



Sentiment Analysis Netizens on Social Media Twitter Against Indonesian Presidential Candidates in 2024 Using Naive Bayes Classifier Algorithm

Ismail*, Mario Adi Nugroho, Rini Sulistiyowati

School of Computing, Informatics Study Program, INABA University, Jakarta, Indonesia
Email: ^{1,*}ismail@inaba.ac.id, ²Marioan749@student.inaba.ac.id, ³rini.sulistiyowati@inaba.ac.id
Correspondence Author Email: ismail@inaba.ac.id

Abstract-The democratic system is highly respected in Indonesia. The Indonesian state government organizes its government in a democratic manner, namely, it is run by, for, and with the consent of the people. Democracy is held through general elections to occupy leadership and power seats, followed by political parties. The high enthusiasm of Twitter social media users with the electability of presidential candidates proposed by political parties in the 2024 general election and the existence of a survey of potential presidents from national institutions, as well as mass and social media coverage of presidential candidates suggested by the political parties, generates opinions, attitudes, and emotions among people from all walks of life through tweets. The availability of abundant data on Twitter and other social media can provide useful information. Tweet data is obtained by crawling data using the help of the Python library, namely sncrape. The sentiment analysis uses a mixed method, namely by using machine learning and Lexicon Based, through the process of fine-grained sentiment analysis using the technique, namely, knowing the level of opinion polarity by grouping netizen responses and opinions into three parts: positive, neutral, and negative, with the help of machine learning and natural language processing. The results of the study were carried out by experimenting with four scenarios by dividing test and training data by 60:40, 70:30, 80:20, and 90:10. Measurement of the accuracy value results in the classification of the Naive Bayes Classifier Algorithm as 68%, 67%, 70%, and 71%. From the tweet data, it is clear that positive sentiment is dominant on the research topic.

Keywords: Democracy; Presidential Candidates; Sentiment Analysis; Machine Learning; Naive Bayes Classifier

1. INTRODUCTION

The system of democracy is strong in Indonesia. The government of Indonesia is organized only democratically, i.e., it is run by, for, and with the consent of the people. Democracy is established through general elections in order to occupy the seat of leadership and power, in which the political parties exist as participants.

In the Indonesian context according to the 1945 Constitution, the President is the head of government and head of state. The president as the head of government has the authority to appoint and dismiss ministers and form a cabinet [1]. Presidential candidate pairs are mostly political groups that are favored by several people. Most pairs of presidential candidates have favorable ratings that far exceed [2]

Each political party proposing a pair of presidential candidates and a vice president proposed by a political party or political party association through an election must obtain at least 20% of the seats in the People's Chamber of Representatives or at least 25% of the national votes in general elections based on the Constitution.

The arrangements related to the Presidential Threshold that have been previously determined are further elaboration in article 6A paragraph (2) of the Constitution which states that "Pairs of presidential candidates and vice-presidential candidates are proposed by a political party or a coalition of political parties. parties participating in general elections prior to holding general elections[3]. The high enthusiasm of the public with the presence of the eligibility of the presidential candidate proposed by the political parties in the general election of 2024 and the existence of presidential polls from the National Institution and the publication of mass and social media to raise the opinions, attitudes, and emotions of people from various backgrounds through tweets on social media.

Sentiment analysis is a component of text mining that serves as a method for extracting opinions from any document or sentence. Currently, sentimental analysis using Indonesian is quite common among academics and professionals in the industrial sector. This can be seen from several papers obtained and business actors conducting sentiment analysis for commercial purposes. Some examples of sentient analysis that have been carried out include forecasting the results of presidential or governor elections[4]

Twitter is the social media platform that has the largest number of users compared to some other social networking platforms, with a total of 326 million users [5]. Social Media Twitter are highly studied to understand behavior among users, groups, organizations or even people and especially to understand people's opinions on various topics such as politics[5].

Text mining is a technique for processing text-based data, including unstructured data, by recognizing words or concepts from the entire text[4]text mining has grown into a significant and preferred study topic. The text mining approach is divided into two categories: lexical-related decoding and machine learning [4]Twitter is a common source of text mining data. This is because Twitter searches are not only based on account names, but also based on certain keywords, hashtags, status dates [6]

As in the sentiment analysis research conducted by the First Researcher, sentiment analysis with airline opinion objects on Twitter conducted by Pravina in 2019. The research aims to classify public comments into



positive or negative classes, in order to help the company and the public. using the Support Vector Machine method, this research has 40% accuracy, 40% precision, 100% recall, and 57.14% f-measure [7]

In the second study, the author classifies Tweets into three classes, namely positive, negative, and neutral. This classification uses the Naive Bayes algorithm. Classification can make it easy for users to see positive, negative and neutral opinions. The results of the accuracy test on the naive Bayes method for classification are 69.97% [8]

The third study used the classification method using the Naïve Bayes Classifier method. The evaluation results obtained are the highest accuracy score of 94.38%, precision of 59.67%, recall of 67.93%, and F-measure of 62.32%. It can be concluded that the results of the sentiment analysis on rising prices of basic commodities have received a negative response from the public[9].

In the fourth study, namely conducting the Naive Bayes classification, it was found that the 2019 election tweet dataset had a pattern of negative perceptions of 52%, far greater than positive perceptions, namely 18% and neutral perceptions, which had a value of 31% higher than positive perceptions. The accuracy of the Naive Bayes classification for the training dataset is 81% and the testing dataset is 76%, the average precision value for positive sentiment is 86.65%, negative sentiment is 77.15%, and neutral sentiment is 80.95% while the average value recall on positive sentiment 36.8%, negative sentiment 93.2% and neutral sentiment 86.8%[10]

Based on the background above, and based on previous research, the problem formulation is how to get a sentiment classification model using the Naïve Bayes Classifier. Based on the problems experienced in previous journals, this study aims to analyze Twitter social media from the perspective of Twitter users about the Indonesian Presidential Candidates for the 2024 period to find out people's opinions, attitudes, and emotions or often called Sentiment Analysis (SA). with the help of a text mining approach, namely machine learning and lexicon based.

2. RESEARCH METHODOLOGY

The sentiment analysis uses a mixed method, namely by using machine learning and Lexicon Based, through the process of fine-grained sentiment analysis using the technique, namely, knowing the level of opinion polarity by grouping netizen responses and opinions into three parts: positive, neutral, and negative, with the help of machine learning and natural language processing., starting with the data crawling stage to collect the required data. Next, do the text preprocessing stage to clean and prepare the text data. Furthermore, the classification process was carried out using the Naive Bayes classification algorithm with the addition of tests and validation tests were carried out using the K-fold Cross Validation. K-Fold Cross Validation is used to test the performance of an algorithm or model. and the appropriate evaluation metrics are used to evaluate the results of the classification model of the algorithm used, namely Naïve Bayes. The word visualization stage is also carried out to analyze and display the words found in the data. The methods used in this study include:

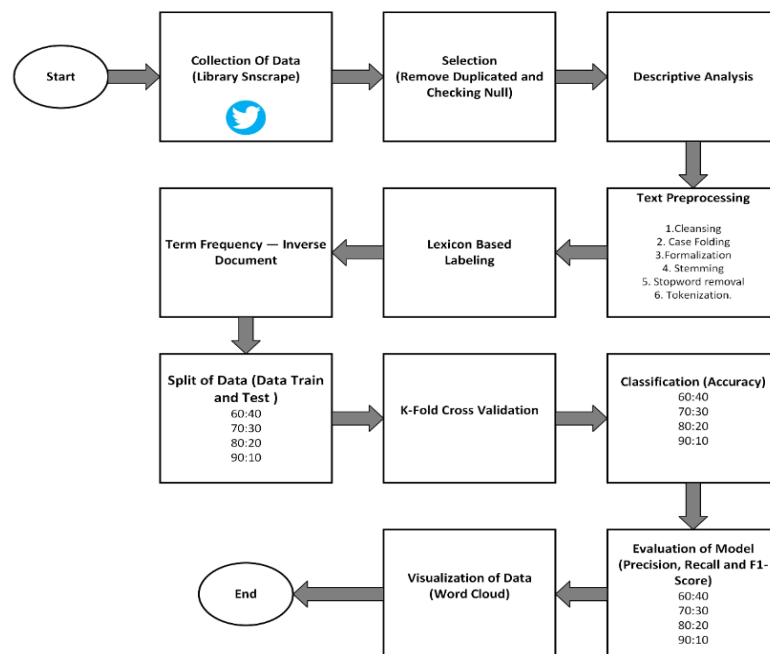


Figure 1. Research Methodology

2.1 Colletion of Data

Data crawling is the process of filtering data from Twitter, including public opinion consisting of positive, non-appropriate, and negative comments, for use in research using the Twitter app. The process begins by visiting the



surveyed Twitter account, then going to the clone folder and continuing to filter data using the Python library by calling an API [8]. Twitter data collection using the Python library is called Snsrape. Data collection from August 1, 2022, to November 1, 2022 The keyword is a key word in a tweet related to the opinions of netizens about the 2024 presidential candidate.

2.2 Dataset

The data set used for this study was 2000 tweets. A keyword on a tweet related to the opinions of netizens about the 2024 presidential candidate. The data is then stored in the comma-separated values file format. (CSV).

Table 1. Dataset

No	Username	Tweet	Date
1	jo_anezt	Pengamat Politik Ray Rangkuti mengatakan Menteri Badan Usaha Milik Negara (BUMN) Erick Thohir laik menjadi bakal calon wakil presiden (Cawapres) 2024. #BangkitBersamaET https://t.co/aoOEeEuahL	2022-10-31
2	zubelsitorus5	PUAN Maharani, calon paling tepat untuk menjadi Presiden RI tahun 2024-2029 🇮🇩	2022-10-31
3	zuhri_saif	@PDI_PerjuanganSaya Mendukung Pak Ganjar Pranowo sebagai Calon Presiden 2024	2022-10-31

2.3 Selection

Lexicon Based has two dictionaries that are used as word lists. A dictionary with a collection of words related to positive emotions and a dictionary with a collection of words related to negative emotions [11]. When the sorting process is performed on one part of the data, the total weight of the data produces positive, negative or neutral values. After sorting the data, it will display the weight of the data[11]. At the data selection stage, namely by viewing and selecting redundant or duplicate tweet data to be deleted and carried out, viewing null data and selecting data that has been labeled with a Lexicon Based dictionary.

2.4 Descriptive Analysis

This study aims to increase general understanding of tweet data related to the 2024 Indonesian presidential candidate. At this stage by providing a general picture in the form of a pie to see negative, neutral and positive data.

2.5 Text Preprocessing

Text preprocessing is an important stage in data mining that aims to clean data from noise or unwanted interference. Data that is cleaned is incomplete data, data that is noisy, and data that is inconsistent [12] On the research to do preprocessing with data in the form of text that will be made into a sentiment analysis by performing several stages of text mining, as follows Cleansing, Case Folding, Formalization, Stemming, Stopword removal and Tokenization. Natural Language Toolkit atau disingkat NLTK, adalah library Python untuk bekerja dengan pemodelan teks. NLTK menyediakan tools yang baik untuk mempersiapkan teks sebelum digunakan untuk analisis lebih lanjut seperti pembuatan model machine learning dan visualisasi [13]

2.6 Lexicon Based Labelling

At the labeling stage, using a lexicon-based method. the Lexicon Based method as a labeling or weighting medium in order to maximize the results obtained. In its application, Calculations on the Lexicon method can be seen in equation 1.

$$Sentence_{Sentiment} \begin{cases} Negative & \text{if } S_{Positive} < S_{Negative} \\ Neutral & S_{Positive} = S_{Negative} \\ Positive & \text{if } S_{Positive} > S_{Negative} \end{cases} \quad (1)$$

After going through the weighting stage it will produce data in the form of positive and negative sentiments provided that if the polarity value is greater than or equal to 0.1 then it will be labeled positive sentiment and if the polarity value is less than or equal to 0.0 it will be labeled negative sentiment. The results of labeling with the Lexicon method will be used as supporting data for classification with Naïve Bayes [14]

2.7 Term Frequency — Inverse Document

Term Frequency- Inverse Document Frequency (TF-IDF) is used to get the weight value of each word in the data used. The TF-IDF values are obtained after the TF and IDF values are calculated first. Then the two values are multiplied to get the TF-IDF value [15]. At this stage, to perform feature extraction using TF-IDF, which aims to improve algorithm performance in building models, we need to calculate the weight or critical value of each word in the data. This is done with the aim of giving more significant value to words that have a greater contribution to



modeling. The word weighing process with TF-IDF uses machine learning with Python script libraries, namely TfidfTransformer and Count Vectorizer.

a. Term Frequency

The term TF is used to determine the frequency of the appearance of words in documents. The TF value can mean the number of words or the frequency of words in the document.

$$Tf_{ij} = \frac{fa(i)}{\text{Max } fa(j)} \quad (2)$$

Fa(i) = The frequency of the word on i that appears in the document j

Max fa(j) = Total of a term in a document j

b. Inverse Document Frequency

IDF is used to determine the frequency of a term in all documents. IDF measures to overcome these terms, with terms that rarely appear in several documents that have a high Inverse Document Frequency.

$$idf(t, d) = \log\left(\frac{N}{df(t)+1}\right) \quad (3)$$

N = Total number of documents

df(t)+1 = The number of documents containing the word t and the value to avoid division against the value 0 if not found on the corpus

2.8 Split of Data (Data Train and Test)

To model and evaluate the performance of the Naive Bayes Classifier algorithm, the data division phase divides training data (train data) and test data (test data). This is very important to ensure that the model that is developed can classify data that has never been seen previously with accuracy.

Table 2. Split of Data

No	Data Train	Data Test
1	60% Data Train	40% Data Test
2	70% Data Train	30% Data Test
3	80% Data Train	20% Data Test
4	90% Data Train	10% Data Test

2.9 K-Fold Cross Validation

K-Fold cross-validation is a highly efficient statistical technique for evaluating the effectiveness of models or algorithms. This method combines two sets of data parts that are used for repetitive learning and evaluation. The benefits of K-Fold Cross Validation are exploited for repeated processes with accurate sub-sampling so that any information is used for the evaluation and validation stages, in addition to any information marked for validation [16] K-fold Cross Validation is the single most effective statistical method for evaluating model or algorithm performance [17]

At this stage, testing and validation are carried out using the cross-fold K-fold method. K-Fold cross-validation is used to test the performance of algorithms or models. This method separates data into two groups, namely learning groups and evaluation groups. One of the main advantages of this technique is that sub-sampling is done randomly and repeatedly, which means that any observation is used for the training and validation processes. We can ensure that the models developed have consistent and reliable performance by using K-Fold cross-validation.

2.10 Classification

The Naïve Bayes classifier is an algorithm that takes part of statistics, namely theory possibility or often referred to as Probability and is one of the data mining algorithm methods classification to solve the data set case on the class or table as the target reference you want is known[18]

In this study, the Naive Bayes Classifier was recommended as a method of classification. To be able to measure and see the presentation of the Naive Bayes Classifier algorithm is the purpose of this classification. Classification is done with test data, which is used to study training data to perform classification testing on test data. In this process, the model or algorithm is studied with the training data and then applied to the test data that has been divided through the data division process. Therefore, test data is used to test the performance of models that have been learned from training data, and separation can be used to evaluate classification results with the help of machine learning.

The Nave Bayes classifier is a method that calculates probability and statistics, classifies data in the form of text or news, and can also be used to make diagnoses in medical science. This method can classify data in large quantities and process results with high accuracy [16] Naïve Bayes for each decision class calculates the probability on the condition that the decision class is true, given the information object vector. This algorithm



assumes that object attributes are independent. The probabilities involved in producing the final estimate are calculated as the sum of the frequencies of the “master” decision table [19]

The Nave Bayes Classifier, also known as the Multinomial Naval Bayes, is a simplified Bayes Method model. This paradigm is suitable for categorizing text and documents. Here is the simplified Bayesian method of the sales model:

$$VMAP = \arg \max P(V_j | a_1, a_2, \dots, a_n) \tag{4}$$

Identifying

a_i = referring to the attribute or feature ke-i

v_j = Refer to Class J.(positif atau negatif)

V = A class of target.

VMAP = It’s a sentiment from a commentary.

2.11 Evaluation of Model

Evaluating Machine Learning models can be very challenging. Typically, data sets are separated into training and test sets. The model is then trained and tested using a training set and a test set, respectively [7] The study aims to evaluate the performance of the model using evaluation methods such as the confusion matrix and the accuracy, precision, recall, and f1-score tables. The evaluation is carried out with the aim of observing the level of accuracy of the model produced by the method. The evaluation method uses the confusion matrix and such tables to evaluate a variety of performance metrics, including accuracy, precision, recall, and f1-Score, which are used to analyze and compare the performance of the model tested.

2.12 Visualization of Data

A word cloud is a visual representation of text data, typically used to describe keyword metadata on websites, or to visualize freeform text. Single words are used and the resulting cloud provides a pictorial representation of the word based on the frequency of word usage in that domain and in principle, the size of a word in a word cloud is determined by the frequency of occurrence [20] In this study, the word cloud is used as a visualization tool to describe free text in word form. A word cloud is used to calculate the frequency of words in tweet data collected from netizens on social media about Indonesian presidential candidates in 2024. Using the word cloud, this research aims to look at the topics or words that most appear in the tweet data.

3. RESULT AND DISCUSSION

In this research, word clouds are used as a tool for visualizing text in the form of kata-kata. Word clouds are used to calculate frequencies of kata in data tweets that dikumpulkan from netizens on social media about the Candidate President of Indonesia in 2024. By using a word cloud, research is meant to look at the topic that most often appears in the data tweet.

3.1 Collection of Data

The researchers used the crawling method of tweet data for data collection using the Python snsrape library and Pyhton programming language. Data is stored in the Comma Separated Values format (CSV). Data crawling was carried out over a 4-month period, starting on August 1, 2022, and ending on November 1, 2022. Data collected as much as 2000 Data from the results of the crawling is processed.

```
import snsrape.modules.twitter as sntwitter
import pandas as pd
# Membuat daftar untuk menambahkan data tweet ke
tweets_list = []
# Menggunakan TwitterSearchScrapper untuk sortir data dan menambahkan tweet ke array
for i,tweet in enumerate(sntwitter.TwitterSearchScrapper('Calon Presiden 2024 & Calon Presiden since:2022-
08-01 until:2022-11-01').get_tweets()):
    if i==1000:
        break
    tweets_list.append([tweet.date, tweet.id, tweet.content, tweet.user.username,
analyze_sentiment(tweet.content)])
# Membuat kerangka data dari daftar tweet di atas
tweet = pd.DataFrame(tweets_list, columns=['Datetime', 'Tweet Id', 'Text', 'Username', 'Label'])
#
tweet
#
tweet.groupby('Label').describe()
```

Figure 2. Collection Of Data

3.2 Dataset

The data set used for this study was 2000 tweets. A keyword on a tweet related to the opinions of netizens about the 2024 presidential candidate. The dataset was taken through the crawling data technique was carried out over a 4-month period, starting on August 1, 2022, and ending on November 1, 2022. The data is then stored in the comma-separated values file format. (CSV).



Table 3. Dataset Example

No	Username	Tweet
1	jo_anezt	Pengamat Politik Ray Rangkuti mengatakan Menteri Badan Usaha Milik Negara (BUMN) Erick Thohir laik menjadi bakal calon wakil presiden (Cawapres) 2024. #BangkitBersamaET https://t.co/aoOEeEuahL
2	zubelsitorus5	PUAN Maharani, calon paling tepat untuk menjadi Presiden RI tahun 2024-2029 🇮🇩
3	zuhri_saif	@PDI_PerjuanganSaya Mendukung Pak Ganjar Pranowo sebagai Calon Presiden 2024

3.3 Selection

the next stage is Data Selection, which is by seeing and selecting redundant or duplicate tweet data to be deleted, and the next step is to do labeling. The data selection process is done manually by sorting the duplicated or redundant data of 651 data tweets so that the data remains 1349.

3.4 Descriptive Analysis

A descriptive analysis showed that netizens on social media sites like Twitter expressed a more positive opinion of the tweet data taken from crawling. i.e., for data tweets that have a positive sentiment, 1045 Tweet data Whereas for data tweets, which have a negative sentiment of 242 And Neutral sentiment total 62, disclosed by netizens through social media, Here is the graphic diagram.

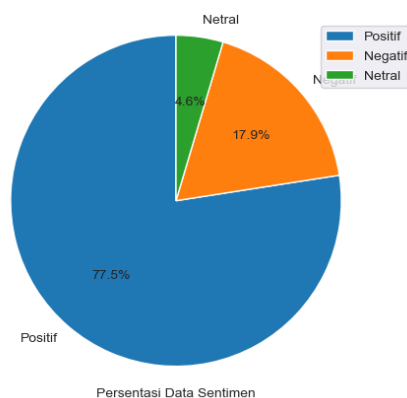


Figure 3. Graphic Pie Diagram Descriptive Analysis

3.5 Teks Preprocessing

The data that is obtained from crawling is a CSV file that contains unstructured text. In order for the data to be structured, text preprocessing is required. The text preprocessing phase includes text data cleaning. There are several steps in this process, among them:

3.5.1 Cleansing

Drawing down in text preprocessing aims to remove unnecessary words from the document.

Table 4. Step of Cleansing

Before	After
@KompasData Mantap kang @ridwankamil sing naek terus meh jadi calon Presiden ke 2024 Siapapun Calon PRESIDEN 2024.Saya setuju. Sebanyak banyaknya.	kompasData Mantap kang ridwankamil sing naek terus meh jadi calon Presiden ke Siapapun Calon PRESIDEN setuju Sebanyak banyaknya

3.5.2 Case Folding

The case folding stage in text preprocessing turns the entire capital letter into small letters, so the words become uniform in writing.

Table 5. Step of Case Folding

Before	After
kompasData Mantap kang ridwankamil sing naek terus meh jadi calon Presiden ke Siapapun Calon PRESIDEN 2024.Saya setuju. Sebanyak banyaknya.	kompasdata mantap kang ridwankamil sing naek terus meh jadi calon presiden ke Siapapun Calon PRESIDEN setuju Sebanyak banyaknya



3.5.3 Tokenization

The tokenization phase in text preprocessing is done to break down a tweet data document into smaller pieces, such as paragraphs or sentences, thus facilitating further analysis and processing.

Table 6. Step of Tokenization

Before	After
kompasdata mantap kang ridwankamil	'kompasdata', 'mantap', 'kang',
sing naek terus meh jadi calon presiden	'ridwankamil', 'sing', 'naek', 'terus',
ke	'meh', 'jadi', 'calon', 'presiden', 'ke'
siapapun calon presiden saya setuju	'siapapun', 'calon', 'presiden', 'saya',
sebanyak banyaknya	'setuju', 'sebanyak', 'banyaknya'

3.5.4 Stopword Removal

At this stage, in order to maintain classification performance, words that rarely appear are filtered. By deleting words that rarely appear, sparse data can be reduced, and feature space can be significantly reduced. This makes it possible to optimize the performance of classification.

Table 7. Step of Stopword Remoyal

Before	After
'kompasdata', 'mantap', 'kang',	'kompasdata', 'mantap',
'ridwankamil', 'sing', 'naek', 'terus',	'kang', 'ridwankamil', 'yang', 'naik', 'terus',
'meh', 'jadi', 'calon', 'presiden', 'ke'	'hampir', 'jadi', 'calon', 'presiden', 'saja'
'siapapun', 'calon', 'presiden', 'saya',	'siapapun', 'calon', 'presiden', 'saya',
'setuju', 'sebanyak', 'banyaknya'	'setuju', 'sebanyak', 'banyaknya'

3.5.5 Stemming

In this phase, a basic word or root word search is performed for each word after the filtration process. This stage involves cutting off the end of the word and removing the compensation on the words that have been filtered, thus resulting in a simpler base word.

Table 8. Step of Stemming

Before	After
'kompasdata', 'mantap',	'kompasdata', 'mantap', 'kang',
'kang', 'ridwankamil', 'yang', 'naik', 'terus',	'ridwankamil', 'yang', 'hampir', 'calon',
'hampir', 'jadi', 'calon', 'presiden', 'saja'	'presiden'
'siapapun', 'calon', 'presiden', 'saya',	['calon', 'presiden', 'tuju', 'banyak']
'setuju', 'sebanyak', 'banyaknya'	

3.6 Lexicon Based Labeling

At the Data Labeling Stage, using the technique of Lexicon Based Labeling, namely labeling with 3 categories, namely positive, negative, and neutral. Where each data label is given labeling based on weight based on the lexicon based dictionary.

Table 9. Step of Lexicon Based Labelling

Tweet	Polarity	Sentiment
[mantap, kang, sing, calon, presiden]	13	Positive
[cebong, panas, wajar, saksi, calon, presiden, iri]	-2	Negative
[calon, pimpin, barakallah, anis]	0	Neutral

3.6 Term Frequency — Inverse Document

At this stage do plotting the words that often appear on the data tweet netizen on the social media twitter on plotted the data-data tweet-tweet used using pandas to manage the data in the form of tables and matplotlib to describe the visualization of the data obtained with the form table distribution, from the results plott the word tweet get words that frequently appear.

Table 10. Frequently Occurring Words

No	Frequently occurring words
1	presiden
2	calon
3	pilih



No	Frequently occurring words
4	partai
5	indonesia
6	ganjar
7	anies
8	wakil
9	dukung
10	pranowo

In this phase, the Term Frequency (TF) method is used to determine how often words appear in a document. TF refers to the number of occurrences or frequencies of words in the document. In this phase, CountVectorizer, a library in scikit-learn using the Python programming language, is used to calculate the occurrence of words. CountVectorizer allows automatic calculation of word frequencies in vector form, which can be used in the process of text analysis and modeling.

Table 11. Word Occurrence

Frequently occurring words	Word Occurrence
presiden	2423
calon	1871
pilih	662
partai	596
indonesia	390
ganjar	383
anies	350
wakil	318
dukung	256
pranowo	246

In this phase, the Inverse Document Frequency (IDF) method is used to calculate how common or rare a word appears in the entire document. IDF measures the quantity of a term that rarely appears throughout the document and affects its weight. High IDF values indicate that the term rarely appears in other documents. At this stage, it used TfidfTransformer, a library in Scikit-Learn using the Python programming language, to calculate the weight of words based on IDF.

Table 12. Word Weighting

Frequently occurring words	Word Weighting
presiden	0.110459
calon	0.090011
pilih	0.027845
partai	0.023055
ganjar	0.019616
anies	0.017926
wakil	0.014146
indonesia	0.014002
dukung	0.013111
pranowo	0.012599

3.7 Split of Data (Data Train and Test)

At this stage, previously processed initial data will be divided into two parts: test data and training data. Training data is used to train classification models, which will study patterns and attributes of training data. With more and more data being trained, classification algorithms become better at understanding data and finding patterns. At the same time, test data is used to evaluate the performance of classification algorithms. Test data is data that is not used in the training process, so it can demonstrate the ability of the classification algorithm to classify data that has never been seen before.

Table 13. Data Train and Test

No	Split of Data	Data Train	Data Test	Total of data
1	60:40	809	540	1349
2	70:30	944	405	1349
3	80:20	1079	270	1349
4	90:10	1214	135	1349



3.8 K-Fold Cross Validation

In this test, 525 tweet data, including test data, were divided into 5 sections with k=5 according to a cross-K-fold validation scheme. Test data will be given to each component in sequence in five test iterations. For example, the test data is in the initial position in the first iteration, then in the second position on the second, and so on. To obtain more accurate results, tests using already partitioned data will be repeated as many as five times.

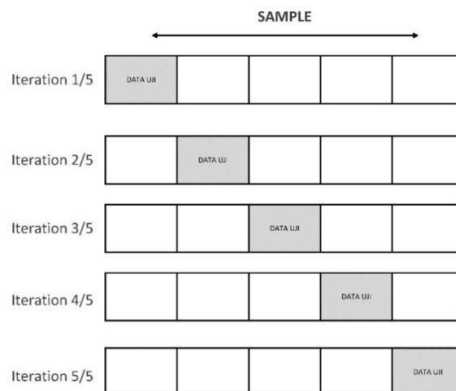


Figure 4. Cross-K-fold Validation

3.9 Classification

In this study, classification testing was performed using the Naive Bayes Classifier Algorithm used with Python libraries using Scikit Learn. The purpose of this test is to measure and see the presentation of the Naive Bayes Classifiers algorithm. By using test data and learning from training data, classification is done. The result was that the accuracy of sentiment classification on tweet data became higher.

Table 14. Naive Bayes Classifier Classification

No	Scenario	Methods	K-Fold Validation Accuracy					Accuracy Without K-Fold Validation
			1	2	3	4	5	
1	(60:40)	Naive Bayes Classifier	86,7%	88,8%	87,8%	90,4%	88,3%	68%
2	(70:30)	Naive Bayes Classifier	88,8%	88,2%	90%	87%	89,3%	67%
3	(80:20)	Naive Bayes Classifier	86,5%	89,5%	89,5%	87,9%	88,5%	70%
4	(90:10)	Naive Bayes Classifier	84,3%	91,0%	91,3%	87,4%	86,7%	71%

3.10 Evaluation of Model

To find out the model performance results from the Naive Bayes Classifier algorithm, the model evaluation calculation is carried out after completion of the classification process. In addition, model evaluation has the ability to calculate the values of accuracy, recall, precision, and F-Measure values on the confusion matrix table.

Table 15. Actual Confusion Matrix

The Scenario	Actual	Predicted		
		Positive	Negative	Neutral
(60:40)	Positive	302	67	51
	Negative	25	25	19
	Neutral	5	7	10
(70:30)	Positive	217	56	39
	Negative	16	41	15
	Neutral	5	4	11
(80:20)	Positive	302	72	50
	Negative	22	56	22
	Neutral	5	8	9
(80:20)	Positive	152	34	19
	Negative	11	25	10
	Neutral	4	4	10

After obtaining the Value of the actual Confusion matrix from 4 scenarios, then confusion measures the accuracy values of the classification to calculate the model evaluation values and evaluate the results of predictions from classification methods such as Recall, Precision and F-Measure with machine learning as follows:



Table 16. Positive Sentiment Assessment Model

Skenario	(60:40)	(70:30)	(80:20)	(90:10)
Precision	0.91	0.92	0.91	0.91
Recall	0.72	0.70	0.74	0.79
F-1 Score	0.80	0.79	0.82	0.84

Table 17. Negatif Sentiment Assessment Model

Skenario	(60:40)	(70:30)	(80:20)	(90:10)
Precision	0.43	0.41	0.40	0.35
Recall	0.56	0.57	0.54	0.41
F-1 Score	0.49	0.48	0.35	0.38

Tabel 18. Neutral Sentiment Assessment Model

Skenario	(60:40)	(70:30)	(80:20)	(90:10)
Precision	0.12	0.18	0.26	0.30
Recall	0.45	0.60	0.56	0.60
F-1 Score	0.20	0.28	0.35	0.40

3.11 Visualization of Data

By gathering information from the tweet data most often discussed or rated by netizens on social media Twitter about the 2024 presidential candidate, we can get information that is considered important from many previous tweet texts. This research will use word cloud to visualize the results of classification analysis.

3.11.1 Data of Sentimen Positive

Word cloud is used to visualize positive sentiment tweet data. Using library python, the word cloud can be used to identify positive sentiment tweets based on the frequency of the words in the tweets. Based on Word Cloud obtained information that in the class of positive sentiment the words that most often appear are the words Candidate, President, Ganjar pranowo, puan maharani, partai demokrasi, Indonesia, juang, calon wakil, etc.

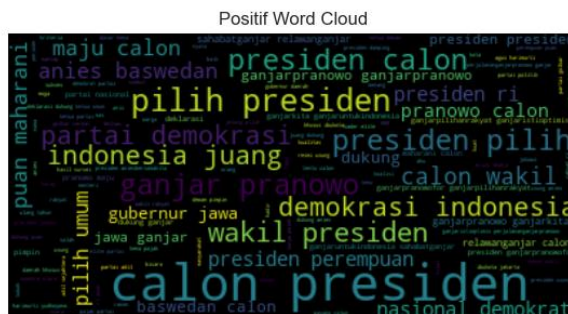


Figure 5. Data of Sentimen Positive

3.11.2 Data of Sentimen Negative

Negative sentiment tweet data is visualized using word cloud. Based on the word cloud obtained information that in the class of negative sentiment the words that most often appear are the words Calon President, nasdem, cebong, pilpres, Demokrat, anies baswedan, nasional, etc.



Figure 6. Data of Sentimen Negative

3.11.3 Data of Sentimen Neutral

Negative sentiment tweet data is visualized using word cloud. Based on the word cloud obtained information that



- [10] S. Juanita, "Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes," *J. MEDIA Inform. BUDIDARMA*, vol. 4, no. 3, p. 552, Jul. 2020, doi: 10.30865/mib.v4i3.2140.
- [11] P. Amira Sumitro *et al.*, "Analisis Sentimen Terhadap Vaksin Covid-19 di Indonesia pada Twitter Menggunakan Metode Lexicon Based," *J. Inform. Dan Teknol. Komput.*, vol. 2, no. 2, pp. 50–56, 2021, doi: 10.33059/j-icom.v2i2.4009.
- [12] S. A. Dainamang, N. Hayatin, and D. R. Chandranegara, "Analisis Sentimen Media Sosial Twitter terhadap RUU Omnibus Law dengan Metode Naive Bayes dan Particle Swarm Optimization," *Komputika J. Sist. Komput.*, vol. 11, no. 2, pp. 211–218, Aug. 2022, doi: 10.34010/komputika.v11i2.6037.
- [13] T. M. Fahrudin *et al.*, "Analisis Speech-to-Text pada Video Mengandung Kata Kasar dan Ujaran Kebencian dalam Ceramah Agama Islam Menggunakan Interpretasi Audiens dan Visualisasi Word Cloud," *SKANIKA Sist. Komput. dan Tek. Inform.*, vol. 5, no. 2, pp. 190–202, 2022.
- [14] A. Hadiy, D. Fatra, N. Hayatin, C. Sri, and K. Aditya, "Analisa Sentimen Tweet Berbahasa Indonesia Dengan Menggunakan Metode Lexicon Pada Topik Perpindahan Ibu Kota Indonesia," *J. Repos*, vol. 2, no. 7, pp. 977–984, 2020, doi: 10.22219/repositor.v2i7.937.
- [15] M. Wongkar and A. Angdresey, "Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler : Twitter," *IEEE*, pp. 1–5, 2020, doi: <https://doi.org/10.1109/ICIC47613.2019.8985884>.
- [16] W. Yulita *et al.*, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naive Bayes Classifier," *JDMSI*, vol. 2, no. 2, pp. 1–9, 2021, doi: <https://doi.org/10.33365/jdmsi.v2i2.1344>.
- [17] D. W. Seno and A. Wibowo, "Analisis Sentimen Data Twitter Tentang Pasangan Capres-Cawapres Pemilu 2019 Dengan Metode Lexicon Based Dan Support Vector Machine," *J. Ilm. FIFO*, vol. 11, no. 2, p. 144, Nov. 2019, doi: 10.22441/fifo.2019.v11i2.004.
- [18] H. Harmayani and L. Sitorus, "Diagnosa Penyakit Ginjal Kronis Menggunakan Metode Klasifikasi Naive," *J. MEDIA Inform. BUDIDARMA*, vol. 4, no. 3, p. 850, Jul. 2020, doi: 10.30865/mib.v4i3.2292.
- [19] E. F. Saraswita, D. P. Rini, and A. Abdiansah, "Analisis Sentimen E-Wallet di Twitter Menggunakan Support Vector Machine dan Recursive Feature Elimination," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 4, p. 1195, Oct. 2021, doi: 10.30865/mib.v5i4.3118.
- [20] Suryasatria Trihandaru, Parhusip Hanna Arini, Susanto Bambang, and Putri Carolina Febe Ronicha, "Word Cloud of UKSW Lecturer Research Competence Based on Google Scholar Data," *J. Ilmu Komput. Dan Inform.*, vol. 7, no. 2, pp. 1–7, 2021, doi: <https://doi.org/10.23917/khif.v7i2.13123>.