

Abstrak

Dalam pembangunan *WordNet* Bahasa Indonesia, *synonym sets* merupakan bagian penting yang merepresentasikan kesamaan makna antar kata. *Synonym sets* dibangun dengan menggunakan Tesaurus Bahasa Indonesia sebagai basis data leksikal yang digunakan. Setelah melalui proses ekstraksi dari Tesaurus Bahasa Indonesia maka akan didapatkan *synonym set* yang memiliki kemiripan/kesamaan makna (*word sense*) antar kata. Secara umum perbedaan *WordNet* dan kamus adalah pada fokus utamanya, kamus biasanya fokus pada satu kata saja, sedangkan dalam *WordNet* fokus pada makna kata dan keterhubungan dengan kata-kata lain. Dalam penelitian-penelitian sebelumnya pembangunan *synonym set* dilakukan dengan menggunakan beberapa pendekatan salah satunya yaitu *clustering* untuk menghasilkan *synonym set* dan *WSD (Word Sense Disambiguation)*. Dalam penelitian ini pendekatan yang digunakan untuk menghasilkan *synonym set* adalah *clustering* dengan menggunakan *ROCK (Robust Clustering Using Links)* yang menggunakan nilai *similarity* dan *link*. *Synonym set* yang telah dihasilkan kemudian akan digunakan untuk pembangunan basis data leksikal. Oleh karena itu fokus utama dari penelitian ini adalah untuk menghasilkan *synonym set* melalui proses *clustering* dan mengitung nilai akurasinya, untuk perhitungan performansi dan evaluasinya menggunakan metode *F-measure* dengan melibatkan *gold standard*.

Kata kunci: *WordNet, ROCK Clustering, Synonym Sets.*