1. Overview

- The overall approach and the architecture
  - Retrieval-based approach, utilizing all the 1.2M comments in the training data
  - LSTM-DSSM
- Two runs
  - YJTI-J-R1
    - Trained by Twitter conversation data
  - YJTI-J-R2
    - Mainly trained by Yahoo Chiebukuro QA data

2. Runtime System

3. Model

- 3-layer LSTM RNN with a fully-connected layer
  - Two models: a comment encoder model and a reply encoder model
  - LSTM's hidden layer size: 1024
  - Embedding layer size: 256
  - Representation size: 1024

4. Data

<table>
<thead>
<tr>
<th>name</th>
<th>type</th>
<th>no. of records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twitter LM</td>
<td>posts</td>
<td>100.0M</td>
</tr>
<tr>
<td>Twitter conversation</td>
<td>pairs</td>
<td>65.1M</td>
</tr>
<tr>
<td>Y! Chiebukuro LM</td>
<td>posts</td>
<td>202.0M</td>
</tr>
<tr>
<td>Y! Chiebukuro QA</td>
<td>pairs</td>
<td>66.3M</td>
</tr>
</tbody>
</table>

5. Analysis and Conclusions

- Official STC results of our runs (Rule-2)
  - Effectiveness of DSSM-like approaches combined with large-scale linguistic resources
  - Social QA data can be useful for modeling topic-oriented conversations