Abstract

Products are something that producers offer to consumers, which each product has a different opinion for each consumer. Different product opinions can be either positive or negative. To analyze opinions in the form of positive or negative opinions automatically, it needs a system that can be used by producers and consumers. In this final project, conducted a research on the classification of opinion by using similarity values between words by using two graphical approach i.e. Word2Vec and WordNet. Word2Vec is a word representation in vector form that is used to generate word embeddings [1]. While WordNet is an English dictionary database that has a hierarchy of connectivity between words through the path it has [2]. This research shows that the opinion classification for product feature using Word2Vec has a higher percentage of accuracy when compared with the opinion classification for product feature using WordNet with average percentage difference of accuracy in 6 datasets is 2.07%. That is because in the Word2Vec approach, owned vocabulary can be developed alone based on used training data. While in the WordNet, the collection of words contained in the corpus are data from WordNet itself, so it can't be developed on its own like in Word2Vec.

Keywords*: sentiment analysis, word2vec, wordnet, graph-based classification, word embeddings.*