negative cases predicted as Neutral, and 19 Negative cases predicted as Positive.
2. Neutral category: The model correctly predicted 1039 cases as True Neutrals, but there were 1 Neutral case predicted as Negative and 10 Neutral cases predicted as Positive.
3. Positive category: The model correctly predicted 1239 cases as True Positives, but there were 7 Positive cases predicted as Negative and 53 Positive cases predicted as Neutral.

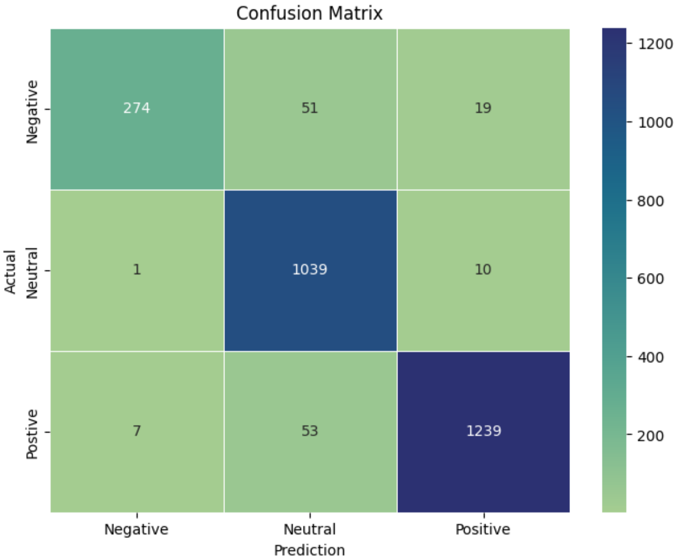

Figure 7. Confusion matrix of SVM implementation result

At this stage, the confusion matrix value of each algorithm will be evaluated such as algorithm accuracy value and F1-score, recall, and precision for each sentiment category which are employed to assess the performance of the implemented model [47]. From the values obtained in the confusion matrix, Equations 2-5 can be utilized to assess the performance of the model in terms of algorithm accuracy as well as Precision, Recall, and F1-Score for each sentiment class [48]. This process aims to gauge the extent to which the constructed model can accurately classify sentiments.

1. Accuracy: It measures how often the model makes correct predictions. It divides the total correct predictions with the total number of cases.

$$Accuracy = \frac{(TP_{positive} + TP_{negative} + TP_{neutral})}{(Total\ Cases)} \quad (2)$$

2. Precision: It measures how accurate the model is when making positive or correct predictions. It is computed by dividing the number of TP by the total number of positive predictions (TP and FP).

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

3. Recall: It measures how well the model identifies all true positive cases. It is computed by dividing the number of TP by the actual number of positive classes (TP and FN).

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

4. F1-Score: It is the harmonic mean of Precision and Recall, providing a single score that balances both metrics. It is particularly useful to know a balance between Precision and Recall.

$$F1\ score = 2\ x \frac{(Precision\ x\ Recall)}{(Precision + Recall)} \quad (5)$$

Following the extraction of data from the confusion matrix presented in Fig. 6 acquired from the implementation outcomes of the algorithm, the researchers embarked on determining the accuracy of the Naive Bayes algorithm. They conducted computations to derive the F1-score, recall, and precision values for each sentiment class from the Naive Bayes algorithm. The researchers performed manual calculations to ensure the accuracy of the values produced by the code. To facilitate this process, the results comparing the performance evaluation of the Naive Bayes and SVM models using the confusion matrix with a total case of 2693 data are presented in Table 8.

Table 8. Comparison of Confusion Matrix for Naive Bayes and SVM

	Naive Bayes	**SVM**
TP positive	1227	1239
TP negative	213	274
TP neutral	913	1039
FP positive	201	29
FP negative	34	8
FP neutral	105	104
FN positive	72	60
FN negative	131	70
FN neutral	137	11

Based on the values in Table 8 above, the following are the outcomes of the calculations and evaluation tests conducted on the Naive Bayes and SVM algorithms.

Naive Bayes:

$$Accuracy = \frac{(TP_{positive} + TP_{negative} + TP_{neutral})}{(Total\ Cases)}$$

$$Accuracy = \frac{1227+213+913}{2693}$$

$$Accuracy = 0.87374\ or\ 87.37\ \%$$

Sentiment Positive Naive Bayes:

$$Precision_{positive} = \frac{(TP_{positive})}{(TP_{positive}+FP_{positive})}$$

$$Precision_{positive} = \frac{1227}{1227+201}$$

$$Precision_{positive} = 0.85924\ or\ 85.92\%$$

$$Recall_{positive} = \frac{(TP_{positive})}{(TP_{positive}+FN_{positive})}$$

$$Recall_{positive} = \frac{1227}{1227+72}$$

$$Recall_{positive} = 0.94457\ or\ 94.46\%$$

$$F1\ score_{positive} = 2\ x \frac{(Precision_{positive}\ x\ Recall_{positive})}{(Precision_{positive} + Recall_{positive})}$$

$$F1\ score_{positive} = 2\ x \frac{(0{,}85924\ x\ 0.94457)}{(0{,}85924 + 0.94457)}$$

$$F1\ score_{positive} = 0.89989\ or\ 89.99\%$$

Sentiment Negative Naive Bayes:

$$Precision_{negative} = \frac{(TP_{negative})}{(TP_{negative}+FP_{negative})}$$

$$Precision_{negative} = \frac{213}{213+34}$$

$$Precision_{negative} = 0.86234\ or\ 86.23\%$$

$$Recall_{negative} = \frac{(TP_{negative})}{(TP_{negative}+FN_{negative})}$$

$$Recall_{negative} = \frac{213}{213+131}$$

$$Recall_{negative} = 0.61918\ or\ 61.92\%$$

$$F1\ score_{negative} = 2\ x \frac{(Precision_{negative}\ x\ Recall_{negative})}{(Precision_{negative} + Recall_{negative})}$$

$$F1\ score_{negative} = 2\ x \frac{(0.86234\ x\ 0{,}61918)}{(0.86234 + 0{,}61918)}$$

$$F1\ score_{negative} = 0.72081\ or\ 72.08\%$$

Sentiment Neutral Naive Bayes:

$$Precision_{neutral} = \frac{(TP_{neutral})}{(TP_{neutral}+FP_{neutral})}$$

$$Precision_{neutral} = \frac{913}{913+105}$$

$$Precision_{neutral} = 0.89685\ or\ 89.69\%$$

IJID (International Journal on Informatics for Development), *e-ISSN: 2549-7448*
Vol. 13, No. 1, June 2024, Pp. 434-447

$$Recall_{neutral} = \frac{(TP_{neutral})}{(TP_{neutral}+FN_{neutral})}$$

$$Recall_{neutral} = \frac{913}{913+137}$$

$$Recall_{neutral} = 0.86952 \ or \ 86.95\%$$

$$F1\ score_{neutral} = 2\ x \frac{(Precision_{neutral}\ x\ Recall_{neutral})}{(Precision_{neutral} + Recall_{neutral})}$$

$$F1\ score_{neutral} = 2\ x \frac{(0.89685\ x\ 0.86952)}{(0.89685 + 0.86952)}$$

$$F1\ score_{neutral} = 0.88297 \ or \ 88.30\%$$

Accuracy SVM:

$$Accuracy = \frac{(TP_{positive}+ TP_{negative}+TP_{neutral})}{(Total\ Cases)}$$

$$Accuracy = \frac{25+ 250+267}{674}$$

$$Accuracy = 0.80415 \ or \ 80.42\ \%$$

Sentiment Positive SVM:

$$Precision_{positive} = \frac{(TP_{positive})}{(TP_{positive}+FP_{positive})}$$

$$Precision_{positive} = \frac{1239}{1239+29}$$

$$Precision_{positive} = 0.97712 \ or \ 97.71\%$$

$$Recall_{positive} = \frac{(TP_{positive})}{(TP_{positive}+FN_{positive})}$$

$$Recall_{positive} = \frac{1239}{1239+60}$$

$$Recall_{positive} = 0.95281 \ or \ 95.28\%$$

$$F1\ score_{positive} = 2\ x \frac{(Precision_{positive}\ x\ Recall_{positive})}{(Precision_{positive} + Recall_{positive})}$$

$$F1\ score_{positive} = 2\ x \frac{(0.97712\ x\ 0.95281)}{(0.97712 + 0.95281)}$$

$$F1\ score_{positive} = 0.93199 \ or \ 93.20\%$$

Sentiment Negative SVM:

$$Precision_{negative} = \frac{(TP_{negative})}{(TP_{negative}+FP_{negative})}$$

$$Precision_{negative} = \frac{274}{274+8}$$

$$Precision_{negative} = 0.97163 \ or \ 97.16\%$$

$$Recall_{negative} = \frac{(TP_{negative})}{(TP_{negative}+FN_{negative})}$$

$$Recall_{negative} = \frac{274}{274+70}$$

$$Recall_{negative} = 0.79651 \ or \ 79.65\%$$

$$F1\ score_{negative} = 2\ x \frac{(Precision_{negative}\ x\ Recall_{negative})}{(Precision_{negative} + Recall_{negative})}$$

$$F1\ score_{negative} = 2\ x \frac{(0.97163\ x\ 0.79651)}{(0.97163 + 0.79651)}$$

$$F1\ score_{negative} = 0.87539 \ or \ 87.54\%$$

Sentiment Neutral SVM:

$$Precision_{neutral} = \frac{(TP_{neutral})}{(TP_{neutral}+FP_{neutral})}$$

$$Precision_{neutral} = \frac{1039}{1039+104}$$

$$Precision_{neutral} = 0.90901 \ or \ 90.90\%$$

$$Recall_{neutral} = \frac{(TP_{neutral})}{(TP_{neutral}+FN_{neutral})}$$

$$Recall_{neutral} = \frac{1039}{1039+11}$$

$$Recall_{neutral} = 0.98952 \ or \ 98.95\%$$

$$F1\ score_{neutral} = 2\ x \frac{(Precision_{neutral}\ x\ Recall_{neutral})}{(Precision_{neutral} + Recall_{neutral})}$$

$$F1\ score_{neutral} = 2\ x \frac{(0.90901\ x\ 0.98952)}{(0.90901 + 0.98952)}$$

$$F1\ score_{neutral} = 0.94756 \ or \ 94.76\%$$

The Naive Bayes evaluation results obtained from program testing are shown in Fig. 8 which displays the accuracy algorithm, precision, recall and F1-scores for each sentiment category.

Based on the results shown in Fig. 8, the evaluation testing from the program code yielded an accuracy of 87% for the Naive Bayes algorithm. For the positive sentiment, the precision was 86%, the recall was 94%, and the F1-score was 90%. For the negative sentiment, the precision was 86%, the recall was 62%, and the F1-score was 72%. For the neutral sentiment, the precision was 90%, the recall was 87%, and the F1-score was 88%. These results indicate that the developed code effectively assesses the Naive Bayes algorithm's accuracy, precision, recall, and F1-score, aligning well with manual calculations.

```
              precision    recall  f1-score   support

    Negative       0.86      0.62      0.72       344
     Neutral       0.90      0.87      0.88      1050
    Positive       0.86      0.94      0.90      1299

    accuracy                           0.87      2693
   macro avg       0.87      0.81      0.83      2693
weighted avg       0.87      0.87      0.87      2693
```

Figure 8. Result accuracy, precision, recall, f1-score of Naive Bayes

The SVM evaluation results obtained from program testing are shown in Fig. 9 which displays the accuracy algorithm, precision, recall and F1-scores for each sentiment category.

```
              precision    recall  f1-score   support

    Negative       0.97      0.80      0.88       344
     Neutral       0.91      0.99      0.95      1050
    Positive       0.98      0.95      0.97      1299

    accuracy                           0.95      2693
   macro avg       0.95      0.91      0.93      2693
weighted avg       0.95      0.95      0.95      2693
```

Figure 9. Result accuracy, precision, recall, f1-score of SVM

Based on the results shown in Fig. 9, the evaluation testing from the program code yielded an accuracy of 95% for the SVM algorithm. For the positive sentiment, the precision was 98%, the recall was 95%, and the F1-score was 97%. For the negative sentiment, the precision was 97%, the recall was 80%, and the F1-score was 88%. For the neutral sentiment, the precision was 91%, the recall was 99%, and the F1-score was 95%. These results demonstrate that the developed code accurately evaluates the SVM algorithm's accuracy, precision, recall, and F1-score, consistent with manual calculations.

The researcher presents a comparison in Table 9 to facilitate an understanding of the comparative evaluation results between the Naive Bayes and SVM algorithms.

Table 9. Comparison of Evaluation Testing for Naive Bayes and SVM

Metrix	**Naive Bayes**	**SVM**
Accuracy	87%	95%
Sentiment Positive		
Precision	86%	98%
Recall	94%	95%
F1-Score	90%	97%
Sentiment Negative		
Precision	86%	97%
Recall	62%	80%
F1-Score	72%	88%
Sentiment Neutral		
Precision	90%	91%
Recall	87%	99%
F1-Score	88%	95%

Based on the results of the values displayed in Table 9 above, it can be concluded that the SVM algorithm demonstrates superior performance compared to Naive Bayes in sentiment analysis of public opinions on the social media platform X regarding the participation of the Indonesian National Team in the Asian Cup U23 2024. SVM achieved an accuracy of 95% compared to Naive Bayes' 87%, indicating that SVM is more reliable in classifying overall sentiment. For positive sentiment, SVM shows a precision of 98% (minimal error in detecting positive sentiment), a recall of 95% (high capability in capturing positive sentiment), and an F1-score of 97% (a good balance between precision and recall). In detecting negative sentiment, SVM also performs better with a precision of 97% and recall of 80%, indicating SVM's effectiveness in recognizing negative sentiment compared to Naive Bayes. For neutral sentiment, SVM has a precision of 91% and recall of 99%, demonstrating an excellent ability to detect almost all neutral sentiments with high accuracy. Thus, SVM proves to be more reliable and consistent in classifying diverse sentiments from user opinions on social media, making it a better algorithm for sentiment analysis in this study.

Based on the established success and failure scenarios, the results of this study can be considered successful because the algorithms used, namely Naive Bayes and SVM, are capable of classifying sentiment with an accuracy well above 50%. Additionally, the precision, recall, and F1-score values for each type of sentiment (positive, negative, and neutral) are also above 50%, demonstrating the algorithms' ability to distinguish between positive, negative, and neutral sentiments effectively. The manual calculation results are also consistent with the program code results, further strengthening the validity of this research.

4 CONCLUSION

This study successfully analyzed public sentiment toward the participation of the Indonesian National Team in the 2024 U-23 Asian Cup using Naive Bayes and SVM algorithms. The findings indicate that the SVM algorithm outperformed Naive Bayes' accuracy by 95%, compared to Naive Bayes' accuracy of 87%. Positive sentiment dominated public opinion, followed by neutral and negative sentiments. The analysis of positive sentiment demonstrates strong public support, which team managers and policymakers can leverage to develop more effective communication and promotional strategies, strengthen relationships with fans, and enhance support for the national team. Conversely, criticisms identified in the negative sentiments can serve as constructive feedback for improving team performance in the future. The study's success is underscored by high values of accuracy, precision, recall, and F1-score for both algorithms, all of which exceeded 50%. This indicates that the algorithms used were adept at distinguishing between positive, negative, and neutral sentiments. Furthermore, the consistency between manually calculated results and program outputs reinforces the validity of this research. For future research, it is recommended to explore the use of more advanced techniques, such as deep learning or pre-trained models, which could enhance the accuracy and depth of sentiment analysis. This study makes a significant contribution by demonstrating how sentiment analysis can be applied in social and sports contexts, providing a solid foundation for future research across various fields.

AUTHOR'S CONTRIBUTION

Sewin Fathurrohman conceptualization of the research, data collection and processing, implementation of the Naive Bayes and SVM algorithms, and writing and revising the draft article. Irfan Ricky Afandi assisted in data collection and preprocessing, wrote the introduction, conducted a literature review, and composed the research conclusion. Firman Noor Hasan validated the model and conducted a statistical analysis of the classification results, structured the methodology and results, and assisted in manuscript editing.

COMPETING INTERESTS

By the publication ethics of this journal Sewin Fathurrohman, Irfan Ricky Afandi, Firman Noor Hasan as the authors of this article declare that this article is free from conflict of interest (COI) and conflict of interest (CI).

ACKNOWLEDGMENT

The authors would like to express sincere gratitude to all parties who have provided support and contributions to this research. We deeply appreciate the assistance provided by the institutions where we are affiliated, as well as our fellow researchers who have offered valuable advice and input throughout the research process. We also extend our thanks to the X users who participated by providing their opinions, and to those who granted access to the necessary data. Without the contributions and support from all these parties, this research would not have been possible.

REFERENCES

[1] I. R. Afandi, N. Pratiwi, A. A. Rizki, M. Irva, and M. F. Aulia, "Perancangan Sistem Informasi Pelayanan Pembuatan Surat Online Di Desa Ciangsana Bebasis Website," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 571–577, 2022, doi: 10.36040/jati.v6i2.5318.

[2] O. B. Islamuddin and I. Yuadi, "Analisis Text Clustering Pada Data Mining Menggunakan Metode K- Nearest Neighbor (KNN) Dan Decision Tree," *J. KHATULISTIWA Inform.*, vol. 11, no. 2, pp. 128–134, 2023, doi: https://doi.org/10.31294/jki.v11i2.16072.g6323.

[3] I. M. Noor and M. Turan, "Sentiment Analysis on New Currency in Kenya using Twitter Dataset," *IJID (International J. Informatics Dev.*, vol. 8, no. 2, p. 81, 2020, doi: 10.14421/ijid.2019.08206.

[4] P. Afikah, I. R. Afandi, and F. N. Hasan, "Implementasi Business Intelligence Untuk Menganalisis Data Kasus Virus Corona di Indonesia Menggunakan Platform Tableau," *Pseudocode*, vol. 9, no. 1, pp. 25–32, 2022, doi: 10.33369/pseudocode.9.1.25-32.

[5] E. R. Setyaningsih, "Sentiment Classification untuk Opini Berita SepakBola," *J. Intell. Syst. Comput.*, vol. 3, no. 2, pp. 93–98, 2021, doi: 10.52985/insyst.v3i2.193.

[6] M. Habibi, D. Handoko, D. Kurniawan, Rasman, and R. Anggriani, "Analisis Framing Robert Entman Pemberitaan Tragedi Kanjuruhan di Media Asing," *Perspekt. Komun. J. Ilmu Komun. Polit. dan Komun. Bisnis*, vol. 7, no. 1, pp. 43–64, 2023, doi: https://doi.org/10.24853/pk.7.1.43-64.

[7] A. Damayanti and A. K. Maulidiyah, "Analisa Sentimen Pada Pembatalan Piala Dunia U20 di Indonesia Menggunakan Metode Support Vector Machine," *JITSI J. Ilm. Ter. Sains dan Teknol.*, vol. 1, no. 2, pp. 97–103, 2023, doi: 10.25139/jitsi.v1i2.6607.

[8] M. Kholilullah, M. Martanto, and U. Hayati, "Analisis Sentimen Pengguna Twitter(X) Tentang Piala Dunia Usia 17 Menggunakan Metode Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 392–398, 2024, doi: 10.36040/jati.v8i1.8378.

[9] R. P. Ratnaningtyas and Y. A. Muhammad, "Analisis Pemberitaan Timnas Indonesia pada Media Daring," *MUKASI J. Ilmu Komun.*, vol. 2, no. 1, pp. 45–52, 2023, doi: 10.54259/mukasi.v2i1.1492.

[10] J. Annissa and A. Ariesta, "Analisis Semiotika Makna Nasionalisme melalui Text Mining pada Media Sosial Twitter di Kejuaraan AFF Tahun 2020," *J. IPTEKKOM J. Ilmu Pengetah. Teknol. Inf.*, vol. 25, no. 1, pp. 69–84, 2023, doi: 10.17933/iptekkom.25.1.2023.69-84.

[11] R. Sulastiyono, A. Setiawan, and S. Nugroho, "Sentimen Analisis Pembatalan Indonesia Menjadi Tuan Rumah Piala Dunia U-20 Menggunakan Metode Naïve Bayes," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1387–1394, 2023, doi: 10.47065/josh.v4i4.3737.

[12] A. D. Dayani, Yuhandri, and W. G. Nurcahyo, "Analisis Sentimen Terhadap Opini Publik pada Sosial Media Twitter Menggunakan Metode Support Vector Machine," *J. KomtekInfo*, vol. 11, no. 1, pp. 1–10, 2024, doi: 10.35134/komtekinfo.v11i1.439.

[13] L. Lestari, M. D. R Wahyudi, and U. Kiftiyani, "Veil and Hijab: Twitter Sentiment Analysis Perspective," *IJID (International J. Informatics Dev.*, vol. 9, no. 1, p. 52, 2020, doi: 10.14421/ijid.2020.09108.

[14] T. Y. Widiyanti, T. B. Adji, and I. Hidayah, "The Social Engagement to Agricultural Issues using Social Network Analysis," *IJID (International J. Informatics Dev.*, vol. 10, no. 1, pp. 1–7, 2021, doi: 10.14421/ijid.2021.2185.

[15] F. Maulana, M. A. Abdullah, J. Sari, D. Z. Siddik, M. Agustinus, and D. D. Saputra, "Sentiment Analysis on the Twitter Pssi Performance Using Text Mining With the Naïve Bayes Algorithm," *J. Comput. Inf. Syst.*, vol. 18, no. 2, pp. 211–216, 2022, doi: 10.33480/pilar.v18i2.3938.

[16] R. Pebrianto, S. N. Nugraha, A. Latif, and M. R. Firdaus, "Analisis Sentimen Twitter Terhadap Menteri Indonesia Dengan Algoritma Support Vector Machine Dan Naive Bayes," *E-Link J. Tek. Elektro dan Inform.*, vol. 17, no. 1, p. 1, 2022, doi: 10.30587/e-link.v17i1.3965.

[17] M. A. Hanafi and A. Solichin, "Analisis Sentimen Terhadap Pssi Atas Tragedi Kanjuruhan Menggunakan Multinomial Naïve Bayes," *Telemat. MKOM*, vol. 15, no. 1, p. 21, 2023, doi: 10.36080/telematikamkom.2292.

[18] S. Rahmawati and M. Habibi, "Public Sentiments Analysis about

Indonesian Social Insurance Administration Organization on Twitter," *IJID (International J. Informatics Dev.*, vol. 9, no. 2, pp. 87–93, 2020, doi: 10.14421/ijid.2020.09205.

[19] A. E. Nanada, A. N. Sihananto, and A. M. Rizki, "Analisis Sentimen Pada Pembatalan Tuan Rumah Indonesia Di Piala Dunia U-20 Menggunakan Fasttext Embeddings Dan Algoritma Recurrent Neural Network," *J. Tek. Inform. Sains dan Ilmu Komun.*, vol. 2, no. 2, 2024, doi: https://doi.org/10.59841/saber.v2i2.1000.

[20] F. Matheos Sarimole and K. Kudrat, "Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine," *J. Sains dan Teknol.*, vol. 5, no. 3, pp. 783–790, 2024, doi: 10.55338/saintek.v5i3.2702.

[21] L. A. Fudholi, N. Rahaningsih, and R. D. Dana, "Sentimen Analisis Perilaku Penggemar Coldplay Di Media Sosial Twitter Menggunakan Metode Naive Bayes," *J. Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 4150–4159, 2024, doi: https://doi.org/10.36040/jati.v8i3.9827.

[22] A. Ardiyanto and E. I. Sela, "Analisis Sentimen Opini Mahasiswa Terhadap Aplikasi Portal Mahasiswa UTY Menggunakan Metode Naïve Bayes Classifier," *J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no. 1, pp. 525–540, 2024, doi: 10.35870/jimik.v5i1.508.

[23] Y. B. Seran and Supatman, "Analisis Sentimen Masyarakat Terhadap Kinerja Kerja Presiden Joko Widodo Enggunakan Algoritma Support Vector Machine," *J. Mhs. Tek. Inform.*, vol. 8, no. 4, pp. 7190–7195, 2024, doi: https://doi.org/10.36040/jati.v8i4.10171.

[24] S. A. Sutresno, "Analisis Sentimen Masyarakat Indonesia Terhadap Dampak Penurunan Global Sebagai Akibat Resesi di Twitter," *Build. Informatics, Technol. Sci.*, vol. 4, no. 4, pp. 1959–1966, 2023, doi: 10.47065/bits.v4i4.3149.

[25] E. L. Rara and E. Mailoa, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Perubahan Piala Dunia U-20," *Progresif J. Ilm. Komput.*, vol. 20, no. 1, pp. 259–268, 2024, doi: https://doi.org/10.35889/progresif.v20i1.1550.

[26] H. Setiawan and I. Zufria, "Analisis Sentimen Pembatalan Indonesia Sebagai Tuan Rumah Piala Dunia FIFA U-20 Menggunakan Naïve Bayes," *J. Media Inform. Budidarma*, vol. 7, no. 3, pp. 1003–1012, 2023, doi: 10.30865/mib.v7i3.6144.

[27] M. A. Kurniawan, A. S. Yuniarti, and S. Makin, "Analisis Sentimen AFC U23 Asian Cup 2024 menggunakan Algoritma Naïve Bayes," *J. IPSIKOM*, vol. 12, no. 1, pp. 2478–2483, 2004, doi: https://doi.org/10.58217/ipsikom.v12i1.275.

[28] N. Dalifah, N. Suarna, and W. Prihartono, "Analisis Data Sentimen Negatif Pada Opini Pengguna Twitter Terhadap Berita Sepak Bola Liga 1 Tahun 2022 Dengan Penerapan Support Vector Mechine," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 209–214, 2024, doi: 10.36040/jati.v8i1.8303.

[29] M. Parhusip, S. Sudianto, and T. G. Laksana, "Sentiment Analysis of the Public Towards the Kanjuruhan Tragedy with the Support Vector Machine Method," *JUITA J. Inform.*, vol. 11, no. 2, p. 241, 2023, doi: 10.30595/juita.v11i2.17405.

[30] F. P. Herlambang and D. Avianto, "Analisis Sentimen Opini Pengguna Twitter Terhadap Tragedi Kanjuruhan Malang dengan Metode Support Vector Machine," *J. Media Inform. Budidarma*, vol. 7, no. 4, pp. 1727–1739, 2023, doi: 10.30865/mib.v7i4.6332.

[31] R. Ardianto and S. K. Wibisono, "Analisis Deep Learning Metode Convolutional Neural Network Dalam Klasifikasi Varietas Gandum," *J. Kolaboratif Sains(JKS)*, vol. 6, no. 12, pp. 2081–2092, 2023, doi: 10.56338/jks.v6i12.4938.

[32] B. A. Maulana, M. J. Fahmi, A. M. Imran, and N. Hidayati, "Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 375–384, 2024, doi: 10.57152/malcom.v4i2.1206.

[33] A. T. Susilawati, N. A. Lestari, and P. A. Nina, "Analisis Sentimen Publik Pada Twitter Terhadap Boikot Produk Israel Menggunakan Metode Naïve Bayes," *J. Ilm. Mhs.*, vol. 2, no. 1, pp. 26–35, 2024, [Online]. Available: https://doi.org/10.59603/niantanasikka.v2i1.240

[34] I. R. Afandi, F. Noor Hasan, A. A. Rizki, N. Pratiwi, and Z. Halim, "Analisis Sentimen Opini Masyarakat Terkait Pelayanan Jasa Ekspedisi Anteraja Dengan Metode Naive Bayes," *J. Linguist. Komputasional*, vol. 5, no. 2, pp. 63–70, 2022, doi: https://doi.org/10.26418/jlk.v5i2.107.

[35] D. A. Nugroho and F. N. Hasan, "Analisis Sentimen Kegiatan Pembersihan Sampah Pada Media Sosial X Menggunakan SVM dan Naïve Bayes," *J. MEDIA Inform. BUDIDARMA*, vol. 8, no. April, pp. 843–853, 2024, doi: 10.30865/mib.v8i2.7562.

[36] M. Samantri and Afiyati, "Perbandingan Algoritma Support Vector Machine dan Random Forest untuk Analisis Sentimen Terhadap Kebijakan Pemerintah Indonesia Terkait Kenaikan Harga BBM Tahun 2022," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 8, no. 1, pp. 1–9, 2024, doi: 10.35870/jtik.v8i1.1202.

[37] G. Radiena and A. Nugroho, "Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Kai Access Menggunakan Metode Support Vector Machine," *J. Pendidik. Teknol. Inf.*, vol. 6, no. 1, pp. 1–10, 2023, doi: 10.37792/jukanti.v6i1.836.

[38] R. Wati, S. Ernawati, and H. Rachmi, "Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH," *J. Manaj. Inform.*, vol. 13, no. 1, pp. 84–93, 2023, doi: 10.34010/jamika.v13i1.9424.

[39] Fauzan Baehaqi and N. Cahyono, "Analisis Sentimen Terhadap Cyberbullying Pada Komentar Di Instagram Menggunakan Algoritma Naïve Bayes," *Indones. J. Comput. Sci.*, vol. 13, no. 1, pp. 1051–1063, 2024, doi: 10.33022/ijcs.v13i1.3301.

[40] I. A. Ricky, I. F. Hanif, F. N. Hasan, E. S. Sinduningrum, Z. Halim, and N. Nunik, "Analisis Sentimen Opini Masyarakat Terkait Penyelenggaraan Sistem Elektronik Menggunakan Metode Logistic Regression," *J. Linguist. Komputasional*, vol. 5, no. 2, p. 77, 2022, [Online]. Available: https://t.co/23c4krbjp

[41] J. E. B. Sinulingga and H. C. K. Sitorus, "Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF," *J. Manaj. Inform.*, vol. 14, no. 1, pp. 42–53, 2024, doi: 10.34010/jamika.v14i1.11946.

[42] A. K. Sari, A. Irsyad, D. N. Aini, Islamiyah, and S. E. Ginting, "Analisis Sentimen Twitter Menggunakan Machine Learning untuk Identifikasi Konten Negatif," *J. Adopsi Teknol. dan Sist. Inf.*, vol. 3, no. 1, pp. 64–73, 2024, doi: https://doi.org/10.30872/atasi.v3i1.1373.

[43] A. Karimah, G. Dwilestari, and Mulyawan, "Analisis Sentimen Komentar Video Mobil Listrik Di Platform YOUTUBE DENGAN METODE NAIVE BAYES," *J. Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 767–773, 2024, doi: https://doi.org/10.36040/jati.v8i1.8373.

[44] L. A. Susanto, "Komparasi Model Support Vector Machine Dan K-Nearest Neighbor Pada Analisis Sentimen Aplikasi Polri Super App," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4152.

[45] A. T. Mukti and F. N. Hasan, "Analisis Sentimen Warganet Terhadap Keberadaan Juru Parkir Liar Menggunakan Metode Naive Bayes Classifier," *J. Media Inform. Budidarma* , vol. 8, no. 1, pp. 644–653, 2024, doi: 10.30865/mib.v8i1.6982.

[46] S. Nurhaliza, D. Febriawan, and F. N. Hasan, "Sentiment Analysis of Society Towards the Child-free Phenomenon (Life Without Children) on Twitter Using Naïve Bayes Algorithm," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 13, no. 1, pp. 40–48, 2024, doi: 10.32736/sisfokom.v13i1.1944.

[47] D. Ananda and R. R. Suryono, "Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode

Support Vector Machine dan Naïve Bayes," *J. Media Inform. Budidarma*, vol. 8, no. 2, pp. 748–757, 2024, doi: 10.30865/mib.v8i2.7517.

[48] P. W. Hardjita, Nurochman, and R. Hidayat, "Sentiment Analysis of Tweets on Prakerja Card using Convolutional Neural Network and Naive Bayes," *IJID (International J. Informatics Dev.*, vol. 10, no. 2, pp. 82–91, 2022, doi: 10.14421/ijid.2021.3007.