

ABSTRAK

Nama : Triastuti Chandrawati
Program Studi : Ilmu Komputer
Judul : Pengembangan *Part of Speech Tagger* untuk Bahasa Indonesia Berdasarkan Metode *Conditional Random Fields* dan *Transformation Based Learning*

Part of Speech Tagging (POS *Tagging*) adalah kegiatan pemberian label kelas kata pada suatu kata. Proses ini pada awalnya dilakukan secara manual. Namun proses *part of speech tagging* secara manual menghabiskan banyak waktu dan tenaga karena membutuhkan banyak ahli bahasa untuk memberikan *tag* pada setiap kata. Masalah ini kemudian menjadi pendorong bagi para peneliti untuk membangun metode dan aplikasi yang dapat melakukan *part of speech tagging* secara otomatis.

Penelitian di bidang *part of speech tagger* untuk Bahasa Indonesia belum terlalu mendapat perhatian, karena sampai saat ini, belum ada suatu aplikasi *part of speech tagger* untuk Bahasa Indonesia. Maka penulis melakukan penelitian untuk menghasilkan suatu aplikasi *part of speech tagger* untuk Bahasa Indonesia dengan memanfaatkan tiga metode yaitu *Conditional Random Fields*, *Transformation Based Learning*, dan kombinasi kedua metode ini. Penelitian ini menggunakan korpus Bahasa Indonesia yang tersusun atas 49 buah artikel surat kabar dan terdiri atas total 13.465 buah token. *Tagset* yang digunakan pada penelitian ini terdiri atas 21 jenis *tag* (21 jenis kelas kata). *Tagset* ini merupakan modifikasi dari *Penn Treebank Tagset*.

Dari hasil penelitian diketahui bahwa hasil tagging tertinggi dengan metode *Conditional Random Fieds* mencapai 80,21%, dengan metode *Transformation Based Learning* 90,08%, dan dengan kombinasi kedua metode 86,24%. Berdasarkan hasil penelitian, metode *Transformation Based Learning* adalah

metode yang paling cocok untuk diterapkan dalam pembuatan *part of speech tagger* untuk Bahasa Indonesia jika dibandingkan dengan metode *Conditional Random Fields* dan kombinasi metode *Conditional Random Fields* dan *Transformation Based Learning*.

Kata Kunci:

Part of Speech Tag, Bahasa Indonesia, *Conditional Random Field*, *Transformation Based Learning*

