

ABSTRAK

Kesamaan semantik memiliki peran penting dalam beberapa tugas Natural Language Processing (NLP) dan beberapa bidang terkait salah satunya Model Semantik Distribusi untuk kesamaan dan keterkaitan kata seperti pada penelitian ini. Pada penelitian sebelumnya, menggunakan dataset terjemahan Al-Qur'an untuk mencari nilai semantik pada surah tertentu, evaluasi dataset kesamaan dan keterkaitan semantik namun penerapannya pada Bahasa Turki, dan penerapan evaluasi dataset untuk kata kerja saja. Tujuan penelitian ini untuk mengevaluasi model semantik dengan dataset Bahasa Indonesia untuk kata benda dan kata kerja. Pasangan kata dipilih untuk mengaktifkan evaluasi distribusi model semantik dengan beberapa atribut kata dan hubungan pasangan kata seperti synonym dan antonym. Dataset yang digunakan terdiri dari 500 pasang kata yang dibantu oleh 15 responden untuk memberi nilai kesamaan dan keterkaitan dari setiap pasangan kata. Metode yang digunakan adalah Sim-Rel vektor, kuisisioner dan perhitungan standar emas, hingga hasil perhitungan performansi menggunakan korelasi Spearman Rank sebesar 0,909. Sumbu vektor Sim-Rel mendapatkan hasil dengan 4 kategori yaitu SU = 4 pasang kata, SR = 412, DU = 1, and DR = 83.

Kata kunci : *Natural Language Processing; Model Semantik Distribusi; Sim-Rel Vector; Spearman Rank*