Nathaphol Leungsuwan 526874 **KAGGLE**: nleungsuwan **Team name**:Dsc2017

Contribution report

I looked after the entity extraction and answer ranking for the basic system. The other two team members focused on sentence retrieval. When the result of sentences retrieval was well enough, I merged it together with answer ranking and tried to improve the basic system until it reached the baseline. At the same time, the other two team members were trying to improve their sentence retrieval part. Some of the key issues of the basic system were speed and precision of TF-idf; therefore, there was a need for a more powerful and robust method. I searched through lecture notes and internet for an improvement. Then, I found BM25 method and believed that the method could improve sentence retrieval. Therefore, I suggested the others to implement the method. Apart from BM25, they also found Language Model method, but the performance of BM25 was the best. Rather than using pure BM25, my team members were afraid of performance issues, so they implemented inverted index on BM25. Once, the basic system reached the baseline, our team had a discussion on which direction the enhancement should be. The conclusion of the discussion was to go out and tried out methods as many as we could, before assembling the best methods into the enhanced system. For sentence retrieval, BM25 was agreed to be the best option. I also tried to improve BM25 by looking at the context around the retrieved sentence, but there was no significant improvement. Then, I was interested in question type classification using various methods: Multinomial Naïve Bayes, Logistic Regression, Decision Tree, Random Forest and Neural network. I tested on training set and evaluated using development set. Out of these 5 methods, Logistic regression has the highest performance (f-1 score) and the other team members also found the result in the same fashion. Therefore, Logistic Regression was selected.

Towards the end of the project, a member was responsible for assembling every part together as a prototype enhanced system. The other member was responsible for cleaning the code and compatibilities and make sure that the system was runnable on computer lab's environment. I was trying to improve the system till the last minute. What I was looking at was trying to see relationships among scores assigned to potential answers, but there was not enough time to use a machine learning technique to derive patterns. Lastly, every member including me put together ideas and concepts of enhancement into a group report.