ABSTRACT

Getting a semantic argument representation of a sentence is important in text mining, such as information extraction, and question answering. In semantic role labeling to obtain high accuracy results required training data in large quantities. It can also be influenced by the selection of features that will affect the performance, in other words also affect the recall and precision are produced. The problem today is not any training data corresponding to the annotation needed for training are available in large numbers and what features will be added if it can improve the performance of semantic role labeling. The final task is to try to apply a process called adaptation domain that has been done in previous studies that showed an increase performance and add new features in the classification process arguments that have existing and proven to improve performance generated. Adaptation methods domain used in this thesis is the transfer of self-training and features used in this thesis is the Baseline Feature plus Noun Head of PP, First Word in Constituent, syntactic Frame, Argument Order, Constituent, First and Last Word / POS in Constituent that will be used in the classification of a semantic argument. The results of this study indicate that the domain adaptation process can improve the performance of the classification of the argument, although not too significant, and not all of the new features added can improve the performance classification of semantic argument, there are some features that it will decrease the classification performance when added. Based on the average of all scenarios, obtained an accuracy of 76.09% and f-score 74.60%. The results obtained from the use of domain adaptation with transfer self-training methods and the proof reading of new features to the baseline features that exist.

Keywords **:** *Domain adaptation, Baseline Feature, Semantic role labelling, Classification, transfer self-training*