Abstrak

Penelitian terkait pengukuran nilai semantik dan korelasi pada pasangan ayat Al-Quran masih sangat sedikit. Penelitian masih terbatas pada pencarian ayat-ayat yang dibagi menjadi level 0, level 1, dan level 2 pada penelitian Qursim. Oleh karena itu, perlu adanya metode lain salah satunya yaitu metode *monolingual alignment*.

Pada penelitian tugas akhir ini digunakan metode *monolingual alignment* dan penambahan sistem *supervised learning* dalam mengukur nilai kesamaan semantik pada pasangan ayat-ayat Al-Quran. Hasil pengujian penelitian sistem fitur-fitur *alignment* pada *dataset* 2016 mendapatkan nilai korelasi tertinggi yaitu 0,728 dan pada *dataset* 2017 yaitu 0,698. Sedangkan untuk sistem fitur-fitur *alignment* yang ditambahkan sistem *supervised learning*, nilai korelasi tertinggi pada *dataset testing* 2016 yaitu 0,728 dan pada *dataset testing* 2017 yaitu 0,7. Berdasarkan hasil tersebut, disimpulkan bahwa fitur *identical words* sebagai fitur dasar *alignment* masih menjadi fitur terbaik dibandingkan fitur lainnya yaitu PPDB dan WordNet. Nilai korelasi dari penambahan *supervised learning* berhasil mendekati nilai korelasi dari sistem *alignment*.

Kata kunci: semantic textual similarity, monolingual alignment, supervised learning, gold standard, correlation.