# ABSTRACT

This thesis aims to categorize expert into categories of domain specific classification system, to compare the performance of two common term weighting algorithms in selecting features and determine the number of training texts required for building the category model. This thesis introduces the framework of broad classification model and specific classification model in expert classification. Broad classification model is examined with two common term weighting algorithm and various sizes of training texts for building a category model.

The performance of expert classification system improved as the number of training texts increased. When selecting categories related to an expert, we found that 3-cat-per-bib approach is better than 1-cat-per-bib approach. The accuracy of broad classification model is 50% and the accuracy of specific classification model is 48.19%. The main limitation of this research is the number of category used in evaluation does not cover all the expertise of the expert. This affects the accuracy of classification when expert indicates expertise which is out of our selected categories.