Classification Task in QA Lab-PolInfo

- Determining the class (“SUPPORT”, “AGAINST” or “OTHER”) of statements in assembly minutes (Japanese) according to Subtask 1, 2 and 3 (shown in the following table).
- Subtask 1: Identifying whether statements are relevant to a topic or not.
- Subtask 2: Identifying whether statements are fact-checkable or not.
- Subtask 3: Identifying stances (positive, negative or neutral) of statements.

Although the number of examples used in argumentation mining is generally around 1,000 to 2,000, there are over 10,000 annotations provided for this task.

<table>
<thead>
<tr>
<th>Subtask 1</th>
<th>Subtask 2</th>
<th>Subtask 3</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>relevant</td>
<td>fact-checkable</td>
<td>positive</td>
<td>SUPPORT</td>
</tr>
<tr>
<td>relevant</td>
<td>fact-checkable</td>
<td>negative</td>
<td>AGAINST</td>
</tr>
<tr>
<td>relevant</td>
<td>fact-checkable</td>
<td>neutral</td>
<td>OTHER</td>
</tr>
<tr>
<td>relevant</td>
<td>non fact-checkable</td>
<td>any result</td>
<td>OTHER</td>
</tr>
<tr>
<td>not relevant</td>
<td>any result</td>
<td>any result</td>
<td>OTHER</td>
</tr>
</tbody>
</table>

Topic example: 築地市場の豊洲移転 (Tsukiji market should be moved to Toyosu)

Statement: 豊洲は、新市場移転により千客万来の施設ができ、今後、観光客の集客が大いに期待できるエリアであります。Toyosu is an area which, after moving the new market there and building facilities capable of hosting thousands of people, could be expected to gather many tourists in the future.

Introduction

- The three stances (positive, negative or neutral) play an important role in recognizing arguments in a minute.
- In recognizing arguments in a minute, the stances play an important role.
- For fact-checking, it is crucial to understand whether an argument is fact-checkable or not.
- Tested: LSTM and BiLSTM
- Compared: machine learning vs. rule-based
- Dataset ratio: Training: 80%, Test: 20%

Subtask 1: Relevance between the Topic and Statements

- Baseline: All statements marked as “relevant”.
- Common Words: When the number of common words between topics and statements exceeds 2 (except for hiragana and stop words), we regarded them as “relevant”.
- Similarity: If cosine similarity between topics and statements is over a threshold, we regard them as “relevant”.
- LSTM and BiLSTM input: topics and statements

Subtask 2: Finding Verifiable Fact in Statements

- Baseline: All statements marked as “fact-checkable”.
- BiLSTM input: statements

Subtask 3: Identifying Stances of Statements

- Baseline: All statements marked as “neutral”.
- BiLSTM input: statements

Sentiment

- Baseline: 49.19%
- BiLSTM: 87.83%

Discussion

- BiLSTM method yields the highest accuracy in all the subtasks.
- We only used word vectors of topics and statements, so in the future we plan to design better features such as the ones we use in a Semantic Approach.
- In Subtask 3, we did not consider phrases like “賛成の意見 (a supporting opinion)” and “持論 (an argument)”.
- It would be better to annotate statements into five semantic relations: “AGREEMENT”, “CONFLICT”, “CONFIRMATION”, “EVIDENCE” and “OTHER” in [1] rather than “SUPPORT”, “AGAINST” and “OTHER” because it will show the structure of the argument more clearly.