An Abstractive Dialog Summarization System for Japanese Assembly Minutes

Kazuma Kadowaki The Japan Research Institute, Limited

Abstract

I developed a system for topic-aware summarization of assembly member speeches. It consists of:

1. A pre-processor
2. A BERT-based sentence extractor that predicts a topic-aware importance of each sentence;
3. A UniLM-based summary generator whose summary length is controllable.

My model achieved the best performance among all the participants in the Dialog Summarization subtask.

INTRODUCTION

Purpose

Generate a topic-aware short summary of Tokyo Metropolitan Assembly minutes, in order to fact-check and to understand speakers’ policies.

Task

Input: speaker’s name, entire speech, topic, desired length
Output: a topic-aware summary of the speech

Challenges

- Very long, multiple-topic speeches
- Minutes without annotation (not segmented, no importance scores)
- Maximum numbers of characters specified for each summary

MY APPROACH

Pre-processor

- Retrieves an entire “source speech”

Sentence Extractor

- Extracts a “source passage” by predicting ROUGE-1 scores using BERT-based regression

Summary Generator

- Generates an abstractive summary from the passage using UniLM (modified to control the length)

RESULTS

Models submitted

- ID 185: trained only using the datasets from the task organizers
- ID 189: trained also using my own dataset from different years

Results

- Achieved better performance in most of the metrics
- Adding my dataset further contributed to the performance

DISCUSSIONS

Performance of each module

- Each module extracted 51.7% and 75.7% of the available content words successfully

Human evaluation

- No system seems to be always helpful to fact-check
- Future work: Revise the task settings

Model generalization

- Robust enough for changes in topics discussed
- Future work: Mitigate/detect performance degradation

CONCLUSIONS

Contributions

- I developed an assembly minutes summarizer, which consists of a BERT-based extractor and a UniLM-based generator
- My models achieved the best performance, and would generalize for future meetings
- The length of a generated summary can be controlled

Future work

- Add a mechanism to consider a context
- Apply my models to other real-world tasks (including business conversations)
- Revise the task settings for fact-checking
- Investigate summarization from noisy minutes generated by ASR systems

Copyright © 2020 The Japan Research Institute, Limited

QA Lab-PoliInfo-2 (Dialog Summarization) The 15th NTCIR Conference. Dec 8–11, 2020, Tokyo, Japan