TELKOMNIKA, Vol.15, No.3, September 2017, pp. 1354~1361
ISSN: 1693-6930, accredited **A** by DIKTI, Decree No: 58/DIKTI/Kep/2013
DOI: 10.12928/TELKOMNIKA.v15i3.4284

Twitter's Sentiment Analysis on Gsm Services using Multinomial Naïve Bayes

Aisah Rini Susanti[1], Taufik Djatna*[2], Wisnu Ananta Kusuma[3]
[1,3]Computer Science, Mathematic and Natural Science, Bogor Agricultural University, Campuss IPB
Darmaga P.O. Box 220 Bogor, Indonesia. Tel: (0251) 86228448, Fax: (0251) 8622986
[2]Dept. Agro-industrial Technology, Bogor Agricultural University, Campuss IPB Darmaga P.O. Box 220
Bogor, Indonesia (Tel: (0251) 86228448, Fax: (0251) 8622986
*****Corresponding author, e-mail: taufikdjatna@ipb.ac.id

Abstract

Telecommunication users are rapidly growing each year. As people keep demanding a better service level of Short Message Service (SMS), telephone or data use, service providers compete to attract their customer, while customer feedbacks in some platforms, for example Twitter, are their souce of information. Multinomial Naïve Bayes Tree, adapted from the method of Multinomial Naïve Bayes and Decision Tree, is one technique in data mining used to classify the raw data or feedback from customers.Multinomial Naïve Bayes method used specifically addressing frequency in the text of the sentence or document. Documents used in this study are comments of Twitter users on the GSM telecommunications provider in Indonesia.This research employed Multinomial Naïve Bayes Tree classification technique to categorize customers sentiment opinion towards telecommunication providers in Indonesia. Sentiment analysis only included the class of positive, negative and neutral. This research generated a Decision Tree roots in the feature "aktif" in which the probability of the feature "aktif" was from positive class in Multinomial Naive Bayes method. The evaluation showed that the highest accuracy of classification using Multinomial Naïve Bayes Tree (MNBTree) method was 16.26% using 145 features. Moreover, the Multinomial Naïve Bayes (MNB) yielded the highest accuracy of 73,15% by using all dataset of 1665 features. The expected benefits in this research are that the Indonesian telecommunications provider can evaluate the performance and services to reach customer satisfaction of various needs.

Keywords: *Indonesian telecommunication service provider, Multinomial Naïve Bayes, sentiment analysis, service performance, Twitter*

1. Introduction

Sentiment analysis is a technique to evaluate and identify either positive or negative emotions or opinions [1]. Sentiment analysis has been widely explored on documents with Twitter being one popular social media where users can express their opinion objectively about a broad range of topics [2], while 19% of them expressed their comment to brand and products [3], as well their emotion to cellular operator with an accuracy of 80% in predicting the sentiment 80% [4]. Cellular operator companies (provider) are those provide the service of telecommunication, particularly in GSM (*Global System for Mobile communications*) service are: Telkomsel (PT. Telekomunikasi Seluler), Indosat Ooredoo (PT. Satelit Indonesia / Satelindo), XL Axiata (PT XL Axiata Tbk) and Hutchison (PT. Hutchison CP Telecommunications Indonesia/ HCPT), with their different products such as Simpati and Halo (Telkomsel), IM3 and Mentari (Indosat Ooredoo), XL (XL Axiata) and Tri (Hutchison). Telecommunication users are growing each year as the demand of better service level of Short Message Service (SMS), telephone or data use, is also emerging. Consequently, service providers are competing to attract or maintain their customers, while the customers can express their feedback through platform like Twitter.

Naïve Bayes is a method to classify data, is an algorithm of inductive study for the most effective and efficient machine learning and data mining [5]. In the practice, it assumed that features are independent, while in fact each feature may have relation or dependency [6, 7]. Thus, a method of NaiveBayes Tree was proposed as the integration of Naïve Bayes method and Decision Tree method. The basic concept of the decision tree is to convert the data into a tree and decision rules [8]. Naïve Bayes Tree is effectively capable of reducing computation

Received May 15, 2017; Revised July 16, 2017; Accepted August 1, 2017

time by diminishing redundancy in the data, resulting better accuracy compared to Naïve Bayes or Decision Tree method solely [9]. Adapting from the workflow for text classification in Naïve Bayes Tree method, Multinomial Naïve Bayes Tree (MNBTree) was resolved showing a better accuracy compared to Naïve Bayes Tree [10]. Raw data processing from Twitter requires a data preprocessing phase to generate the “standard words” as the feature to result the sentiment analysis [11]. The research on sentiment analysis has been widely conducted, for example in the field of politics [12], economy [13] and product quality survey [14]. Every consumers sentiment in Twitter about a product reflects the quality of the product, as well the sentiment towards service providers about the telecommunication service.

This research about Twitter sentiment analysis employed the algorithm of NaiveBayesto detect the polarity of English tweet to reveal the best performance using biner classifier between two categories of contrast polarity: positiveand negative [15]. Other work with Naïve Bayes method to analyze the sentiment in Twitter about cellular operator in Indonesia had an accuracy of 72,22 % [4], Analysis of the mobile phone service provider quality in social media Twitter using Naïve Bayes shows the provider with the highest customer satisfaction level [16]. Yu and Hatzivassiloglou 2003 reported an accuracy of 97% in document classification of data acquisition of 400 sentences [17] dan memiliki and reported a good result for sentiment data classification using N-gram dan POS-tag as the features [18]. Thus, Naïve Bayes is an accurate, efficient and easily interpreted method of classification [19, 20].

A variant of Naïve Bayes method to administer Multinomial data in text classification is Naïve BayesMultinomial method. Multinomial model yields better accuracy compared to Multi-variate Bernoulli method in text classification with high number of vocabulary [21]. Multinomial Naïve Bayes method is a Naïve Bayes to estimate the frequency of term in a document. Substantially, MNBTree method stamdard can be improved by applying the transformation of TF-IDF (*Term Frequency-Inverse Document Frequency*)for feature and vector normalization with the rate of vector lentght observed in the data [22].

2. Research Method

This research was conducted with the step of data preparation, Twitter data preprocessing and sentiment classification modeling as shown in Figure 1. Data preparation was done by connecting with API (*Application Programming Interface*). The preprocessing comprised the conversion of tweet to lowercase words and removing the tweet from: ReTweet, user ID, punctuation, number, website link, stopwords, stemming, normalization and labeling. After data preprocessing, the data was transformed into matrix document which was the terms representation and frequency in a document. The final step consisted of the model construction of Multinomial Naïve Bayes using *Weka 3.6,* Multinomial Naïve Bayes Tree modeling using *Netbeans IDE 8.0.2*, and the evaluation of model.

2.1. Data Preparation

Twitter provides Application Programming Interface (API) to enable tweets acquisition by third party, which is free feature for a sample of 1% of all tweets [23]. The data used in this research was the tweet of comment in Bahasa, acquired from API on Twitter user status in Indonesian about telecommunication provider in Indonesia.

2.2. Preprocessing

After storing the tweets in data storage, the preprocessing was conducted in the following steps:

1. Converting tweet to lowercase and removing the following from the tweet: ReTweet, user ID, punctuation, number, website link, stopwords, stemming, normalization and labeling. Stemming has been commonly used by some researchers in natural language processing area such as text mining, text classification, and information retrieval [24]. Stemmingis transforming the word into its root word, while normalization is transforming words into the desired words, in this case was the casual words and local language into Bahasa. The result then was stored into a file. This stage resulted keywords of positive and negative sentiments (Table 1) which later were stored in a corpus and used for labelling for each tweet.

2. Labeling using lexicon dictionary and corpus as described in Table 1. Lexicon dictionary used was in English [25] which includes positive and negative sentiments and then translated into Bahasa using Google Translate.
3. Conducting tokenization, which is breaking down a sentence into root words and transforming into matrix using R programming.

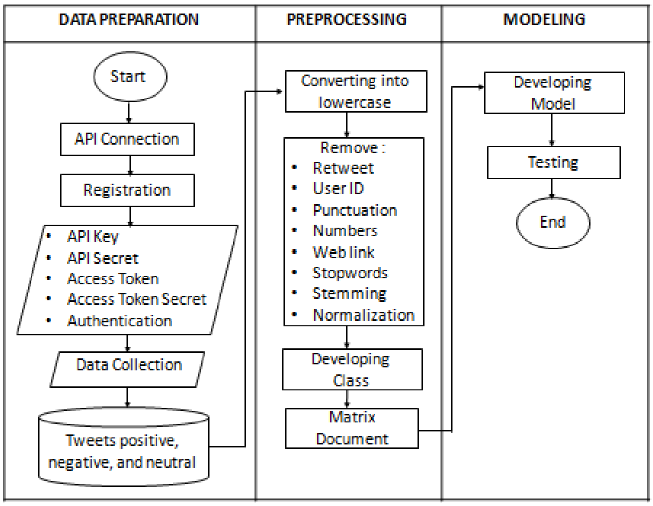

Figure 1. Research Framework

Preprocessing was conducted using R programming. Labeling was conducted by subtracting total score of positive words with total score of negative words. Next, the data is classified into training and testing. Training data was used to provide information about the rule or pattern of a class. The MNBTree method was performed using Netbeans IDE 8.0.2 for training data and testing data.

Table 1. The features in the corpus keywords positive and negative sentiment

Positive				Negative		
active	fast	believe	loyal	wait	complaint	down
good	support	congratulation	succesful	nul	loose	headache
help	easy	spirit	thank you	difficult	obstacle	fake
able	gift	happy	champion	far	error	change
can	suprise	smile	beneficial	off	run out of	problem
bonus	cool	fun	win	bad	disturbance	out
			okay			

2.3. Multinomial Naive Bayes method

Multinomial Naïve Bayes is a Naïve Bayes algorithm which administers Multinomial data in text classification. Data in Multinomial Naïve Bayes is represented as the total of data vector, thus Multinomial Naïve Bayes is appropriate for estimating term frequency in a document. In Multinomial Naïve Bayes, firstly the probability of words in a class (prior) was computed as Equation 1 [26]:

$$P(c|d) \propto P(c) \prod_{1\leq k\leq n_d} P(t_k|c) \qquad (1)$$

where $P(t_k|c)$ is the conditional probability of term $t_k$occurring in a document *d* of a class$c$. $P(c)$is the prior probability of a document occurring in class *c*. The probability of document *d* inc was perfomed using Equation 2 [26]:

$$\hat{P}(c) = \frac{N_C}{N} \tag{2}$$

N_C is the number of documents in class $c$and$N$is the total number of documents. The conditional probability $P(t|c)$ as the relative frequency of term$t$ in documents belonging to class c, as in Equation 3 [26]:

$$\hat{P}(t|c) = \frac{T_{ct}}{\sum_{t'\in V} T_{ct'}{}'} \tag{3}$$

T_{ct} is the number of occurences of term $t$in training documents from class c. $\sum_{t'\in V} T_{ct'}{}'$is the number of all terms in the whole document in class c including redundant term in the same document. The sparseness of term in the document resulted the estimationof frequency $P(t_k|c)$w as zero (0),thus we added *one* or *laplace smoothing*[26] as Equation 4:

$$\hat{P}(t|c) = \frac{T_{ct}+1}{\sum_{t'\in V}(T_{ct'}+1)} = \frac{T_{ct}+1}{(\sum_{t'\in V} T_{ct'})+B'} \tag{4}$$

Where B = | V | is the number of terms in the vocabulary in training data. The algorithm of Multinomial Naïve Bayes for training step is as follows [26]:

```
Multinomial NB Training (C,D)
V ← extract vocabulary (D)
N ← count documents (D)
for each c ∈ C
    do $N_c$ ← count documents in class (D, c)
       prior [c] ← $N_c$ / N
       $text_c$ ← concatenate text of all documents in class (D, c)
       for each t ∈ V
           do $T_{ct}$ ← count tokens of term ($text_c$, t)
       for each t ∈ V
           do condprob[t][c] ← $\frac{T_{ct}+1}{\sum_{t'}(T_{ct'}+1)}$
return V, prior, condprob
```

Training process in Multinomial Naïve Bayes algorithm is the step of the training the stage of data whereas the training sessions to the data describing that V is the number of terms in the vocabulary in the training data then *N* is the number of times of data that sentences Twitter comments. For each class of data is calculated on the amount of data in each class(N_C) divided by the large number of data ($\frac{N_C}{N}$) as listed in the Equation 2 and then for each text feature (*t*) in Vcalculated the odds on each class using the Equation 3. If the conditions are not zero value then do laplace smoothing salty Equation 4. the next step is to apply the testing stages as follows [26]:

```
Apply Multinomial NB (C, V, prior, condprob, d)
W ← extract all tokens from document (V, d)
for each c ∈ C
    do score [c] ← log prior [c]
       for each t ∈ W
           do score [c] + = log condprob [t][c]
return $\mathrm{argmax}_{c\in C}$ score[c]
```

At the stage of testing applied to the testing data for each class by calculating a score using Equation 2, then for each text features were calculated using Equation 3. If there are features worth zero then applied to the Equation 4. Application testing phase resulted in the probability of each class so the highest probability value is the winner of the class that is the biggest opportunity in each document.

2.4. Multinomial Naïve Bayes Tree Method (MNB Tree)

Multinomial Naïve Bayes Tree method develops the text classification in naïve BayesMultinomial for every node on Decision Tree. Multinomial Naïve Bayes Tree method was inspired by Naïve Bayes Tree method which has a high complexity of time in the steps. In the practice, Multinomial Naïve Bayes Tree algorithm is an adaptation from Decision Tree method and Multinomial Naïve Bayes method. Decision Tree is a method for data classification generating a tree-like structure, consisting the root (root node) and leaf (leaf node). Decision Tree method is able to administer categoric and numeric data. In the technique, Multinomial Naïve Bayes Tree developed a binary tree, where the attribute score was determined as zero and nonzero, with a measure of information gain in developing the tree to be efficient in time.

Algorithm: MNBTree Method(D) [10]

```
Input: a training instances set D
Output: The learnedMNBTree text classification
1. Set the minimum size set/to | $D$ | * 40%
2. If | $D$ | is less than/then create a leaf node and build a MNB using the instances falling into this leaf node, and then return
3. To each attribute $W_i$, use Equation 5 to get its information gain [9]:
```

$$Gain(D, w_i) = Entropy(D) - \sum_{V \in \{0, \bar{0}\}} \frac{|D_v|}{|D|} Entropy(D_v) \qquad (5)$$

Where *D* set of documents or data, $w_i$is feature is that the form of each word in the document, $|D_v|$ is the number of instances whose in the partition*v*whose while$|D|$is documents number in the set of documents.

```
4. Let $W_{max}$be the attribute with maximum information gain$G_{max}$
5. If $G_{max}$=0 then create a leaf node and build a MNB using the instances falling into this leaf node, and then return
6. Else for instance $d$ in $D$
    (A) Let $Vw_{max}(d)$ be the value of $W_{max}$in $d$
    (B) If $Vw_{max}(d)$ = 0 then assign $d$ to the left child $Dl$
    (C) Else assign$d$to the right child$Dr$
7. Let $D$ = $Dl$and goto step 2
8. Let $D$ = $Dr$and goto step 2
9. Return the learned MNBTree Text Classifier
```

where | *Dv* | is the number of instances whose value of the attribute *wi*is*v* (*v*∈ {0, $\bar{0}$}), Entropy (*D*) is the entropy of*D*, which can be estimated by Equation 6 [10].

$$Entropy(D) = -\sum_{c \in C} P(c) * log P(c) \qquad (6)$$

Where D is the set of documents while *P (c)* is the probability of class *c* in*D*.

If the value of the resulting information gain is zero, or the number of features less than 40% of the amount of data (documents), the counting is done by using Mutinomial Naïve Bayes. The minimum value of the leaf (l*eaf*) that *l* have empirical value | D | * 40% as the minimum size of leaves in an effort to reduce the consumption of time, overfitting, and reduce the complexity of leaf nodes on the training data [10]. Minimal size of the leaves affects the size of the tree is built.

Having in mind the minimum size of the leaves are then calculated for each attribute information gain W_i. Where W_i is the attribute that is said to feature *i* calculated using Equation 5. Next set W_{max} be an attribute with the highest information value gain (G_{max}). If G_{max} is zero used the left side branch (child left) is D*l* were calculated using the MNB but if G_{max}worth nonzero used right side branches (child right) are D*r*. were calculated using the method Decision Tree.

Table 2. Confusion Matrix

Actualitation	Prediction Class	
	Class = yes	Class = no
Class = yes	tp	fn
Class = no	fp	tn

2.4. Evaluation

This research employed Confusion Matrix to estimate the model accuracy*.* The measure and formula used in Confusion Matrix is as shown in Table 2[27]. TP (True Positive) is the label for the data similar to model prediction result, while TN *(*True Negative*)* is label for data different from prediction result.FP is False Positive*,* while FN is False Negative.

Table 3. The results of the evaluation of the accuracy of the classification model Twitter sentiment opinion with Multinomial Naïve Bayes Tree

Feature	145	181	231	381	1665
K-*Fold* 1	23.32 %	17.99 %	19.94 %	15.17 %	15.07 %
K-*Fold* 2	18.67 %	13.69 %	21.52 %	14.78 %	16.26 %
K-*Fold* 3	22.25 %	19.07 %	19.53 %	13.36 %	14.65 %
K-*Fold* 4	21.75 %	14.21 %	16.48 %	13.12 %	9.57 %
K-*Fold* 5	22.27 %	14.21 %	16.98 %	12.97 %	13.74 %
Average	21.65 %	15.83 %	18.89 %	13.88 %	13.85 %

3. Results and Analysis

3.1 Data Preparation

In the first step, connection to API (Application Programming Interface) was conducted by firstly registering to sign up an account for acquiring data from Twitter, particularly for obtaining API Key, API Secret, Access Token, Access Token Secret. After acquiring API Key, API Secret, Access Token, Access Token Secret, the next step is conducting authentification and registration. This resulted the data in form of Twitter users' comment. Then, the data was retrieved based on the desired keywords. Those data were stored in a frame in R programming and then saved in the file with CSV format (Comma Delimited Fitur). The data included were: *text, favorited, favoriteCount, replyToSN, created, truncated, replyToSID, id, replyToUID, statusSource, screenname, retweetCount, isRetweet, retweeted, longitude* and *latitude.* The feature used in this research was limited to the text column which was the content of the user's comment.

The input in this work was tweet (comment) of Twitter user in Bahasa acquired from API on the status of Twitter user using Bahasa about telecommunication provider in Indonesia. The official account of the providers are @Telkomsel, @kartuas, @simpati (Telkomsel); @Indosat, @Indosatcare, @Indosatmania (Indosat); and @XL, @XLandme, @XLcare (XL). Data preparation yielded 5210 commentsconsisted of the ads from provider and comments from Twitter user. The data mining was conducted from January 3, 2016 to January 5, 2016.

3.2. Multinomial Naïve Bayes Tree method

After the data was input to Multinomial Naïve Bayes Tree method, it generated root with highest information gain of the word **"**aktif" word left child "untung" and right child "negatif" class, while generating 22 Leaf.

The first step is to exercise the data to the Decision Tree method.The number of zero and nonzero value in each feature of the document was accumulated. Next, the probability of the positive, negative and neutral class with zero value as well as the entropy of each terms in the document was estimated.Then, the information gain of each feature *(*W_i*)* in the document was figured using Equation 5, where we obtained the highest value of information gain was 0.035136 of the feature "aktif" *(*W_{max}*).* The feature with highest information gain was assigned as the root of Decision Tree *(*G_{max}*).*

The feature "aktif" was assigned as the root of Decision Tree. After that, iterations were simulated to generate the leaves of feature with zero value (581 features) which then assigned as child left *Dl*, using Multinomial Naïve Bayesresulted negative class had the highest probability of 2.718463. Meanwhile, the feature with nonzero value was assigned as child right

Dr (as much as 4601 documents), estimating the measure of highest information gain yielded the feature “untung” with the value of 0.025782.

3.3. Evaluation

The accuracy of the Twitter sentiment classification model using MNBTree method is shown as in Table 3. Based on Table 3, MNBTree model was evaluated with 5 runs of cross-validation consisting 80% of training data and 20% testing data resulting highest accuracy of 21.652% in 145 features. Meanwhile, for Multinomial Naïve Bayes model had a highest accuracy of 73.15% in 1665 features. Table 4 shows the accuracy comparison between Multinomial Naive Bayes and MNBTree. The amount of feature was simulated from 145 to 1665 to reveal the best highest accuracy, whereas the determination was not based on any method. Based on Table 4, it was concluded that MNB method accuracy were increasing as the increase of the number of features, while MNBTree accuracy tended to increase when the feature was decreasing.

Table 4. The results of the evaluation of the accuracy of the classification model Twitter sentiment opinion with Multinomial Naïve Bayes and Multinomial Naïve Bayes Tree

Feature	145	181	231	381	1665
MNB	56.63 %	66.19 %	67.19 %	69.91 %	73.15 %
MNBTree	21.65 %	15.83 %	18.89 %	13.88 %	13.85 %

4. Conclusion

Based on Twitter sentiment analysis result on the telecommunication provider service performance using Multinomial Naïve Bayes Tree, the root with highest information gain was of the feature “aktif”, in which the probability of the feature "aktif" was from positive class in Multinomial Naive Bayes method. The evaluation showed that the highest accuracy of classification using Multinomial Naïve Bayes Tree (MNBTree) method was 16.26% using 145 features. Moreover, the Multinomial Naïve Bayes (MNB) yielded the highest accuracy of 73,15% by using all dataset of 1665 features. The result of this research of sentiment analysis in Twitter comment using MNBTree method is a decision tere classification model of Twitter users comment as data to telecommunication service provider in Indonesia.For further research, a relevant topic will be the essential of using *Lexicon*dictionary in proper and formal Bahasa as well as the addition of normalization dictionary of casual terms and local language, while the research itself can cover more than one class category in labeling.

References

[1] Wilson TA, Wiebe J, Hoffmann P. Recognizing Contextual Polarity: an exploration of features for phrase-level sentiment analysis. *Computational Linguistics.* 2009; 35(3): 399–433.

[2] Coletta LFS, De Silva NFF, Hruschka ER, Hruschka ER. *Combining classification and clustering for tweet sentiment analysis.* Proceedings-2014 Brazilian Conference on Intelligent Systems (BRACIS). 2014; 210–215.

[3] Jansen BJ, Zhang M, Sobel K, Chowdury A. Twitter power: *Tweet*s as electronic word of mouth. *Journal of the American Society for Information Science and Technolog*y. 2009; 60(11), 2169.

[4] Wijaya H, Erwin A, SoetomoA, Galinium M. *Twitter Sentiment Analysis and Insight for Indonesian Mobile Operators.* Information Systems International Conference (ISICO). 2013; 367.

[5] Zhang H. *The Optimality of Naive Bayes.* Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference FLAIRS*.* 2004; 1(2): 1–6.

[6] Domingos P, Pazzani M. On the Optimality of the Simple Bayesian Classifier under Zero-One Los. *Machine Learning.* 1997; 29: 103–130.

[7] Zheng F, Webb G. *A comparative study of semi-naive Bayes methods in classification learning*. Proceeding 4th Australasian Data Mining Conference*.* 2005; *DM05*(1): 141–156.

[8] Thariqa P, Sitanggang IS, Syaufina L. Comparative Analysis of Spatial Decision Tree Algorithms for Burned Area of Peatland in Rokan Hilir Riau. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2016; 14(2): 684–691.

[9] Veeraswamy A, Alias SA, E Kannan P. An Implementation of Efficient Datamining Classification Algorithm using Nbtree. *International Journal of Computer Applications*. 2013; 67(12), 26–29.

[10] Wang S, Jiang L, Li C. Adapting naive Bayes tree for text classification. *Knowledge and Information Systems*. 2014.

[11] Aziz A. Sistem Pengklasifikasian Entitas Pada Pesan Twitter Menggunakan Ekspresi Regular Dan Naïve Bayes. Bogor (ID): Institut Pertanian Bogor. 2013.

[12] DiGrazia J, McKelvey K, Bollen J, Rojas F. More *tweet*s, more votes: Social media as a quantitative indicator of political behavior. *PLoS ONE.* 2013; *8*(11).

[13] Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. *Journal of Computational Science.* 2011;*2*(1) 1–8.

[14] Chamlertwat W, Bhattarakosol P. Discovering Consumer Insight from Twitter via Sentiment Analysis. *J Ucs.* 2012; 18(8): 973–992.

[15] Gamallo P, Garcia M, Technology CL. *Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets.* Proceedings of the 8th International Workshop on. 2014.

[16] Yu H, Hatzivassiloglou V. Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences. *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*.2003;129–136.

[17] Calvin, & Setiawan, J. Using Text Mining to Analyze Mobile Phone Provider Service Quality (Case Study: Social Media Twitter). *International Journal of Machine Learning and Computing.* 2014; 4(1) 106–109.

[18] Pak A, P Paroubek, D Paris-sud, L Limsi-cnrs, FO Cedex. Twitter Based System : Using Twitter for Disambiguating Sentiment Ambiguous Adjectives. *Comput Linguist.* 2010; 436–439.

[19] Wu X, Kumar V, Ross QJ, Ghosh J, Yang Q, Motoda H, Steinberg D. Top 10 algorithms in data mining. *Knowledge and Information Systems*. 2008; 14(1): 1–37.

[20] Rennie JDM, Shih L, Teevan J, Karger DR. *Tackling the Poor Assumptions of Naive Bayes Text Classifiers*. Proceedings of the Twentieth International Conference on Machine Learning (ICML). 2003; vol 20 no. 1973 pp 616–623.

[21] McCallum A, Nigam K. A Comparison of Event Models for Naive Bayes Text Classification. *AAAI/ICML-98 Workshop on Learning for Text Categorization*. 1998; 41–48.

[22] Kibriya A, Frank E, Pfahringer B, Holmes G. Multinomial Naive Bayes for Text Categorization Revisited. *In AI 2004: Advances in Artificial Intelligence*, 2005;488–499.

[23] Hawwash B. From Tweets to Stories : Using Stream-Dashboard to weave the twitter data stream into dynamic cluster models. 2014;182–197.

[24] Hidayatullah AF. The Influence of Stemming on Indonesian Tweet Sentiment Analysis. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2015; vol. 14, no. August, pp. 19–20.

[25] Liu B, Street SM. *Opinion Observer : Analyzing and Comparing Opinions on the Web*. Proceedings of the 14th International Conference on World Wide Web. 2005; 342–351.

[26] Manning CD, Ragahvan P, Schutze H. An Introduction to Information Retrieval. *Information Retrieval*. 2009; 253–270.

[27] Han J, Pei J, Kamber M. Data Mining : Concepts and Techniques. Journal of Chemical Information and Modeling. 2012; 3. 364-369.