ABSTRACK

Another person's opinion of a movie review in the media is very important in making a decision. To know opinion or polarity someone to a movie review in the media required a system that can facilitate in knowing the polarity of a person. Sentiment classification is one that can help in building the system to know the polarity of a person against a movie review. The dataset used in this sentiment classification process is IMDb (Internet Movie Database). However, the problem of knowing the polarity of an opinion in sentiment classification process on this dataset are their movie review unstructured data, many attributes of data as well as the negation that causes the polarity of a word will be different in the context of different texts. With these problems then the classification process on the dataset will be classified into two classes of polarity that is positive and negative. Before doing the classification process, there are some things to do to reduce the attributes on the dataset. First, there will be a preprocessing process for handling unstructured data. Second, negation handling, to be handled here is the use of the word "not", "no" "n't". The word "not" will be combined with the next word so it becomes a new word. Negation handling of "not","no", "n't" with the technique can determine the polarity of a word in different text context. Third, the TF-IDF process is used as an option to be able to select any terms or words to be used for interior sentiment process. After the three phases, sentiment classification is carried out by using either a machine learning method that is Multinomial Naïve Bayes. Result of reseach show that using preprocessing, negation handling based on punctuation in multinomial naive bayes is 85.16 %.

Keywords: Multinomial Naïve Bayes, TF-IDF, Preprocessing, Negation