Abstract

Qur'an is a holy book that became the guide of life for Muslims. In the Qur'an there is a repetition of the same verse in another verse. One way to understand the Qur'an is to seek similarity and interrelationship between verses. Therefore, it is necessary research that can assess the similarity between verses with other verses. One of the research in aligning words that have similarities is word alignment. Word alignment aligns words based on contextual similarities of identical word sequences, named entities, word dependencies and surrounding words. The research also added a paraphrase database related to the Qur'an. In addition, alignment can be done by representing a sentence into a vector form using word2vec. For the measurement of semantic vectors can use cosine similarity calculation [1]. Evaluation was performed using Support Vector Regression (SVR) to measure the prediction value of value of Quranic verses pair English translation based on alignment and word2vec. The use of word alignment method, word2vec based on SVR in this research resulted pearson correlation 0,81221.

Keywords: Qur'an, semantic similarity, word alignment, word2vec, svr, pearson correlation.