Information Systems International Conference (ISICO), 2 – 4 December 2013

Twitter Used by Indonesian President: An Sentiment Analysis of Timeline

Paulina Aliandu
Department of Informatics, Faculty of Engineering, Widya Mandira Catholic University

Keywords:

Twitter
Sentiment analysis

ABSTRACT

Humans are social beings and they loved to communicate with each other. With the transformation of communication nowadays people interact using social media such as Twitter. Even Indonesian president and political figures used twitter to express their feelings, emotions (sentiments) and ideas of different aspects every day. We interested to find out the feeling or emotion (sentiment) of Indonesian president based on his timeline. Our application can determine @SBYudhoyono's timeline. We used three classes of sentiment: positive, negative and neutral. Model that used to determine timeline sentiment comes from our previous research. We found a good result with accuracy 79, 42% on collected timeline testing set. By using our application we found that Indonesian president timeline dominated by neutral sentiment.

Corresponding Author:

Paulina Aliandu,
Department of Informatics, Faculty of Engineering,
Widya Mandira Catholic University,
Jalan Jend. Ahmad Yani No. 50-52, Merdeka, Kupang, Indonesia.
Email: paulinaaliandu@gmail.com

1. INTRODUCTION

Twitter has become a phenomenon. It is becoming a largest micro blog with 106 million accounts and 180 million unique users every day in April 2010 (Twitter Chirp Conference 2010) and this statistic were growing. Even the Indonesian president and political figures used twitter to express their feelings, emotions (sentiments) and ideas on different aspects every day.

Twitter also becomes a political phenomenon. Many political revolutions in the last year are initiated by tweet and made several countries in Middle East generated a temporal regulation to control traffic data on Twitter. A number of young social media activist who, for almost two years, exchanged information via Facebook and Twitter also helped to initiate the revolution [1]. There were a number of individual activists with sufficient knowledge of social media resources who helped bring the revolution to life. These activists created Facebook groups, personal blogs and Twitter accounts to engage supporters and followers in discussion on current conditions in Egypt [2]. According to the Pew Internet and American Life Project, six in ten U.S. internet users, nearly 44% of American adults went online to get news or information about politics in 2008. Additionally, Americans are taking an active role in the online political discourse, with 20% of internet users contributing comments or questions about the political process to social networking sites, blogs or other online forums [3].

On the other hand, several political figures, congressman even the president are using Twitter as their social media for campaign, communication media, or news center media. Indonesian President, Susilo Bambang Yudhoyono also used Twitter. His account on Twitter is @SBYudhoyono and tweeted since 13 April 2013. This research is a preliminary research in order to gain timeline sentiment of Indonesian president and how it correlation with others account. In this preliminary research we want to know what the majority sentiment of president timeline is.

Opinion mining or sentiment analysis is a computational research about opinion, sentiment and emotion expressed in text [4]. With the growth of internet using and social networking phenomenon, blogs, online review, (called user generated content) makes data source available for opinion mining become no longer limited. This made opinion mining becoming a popular research topic nowadays in Natural Language Processing and web data mining.

2. RELATED WORKS

Research by [5] used machine learning to classify movie reviews. These researches do sentiment classification on movie review and determine if the review was positive or negative. Movie reviews were extracted different features. The generation of classification model was used Naive Bayes and Support Vector Machine (SVM). They got 78, 7% accuracy when used Naive Bayes on unigram. When used SVM on unigram, the accuracy was 72, 8%.

Research by [6] used emoticons to build a training set of sentiment classification. Read collected text contains emoticons from Usenet Newsgroup. Research by [6] did a classification experiment based on topic dependency. For mixed topic on data training and mixed topic on data testing, [6] got 84, 6% accuracy for Naive Bayes and 81, 1% for SVM. Another classification topic, Read got 81, 5% for SVM and 78, 9% for Naive Bayes when used domain dependency on sentiment classification.

Research by [7] used data source from Pang and Lee's movie review corpus, corpus gathered from blogs, discussion blogs and other websites. [7] used Naive Bayes, SVM and Maximum Entropy for its models. While many of the methods show encouraging results, there are still challenges to be overcome when applying them to data gathered from World Wide Web, especially from blogs. They demonstrated circumstances improvements over states of the arts method for sentiment recognition in texts are possible.

Research of [8] were trying to replicate movies review sentiment classification experiment by [5] for Indonesian. They used machine translation tools to translate English corpus that made [5] into Indonesian. Result of this translation used to train classification. Average accuracy that they got were 74,62% for Naive Bayes and 75,62% for SVM. The best result was the same when experimet were did in English.

Research by [9] proved that machine translation can be used to generate a subjectively-annotated corpus and to effectively train tools for subjective analysis by using SVM and Naive Bayes. [9] used Machine Translation (MT) of manually annotated corpora, MT of source language training, MT of target language training data, MT of targeted language training data, MT of target language test data.

Research of [10] and [11] also used emoticons in order to make data labeled easier on English tweet sentiment analysis. [10] classified tweet sentiment on two classes that are positive and negative sentiment class. The accuracy was 81, 3% when used Naive Bayes, 80, 5% when used Maximum Entropy and 82, 2% when used Maximum Entropy on unigram. [11] labeled corpus with positive, negative and neutral sentiment classes. Training data in neutral class were come from tweets English news media account. They used Naive Bayes with n-grams. The best performance when used bigram.

Research of [12] used emoticon to build Indonesian corpus and used Naive Bayes machine learning with unigram with term frequency (Laplace smoothing) and TF-IDF to build classification model. The research built training corpora with three classes sentiment, there are positive, negative and neutral. The results obtained an accuracy of 77,45% using term frequency with Laplace smoothing and 75,86% using TF-IDF on test set that annotated by emoticons. The results of manually marked test set are 70, 68% for term frequency with Laplace smoothing and 71,26% for TF-IDF. The accuracy measurement also used Support Vector Machine. The results obtained an accuracy of 77, 79% using term frequency and 77,57% using TF-IDF. Accuracy of Support Vector Machine is better than Naive Bayes.

Research by [13] perform a series of 3-class sentiment classification experiments on a set of 2,624 tweets produced during the run-up to the Irish General Elections in February 2011. [13] also introduced a new dataset of political tweets which will be made available for use by other researchers. Each tweet in this set has been annotated for sentiment towards of political entity, as well as for the presence of sarcasm. [13] also omitting the sarcastic tweets from its experiments. The highest accuracy achieved was 61,6% using supervised learning and a feature set consisting of subjectivity-lexicon-based scores.

This research was used the probability model which generated by [12]. The object that wants to analyze is Twitter timeline (especially Indonesian president timeline) that is a long stream showing all tweets. Timeline in Twitter is similar to chronology on Facebook. We built an application that can determine latest timeline sentiment of Indonesian president.

3. RESEARCH METHOD

This research built and designed the system that has ability to determine sentiment on tweet according to a Twitter user's timeline (@SBYudhoyono's timeline). Determining sentiment of Twitter user's timeline resulted from tweet classification process by determining all tweet class. Because we used the probability model which generated by [12] so we just need to obtain real time data from @SBYudhono's timeline based on query user. Then, we can conclude what is the sentiment of @SBYudhoyono's timeline based on real time query. We also collected data from @SBYudhoyono's timeline and annotated them in order to check the accuracy of classifier.

3.1. Preprocessing query on @SBYudhoyono's timeline

The process of obtaining clean data from @SBYudhoyono's timeline as follows:

a. Case folding. This stage changes all the capital letter in tweets becomes lower case.
b. Remove repeated tweet. Sometimes Twitter API returns duplicate tweet.
c. Remove any tweet containing both positive and negative emoticon. This may happen if a tweet contains two subjects.
d. Removed *retweets* (response tweet). "RT" commonly abbreviates retweets. Retweeting is the process of copying another user's tweets and posting to other accounts. This usually happens if a user likes another user's tweet.
e. Tokenization
f. Removed URL link and Twitter username
g. Filtering, by eliminating illegal character in tweet such as %, /,* and so on.
h. Removed any replacing slang words with formal words listed in local dictionary.
i. *Stoplist* removal, by removing character and words listed as *stopwords* or words with high frequency availability (such as "*dan*" and "*yang*').
j. Rearrange the token to become a clean tweet. We will determine the sentiment of the @SBYudhoyono's timeline.

3.2. Accuracy Measurement

We used equation (1) in order to calculate accuracy**.** [14] said that classifier accuracy could be estimate by using equation (1).

$$Accuracy = acc\,(M) = \frac{(true\_pos+true\_neg)}{(all\_pos+all\_neg)} \quad \ldots\ldots\ldots\ldots \quad (1)$$

3.3. Data

Data that used in this research was a timeline of Indonesian president account: @SBYudhoyono. We used probability model which generated by [12]. The object that we want to find its sentiment was timeline of account @SBYudhoyono. President account: @SBYudhoyono tweet since 13 April 2013. The data collection of timeline started from 13April 2013 until 5 Mei 2013 and we have 230 tweets. A collection has been manually annotated. We annotated the sentence that doesn't have sentiment into neutral class.

4. RESULTS AND ANALYSIS

We did some experiments by querying on @SBYudhoyono's timeline by using our application. We also checked the accuracy of the classifier on data testing comes from @SBYudhoyono's timeline.

4.1. Experiment by Real Time Query on Timeline

Real time query on @SBYudhoyono's timeline in 21 Mei 2013 at 08.00 am got 31 tweets and application determined 29 tweets as neutral tweets, 2 tweets as positive tweets and no negative tweets. The results show on Figure 1. The twitter API has a limit of 100 tweets in a respond for any request. The Search API is not complete index of all tweets, but instead an index of recent tweets. At the moment that index includes between 6-9 days of tweets.

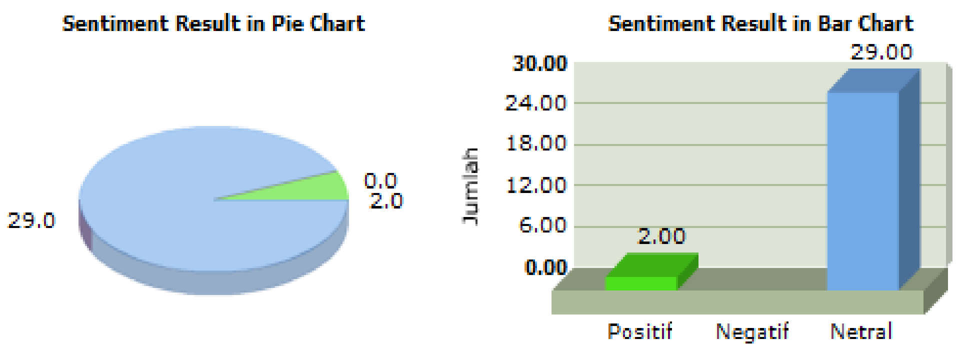

Figure 1. Graphic of @SBYudhoyono's Timeline Sentiment on Application accessed on 21 Mei 2013

4.2. Accuracy on Timeline Data Collection

We have 230 tweets from @SBYudhoyono's timeline and manually annotated. We annotated 60 tweets as positive sentiment, 3 tweets as negative sentiment and 167 tweets as neutral sentiment. Our application determined 56 tweets as positive and 6 tweets as negative and 168 tweets as neutral. We show many way confusion matrix in Table 1. From table 1, we can see that there are 33 positive tweets incorrectly classified into neutral class. Table 1 also show that there are 29 neutral tweets incorrectly classified into positive class. According to my research, tweets that not contain sentiment into neutral class, we consider that [15] declared that it is very confusing even for human to distinguish between neutral opinion and non-opinion bearing sentences.

Table 1. Many ways Confusion Matrix

		Response		
		Positive	Negative	Neutral
Reference	Positive	27	0	33
	Negative	0	0	3
	Neutral	29	6	132

We found accuracy 79, 42% using sum of One vs All Concussion Matrices and 69, 13% using many way confusion matrices. Sum of One vs All Matrices shown in Table 2.

Table 2. Sum of One vs All Matrices

		Response	
		True	False
Reference	True	159	71
	False	71	389

5. CONCLUSION

Model that using Naive Bayes Method that generated by [12] working good to determine on @SBYudhoyono's timeline with accuracy 79,42%. Although tweets have, unique characteristics compared to other corpora, machine-learning algorithms shown to classify tweet sentiment with similar performance.

This research showed that the timeline of @SBYudhoyono dominated by neutral sentiment. In data collection we annotated tweet like: *Bermain volley dengan karyawan Pupuk Kujang dan masyarakat Karawang. pic.twitter.com/yAvPJfw27L* into neutral class. Our suggest it is good if we separate tweets that contain sentiment/opinion and not contain sentiment/opinion. Tweets like We annotated tweet like *Saya menemukan akun @GNFI , senang rasanya membaca hal-hal positif dan prestasi Indonesia. Maju terus GNFI. *SBY** as positive sentiment and *Konflik dan tragedi kemanusiaan di Syria sudah melampaui batasnya. Korban sudah terlalu banyak & terus berjatuhan. *SBY** as negative sentiment. We found that positive tweets most incorrectly classified as neutral sentiment, such as *Saya menemukan akun @GNFI , senang rasanya membaca hal-hal positif dan prestasi Indonesia. Maju terus GNFI.* *SBY*.

We found that sometimes some tweets were the parts of kultwit. Kultwit is a series of successive tweets discussing one particular topic. In that case, sometimes it is hard to determine sentiment of the tweets if we didn't consider the context of the tweets. And it is possible that a single sentence expresses 2 different opinions such as: *The voice quality of this phone is good, but the battery life is short*. It is hard to determine this kind of sentence.

In further research we plan to build a sentence subjectivity classifier for Indonesian. As the future works we plan to build a political corpus in Indonesian and consider what [13] did. By using that political corpus and combining with POS-Tag, hopefully the accuracy will increase. For further research, this application could detect feature of sentiment. It will be more useful if we can determine the feature of sentiment that president mention about.

REFERENCES

[1] D.D. Kirkpatrick and D.E. Sanger, "*A Tunisian-Egyptian link that shook Arab History,* "The New York Times. Retrieved from http://www.nytimes.com/2011/02/14/world/middleeast/14egypt-tunisia-protests.html?_r=1

[2] N. Eltantawy and J. B. Wiest, "Social Media in Egyptian Revolution: Reconsidering Resource Mobilization Theory, "*International Journal of Communication 5,* 2011 High Point University

[3] Pew Internet and American Life Project 2008, "*Social networking and online videos take off : Internet's broader role in campaign 2008*. "Technical report, Pew Research Center.

[4] B. Liu, "Handbook of Natural Language Processing," *chapter Sentiment Analysis*, 2nd Edition, 2010.
[5] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up ? Sentiment classification using machine learning techniques, "*Proceedings of the 7th Conference on Empirical Methods in Natural Language Processing*, 2002.
[6] J. Read, "Using emoticons to reduce dependency in machine learning techniques for sentiment classification," *Proceedings of the ACL Student Research Workshop*, 2005.
[7] E. Boiy, P. Hens, K. Deschacht and M. Moens, "Automatic Sentiment Analysis in On-line Text, " *Proceedings of the Conference on Electronic Publishing (ELPUB)*, 2007
[8] Frangky and R. Manurung, "Machine learning based sentiment analysis of automatic Indonesian translations of English movie reviews," *Proceedings of the International Conference on Advanced Computational Intelligence and Its Application*, 2008.
[9] C. Banea, R. Mihalcea, J. Wiebe and S. Hassan, "Multilingual Subjectivity Analysis Using Machine Translation," *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, 2008.
[10] A. Go, R. Bhayani and L. Huang, "Twitter Sentiment Classification using Distant Supervision," *CS224N Project Report*, Stanford, 2009.
[11] A. Pak and P. Paurobek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," *Proceedings of the International* Conference *on Language Resources and Evaluation*, 2010.
[12] P. Aliandu, "Sentiment Analysis on Indonesian Tweet," *Proceedings of the 7^{th} International Conference on Information and Communication Technology Systems*, 2013.
[13] A. Bakliwal, J.Foster, J. van der Puil, R. O'Brien, L. Tounsi and M. Hughes, "Sentiment Analysis of Political Tweets : Towards an Accurate Classifier, " *Proceedings of the Workshop on Language in Social Media (LASM 2013)*, 2013.
[14] D. Larose, *Data Mining Methods and Models*, John Wiley & Sons, Inc. New Jersey, USA, 2006.
[15] S. Kim and E. Hovy, "Determining the Sentiment of Opinions," *Proceedings of the COLING Conference*, Geneva, 2004.

BIBLIOGRAPHY OF AUTHORS

	Graduated from Sekolah Tinggi Teknologi Telkom Bandung majoring in Industrial Engineering in 2003 and Finished Master Degree in Computer Science in Universitas Gadjah Mada in 2012. Interested in text processing, text mining, social networking and artificial intelligent especially sentiment analysis.