

PERBANDINGAN PENGGUNAAN METODE PART- OF-SPEECH PADA BAHASA INDONESIA MENGUNAKAN METODE N-GRAM, TnT, DAN CLASSIFIER BASED

ABSTRAK

Abstract: *The lack of a corpus in Indonesian makes text processing in Indonesian have little interest in processing the text, therefore the aim of doing POS-Tagging is to do word labeling that takes into account the essence of the sentence based on the words contained in the sentence into a per-word tagset , the tagging process uses the N-gram, TnT, and Classifier Based methods. The three methods are used because the three methods have similar work methods and work by looking at the traces of learning and working on the previous method. The method that produces the optimal tagger of the three methods is the Classifier Based method because this method is a combination of the three methods, so that the tagger results issued are more optimal than the other two methods.*

Keywords: *Part-Of-Speech Tagging, N-gram Tagger, TnT Tagger, Classifier-based Tagger.*

Abstrak: *Kurangnya korpus dalam bahasa Indonesia membuat pengolahan teks pada bahasa Indonesia memiliki peminat yang sedikit untuk dilakukan pengolahan terhadap teks, oleh hal itu tujuan dari melakukan POS-Tagging adalah melakukan pelabelan kata yang memperhitungkan inti dari kalimat berdasarkan kata yang terkandung dalam kalimat menjadi tagset per-kata, proses melakukan tagging menggunakan metode N-gram, TnT, dan Classifier Based. Ketiga metode tersebut digunakan dikarenakan ketiga metode memiliki cara pengerjaan yang mirip dan bekerja dengan cara melihat jejak pembelajaran dan pengerjaan metode sebelumnya. Metode yang menghasilkan tagger optimal dari ketiga metode tersebut adalah metode Classifier Based dikarenakan metode ini merupakan gabungan dari ketiga metode tersebut, sehingga hasil tagger yang dikeluarkan lebih optimal dibandingkan dengan dua metode lainnya.*

Kata Kunci: *Part-Of-Speech Tagging, N-Gram Tagger, TnT Tagger, Classifier-based Tagger.*