Abstract

The increasing of accessing news over the internet makes the easiness of accessing it has become a very important thing for the community. In order to simplify this process, there should be a news processing to make it well organized and efficient that can make the accessing of old news become relatively easier. One of the news processing is classification. Before performing the classification, we need to do the preprocessing data which one of this is stemming.

Stemming is the process for breaking any variants of a word into its basic word. One of stemming method is successor variety. Successor variety is a stemming method which is based on morpheme boundaries and uses the knowledge of structural linguistic. There are several methods of successor variety stemming, some if them are cutoff and entropy method. Both the stemming methods was applied to a system and be observed how it is perform.

Result from implementation of both methods is the comparation of the methods performance which based on the data used, the entropy method is better than the cutoff method from the accuracy, but from the ICF, cutoff is better than entropy. But in term of classification, both methods have an equivalent performance.

Keywords: classification, successor variety stemming, cutoff, entropy