

ABSTRACT

Measurement of semantic similarity has an important role in natural language processing task. The background of this measurement is an issue in which the current computer cannot equate the human perception of semantic similarity measure between words. The use of knowledge bases as a source of information has been implemented using several method for calculating the semantic similarity between words. One of them is path based measurement that is popularly used by fully using the structure of taxonomy. This study was implemented path based measurement with Wang & Hirst measure for calculating the semantic similarity between words using WordNet as a knowledge base. The system performs calculations score in any combination of sense words with alpha parameter weights are set on a scale 0 - 5. Based on the observation which is indicated that the measurement by Wang & Hirst can be able to measure a high enough correlation value in both of dataset, it is 0.59 in SimLex999 dataset and 0.65 for WordSim353 dataset with weight alpha parameter is 0. The result of those correlation include to the moderate positive correlation in SimLex999 dataset and the strong positive correlation in WordSim353 dataset.

Keywords: semantic similarity, path based measure, Wang & Hirst measure, WordNet, SimLex999, WordSim353