

Abstrak

Al-Quran adalah kitab suci yang menjadi pedoman hidup bagi umat islam. Pada Al-Quran terdapat pengulangan ayat yang sama pada ayat lain. Salah satu cara untuk memahami Al-Quran adalah dengan mencari kesamaan dan keterkaitan antar ayat. Oleh karena itu, diperlukan penelitian yang dapat menilai kesamaan antar ayat dengan ayat lainnya. Salah satu penelitian dalam penyejajaran kata-kata yang memiliki kesamaan adalah *word alignment*. *Word alignment* memperhatikan kesamaan konteks dalam penyejajaran berdasarkan *identical word sequence*, *named entities*, *word dependency* dan *surrounding words*. Pada penelitian ini dilakukan penambahan *database* parafrase yang berhubungan dengan Al-Quran. Selain itu dalam penyejajaran ayat dapat dilakukan dengan merepresentasikan kedalam bentuk vektor dengan menggunakan *word2vec*. Untuk pengukuran nilai kemiripan berdasarkan vektor dapat menggunakan perhitungan *cosine similarity* [1]. Evaluasi yang dilakukan menggunakan *Support Vector Regression* (SVR) untuk mengukur nilai prediksi data pasangan ayat Al-Quran terjemahan bahasa Inggris berdasarkan *alignment* dan *word2vec*. Penggunaan metode *word alignment*, *word2vec* berdasarkan SVR pada penelitian ini menghasilkan nilai *pearson correlation* 0,81221.

Kata kunci: Al-Quran, kesamaan semantik, *word alignment*, *word2vec*, svr, *pearson correlation*.