ABSTRACT

Product reviews from consumers can be used by a company to increase sales and product quality, and for consumers product reviews can provide a purchase decision. The increasing number of reviews and user scores from consumers raises a new problem that is when the reviews provided by consumers are not in accordance with the given user score. So, a system is needed to perform a summary to provide value to consumer reviews with the help of user scores from consumers. The first stage to build this system is preprocessing where preprocessing consists of 3 stages, Stopword Removal, Lemmatization, and POS Tagging. The second stage is the feature and opinion extraction using the type dependency parser. The third stage is the determination of feature polarity by using Semantic Orientation – Pointwise Mutual Information. The fourth stage is a summary using feature scoring and Pearson correlation to find out the value of each extracted feature. The result of the extraction process and the polarity determination is a list of features and opinions as well as the polarity of each feature. Document summary using feature scoring generates scores of each feature successfully extracted in each review, followed by a Pearson Correlation process to see trends between scores of datasets and scores generated by feature scoring, and the highest correlation value is on the 16Gb Nook Tablet dataset is the 46th feature, Seller with a correlation value of 1.00 or very string correlated, while the lowest correlation is on the 16Gb Nook Tablet dataset that is the 62nd feature, Unit with a correlation value of -0.94 or very strong inverse correlation.

Keyword : sentiment analysis, product review, feature based opinion summarization, type dependency parser, semantic orientation- pointwise mutual information, feature scoring, pearson correlation