

Abstrak

Kesamaan semantik adalah tugas untuk memperkirakan kekuatan hubungan semantik antara unit bahasa atau konsep, dalam hal ini kesamaan makna yang dimiliki oleh sepasang kata. Kesamaan semantik pada kata bahasa Indonesia dapat diukur dengan menggunakan basis pengetahuan seperti Kamus Besar Bahasa Indonesia. Metode berbasis vektor merupakan salah satu metode yang dapat digunakan untuk mengukur kesamaan semantik. Pada Tugas Akhir ini diimplementasikan kesamaan semantik pada pasangan kata bahasa Indonesia dengan menggunakan metode berbasis vektor, pembobotan *tf-idf*, dan perhitungan kesamaan kosinus, Kamus Besar Bahasa Indonesia sebagai basis pengetahuan, dan dataset yang digunakan dibuat berdasarkan referensi dari *SimLex999* dan *Rubenstein-goodenough* sebanyak 180 pasang kata, *gold standard* yang didapat berdasarkan hasil kuesioner terhadap 31 orang responden. Hasil penelitian yang telah dilakukan, didapatkan nilai korelasi terbaik sebesar 0.5416 dengan menambahkan definisi sinonim dalam pengujian. Parameter terbaik yang mempengaruhi nilai kesamaan semantik pada penelitian ini adalah dengan menambahkan definisi dari sinonim tanpa *stopword removal*.

Kata Kunci: *gold standard*, Kamus Besar Bahasa Indonesia, kesamaan semantik, kesamaan kosinus, metode berbasis vektor, *tf-idf*.