Abstract

Community question answering (CQA), like Qatar Living Forum has become a place for Internet users to get information. One question on the CQA can have many answers, but the user must choose the most appropriate answer manually which takes a long time. This problem can be overcome by establishing a answer ranking system from list of answer.The appropriate answers to the questions above answers are not in accordance with a question, it can help users to get the best answer quickly.

The first step in this study is text preprocessing which includes Elimination tag removal ,case folding, tokenizing, filtering, and stemming. After the text data become more structured next process is feature extraction where the feature are textual features and topic modelling . Textual features is identifying the characteristics of an answer by looking at the elements of the text like to see if an answer containing a question mark (?), emoticons, links or special words. Topic modelling is textual data modeling aimed at finding hidden variables. In this research will focus on the use of topic modelling to find similarity between questions and answers. This feature extraction results will be used as input to classfier to create a model that will be used by the data testing. In this final project using Support Vector Machine (SVM) to get a score classification where these score would rank the answers to every question. The difference this study with research belongs jaist is the process of ranking on the answer. JAIST's own system only until the classification process there is no rating system only answer.

Based on the evaluation of research by the author found that the rating system that is made has a mean avarage precision value of 71.9% . The results obtained using logistic regression as classifier. if compared with the results of SemEval 2016 the results obtained in this study was ranked eighth from 13 participant of SemEval 2015 task 3 .

Key words: community question answering, 1extual feature, topic modelling, feature extraction, rank