OUC at the NTCIR-16 QA Lab-PolilInfo-3 Budget Argument Mining

Keiyo Nagafuchi1, Rin Sasaki2, Seiya Oki2, Yasutomo Kimura2 and Kenji Araki3

Graduate School of Information Science and Technology, Hokkaido University, Japan 2 Otaru University of Commerce, Japan 3 Faculty of Information Science and Technology, Hokkaido University, Japan

1. Our methods

The Budget Argument Mining task consists of two subtasks, including Argument Classification (AC) and linking relatedIDs (RID). We separately proposed several methods to perform AC or RID, and combined them.

2. Our results

2.1 Discussion: AC

• BERT classifiers obtained higher scores than rule-based and BoW-based ones. Considering context was effective for this subtask.

• We counted the number of misclassifications for all methods.
  - Missclassification rates of "Regular classes" were low, but the rates of other classes were high.
  - Training and fine-tuning did not go well because data sets of this task were imbalanced.

3.2 Discussion: RID

• It is likely that poor results of SBERT were attributed to the fact that budget item descriptions were often omitted in the remarks.

• Most utterances contained keywords related to budget items in preceding and following contexts of monetary expressions.

• It is likely that the TF-IDF obtained good results in linking RID.

• Utterances that were answered incorrectly with TF-IDF did not contain keywords.

• Keywords were included in the preceding and following sentences.

• In the future, we should consider a system that also considers the surrounding sentences.

4. Conclusion

• We separately proposed several methods to perform AC or RID, and combined them.

• Among our methods, the combination of BERT base classifier and TF-IDF, modals model obtained the highest score (1.4468).

• This method got 1st place on the leaderboard of overall score (0.5712).

• BERT base classifier got 2nd place on the leaderboard of AC score (0.5696).

• TF-IDF, modals got 1st place on the leaderboards of RID score.

• Because only one utterance sentence was used as input for our systems in this work, it is necessary to develop a system that could consider the surrounding context in the future.