

Abstract

Part of Speech Tagging is determining tag process of the language. Determining tag process become very important. It related for the user requirement to access information. So, the author will produces POS Tagging system, that is POS Tagger for Indonesian Language with Latent Clustering Descriptor method. Latent Descriptor Clustering method makes an approach where every token contained in the corpus has been defined and processed (preprocessing) used a latent descriptor vector. Descriptor is accessed from the immediate left context and right context of each token in the corpus. The method will produces tag output and tag accuracy. The result of tag accuracy is affected by parameter input σ and constant value. However, iteration input doesn't affect the result of tag accuracy. In addition, the number of tokens can produces tag accuracy which it becomes very great.

Keywords: *POS Tagging, latent descriptor clustering, preprocessing, left context, right context, latent descriptor vector*