

Pembangunan Grouping Sinonim Set Tesaurus KosaKata Al-Qur'an dalam Bahasa Inggris Menggunakan Algoritma Hierarchical Clustering

Salma Fauziah¹, Moch Arif Bijaksana

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹salmaf@students.telkomuniversity.ac.id, ²arifbijaksana@telkomuniversity.ac.id

Abstrak

Penelitian di bidang penambangan teks untuk memproses entri atau kata-kata dari Al-Qur'an sangat bermanfaat bagi umat Islam. Penelitian ini bertujuan untuk membangun grouping sinonim set untuk tesaurus dalam kata-kata Al-Qur'an. Penelitian ini digunakan karena sumber pengetahuan tentang sains Alquran masih kurang. Dataset dalam penelitian ini menggunakan Corpus Qur'an dan Terjemahan Bahasa Inggris. Penelitian ini merupakan penelitian pengembangan dari sebuah artikel yang telah dipublikasikan, yaitu "Pengembangan Kosakata Al-Qur'an Set Sinonim dengan Pendekatan WordNet" oleh Laras Gupitasari. Input kata dari sistem penelitian ini menggunakan kata benda dari kata bahasa Inggris dalam Al-Quran. Output dari sistem menghasilkan beberapa kelompok yang memiliki tingkat kedekatan makna yang sama, kelompok pertama berarti kata dalam kelompok memiliki arti yang dekat. Untuk menghasilkan output, penelitian ini menggunakan pengelompokan kata dengan algoritma hierarchical clustering dan menghitung jarak menggunakan Path Similarity, kemudian mengelompokkan hasil sesuai dengan makna dari entri kata. Evaluasi dalam penelitian ini menghasilkan nilai F-Measure 76%, F-Measure adalah evaluasi untuk mengukur akurasi prediksi yang dikeluarkan oleh sistem.

Kata kunci : Tesaurus, Sinonim Set, Al-Qur'an, *Text Mining*, *Hierarchical Clustering*