Abstract

Semantic Relatedness measurement at this point is used to estimate the values of closeness between couples of words and text. Couples word or text has a concept or meaning which can be obtained from the encyclopedia. Wikipedia provides tens of thousands of concepts and meaning to the whole article in it. By using Wikipedia as a provider of the concept of closeness between couples can be estimated. In this thesis using the link on wikipedia that indicates the relationship between the article as a dataset in the calculation of semantic relatedness with input in the form of two words. Measurement of semantic relatedness done by using Wikipedia Link Based and two measurements are TFxIDF Inspired and Normalized Google Distance. Testing is done by correlating the value obtained on the dataset WordSim 353 as a standard dataset. Measurement of semantic relatedness on wikipedia done by forming relationships among interrelated article link that led to a concept related to the two input words. Results obtained correlation system that is equal to 50.2% by using a collection of 353 word pairs, while the reference value is the value of the correlation between the dataset Miner Simple system WS 353 is equal to 69%. To be able to maximize the value of the correlation of each pair should have a concept and word relationships respectively.

Keywords*: Word Relatedness, Semantic Relatedness, Wikipedia Link Based, TFxIDF, Normalized Google Distance.*