

## Analysis of Emoticon and Sarcasm Effect on Sentiment Analysis of Indonesian Language on Twitter

Debby Alita<sup>1)\*</sup>, Sigit Priyanta<sup>2)</sup>, Nur Rokhman<sup>3)</sup>

<sup>1)</sup> Universitas Teknokrat Indonesia, Indonesia

Jl. H. Zainal Abidin Pagaralam 9-11 Labuhanratu, Bandar Lampung

<sup>1)</sup> [debbyalita@gmail.com](mailto:debbyalita@gmail.com)

<sup>2) 3)</sup> Departemen Ilmu Komputer dan Elektronika, FMIPA Universitas Gadjah Mada, Indonesia

Sekip Utara Bulaksumur Kabupaten Sleman, Yogyakarta

<sup>2)</sup> [sigitpriyanta@ugm.ac.id](mailto:sigitpriyanta@ugm.ac.id); <sup>3)</sup> [nurrokhman@gmail.com](mailto:nurrokhman@gmail.com)

### Article history:

Received 22 Februari 2019  
Revised 16 May 2019  
Accepted 16 May 2019  
Available online 28 October 2019

### Keywords:

Emoticon, Naïve Bayes Classifier,  
Random Forest Classifier, Sarcasm,  
Support Vector Machine

### Abstract

**Background:** Indonesia is an active Twitter user that is the largest ranked in the world. Tweets written by Twitter users vary, from tweets containing positive to negative responses. This agreement will be utilized by the parties concerned for evaluation.

**Objective:** On public comments there are emoticons and sarcasm which have an influence on the process of sentiment analysis. Emoticons are considered to make it easier for someone to express their feelings but not a few are also other opinion researchers, namely by ignoring emoticons, the reason being that it can interfere with the sentiment analysis process, while sarcasm is considered to be produced from the results of the sarcasm sentiment analysis in it.

**Methods:** The emoticon and no emoticon categories will be tested with the same testing data using classification method are Naïve Bayes Classifier and Support Vector Machine. Sarcasm data will be proposed using the Random Forest Classifier, Naïve Bayes Classifier and Support Vector Machine method.

**Results:** The use of emoticon with sarcasm detection can increase the accuracy value in the sentiment analysis process using Naïve Bayes Classifier method.

**Conclusion:** Based on the results, the amount of data greatly affects the value of accuracy. The use of emoticons is excellent in the sentiment analysis process. The detection of superior sarcasm only by using the Naïve Bayes Classifier method due to differences in the amount of sarcasm data and not sarcasm in the research process.

## I. INTRODUCTION

Twitter is a social media that can be used by all people to express themselves freely, so Twitter has a significant user increase [1]. There are several reasons why Twitter is superior. Firstly, there are many third-party software that can be used to explore Twitter. Also, Twitter has a user-friendly interface which makes it easier to use and adapt. Furthermore, Twitter is a real-time search engine which makes it as reliable research tool and therefore faster data analysis can be applied. One of the most frequently used forms of data analysis is sentiment analysis.

Sentiment analysis is the process of classifying sentiments expressed in documents to obtain information [2] and the purpose is to convert that information into natural language [3]. Basically sentiment analysis is a classification, but in reality it is not as easy as the usual classification process because it is related to the use of language [4], where there is ambiguity in the use of words, the absence of intonation in a text, and the development of the language itself. In face-to-face communication, sentiments can be inferred from visual signals such as smiles that can be described with emoticons. Using emoticons in communication is very popular nowadays because it is considered to be able to dilute formal communication, making communication more open and making it easier for someone to express their feelings through messages or comments on social media [5]. In the past few years, the approach to sentiment analysis has been widely proposed. However, many other approaches ignore emoticons as part of the features. Instead, they use texts since emoticons are seen as possible interference to the sentiment analysis process

\* Corresponding author



[6]. However, it is found that there are many comments on social media which use emoticons to give positive or negative opinions. There are also other opinions in the form of innuendo to political figures, product brands, companies, etc. which are often referred to sarcastic sentences. Sarcasm is a word that has the opposite meaning of what is said that is used to mock or show resentment. Almost all sarcastic sentences take the form of positive sentences that have negative meanings, even though it is not impossible that there is a sarcastic sentence that has positive meaning [7].

In the study [8], analyzing the use of emoticons in Dutch-language forum messages based on lexicons and providing an increase in the value of accuracy on the classification results. Meanwhile, in research [9] used 3 features to detect sarcasm, namely unigram, negativity and number of exclamations. The results showed that features negativity and the number of exclamations opinions can increase accuracy by 6% compared to just using the unigram feature. Tests and analysis were conducted to find out how much influence emoticons and sarcasm in the sentiment analysis process were classified using the Naïve Bayes Classifier method, Support Vector Machine and random forest classifier.

## II. METHODS

This study used the Naïve Bayes Classifier and Support Vector Machine method for the sentiment data classification process to produce positive, negative, and neutral labeled data. Afterward, sarcasm detection was carried out using random forest classifier method which changed the positive data to negative labeled data. The sentiment analysis process consisted of 4 stages, namely data collection, preprocessing, classification, and detection of sarcasm.

The test carried out in this study was the test of fold cross-validation. It consisted of three testing processes, namely the process of sentiment analysis without the detection of sarcasm, the sarcasm detection process, and the sentiment analysis process with the detection of sarcasm. In the sentiment analysis process, there were two comparisons of testing namely test data by utilizing emoticons and test data without utilizing emoticons.

The test results were analyzed to ascertain the extent of the effect of using emoticons and sarcasm detection on the sentiment analysis process. The data used was the Indonesian language commentary obtained from twitter. From several classification methods, information about the values of accuracy, precision, recall, and F1score was also attained.

The proposed process in the system consisted of tweet crawling, preprocessing, feature extraction, classification, and detection of sarcasm. The following steps can be seen in Fig.1.

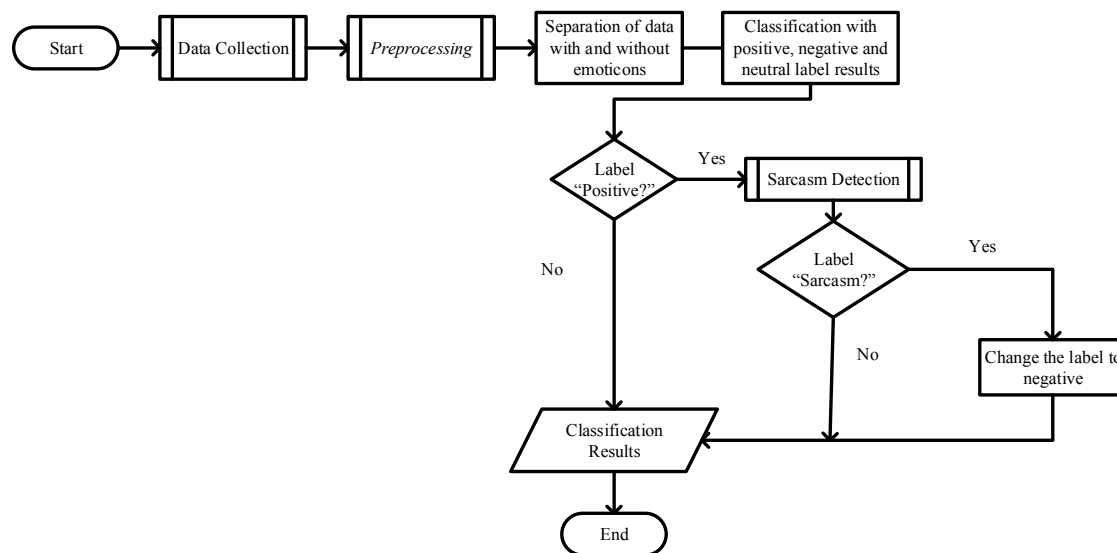


Fig.1. General Process Design

### A. Data Collection

In the process of data collecting, the data were obtained from twitter.com. That is from scraping the twitter search page used to retrieve tweets. Data collection starts from the period beginning in January 2018 to March 2018 with data in Indonesian language tweets about “public services”. The tweet data will be in the form of groups according



to the group name created when searching for data based on product brands, companies and governments. The tweet data obtained is then given a positive, neutral and negative label for the classification of sentiment analysis and label sarcasm or not sarcasm in a positive label for the process of detecting sarcasm.

The total data generated through the scraping process are 2,281 data which then through the preprocessing process the data is divided into two parts, namely the first part is the emoticon data, the data is done convert emoticons with 2,072 data numbers. The second part is non emoticon data which ignores data emoticons with 2,072 data numbers. The data consists of three sentiment polarity namely positive, negative and neutral sentiments. The number of emoticon data that is labeled positive is 1023, negative is 587, and neutral is 462, while the number of non emoticon data that is labeled positive is 557, negative is 406, and neutral is 1109.

The labeling of the data are manually listed by involving expert respondents. Examples of the results of data retrieval stages can be seen in Table 1.

TABLE 1  
EXAMPLE OF CRAWL TWEET RESULT

|           | Doc | Tweet   | Class    |
|-----------|-----|---|----------|
| Indosat   | D1  | Thanks @IndosatCare follow up kendala jaringannya kemarin. Jaringan & data seluler udah semakin baik. Semoga kualitas jaringan tambah bagus ☺ | positive |
|           | D2  | Jaringan indosat @IndosatCare tuh awal tahun doang bagus, kesininya jaringan internet4G BGZD bikin emosi. Mending pake yang lain aja guys :-/ | negative |
| PLN       | D3  | Terimakasih @PLN_123 sudah mematkan listrik rumah saya padahal saya sudah rajin bayar listrik setiap bulan :) Anjanyyy                        | negative |
| Telkomsel | D4  | @Telkomsel @agus_pinoa tutup halo aja gus, ganti simpati. Lebih hemat. Kualitas sinyal sama. Percayalah :-D                                   | positive |
|           | D5  | Simpati sinyalnya naik turun kadang bagus kadang jelek.. hufft @Telkomsel ☺   | neutral  |
|           | D6  | Alhamdulillah akhirnya ada juga yang hack @Telkomsel, sungguh kasiannya :( .... HAHHAHA   | negative |

### B. Preprocessing

In conducting text mining, the text of the document used must be prepared in advance, after that it can only be used for the main process. The process of preparing text in a document or raw dataset is also called a text preprocessing process. Text preprocessing serves to convert unstructured or arbitrary text data into structured data [10].

Preprocessing stage is to manage tweet data to make it easier to carry out the classification process to be more accurate and preprocessing is the initial stage used before taking steps in data mining to get better data quality [11]. In the preprocessing process there are five stages, namely tokenizing, case-folding, stopword, convert emoticons, and slangword.

### C. Emoticon

Emoticon is a symbol or a combination of symbols used to describe facial expressions and feelings in the form of messages or writing. Paralanguage emoticons are used to express positive, negative and neutral people on Twitter, Facebook, or other communication and social media [12]. Some examples of emoticons are :-> >:) :> :o) :] :3 :c) :>=] 8) =) :> ^) :D >:D :D 8-D x-D xD XD =-D =D =-3 =3 >:P :~P :P X-P x-p xp XP :~p :p =p :~b :b >:o >:O :~O :O o\_O o.O 8-0 for positive, >:[ :(: :~c :c :< :< :~[ :[ :> ,>., >.,< :?(>:\>:/ :/~ :./ :~\ =/= :S D :< D : D 8 D;D =DX v.v D-~; for negative and :| :~| for neutral.

#### D. Classification

The Naïve Bayes Classifier algorithm is a simple component for the classification process and looks for the highest probability value by applying the Bayes theorem which will be assumed to be in the learning machine [13]. There are two stages in the classification of documents. The first stage is training on documents that have known categories. While the second stage is the process of classifying documents that are not yet known[14]. At the time of classification the algorithm will look for the highest probability of all categories of documents tested (Vmap), where the equations are as follows (1):



$$V_{map} = \underset{V_j \in \{"Positif", "Negatif"\}}{\operatorname{argmax}} P(V_j) \prod P(x_i | V_j) \quad (1)$$

Information :

$V_j$  : Tweet category  $j=1,2,3,\dots,n$ . when in this research  $j_1$  = Positive sentiment tweet category ,  $j_2$  negative sentiment *tweet* category.

$P(x_i | V_j)$  : Probability  $x_i$  on  $V_j$  category

$P(V_j)$  : Probability from  $V_j$

Support Vector Machine (SVM) has proven to be a useful learning machine, especially for multiclass data classification. Basically, there are two kinds of approaches for multiclass SVM. First, directly process all data in one optimization formulation. the second describes multiclass into a binary SVM series. The idea behind binary SVM is to build a multiclass classification from binary with the one against one technique [15]

This method with one against one technique was built  $(k(k-1))/2$  the binary classification model fruit ( $k$  is the number of classes). Each classification model is trained on data from two classes. There are several methods for carrying out testing after the whole  $k(k-1)/2$  the classification model was completed. One of them is the voting method [16]. For training used 3 pieces of binary SVM and their use in classifying new data can be seen in Fig.2.

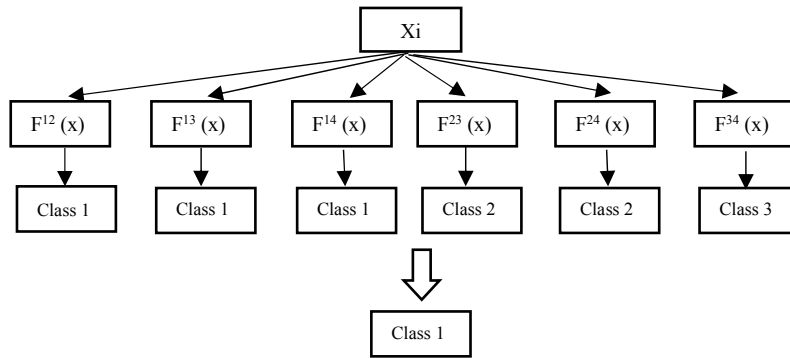


Fig. 2. Examples of classification with the SVM method

### E. Sarcasm Detection

Sarcasm detection is done on tweets that have a positive classification result in sentiment analysis. The features that will be extracted in the detection of sarcasm consist of sentiment-relate, punctuation-relate, lexical and syntactic, and pattern-relate [17]. Sarcasm detection will be classified using the Random Forest Classifier.

In the sentiment-relate feature set, there are 10 features that will be extracted. The first feature is to determine the number of words in the form of opinion words that have a tendency to be positive or negative. The second, third, and fourth feature extracted is the number of positive, negative, and sarcasm emoticons. The fifth and sixth feature is to determine the number of hashtags on a tweet. The number of positive and negative hashtags is calculated and made a feature. The seventh feature up to the tenth feature describes the contrast between components, where there are negative and positive components in one tweet in the first order are the word components, the two hashtag components with the hashtag, the three word components with hashtag, and finally the word component with emoticons.

In the set of punctuation-relate feature, there are 6 features that will be extracted. The first feature is counting the number of exclamation marks, the second feature calculating the number of question marks, the third feature calculating the number of points, the fourth feature calculating the number of letters, the fifth feature calculating the number of words, and the last feature calculating the number of repetitions of more than 2 times on a word.

In the lexical and syntactic feature set, there are 4 features that will be extracted [18], namely Count of Laughter, Laughter is a word that describes joy, pleasure or amusement that is described through writing on a tweet. Examples of laughter that are often found in Indonesian-language tweets are hehe, haha, hihi, hoho, wkwk, wkwko, wkewke, and LOL. Count of Exclamation, Exclamations are words that express one's feelings and intentions. Exclamations are used to show admiration, wonder, disgust, or fear of something. Examples of exclamations that can be recognized are, huh, uh, uh, oh, aw, iw, uw, ew, ow, waw, wow, wah, woh, idih, dihi, aih, duh, huh, hm, um, ups , cie, loh, pft, and o. The existence of words that are rarely used, Words that are rarely used are derived from the



words in the data tweet, which only appear once. The existence of a general pattern of sarcasm sentences. The general pattern of sarcasm sentences is obtained by extracting all possible sentence patterns with word lengths that vary from 3 to 6. Words that have a POS CD, FW, NN, PRP, SYM, UH, MD, RB, WDT, WP will be converted into POS on a tweet. Sentences that appear at least 2 times will be selected and checked manually to eliminate irrelevant patterns. Then the sentence pattern that appears on the tweet does not sarcasm also be deleted.

The way to extract the pattern-relate feature set is the same as in the general pattern of sarcasm and then list the patterns of sarcasm and group the number of sarcasm patterns based on the length of the sarcasm pattern.

#### F. Sarcasm Classification

Sarcasm detection will be classified using the Random Forest Classifier, Naïve Bayes Classifier and Support Vector Machine by utilizing 23 features that have been extracted previously. Positive Tweets classified as sarcasm will be changed to negative. Here are examples of tweets and extraction results of sarcasm detection features with manually determined label values that can be seen in Table 2.

TABLE 2  
AN EXAMPLE OF EXTRACTION OF THE SARCAISM DETECTION FEATURE

| No | Tweet  | Extraction Feature Result  | Label   |
|----|--|--|---------|
| 1  | Terimakasih @PLN_123 sudah mematikan listrik rumah saya padahal saya sudah rajin bayar listrik setiap bulan :) Anjayyy | [0.28, 1, 0, 1, 0, 0, True, False, False, True, 0, 0, 0, 0, 0, True, True, False, 0, 0, 0.56, 1.24, 0.012]   | Sarcasm |
| 2  | Alhamdulillah akhirnya ada juga yang hack @Telkomsel, sungguh kasiannya :( .... Makanya jangan mahal-mahal HAHHAHA     | [-0.5, 0, 1, 0, 0, 0, True, False, False, True, 0, 0, 0, 4, 1, False, True, False, 1, 0, 1.04, 0.627, 0.678] | Sarcasm |
| 3  | Pelayanan @BPJSKesehatanRI selalu dinomor belakangan :(  | [-1, 0, 1, 0, 0, 0, False, False, False, False, 0, 0, 0, 0, 0, False, False, False, 0, 0, 1.03, 0.61, 0.66]  | no      |

Making Decision Trees is obtained from the results of sarcasm detection extraction consisting of several Trees that contain 23 features. Each Tree will determine whether a tweet is a sarcasm sentence or not. Examples of vote collection can be seen in Table 3.

TABLE 3  
EXAMPLE OF TAKING VOTES ON THE RANDOM FOREST CLASSIFIER

| Tweet  | Extraction Feature   | Tree Vote |    |    |    |
|--|--|-----------|----|----|----|
|  |  | 1         | 2  | 3  | 4  |
| Admin @Indosatcare balesin pesan lebih lama daripada mantan :D | [0.25, 1, 0, 0, 0, 0, True, False, False, False, True, 0, 0, 0, 0, 0, False, True, False, 0, 0, 1.54, 0.04, 0.018] | no        | no | no | no |

The results of the sarcasm detection classification can be seen from the highest number of votes on the Random Forest Classifier using Decision Trees, the result of which is "No", the sentence of the tweet is not sarcasm.

### III. RESULTS

#### A. Results of testing sentiment analysis

In the sentiment analysis test without sarcasm detection using 2 methods namely Naïve Bayes Classifier and Support Vector Machine. The test uses 2-fold cross validation to 10-fold cross validation and then compares the effect of using emoticons with no emoticons and obtains a quality measure of the sentiment analysis test that can be seen in Table 4.

From Table 4, it can be seen that the results of the Naïve Bayes Classifier method with the use of emoticons provide the best accuracy, precision, recall and F1score values through 10-Fold cross validation with an accuracy value of 74.57%, precision value of 47.30%, value 53.21% recall and F1Score value 50.00%.

The results of testing the Naïve Bayes Classifier method without the use of emoticons provide the highest accuracy value in testing 6-Fold cross validation and the precision, recall and F1score values are best done through 10-Fold cross validation with an accuracy of 52.91%, precision values of 50.32 %, 46.91% recall value and F1Score value 46.75%.



TABLE 4  
THE MEASURE OF THE QUALITY OF SENTIMENT ANALYSIS USING THE NAÏVE BAYES CLASSIFIER METHOD

| Testing             | Fold Cross Validation |         |         |         |         |         |         |         |         |
|---------------------|-----------------------|---------|---------|---------|---------|---------|---------|---------|---------|
|                     | 2                     | 3       | 4       | 5       | 6       | 7       | 8       | 9       | 10      |
| <b>Emoticon</b>     |                       |         |         |         |         |         |         |         |         |
| Accuracy            | 50.35%                | 50.35%  | 51.88%  | 51.49%  | 52.91 % | 51.60 % | 51.55 % | 51.93 % | 52.04 % |
| Precision           | 39.27 %               | 41.07 % | 44.37 % | 45.80 % | 47.18 % | 47.57 % | 47.96 % | 48.85 % | 50.32 % |
| Recall              | 36.15 %               | 38.26 % | 41.86 % | 43.89 % | 44.72 % | 45.00 % | 45.43 % | 46.36 % | 46.91 % |
| F1Score             | 35.07 %               | 37.75 % | 41.50 % | 43.71 % | 44.34 % | 44.47 % | 45.11 % | 46.07%  | 46.75%  |
| <b>non Emoticon</b> |                       |         |         |         |         |         |         |         |         |
| Accuracy            | 47.74 %               | 47.80 % | 48.99 % | 48.67 % | 50.03 % | 49.32 % | 48.78 % | 49.81 % | 49.21 % |
| Precision           | 34.72 %               | 37.36 % | 42.01 % | 46.52 % | 46.37 % | 49.86 % | 48.34 % | 49.46 % | 49.38 % |
| Recall              | 33.24 %               | 34.78 % | 39.81 % | 43.40 % | 43.66 % | 46.68 % | 45.44 % | 46.58 % | 46.37 % |
| F1Score             | 31.31 %               | 32.64 % | 37.26 % | 40.86 % | 41.18 % | 44.02 % | 42.83 % | 43.65 % | 43.39 % |

TABLE 5  
THE VALUE OF THE QUALITY OF SENTIMENT ANALYSIS USING THE SVM METHOD

| Testing             | Fold Cross Validation |         |         |         |         |         |         |         |         |
|---------------------|-----------------------|---------|---------|---------|---------|---------|---------|---------|---------|
|                     | 2                     | 3       | 4       | 5       | 6       | 7       | 8       | 9       | 10      |
| <b>Emoticon</b>     |                       |         |         |         |         |         |         |         |         |
| Accuracy            | 75.82 %               | 77.79 % | 77.49 % | 77.01 % | 77.61 % | 77.26 % | 77.20 % | 77.19 % | 77.67 % |
| Precision           | 52.44 %               | 58.84 % | 58.52 % | 59.68 % | 62.10 % | 62.23 % | 63.19 % | 62.56 % | 64.01 % |
| Recall              | 50.65 %               | 56.90 % | 56.89 % | 58.80 % | 60.20 % | 60.96 % | 61.83 % | 61.42 % | 62.45 % |
| F1Score             | 50.52 %               | 56.90 % | 56.69 % | 58.59 % | 60.17 % | 60.71 % | 61.45 % | 61.09 % | 62.32 % |
| <b>Non Emoticon</b> |                       |         |         |         |         |         |         |         |         |
| Accuracy            | 47.74 %               | 47.80 % | 48.99 % | 48.67 % | 50.03 % | 49.32 % | 48.78 % | 49.81 % | 49.21 % |
| Precision           | 34.72 %               | 37.36 % | 42.01 % | 46.52 % | 46.37 % | 49.86 % | 48.34 % | 49.46 % | 49.38 % |
| Recall              | 33.24 %               | 34.78 % | 39.81 % | 43.40 % | 43.66 % | 46.68 % | 45.44 % | 46.58 % | 46.37 % |
| F1Score             | 31.31 %               | 32.64 % | 37.26 % | 40.86 % | 41.18 % | 44.02 % | 42.83 % | 43.65 % | 43.39 % |

From Table 5 it can be seen that the results of testing the Support Vector Machine method with the use of emoticons provide the highest accuracy value in testing 3-Fold cross validation and the best precision, recall and F1score values are done through 10-Fold cross validation with an accuracy value of 77.79%. The precision value is 64.01%, the recall value is 62.45% and F1Score value is 62.32%.

The results of testing the Support Vector Machine method without the use of emoticons provide the highest accuracy value in testing 9-Fold cross validation and the best precision, recall and F1score values are done through 7-Fold cross validation with an accuracy value of 49.81%, precision value of 49.86 %, recall value of 46.68% and F1Score value of 44.02%.

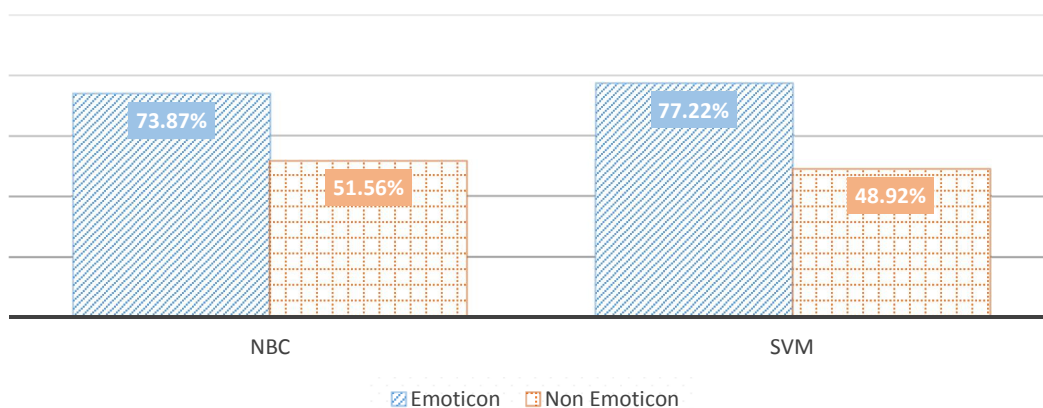


Fig. 3. A comparison chart of quality measures in sentiment analysis without detection of sarcasm

In Fig. 3 shows that the Naive Bayes Method without the use of emoticons has a higher average value of accuracy of 51.56% compared to the SVM method, while the highest use of emoticons is achieved by the SVM method which has a value of 77.22%.



### B. Sarcasm detection test results

The results of sarcasm detection tests can be seen in Table 6. This sarcasm detection testing using random forest classifier method, Naïve Bayes Classifier and Support Vector Machine. It extracted with 4 features, namely relate feature sentiment, relate feature punctuation, lexical and syntactic features, and relate feature pattern obtained by the quality measure of sarcasm detection tests

TABLE 6  
RESULTS OF SARCASM DETECTION TESTS

| Testing                  | Fold Cross Validation |         |         |         |         |         |         |         |         |
|--------------------------|-----------------------|---------|---------|---------|---------|---------|---------|---------|---------|
|                          | 2                     | 3       | 4       | 5       | 6       | 7       | 8       | 9       | 10      |
| Naïve bayes classifier   |                       |         |         |         |         |         |         |         |         |
| Accuracy                 | 54.97%                | 54.53 % | 54.97 % | 54.64 % | 55.18 % | 55.30 % | 54.86 % | 54.96%  | 54.59 % |
| Presision                | 62.85 %               | 62.60 % | 62.82 % | 62.39 % | 63.07 % | 63.13 % | 62.87 % | 62.99 % | 62.67 % |
| Recall                   | 64.32 %               | 63.97 % | 64.31 % | 63.64 % | 64.54 % | 64.68 % | 64.35 % | 64.29 % | 64.04 % |
| F1Score                  | 54.78 %               | 54.35 % | 54.76 % | 54.28 % | 54.97 % | 54.97 % | 54.65 % | 54.65 % | 54.35 % |
| Support vector machine   |                       |         |         |         |         |         |         |         |         |
| Accuracy                 | 61.30 %               | 59.44 % | 60.04 % | 60.04 % | 60.15 % | 61.13 % | 60.48 % | 60.26 % | 60.26 % |
| Presision                | 58.28 %               | 59.48 % | 58.54 % | 59.55 % | 59.89 % | 60.45 % | 59.24 % | 59.53 % | 59.63 % |
| Recall                   | 59.90 %               | 61.69 % | 60.42 % | 61.70 % | 62.09 % | 62.69 % | 61.45 % | 61.52 % | 61.75 % |
| F1Score                  | 57.50 %               | 57.45 % | 57.10 % | 57.73 % | 57.96 % | 58.80 % | 57.89 % | 57.75 % | 57.84 % |
| Random forest classifier |                       |         |         |         |         |         |         |         |         |
| Accuracy                 | 59.50 %               | 62.55 % | 61.19 % | 60.10 % | 60.10 % | 61.30 % | 60.21 % | 59.44 % | 61.14 % |
| Precision                | 56.95 %               | 59.39 % | 58.72 % | 58.54 % | 58.50 % | 59.53 % | 58.80 % | 58.09 % | 58.69 % |
| Recall                   | 58.22 %               | 61.09 % | 60.59 % | 60.36 % | 60.28 % | 61.56 % | 60.71 % | 59.92 % | 60.60 % |
| F1Score                  | 55.95 %               | 58.80 % | 57.90 % | 57.24 % | 57.12 % | 58.40 % | 57.25 % | 56.56 % | 57.75 % |

From Table 6, it can be seen that the results of the sarcasm detection test with the Naïve Bayes Classifier method provide the best accuracy, precision, recall and F1score values through 7-Fold cross validation with an accuracy value of 55.30%, precision value of 63.13%, value recall of 64.68% and F1Score value of 54.97%.

The results of the sarcasm detection test with the Support Vector Machine method provide the best value of accuracy, precision, recall and F1score done through 7-Fold cross validation with an accuracy value of 61.13%, a precision value of 60.45%, a recall value of 62.69% and F1Score value of 58.80%.

The results of the sarcasm detection test with random forest classifier method provide the highest accuracy of 62.55% and F1 score value of 58.80% by testing 3-Fold cross validation. The highest precision and recall values were obtained and through testing 7-Fold cross validation with a precision value of 59.53% and a recall value of 61.56%.

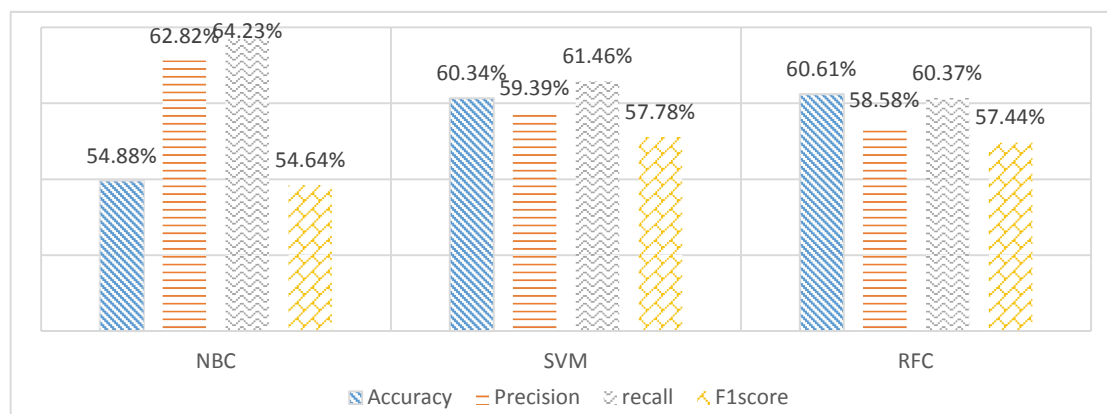


Fig. 4. Graph of comparison of quality measures in the detection of sarcasm

Fig. 4. shows that the random forest classifier method has a higher average value of accuracy of 60.61% compared to other methods, while the highest value of precision and recall is obtained by the Naïve Bayes Classifier



method with a precision value of 62.82% and value recall of 64.23% and for F1score value with a value of 57.78% with the Support Vector Machine method.

### C. The results of testing sentiment analysis with the detection of sarcasm

In testing sentiment analysis with sarcasm detection tweets that have been labeled positive, negative, neutral are used, then for sentences that are labeled positively a sarcasm and no sarcasm labeling process will then be carried out by the classification process and the quality measure of the sentiment analysis test with the detection of sarcasm can be seen in Table 7.

TABLE 7  
THE MEASURE OF THE QUALITY OF SENTIMENT ANALYSIS WITH THE DETECTION OF SARCASM USING THE NAÏVE BAYES CLASSIFIER METHOD

| Testing             | Fold Cross Validation |         |         |         |         |         |         |         |         |
|---------------------|-----------------------|---------|---------|---------|---------|---------|---------|---------|---------|
|                     | 2                     | 3       | 4       | 5       | 6       | 7       | 8       | 9       | 10      |
| <b>Emoticon</b>     |                       |         |         |         |         |         |         |         |         |
| Accuracy            | 60.89 %               | 61.49 % | 61.37 % | 61.73 % | 61.67 % | 61.80 % | 61.80 % | 61.73 % | 61.85 % |
| Precision           | 32.04 %               | 37.41 % | 37.68 % | 40.02 % | 39.65 % | 40.46 % | 40.85 % | 41.22 % | 41.50 % |
| Recall              | 35.68 %               | 41.76 % | 41.93 % | 44.33 % | 44.14 % | 44.80 % | 45.39 % | 45.59 % | 46.02 % |
| F1Score             | 32.18 %               | 37.54 % | 37.71 % | 39.96 % | 39.61 % | 40.36 % | 40.73 % | 40.96 % | 41.24 % |
| <b>Non Emoticon</b> |                       |         |         |         |         |         |         |         |         |
| Accuracy            | 47.27 %               | 46.01 % | 48.14 % | 47.27 % | 48.14 % | 46.88 % | 47.82 % | 47.70 % | 47.81 % |
| Precision           | 32.97 %               | 37.21 % | 39.92 % | 40.76 % | 43.06 % | 43.82 % | 44.41 % | 45.02 % | 45.56 % |
| Recall              | 31.72 %               | 36.00 % | 38.85 % | 39.88 % | 41.85 % | 42.86 % | 43.10 % | 43.68 % | 44.52 % |
| F1Score             | 31.17 %               | 35.55 % | 38.18 % | 39.31 % | 41.36 % | 41.75 % | 42.42 % | 42.87 % | 43.74 % |

From the results of Table 7, it can be seen that the results of testing the Naïve Bayes Classifier method with the use of emoticons provide the best accuracy, precision, recall, and F1score values through 10-Fold cross validation with an accuracy value of 61.85%, precision value of 41.50%, recall value of 46.20%, and F1Score value of 41.24%.

The results of testing the Naïve Bayes Classifier method without the use of emoticons provide the highest accuracy value in testing 4- Fold cross validation and 6- Fold cross validation and the best value of precision, recall and F1score is done by testing 10-Fold cross validation with an accuracy value of 47.81%, precision value of 45.56%, recall value of 44.52% and F1Score value of 43.74%.

TABLE 8  
THE MEASURE OF THE QUALITY OF SENTIMENT ANALYSIS WITH THE DETECTION OF SARCASM USING THE SVM METHOD

| Testing      | Fold Cross Validation |         |         |         |         |         |         |         |         |
|--------------|-----------------------|---------|---------|---------|---------|---------|---------|---------|---------|
|              | 2                     | 3       | 4       | 5       | 6       | 7       | 8       | 9       | 10      |
| Emoticon     |                       |         |         |         |         |         |         |         |         |
| Accuracy     | 62.95 %               | 61.80 % | 61.92 % | 62.40 % | 61.92 % | 62.09 % | 62.28 % | 62.34 % | 62.59 % |
| Precision    | 42.35 %               | 51.37 % | 52.56 % | 55.66 % | 54.83 % | 55.25 % | 57.27 % | 57.38 % | 57.84 % |
| Recall       | 57.84 %               | 47.36 % | 48.05 % | 50.85 % | 50.18 % | 51.79 % | 52.32 % | 52.89 % | 52.65 % |
| F1Score      | 36.33 %               | 45.53 % | 46.32 % | 49.14 % | 48.32 % | 49.39 % | 50.37 % | 50.84 % | 50.70 % |
| Non Emoticon |                       |         |         |         |         |         |         |         |         |
| Accuracy     | 46.12 %               | 47.05 % | 47.10 % | 46.01 % | 46.78 % | 47.49 % | 46.88 % | 46.51 % | 47.16 % |
| Precision    | 35.96 %               | 42.88 % | 42.04 % | 44.28 % | 45.56 % | 48.03 % | 47.97 % | 48.75 % | 48.94 % |
| Recall       | 32.97 %               | 38.59 % | 38.74 % | 40.36 % | 41.78 % | 43.74 % | 43.95 % | 44.55 % | 44.74 % |
| F1Score      | 30.08 %               | 35.45 % | 35.71 % | 36.74 % | 38.46 % | 40.18 % | 40.24 % | 40.90 % | 41.27 % |

From Table 8, it can be seen that the results of testing the Support Vector Machine method with the use of emoticons provide the highest accuracy value in testing 2-Fold cross validation and the highest precision value in testing 10-Fold cross validation, while the best recall and F1score values are tested 9- Fold cross validation with an accuracy value of 48.14%, precision value of 45.56%, recall value of 44.52% and F1Score value of 43.74%.

The results of testing the Support Vector Machine method without the use of emoticons provide the highest accuracy value in testing 7-Fold cross validation and the best precision, recall and F1score values are done by testing 10-Fold cross validation with an accuracy of 47.49%, precision values of 48.94%, recall value 44.74% and F1Score value 41.27%.

In Fig. 5 shows that the SVM method with the use of emoticons has the highest average value of accuracy performed with a value of 62.25% for the highest average accuracy value without the use of emoticons obtained through the Naïve Bayes Classifier method with a value of 47.45%.



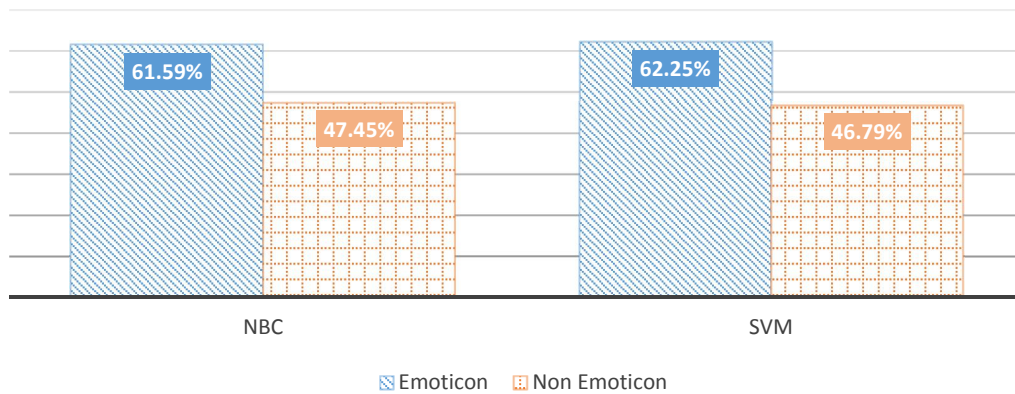


Fig. 5. Graph of comparison of quality measures in sentiment analysis with detection of sarcasm

#### IV. DISCUSSION

Sentiment classification stage uses the Naïve Bayes Classifier and Support Vector Machine method. The Tweet will be classified into three classes: positive, negative, and neutral [19]. Before classification, a feature extraction process consisting of unigram, POS Tag, TF, and TF-IDF will be carried out. The final value of the classification is in the form where the largest value will determine the class of tweets, before getting the final grade the training process will be carried out on the training data which will be used to determine the class in the test data. The SVM method uses votes to determine the sentiment class of each of the two class classification results. The classification results with a positive label will then be used for the process of detecting sarcasm because sarcasm in this study is sarcasm in positive but negative meaning sentences which are often used to insinuate someone, product brand, company, etc.

At the sarcasm detection stage using the Random Forest Classifier method, the Support Vector Machine and the Naïve Bayes Classifier contain a set of decision trees or trees. Each tree uses different random features that can be used to determine which votes and votes with the most classes will be used as a result of the sarcasm classification or not sarcasm [20].

Testing is done by measuring accuracy from the results of sentiments regarding public opinion carried out by the system. Based on the results of the tests carried out, it can be seen which parameters produce the best accuracy.

This study produced a system to analyse the influence of emoticons and sarcasm in sentiment analysis, the test results using k-fold cross validation indicate that the sentiment analysis system with the use of emoticons produces better system accuracy compared to the sentiment analysis system without the use of emoticons.

The use of emoticons without sarcasm detection can increase the accuracy value of the sentiment analysis process with a slightly higher value than the use of emoticons with sarcasm detection which also increases the value of accuracy in the sentiment analysis process. The difference in the amount of sarcasm data and not sarcasm is the cause of the difference in increasing the value of accuracy with the sentiment analysis process and the difference in the amount of data on positive, negative and neutral sentiments is also one of the causes of the value less than 70%.

#### V. CONCLUSIONS

Based on the research that has been done, it can be concluded that the use of emoticons without the detection of sarcasm can increase the value of accuracy in the sentiment analysis process by 25.30% and the use of emoticons with sarcasm detection can increase the value of accuracy in the sentiment analysis process of 14.80%. The best method in the sentiment analysis process with the use of emoticons is the SVM method with a value of 69.74% and without the use of emoticons is the Naïve Bayes Classifier method with a value of 49.51%.

#### REFERENCES

- [1] A. F. Hidayatullah and A. Sn, "Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter," *Semin. Nas. Inform.* 2014, vol. 2014, no. August 2013, pp. 0–8, 2014.
- [2] T. Parlar, E. Sarac, and S. A. Ozel, "Comparison of feature selection methods for sentiment analysis on Turkish Twitter data," in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, 2017, pp. 1–4.
- [3] F. A. Pozzi, E. Fersini, E. Messina, and B. Liu, *Sentiment Analysis in Social Networks*. 2016.



- [4] R. Habibi and D. B. Setyohadi, "Analisis Sentimen Pada Twitter Mahasiswa Menggunakan Metode Backpropagation," vol. 12, no. 1, pp. 103–109, 2016.
- [5] F. Jiang, "Microblog Sentiment Analysis with Emoticon Space Model," vol. 30, pp. 1120–1121, 2015.
- [6] G. A. Buntoro, T. B. Adji, and A. E. Purnamasari, "Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation," no. OCTOBER 2014, pp. 7–8, 2014.
- [7] E. Inderasari, F. Achsani, and B. Lestari, "Bahasa Sarkasme Netizen Dalam Komentar Akun Instagram ' Lambe Turah,'" vol. 8, no. 1, 2019.
- [8] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. de Jong, and U. Kaymak, "Exploiting emoticons in sentiment analysis," in *Proceedings of the 28th Annual ACM Symposium on Applied Computing - SAC '13*, 2013, p. 703.
- [9] E. Lunando and A. Purwarianti, "Indonesian social media sentiment analysis with sarcasm detection," in *2013 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2013*, 2013, pp. 195–198.
- [10] H. Jiawei, M. Kamber, J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2006.
- [11] S. Adinugroho and Y. A. Sari, "Implementasi Data Mining Menggunakan WEKA", 2018, UB Press ISBN: 976-602-432-445-2.
- [12] A. R. Sentiaji and Adam Mukaharil Bachtiar, "Analisis Sentimen terhadap Acara Televisi berdasarkan Opini Public," *J. Ilm. Komput. dan Inform.*, vol. 2, no. 1, pp. 1–4, 2013.
- [13] R. Mehra, M. K. Bedi, G. Singh, R. Arora, T. Bala, and S. Saxena, "Sentimental analysis using fuzzy and naive bayes," in *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, 2017, pp. 945–950.
- [14] S. Vijayarani and S. Dhayanand, "Liver Disease Prediction using SVM and Naïve Bayes Algorithms," vol. 4, no. 4, pp. 816–820, 2015.
- [15] A. A. Esmael, J. A. Dos Santos, and R. da Silva Torres, "On the ensemble of multiscale object-based classifiers for aerial images: a comparative study," *Multimed. Tools Appl.*, pp. 1–28, 2018.
- [16] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.
- [17] M. Bouazizi and T. Ohtsuki, "Opinion Mining in Twitter How to Make Use of Sarcasm to Enhance Sentiment Analysis," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015 - ASONAM '15*, 2015, pp. 1594–1597.
- [18] F. Prawira, "The Effect Of Sarcasm Detection On Sentiment Analysis Quality Measurement In Twitter The Effect Of Sarcasm Detection On Sentiment," 2017.
- [19] A. Agarwal, B. Xie, I. Vovsha, O. Rambow and R. Passonneau "Sentiment analysis of Twitter data," *LSM '11 Proceedings of the Workshop on Languages in Social Media*, June, pp. 30–38, 2011.
- [20] N. Horning, "Random Forests : An algorithm for image classification and generation of continuous fields data sets," 2010.