STARS at NTCIR-14 QA Lab-PoliInfo Classification Task

Daiki Shirafuji[1], Sho Takishita[1], Patrycja Swieczkowska[1, 2], Rafal Rzepka[1, 2], and Kenji Araki[1]

[1]Hokkaido University, Sapporo, Japan

[2]RIKEN Center for Advanced Intelligence Project (AIP)

Classification Task in QA Lab-PoliInfo

- 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 about 1,000 to 2,000, there are over 10,000 annotations provided for this task.

Subtask 1	Subtask 2	Subtask 3	Class
relevant	fact-checkable	positive	SUPPORT
relevant	fact-checkable	negative	AGAINST
relevant	fact-checkable	neutral	OTHER
relevant	non fact-checkable	any result	OTHER
not relevant	any result	any result	OTHER

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るなど、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.)*

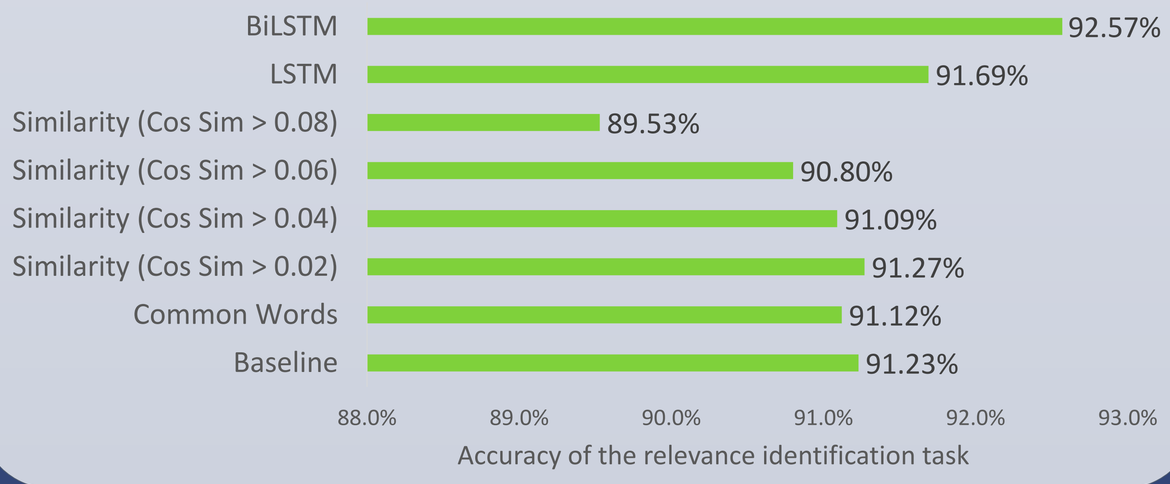
Subtask 1: relevant
Subtask 2: fact-checkable
Subtask 3: positive

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

築地市場の豊洲移転で、200億円を超える3ヶ所の建築物が入札不成立になる
(Three buildings costing over 20 billion yen were not tendered due to the moving of Tsukiji Market to Toyosu) […]

When numerals are included, we regard statements as "fact-checkable".

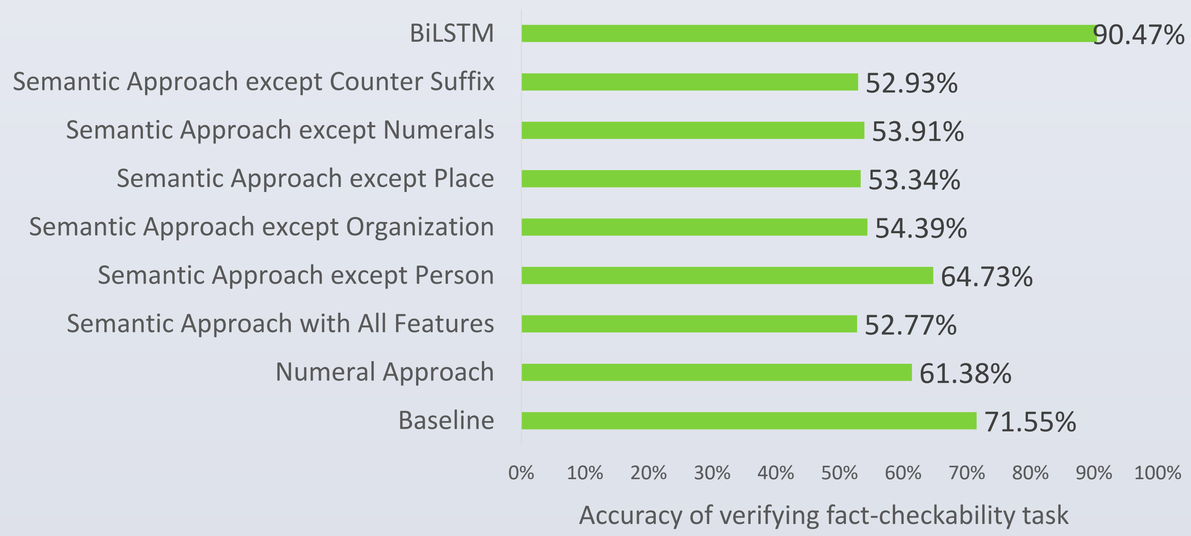
Numeral Approach

Semantic Approach

Feature Name	Example
Place	築地 *(Tsukiji)*, 京都 *(Kyoto)*
Person	Abe, Ueda
Organization	House of Representatives, Cabinet
Numerals	1, 一 (one)
Counter Suffix	円 (Yen), 人 (number of people)

Can provide useful clues

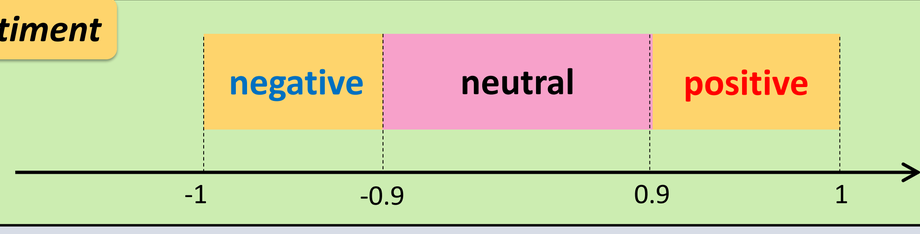
When any of the features are included, we regard statements as "fact-checkable".

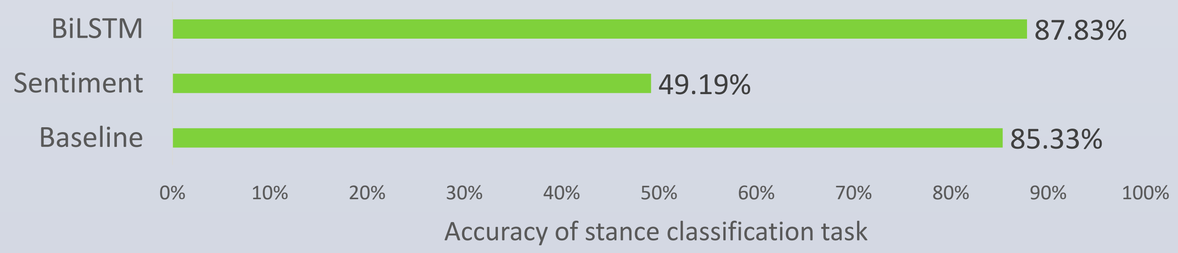


Subtask 3: Identifying Stances of Statements

- Baseline: All statements marked as "neutral".
- BiLSTM input: statements

Sentiment





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 否の立場です (*I dissent*). In the next step, we will employ this feature into BiLSTM.
- It would be better to annotate statements into five semantic relations: "AGREEMENT", "CONFLICT", "CONFINEMENT", "EVIDENCE" and "OTHER" like in [1] rather than "SUPPORT", "AGAINST" and "OTHER" because it will show the structure of the argument more clearly.

[1] Koji Murakami, Eric Nichols, Junta Mizuno, Yotaro Watanabe, Shouko Masuda, Hayato Goto, Megumi Ohki, Chitose Sao, Suguru Matsuyoshi, Kentaro Inui and Yuji Matsumoto.: Statement Map: Reducing Web Information Credibility Noise through Opinion Classification. In: Proceedings of the fourth workshop on Analytics for noisy unstructured text data. ACM, 2010. p. 59-66.