A text mining application of emotion classifications of Twitter's users using Naïve Bayes method

Conference Paper · November 2015

DOI: 10.1109/ICWT.2015.7449218

Liza Wikarsa
Universitas Katolik De La Salle Manado

A Text Mining Application of Emotion Classifications of Twitter's Users Using Naïve Bayes Method

Liza Wikarsa, BCS, MComp
Informatics Engineering
University of De La Salle Manado
Manado, Indonesia
lwikarsa@unikadelasalle.ac.id

Sherly Novianti Thahir, ST
Informatics Engineering
University of De La Salle Manado
Manado, Indonesia
sherlythahir@gmail.com

Dr. Ir. Rila Mandala, MEng
The School of Electrical and Informatics Engineering
Institut Teknologi Bandung
Bandung, Indonesia
rila@informatika.org

***Abstract*— Twitter is one of social media with more than 500 million users and 400 million tweets per day. In any written tweet of twitter users it contains various emotions. Most research on the use of social media classifies sentiments into three categories that are positive, negative, and neutral. However, none of these studies has developed an application that can detect user emotions in the social media, particularly on Twitter. Hence, this research developed a text mining application to detect emotions of Twitter users that are classified into six emotions, namely happiness, sadness, anger, disgust, fear, and surprise. Three main phases of the text mining utilized in this application were preprocessing, processing, and validation. Activities conducted in the preprocessing phase were case folding, cleansing, stop-word removal, emoticons conversion, negation conversion, and tokenization to the training data and the test data based on the sentiment analysis that performed morphological analysis to build several models. In the processing phase, it performed weighting and classification using the Naïve Bayes algorithm on the validated model. The process for measuring the level of accuracy generated by the application using 10-fold cross validation was done in the validation phase. The findings showed that this application is able to achieve 83% accuracy for 105 tweets. In order to get a higher accuracy, one requires a better model in training data.**

Keywords—text mining; Twitter; emotion; classification; naïve bayes.

I. INTRODUCTION (*HEADING 1*)

Social networking websites, like Twitter and Facebook, create enormous opportunities for users to communicate with one another without having to worry about differences in moral and social values. They also enable mutual learning and sharing of valuable knowledge with no regard to geographical distance, time barrier, and language skills. People thus join and engage in various communities and discussion groups that best suit their needs.

Twitter has at least 500 million users and 400 million tweets posted in its site every day [1]. Tweets are written messages in the form of texts that contain opinions, expressions, and emotions of users. Data in Twitter's site is inherently unstructured because users do not care about spelling and grammatical construction of sentences when posting their tweets. With a large amount of user-generated data on Twitter every day, extracting logical emotional patterns of this data with accurate information from such unstructured form is considered as a critical task to perform.

Text mining can be used to overcome this problem as it provides computational intelligences of multidisciplinary disciplines like information retrieval, artificial intelligence, statistics, database systems, and others [2]. Application of text mining techniques on social networking sites can further reveal results related to human thinking patterns, group identification and recommendation, and also opinion about any specific topics of interests. In the previous studies, many researchers classified emotions into two classes that are 'positive and negative' or 'happy and unhappy'. Hence, this excludes user's other basic emotions like anger, fear, disgust and surprise [3]. Furthermore, it was noticeable that those existing techniques of emotion prediction in text just work on articles which contain only the direct emotional words and neutral words. In addition, the existing research papers do not particularly target the datasets in social networking sites and they hardly mention the pre-processing phase that is important to simplify the text mining process [4].

Based on the problems mentioned above, this study will try to develop an application that is able to extract data from Twitter and then classify the data into six categories of emotions, such as happiness, sadness, fear, anger, surprise, and disgust, using Naïve Bayes classification method [5]. Classification is "the process of learning a set of rules from a set of examples in a training set. Text classification is a mining method that classifies each text to a certain category." [2, p. 5]. This study is divided into three phases to extract information from datasets and transform it into an understandable structure for further use. The three phases are pre-processing phase, text mining phase, and results validation phase that help this study to conduct a thorough analysis on how to discover groups and structures in the data set.

II. RESEARCH QUESTION AND OBJECTIVES

A. Research Question

How to build a text mining application that can classify emotions of Twitter's users into six categories, such as happiness, sadness, fear, anger, surprise, and disgust, using Naïve Bayes classification?

B. Research Objectives

Objectives of this research are as follows:

1. To be able to extract data from Twitter that is large, unstructured, and dynamic.
2. To organize the collected data into pre-defined categories that will be used for text analysis by performing pre-processing process.
3. To construct suitable models based on the datasets by assessing how well the model fits, adjusting some models, and finally selecting the best model from among those that have been tried.
4. To present the classified data accordingly to the emotion categories in a useful format, such as a graph or table.

III. LITERATURE REVIEW

A. Definition of Text Mining

According to Berry and Kogan, text mining is a technique used to discover textual patterns from unstructured, big, and dynamic data [6]. Text mining focuses on categorizing texts, grouping texts, extracting concepts, and reducing texts. Text mining is an extension of data mining technique. Irfan *et al.* pointed out that data mining is a technique to extract logical patterns from structured database [2].

B. Phases in Text Mining

As an extension of the data mining technique, text mining uses phases in data mining to extract unstructured text that can be further used for finding opinion about certain subjects, patterns, and other purposes. The phases are as follows:

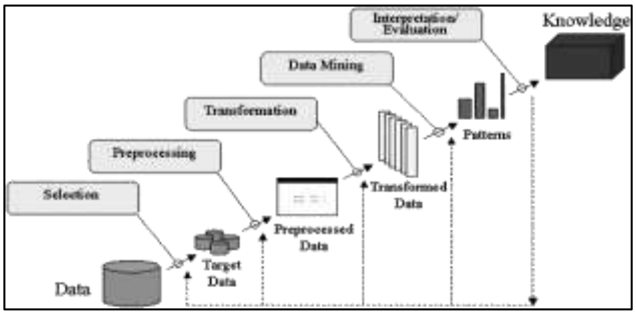

Figure 1. Phases in Data Mining [7]

- Data selection

  Data selection from the collection of operational data needs to be done before the stage of extracting information in Knowledge Discovery in Databases (KDD) begins. Data resulted from the selection in the data mining process, is stored in a file, separated from operational databases.

- Pre-processing/cleaning

  Before the process of data mining can be implemented, it is necessary to perform data cleansing that is the focus of KDD. Cleaning processes include checking inconsistent data, removing duplication of data, and correcting errors in the data like printing errors.

- Transformation

  Is the process of transformation on the data that has been selected to fit the process of data mining. The coding process in KDD is considered as a creative process and is highly dependent on the type or pattern of information to be searched in the database.

- Data mining

  Is the process to find patterns or information in the selected data using techniques or methods that are varied widely. Selection for suitable methods or algorithms to use is fully dependent on the purpose(s) and KDD processes as a whole.

- Interpretation/evaluation

  Information patterns resulted from the data mining process must be presented in the comprehensive form that can be understood by the involved parties. This phase will check whether or not the pattern or information found is conflicting with the previous facts or hypothesis.

C. Classification

Data mining algorithms are basically divided into supervised learning, unsupervised learning, and semi supervised learning. Classification is a common example of supervised learning approach whereby a given data set is typically divided into three parts that are training, validation, and testing data sets with known class labels. Classification is the process of providing an algorithm with records in which an output variable of interest is known and the algorithm "learns" how to predict this value with new records where the output is unknown.

Gundecha and Liu explained that "supervised algorithms build classification models from the training data and use the learned models for prediction. To evaluate a classification model's performance, the model is applied to the test data to obtain classification accuracy" [8, p. 2]. Several methods that can be used for supervised learning are decision tree induction, Naïve Bayes, Support Vector Machines (SVM), and k-nearest neighbors. In light of this, this particular research will use Naïve Bayes classification method to classify emotions of Twitter's users.

D. Naïve Bayes Classification Method

Naïve Bayes is one of the fastest and simplest Bayesian Learning methods. It is derived from Bayes theorem and the hypothesis of freedom, producing statistic classification based on opportunities. This is a simple technique and should be used before attempting a more complex method. Barry and Kogan explained that chances of a message d in class c, P cd, calculated as [6]:

$$P(c|d)\ P(c)\Pi P(tk|c)mk=1 \tag{i}$$

where (|) is the conditional probability of a feature that occurs in the classroom message c. $P(c)$ is probability from the previous message happened in class c. $P(tk|c)$ can be used to measure how much evidence the contribution that c is the correct class. In the email classification, classes from the messages are determined by finding maximum a posteriori (MAP) classes that are mostly defined by:

$$cmap=a\ gmaxc\ {*c}\ cs+\ P(c|d)\ =a\ gmaxc\ {*c}\ cs+\ P(c)\ \Pi P(tk|c)\ =1 \tag{ii}$$

Formula 2 involves multiplication many conditional probability, one for each feature, then the resulting calculations are in floating point underflow. In practice, multiplying the opportunities are often converted into an additional logarithmic probability and therefore, to maximize the equation can use the following alternative.

$$cmap=a\ gmaxc\ {*c}\ cs+,\ P(c)\ \Sigma\ P\ tk\ c\ =1 \tag{iii}$$

All parameters of the model, the probability distribution classes and features, can be estimated with relative frequency of training data D. For example, when the given class and messaging features do not occur together in the training data, the estimated probability based on the appropriate frequency will be zero, which would make the right side of Formula 3 is undefined. This problem can be overcome by incorporating some corrections as Laplace smoothing in all probability estimates, so the chances of each feature can be calculated by the following equation.

$$P(t|c)=\ T\ 1\Sigma\ T\ V=\ T\ .\Sigma T\ V \tag{iv}$$

where | | is the number of terms in vocabulary.

E. Emotions

Emotion is defined as a reaction to certain situations that can be seen from the face or body [9]. According to Ekman, there are six categories of emotions such as happiness, sadness, fear, anger, disgust, and surprise [5]. In extending this viewpoint, Robbs made a wheel of feelings as shown below.

Figure 2. Wheel of Felling [10]

This wheel of feelings depicts the six emotions that thus are formed in the center of the wheel and the outside is synonymous more complex (e.g. loving, remorseful, and alienated). This helps to narrow down the best word used to express the emotional state at this time.

F. Twitter

Twitter offers a form of micro-blogging social network that allows users to make updates of writing texts with a maximum length of 140 characters and read messages [11]. Twitter updates are known as tweets. Several features provided in Twitter's site that are following, followers, mentions, favorites, hash tag (#), direct message, list, and more.

Tweets can be seen by the public, but the sender can restrict delivery of messages to their friends list only. Users can see the tweets of other users, known as followers (followers).

IV. Design and Implementation

A. Requirements

The following will enlist the specification requirements for the text mining application of emotion classifications of Twitter's users.

1. Applications can extract data from Twitter online using the Twitter API (Application Program Interface).
2. Applications can search tweets by username or keywords entered in real-time using the Twitter API.
3. Applications can clear the data that has been extracted.
4. Applications can perform training of the data that has been extracted.
5. Applications may classify tweets or sentences into six emotional categories, namely happiness, sadness, fear, anger, surprise, and disgust.
6. Applications can validate and provide accuracy values obtained from the training data and testing data.

B. Design Modeling

This particular section will briefly explain the modeling of the text mining application by showing all the activities and datasets used in the application. Once the tweets data have been successfully extracted from the Twitter API, these data will go to the pre-processing phase where there are numerous activities performed including case folding, cleansing, stop-word removal, emoticon conversion, negation conversion, and tokenization. When the pre-processing is done, it later moves to the training part where one should use the training dataset to generate models based on the rules set on the data. The models are evaluated using the Naïve Bayes algorithm that results. After the best model is selected, it can automatically classify the emotions and thus display the results by showing the level of accuracy of classifying emotions that exist in the tweets. Of note here is that test data are used only at the end of the model building and selection process to assess how well the final model might perform on additional data.

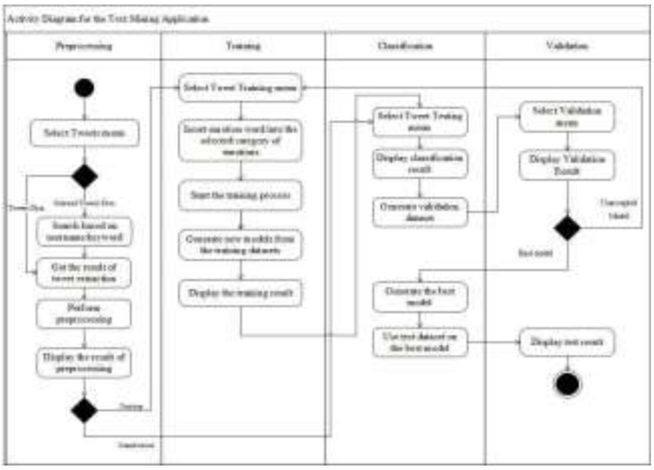

Figure 3. Activity Diagram for the Text Mining Application

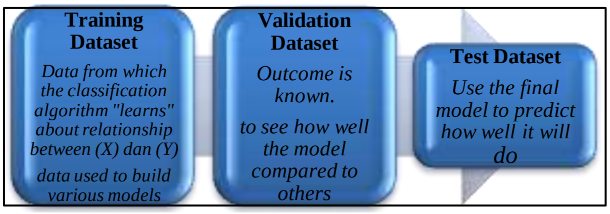

Figure 4. Datasets for the Text Mining Application

Note:
X = input variable (independent variable)
Y = output variable (dependent variable)

In the design process, the data model is divided into three datasets that are training dataset, validation dataset and test dataset. A series of rules are set on the training dataset. This dataset is obtained when entering into the application to determine the category of each tweet that is later weighted for each word in that particular tweet. Several models are emerged in this training process. Another dataset, also known as the validation dataset, is thus applied to these models. This is to compare which model works the best using the Naïve Bayes algorithm. Once the best model has been selected, the test dataset is used to predict how well it can classify the emotions contained in those tweets by showing the accuracy results.

C. Development Phases

Figure 5 shows steps in developing the text mining application.

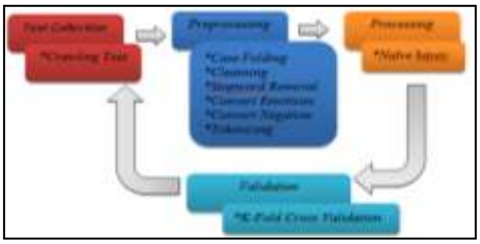

Figure 5 Development Phases for the Text Mining Application

1. Text Collection
   It is done using the streaming API and Twitter Search with additional filters based on username/keyword as shown in the figure below.

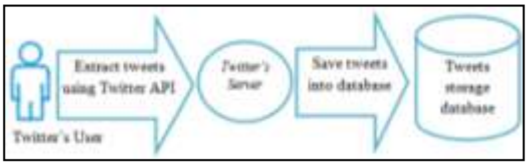

Figure 6. Process of Text Collection

2. Preprocessing:
   Once the entire data tweets are successfully retrieved, then one needs to separate the data into two datasets that are the training data and data testing. Furthermore, the second part of the data is then performed text preprocessing of the data. The stages of preprocessing, among others:
   1. Conversion to lowercase.
   2. Removing URL (e.g. http://bit.ly/mEibnV).
   3. Removing mention (e.g. xoxo).
   4. Delete a character other than a- z.
   5. Eliminate stop-word.
   6. Convert the symbols/emoticons into texts.
   7. Convert the negation and combine it with the word after negation.

   After the training data and testing are clean, the next step to do is to apply the Naïve Bayes classifier algorithm in the training process to build a model of the probability of training data.

3. Processing:
   Using Naïve Bayes algorithm to classify tweets in the application.
   Naive Bayes classification on every tweet represented in a pair of attributes $<a_1,a_2,a_3,\dots,a_n>$ where a_1 is the first sentence, a_2 is the second sentence, and so on. V is the class set. At the time of classification, this method will produce a category/class of highest probability (V_{MAP}) by inserting attributes $<a_1,a_2, a_3,\dots,a_n>$. Formula for V_{MAP}:

$$V_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} P(v_j | \, a_1, a_2, a_3, \dots, a_n) \quad \text{(v)}$$

   By using the Bayes theorem, the equation (v) can be written as:

$$V_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} \frac{P(a_1, a_2, a_3, \dots, a_n | v_j) \times P(v_j)}{P(a_1, a_2, a_3, \dots, a_n)} \quad \text{(vi)}$$

   $P(a_1,a_2, a_3,\dots,a_n)$ is a constant value for all, v_j, so that the equation (vi) can be expressed into the equation (vii).

$$V_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} P(a_1, a_2, a_3, \dots, a_n | v_j) \times P(v_j) \quad \text{(vii)}$$

Naive Bayes algorithm simplifies this equation by assuming that every attribute is conditionally independent of each other for each category. In other words,

$$P(a_1, a_2, a_3, .., a_n | v_j) = \prod_i P(a_i | v_j) \quad \text{(viii)}$$

If the equation (vii) substitutes to the equation (viii), the result is:

$$V_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} P(v_j) \times \prod_i P(a_i | v_j) \quad \text{(ix)}$$

a_i

$P(v_j)$ and word probabilities for each category P($a_i | v_j$) counts in training process. Where,

$$P(v_j) = \frac{docs_j}{training}$$
$$P(a_i | v_j) = \frac{n_i + 1}{n + kosakata} \quad \text{(x)}$$

$$P(a_1, a_2, a_3, .., a_n | v_j) = \prod_i P(a_i | v_j) \quad \text{(xi)}$$

Where $docs_j$ is the total of documents in j category and training is the total of documents used in the training process. n_i is the total of words appeared in v_j category, n is the total of glossaries existed in each category v_j and glossary is total of unique words in all of training data.

4. Validation: using 10-fold cross validation to determine the accuracy of classifying the emotions existed in the tweets generated by this application.

$$Accuracy = \frac{Total\ of\ True\ Classification}{Total\ of\ Testing\ Data}\ x\ 100\% \quad \text{(xii)}$$

D. Implementation of User Interfaces

Figure 7. Preprocessing Phase

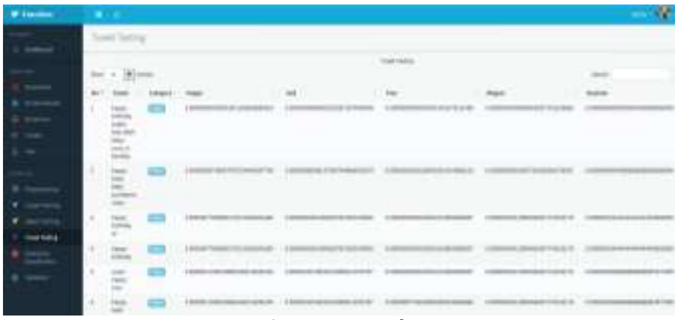

Figure 8. Processing Phase

Figure 9. Validation Phase

E. Implementation of Database

Figure 10. Database for the Application

V. TESTING

A series of testing performed on the application to measure the accuracy of classifying emotions of Twitter's users. The testing is divided into three parts as follows:

1. Using 10-fold cross validation

Table 1. Testing Performed Using 10-fold Cross Validation

Number of unique words	1.366 words
Training data	423 data
Test data	105 words
Accuracy	83.1034482759%

2. Based on the number of training data
   a. Number of unique words : 557 words
   b. Test data : 152 data

Table 2. Testing Based on Training Dataset

	Training Data	**Accuracy**
Test #1	116	71,3 %
Test #2	229	60,45 %
Test #3	280	62,04 %

3. Based on unique words and training dataset
   Number of dataset: 152 data

Table 3. Testing Based on Unique Words and Training Dataset

	Unique Words	Training Data	Accuracy
Test #1	1.164	280	63,7 %
Test #2	1.362	359	77,55 %
Test #3	1.495	404	77,14 %

The results of testing performed using 10-fold cross validation and training data showed that the ratio of the first and second test resulted in decreased levels of accuracy. Meanwhile, the second and third tests yield increased accuracy. Testing based on the number of unique words and training dataset revealed that the accuracy of classifying the user emotions is better as can be seen from the first, second, and third testing in Table 3. In summary, the number of unique words and training datasets have significant impacts on the level of accuracy achieved as it may increase or decrease the accuracy to certain levels.

VI. CONCLUSION AND RECOMMENDATION

A. *Conclusion*

1. This text mining application can successfully extract data from Twitter online using the Twitter API.
2. This application can classify emotion words and examples of emotion sentences that are suitable to be used to express the emotions of Twitter's users.
3. The application can classify emotions into 6 categories such as happiness, sadness, fear, anger, surprise, and disgust.
4. It has managed to build a model to classify tweets based on sentiment and categories using Naïve Bayes algorithm.
5. The test results showed that unique words and a larger training data will lead to a higher accuracy for the identification of emotions because it can provide a better and wider coverage of the emotional moments in our daily lives.

B. *Recommendation*

1. To conduct further research on this area of interest by developing this application using other classification methods such as K-Nearest Neighbor and Support Vector Machine.
2. Can add #hashtags as one of the components for the classification of emotions.
3. To automatically delete or omit a tweet that contains no emotion.
4. In the preprocessing process, one can add functionalities to detect and eliminate duplication of tweets (spams) as well as to reduce and eliminate the number of letters in a row.
5. The language used in the stop-word is not only English but could use Indonesian, local language or other foreign languages.
6. Add the Feature Extraction in particular for the Semantic Analysis and syntactical Analysis. Included in the syntactical Analysis is POS tagging and parsing, Semantic Analysis can detect concepts, events, and relationships between them.
7. Add the Feature Selection in particular for frequency-based feature selection, latent semantic indexing (LSD), and random mapping to eliminate irrelevant information and information which is repeated from the tweet.

References

[1] D. Farber, "Twitter Hits 400 Million Tweets Per Day, Mostly Mobile," 06th June 2012. [Online]. Available: www.cnet.com/news/twitter-hits-400-million-tweets-per-day-mostly-mobile/. [Accessed 10th September 2015].

[2] R. Irfan, C. K. King, D. Grages, S. Ewen, S. U. Khan, S. A. Madani, J. Kolodziej, L. Wang, D. Chen, A. R. Rayes, N. Tziritas, C. Z. Xu, A. Y. Zomaya, A. S. Alzahrani and H. X. Li, "A Survery on Text Mining in Social Networks," *The Knowledge Engineering Review*, pp. 1-24, 2004.

[3] T. Patil and S. Patil, "Automatic Generation of Emotions for Social Networking Websites using Text Mining," *Journal of IEEE Computer Society*, vol. 4, p. 31661, 2013.

[4] M. Adedoyin-Olowe, M. M. Gaber and F. Stahl, "A Survey of Data Mining Techniques for Social Media Analysis," 17th December 2013. [Online]. Available: http://arxiv.org/abs/1312.4617.

[5] P. Ekman, Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life, New York: Times Books, 2003.

[6] M. W. Berry and J. Kogan, Text Mining Application and Theory, United Kingdom: Wiley, 2010.

[7] J. Han, M. Kamber and J. Pei, Data Mining: Concept and Techniques, Waltham: Morgan Kaufmann Publishers, 2012.

[8] P. Gundecha and H. Liu, "Mining Social Media: A Brief Introduction," INFORMS, Catonsville, USA, 2012.

[9] S. W. Sarwano, Pengantar Psikologi Umum, Jakarta: PT. Raja Grafindo Persada, 2010.

[10] "Dramatic Languages," 2014. [Online]. Available: https://robbsdramaticlanguages.wordpress.com/2014/07/31/. [Accessed 10th August 2015].

[11] D. Juju and F. Sulianta, Broading Promotion with Social Network, Jakarta: PT. Elex Media Komputindo, 2010.

[12] M. Yassine and H. Hajj, "A Framework for Emotion Mining from Text in Online Social Networks," *Journal of IEEE Computer Society*, vol. 10, pp. 1136-1142, 2010.