Abstract

Research in the field of linguistics about Al-Quran is very limited, including on Natural Language Processing (NLP). In the recent years, semantic similarity calculation becomes a very popular topic in NLP field. Which is interisting to apply this application and using Al-Quran as the dataset. Semantic similarity is a comparation between two unit in language, then assessed how these two units share the same attributes or properties. The more these two units have the same characteristics, then the higher value of semantic similarity between these two. In this final project we build the system to calculate the semantic similarity of words in the Quran. The selected approach for calculating the semantic similarity is to use the *Wordnet* as its knowledge-Based. As the selected method are Lin, Resnik and Wu & Palmer. The value obtained will be compared with the Gold Standard as the level of accuracy using Pearson correlation. The Result from this research is Resnik method is appeared to be the better method than the others. However, due to the low correlation score that being obtained from the three methods, we conclude that using wordnet as knowledge-based for this research is not suitable.

KeyWord : Semantic Similarity, Al-Quran, *Wordnet*, Knowledege-based, Gold Standard,Correlation