Abstract

The increasing of Indonesian electronics news articles has made providing of the information in large number and unstructured. This condition make increasing of needs for searching and managing information well so can catch the valuable knowledge. Text mining is effort to take text form the data that is come from document, to search term that is represent content of documents so that can analyzing it.

NLP is used as preprocessing on text mining because user decide the releavance documents by reading and analyzing it. In this reseach, preprocessing are divided in 3 steps: POS tagging, syntax parsing, and semantic role labeling.

In this research, the impelemntation of semantic role labeling based on case grammar theory. Label that is used in this research is based on PropBank annotation. Semantic role labeling is consist of two main task, that is frame identification and labeling sentences with frame elements which is identified.

The result of semantic role labeling implementation can solve the problem of semantic role labeling. Compared with hand-labeled, testing with 52 sentences taken from Harian Kompas (300 articles) gives accurate labeling results for all words in that sentences.
.

Keywords: text mining, preprocessing, natural language processing, semantic role labeling, case grammar.