

The Development of Indonesian Sentiment Analysis with Negation Handling

Novi YUSLIANI^{1*}, Mastura Diana MARIESKA², Endang LESTARI³, M. Ridho
Putra SULFA⁴, and Wibi ARIMURTI⁵

^{1,2,3,4,5}*Universitas Sriwiaya, Indonesia*

**Corresponding author: novi_yusliani@unsri.ac.id*

ABSTRACT

The polarity of a text can be seen based on the words used in the text. The use of negation words in a sentiment can change the polarity of the sentiment. One example of a negation word in Indonesian is the word 'tidak'. If a sentiment has a word "tidak" combined with a word "jelek" to form "tidak jelek", then the sentiment is classified into positive class. However, if the combination of the word "tidak jelek" is processed as a word that has no relationship, then the sentiment will be falsely classified into negative class. The existence of negation word in a sentiment is one of the challenges in sentiment analysis, because it can cause ambiguity and misclassification when classifying text based on its polarity. This research shows that negation handling cannot be done only by doing syntactic analysis, but it also requires semantic analysis to overcome the problem of ambiguity.

Keywords: *sentiment analysis, negation handling, Multinomial Naïve Bayes*

INTRODUCTION

Sentiment analysis is a process to study the opinions, emotions, and expressions of a person in written form or text, with the aim to conclude whether the text is positive or negative. In sentiment analysis, there is a challenge in classifying, namely the use of negative words in a text that can change the polarity of the text. One example of a negative word in Indonesian is the word 'tidak'. If a text has the word 'tidak' combined with the word 'jelek' to 'tidak jelek', then the text is classified as a positive class. However, if the combination of the word 'tidak jelek' is processed as a word that has no relationship, the text will be wrongly classified as a negative class.

According to the research by Dekharghani et.al. [1], one of the sentiment analysis issues shared across languages and domains is the need to handle negation, this is because negation handling is language-specific problem and need to be handled by different methods in each language. On another previous research, Sharif et.al. [2] discusses the impact of the word negation on customer reviews. According to the study, the existence of the word negation can lead to misclassification, due to their customer reviews that should be negative but incorrectly classified into positive reviews. Based on [3] by Mohey and Husein, negation has the highest accuracy percentage that can support the result of the sentiment analysis because researchers in sentiment do not need to understand whether the negative reviews are explicit or implicit.

In this research, the negation word known in a sentiment will be processed using Negation Handling, through this process, the word negation will be removed from the sentiment, after that, the polarity of the sentiment is changed from positive to negative or vice versa.

LITERATURE REVIEW

Related Works

Problems related to the use of negation word in sentiment analysis are the main focus in the study of Wiegman et.al. [4], Hogenboom et.al. [5], Amalia et.al. [12] and Arabic sentiment analysis by Duwairi and Alshboul [6]. Based on a survey conducted, Wiegman et.al. concluded that an effective way to deal with negation word is to have knowledge related to the form of sentence expressions of each positive class and negative class. A negation does not negate every word in a sentence, therefore, using syntactic knowledge to model scope of negation expressions is useful [4].

Hogenboom et.al. compared several approaches to accounting for negation in sentiment analysis, differing in their methods of determining the scope of influence of a negation keyword. Hogenboom et.al. concluded that the method that has the best performance in handling negation words in English sentences is to consider two words after a negation word is declared as negation word [5]. Duwairi and Alshboul [6] addressed negation effects on the polarity of sentences written in Arabic. In Arabic, the handling of negation words in sentences is related to culture and language morphology. Therefore, a set of rules is generated based on the morphology of negative sentences in Arabic [6]. Amalia et.al [12] stated that the presence of the word negation is able to change the polarity of the text if it is not handled properly. They proposed a rule based method to determine what words negated by adapting the rules of Indonesian language syntactic of negation to determine the scope of negation.

POS-Weighted TF-IDF Algorithm

POS-Weighted TF-IDF algorithm is a method derived from the TF-IDF algorithm. In this process, the weighting value will vary depending on the POS-Tagging. In the POS-Weighted TF-IDF algorithm, the frequency of each term will be calculated by the value or weight according to the part of speech of each word. In the POS-Weighted TF-IDF method based on previous research [7], there are 3 types of calculations will be performed, namely calculating TF (Term Frequency), then calculating IDF (Inverse Document Frequency), and finally calculating the TF-IDF. The following are the calculations for each calculation:

1) TF (Term Frequency) using equation (1).

$$tf_{POS}(t, d) = \frac{c(t,d) \times w_{POS}(t)}{\sum_i c(t_i,d) \times w_{POS}(t_i)} \quad (1)$$

Where t is the symbol of *term*, d is the symbol of *document*, $c(t, d)$ is the total number of term frequency in document d , while $c(t_i, d)$ is the total number of specific term in document d , $w_{POS}(t)$ is the weight of term that have been determined in each part-of-speech category, and $w_{POS}(t_i)$ is the weight of specific term that have been determined in each part-of-speech category.

(2) IDF (Inverse Document Frequency) using equation (2).

$$idf(t) = 1 + \log \left(\frac{N}{df(t)} \right) \quad (2)$$

Where N is the total number of all document, and $df(t)$ is the total number of document that contains of *term* (t).

(3) Thus the formula for POS Weighted TF-IDF is calculated using equation (3) and equation (4).

$$w_{POS}(t, d) = tf_{POS}(t, d) \times idf(t) \quad (3)$$

$$w_{POS}(t, d) = \frac{c(t,d) \times w_{POS}(t)}{\sum_i c(t_i,d) \times w_{POS}(t_i)} \times \left(1 + \log \left(\frac{N}{df(t)} \right) \right) \quad (4)$$

According to Xu [7], in general, words with POS-Tag, namely nouns and verbs, are more meaningful than adjective and adverb tags, which are more important than conjunctions, preposition, etc., so that the more meaningful the POS-Tag are, the greater the weighting value, therefore weighting knowing the POS-Tag in advance can drastically improve the results of sentiment analysis research. In other studies, Yang [8] explains that by performing this post-weighted TF-IDF algorithm, the average accuracy is 80.2% for the original TF-IDF, and 87.2% for the post-weighted TF-IDF.

Multinomial Naïve Bayes

There are two different generative models of the “naïve Bayes” classifier, the first one is called Multivariate Bernoulli event model and the second is called Multinomial event model. Multinomial Naïve Bayes is a frequency-based model proposed for text classification, where this model takes into account the frequency of each word occurrence in a document and its probability by calculating prior probability $P(c)$ and calculating conditional probability $P(w_i|c)$ [9]. The following is the calculation of the text classification with the Multinomial Naïve Bayes according to [10]:

(1) Calculating prior probability $P(c)$ using equation (5).

$$P(c) = \frac{\sum_{j=1}^n \delta(c_j, c) + 1}{n + l} \quad (5)$$

Where $\delta(c_j, c)$ is the total number of *class* c in *training document*, n is the total number of *training document*, and l is the sum of *class*.

(2) Calculating conditional probability $P(w_i|c)$ using equation (6).

$$P(w_i|c) = \frac{\sum_{j=1}^n f_{ji} \delta(c_j, c) + 1}{\sum_{i=1}^m \sum_{j=1}^n f_{ji} \delta(c_j, c) + m} \quad (6)$$

Where $\sum_{j=1}^n f_{ji} \delta(c_j, c)$ is the total number of occurrences of certain words in a *training document* from *class* c , $\sum_{i=1}^m \sum_{j=1}^n f_{ji} \delta(c_j, c)$ is the total number of words in a *training document* from *class* c , m is the total number of different (unique) words in the *document*, binary function of $\delta(c_j, c)$ is defined from equation (7).

$$\delta(c_j, c) = \begin{cases} 1, & \text{if } c_j = c \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

(3) Classifying *test documents* d shown by word vectors $\langle w_1, w_2, \dots, w_m \rangle$ using equation (8).

$$c(d) = \arg \max_{c \in C} [\log P(c) + \sum_{i=1}^m f_i \log P(w_i|c)] \quad (8)$$

Where $P(w_i|c)$ is the conditional probability of a document in *class* c , $P(c)$ is the prior probability of a document in *class* c , $w_i (i = 1, 2, \dots, m)$ is word that appear in the document d , $f_i (i = 1, 2, \dots, m)$ is frequency of w_i in the document d

4. Then determine the class of document by selecting the highest probability value.

Based on the research by McCalum & Nigam [9], the Multinomial model is better than Multivariate Bernoulli, because their result shown that with Multinomial model, it can reduces the error rate by an average of 27%. On another research by Song et.al [10], using Multinomial Naïve Bayes, they succeeded in conducting sentiment analysis with a maximum accuracy of 85.33%, from the results they stated that, this method succeeded in exceeding the Maximum Entrophy method and the Support Vector Machine method in conducting sentiment analysis research.

Negation Handling

Based on research conducted by Yulietha et. al. [11], negation handling has a significant influence in increasing accuracy in conducting sentiment analysis, with an average accuracy of 82.4% without negation handling, and 84.8% with negation handling. This research also concluded that, the more data used to conduct training data, the results of the F-Measure value would be higher. Better accuracy results were also obtained in studies conducted by Sharif et.al. [2], negation handling in sentiment analysis can improve the accuracy of the research they have done, with an accuracy of 84.81% without negation handling, and an accuracy of 91.8% with negation handling.

PROPOSED METHOD

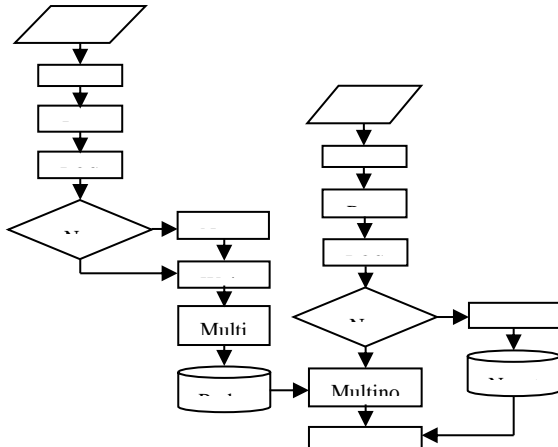


Figure 1. Negation Handling Process for Proposed Method

The method proposed for this research is the Negation Handling process, where with the Negation Handling process, negative word in the sentiment will be eliminated. Then, changing the polarity of the sentiment into negative from positive, and vice versa, in the training phase. This negation handling process will be implemented in sentiment analysis software.

Stages of the software flow begins with entering training data in the form of text from Twitter relating to the construction of the Light Rail Transit (LRT) in the city of Palembang. This data has been manually grouped into positive and negative sentiments. Furthermore, the training data passes through the normalization and pre-processing stages. After the pre-processing stages are completed, the

POS-Tagging process will be carried out which aims to get the word class for each word. Then eliminate words with negation / NEG word classes in the training data through the Negation Handling stage. Then assign weights with certain values depending on the word class on each word using the TF-IDF method. Furthermore, the results of word weighting are used for classification with Multinomial Naïve Bayes. The system architecture is shown in Figure 1. The following is an explanation along with an analysis at each stage in the diagram:

Data

The data used in this research is a collection of tweets about LRT development in Palembang. This data is used for training and testing. There are 500 tweets used in this research, consists of 293 positive tweets and 207 negative tweets. From 207 negative tweets, there are 155 tweets contain negation word ('tidak', 'jangan', 'bukan', 'tak', 'belum'). This data will then pass through the stages of normalization, preprocessing, POS-Tagging, negation handling, and classification to get a probabilistic model from the classification results using Multinomial Naïve Bayes (MNB).

Normalization

At this stage, the normalization stage will be carried out on the training data, this normalization phase consists of three processes:

1. removal of symbols and punctuation,
2. delete unicode on Twitter, namely username, hashtags, and URL,
3. and change words that are not standard, into standard words. This process uses additional data in the form of text, namely standard and non-standard language dictionaries.

Pre-processing

At this stage, the text data is converted to numeric form, so that the text data can be easily processed at later stages. Pre-processing has several processes, namely casefolding, tokenizing, stopword removal, and stemming. Casefolding is used to convert all letters into lower case forms. While, tokenizing is used to break a sentence into a single string or tokens. At the stopword removal process, the term or word is deleted if it is in the stopword list in the dictionary. The additional data used in this process is a stopword dictionary. Last stemming is the process of changing words that have affixes into basic words. The method used for stemming is the Nazief Andriani method.

POS-Tagging

The next stage is the POS-Tagging stage which is used to give the word class to the chirp text. This POS-Tagging process uses the Hidden Markov Model method, using the library of NLP ITB. In this process, the word class displayed is limited to only the following 6 word classes, namely: JJ (adjective), RB (adverb), VB (verb), NN (noun), NEG (negation word), and OTHER. These classes will be used for determining the weighting value for POS-Weighted TF-IDF algorithm.

Negation Handling

The next stage is the Negation Handling stage which is the main objective in this research. At this stage, the data that has the negation/ NEG word class will be changed in polarity, from positive to negative sentiment, and vice versa. After that, words with the negation/ NEG word class will be omitted in the data. This stage is used to reduce the existence of ambiguity during the sentiment classification stage.

Weighting

The next step is the weighting stage using POS-Weighted TF-IDF method. This method is a way to provide word weighting based on its word class. The results of the weighting will then be used as the weight value of each word in the classification feature. The scoring on this weighting differs slightly depending on the part-of-speech class related to the term given a weighting value.

Multinomial Naive Bayes

Multinomial Naïve Bayes is used to produce a probabilistic model which will later be used as a reference for classifying sentiments with new testing data. Previous probabilities and conditional probabilities for each class are modeled in this classification stage.

Probabilistic Model

Previous probability calculation results and probability calculation results in training data will be stored into the system. Both of these calculation results can also be referred to as probabilistic models. This probabilistic model will be used at the testing stage as a reference to the weight value of each term.

RESULT

This research used 10-fold cross validation to validate the system. The data is divided into ten group. Each group contains randomly 50 tweets, negative and positive. There are 155 tweets that contain negative word ('tidak', 'bukan', 'belum', 'tak', 'jangan'). The performance measurements displayed in table 1, were the result of accuracy, precision, recall, and f-measure (in percentage) of sentiment analysis that did not use negation handling, and from sentiment analysis using negation handling. The value of Performance Measurements taken is the average value of the results of data testing using 10-Fold Cross Validation.

Table 1. Average Performance Measurements

Experiments	Performance Measurements			
	Accuracy	Recall	Precision	F-Measure
Without Negation Handling	89.6%	91.29%	90.64%	90.85%
With Negation Handling	74.6%	78.44%	78.08%	78%

DISCUSSION

The proposed method, used negation handling to classify the polarity of sentiment. All training data in this method, have to pass the pre-processing and the normalization stage. Both of this stage, used to clean the noise in the text. After that, the data will pass to POS Tagging stage, to detect the negative word. Then eliminating negative word in the sentiment and changing the polarity of the sentiment in training stage. The polarity of a sentiment will not ambiguous again, and the meaning of words in sentiments other than negative words, the polarity also will not ambiguous. The polarity of the sentiment will not be changed if there is no negative word in the sentiment.

After the negation handling stage, all words in the sentiment will be assigned a weight based on their word class. POS Tagging stage is used to determine word class each word in sentiment. After the weighting value has been determined, the next step is the weighting stage using POS-Weighted TF-IDF algorithm. In the POS-Weighted TF-IDF algorithm, the frequency of each term will be calculated by the value or weight according to the part of speech of each word. In this stage, there are 3 types of calculations will be performed, namely calculating TF (term frequency), with the value of the appearance of a term will differ depending on the weighting value that has been determined, then calculating IDF (Inverse Document Frequency), and finally calculating the TF-IDF. In this weighting stage, by knowing the part-of-speech in advance can drastically improve the results of sentiment analysis research.

After weighting stage, all of the weight value will be trained using Multinomial Naïve Bayes and the probability model will be stored. In the classification phase, the frequency of each word occurrence in a sentiment and its probability is calculated by calculating prior probability and conditional probability in training phase, then determine the polarity of a sentiment (positive or negative) by selecting the highest probability value. The polarity of a sentiment resulted from the Multinomial Naive Bayes will be checked, if there is a negative word in the sentiment, then the polarity will be changed, from positive to negative, and vice versa.

The performance of this research can be seen in table 2. The results show that the negation handling by removing the word negation and changing the sentiment polarity cannot improve the accuracy of the sentiment analysis system using Multinomial Naive Bayes. The negation handling used in this research is only based on the surface expression of sentiments, the presence or absence of the negation word in the sentiment.

The main factor causing the low accuracy is the absence of semantic analysis. The purpose of semantic analysis in sentiment analysis systems is to deal with the ambiguity of the meaning of words in sentiment due to the negation word. Some examples of tweets that are incorrectly classified by the system:

1. In a single sentence, word with negation/NEG post tags are followed by words with noun/NN heading tags, if the negation word is deleted, it does not change the sentiment of the tweet, for example:
 - a. Sy org palembang.. proyek LRT adalah proyek yg mubadzir.. kurang manfaat. Nyata nya kereta sepi dr penumpang. Wong palembang butuhny jembatan bukan/NEG kereta/NN.
 - b. Kendaraan yang pas buat mahasiswa apalagi kalau bukan/NEG LRT/NN ini Keren banget lah pokoknya.
2. In compound sentences, special handling is needed, because deleting words with negation/NEG posts can make the sentence ambiguous, for example:
 - a. Kini Terbukti Merugi, LRT Palembang Belum/NEG Terlalu Dibutuhkan
 - b. Lrt Palembang proyek gagal, kurang diminati, bukan/NEG solusi terbaik utk masalah yg dihadapi rakyat
3. In some compound sentences, deleting words with negation / NEG posts, meaningfully, still does not change the tweet's sentiment, for example:
 - a. LRT Palembang mogok terus. Padahal pengadaan keretanya gak/NEG makan biaya sedikit.
 - b. LRT di Palembang bukan/NEG hanya mengatasi masalah transportasi tp jd ikon kota yg memiliki dampak positif bagi masyarakatnya #AlexNoerdin
4. Compound sentences consist of parent sentence and child sentence, in compound sentences, most negations that are considered inappropriate occur because the word negation/NEG is exist in child sentence, for example:
 - a. Lrt sangat mengasyikan, semoga ada mobil angkutan yang membawa kita menuju stasiun

terdekat sehingga nggak/NEG bingung ke sana nya #lrtpalembangpunyokito"

Parent sentence: Lrt sangat mengasyikan

Child sentence: semoga ada mobil angkutan yang membawa kita menuju stasiun terdekat sehingga nggak/NEG bingung ke sana nya

- b. Belum/NEG genap satu bulan uji operasi, LRT Palembang sudah tiga kali mogok.

Parent sentence: LRT Palembang sudah tiga kali mogok

Child sentence: Belum/NEG genap satu bulan uji operasi

- c. Lrt ini sangat memudahkan saya, saya tidak/NEG perlu lagi merasakan kemacetan

Parent sentence: Lrt ini sangat memudahkan saya

Child sentence: saya tidak/NEG perlu lagi merasakan kemacetan

CONCLUSION

The proposed method in this research can't increase the performance of sentiment analysis system. This method only analyzes sentiment based on the surface expression of the sentiment, the existence of the word negation in the sentiment. In fact the word negation is also used to express positive sentiment. If the word negation is deleted, the sentiment is still positive. In this study, if the word negation is found in the sentiment the word negation will be deleted and the polarity of the sentiment will be changed. Therefore, the words that are in the sentiment will be classified into the last polarity. In the Multinomial Naive Bayes method, the classification of sentiments depends on the probability value of the sentiment. Sentiment probability value depends on the probability value of each word in each class, positive and negative. If the probability value of sentiments in the positive class is greater than the probability value in the negative class, then the sentiment will be classified into positive. Therefore, negation handling in sentiment analysis systems cannot be overcome only by using syntactic analysis. It also needs semantic analysis to overcome the problem of ambiguity.

REFERENCES

- [1] Dehkharghani R, Yanikoglu B, Saygin Y and Oflazer K 2016 Sentiment analysis in Turkish at different granularity levels *Nat. Lang. Eng.* 1–25
- [2] Sharif W, Samsudin N A, Deris M M and Naseem R 2017 Effect of negation in sentiment analysis 2016 *6th Int. Conf. Innov. Comput. Technol. INTECH 2016* 718–23
- [3] Mohey D and Hussein E M 2018 A survey on sentiment analysis challenges *J. King Saud Univ. - Eng. Sci.* 30 330–8
- [4] Wiegand M, Balahur A, Roth B, Klakow D and Montoyo A 2010 A Survey on the Role of Negation in Sentiment Analysis *Proc. Work. Negation Specul. Nat. Lang. Process.* 60–8

- [5] Hogenboom A, Van Iterson P, Heerschop B, Frasinca F and Kaymak U 2011 Determining negation scope and strength in sentiment analysis *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.* 2589–94
- [6] Duwairi R M and Alshboul M A 2015 Negation-Aware Framework for Sentiment Analysis in Arabic Reviews *Proc. - 2015 Int. Conf. Futur. Internet Things Cloud, FiCloud 2015 2015 Int. Conf. Open Big Data, OBD 2015* 731–5
- [7] Xu R 2015 POS weighted TF-IDF algorithm and its application for an MOOC search engine *ICALIP 2014 - 2014 Int. Conf. Audio, Lang. Image Process. Proc.* 868–73
- [8] Yang Y 2018 Research and Realization of Internet Public Opinion Analysis Based on Improved TF - IDF Algorithm *Proc. - 2017 16th Int. Symp. Distrib. Comput. Appl. to Business, Eng. Sci. DCABES 2017* 2018-Sept 80–3
- [9] McCallum A and Nigam K 1998 A Comparison of Event Models for Naive Bayes Text Classification *AAAI/ICML-98 Work. Learn. Text Categ.* 41–8
- [10] Song J, Kim K T, Lee B, Kim S and Youn H Y 2017 A novel classification approach based on Naïve Bayes for Twitter sentiment analysis *KSII Trans. Internet Inf. Syst.* 11 2996–3011
- [11] Yulietha I M, Faraby S A, Adiwijaya and Widyaningtyas W C 2018 An implementation of support vector machine on sentiment classification of movie reviews *J. Phys. Conf. Ser.* 971
- [12] Amalia R , Bijaksana M A, and Darmantoro D 2018 Negation Handling in Sentiment Classification using Rule-Based Adapted from Indonesian Language Syntactic for Indonesian Text in Twitter *J. Phys. Conf. Ser.* 971