International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I2P125
Volume 72 Issue 2, 242-253, February 2024

Original Article

Perception of the Indonesian Society on the Performance of the Ministry of Investment/BKPM Based on Sentiment Analysis

Moehamad Taufik Hidayat[1], Tuga Mauritsius[2]

[1,2]*Information System Management Department, BINUS Graduate Program – Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia.*

[1]*Corresponding Author: moehamad.hidayat@binus.ac.id*

Received: 17 August 2023 Revised: 18 November 2023 Accepted: 06 January 2024 Published: 03 February 2024

Abstract *- Social media supports various performance or service dissemination activities by ministries and organizations. Social media such as Twitter can be used as material to get an overview from Indonesian netizens regarding perceptions of the Ministry of Investment/BKPM. Sentiment analysis is done by collecting data on the Ministry of Investment/BKPM. The results of this study are helpful as evaluation material and input for the performance of the Ministry of Investment/BKPM. The method developed in this study resulted in a percentage of sentiment analysis using Naïve Bayes of 82% with a ruled-based analysis approach. System evaluation obtained the best results from testing the highest test data when testing k-fold with 80% accuracy, 92.3% recall, 85.7% precision, and 88% f-measure.*

Keywords *- Sentiment analysis, Rule-Based, k-fold, Naïve Bayes, Investment.*

1. Introduction

Investment realization in Indonesia since mid-2020 has grown 2% y-o-y to 826.3 trillion rupiah, even though Indonesia was being hit by the Covid-19 pandemic. Meanwhile, in 2021, the growth in investment realization is 9% year-to-year to 901 trillion rupiah. This growth is also accompanied by investment equity between Java Island and Outside Java Island. Consecutively, from 2019 to 2021, the investment portion outside Java will be 46%, 51%, and 52%. Indicators of healthy investment realization can also be seen from the positive growth of the manufacturing sector.

From 2019 to 2021, the manufacturing portion is 7.6%, 11.5%, and 13%. The more rapidly existing technology develops, the more responses or comments are submitted by the public to these agencies. We can see comments and reactions on various social media platforms, for example, Twitter (Currently X). Twitter is a social media networking site that is currently proliferating. With 280 characters to write responses on Twitter, customers or the public can write their responses to a particular Ministry of Investment/BKPM. Comments often written to the Ministry of Investment/BKPM are regarding services at each existing Ministry of Investment/BKPM. With this in mind, it can help the Indonesian people know and adapt various kinds of assessments of different service forms from the existing Ministry of Investment/BKPM.

The process is carried out to assess the public's opinion on the tweets issued by BKPM or those related to the mention of the Ministry of Investment/BKPM. These comments were collected within a certain period to be used as data to be labeled with a technological approach to provide positive, neutral, and negative sentiment values. This research focuses on labeling automatically, followed by providing naive Bayes classification to get the appropriate analyst sentiment results.

2. Related Works

Opinion Mining / Sentiment Analysis is something the same term [4] which is a research branch in the Text Mining domain. According to Bing Liu in his book entitled "Sentiment Analysis and Opinion Mining," opinion mining is a field of study that analyzes a person's opinions, sentiments, evaluations, attitudes, and emotions from written language [6]. Expression or sentiment refers to the focus of a particular topic; a statement on one issue may have a different meaning from the same information on a different subject. Therefore, in several studies, especially in product reviews, work is preceded by determining the product elements being discussed before starting the opinion mining process [6]. Sentiment analysis consists of 3 major sub-processes. We can make each of these sub-processes separate research topics/materials because each of these sub-processes requires techniques that are not easy to implement:

2.1. Subjectivity Classification

Determine which sentence is an opinion
Example:

- A bicycle has 2 wheels (not an opinion)
- That's a great bike! (opinion)

2.2. Orientation Detection

Classifying opinion basically determines whether the opinion is positive or negative, or neutral.

Example:

- That's a great bike! (positive)
- Ah, that's a lousy bike! (negative)

2.3. Opinion Holder and Target Detection

Determine which part is the opinion holder and which part is the target. Example: Harry said it was a nice bike. The bicycle in the example above is the target of an opinion.

To support this research, a literature study was conducted to get an overview of sentiment analysis; here are some related studies: Mustofa et al. [1] analyzed the Indonesian people's response to AFTA from social media. This study uses a dictionary of word degrees and design rules to determine opinion sentiment in the opinion mining section. Making rules using impression techniques, where this impression technique is more inclined to analyze the word order in one sentence. The impression technique, which has 9 rules, analyzes the location of adjectives, verbs, and prepositions in a sentence.

Jie Li, and Lirong Qiu [2] showed a way to reduce problems with sentiment structure and calculation rules. The sentiment structure is obtained from parsing dependencies with the migration of modified relations and distances, which is suitable for understanding the sentiments of short texts. Short text sentiment is accumulated according to the influence of the relationship between the modifier and the word emotion and the contribution of each sentence to the short text sentiment calculation. Suhariyanto et al. [3] proposed a method for predicting film sentiment on The Rotten Tomatoes website by combining sentiment scores from SentiWordnet and original scores from experts. The experimental results show that the proposed method provides a better F measure compared to other methods with a value of 0.97.Pamungkas and Putri [4] use SentiWordnet to classify the play store review dataset. They use first sense to select a synset from Wordnet. The results of their experiments were around 67.93%.

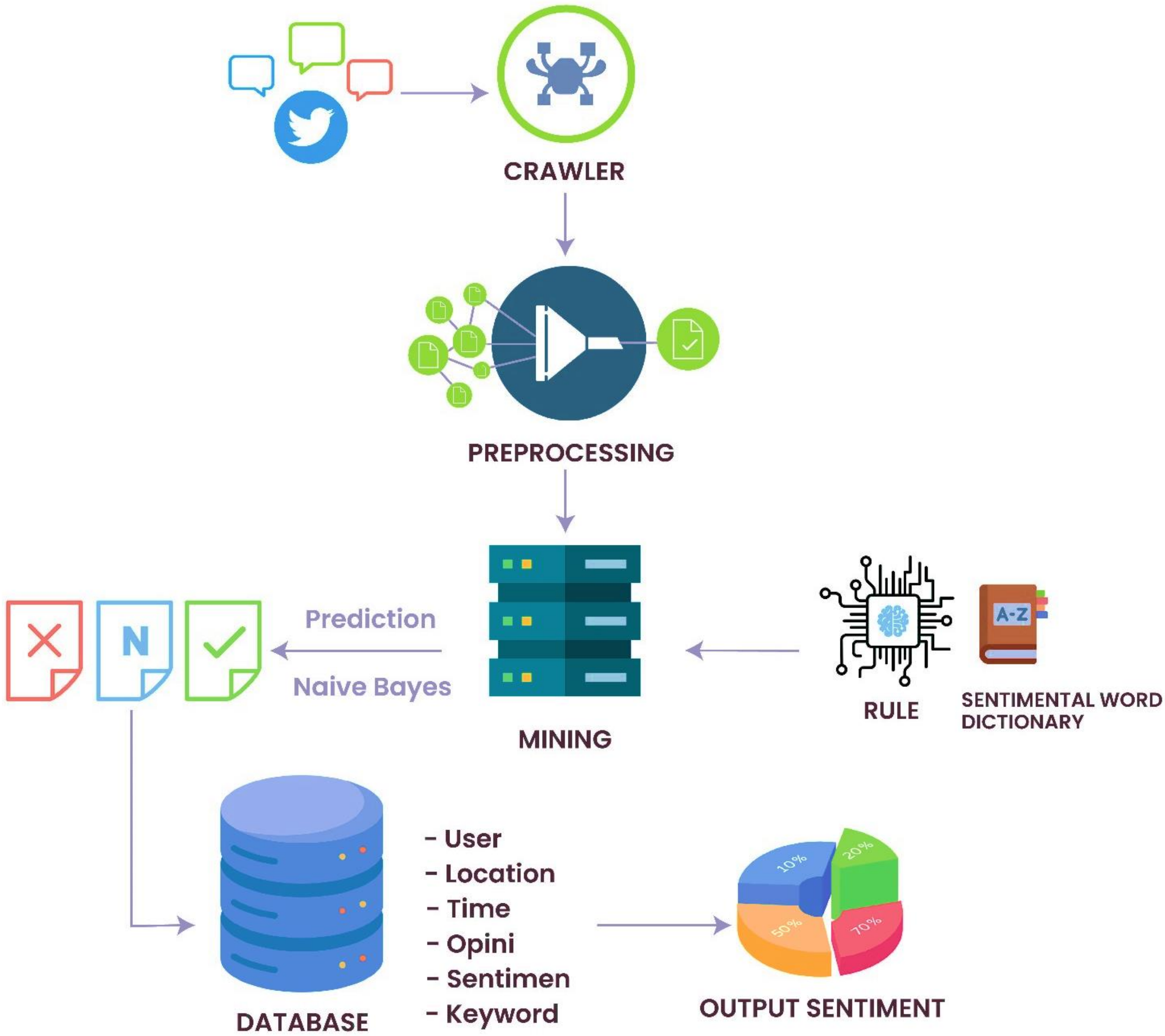

Fig. 1 System diagram design

The method proposed by Sheeba Naz et al. [6] relates to classifying Twitter sentiments using a machine learning classification model that utilizes different textual features, viz. n-gram Twitter data. In addition, they used three different weighting schemes to understand the impact of weighting on classifier accuracy. Furthermore, the tweet sentiment score vector is used to provide external knowledge to improve the performance of the SVM classifier. Asmara and Rengga [11] also used the impression method to analyze culinary reviews in Surabaya. Results from this research are positive, negative, and neutral in every sentence. Based on the comparison of previous research, the researcher combines the word degree dictionary approach and design rule in determining opinion sentiment added by Naive Bayes for predictions.

3. Methodology

This chapter describes the research design, which is the steps of research work, starting from the beginning of the research until the final objectives are obtained. This study uses a rule-based approach for calculating sentiment and a correlation measurement algorithm for categorization. For the development of information systems using PHP technology and the Laravel framework. For databases, use MySQL and visualization using highchart tools.

3.1. System Design

The following is a research system design that is being carried out. The following system design describes the data collection process up to the sentiment analysis process and the classification of each service.

The first process is data collection, done by crawling Twitter data, followed by preprocessing, which aims to clean the crawled data. Then, proceed with the process of sentiment analysis using a rule base. After getting sentiment from each opinion, it will go into the categorization process to get the category of each statement so that the results will be visualized using the web.

3.2. Data Collection

Data is obtained by crawling queries by entering the Indonesian Ministry of Investment/BKPM account name. We use the GetOldTweets-python library created by Jefferson-Henrique for crawling data on Twitter. This library allows us to retrieve old tweet data from more than one week ago.

Because Twitter has a policy on its API, it only provides old tweet data with a maximum range of 1 week ago. The data used is tweet data from January 2023 to July 2023. This study carried crawling data based on queries, such as popular hashtags related to investment policies or tweets mentioning official accounts from the Ministry of Investment/BKPM. The collected data is stored in a database that contains the detailed attributes needed to obtain complete information.

3.3. Preprocessing

After the opinion data from Twitter is obtained, a preprocessing process will be carried out before the opinion mining process. This preprocessing stage aims to remove noise, uniform word forms, and reduce vocabulary volume. The preprocessing steps in this study will be explained using one of the opinions from Twitter.

3.3.1. Delete URLs

Because there are times when the URL, in my opinion, has a problem, namely having space, this will interfere with the regular expression process to delete the URL; therefore, before deleting the URL, the process of removing space in the URL is done first.

Remove URLs

@anen36 @yunartowijaya @CCICPolri Faktanya: "Menteri Investasi/Kepala BKPM Bahlil Lahadalia mengungkapkan, akses ekonomi di Indonesia dikuasai oleh kelompok yang jumlahnya tidak lebih dari 1%." https://t.co/5f30lyKdZO”

3.3.2. Removed Twitter Special Characters

This process is done by removing special Twitter characters such as hashtags (#hashtag), usernames (@username), and special characters (e.g., RT, which indicates that the user is retweeting).

Delete @anen36 @yunartowijaya @CCICPolri

”Faktanya: "Menteri Investasi/Kepala BKPM Bahlil Lahadalia mengungkapkan, akses ekonomi di Indonesia dikuasai oleh kelompok yang jumlahnya tidak lebih dari 1%."

3.3.3. Clear Numbers

According to Aqsath and Ayu (2011), the appearance of numbers at the front and the end of words has an insignificant effect on the sentiment value of a tweet.

3.3.4. Remove Punctuation Symbols

This step is carried out to remove symbols and punctuation in tweets, except for commas (,) and periods (.) because commas and periods are used in the opinion mining rule.

3.3.5. Case Folding (lower case)

The case folding process standardizes the shape of letters into uppercase or lowercase letters; in this study, all letters were converted to lowercase.

After lower case:

“ faktanya menteri investasi kepala bkpm bahlil lahadalia mengungkapkan, akses ekonomi di indonesia dikuasai oleh kelompok yang jumlahnya tidak lebih dari 1”

3.3.6. Tokenize

Tokenize or tokenization is the stage of cutting the input string based on the word that composes it. Usually, the process of cutting text is done based on the space in each word, but the tokenization process in this study is a bit unique; that is, not every word space will be cut, but it will be checked first whether there are words consisting of several words. For example, in this sentence :

“faktanya menteri investasi kepala bkpm bahlil lahadalia mengungkapkan, akses ekonomi di indonesia dikuasai oleh kelompok yang jumlahnya tidak lebih dari 1”

"tidak lebih dari" is a single word, namely a word with one meaning consisting of several words; if it is cut based on spaces, it will have a different purpose. The writer knows this fundamental thing. Therefore, the writer creates a function to check in advance whether there are words that consist of several words. Word data containing several words is taken from the word sentiment dictionary database.

3.3.7. Stopword Removal

Stopwords are common words that often appear, so they cannot represent the topic in a sentence, such as personal pronouns and adverbs. Such comments should be removed to reduce noise.

3.3.8. Normalization

In Twitter data, many netizens use non-standard words such as alay, slang, and abbreviations. Some of these words have important meanings, which, if they are not converted into standard words, they will become meaningless words and will not be processed by the system. Therefore, checking non-standard words and replacing them with standard terms is necessary.

3.3.9. Post-Tagging

Part of Speech Tagging (POS Tagging) is labelling each word in a sentence with a POS or tag corresponding to the class of words such as verbs, adverbs, adjectives, and others. The problem faced in the POS Tagging process is finding ambiguous words or words with two or more meanings in a dialogue, making it challenging to find the correct tag or class of terms for a word in a sentence.

In this study, post-tagging was carried out by checking each token with a word dictionary that already exists in the database and then labeling it.

3.4. Analysis Sentiment

The next stage is the opinion-mining process. Opinion mining is the main stage of this research. In NLP (Natural Language Processing), Opinion mining is often called Sentiment Analysis. Sentiment analysis is the development of text mining to understand, extract, and process textual data automatically to obtain sentiment information in an opinion sentence.In this study, sentiment analysis refers to classifying opinions into 3 classes: positive, negative, and neutral. To carry out sentiment analysis, it is necessary to have word sentiment dictionary data and rules. The word sentiment dictionary is used as a reference for word sentiment, and the rule is used as a sentiment calculation technique.

3.4.1. Word Sentiment Dictionary

The word sentiment dictionary is used to assign a value to each word. This value is 1, -1, and 0, where 1 is positive, -1 is negative, and 0 is neutral. In addition to sentiment values, the word sentiment dictionary also stores word types to make it easier for the system to make opinion sentiment assessment rules later. In making a word sentiment dictionary, the initial data came from Mustofa's research in 2015; from Mustofa's research in 2015, the author managed to get 263 words. Then, the author adds his own in line with ongoing research. To add data to the word sentiment dictionary, author created an unknown word system design to assist in adding data to the word sentiment dictionary. The method of this system serves to check whether a comment already exists in the word sentiment dictionary database or not; if not, then the word will be stored in the database.

←T→	code	name	description
Edit Copy Delete	1	Nomina	Kata benda
Edit Copy Delete	2	Verba	Kata kerja active or basic
Edit Copy Delete	3	Verba-di	Kata kerja passive
Edit Copy Delete	4	Adjektiva	Kata sifat / keadaan terhadap nomina
Edit Copy Delete	5	Preposisi	Kata depan, biasanya diikuti nomina / pronomina
Edit Copy Delete	6	Konjungsi	Kata sambung
Edit Copy Delete	7	Interjeksi	Kata pengungkap ekspresi perasaan
Edit Copy Delete	8	Numeralia	Kata yang mengatakan jumlah
Edit Copy Delete	9	Keyword	Just keyword
Edit Copy Delete	10	Adverbia	Kata keterangan untuk bukan kata benda
Edit Copy Delete	11	Pronomina	Kata ganti nomina
Edit Copy Delete	12	Lain-lain	just the others

Fig. 2 Types of indonesian words

No	Kata	Jenis	Nilai
1	difasilitasi	Verba-di	Positif
2	diundur	Verba-di	Negatif
3	lord	Interjeksi	Negatif
4	rangkaian	Verba	Positif
5	oss	Nomina	Netral
6	maharani	Nomina	Netral

Fig. 3 Example word dictionary

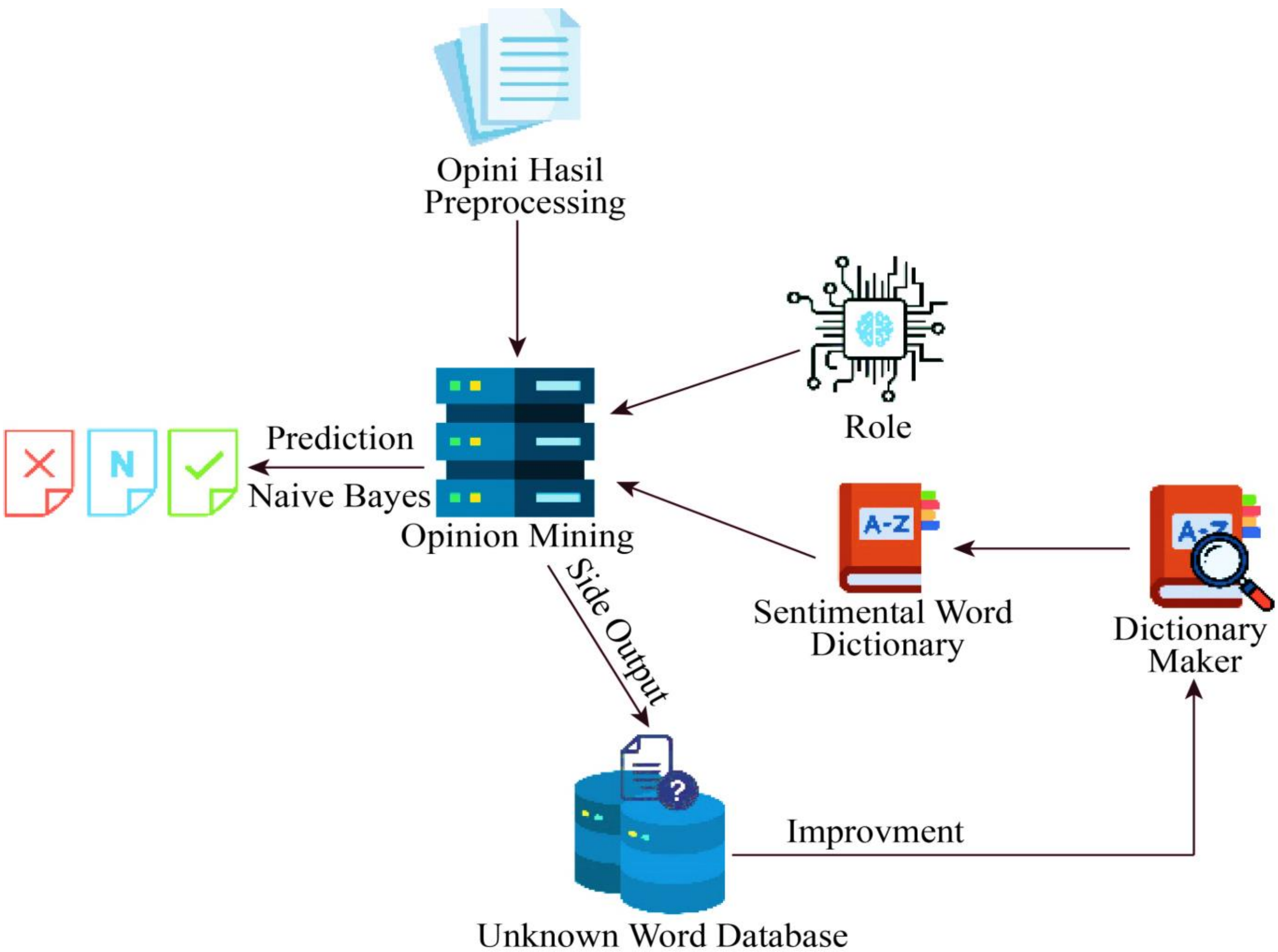

Fig. 4 Unknown word system design

Table 1. Active verb combination rule

No	Active Verb Word Combination	Examples
1	Verb	Saya Setuju
2	Verb + Noun	Mengembangkan Indonesia
3	Verb + Noun + Adj	Mengembangkan Indonesia Dengan Buruk
4	Verb + Adj	Mengembangkan Dengan Baik
5	Verb + Adj + Noun	Mengembangkan Dengan Baik Indonesia
6	Pre + Verb	Saya Sangat Melindungi
7	Pre + Verb + Noun	Tidak Memajukan Indonesia
8	Pre + Verb + Noun + Adj	Tidak Memajukan Indonesia Dengan Baik
9	Pre + Verb + Adj	Tidak Memajukan Dengan Baik
10	Pre + Verb + Adj + Noun	Tidak Memajukan Dengan Baik Indonesia
11	Noun + Verb	Indonesia Melindungi
12	Noun + Verb + Noun	Indonesia Menjaga Kedamaian
13	Noun + Verb + Noun + Adj	Indonesia Menjaga Kedamaian Dengan Baik
14	Noun + Verb + Adj	Indonesia Memajukan Dengan Baik
15	Noun + Verb + Adj + Noun	Indonesia Menjaga Dengan Baik Kedamaian
16	Noun + Pre + Verb	Indonesia Sangat Melindungi
17	Noun + Pre + Verb + Noun	Indonesia Tidak Menjaga Kedamaian
18	Noun + Pre + Verb + Noun + Adj	Indonesia Tidak Menjaga Kedamaian Dengan Baik
19	Noun + Pre + Verb + Adj	Indonesia Tidak Menjaga Dengan Baik
20	Noun + Pre + Verb + Adj + Noun	Indonesia Tidak Menjaga Dengan Baik Kedamaian

Table 2. Passive verb combination rule

No	Passive Verb Word Combination	Examples
1	Verb_di	Didukung
2	Verb_di + Noun	Didukung oleh Indonesia
3	Verb_di + Noun + Adj	Didukung oleh Indonesia dengan baik
4	Verb_di + Adj	Didukung dengan baik
5	Verb_di + Adj + Noun	Didukung dengan baik oleh Indonesia
6	Pre + Verb_di	Tidak didukung
7	Pre + Verb_di + Noun	Tidak didukung oleh Indonesia
8	Pre + Verb_di + Noun + Adj	Tidak didukung oleh Indonesia dengan baik
9	Pre + Verb_di + Adj	Tidak didukung dengan baik
10	Pre + Verb_di + Adj + Noun	Tidak didukung dengan baik oleh Indonesia
11	Noun + Verb_di	Indonesia dikembangkan
12	Noun + Verb_di + Noun	kedamaian dijaga oleh Indonesia
13	Noun + Verb_di + Noun + Adj	kedamaian dijaga oleh Indonesia dengan baik
14	Noun + Verb_di + Adj	kedamaian dijaga dengan baik
15	Noun + Verb_di + Adj + Noun	kedamaian dijaga dengan baik oleh Indonesia
16	Noun + Verb_di + Verb	Indonesia tidak didukung
17	Noun + Pre + Verb_di + Noun	kedamaian tidak dijaga oleh Indonesia
18	Noun + Pre + Verb_di + Noun + Adj	kedamaian tidak dijaga oleh Indonesia dengan baik
19	Noun + Pre + Verb_di + Adj	kedamaian tidak dijaga dengan baik
20	Noun + Pre + Verb_di + Adj + Noun	kedamaian tidak dijaga dengan baik oleh Indonesia

Table 3. Rule of adjective word combinations

No	Adjective Verb Word Combination	Examples
1	Adj	Baik
2	Adj + Verb	Dengan Baik Menumbuhkan
3	Adj + Verb + Noun	Dengan Baik Memajukan Indonesia
4	Pre + Adj	Tidak Baik
5	Pre + Adj + Verb	Belum Baik Memajukan
6	Pre + Adj + Verb + Noun	Tidak Dengan Baik Memajukan Indonesia
7	Noun + Adj + Verb	Indonesia Dengan Baik Memajukan
8	Noun + Adj + Verb + Noun	Indonesia Dengan Baik Menjaga Kedamaian
9	Noun + Pre + Adj	Indonesia Tidak Baik
10	Noun + Pre + Adj + Verb	Indonesia Tidak Dengan Baik Melindungi

In the opinion process, there is a word labeling process; placing the word checking process in the opinion mining process will shorten the process with the system not working twice because labeling and word checking access the same database, namely the word sentiment dictionary database. Words that are not in the phrase sentiment dictionary will be stored in the unknown word database, and then from the unknown word database, the process of labeling the unknown word will be carried out. There are two ways to label words, namely automatically and manually.

3.4.2. Automatic Unknown Word Labeling

Creating a word dictionary is automatically carried out by taking words in unknown words whose number of occurrences has exceeded a predetermined threshold. Then, do mining sentiment words. This process assigns a sentiment value and word type to each word. Word sentiment mining process uses English sentiwordnet. Because it is difficult to find Sentiwordnet in Indonesian, the author must first translate the words into English using API assistance from Google Translate to compare with Sentiwordnet in English at Sentiwordnet. is.cnr.it to determine the sentiment value and word type.

3.4.3. Unknown Word Manual Labeling

Another way that can be done is manually labeled by the author because not all words can be labeled by sentiwordnet.

In this manual process, the author carries out labeling subjectively because there is no open-source Indonesian word sentiment dictionary. Because the existence of a word dictionary is essential in this study for the opinion mining process, this method must be carried out to complete the word sentiment dictionary data. The number of words in the word sentiment dictionary that has been made is as many as ... words with ---- positive words, ---- negative words, and --- neutral words.

3.4.4. Word Combination Rules

After matching the word to the sentiment dictionary database and labeling the word, the rule design process begins. Rule is used to provide comment sentiment assessment rules. This process does not use a special algorithm but uses an impression technique. This technique is more straightforward than using an algorithm. The impression technique is more inclined to analyze the word order in a sentence [2].

This technique analyzes the location of adjectives, verbs, and prepositions in a sentence. Prepositions make up words or parts of sentences and are usually followed by a noun or pronoun, for example, not, yet, remarkably, etc.

There are 31 rules used in this study, consisting of 3 main combination rules where each combination of word types uses a different logic rule to get different results [5].

Table 3.1 is a list of rule bases where the main component of the sentence is an active verb. The rule of this word combination will be used to calculate the sentiment of an opinion. In the rules in Table 1, active verbs have an important role in determining the value of a sentiment, so what is checked first is whether the verb is active or not. If there is, the system will continue to check the combination of opinions according to the rules in this table so that the sentiment value of an opinion is obtained. If an opinion does not meet the combinations in Table 1, it will be checked on the combinations in other tables, namely Table 2 and Table 3. Opinions that contain passive verbs will use the combination rule in Table 2. If it does not meet, then it is possible to calculate the sentiment value using the rule combination in Table 3.

3.4.5. Sentiment Calculation

The sentiment calculation process is carried out by implementing the combination rule above, and some of the calculation determinations refer to Mustofa's research (2015) and with the help of truth tables for logical calculations.

Remember, the impression rule technique analyzes the location of adjectives and verbs; therefore, look for verb or adjective-type words sequentially in a sentence. After getting it, the calculation process starts for only five words: two words before the main dish (verb/adjective) and two words after it.

The following is an example of calculating sentiment on a sentence fragment consisting of 4 words with an active verb as the main word.

Examples of sentence fragments:

"*tidak memajukan* Indonesia dengan baik"

In this process, the word fragments are obtained along with the labels.

If you look closely at this wording, it forms the word combination rule no. 8 in the active verb category, namely:

Pre + Verb + Noun + Adj

We can see the negative result by carrying out the calculation process according to the rules.

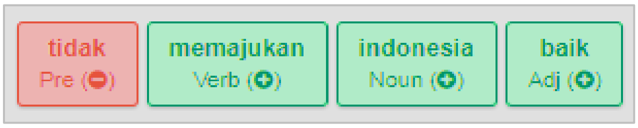

Fig. 5 Preprocessing results

Fig. 6 Final preprocessing results

Table 4. Truth table

p	q	~p	~q	p^q	p∨q	p→q	p↔q
B	B	S	S	B	B	B	B
B	S	S	B	S	B	S	S
S	B	B	S	S	B	B	S
S	S	B	B	S	S	B	B

3.4.6. Naive Bayes

In addition to using ruled-based impressions, this research implements the Naive Bayes Classifier method into a system for classifying data into positive and negative sentiments based on the opinion data that has been collected. Of course, the choice of Naive Bayes also has a particular reason apart from the previous research conducted, which was only at the ruled-based stage to analyze the sentiment analysis. To increase the method's accuracy, other methods can be added to support the classification process. Researchers try various possible references, for example, by selecting Naive Bayes based on the impressions produced.

Another approach was chosen to ensure the chosen method; the authors also compared the existing algorithms in the library tools issued by Python. A unique library that many researchers often use is the sklearn library, and based on references, it is also known as Naive Bayes or Naïve Bayes Multinominal, which is often the flagship for these libraries. Using the same training data will make this method a benchmark for why the author chose the Naive Bayes method as a reference. Broadly speaking, the NBC (Naïve Bayes Classifier) model is as follows:

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)} \qquad (1)$$

Where:

$P(X|Y)$ = Percentage of X in Y
$P(X \cap Y)$ = Specific word in class (tweet)
$P(Y)$ = Number of occurrences of a particular labeled word (Positive, Negative, Neutral)

4. Experiment and Analysis

In this chapter, we conduct experiments and analyze the system created. This aims to determine the extent to which the accuracy of the execution of the plan has been made. So from here, it can be concluded whether the software created can run correctly and according to the expected criteria.

4.1. Sentiment Classification Test

The following experiment is to test the opinion or sentiment analysis built by the system. The system will be tested by calculating sentiment values from some commentary data. Of course, this process is done after preprocessing. The following is the definition of calculation and the results of sentiment values that the system has processed.

4.1.1. Opinion 1

"Salah satu investasi mangkrak yang telah difasilitasi oleh #BKPM yaitu proyek pembangunan kompleks kilang minyak dan petrokimia di Desa Sumurgeneng, Kabupaten Tuban, Jawa Timur, yang merupakan kerja sama antara @pertamina dengan Rosneft dengan nilai investasi sebesar Rp211,9 T."

The system has created a view to show calculations in sentences according to the word combination rules that have been made. Following are the results of sentiment analysis in comment 1,In opinion 1, there is an Adv/Pre(-) + Noun(+) rule for the word "salah" investors, which results in negative sentiment. So the final sentiment value generated in opinion 1 is:

$$-1 + 0 = -1$$

The result is -1, so it can be concluded that the opinion is negative.

4.1.2. Opinion 2

"Mantap bang ilmunya? Konferensi Pers Menteri Investasi/Kepala BKPM https://t.co/guNI391STK"

Following are the results of sentiment analysis on the 2nd opinion. In opinion 2, there is a Verb(+) + Noun(o) rule in the words "mantap bang ilmunya" there is a Verb(+) + Noun(o) rule in the sentence fragment, which produces positive sentiment. So, the final sentiment value generated in Opinion 2 is

$$1 + 0 = -1$$

The result is -2, so the 2nd opinion sentiment is positive.

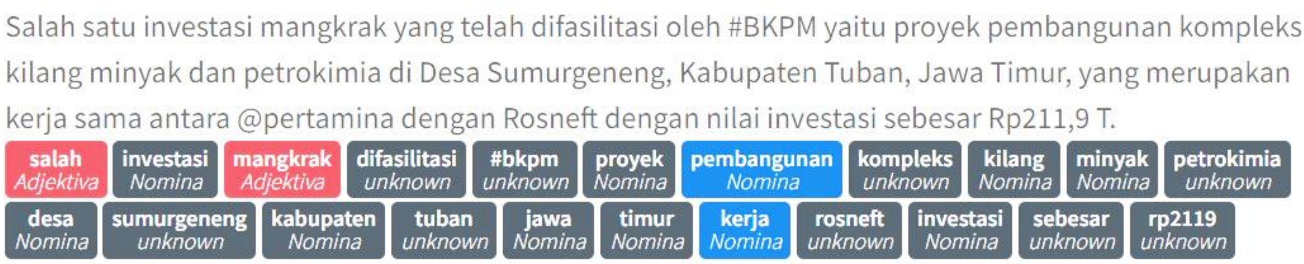

Fig. 7 Opinion calculation results 1

Fig. 8 Opinion calculation results 2

Table 5. Result from system

Topic	Opinion	Result			Accuracy %	
		Positive	Negative	Normal	Impresion	Naïve Bayes
Ministry of Investment/BKPM	88	18	22	38	88	88

Fig. 9 Opinion calculation results 3

Fig. 10 Opinion calculation results 4

4.1.3. Opinion 3

"Kan @Kemenperin_RI @bkpm @kemenkomarves @KementerianESDM @DivHumas_Polri Bisa tutup smelter tersebut yang berlaku culas dan @KemenkeuRI @beacukaiRI Bisa pecat oknumnya... Segitu mudahnya... Masa negara kalah sama MALING Cc @jokowi https://t.co/F2FGRguFW8"

Following are the results of sentiment analysis on the 3rd opinion. In opinion 3, there is a Verb(+) + Noun(-) rule in the sentence "kan tutup kalah," which produces a negative sentiment, then there is the Adj(-) combination rule in the sentence "Masa negara kalah" which produces negative sentiment. So the final sentiment value generated in opinion 3 is:

1 + - 1 + -1 = -1

The result is -1, so the 3rd opinion sentiment is negative.

4.1.4. Opinion 4

"Kepala BKPM Bahlil Lahadalia mengatakan pemerintah sudah melunasi seluruh utang Indonesia kepada IMF pada era Presiden SBY. Berapa banyak? https://t.co/lHEDam8iAQ " Following are the results of sentiment analysis on the 4th opinion, In opinion 4, there is a single (+) rule verb, namely "melunasi," so the temporary sentiment results are positive. There is a Verb(-) + Noun(o) rule in the sentence fragment "melunasi seluruh hutang," which generates positive sentiment. So the final sentiment value generated in opinion 3 is:

1 + 0 = 1

The result is 1, so the 4th opinion sentiment is positive.

4.2. Naive Bayes

After the preprocessing stage is complete, then the implementation of the Naive Bayes algorithm is carried out from the impression results that have been stored in the database. The training process is carried out to build a probability model from training data. The purpose of this test is to be able to classify tweets as positive or negative tweet sentiment automatically. In the Naive Bayes Classifier method, the process of categorizing text is carried out based on training data that has previously been stored. The result is shown in Table 5.

4.3. Performance Testing of Sentiment Classification

In this section, we will test the method used to calculate sentiment from tweet data, namely the rule-based approach. The process is measured using a confusion matrix, which produces accuracy, precision, and recall values with details as shown in Table 7. Calculations performed on 600 data that have been labeled manually:

- Accuracy is the comparison of correctly identified cases with the total number of issues. There are 3 cases, namely positive, negative, and neutral cases.

  The formula of accuracy = $\frac{TN+TP}{FN+FP+TN+TP}$ (2)

- Precision is the level of accuracy between the information requested by the user and the answers given by the system.

  Formula from precision = $\frac{TP}{FP+TP}$ (3)

- Recall is the success rate of the system in retrieving information.

  Formula from recall = $\frac{TP}{FN+TP}$ (4)

Table 6. Confusion matrix

Data Class	Classified as Positive	Classified as Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP2)	True Negative (TN)

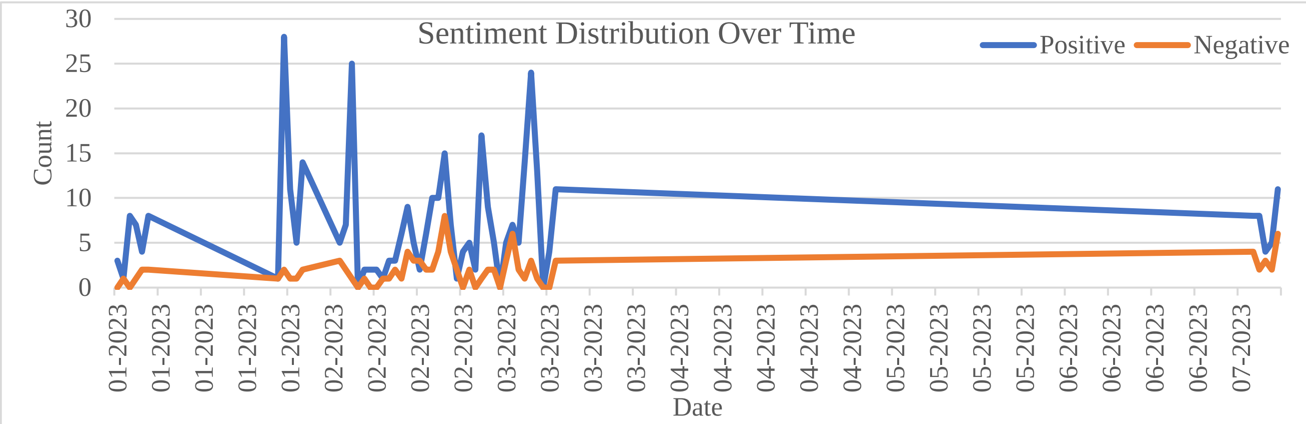

Fig. 11 Positive and negative sentiment trend charts

Information :

- TP = if all training data label A predicts the same result as A.
- TN = if all training data is not labeled A, the predicted result is not the same as A.
- FP = if all training data is not labeled A, the predicted result is equal to A.
- FN = if all label A training data predicted results are not the same as A.

4.4. Application Trial

This trial will display data calculated by sentiment and its categories in statistical form.

1. Showing Trends Figure 11 shows positive and negative sentiments about the Ministry of Investment/BKPM in six months. The visualization above is a representation of Twitter data that has calculated sentiment. It can be seen that based on system predictions, the percentage of positive sentiment is higher. This shows that people express more positive opinions.
2. Displays the Ministry of Investment/BKPM's Positive and Negative Sentiment Charts The visualization above Figure 12 is a representation of Twitter data that has calculated sentiment. It can be seen that based on system predictions, the percentage of positive sentiment is higher. This shows that people express more positive opinions.
3. Display sentiment distribution based on location. The visualization above Figure 13 is a representation of sentiment distribution based on locations. Positive using green color and red represented negative sentiment.
4. Display statistics on the development of sentiment words that often appear figure 14 shows a word cloud result for filtering famous words' positive and negative sentiments about the Ministry of Investment/BKPM in six months.

Table 7. Result sentiment performance from the system

Data (TP+FP+FN)	Accuracy %	Recall %	Precision %	f-measure %
88	80	92,30	85,71	88

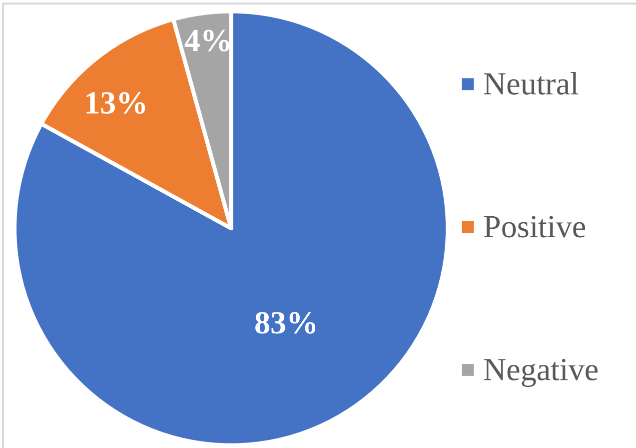

Fig. 12 Positive sentiment percentage chart

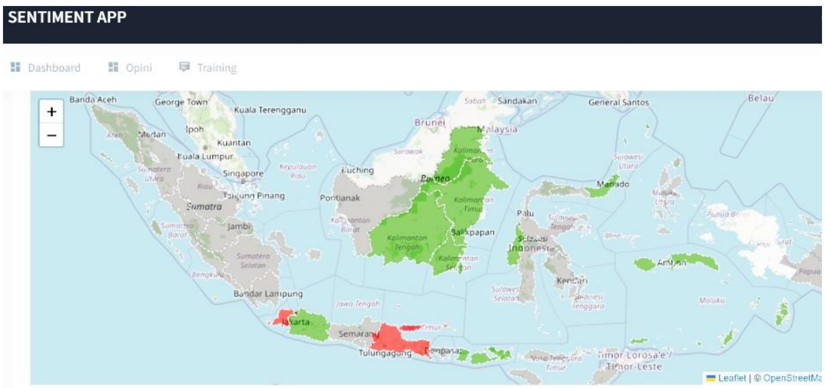

Fig. 13 Distribution of positive and negative sentiment

Fig. 14 Words that are often mentioned in Tweets.

4.5. Naïve Bayes Performance Analysis using k-fold

The test results show that the results achieved are 82%; by looking at the numbers obtained, the authors try to use k-fold as a cross-validation method to try to increase the accuracy of the results. k-fold is one of the popular cross-validation methods that fold data as much as k and repeats (iteration) the experiment as much as k as well. In this test, the data to be used is 88 tweets, including test data, which will later be divided into five parts or k = 5 so that the data obtained is 88 data divided into five folds with 15, 15, 15, 15, 18 and 10, respectively. Testing using partitioned data will be repeated five times (k = 5) with different test data positions in each iteration. Suppose the first iteration of the test data is in the initial position, the second iteration is in the second position, and so on.

The description of the k-fold iteration up to the fifth iteration can be seen in the following figure.

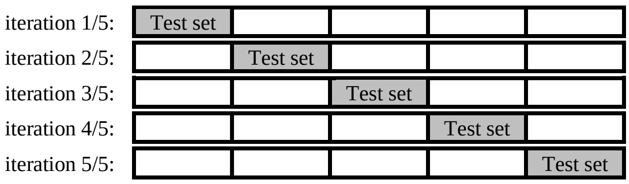

Fig. 15 Overview of k-fold iterations

For testing the 1st fold to the 5th fold, the number of test data entered is 100 tweets with the position of the test data as depicted in Figure 15, and the results can be seen in Table 8 below

Table 8. Test Results in k Fold Naive Bayes

Fold	Total Data (TP+FP+FN)	Accuracy %	Recall %	Precision %	f-measure %
I	15	73,33	84,61	84,61	84,61
II	15	26,67	66,67	30,77	42,10
III	15	80	92,30	85,71	88,89
IV	15	73,3	84,61	84,61	84,61
V	18	55,55	71,42	71,42	71,42
AVG	-	61,77	79,922	71,424	74,326

The best results were obtained from testing the highest test data during the 3rd fold test, namely 80% accuracy, 92.3% recall, 85.7% precision, and 88% f-measure. Evaluation of the research method, starting from collecting tweet data preprocessing to calculating probability values using the Naive Bayes Classifier, there are still various problems that make the system in this study work not optimally, namely:

Lack of Training Data: The training data used in this study is lacking because only 400 data are used, so when the entered test data is not recognized in the training data, the predicted results and actual labels do not match. The more training data used, the higher the resulting accuracy, and the accuracy of the system in recognizing test data will be better because the system can recognize many sentences and varied vocabulary, which is used as learning by the system.

There is an incorrect dataset: An error in the system classification process was caused by an inaccurate dataset. This situation causes many data appearance features that do not belong to the category in the test data used. These problems can also affect system accuracy, resulting in system performance failing to run optimally.

Constraints understanding sentiment classification: The obstacle often encountered in understanding sentences to be processed in the sentiment classification process is when encountering an initial sentence, which gives the perception that the sentence is a positive sentence. The end of the sentence provides the perception that the sentence is a negative sentence or a sentence that initially gives the perception that the sentence is negative. The end of the sentence provides the perception that the sentence is positive. These constraints result in system performance failing to run optimally because the system only detects sentiment classifications in sentences at the beginning. One example is the sentence in the training data, "Bagus Sekali Pak Menteri BKPM." Those who only talk capital and owe money for campaigns are told to become lecturers". The sentence is included in the negative sentiment category in manual data classification. Still, when using the system classification, the sentence is included in the positive sentiment category because, at the beginning of the sentence, there are words that contain positive sentiment.

5. Conclusion

Opinion analysis regarding government policy is one of the right ways to determine the extent of public response to government policy. Because it uses Twitter social media as media mining, sentiment analysis is an opinion analysis approach suitable for textual data. This approach is quite powerful in assessing a text. By using the impression as a basic technique, the opinion-mining process becomes more accessible, and the addition of rules to the impression technique produces better results. The impression technique is very dependent on the word sentiment dictionary; the more complete the vocabulary is in the word sentiment dictionary, the better the results are produced by this technique. The method developed in this study resulted in a percentage of sentiment analysis using Naïve Bayes of 82% with a ruled-based analysis approach. It is possible to produce even higher accuracy by adding data.

References

[1] Mustofa Kamal, Ali Ridho Barakbah, and Nur Rosyid Mubtadai, "Temporal Sentiment Analysis for Opinion Mining of ASEAN Free Trade Area on Social Media," *2016 International Conference on Knowledge Creation and Intelligent Computing (KCIC)*, Manado, Indonesia, pp. 207-212, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[2] Jie Li, and Lirong Qiu, "A Sentiment Analysis Method of Short Texts in Microblog," *2017 International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, Guangzhou, China, pp. 776-779, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[3] Suhariyanto, Ari Firmanto, and Riyanarto Sarno, "Prediction of Movie Sentiment based on Reviews and Score on Rotten Tomatoes Using SentiWordnet," *International Seminar on Application for Technology of Information and Communication (iSemantic)*, Semarang, Indonesia, pp. 202-206, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[4] Endang Wahyu Pamungkas, and Divi Galih Prasetyo Putri, "An Experimental Study of Lexicon-Based Sentiment Analysis on Bahasa Indonesia," *Proceedings - 2016 6th International Annual Engineering Seminar*, Yogyakarta, Indonesia, pp. 28-31, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[5] Bagus Setya Rintyarna, Riyanarto Sarno, and Chastine Fatichah, "Enhancing the Performance of Sentiment Analysis Task on Product Reviews by Handling Both Local and Global Context," *International Journal of Information and Decision Science*, vol. 12, no. 1, pp. 75-101, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Sheeba Naz, Aditi Sharan, and Nidhi Malik, "Sentiment Classification on Twitter Data Using Support Vector Machine," *IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Santiago, Chile, pp. 676-679, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[7] Victoria Ikoro et al., "Analyzing Sentiments Expressed on Twitter by UK Energy Company Consumers," *Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, Valencia, Spain, pp. 95-98, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[8] Endah Susilawati, "Public Services Satisfaction Based on Sentiment Analysis: Case Study: Electrical Services in Indonesia," *2016 International Conference on Information Technology System and Innovation (ICITSI)*, Bandung, Indonesia, pp. 1-6, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[9] Quanzhi Li et al., "Tweet Sentiment Analysis by Incorporating Sentiment-Specific Word Embedding and Weighted Text Features," *2016 IEEE/WIC/ACM International Conference on Web Intelligence*, Omaha, NE, USA, pp. 568-571, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[10] Ahmad Fathan Hidayatullah, and Azhari SN Azhari, "Sentiment Analysis and Category Classification of Public Figure on Twitter," *National Informatics Seminar*, vol. 1, no. 1, pp. 115-118, 2014. [Google Scholar] [Publisher Link]

[11] Rengga Asmara, Achmad Basuki, and M. Udin Harun Al Rasyid, "Gender Based Temporal Sentiment Analysis in Indonesian on Culinary Places in Surabaya City," *Technomedia Journal*, vol. 5, no. 1, pp. 67-81, 2020. [CrossRef] [Google Scholar] [Publisher Link]