Abstract

Product reviews from consumers is one of important factor for saling product. Analyzing and summarizing sentiment classification of product reviews has three steps that must be done. First step is feature extraction using frequent itemset mining with apriori algorithm. Second, opinion classification using SentiWordnet to determine opinion word polarity of a sentence. Last, summarizing document with semantic similarity scoring and sentence clustering. The results of this study found that filtering corresponding word also affect the performance of the extraction in this study. Word filtering that used are Noun, Noun Phrase, intersection and union of both, four of that filtering words used have a fairly balanced outcome, union and intersection that are expected to improve performance still got the result which is not too far away with just Noun and Noun Phrase. Results of the extraction performance in this study was about 20-40% in the dataset used. Classification using SentiWordNet show a good results, but in some datasets that have fairly high complexity sentence also decreased, although not much different and the performance still at 40-90%. Sumarization documents can be done well on a provided dataset for the dataset has a number of sentences adequate product reviews and summary with this methods some sentences representation of clustering are done well.

Keywords: sentiment analysis, product reviews, frequent itemset mining, association mining, semantic smimilarity scoring, sentence clustering.