We report a Japanese subtask for NTCIR-13 STC-2 for which we made a dialogue system and introduced neural network-based retrieval models (LSTM, ESIM and CNN) to rank the dialogue replies in the training dataset. We used data from Yahoo! News comments data and introduced LSTM and ESIM to effectively capture sequential information from the given comments. To evaluate the effectiveness, we compared systems using LSTM or ESIM with systems that use CNN. We also introduced an n-gram-based statistical filter into our systems to reduce the number of reply candidates.

Method: Candidate ranking

Ranking Model (Dual Encoder (R Lowe 2017))

We can see attention mechanism is improving the result, but the candidate filtering we proposed decreased the accuracy. Probability calculation method was inappropriate.

Method: Reply candidates filtering

1. Candidates Filtering with simple rules
   1. Remove the candidates that contain the same words more than two times
   2. Remove the candidates that contain more than two sentences
2. N-gram based candidates filtering
   1. Extract top 100 tri-gram combinations which appears at the end of the sentence
   2. Extract top 100 sentences containing ones in list from 1.

References: