Accredited by National Journal Accreditation (ARJUNA) Managed by
Ministry of Research, Technology, and Higher Education, Republic Indonesia with Second Grade (Peringkat 2, Sinta 2)
since year 2017 to 2021 according to the decree No. 10/E/KPT/2019.

Published online on the journal's webpage: **http://jurnal.iaii.or.id**

RESTI JOURNAL

(System Engineering and Information Technology)

Vol. 4 No. 5 (2020) 782 - 789 ISSN Electronic Media: 2580-0760

Implementation of Rumor Detection on Twitter Using the SVM Classification Method

Annisa Rahmaniar Dwi Pratiwi[1], Erwin Budi Setiawan[2*]
[12]Informatics, School of Computing, Telkom University
[1]annisarahmaniar@student.telkomuniversity.ac.id, [2]erwinbudisetiawan@telkomuniversity.ac.id
[*]corresponding author

Abstract

Twitter is one of the popular social network sites, that was first launched in 2006. This service allows users to spread real-time information. However, the information obtained is not always based on facts and sometimes deliberately used to spread rumors that cause fear to the public. So detection efforts are needed to overcome and prevent the spread of rumors on Twitter. Much research regarding the detection of rumors but is limited to English and Chinese. In this study, the authors built a system to detect Indonesian-language rumors based on the implementation of the SVM classification and feature selection using the TF-IDF weighting. Data collection was conducted in November 2019 to February 2020 using crawling methods by keywords and manual labeling process. Research data used topics around government and trending with 47,449 records and features combination based on users and tweets. Stages of research include the process of collecting data on the Twitter social networking site which is then carried out preprocessing consists of case-folding, URL removal, normalization, stopwords removal, and stemming. The next stage is feature selection, N-Gram modeling, classification, and evaluation using a confusion matrix. Based on the results of the study, the system gets good performance in the test scenario using 10% of testing data and unigram features with the highest accuracy value of 78.71%. As for features twitter that affected the detection of rumors covering the number of following, the number of like and mention.

Keywords: SVM, TF-IDF, Crawling, rumor, Twitter.

1. Introduction

Twitter is a microblogging-based social networking site that was first launched in 2006. The survey results [1] place the Twitter platform in the fourth position in Indonesia as the most accessed social networking site. Twitter relies on internet connections or devices by providing communication services in the form of message exchanges, so that users are connected. This service provides users the ability to create various information in real-time, users can write and share information in the immediate environment.

Information obtained from the social networking site Twitter is not all true and is sometimes used to spread false information or rumors. Aligned with monitoring results until December 2018 by the Ministry of Communication and Information (Kemenkominfo), Twitter is the most widely reported social networking site [2] . Reporting intended to complain about the negative content in the form of fraud, false news, and does not preclude the possibility of rumors

Rumors are information that has spread and has not been verified about its authenticity by deliberately seeking attention and spreading fear [3]. The credibility of the information becomes important, so detection efforts needed to overcome and prevent the spread of rumors on Twitter and increase user confidence in the Twitter platform.

The detection of rumors on Twitter is quite crucial. Based on the graphic publication by [4], illustrating that from 2015 to 2018 the number of studies related to the detection of published rumors continues to increase. The survey results provide an overview of the techniques used in detecting rumors. The classification approach is divided into sections based on machine learning and deep learning.

The detection of rumors seen as a classification problem to determine whether the tweets are rumor or not and constituting critical topic research to information credibility [5]. According to research by Akshi Kumar et al. 2019, Machine Learning-based techniques emerged as a promising approach to detect rumors on social

Accepted by editor: 13-06-2020 | Final Revision: 26-08-2020 | Online Publication : 30-10-2020

media with popular techniques including Support Vector Machine (SVM), Naïve Bayes and Decision Tree [3].

Research related to the detection of rumors by Yang Liu et al. 2012, about "Automatic Detection of Rumors on Sina Weibo" with the conclusion of the experimental results illustrate that the detection method based on the proposed SVM classification use the feature based on account, content, and propagation are effective in detecting rumors on social networks with an accuracy of up to 72% [6].

Research by Qiao Zhang et al. 2015 related to "Automatic Detection of Rumors on Social Network" the results of the detection method of rumors using the proposed SVM classification based on the implicit features of the content and users achieve better performance by a comparison based on basic features [5]. So that the implicit feature can identify better rumor detection.

Another research by Gang Liang et al. 2015 related to "Identification of Rumors in Microblogging Systems Based on User Behavior" identify rumors on Sina Weibo platform based on user behavior using the Machine Learning approach. User behavior is treated as a clue to show who tends to be rumor maker. The results showed that the proposed methods and features have good performance in identifying [7].

The research data is based on the Indonesian language with limited topics around government and trending topics. The current issue of government has been widely discussed with various public comments to give opinions and expressions on the issue.

Based on research that has been done before, the method implemented by the author on the system to classify tweet data using Support Vector Machine (SVM) with the concept to find the most optimum hyperplane [8]. The SVM classification method will try to find the best hyperplane that has a maximum margin so that it can classify all data accurately.

Feature selection implemented in research related to the detection of rumors is using the Weighting Frequency Inverse Document Frequency (TF-IDF). TF-IDF method which is considered easy to implement and is used to determine the frequency of words in certain documents [9].

The purpose of this research is to implement the SVM classification methods and TF-IDF weighting in the problem of detection rumors on Twitter and find out the level of accuracy produced and features of Twitter that affect the detection of rumors using the SVM classification method with a combination of TF-IDF.

2. **Research Method**

2.1. Rumor Detection Model Approach

In this research the system built can process data consisting of training data and testing with various scenarios so as to determine which scenario has the best performance. The system workflow can be seen in Figure 1.

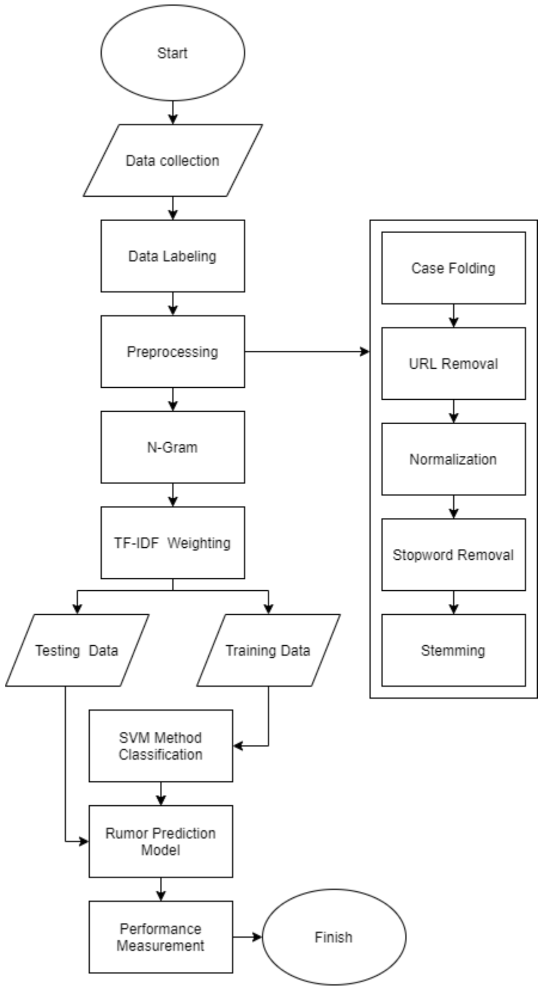

Figure 1. Rumor Detection Model

The initial stage of research was collecting data on the Twitter platform using the developed Twitter Crawler. The crawling process is based on the search method by keyword with the results obtained from fragments of words and hashtags which are indexes in the form of keywords or topics. Labeling the research data is done manually using human intelligence. Before the classification process, the data is first carried out preprocessing include case-folding, URL removal, normalization, stopwords removal, and stemming. The data obtained was then performed N-Gram modeling and feature selection using TF-IDF weighting and then divided into training and testing data. At the generalization stage, the SVM classification method is used to build a classification model based on the results of the training data. The classification model that has been built is used to predict or classification on testing of new tweets unknown the category. The category results obtained will be evaluated through performance measurements on systems built using a confusion matrix.

2.2. Data collection

Data collection in the study was carried out using the crawling method. Twitter Crawler used is a research innovation by Jaka Eka Sembodo, et al. 2016. This application collects tweets through Application Programming Integration (API) with the crawling process conducted in research using the by keyword method [10]. The results of crawling will be saved and entered into a database using SQL queries which are then exported in the .xlsx format.

2.3. Rumor Feature

Features are attributes that describe the characteristics of an object. In the classification problem, the attribute is used as a reference to determine the classification results category. In this research, the Twitter features used are a combination of user and tweets based. Twitter features are used in research.

Table 1. Rumor Features

Feature	Description
User Based	
Verified	User verification type
Number of Followers	Number of followers from users
Number of Following	The number of other users followed
Number of Tweets	Number of tweets made by the user
Tweet Based	
Number of Likes	Number of users who like a tweet
Number of Emoticons	The number of emojis in the tweet
Retweet	User tweet marks are included as tweets or retweets
Hashtag	The number of hashtags in a tweet
Mention	The number of mentions in the tweet

2.4. Data Labeling

Labeling of research data is done manually using human intelligence with the categories of Rumors and Non-Rumors. The label provided is based on the Rumor feature that has been defined. Illustration of labeling will be shown in Table 2.

Table 2. Illustration of labeling

Feature	Data	Data
Verified	1	0
Number of Followers	15442206	111
Number of Following	28	216
Number of Tweets	1550802	2007
Number of Likes	35	314
Number of Emoticons	0	0
Retweet	1	1
Hashtag	2	1
Mention	0	0
Tweet	"Pukat UGM menilai praktik bagi-bagi kekuasaan di kabinet yang dilakukan Presiden Jokowi berpotensi menumbuhkan korupsi di segala lini. Begini penjelasannya: #KabinetJokowi #Korupsi https:\/\/t.co\/dcAP4fi3lj"	"Saya tadi lewat depan gedung DPR, kok banyak polisi ya, pada hal gak ada demo. Jalan pada sepi, ada apa ya \n#MatikantvTgl20"
Class	Non Rumor	Rumor

2.5. N-Gram

N-Gram is a series of words or characters that contain n elements [11]. In various extraction of information retrieval and natural language processing n-gram has many uses. N-gram provides useful representations for tasks such as text categorization and machine translation. An example illustration of N-Gram can be seen in Table 3.

Table 3. Illustration of N-Gram

Tweet	N-Gram	Results
hari minggu masuk kerja nonton lantik jokowiasiik matikantvtgl	Unigram	hari', minggu, masuk, kerja, nonton, lantik, jokowiasiik, matikantvtgl
	Bigram	hari minggu, minggu masuk, masuk kerja, kerja nonton, nonton lantik, lantik jokowiasiik, jokowiasiik matikantvtgl
	Trigram	hari minggu masuk, minggu masuk kerja, masuk kerja nonton, kerja nonton lantik, nonton lantik jokowiasiik, nonton lantik jokowiasiik matikantvtgl

2.6. Preprocessing

The first step in processing the data to be used in classification systems. This process can assess the input data used whether quality or not. Low data quality will affect the performance of the implemented model [12]. In this research, preprocessing includes case folding, stemming, stopword removal, URL removal, and normalization.

Case Folding is a process that has a purpose forchanging all letters in the document to lowercase and eliminate characters such as numbers and punctuation [13]. The case folding illustration is shown in Table 4.

Table 4. Illustration of Case Folding

Tweet	Case Folding Results
Hari minggu msh msk kerja sambil nonton pelantikan jokowi..asiik #MatikantvTgl20	hari minggu msh msk kerja sambil nonton pelantikan jokowiasiik matikantvtgl

URL Removal is the process of removing URL or website address from twitter data. The appearance of a URL from Twitter data makes the data ineffective. The URL removal illustration is shown in Table 5.

Table 5. URL Removal Illustration

Tweet	Url Removal Results
hari minggu msh msk kerja sambil nonton pelantikan jokowiasiik matikantvtgl https:\\/\\/t.co\\/3t4vlmv15j	hari minggu msh msk kerja sambil nonton pelantikan jokowiasiik matikantvtgl

Normalization is the process of converting non-standard words in Twitter data into standard words that are done by matching tokens with word lists. Illustration of normalization will be shown in Table 6.

Table 6. Normalization Illustration

Tweet	Normalization Result
hari minggu msh msk kerja sambil nonton pelantikan jokowiasiik matikantvtgl	hari minggu masih masuk kerja sambil nonton pelantikan jokowiasiik matikantvtgl

Stopwords are terms or words that have no meaning and are not related to the contents of the document. Stopwords Removal is the process of deleting words from documents because they are not measured as keywords [14]. The removal of these words will not change the meaning and content of rumor and non-rumor information. Illustration of Stopwords Removal will be displayed in Table 7.

Table 7. Illustration of Stopwords Removal

Tweet	Stopwords Removal Results
hari minggu masih masuk kerja sambil nonton pelantikan jokowiasiik matikantvtgl	hari minggu masuk kerja nonton pelantikan jokowiasiik matikantvtgl

Stemming is the process of changing words that have affixes into basic words. The purpose of this process is to erase various affix words so as to reduce the number of words [14]. Implementation of stemming on the system using the library in Python Sastrawi. Illustration of stemming will be shown in Table 8.

Table 8 . Illustration of Stemming

Tweet	Stemming Results
hari minggu masuk kerja nonton pelantikan jokowiasiik matikantvtgl	hari minggu masuk kerja nonton lantik jokowiasiik matikantvtgl

2.7 Term Frequency – Inverse Document Frequency (TF-IDF)

TF-IDF is a method of weighting words with calculations that illustrate the importance of a term in a document. This process is used to assess the weight of words (terms) in all documents [15]. Term Frequency is a measure of the appearance of a term in a document and all documents in the corpus. While Inverse Document Frequency is a logarithmic ratio based on the total number of documents in the corpus with the number of documents that have the term. The TF-IDF method will calculate the weight of w_i in document d from the word frequency value or TF (t, d) and document frequency or DF (t) [16].

$$w_i = TF(t,d) \times log\frac{|D|}{DF(t)} \quad (1)$$

The TF (t, d) is the frequency of occurrence of the word t in document d. While DF (t) is a calculation of several documents | D | divided by the number of documents containing the word t. An illustration of the TF-IDF weighting process will be shown in Table 9.

Table 9. TF-IDF Illustration

	TF			IDF	W = TF × IDF	
Term	D1	D2	DF	Log |D| / DF	D1	D2
nonton	1		1	0.3010	0.3010	0
pelantikan		1	1	0.3010	0	0.3010
jokowiasiik	1	1	2	0	0	0

2.8 Support Vector Machine (SVM)

This is one of the Pattern Recognition methods developed by Boser, Guyon, Vapnik, and was first presented in 1992. SVM is part of Machine Learning that has the goal of finding the best hyperplane which generally can separate data sets into two classes.

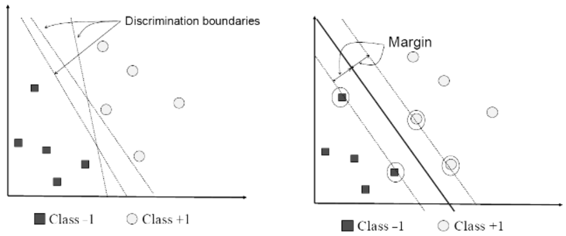

Figure 2. SVM trying to find the best hyperplane [17]

SVM tries to find the best hyperplane, in contrast to a neural network strategy that only tries to find hyperplane separators between classes. A hyperplane is said to be the best if it has the largest margin which will be able to classify new data accurately without errors. For two-dimensional data, a hyperplane in form straight line.

In real-world problems, data sets are generally separated non linearly, SVM is considered to be successfully implemented and generally provides better solutions than other methods. In non-linear SVM, data is mapped by the function Φ (x) to a higher dimensional vector space and then hyperplane as a separator between classes can be constructed [17].

The SVM learning process for finding support vector only depends on the dot product taken from the data that has been transformed, namely $\Phi_i \cdot \Phi_k$. Transformation Φ is difficult to transform so the calculation of dot products can be replaced with kernel $K(x_i \cdot x_k)$ This function is known as Kernel Trick with the following formula.

$$K(x_i \cdot x_k) = \Phi_j \cdot \Phi_k \quad (2)$$

The use of Kernel Trick can simplify the process because in determining support vectors only need to know the kernel functions to be implemented. Here are four types of kernels that can be used.

Table 10. Kernel Functions [8]

No.	Jenis Kernel	Rumus
1.	Kernel Linear	$K(x_i, x_k) = x_k^T x$
2.	Kernel Polynomial	$K(x_i, x_k) = (x_k^T + 1)^d$
3.	Kernel Gaussian (*radial basis function, RBF*)	$K(x_i, x_k) = \exp\{-\|\|x - x_k\|\|^2 / 2\sigma^2\}$
4.	Kernel Sigmoid	$K(x_i, x_k) = \tanh\,[k\, x_k^T x + \theta]$

In this research, the kernel that will be implemented on the system built is the RBF Kernel which has the same performance as the Linear kernel and behaves like the Sigmoid kernel function.

2.9 Measuring Performance

To find out the quality of the system needed an evaluation of the performance of the system designed. The Confusion Matrix is a useful tool for analyzing how well the classification is implemented to recognize tuples from various classes [18].

Table 11. *Confusion Matrix*

		Predicted Class	
Actual Class		*Positive* (P)	*Negative* (N)
	True (T)	*True Positive* (TP)	*False Negative* (FN)
	False (F)	*False Positive* (FP)	*True Negative* (TN)

Based on the general form of the Confusion Matrix there are several terms including True Positive (TP) that refer to positive tuples that are correctly labeled by the classifier. True negative (TN), negative tuples are correctly labeled by the classifier. False Positive (FP), negative tuples that are incorrectly classified as positive and then False Negative (FN), positive tuples that are incorrectly classified as negative. In this research, evaluation measurements on the system use Accuracy, Precision, Recall, and F-Scores.

Accuracy is a measure that shows the closeness of the measurement results with the real value. Precision is measurements based on the ratio of positive observations predicted correctly to the total predicted positive observations. The recall is a measure of completeness with measurements based on the ratio of positive observations correctly predicted with all observations in the actual class - True. F1-Score represents a harmonic average of precision and recall with a high F-Score indicates high classification performance [19].

Table 12. Performance Measurement

No.	Performance Measurement	Formulas
1	Akurasi	$\frac{TP + TN}{TP + FP + TN + FN}$
3	*Precision*	$\frac{TP}{TP + FP}$
4	*Recall*	$\frac{TP}{TP + FN}$
5	*F1 - Score*	$\frac{2 \times (Recall \times Precision)}{(Recall + Precision)}$

3. Result and Discussion

The process of testing data by the system is divided into 3 scenarios namely, scenario 1 by changing the composition of training and testing data in the classification process. Furthermore, scenario 2 adds an N-gram feature to determine its effect on the results of classification, and scenario 3 determines the Twitter feature that influences the detection of rumors.

3.1. Dataset

The dataset used is data on the Indonesian-language Twitter platform with a total of 47,449 records. Labeling is done manually with a class about rumors and non-rumors. The file format is in the .xlsx format. The data collection period is conducted from November 2019 to February 2020. The feature used is a combination of Twitter and TF-IDF features. The number of features used in each test is 1000 features. The selection of the number of features is based on the results of observations and limitations related to the time and capacity of the device. Table 12 is some of the hashtags and keywords used in the crawling process.

Table 13. Keyword

Keyword
Presiden Jokowi
#4niesIllegalLogging
#coronavirus
#AniesBaswedanJuaraBOHONG
#NegaraDijajahKoruptor
#100HariUnfaedah
#CongratsJokowiMarufAmin
#IbuKotaBaru
#AwasHTImasihEksis
#JiwasrayaMeradang

3.2. Test Result

Table 14. Testing Scenarios 1

Data Comparison	Accuracy %	Precision %	Recall %	F Score
90 : 10	78.71	78.18	95.94	0.8615
80 : 20	77.31	76.82	95.48	0.8514
60 : 40	76.21	75.82	95.49	0.8453
50 : 50	75.78	75.31	95.86	0.8435

Based on the results of testing the system with several composition data scenarios in Table 14, the detection of rumors using the SVM classification model and the TF-IDF weighting gives the highest accuracy results in the 90:10 data composition of training and testing data. The test results obtained are an accuracy value of 78.71% with Precision, Recall, and F Score of 78.18%, 95.94%, 86.15%.

After testing on scenario 1, the authors conducted a test to determine the effect of the TF-IDF weighting method on classification using a data composition with a ratio of 90: 10 to training and testing data.

Table 15. Testing without TF-IDF Features

Without Features	Accuracy %	Precisison %	Recall %	F Score
TF-IDF	76.26	77.11	95.00	0.8512

Based on the system test results in Table 15 it can be seen that the TF-IDF feature can be used to detect rumors on Twitter. The results of testing the system by removing the TF-IDF feature get an accuracy level of 76.26% with Precision, Recall, and F Score of 77.11%, 95.00%, 85.12%.

Testing in scenario 2 is based on the highest accuracy of previous tests with the addition of the N-gram feature. This is the number of N-Gram features used in scenario 2.

Table 16. Number of N-Gram features

N-Gram	Number of Features
Unigram	1000
Bigram	1000
Trigram	1000
Unigram + Bigram	500, 500
Unigram + Trigram	500, 500
Bigram + Trigram	500, 500
Unigram + Bigram + Trigram	400, 300, 300

Table 17. Testing Scenario 2

N-Gram	Accuracy %	Precision %	Recall %	F Score
Unigram	78.71	78.18	95.94	0.8615
Bigram	77.21	76.56	95.43	0.8496
Trigram	76.85	76.22	95.29	0.8470
Unigram + Bigram	77.95	77.02	96.14	0.8552
Unigram + Trigram	77.68	77.02	95.58	0.8530
Bigram + Trigram	78.10	78.06	95.40	0.8586
Unigram + Bigram + Trigram	77.63	76.95	95.74	0.8533

Based on the results of system testing with the scenario of adding N-Gram features in Table 17, rumor detection using the SVM classification model and TF-IDF weighting gives the highest accuracy results using the Unigram feature. The test results obtained are an accuracy value of 78.71%% with Precision, Recall, and F Score of 78.18%, 95.94%, 86.15%.

Based on the results of testing the system with scenarios to determine the features that affect the detection of rumors in Table 18, the results show that the following number of features, the number of likes and mentions most affect the classification. Testing without the number of following features, accuracy levels obtained is 76.35% with Precision, Recall, and F Score of 76.37%, 94.69%, 84.55%. While for the test without the number of likes feature, the accuracy obtained is 76.56% with Precision, Recall and F Score of 76.58%, 94.80%, 84.72% and for testing without mention the level of accuracy obtained is 77.17% with Precision, Recall and F Score of 77.11%, 95.00%, 85.12%.

Table 18. Testing Scenario 3

Without Features	Accuracy %	Precision %	Recall %	F Score
Verified	77.34	76.80	95.64	0.8519
Number of Followers	77.91	77.74	95.13	0.8556
Number of Following	76.35	76.37	94.69	0.8455
Number of Tweets	78.20	77.93	95.10	0.8566
Number of Likes	76.56	76.58	94.80	0.8472
Number of Emoticons	78.25	77.83	95.24	0.8566
Retweet	77.74	77.25	95.48	0.8541
Hashtag	77.49	76.96	95.57	0.8526
Mention	77.17	77.11	95.00	0.8512

3.3. Analysis of Test Results

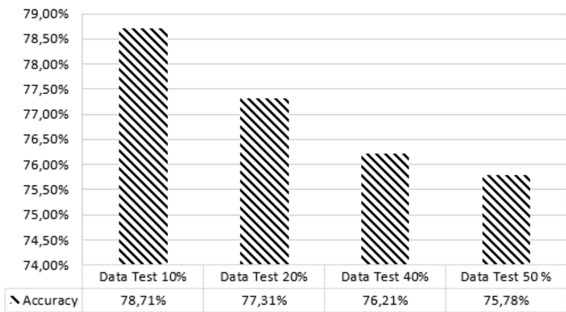

Figure 3. Accuracy of Scenario 1

The data sharing scenario implemented consists of several tests, namely 10%, 20%, 40%, and 50% of the testing data. This test uses a combination of Twitter features and 1000 TF-IDF features. The level of accuracy obtained by the system is 78.71%, 77.31%, 76.21%, and 75.78%. The system gets the best performance in testing using 10% testing data which is marked with the highest level of accuracy compared to other data sharing scenarios. While the lowest accuracy is obtained through testing of 50% of testing data. The accuracy level decreases when the amount of data trains also decreases.

This shows that more data train is used, the better the learning of the algorithm so that it affects the resulting accuracy. SVM is a Supervised Learning that requires a learning process in doing classification.

Features also affect the performance generated by the system in classifying data. To determine the effect of the TF-IDF feature in the system, two tests were performed by applying a combination of Twitter and TF-IDF features and eliminating the TF-IDF features in further tests.

The test results using a combination of Twitter features and TF-IDF get an accuracy rate of 78.71%. The level of accuracy generated without using the TF-IDF feature gets 76.26% accuracy. The level of accuracy generated by testing without the TF-IDF feature has decreased by 2.45%. This indicates that the use of the TF-IDF feature affects the classification process.

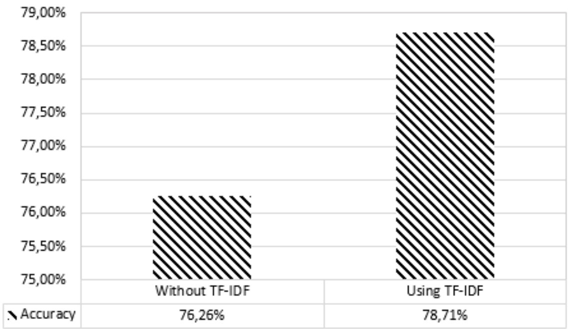

Figure 4. Accuracy of Feature Comparison

The number of features used can also affect system performance. In testing without the TF-IDF feature, the number of features used is less when compared with a combination of Twitter and TF-IDF features. The more features, the learning algorithm does the better so that the performance generated by the system in doing data classification is also getting better.

The selection of the number of TF-IDF features of 1000 is based on the results of observations. The first test was performed using 500 TF-IDF features which were then compared using 1000 TF-IDF features. The test results obtained that the number of features of TF-IDF of 1000 has higher accuracy. The selection of the number of TF-IDF features is done because of the limitations that are owned related to the device and time.

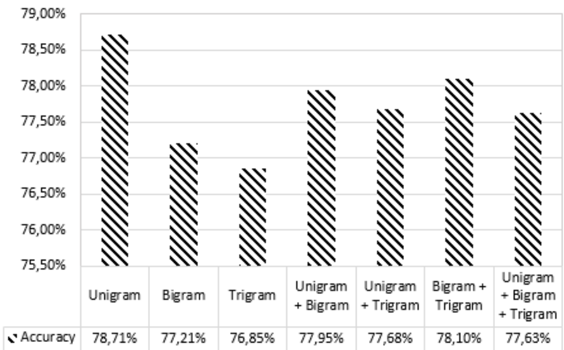

Figure 5. Accuracy of Scenario 2

The testing scenario with the addition of N-Gram consists of several N-Gram compositions, namely Unigram, Bigram, Trigram, and a combination of Unigram, Bigram, Trigram. The system gets the best performance in testing with Unigram feature composition which gets the highest accuracy rate of 78.71% compared to testing using the Bigram feature which only gets an accuracy of 77.21% and Trigram of 76.85%. However, in testing by applying a combination of N-Gram the system gets the best performance in testing by applying a combination of Bigram and Trigram which is equal to 78.10%. As for testing on other N-Gram combinations get lower accuracy that is equal to 77.95% in the Bigram Unigram combination, 77.68% in the Triram Unigram combination, and 77.63% in the testing with the Unigram Bigram Trigram combination.

This shows that the average dataset has a similarity with word unigram thus affecting its level of accuracy. In the N-Gram combination, the accuracy obtained is also close to the results of the Unigram test which indicates that the dataset has the same words as the Bigram Trigram word combination. While the lowest accuracy is produced in testing with the Trigram feature which shows the word Trigram only exists in a few tweets, not the whole dataset.

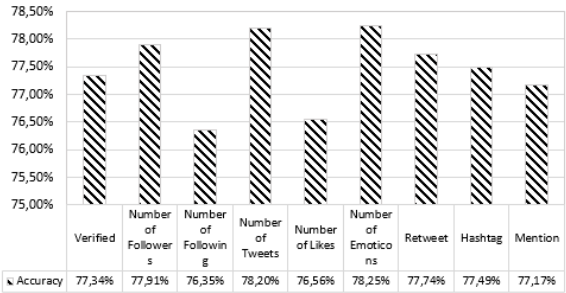

Figure 6. Accuracy of Scenario 3

Twitter features that are used on the system consist of verified, number of following, number of followers, number of tweets, number of likes, number of emojis, retweets, hashtags, and mentions. To find out the Twitter features that affect the classification process of testing is done by eliminating Twitter features one by one so that it can determine its effect on accuracy.

Based on the Twitter features used, there are three Twitter features that are most affect, namely the number of following, the number of likes and mentions. The highest level of accuracy obtained by the system in scenarios 1 and 2 is 78.71% while in scenario 3 with testing without features number of following number gets an accuracy of 76.35%, 76.56% in the number of likes, and 77.1% in the mention. So that these features affect the classification process of rumors which can be seen from the decrease in the level of accuracy obtained by the system by 2.36%. This shows that the lower the accuracy obtained in testing scenario 3, the Twitter feature will increasingly affect the class of rumors

Preprocessing and labeling also greatly affect system performance. There are still many non-standard words that have not been included in the normalization dictionary. This results in the word that cannot be processed in preprocessing and becomes a separate feature so that the number of features used is increasing. Labeling is also still done manually using human intelligence so that the error rate in labeling the data can be said to be quite high.

4. Conclusion

Based on tests that have been conducted on the Indonesian language rumor detection system on Twitter, it can be concluded that the SVM classification model and TF-IDF feature selection can be used to detect rumors on Twitter. Tests get the highest accuracy of 78.71% against 10% of testing data by implementing a combination of Twitter and TF-IDF features compared without the TF-IDF feature which only gets an accuracy of 76.26%. Whereas in the addition scenario N-Gram gets the highest level of accuracy in the combination of Unigram features that is equal to 78.71%. The Twitter features that affect the detection of rumors on Twitter include the number of following, the number of likes and mentions with the accuracy obtained at 76.35%, 76.56%, and 77.17%.

The suggestion from this research is that future studies use other classification models with the additional features to find out which model has the best performance and influential features in detecting rumors on Twitter.

References

[1] datareportal.com, "Digital 2019: Indonesia — DataReportal – Global Digital Insights," 2019. [Online]. Available: https://datareportal.com/reports/digital-2019-indonesia. [Accessed: 09-Oct-2019].

[2] kominfo.go.id, "Warganet Paling Banyak Laporkan Akun Twitter," 2019. [Online]. Available: https://kominfo.go.id/content/detail/15852/siaran-pers-no-08hmkominfo012019-tentang-warganet-paling-banyak-laporkan-akun-twitter/0/siaran_pers. [Accessed: 09-Oct-2019].

[3] A. Kumar and S. R. Sangwan, "Rumor Detection Using Machine Learning Techniques on Social Media," *Int. Conf. Innov. Comput. Commun. Lect. Notes Networks Syst.*, vol. 56, pp. 443–451, 2019.

[4] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Inf. Sci. (Ny).*, vol. 497, pp. 38–55, 2019.

[5] J. Li, H. Ji, D. Zhao, and Y. Feng, "Automatic Detection of Rumor on Social Network," *Nat. Lang. Process. Chinese Comput. NLPCC 2015. Lect. Notes Comput. Sci.*, vol. 9362, pp. 113–122, 2015.

[6] F. Yang, Y. Liu, X. Yu, and M. Yang, "Automatic Detection of Rumor on Sina Weibo Categories and Subject Descriptors," *Proc. ACM SIGKDD Work. Min. Data Semant.*, vol. 2, pp. 13:1--13:7, 2012.

[7] G. Liang, W. He, C. Xu, L. Chen, and J. Zeng, "Rumor Identification in Microblogging Systems Based on Users' Behavior," *IEEE Trans. Comput. Soc. Syst.*, vol. 2, no. 3, pp. 99–108, 2015.

[8] Suyanto, *Machine Learning: Tingkat Dasar dan Lanjut*. Informatika Bandung, 2018.

[9] B. Trstenjak, S. Mikac, and D. Donko, "KNN with TF-IDF based framework for text categorization," *Procedia Engineering*, vol. 69. pp. 1356–1364, 2014.

[10] J. Eka Sembodo, E. Budi Setiawan, and Z. Abdurahman Baizal, "Data Crawling Otomatis pada Twitter," no. September, pp. 11–16, 2016.

[11] G. Giannakopoulos, V. Karkaletsis, G. Vouros, and P. Stamatopoulos, "Summarization system evaluation revisited: N-gram graphs," *ACM Trans. Speech Lang. Process.*, vol. 5, no. 3, pp. 1–39, 2008.

[12] S. García, *Intelligent Systems Reference Library 72 Data Preprocessing in Data Mining*. 2015.

[13] A. Rahman, "Online News Classification Using Multinomial Naive Bayes," *Itsmart*, vol. 6, no. 1, pp. 32–38, 2017.

[14] S. Vijayarani, M. J. Ilamathi, M. Nithya, A. Professor, and M. P. Research Scholar, "Preprocessing Techniques for Text Mining -An Overview," *Int. J. Comput. Sci. Commun. Networks*, vol. 5, no. 1, pp. 7–16, 2015.

[15] M. N. Saadah, R. W. Atmagi, D. S. Rahayu, and A. Z. Arifin, "Information Retrieval Of Text Document With Weighting TF-IDF And LCS," *J. Ilmu Komput. dan Inf. (Journal Comput. Sci. Information)*, vol. Vol 6, No, pp. 34–37, 2013.

[16] T. B. Adji, Z. Abidin, and H. A. Nugroho, "System of negative Indonesian website detection using TF-IDF and Vector Space Model," *Proc. 2014 Int. Conf. Electr. Eng. Comput. Sci. ICEECS 2014*, no. February, pp. 174–178, 2014.

[17] D. Nugroho, A.S., Witarto, A.B., Handoko, "Application of Support Vector Machine in Bioinformatics," *Proceeding Indones. Sci. Meet. Cent. Japan, December 20, 2003, Gifu-Japan*, pp. 842–847, 2011.

[18] S. Agarwal, *Data mining: Data mining concepts and techniques*. 2014.

[19] A. Tharwat, "Classification assessment methods," *Appl. Comput. Informatics*, 2018.