Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

Analysis Of Indonesian People's Sentiment Towards 2024 Presidential Candidates On Social Media Using Naïve Bayes Classifier and Support Vector Machine

Nia Mardiah*, Leni Marlina, Khairul, Zulham Sitorus, Muhammad Iqbal

Faculty of Postgraduate, Master of Information Technology Study Program, Universitas Pembangunan Panca Budi, Medan, Indonesia
Email: niamardiah02@gmail.com, lheny@pancabudi.ac.id, khairul@dosen.pancabudi.ac.id, zulhamsitorus@dosen.pancabudi.ac.id, muhammadiqbal@dosen.pancabudi.ac.id
Correspondence Author Email: niamardiah02@gmail.com
Submitted: **10/08/2024**; Accepted: **10/09/2024**; Published: **11/09/2024**

Abstract−This research aims to analyze the sentiment of the Indonesian public towards the 2024 presidential candidates on social media platforms X and Instagram. The main issue addressed is how to determine public opinion as disseminated on social media regarding the presidential candidates. To address this issue, two classification methods are used: Naïve Bayes Classifier and Support Vector Machine (SVM). The objective of this research is to measure public sentiment, both positive and negative, towards the 2024 presidential candidates using these two methods. The research findings indicate that the implementation of the Naïve Bayes method with manual labeling achieved the highest accuracy of 86% for X data and 85% for Instagram comments data. Meanwhile, with lexicon-based labeling, the highest accuracy was 60% for both X and Instagram data. The SVM method with manual labeling also achieved the highest accuracy of 86% for X data and 85% for Instagram data. With lexicon-based labeling, the highest accuracy was 60% for X data and 70% for Instagram data. This research concludes that both Naïve Bayes and SVM demonstrate strong performance in sentiment analysis on social media, with SVM slightly outperforming in some scenarios. The implementation of these two methods provides valuable insights into public opinion towards the 2024 presidential candidates on social media.

Keywords: Sentiment Analysis; Naïve Bayes Classifier; Support Vector Machine (SVM); Social Media (X and Instagram); 2024 Presidential Candidates

1. INTRODUCTION

Indonesia, as a democratic country, holds general elections for president and vice president. The history of elections in Indonesia has taken place several times, but direct elections by citizens only began in the reform era after the fall of the New Order in 2004[1][2]. General elections are usually held periodically. The presidential and vice presidential elections planned for 2024 are important in realizing democracy in the Unitary State of the Republic of Indonesia. Presidential candidates and campaign teams can utilize social media platforms such as X and Instagram to spread campaign messages. The popularity of presidential and vice presidential candidates is often assessed based on public opinion.

In this modern era, social media such as X and Instagram have become a platform for people to express their views, especially regarding the 2024 Presidential Candidates. Indonesia ranked fourth in terms of active users of Instagram social media in January 2023, reaching a total of 89.15 million users. This ranking is after India, the United States, and Brazil. Meanwhile, in terms of active X users in January 2023, Indonesia ranked fifth with 24 million users after the United States, Japan, India, and Brazil[3]. Therefore, it can be concluded that Instagram and X have become the main channels for Indonesians to communicate and voice their opinions about the 2024 Presidential Candidates. This is reflected in the stories shared by these social media users.

Sentiment analysis or opinion mining is a computational process that parses people's opinions conveyed in the form of text or documents[4]. The purpose of sentiment analysis is to assess the view or tendency of opinion on an issue or object, whether it is positive or negative. The technique used in extracting keywords, opinions, and reviews or information is Text Mining. The application of Text Mining helps in understanding people's point of view contained in text data[5][6].

Some techniques that can be used in sentiment analysis are Naïve Bayes, Support Vector Machine, and Decision Tree. In this study, the Naïve Bayes Classifier classification method was chosen because it has the advantage of high speed and accuracy when applied to large and diverse datasets[7][8]. This choice is in line with the findings of several researchers who stated that the Naïve Bayes Classifier shows high speed and accuracy when used in large and varied datasets. In addition, this method has the advantages of being simple, fast, and high accuracy[9].

This research also uses the Support Vector Machine (SVM) method. This method was chosen because it is the most accurate method for text classification[10][11]. In previous research on the comparison of the use of several methods in text classification shows that SVM is an effective method for text classification. The test results of the SVM method on text classification which is divided into four categories show an average accuracy rate of 90.72%, recall rate of 86.37%, and F1-Value of 88.97%. Based on these values, SVM can be said to have stable performance compared to other methods. In previous research on the comparison of the use of several methods in text classification shows that SVM is an effective method for text classification. The test results of the SVM method on text classification which is divided into four categories show an average accuracy rate of 90.72%, recall rate of 86.37%, and F1-Value of 88.97%. Based on these values, SVM can be said to have stable performance compared to other methods[12].

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

The sources of relevant research are journals on sentiment analysis of 2024 presidential candidates using Naïve Bayes Classifier and Support Vector Machine. One of them is research conducted by Abdillah and Hasan in 2023 with a discussion of sentiment analysis of presidential candidates based on tweets on social media using Naïve Bayes Classifier, which shows that Anies Baswedan received 72 negative sentiments and 201 positive sentiments from 273 data analyzed, Ganjar Pranowo received 142 negative sentiments and 160 positive sentiments from a total of 302 data, Prabowo Subianto received 394 negative sentiments and 188 positive sentiments from a total of 582 data, and Sandiaga Uno received 170 negative sentiments and 228 positive sentiments from a total of 398 data[13]. In addition, research by Findawati et al in 2023 discussed sentiment analysis of X-based potential 2024 presidential candidates, showing that X users express sentiments that tend to be positive towards each potential candidate, with SVM having an accuracy of 84%, Bernoulli Naïve Bayes 77%, and Logistic Regression 84%[14]. Another study by S. L. Ramadhan and Windarto in 2023 discussed the application of Naïve Bayes to analyze the sentiment of X users towards the determination of the 2024 PDIP presidential candidate, showing a positive sentiment of 84.18% and a negative of 15.82% with an accuracy of 69%, precision of 91%, and recall of 72%[15]. Furthermore, research by Asno Azzawagama Firdaus et al in 2023 discussed sentiment analysis on the projection of the 2024 presidential election using Support Vector Machine, showing the accuracy of the method owned by the dataset Anies Baswedan 73%, Ganjar Pranowo 79%, and Prabowo Subianto 79%, with the popularity of words that appear in X related to President 2024 are "Prabowo Subianto", "president ri", "presidential candidate", "Ganjar Pranowo", and "Anies Baswedan"[16]. Finally, research by F. A. Ramadhan and Gunawan in 2023 discussed sentiment analysis of 2024 presidential candidates in Indonesia using a logistic regression approach, showing an overall average accuracy score of 79%, with the highest accuracy of 93% in the Prabowo Subianto dataset and the lowest of 70% in the Anies Baswedan and Muhaimin Iskandar datasets. This analysis provides valuable insights into public sentiment towards potential candidates for the 2024 Presidential Election in Indonesia[17].

The purpose of this research is to find out the implementation of the Naïve Bayes Classifier method in analyzing public sentiment towards 2024 Presidential Candidates on social media X and Instagram and determine the level of accuracy resulting from the implementation of the method. In addition, this study also aims to determine the implementation of the Support Vector Machine method in analyzing public sentiment towards Presidential Candidates 2024 on social media X and Instagram and evaluating the level of accuracy resulting from the implementation of these methods. Based on the research background that has been described, the researcher will take the title of the proposal "Analysis of Indonesian Public Sentiment towards 2024 Presidential Candidates on Social Media Using Naïve Bayes Classifier and Support Vector Machine".

2. RESEARCH METHODOLOGY

2.1 Research Stages

The stages of research carried out in this study were made to make it easier to understand the direction of the research and what stages were carried out in this study. The chart of the research stages can be seen in Figure 1 below.

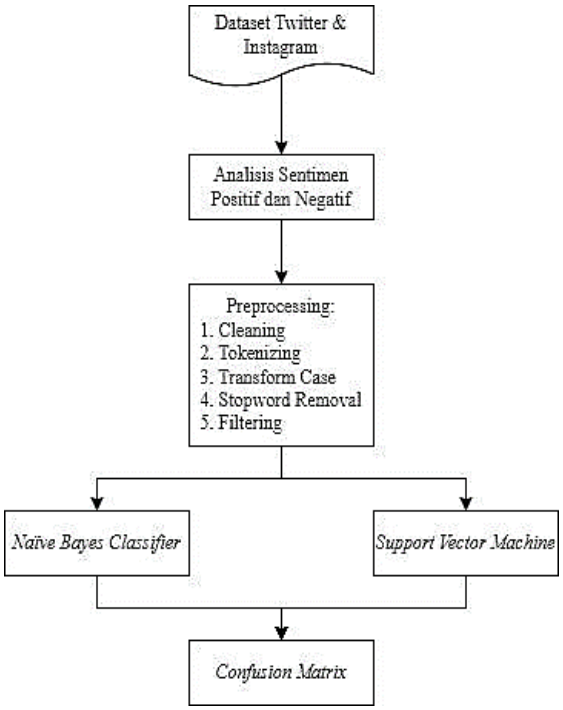

Figure 1. Research Stages

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

The following is an explanation of the research stages depicted in Figure 1 data collection, sentiment analysis, preprocessing, application of Naive Bayes and SVM algorithms, and evaluation of results using confusion matrix.

1. Dataset collection from Twitter and Instagram
   In this stage, data related to the 2024 presidential candidates was collected from two major social media platforms, namely Twitter and Instagram. This data includes various types of posts, including tweets, retweets, comments, and captions, which contain public opinions and sentiments towards presidential candidates.
2. Positive and negative sentiment analysis
   This stage involves sentiment analysis to determine whether the opinions expressed in the data are positive or negative. This analysis aims to understand how the public feels about a particular presidential candidate based on social media content.
3. Data preprocessing
   This stage includes several important steps to prepare the data before further analysis, namely cleaning, tokenizing, case transformation, stopword removal, and filtering.
4. Application of algorithms for data analysis
   Two main algorithms are applied to the preprocessed data, namely naïve bayes classifier and support vector machine (SVM).
5. Evaluation of results using confusion matrix
   This stage involves evaluating the performance of the algorithms using confusion matrix. Confusion matrix is an evaluation tool that shows the number of correct and incorrect predictions made by the model, allowing the calculation of evaluation metrics such as accuracy, precision, recall, and F1-score to assess the performance of the applied algorithm.

2.2 Text Preprocessing

In conducting text mining, the document text used must be prepared first, then it can be used for the main process. Text preprocessing is the initial stage of text mining, namely converting information from each data source into a standardized form or format. In text mining, raw data containing information is unstructured data, so it is necessary to change the form into structured data as needed, which will usually become numerical values[18][19]. In general, the process carried out in the preprocessing stage is as follows[20]:

1. Cleaning
   Cleaning is a stage to clean the data in the set and select words that are not needed, have no meaning, or meanings that affect sentiment such as html, links, mentions, and hastags[21].
2. Tokenizing
   Tokenization is the breakdown of data sets into tokens or pieces of words to make the next stage easier. For example, the sentence 'I want to eat' is decapitated into ['me', 'want', 'eat'][22].
3. Transform Case
   Transform case is the stage of converting text data sentences into uniform text. This stage is always present in the text preprocessing process because the existing data is not always structured in its use of letters. With this stage, it can play a role in leveling the use of capital letters. For example, the words "Data" and "data" will be read as two different words, so that through this process the system can read effectively[23].
4. Stopword Removal
   Stopword is a stage to remove words that often appear but have no important meaning and their meaning has no effect on the system, such as 'oh', 'at', 'on', and so on[24].
5. Filtering
   Filtering is a stage to remove words that are too short and too long with a minimum of 3 letters and a maximum of 25 letters[25].

2.3 Naïve Bayes Classifier

Naïve Bayes Classifier is one of the algorithms used to classify text using Machine Learning methods[26]. This algorithm utilizes probability and statistical calculations proposed by Thomas Bayes. The Naïve Bayes approach is based on classification techniques that have been tested for effectiveness and efficiency in large databases[27]. In large amounts of data, Naïve Bayes shows a high level of accuracy and efficiency. In addition, Naïve Bayes performance also has a short classification time, speeding up the sentiment analysis process of the system. The Naïve Bayes method assumes that each variable is independent of each other and not related to any other variable. Thus, a document is considered as a collection of words that make up the document regardless of the order in which the words appear in the document[28][29][30]. Perhitungan The probability calculation can be considered as the product of the probability of occurrence of words in the document. The Bayes theorem equation is as follows[31]:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Description:
X = Data with unknown class
H = Hypothesized data

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

P(H|X) = Probability of hypothesis H based on condition X (a posteriori probability)
P(H) = Probability of hypothesis H (prior probability)
P(X|H) = Probability of X based on the condition in hypothesis H
P(X) = Probability of X

2.4 Support Vector Machine

Support Vector Machine (SVM) was developed by Boser, Guyon, and Vapnik, first introduced in 1992 at the Annual Workshop on Computational Learning Theory[32]. The basic concept of the SVM method is actually a combination of computational theories that have existed in previous years, such as hyperplan margins, kernels introduced by Aronszajn in 1950, Lagrange Multiplier discovered by Joseph Louis Lagrange in 1766, and so on with other supporting concepts[33][34]. SVM is a technique for making predictions, both predictions in the case of regression and classification. The SVM technique is used to obtain an optimal separation function (hyperplane) to separate observations that have different target variable values[35][36]. This hyperplane can be a line in two dimensions and can be a flat plane in multiple dimensions.

2.5 Confusion Matrix

Confusion matrix is a matrix used to evaluate the classification model process in the form of the number of correct and incorrect test data[37]. This matrix can determine the quality of the classification model performance. This matrix contains predicted target data compared to actual target data. Prediction data is the value obtained from machine learning modeling results, while actual data is the actual value owned.The existence of a confusion matrix to determine the extent to which machine learning works as desired[38].

1. Accuracy is the ratio value of the tweet data that has been detected in the test. The accuracy value can show the closeness between the system prediction value and human prediction. The following formula can be seen in formula 2:

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

2. Precision is the value of the system's accuracy regarding system information to show correct positive data and negative data. This precision value is generated from the positive prediction value compared to the number of positive values, which can be seen in formula 3:

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

3. Recall is a value that shows the success rate of retrieving information about true positive and negative data. Recall is generated from the number of true positive values compared to the actual positive values, which can be seen in formula 4:

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

4. F1 Score is a weighted average comparison of precision and recall, with formula 5:

$$\text{F1 Score} = 2 * \frac{\text{Recall}*\text{Precision}}{\text{Recall+Precision}} \quad (5)$$

Description:
TP (True Positive) = The number of data on the actual value of the positive class and the predicted value of the positive class.
TN (True Negative) = The number of negative actual value data and negative predicted values.
FP (False Positive) = The number of positive actual value data and negative prediction values.
FN (False Negative) = The number of negative actual value data and positive predicted values.

3. RESULT AND DISCUSSION

3.1 Crawling Data

The dataset sources in this study were obtained from social media platforms X and Instagram using different methods.

1. Data X
   X data for sentiment analysis is obtained through the Python programming language using tweet harvest. Tweet harvest is a special tool for crawling data on social media X using the Application Programming Interface (API). Here is Figure 2, which is an example of the results of data retrieval using Tweet Harvest.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

created_at	id_str	full_text	quote_count	reply_count	retweet_count	favorite_count	lang	user_id_str	conversation_id_str	username	tweet_url
Wed Jan 10 13:40:45 +0000 2024	1,75E+18	@B_nin927 Nampaknya beliau sudah memperkirakan, bahwa segala hal bisa terjadi. Menang atau syahid. Tapi langkah tak boleh surut. Darah kakeknya mengalir dan merasuki jiwa Pak #aniesbaswedan . Semoga Allah jaga beliau.	0	0	0	0	in	,68E+1	1,75E+18	NifMuh01	https://twitter.com/NifMuh01/status/1745078223782678868

Figure 2. Example of Data Capture Results Using Tweet Harvest

2. Data Instagram
   Instagram data for sentiment analysis was obtained using Data Scraper extensions. The extension is a tool that makes it easy to extract data from HTML web pages and import it into a Microsoft Excel spreadsheet. The following figure 3 is an example of the results of data retrieval using Data Scraper.

Column 1
ilhamfmsyh========== =====>>>>>KITA PERGI KE GBK DULU AMBIL AMPLOPNYA TERUS KE JIS DUKUNG AM1N SETUJU?2w5,896 likesReplySee translationComment OptionsLike

Figure 3. Example of Data Retrieval Results Using Data Scraper

3.2 Explore

The results of crawling using X contain 12 columns regarding the details of a tweet which can be seen in Figure 2. In this study, the X data attributes used to facilitate data analysis are only full_text attributes as shown in Table 1 below.

Table 1. X Data on full_text Attribute

full_text
P Info cowo yg kaya abah anies dong aku mau memaksakan diri
@UbahBareng Desak Anies yang terbaru konsepnya bagaimana yamengkritis kinerja pemerintah atau model anak sekolahan antara dosen dan murid tapi dialog 2 arah

The crawling results using Instagram only contain 1 column regarding the details of a comment on a post on Instagram which can be seen in Figure 3. This crawling result data needs to be cleaned manually, leaving only the actual comments. After the data is cleaned, the results obtained can be seen in table 2.

Table 2. Instagram Data on Comment Attributes

Comment
semoga allah meridhoi abah untuk menjadi pempimpin negeri iniðŸ˜ðŸ¤• ðŸ¥¹
Pak Anies saya mahasiswa Yaman dan di Yaman hari ini sudah melalukan pencoblosan. Dan saya sudah menyumbangkan satu suara saya buat bapak. Bismillah AMIN menang

3.3 Modify

a. Cleaning
   The first step is to perform cleaning, where special characters, URLs, and others that are not needed in the analysis are removed. The results of data cleaning can be seen in Table 3 below.

Table 3. Cleaning Data

Before	After
P Info cowo yg kaya abah anies dong aku mau memaksakan diri	P Info cowo yg kaya abah anies dong aku mau memaksakan diri
@UbahBareng Desak Anies yang terbaru konsepnya bagaimana ya mengkritis kinerja pemerintah atau model anak sekolahan antara dosen dan murid tapi dialog 2 arah	Desak Anies yang terbaru konsepnya bagaimana ya mengkritis kinerja pemerintah atau model anak sekolahan antara dosen dan murid tapi dialog arah

b. Transform Case
   The next step is transform case which aims to reduce data complexity by eliminating variations in capitalization. The transform case results can be seen in table 4 below.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

Table 4. Transform Case

Before	After
P Info cowo yg kaya abah anies dong aku mau memaksakan diri	p info cowo yg kaya abah anies dong aku mau memaksakan diri
Desak Anies yang terbaru konsepnya bagaimana ya mengkritis kinerja pemerintah atau model anak sekolahan antara dosen dan murid tapi dialog arah	desak anies yang terbaru konsepnya bagaimana ya mengkritis kinerja pemerintah atau model anak sekolahan antara dosen dan murid tapi dialog arah

c. Tokenizing
The next stage of text preprocessing is tokenizing, which separates the text into individual tokens such as words. The results of tokenizing can be seen in table 5 below.

Table 5. Tokenzing

Before	After
p info cowo yg kaya abah anies dong aku mau memaksakan diri	['p', 'info', 'cowo', 'yg', 'kaya', 'abah', 'anies', 'dong', 'aku', 'mau', 'memaksakan', 'diri']
desak anies yang terbaru konsepnya bagaimana ya mengkritis kinerja pemerintah atau model anak sekolahan antara dosen dan murid tapi dialog arah	['desak', 'anies', 'yang', 'terbaru', 'konsepnya', 'bagaimana', 'ya', 'mengkritis', 'kinerja', 'pemerintah', 'atau', 'model', 'anak', 'sekolahan', 'antara', 'dosen', 'dan', 'murid', 'tapi', 'dialog', 'arah']

d. Stopword Removal
Stopword removal can help reduce data dimensionality and improve analysis accuracy. The results of stopword removal can be seen in Table 6 below.

Table 6. Stopword Removal

Before	After
['p', 'info', 'cowo', 'yg', 'kaya', 'abah', 'anies', 'dong', 'aku', 'mau', 'memaksakan', 'diri']	['p', 'info', 'cowo', 'yg', 'kaya', 'abah', 'anies', 'memaksakan']
[‘desak’, ‘anies’, ‘yang’, ‘terbaru’, ‘konsepnya’, ‘bagaimana’, ‘ya’, ‘mengkritis’, ‘kinerja’, ‘pemerintah’, ‘atau’, ‘model’, ‘anak’, ‘sekolahan’, ‘antara’, ‘dosen’, ‘dan’, ‘murid’, ‘tapi’, ‘dialog’, ‘arah’]	[‘desak’, ‘anies’, ‘terbaru’, ‘konsepnya’, ‘bagaimana’, ‘mengkritis’, ‘kinerja’, ‘pemerintah’, ‘model’, ‘anak’, ‘sekolahan’, ‘antara’, ‘dosen’, ‘murid’, ‘dialog’, ‘arah’]

e. Stemming
Stemming on Indonesian datasets is done using the Sastrawi library. The stemming results can be seen in Table 7 below.

Table 7. Stemming

Before	After
['p', 'info', 'cowo', 'yg', 'kaya', 'abah', 'anies', 'memaksakan']	p info cowo yg kaya abah anies memaksakan
[‘desak’, ‘anies’, ‘terbaru’, ‘konsepnya’, ‘bagaimana’, ‘mengkritis’, ‘kinerja’, ‘pemerintah’, ‘model’, ‘anak’, ‘sekolahan’, ‘antara’, ‘dosen’, ‘murid’, ‘dialog’, ‘arah’]	desak anies terbaru konsepnya bagaimana mengkritis kinerja pemerintah model anak sekolahan antara dosen murid dialog arah

3.4 Model

The modeling stage of this research uses two models to label sentiment on the dataset, namely the manual method and the lexicon-based method. Data labeling was conducted on 1000 data divided into tweet data and Instagram comment data on three 2024 presidential candidates, namely Anies Baswedan, Prabowo Subianto, and Ganjar Pranowo. The data division can be seen in Table 8 below.

Table 8. Data Sharing

Presidential Candidates	Total Data	
	X	Instagram
Anies Baswedan	233	100
Prabowo Subianto	234	100
Ganjar Pranowo	233	100

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

Total	1000

3.5 Interpretation of Results

Based on the testing process by applying the Naïve Bayes and Support Vector Machine methods to X data and Instagram data, the accuracy data of the two methods is obtained as follows.

3.5.1 Manual Labeling

Table 9. Anies Baswedan X Data with Manual Labeling

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
90:10:00	79,17%	26,39%	33,33%	29,45%	79,16%	43,93%	38,24%	40,87%
80:20:00	59,57%	20,74%	31,11%	24,84%	65,95%	71,33%	47%	56,65%
70:30:00	55,71%	32,30%	33,05%	32,69%	54%	45%	39%	41,79%
60:40:00	51,61%	24,46%	31,45%	62%	54%	63,66%	39%	48,35%

Based on table 9, it can be seen that Anies Baswedan's X data which is labeled manually has the highest accuracy of 79.17% with the Naïve Bayes method and 79.16% with the Support Vector Machine method.

Table 10. Anies Baswedan's Instagram Data with Manual Labeling

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
80:20:00	85,00%	29,66%	33,33%	31,38%	85%	29,66%	33,33%	31,38%
70:30:00	77,00%	25,66%	33,33%	28,99%	80%	44%	39%	41,35%
60:40:00	74,00%	24,66%	33,33%	28,39%	77%	42,66%	37,33%	39,86%

Based on table 10, it can be seen that Anies Baswedan's Instagram data that is manually labeled has the highest accuracy of 85.00% with the Naïve Bayes method and the Support Vector Machine method.

Table 11. Prabowo Subianto X Data with Manual Labeling

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
90:10:00	61,00%	37,00%	39,66%	38,21%	57%	36%	34,33%	35,15%
80:20:00	58,00%	37,33%	42,00%	39,56%	51%	34,66%	36,33%	35,00%
70:30:00	58,00%	41,66%	44,00%	42,75%	52%	43,66%	42%	42,77%
60:40:00	51,00%	36,33%	38,00%	37%	46,67%	44,33%	49,33%	52%

Based on table 11, it can be seen that data X belonging to Prabowo Subianto that was manually labeled obtained the highest accuracy of 61.00% with the Naïve Bayes method and 57% with the Support Vector Machine method.

Table 12. Prabowo Subianto's Instagram Data with Manual Labeling

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
80:20:00	50,00%	17,66%	30,33%	22,30%	55%	35,33%	34,33%	34,86%
70:30:00	57,00%	35,66%	51,00%	42,00%	63%	45,66%	39,66%	42,42%
60:40:00	55,00%	35,00%	33,66%	34,30%	57%	39%	36,66%	37,84%

Based on table 12, it can be seen that Prabowo Subianto's Instagram data that was manually labeled obtained the highest accuracy of 57% with the Naïve Bayes method and 63% with the Support Vector Machine method.

Table 13. Ganjar Pranowo X Data with Manual Labeling

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
90:10:00	90:10:00	78,00%	26,00%	33,33%	29,22%	78%	26%	33,33%
80:20:00	80:20:00	83,00%	27,66%	33,33%	30,25%	83%	27,66%	33,33%
70:30:00	70:30:00	81,00%	27,00%	33,33%	29,81%	81%	27%	33,33%
60:40:00	60:40:00	86,00%	28,66%	33,33%	30%	86%	28,66%	33,33%

Based on table 13, it can be seen that Ganjar Pranowo's X data which is manually labeled has the highest accuracy of 86% with Naïve Bayes method and Support Vector Machine method.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

Table 14. Ganjar Pranowo's Instagram Data with Manual Labeling

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
80:20:00	70,00%	23,33%	33,33%	27,41%	75%	58%	50%	53,70%
70:30:00	67,00%	22,33%	33,33%	26,75%	70%	56,33%	41,66%	47,45%
60:40:00	75,00%	25,00%	33,33%	28,55%	78%	59%	41,66%	48,65%

Based on table 14, it can be seen that Ganjar Pranowo's Instagram data which is manually labeled has the highest accuracy of 75% with Naïve Bayes method and 78% with Support Vector Machine method.

From the results of the above research on manually labeled data, the following conclusions can be drawn.

a. Effect of Data Sharing
   The variation in model accuracy depends on the division ratio of training and testing data. In this study, it is seen that the more data allocated to training (higher ratio), the model performance tends to improve. However, this does not always apply consistently for every combination of model and data type.
b. Model Consistency
   It can be seen that Support Vector Machine tends to be more consistent in its performance compared to Naive Bayes, especially in the context of variations in data sharing ratios.
c. Reliance on Subjectivity of Labeling
   Manual labeling by researchers can have a significant impact on model accuracy. This can be indicated by the variation in model performance between the X data and Instagram comment data, where there may be differences in the subjectivity or quality of the labels assigned.

In conclusion, model selection and data partitioning are critical to good performance in text classification. In addition, consistency in the labeling process and objective evaluation can help minimize variations in model accuracy results.

3.5.2 Lexicon Based Labeling

Table 15. Anies Baswedan X Data with Lexicon Based

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
90:10:00	54,00%	35,66%	37,00%	36,35%	46%	32%	32,33%	32,15%
80:20:00	53,00%	55,66%	46,00%	50,53%	43%	36,66%	37%	36,84%
70:30:00	41,00%	42,66%	37,33%	39,29%	42,65%	40,33%	45,33%	43%
60:40:00	37,00%	25,66%	24,00%	12%	40%	44,33%	37,66%	40,19%

Based on Table 15, it can be seen that Anies Baswedan's X data, which is labeled with lexicon-based, has the highest accuracy of 54% with the Naïve Bayes method and 46% with the Support Vector Machine method.

Table 16. Anies Baswedan's Instagram Data with Lexicon Based

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
80:20:00	50,00%	49,00%	44,33%	46,47%	55%	83,33%	48,33%	61,18%
70:30:00	53,00%	50,66%	40,00%	44,71%	63%	86%	53,33%	65,82%
60:40:00	59,00%	52,66%	39,00%	44,88%	67%	79,66%	51,66%	62,71%

Based on Table 16, it can be seen that Anies Baswedan's X data, which is labeled with lexicon-based, has the highest accuracy of 59% with the Naïve Bayes method and 67% with the Support Vector Machine method.

Table 17. Prabowo Subianto X Data with Lexicon Based

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
90:10:00	52,00%	51,00%	54,00%	52,46%	57%	57%	57%	57%
80:20:00	40,00%	40,66%	40,66%	40,66%	40%	42,33%	40,66%	40,99%
70:30:00	46,00%	45,33%	47,00%	46,14%	48%	52%	48%	49,92%
60:40:00	46,00%	46,00%	47,00%	46%	47%	50,66%	46,66%	48,58%

Based on Table 17, it can be seen that Anies Baswedan's X data, which is labeled with lexicon-based, has the highest accuracy of 52% with the Naïve Bayes method and 57% with the Support Vector Machine method.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

Table 18. Prabowo Subianto Instagram Data with Lexicon Based

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
80:20:00	55,00%	18,33%	33,33%	23,64%	39,72%	41,66%	38%	60%
70:30:00	60,00%	52,33%	40,66%	45,77%	67%	70,33%	54,66%	61,60%
60:40:00	50,00%	16,66%	33,33%	22,22%	60%	62,66%	50%	56,65%

Based on the table above, it can be seen that Anies Baswedan's X data, which is labeled with lexicon-based, has the highest accuracy of 60% with the Naïve Bayes method and 67% with the Support Vector Machine method.

Table 19. Ganjar Pranowo's X Data with Lexicon Based

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
90:10:00	33,00%	11,00%	33,33%	16,55%	33%	11%	33,33%	16,55%
80:20:00	44,00%	14,66%	33,33%	20,37%	44%	15,66%	33,33%	21,34%
70:30:00	56,00%	18,66%	33,33%	23,94%	52%	18,66%	31%	23,29%
60:40:00	60,00%	20,00%	33,33%	25%	60%	20%	33,33%	25%

Based on table 19, it can be seen that Anies Baswedan's X data, which is labeled with lexicon-based, has the highest accuracy of 60% with the Naïve Bayes method and the Support Vector Machine method.

Table 20. Ganjar Pranowo's Instagram Data with Lexicon Based

Ratio	Naïve Bayes				Support Vector Machine			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
80:20:00	60,00%	41,66%	57,33%	48,36%	70%	78,33%	68,33%	73,05%
70:30:00	53,00%	38,33%	53,33%	44,62%	57%	72,33%	56,66%	63,57%
60:40:00	53,00%	39,33%	50,00%	44,00%	53%	39,33%	50%	44%

Based on the table above, it can be seen that Anies Baswedan's X data, which is labeled with lexicon-based, has the highest accuracy of 54% with the Naïve Bayes method and 46% with the Support Vector Machine method.

From the results of the above research on data labeled with lexicon-based, the following conclusions can be drawn.

a. Effect of data division
   As in the previous approach, there are variations in the accuracy of the model depending on the split ratio of training and testing data. In general, the more data allocated for training, the model performance tends to improve.
b. Model performance
   In lexicon-based labeling, using both Naive Bayes and Support Vector Machine, it is seen that the accuracy is relatively low compared to the previous approach using manual labeling. This suggests that classification relying on lexical analysis may not be robust enough to handle the complexity of Instagram X and comment data.
c. Performance comparison between models
   There is variation in performance between Naive Bayes and SVM depending on the type of data and data sharing scenario. However, neither model consistently excels in all conditions. For example, in X data, Support Vector Machine has the highest accuracy for Anies Baswedan data, while Naive Bayes has the highest accuracy for Ganjar Pranowo data at a 60:40 ratio.
d. Differences between X data and Instagram comments
   There are differences in model accuracy between X data and Instagram comments, which may be due to differences in structure, language style, and content type between the two.

In conclusion, lexicon-based labeling may be less effective in handling the complexity and variation in X data and Instagram comments, especially if the lexicon used is not robust enough. Other approaches may be needed to improve text classification performance.

4. CONCLUSION

Based on the results of the research that has been conducted, several important conclusions are obtained regarding the implementation of the Naïve Bayes and Support Vector Machine (SVM) methods in analyzing public sentiment towards 2024 Presidential Candidates on social media X and Instagram. The implementation of Naïve Bayes method with manual labeling shows the highest accuracy of 86% for X data and 85% for Instagram comment data. Meanwhile, with lexicon-based labeling, the highest accuracy obtained was 60% for both X data and Instagram comments.This

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

variation in accuracy appears to depend on the type of data and the data sharing scenario used, where the highest accuracy with manual labeling was recorded at 86% for X data at a ratio of 60:40 (Ganjar Pranowo) and 85% for Instagram comment data at a ratio of 80:20 (Anies Baswedan). On the other hand, lexicon-based labeling gave the highest accuracy of 60% for X data at a ratio of 60:40 (Ganjar Pranowo) and for Instagram comment data at a ratio of 80:20 (Ganjar Pranowo) and 70:30 (Prabowo Subianto). The implementation of the SVM method also produced similar accuracy to Naïve Bayes, which was 86% for X data and 85% for Instagram comment data with manual labeling, while with lexicon-based labeling, the highest accuracy achieved was 60% for X data and 70% for Instagram comment data. The accuracy obtained from the SVM method also shows variations depending on the type of data and the data sharing scenario, with the highest results in manual labeling of 86% for X data at a ratio of 60:40 (Ganjar Pranowo) and 85% for Instagram comment data at a ratio of 80:20 (Anies Baswedan). Meanwhile, with lexicon-based labeling, the highest accuracy was recorded at 60% for X data at a ratio of 60:40 (Ganjar Pranowo) and 70% for Instagram comment data at a ratio of 80:20 (Ganjar Pranowo).

REFERENCES

[1] M. D. J. Wardana et al., "Pengaruh Sejarah Pemilihan Umum Terhadap Sistem Ketatanegaraan Dan Hubungannya Dengan Demokrasi," Jurnal Ilmu Hukum dan Tata Negara, vol. 2, no. 2, pp. 84–93, 2024, doi: 10.55606/birokrasi.v2i2.1168.

[2] M. Yusuf, M. Jannah, N. Rahmi, P. A. Dewi, and U. F. Husaini, "Menggali Makna Pemilihan Umum : Peran, Sejarah, dan Tantangan Demokrasi," Jurnal Puan Indonesia, vol. 5, no. 2, pp. 656–667, 2024, doi: 10.37296/jpi.v5i2.231.

[3] R. S. Agustin, "INSTAGRAM SEBAGAI MEDIA KOMUNIKASI E-GOVERNMENT DALAM PELAYANAN PUBLIK DAN PEMBANGUNAN PEMERINTAH KOTA MADIUN," Retorika: Jurnal Komunikasi, Sosial dan Ilmu Politik, vol. 1, no. 1, pp. 90–97, 2024, Accessed: Aug. 09, 2024. [Online]. Available: https://jurnal.kolibi.org/index.php/retorika/article/view/1249

[4] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," SMATIKA Jurnal: STIKI Informatika Jurnal, vol. 10, no. 02, pp. 71–76, 2020, doi: 10.32664/smatika.v10i02.455.

[5] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI," Jurnal Teknoinfo, vol. 14, no. 2, pp. 116–124, Jul. 2020, doi: 10.33365/jti.v14i2.679.

[6] A. P. Nardilasari, A. L. Hananto, S. S. Hilabi, T. Tukino, and B. Priyatna, "Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM Pada Media Sosial Twitter," JOINTECS: Journal of Information Technology anda Computer Science), vol. 8, no. 1, pp. 11–18, 2023, doi: 10.31328/jointecs. v8i1.4265.

[7] M. N. Zarti, E. Sahputra, A. Sonita, and Y. Apridiansyah, "Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes Untuk Memprediksi Partisipasi Minat Mayarakat Pada Pemilu 2024," Jurnal Komputer, Informasi dan Teknologi, vol. 3, no. 1, pp. 105–114, 2023, doi: 10.53697/jkomitek.v3i1.

[8] D. S. Nugroho, I. F. Hanif, M. A. Hasbi, F. Fredianto, A. M. Saputra, and R. Zildjian, "Analisis Sentimen Dugaan Pelanggaran Pemilu 2024 Berdasarkan Tweet Menggunakan Algoritma Naïve Bayes Classifier," MALCOM: Indonesian Journal of Machine Learning anda Computer Science, vol. 4, no. 3, pp. 1169–1176, 2024, doi: 10.57152/malcom.v4i3.1496.

[9] D. Faroek, M. Yusuf, and G. Syatauw, "Sentiment Analysis of the Popularity of Parties Supporting the 2024 Presidential Candidates on Twitter Using the Naive Bayes Classifier Algorithm," Antivirus : Jurnal Ilmiah Teknik Informatika, vol. 17, no. 2, pp. 216–227, Nov. 2023, doi: 10.35457/antivirus.v17i2.3261.

[10] B. M. Iqbal, K. M. Lhaksmana, and E. B. Setiawan, "2024 Presidential Election Sentiment Analysis in News Media Using Support Vector Machine," Journal of Computer System and Informatics (JoSYC), vol. 4, no. 2, pp. 397–404, Feb. 2023, doi: 10.47065/josyc.v4i2.3051.

[11] F. Nurrizky and S. Dwiasnati, "Comparison of Naive Bayes and Support Vector Machine (SVM) Algorithms Regarding The Popularity of Presidential Candidates In The Upcoming 2024 Presidential Election," Computer Engineering and Applications (ComEngApp) Journal, vol. 13, no. 1, pp. 17–28, 2024, doi: 10.18495/comengapp.v13i1.459.

[12] O. H. Rahman, G. Abdillah, and A. Komarudin, "Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 5, no. 1, pp. 17–23, Feb. 2021, doi: 10.29207/resti.v5i1.2700.

[13] A. R. Abdillah and F. N. Hasan, "Analisis Sentimen Terhadap Kandidat Calon Presiden Berdasarkan Tweets Di Sosial Media Menggunakan Naive Bayes Classifier," SMATIKA JURNAL: STIKI Informatika Jurnal, vol. 13, no. 01, pp. 117–130, Jul. 2023, doi: 10.32664/smatika.v13i01.750.

[14] Y. Findawati, U. Indahyanti, Y. Rahmawati, and R. Puspitasari, "Sentiment Analysis of Potential Presidential Candidates 2024: A Twitter-Based Study," Academia Open, vol. 8, no. 1, Aug. 2023, doi: 10.21070/acopen.8.2023.7138.

[15] S. L. Ramadhan and W. Windarto, "Penerapan Naive Bayes untuk Menganalisis Sentimen Pengguna Twitter Terhadap Penetapan Calon Presiden 2024 PDIP," Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI), vol. 2, no. 2, pp. 716–725, 2023, [Online]. Available: https://senafti.budiluhur.ac.id/index.php/senafti/article/view/846

[16] A. A. Firdaus, A. Yudhana, and I. Riadi, "Analisis Sentimen Pada Proyeksi Pemilihan Presiden 2024 Menggunakan Metode Support Vector Machine," Decode: Jurnal Pendidikan Teknologi Informasi, vol. 3, no. 2, pp. 236–245, Jun. 2023, doi: 10.51454/decode.v3i2.172.

[17] F. A. Ramadhan and P. H. Gunawan, "Sentiment Analysis of 2024 Presidential Candidates in Indonesia: Statistical Descriptive and Logistic Regression Approach," 2023 International Conference on Data Science and Its Applications (ICoDSA), pp. 327–332, 2023, doi: 10.1109/ICoDSA58501.2023.10276417.

[18] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 5, no. 2, p. 406, Apr. 2021, doi: 10.30865/mib.v5i2.2835.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 950−960
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5766

[19] A. Mustofa and R. Novita, "Klasifikasi Sentimen Masyarakat Terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Text Mining Pada Twitter," Building of Informatics, Technology and Science (BITS), vol. 4, no. 1, pp. 200–208, Jun. 2022, doi: 10.47065/bits.v4i1.1628.
[20] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," J-SAKTI (Jurnal Sains Komputer dan Informatika), vol. 5, no. 2, pp. 697–711, 2021, doi: 10.30645/j-sakti.v5i2.369.
[21] S. Juniarsih, E. F. Ripanti, and E. E. Pratama, "Implementasi Naive Bayes Classifier pada Opinion Mining Berdasarkan Tweets Masyarakat Terkait Kinerja Presiden dalam Aspek Ekonomi," JUSTIN (Jurnal Sistem dan Teknologi Informasi), vol. 8, no. 3, pp. 239–249, 2020, doi: 10.26418/justin.v8i3.39118.
[22] B. Kurniawan, A. A. Aldino, and A. R. Isnain, "Sentimen Analisis Terhadap Kebijakan Penyelenggara Sistem Elektronik (Pse) Menggunakan Algoritma Bidirectional Encoder Representations From Transformers (Bert)," J. Teknol. dan Sist. Inf, vol. 3, no. 4, pp. 98–106, 2022, doi: 10.33365/jtsi.v3i4.2204.
[23] H. Syah and A. Witanti, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm)," Jurnal Sistem Informasi Dan Informatika (Simika), vol. 5, no. 1, pp. 59–67, 2022, doi: 10.47080/simika.v5i1.1411.
[24] N. Q. Rizkina and F. N. Hasan, "Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naive Bayes," Journal of Information System Research (JOSH), vol. 4, no. 4, pp. 1136–1144, Jul. 2023, doi: 10.47065/josh.v4i4.3803.
[25] F. Taufiqurrahman, S. Al Faraby, and M. D. Purbolaksono, "Klasifikasi teks multi label pada hadis terjemahan bahasa indonesia menggunakan chi-square dan svm," eProceedings of Engineering, vol. 8, no. 5, 2021.
[26] E. Hasibuan and E. A. Heriyanto, "Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping Di Google Play Store Menggunakan Naive Bayes Classifier," Jurnal Teknik dan Science, vol. 1, no. 3, pp. 13–24, 2022, doi: 10.56127/jts.v1i3.434.
[27] A. Y. Simanjuntak, I. S. S. Simatupang, and A. Anita, "Implementasi Data Mining Menggunakan Metode Naïve Bayes Classifier untuk Data Kenaikan Pangkat Dinas Ketenagakerjaan Kota Medan," Journal of Science and Social Research, vol. 5, no. 1, pp. 85–91, 2022, doi: 10.54314/jssr.v5i1.804.
[28] J. Sihombing, "Klasifikasi Data Antroprometri Individu Menggunakan Algoritma Naïve Bayes Classifier," BIOS: Jurnal Teknologi Informasi dan Rekayasa Komputer, vol. 2, no. 1, pp. 1–10, 2021, doi: 10.37148/bios.v2i1.15.
[29] A. C. Khotimah and E. Utami, "COMPARISON NAÏVE BAYES CLASSIFIER, K-NEAREST NEIGHBOR AND SUPPORT VECTOR MACHINE IN THE CLASSIFICATION OF INDIVIDUAL ON TWITTER ACCOUNT," Jurnal Teknik Informatika (JUTIF), vol. 3, no. 3, pp. 673–680, 2022, doi: 10.20884/1.jutif.2022.3.3.254.
[30] F. Amirullah, S. Alam, and M. I. Sulistyo S, "Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naïve Bayes," Jurnal Ilmiah Teknik dan Ilmu Komputer, vol. 2, no. 3, pp. 69–76, 2023, doi: 10.55123/storage.v2i3.2293.
[31] D. P. Ray, F. N. Hasan, and A. R. Dzikrillah, "Analisis Sentimen Terhadap KPU 2024 Berdasarkan Tweet Media Sosial Twitter Menggunakan Algoritma Naïve Bayes," KLIK: Kajian Ilmiah Informatika dan Komputer, vol. 4, no. 4, pp. 2235–2243, 2024, doi: 10.30865/klik.v4i4.1587.
[32] G. Gunawan and Y. Reswan, "DESAIN APLIKASI PENGENALAN POLA TANDA TANGAN MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)," Jurnal Media Infotama, vol. 17, no. 1, pp. 8–12, 2021, doi: 10.37676/jmi.v17i1.1311.
[33] R. Pramestiawan, "KOMPARASI ALGORITMA SUPORT VECTOR MACHINE DAN NAIVE BAYES DALAM ANALISIS SENTIMEN KEBIJAKAN PPKM," Jurnal Sistem Informasi dan Manajemen Basis Data, vol. 5, no. 2, pp. 23–32, 2022, doi: 10.30873/simada.v5i2.3427.
[34] A. I. Nurhidayat, A. Asmunin, and D. Fatrianto, "Prediksi Kinerja Akademik Mahasiswa Menggunakan Machine Learning dengan Sequential Minimal Optimization untuk Pengelola Program Studi," Journal of Information Engineering and Educational Technology, vol. 5, no. 2, 2021, doi: 10.26740/jieet.v5n2.p84-91.
[35] F. Listanto, M. Fatchan, and W. Hadikristanto, "Prediction of Casting Product Defects Using SVM Algorithm Based on RBF and Linear," Jurnal Ilmiah Intech : Information Technology Journal of UMUS, vol. 5, no. 2, pp. 109–119, 2023, doi: 10.46772/intech.v5i2.1376.
[36] A. R. Damanik, S. Annisa, A. I. Rafeli, A. S. Liana, and D. S. Prasvita, "Klasifikasi Jenis Buah Cherry Menggunakan Support Vector Machine (SVM) Berdasarkan Tekstur dan Warna Citra," Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia, pp. 302–311, 2022, [Online]. Available: https://www.kaggle.com/moltean/fruits
[37] M. N. Winnarto, M. Mailasari, and A. Purnamawati, "KLASIFIKASI JENIS TUMOR OTAK MENGGUNAKAN ARSITEKTURE MOBILENET V2," Jurnal SIMETRIS, vol. 13, no. 2, pp. 1–12, 2022.
[38] M. Sholawati, K. Auliasari, and F. X. Ariwibisono, "PENGEMBANGAN APLIKASI PENGENALAN BAHASA ISYARAT ABJAD SIBI MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN)," Jurnal Mahasiswa Teknik Informatika, vol. 6, no. 1, pp. 134–144, 2022, doi: 10.36040/jati.v6i1.4507.