

Abstraksi

Alquran merupakan pedoman hidup umat Islam. Alquran memiliki beberapa karakteristik yang unik diantaranya pengulangan ayat yang sama pada ayat lain dan pembahasan suatu topik pada ayat tertentu juga dibahas pada ayat lain dalam surat yang berbeda. Oleh karena itu, berdasarkan kebutuhan pemahaman Alquran oleh umat Islam maka diperlukan adanya penelitian pada teks Alquran. Salah satu penelitian dasar dalam *text mining* adalah penelitian tentang STS. Hasilnya dapat digunakan untuk penelitian lainnya dibidang *text mining* dan NLP seperti, *clustering*, *word sense disambiguation*, *question answering* dan penelitian lainnya. Pada penelitian ini dilakukan pengukuran kesamaan semantik antar dua potongan ayat Alquran dalam terjemahan bahasa Indonesia. Pengukuran kesamaan semantik pada penelitian ini mengimplementasikan metode *word alignment*. Selain itu, dilakukan juga pembuatan *gold standard* dan pengumpulan data parafrase Alquran yang menjadi kontribusi penulis dan *gold standard* dapat dijadikan sebagai tolok ukur untuk mengetahui nilai korelasi antara hasil pengukuran sistem dengan hasil dari intuisi manusia sebagai domain ahli. Pada penelitian ini ingin diketahui bagaimana nilai korelasi yang dihasilkan sistem, serta bagaimana pengaruh *feature alignment* yang diimplementasikan pada sistem. Pada hasil pengujian didapatkan bahwa nilai korelasi paling tinggi mencapai 75% dan diketahui *feature alignment* yang paling berpengaruh dalam implementasi STS pada ayat Alquran adalah *feature identical word alignment*.

Kata Kunci : Alquran, ayat Alquran, kesamaan semantik, *Word Alignment*, korelasi, *gold standard*, *feature alignment*