Abstract

Product reviews based on other people's opinions are very important to be a source of information to build the trust of potential buyers to judge the product. Even though there have been many reviews about restaurant products, there are still many reviews that are not convincing or biased, whether they are good or bad reviews. This problem is essential to be solved because biased judgments can affect the judgment of potential buyers and affect the business side of the restaurant. To solve this problem, a document-level sentiment analysis system was built with the input of a single review sentence and produces a positive or negative class output. The system was built using TF-IDF as a feature extraction method with the K-Nearest Neighbor (KNN) classification algorithm. The selection of the K-Nearest Neighbor (KNN) classification algorithm is because based on existing research, KNN results can be better than others because KNN performance is based on good features and the appropriate number of features. The final result of this study shows that the TF-IDF combined with the KNN classification algorithm with the Euclidean Distance calculation method without going through the Stopwords Removal preprocessing technique and with a ratio of 80:20 produces the best performance accuracy and f1-measure, namely 93.09% and 93.08% by using the value of k = 53.

Keywords: product reviews, sentiment analysis, preprocessing, TF-IDF, K-Nearest Neighbor, classification