

Abstract

Natural Language Processing (NLP) is a chapter of computer science which is focused on natural language/human language processing. Mainly task in NLP like Question Answering, Semantic Role Labeling, and Information Extraction require 5W (Who, What, Where, When, Why) and 1H (How) for extracting needed information. Semantic argument classification is arguments labeling process based on semantic role.

Some features are needed in semantic argument classification process. In this research, base feature and 3 additional features i.e. Named Entities in Constituent, Head Word POS, and Syntactic Frame are used. Those features have been proven to improve system accuracy, For the classification process, Sequential Minimum Optimization (SMO) algorithm is used, SMO is development algorithm from Support Vector Machine (SVM). SMO algorithm can handle multi-class problem and be able to do learning process faster than SVM.

Based on testing that was done, the use of three additional features Named Entities in Constituent, Head Word POS, and Syntactic Frame can improve accuracy results with an increase of 11.82%.

Keywords : semantic argument classification, Sequential Minimum Optimization (SMO), Natural Language Processing (NLP), Named Entities in Constituent, Head Word POS, Syntactic Frame