

Abstract

Ambiguous words or sentences are a problem found in the use of language both in verbal and textual contexts. A word is ambiguous if it has more than one meaning. In a sentences, the meaning of words can be determined by the structure of the sentence. In Natural Language Processing there is a method that can search for a word's meaning with an exact calculation process that is called word sense disambiguation. This research focuses on the advance of naïve Bayes algorithm in Indonesian sense of word disambiguation. The choices of naïve bayes algorithm is because naïve bayes can produce good accuracy as part of the supervised method. This research collected data from the Indonesian corpus. The corpus was taken from social networks, online newspapers and made own ambiguous sentences. In simplifying the classification of this research using the python programming language. In python there is a module that is used to find certain string patterns called regex. The purpose of this study is to classify, also to analyze the application of the Naïve Bayes algorithm into word sense disambiguation in Indonesian language.

Keywords: ambiguous words, Indonesia language, natural language processing, word sense disambiguation, supervised, naïve bayes classification