



HOKKAIDO UNIVERSITY

Summarizing Utterances from Japanese Assembly Minutes using Political Sentence-BERT-based Method for QA Lab-PoliInfo-2 Task of NTCIR-15

SKRA team

**Daiki Shirafuji¹, Hiromichi Kameya²,
Rafal Rzepka³, Kenji Araki³**

¹ Graduate School of Information Science and Technology,
Hokkaido University, Japan

² School of Engineering, Hokkaido University, Japan

³ Faculty of Information Science and Technology,
Hokkaido University, Japan



Introduction

Our Assumptions for the Summarization Subtask

1. Embedding model which is **trained on political domain** could be useful for this task;
2. MMR score could be useful for extracting **diverse sentences** from documents



Main Contribution

1. **Proposing Japanese Political Sentence-BERT for sentence embedding model;**
2. Adapting an embedding-based unsupervised key-phrase extraction, EmbedRank++, to summarization
3. Adding **two similarity** functions to the MMR score which is used in EmbedRank++

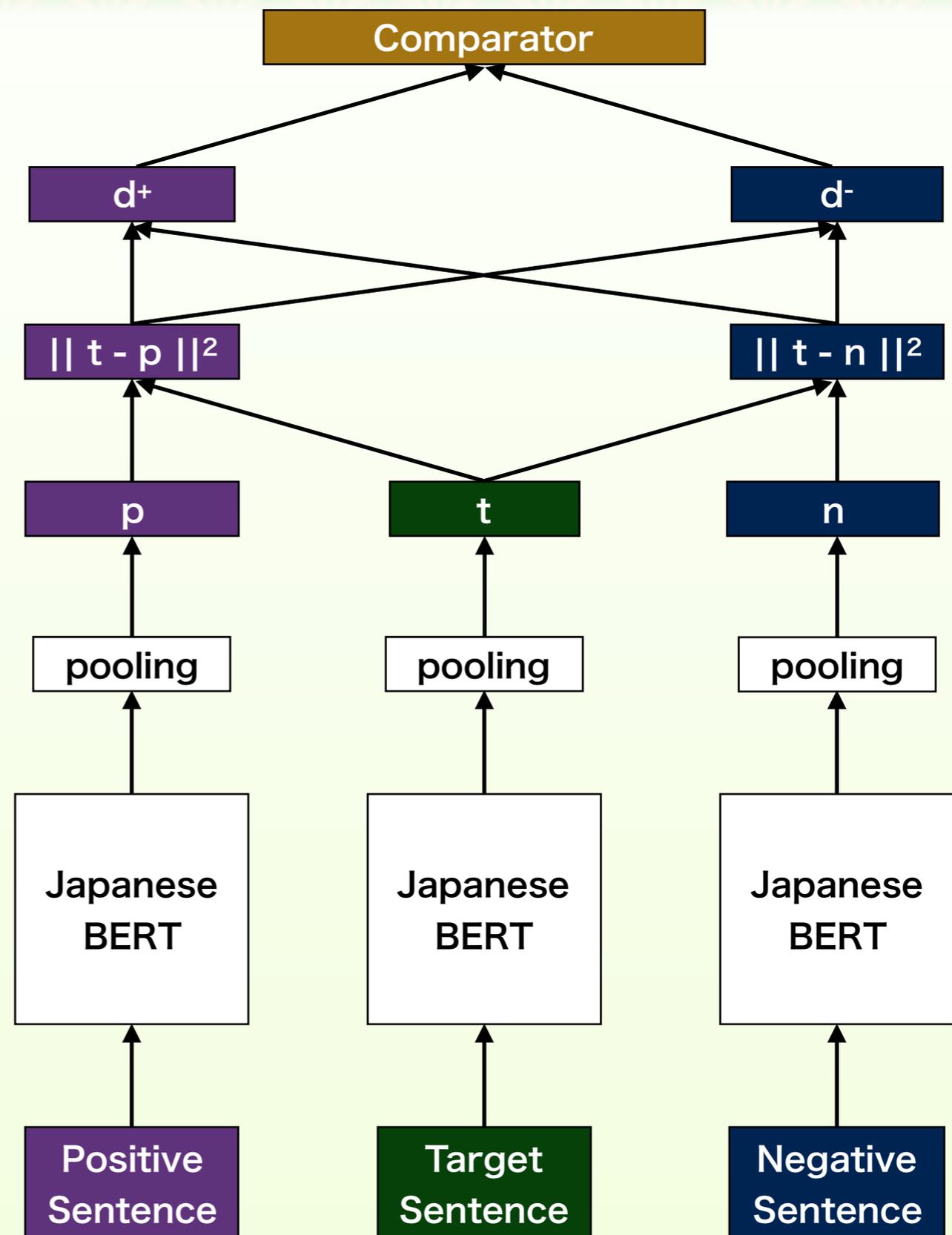
Japanese Political Sentence-BERT Creation

JPSB (Sentence Embedding Model)

1. Insert Japanese BERT into a triplet network
2. Input target and positive / negative sentence
3. Embed inputted sentences with the BERT modules, inside a triplet network, using **MEAN** pooling;
4. Training on dataset via making **the distance $d+$ as 1 (positive)**, and **the distance $d-$ as 0 (negative)**.

Dataset

Based on [the dataset provided by the QA LabPolInfo-2 organizers](#).
27,078 triplets of [target, positive, negative].



JPSB Model Evaluation

Evaluation Metrics:

1. *diff*:

take the difference between the cosine similarity of [positive sentence, target sentence] and [negative sentence, target sentence]. *The larger *diff* indicates that the model can identify a positive sentence and a negative sentence better.

2. Accuracy

The results show that our JPSB works better than one of common sentence embedding models, i.e. USE.

	<i>diff</i>	Accuracy
USE	0.2441	0.8674
JPSB	0.3705	0.9849

Our Model for Summarizing Utterances

EmbedRank++ based Method

- ▶ To generate various sentences as summaries of utterances, we adopt EmbedRank++, which is based on MMR score.
- ▶ To compute similarities, **JPSB** is adopted.
- ▶ We newly **add cosine similarities between [a given utterance, Main Topic] and [a given utterance, Subtopic]**.

JPSB is used to compute

$$\begin{aligned} MMR = \arg \max_{D_i \in R \setminus S} \{ & k * \{ 0.5 * \text{CosSim}(D_i, Q) \\ & - 0.5 * \max_{D_j \in S} \text{CosSim}(D_i, D_j) \} \\ & + m * \text{CosSim}(D_i, MT) \\ & + s * \text{CosSim}(D_i, ST) \}, \end{aligned}$$

Additional Functions

$$\begin{aligned} \text{key_size}(\text{Answer}) &:= \frac{\text{Answer length}}{50}, \\ \text{key_size}(\text{Question}) &:= \frac{\text{Question length}}{50}. \end{aligned}$$

Parameters k , m and s are set to 0.2, 0.3, and 0.5, respectively.

R : the ranked list of sentences retrieved by an algorithm

S : the subset of sentences in R , D_i and D_j are retrieved sentences,

Q : the averaged vector of inputted all sentences

$\text{key_size}(\text{Answer})$ and $\text{key_size}(\text{Question})$ refer to the number of outputted sentences

Our Model for Summarizing Utterances

EmbedRank++ based Method

- ▶ To generate various sentences as summaries of utterances, we

**Extract the top-*[key size]*
sentences as outputs**

$+s * CosSim(D_i, ST)\}$,

$$key_size(Answer) := \frac{Answer\ length}{50},$$
$$key_size(Question) := \frac{Question\ length}{50}.$$

Additional
Functions

Q: the averaged vector of inputted all sentences

$key_size(Answer)$ and $key_size(Question)$ refer to the number of outputted sentences

Evaluation Results

- ▶ Regarding ROUGE-1 scores, Sentence-BERT-based approach does not outperform USE-based one.
- ▶ We cannot say that JPSB is also useful for summarization task.

	ROUGE-1 (Recall)
USE-based EmbedRank++	0.0846
JPSB-based EmbedRank++	0.0696

Evaluation Results

- ▶ **Blue sentences** express good summaries, but **red one** does not.
- ▶ This is because the **red sentence** is the closest to central embedding.
→ We need to tune the parameters for EmbedRank++ scores.

Reference Summary #1	都の産業政策やインフラ整備との一体的取組、川崎港・横浜港との連携に加え、都が責任を持って経営に関わる体制を確保。(We should ensure that the Tokyo Metropolitan Government will take responsibility for management of the Port of Kawasaki and the Port of Yokohama, in addition to coordinated efforts with Tokyo's industrial policies, infrastructure development and cooperation with the ports of Kawasaki and Yokohama.)
Output #1	東京の産業政策やインフラ整備と一体的に取り組むことが求められております。/現場の実態を熟知した東京都が責任を持って港湾経営にかかわっていく体制を確保してまいります。(At the same time, we need to work on Tokyo's industrial policies and infrastructure development. We should ensure that the Tokyo Metropolitan Government, which knows well the actual situation in the field, will take responsibility for port management.)
Reference Summary #2	都道について新たな整備目標示す。区市町村支援、国への財源拡充要求、電線事業者等との連携を強化し積極的に推進する。(Let us explain the new purposes for Tokyo metropolitan roads. We will actively promote support for municipalities, also demand expansion of financial resources from the government, and strengthen cooperation with wire-line operators.)
Output #2	一層の支援に努めてまいります。/風格ある都市景観の形成と高度防災都市の実現を目指し世界に誇れる都市空間を創出してまいります。(We will make every effort to support the city. We will create a world-class urban space with the aim of creating a distinctive urban landscape and a highly disaster-resistant city.)

Conclusion

- ▶ We created Sentence-BERT for Japanese political texts.
- ▶ We adopt the Japanese Political Sentence-BERT to utterances included in the Japanese minute data summarization task.
- ▶ Our sentence embedding model did not work well for this subtask (both ROUGE and human evaluation).
- ▶ **It is necessary to tune the parameters of the score function in our method.**



Thank You for Your Attention!

