

Abstrak

Synonym set merupakan salah satu bagian dalam pembangunan *WordNet* Bahasa Indonesia. *Synonym sets* yang dibangun bersumber dari Tesaurus Bahasa Indonesia yang diproduksi oleh Pusat Bahasa. Proses pengekstraksian Tesaurus menghasilkan beberapa *synonym sets* yang memiliki kesamaan makna. Oleh karena itu, *hierarchical clustering* digunakan untuk mengelompokkan *synonym set* yang mirip berdasarkan nilai *similarity* dan *distance value*. *Synonym set* yang telah melalui proses *clustering* dapat digunakan sebagai salah satu pembangun *lexical database*.

Kata kunci : *hierarchical clustering, synonym set, Tesaurus bahasa Indonesia, WordNet.*

