

---

# Sentiment Analysis to Prediction DKI Jakarta Governor 2017 on Indonesian Twitter

Ghulam Asrofi Buntoro<sup>a</sup>

<sup>a</sup>Informatics Engineering, Department of Engineering, Universitas Muhammadiyah Ponorogo Jl. Budi Utomo 10, Ponorogo, 63471, Indonesia

---

## ABSTRACT

This study was conducted to test opinion data tweet of three candidates for governor Jakarta, 2017. Data only in Indonesian tweet, data tweet 100 tweets with keywords AHY, 100 tweets with keywords Ahok, and 100 tweets with keywords Anies. Data taken by random either from a normal user or online media at Twitter. Indonesian tweet opinion with three candidates for governor Jakarta in 2017 divided into three sentiment: positive, neutral and negative. Preprocessing data is, Lower Case Tokens, Normalization, Tokenization, Cleansing and Filtering. Classification method in this study using Naïve Bayes classifier (NBC), because this method is the most widely performed for sentiment analysis and proven always produce highest accuracy. Results of classification, Precision AHY data scored the highest with 95% and 95% Recall, while Ahok data lowest Precision scores with 81.6% and 82% recall.

**Keywords:** data mining; machine learning; sentiment analysis; jakarta governor candidate 2017; naïve bayes classifier.

---

## Article History

Received 13 August 17

Received in revised form 12 September 17

Accepted 12 November 17

---

## 1. Introduction

Social media, especially Twitter lately bustling discussing Jakarta Governor Election will be held soon in 2017. Schedule Jakarta Governor Election has been set [1]. Since registration and set a candidate for Governor Jakarta in 2017, names of candidates many discussed, in real world or the virtual world. Everyone is think free opine on Jakarta governor candidate in 2017 so many opinions, not only positive or neutral opinion but negative. Media development so quickly, much online media from news media to social media. Social media alone so much, from Facebook, Twitter, Path, Instagram, Google+, Tumblr, LinkedIn and so many more [2]. Social media today is not only used as a means of friendship, make friends, but it already used to other activities. Promo merchandise, buying and selling up promo political or campaigns candidate legislative and presidential candidate,

Social media especially Twitter today being a communication device is very popular in internet users. Official Twitter developer conference Chirp, 2010, the company delivered the statistics regarding site and Twitter users. The statistics said in April 2010, Twitter has 106 million accounts and 180 million unique visitors each month. Number of Twitter users mentioned user continues increase to 300,000 per day [3]. Digital Buzz blog a site provides statistics infographic mentions same statistical data. Successor team of candidates the governor or mayor can justify any means in campaign of the candidate. Especially on governor election campaign or mayor the region, there is the term Black Campaign, especially in social

media. Because the current promotions, imaging campaigns not only in real world but also in cyberspace. Social media, especially Twitter now become one of the promotions or campaigns are effective and efficient.

Sentiment analysis or opinion mining is the process of understanding, extract and process the textual data automatically to get the sentiment of information contained in an opinion sentence. The magnitude of effect and benefits to research sentiment analysis, sentiment analysis based applications is growing rapidly. Even in America, there are about 20-30 companies that focus on sentiment analysis service [4]. Social network mining (SNM) has become one of main themes on agenda of research data mining lately. Partly resulting from this research, we can extract information from a variety of social media, but the sources of information to develop dynamically requires a flexible approach [5]. In addition to sentiment analysis, data mining has also been developed for the prediction of precipitation such as those used to predict rainfall in Malaysia [6]. In this study used the data mining sentiment analysis of public opinion in Twitter to Jakarta governor candidate in 2017. There are three data for candidates for Governor Jakarta in 2017 were categorized into positive opinion, neutral and negative. The amount of sentiment to Jakarta governor candidate in 2017 could be one of the parameters of victory or defeat of a candidate.

Research by Merfat M. Altawaier [7] uses machine learning to classify sentiment tweet in Arabic Twitter. Three methods of classification used is Naïve Bayes classifier (NBC), Decision Tree (DT) and Support Vector Machine (SVM). This study determined classification of sentiment tweet Arabic Twitter, tweet about politics and arts with two sentiment: positive

---

\* Corresponding author.

E-mail address: [ghulamasrofibuntoro@gmail.com](mailto:ghulamasrofibuntoro@gmail.com).

and negative. Weighting was conducted using TF-IDF (Term Frequency – Inverse Document Frequency) and use a Arabic stemming algorithm. of three techniques or classification methods, Decision Tree (DT) method best the other method with f-measure of 78%.

Research by Mesut Kaya [8] uses machine learning to classify the Turkish political news. This study determined classification of sentiment Turkish political news and political news to determine whether Turkey has a positive or negative sentiment. feature of Turkish political news machine learning algorithms extracted with Naïve Bayes classifier (NBC), Maximum Entropy (ME) and Support Vector Machine (SVM) to generate a classification model. This study obtained Accuracy 72.05% for the Naïve Bayes classifier (NBC), Maximum Entropy Accuracy 69.44% and 66.81% for the SVM to use bigram tokenization.

Researches other using machine learning made by Pak. and Paurobek [9]. Use emoticons to build a corpus of English Twitter with sentiment positive, negative and neutral. For the class of neutral Pak and Paurobek take data tweet from account English media. Classification used Naïve Bayes classifier (NBC), tokenization method used unigram, Bigram and n-grams. Best performasi generated when using classification Naive Bayes classifier (NBC) and tokenization method Bigram.

Research machine learning for sentiment classification the 2014 presidential candidates G.A.Buntoro [10]. This research holds views and opinions of the people by dividing them into five class attribute, which is very positive, positive, neutral, negative and very negative. The classification process in this study using the classification method Naive Bayes classifier (NBC) with the preprocessing data using tokenization, cleansing and filtering. data tweet in Indonesian Twitter about Indonesia Presidential candidate 2014, dataset of 900 tweets were distributed evenly into five class attribute. The highest of results accuracy obtained when using a Naive Bayes classifier (NBC) tokenization n-gram, stopwords list WEKA and emoticons, with accuracy 71.9%, 71.6% Precision, 71.9% recall, 66,1% TP rate and 65% TN rate.

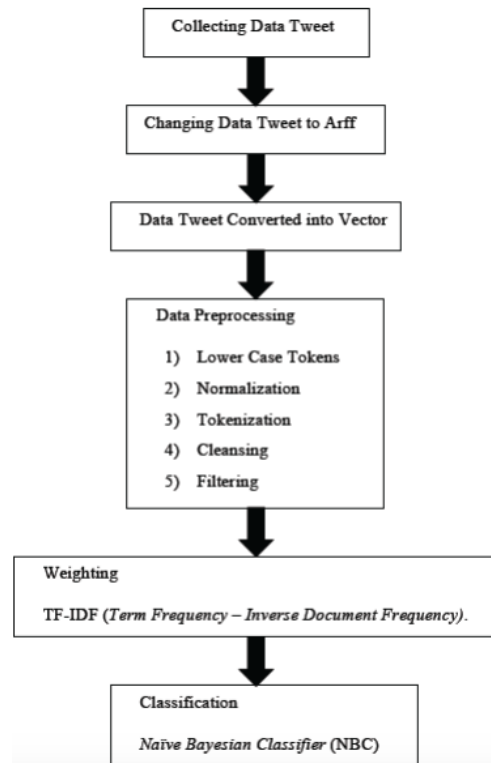
With machine learning research by Frangky [11] tried to repeat experiment sentiment classification movie review by Pang to Indonesian. In connection with not be training corpora for Indonesian, then applied to machine translation tools to translate English corpus created Pang native to ndonesian and the results are used to train translate classification. Wide choice of machine translation is used from commercial tools to simple translation word by word and text classification methods attempted. Average accuracy of the results obtained for the Naïve Bayes method was 74.6% and 75.62% for SVM method. Best results are obtained together with that obtained when using experiments in English.

## 2. Methodology

This study was conducted to test data tweet to opinion of three candidates for governor Jakarta, 2017. Data only tweet in Indonesian, which is 100 tweets with keywords AHY, 100 tweets with keywords Ahok, and 100 tweets with keywords Anies. Data taken at random either from a normal user or online media at Twitter. Tweet Indonesian opinion with the three candidates for governor Jakarta in 2017 and then divided into three sentiment: positive, neutral and negative. Preprocessing Data is, Lower Case Tokens, Normalization, Tokenization, Cleansing and Filtering. The Classification method in this study using Naïve Bayes classifier (NBC), because this method the most powerfully performed for sentiment analysis and proven always produce the highest accuracy. Research steps in accordance with the flow of the study are:

Figure-1 Flowchart Method

### 2.1. Collect data tweet



Data taken with Crawling tweet from social media Twitter. Data only tweet in Indonesian, which is 100 tweets with keywords AHY, 100 tweets with keywords Ahok, and 100 tweets with keywords Anies. Data taken at random either from a normal user or online media at Twitter.

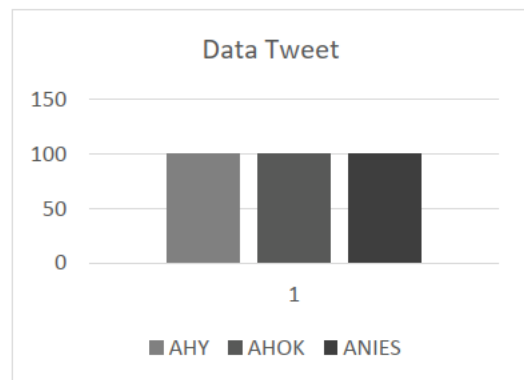


Figure-2 Data tweet

### 2.2. Converts data into a format tweet ARFF

Data tweet collected in text, then converts to ARFF file (Attribute Relation File Format) [12]. To manufacture file ARFF by manual.

### 2.3. Data tweet converted into vector

Data tweet has been shaped ARFF, then converted into a file vector [14]. How to turn data into vector by selecting StringToWordVector in WEKA tool.

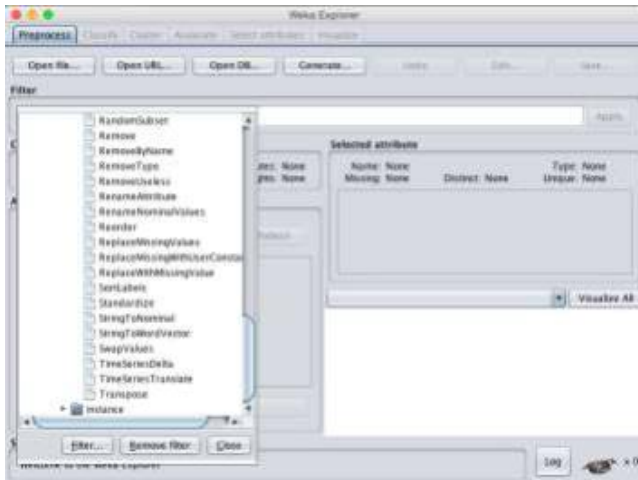


Figure-3 Data tweet converted into vector

The result of the conversion of the ARFF file into vector shapes can be seen in Figure 3.

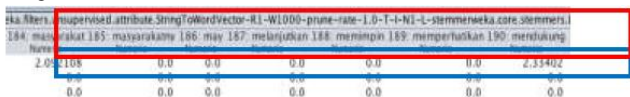


Figure-4 Example vector from tweet data

In Figure 3. The part which inside red box are words that exist in the data tweet. For each row of data representing each tweet. On line 1 the blue box can be seen word " masyarakat" has a value of 2.092108 and the word "mendukung" has a value of 2.33402. While others worth 0.0, it means the word "masyarakat" and "mendukung" contained in first data tweet.

### 2.4. Preprocessing Data

Perform preprocessing data tweet. Preprocessing data include lower case tokens, normalization, tokenization cleansing and filtering. All stages of preprocessing data using WEKA 3.8.1 tools. Stages preprocessing conducted as:

- 1) Lower Case Tokens tweet is to make the data be all lowercase, for example of capital letters to lowercase.
- 2) Normalization is to normalize words that are not standard, for example slang word, Alay word in Indonesian. 3) Tokenization is to break down the tweet into some word or set of words that stand alone. This study uses three methods tokenization, unigram, Bigram, and n-gram with a minimum value n = 1 and maximum n = 3. The process tokenization uses existing menu in WEKA. In tokenizer select tokenization choose and select the method that will be used.
- 3) Cleansing is process removing the symbols of little importance in a tweet, that the data could interfere with the classification process will be. This process is by using the menu on WEKA delimiters.
- 4) Filtering is to remove the words are less important or less affect the classification process will be. This process is by using stop word list. Stop word list in this study is stop word list WEKA and Indonesian stop word list by Tala [13].

### 2.5. Weighting

In next stage is to give weight to each word (term). Weighting is to get the value of a word successfully extracted. The method used for assigning weights in this study is TF-IDF (Term Frequency - Inverse Document Frequency). Because this method works best when combined with the classification method Naive Bayes classifier (NBC).



Figure-5 TF-IDF (Term Frequency – Inverse Document Frequency) Classification

In this study, classification data using WEKA 3.8.1 tools. Classification methods used in this study the Naïve Bayes classifier (NBC). Naïve Bayes Classifier (NBC) is a method of classifier based on Bayesian probability theorem and assuming that each variable X is free (independence). In other words, Naïve Bayesian Classifier (NBC) assuming the presence of an attribute (variable) has nothing to do with being of attributes (variables) to another. Here is the formula.

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

In the process of classification data tested using 10-fold cross validation [15]. The dataset will be divided into two, namely 10 parts by 9/10 parts used for the training process and 1/10 part is used for the testing process. Iteration lasts 10 times with a variety of data, training and testing using a combination of 10 parts data.

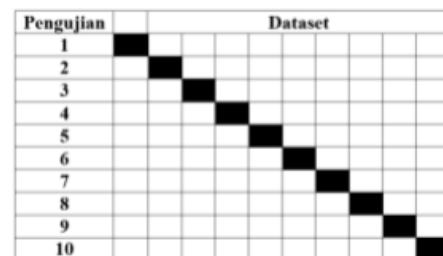


Figure-6 Illustration 10 fold cross validation

## 2.6. Evaluation Results

To evaluate the performance of TP rate, FP rate, Precision, Recall and F-measure of the experiments that have been using. Evaluation using the Confusion Matrix is true positive rate (TP rate), true negative rate (TN rate), false positive rate (FP rate) and false negative rate (FN rate) as an indicator. TP rate is the percentage of positive successful class are classified as positive grade, while TN rate is the percentage of negative class who succeeded classified as negative class. FP rate is negative class are classified as positive class. FN rate is a positive grade classified as negative class [16].

**Table 1. Confusion Matrix.**

		Predicted	
		Negative	Positive
Actual	Negative	a	b
	Positif	c	d

## 3. Result and Discussion

The dataset in this study using ARFF format collected from Twitter. Data taken only tweet in Indonesian. Tweet opinion on the three candidates for governor in 2017 DKI well drawn randomly from a normal user or online media at Twitter.

The dataset used by 300 Tweets, data is split equally (balanced) each class, because the data is not balanced (imbalanced), a classification that is built has a tendency to ignore the minority class [15]. Data is divided into about 100 for AHY tweet, tweet to Ahok 100, and 100 tweets for Anies. Labelling is done manually with the help of experts Indonesian.

**Table 2. Dataset Details**

Sentimen	AHY	Ahok	Anies
Positive	62	33	62
Neutral	25	10	26
Negative	13	57	12

### 3.1. Candidates for Governor AHY

The first experiments with data tweet AHY using the classification method Naïve Bayesian Classifier (NBC) produces Precision, Recall, and F-measure

**Table 3. Result AHY Dataset**

	TP	FP	FN	Precision	Recall	F-measure
Positif	60	2	2	0,952	0,968	0,96
Netral	24	1	1	1	0,96	0,98
Negatif	11	2	2	0,846	0,846	0,846
Average	95	5		0,95	0,95	0,95

From Table 3, we can see Precision of positive sentiment was 95.2% and for Recall was 96.8%. For the neutral sentiment, Precision 100% and Recall

96%. For negative sentiment, Precision and Recall same is 84.6%. In experiments with this data is not much going misclassification, as evidenced by the value of precision and recall is very high.

### 3.2. Candidates for Governor Ahok

The second experiments with the data tweet Ahok using the classification method Naïve Bayesian Classifier (NBC) produces Precision, Recall, and F-measure.

**Table 4. Result AHOK Dataset.**

	TP	FP	FN	Precision	Recall	F-measure
Positif	29	4	4	0,853	0,879	0,866
Netral	4	1	5	0,444	0,4	0,421
Negatif	49	4	4	0,86	0,86	0,86
Average	82	9	13	0,816	0,82	0,818

From Table 4, we can see Precision of positive sentiment was 85.3% and for Recall is 87.9%. For the neutral sentiment, Precision 44.4 % and Recall 40%. For the negative sentiment Precision and Recall same is 86%. In experiments with this data, the error occurred at a neutral sentiment, as evidenced by the more neutral sentiment data is classified into another sentiment. So the value of precision and recall of neutral sentiment is quite small.

### 3.3. Candidates for Governor Anies

The third experiments with data tweet Anies using the classification method Naïve Bayesian Classifier (NBC) produces Precision, Recall, and F-measure.

**Table 5. Result Anies Dataset**

	TP	FP	FN	Precision	Recall	F-measure
Positif	54	1	0	0,885	0,871	0,878
Netral	19	6	1	0,655	0,731	0,691
Negatif	9	1	0	0,9	0,75	0,818
Average	82	8	1	0,827	0,82	0,822

From Table 5, we can see Precision of positive sentiment was 88.5% and for Recall is 87.1%. For the neutral sentiment Precision 65.5% and Recall 73.1%. For the negative sentiment Precision 90% and Recall 75%. Experiment with this data, pretty much an error occurred while neutral sentiment classification, although the value is not more neutral sentiment which was classified as neutral sentiment. So the value of precision and recall of neutral sentiment has been quite high.

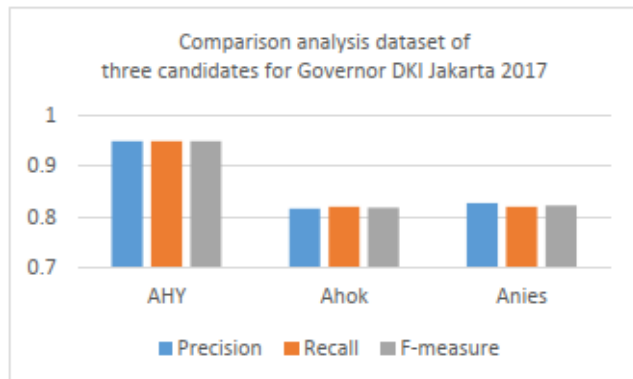
### 3.4. Comparison of the data analysis of three candidates for Governor of DKI 2017

After experiments with the three dataset candidates for governor DKI Jakarta in 2017, the following average values Precision, Recall, and F-measure.

**Table 6. Result Dataset Analysis**

Candidates	Precision	Recall	F-measure
AHY	0,95	0,95	0,95
Ahok	0,816	0,82	0,818
Anies	0,827	0,82	0,822

From Table 6, we can see the highest scored Precision, Recall and F-measure is AHY dataset with Precision 95.2%, Recall 96.8% and F-measure 95%. The Lowest Scored Precision, Recall and F-measure is Ahok dataset with Precision 81.6%, Recall 82% and F-measure 81.8%. Data AHY get the highest score due to positive sentiment on the data AHY most successful classified as positive sentiment. As for the data Ahok be the lowest for the most successful negative sentiment classified as negative sentiment.



**Figure-7** Comparison analysis dataset of three candidates for Governor DKI Jakarta 2017

#### 4. CONCLUSIONS

This study was conducted to test the data tweet to the opinion of three candidates for governor of Jakarta, 2017. Data taken only tweet in Indonesian, which is 100 tweets with keywords AHY, 100 tweets with keywords Ahok, and 100 tweets with keywords Anies. Data taken at random either from a normal user or online media at Twitter. Tweet Indonesian opinion with the three candidates for governor of Jakarta in 2017 and then divided into three sentiments: positive, neutral and negative. Preprocessing Data is Lower Case Tokens, Normalization, tokenization, Cleansing and Filtering. For the method of classification in this study using Naïve Bayes classifier (NBC), because this method is the most powerfully performed for sentiment analysis and proven always produce the highest accuracy. From the results of the classification, the AHY dataset scored the highest with Precision 95% and Recall 95%, while Ahok dataset lowest scores with Precision 81.6% and Recall 82%. In this study also showed that the positive sentiment on the data AHY most successful classified as positive sentiment, as evidenced by the value of precision and recall is very

high. For the most successful negative sentiment classified as negative sentiment contained in the data Ahok. For further research needs to be tested was developed using more data and Real Time. Need to develop also stop word lists and Indonesian stemming were able to improve the accuracy of the Indonesian sentiment analysis.

#### REFERENCES

- [1] KPUD DKI Jakarta (2016) Agenda Pemilihan Gubernur DKI Jakarta <http://kpujakarta.go.id/agenda/>
- [2] Top Media Sosial <http://www.evadollz.com/2014/09/top-10-social-networkings-terpopuler.html>
- [3] Marian Radke Yarrow, John A. Clausen and Paul R. Robbins (2010). The Social Meaning of Mental Illness. *Journal of Social Issues*. Volume 11, Issue 4, pages 33–48, Fall 1955.
- [4] Go, A., Huang, L., & Bhayani, R. (2009). Twitter Sentiment Analysis. Final Project Report, Stanford University, Department of Computer Science.
- [5] Mahyuddin K. M. Nasution. Social Network Mining (SNM): A Definition of Relation between The Resources and SNA. *International Journal on Advanced Science, Engineering and Information Technology*. Vol.6 (2016) No. 6, ISSN: 2088-5334
- [6] Suhaila Zainudin, Dalia Sami Jasim, and Azuraliza Abu Bakar. Comparative Analysis of Data Mining Techniques for Malaysian Rainfall Prediction. *International Journal on Advanced Science, Engineering and Information Technology*. Vol.6 (2016) No. 6, ISSN: 2088-5334
- [7] Merfat M. Altawaier, Sabrina Tiun. Comparison of Machine Learning Approaches on Arabic Twitter Sentiment Analysis. *International Journal on Advanced Science, Engineering and Information Technology*. Vol.6 (2016) No. 6, ISSN: 2088-5334
- [8] Mesut Kaya, Guven Fidan, Ismail H. Toroslu (2012). Sentiment Analysis of Turkish Political News. *IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*.
- [9] Pak, A., dan Paurobek, P., (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining, *Universite de Paris-Sud, Laboratoire LIMSI-CNRS*.
- [10] G. A. Buntoro, (2016). " Sentiment Analysis Candidates of Indonesian Presiden 2014 with Five Class Attribute" in *International Journal of Computer Applications* (0975 – 8887), Volume 136 – No.2, February 2016.
- [11] Franky dan Manurung, R., (2008). Machine Learning-based Sentiment Analysis of Automatic Indonesian Translations of English Movie Reviews. In *Proceedings of the International Conference on Advanced Computational Intelligence and Its Applications*.
- [12] ARFF files from Text Collections. <http://WEKA.wikispaces.com/ARFF+files+from+Text+Collections>.
- [13] Tala, F. Z. (2003). A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia. M.S. thesis. M.Sc. Thesis. Master of Logic Project. Institute for Logic, language and Computation. Universiteti van Amsterdam The Netherlands.
- [14] `ClassStringToWordVector`.
- [15] <http://WEKA.sourceforge.net/doc.de.v/WEKA/filters/unsupervised/attribute/StringToWordVector.html>.
- [16] Ian H. Witten. (2013) *Data Mining with WEKA*. Department of Computer Science University of Waikato New Zealand.
- [17] Kohavi,&Provost.(1998)*ConfusionMatrix* [http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion\\_matrix/confusion\\_matrix.html](http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html)