ABSTRACT

Technological advances, high number of internet users, and human nature as a social being become several factors in emersion of online forum discussion. Those kind of websites ease people in finding solution to their problem. But unfortunately, there are so many bad answers. Furthermore, the rating of an answers is given by human that makes it kind of subjective. This research will try to develop an answer ranking system. Dataset is obtained from SemEval 2017 Task 3 Subtask A. Ranking to an answers will be given through several processes, such as preprocessing data, and feature extraction. Then, preprocessed data will be classified using SVM. The score that is obtained from the classification process will be used as the ranking. Performance evaluation in this research will use the same as what is being used by SemEval 2017, that is Mean Average Precision (MAP). The system that is successfully developed here has a quite good performance, with an MAP score of 72.3%. Compared to the official SemEval 2016 result, this system is ranked 8th out of 13 contestants.

Keywords: *question answering system, semantic similarity, answer ranking, support vector machine, mean average precision*