

---

## Abstrak

*Word sense disambiguation* (WSD) merupakan metode pencarian makna asli dari sebuah kata ambigu dalam konteks tertentu. Berbagai jenis *classifier* dapat digunakan di WSD, salah satunya adalah pendekatan *supervised* dengan metode *decision list*. Metode klasifikasi *decision list* mampu menjadi kompetitor terbaik pada senseval 1 diantara partisipan *supervised*. Pendekatan *supervised*, tidak diragukan lagi bekerja lebih baik daripada pendekatan lain. Namun, pendekatan ini perlu mengandalkan banyaknya ketersediaan korpora yang digunakan untuk membuat dataset. Semakin banyak korpora yang digunakan maka semakin banyak atribut yang terdapat pada dataset. Banyaknya atribut yang diolah *classifier* akan berdampak pada menurunnya kinerja *classifier*. *Feature selection* dapat digunakan untuk mengoptimalkan kinerja *classifier* dengan cara mengurangi atribut yang kurang relevan pada dataset. *Information gain* merupakan salah satu seleksi fitur terbaik dibanding seleksi fitur lainnya pada penelitian yang telah dilakukan oleh Yang dan Pederson, Tan dan Yang serta Forman dalam hal klasifikasi dokumen. Karena keunggulan seleksi fitur *information gain* dan keunikan metode klasifikasi *decision list* tersebut, penelitian ini mengimplementasikan *information gain* sebagai seleksi fitur pada WSD bahasa Indonesia menggunakan metode klasifikasi *decision list*. Hasil penelitian ini, *information gain* dapat meningkatkan akurasi dengan selisih 0.5% dan selisih presisi 1.3% pada pengujian range *collocation* 2, serta selisih akurasi 0.3% dan selisih presisi 0.7% pada pengujian range *collocation* 3.

Kata kunci : word sense disambiguation, decision list, feature selection, information gain

---