

Abstrak

Part of Speech Tagging merupakan proses penentuan *tag* kelas kata pada bahasa tertentu. Penentuan *tag* kelas kata menjadi hal yang sangat penting. Hal ini berkaitan dengan tingkat kebutuhan *user* dalam mengakses informasi. Maka, penulis melakukan penelitian untuk menghasilkan sistem *POS Tagging* berupa *POS Tagger* Bahasa Indonesia dengan menggunakan metode *Latent Descriptor Clustering*. Metode *Latent Descriptor Clustering* melakukan sebuah pendekatan dimana tiap token yang terdapat pada korpus akan didefinisikan dan diproses (*preprocessing*) menjadi vektor *latent* deskriptor. Deskriptor tersebut diakses dari konteks kiri dan konteks kanan pada tiap token di dalam korpus. Metode tersebut akan menghasilkan *output tag* kelas kata dan akurasi *tag* kelas kata. Hasil akurasi *tag* kelas kata dipengaruhi oleh *input* parameter σ dan konstanta. Akan tetapi, *input* parameter iterasi tidak berpengaruh terhadap hasil akurasi *tag* kelas kata. Selain itu, jumlah token yang semakin besar dapat menghasilkan hasil akurasi *tag* kelas kata menjadi lebih besar.

Kata kunci: *POS Tagging*, *latent descriptor clustering*, *preprocessing*, konteks kiri, konteks kanan, vektor *latent* deskriptor