SumRec: A Framework for Recommendation using Open-Domain Dialogue

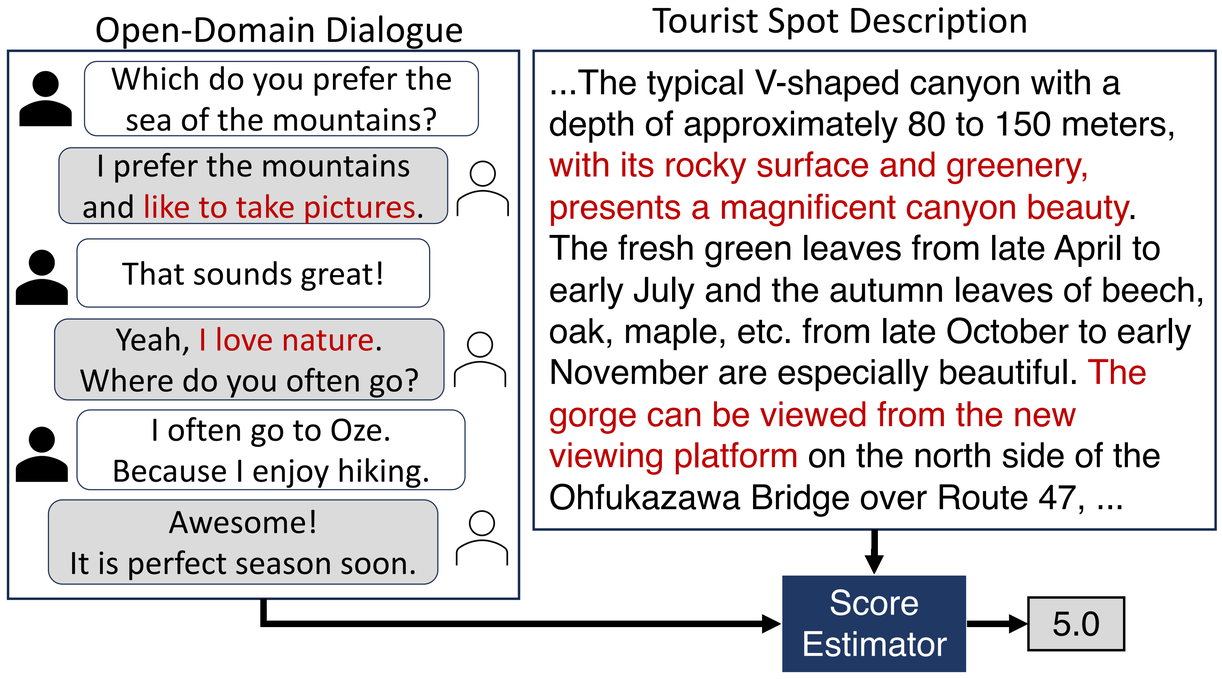
Ryutaro Asahara, Masaki Takahashi, Chiho Iwahashi, Michimasa Inaba.
The University of Electro-Communications.

Introduction

- Chat dialogue has a lot of the speaker's information such as interests, preferences, and experiences.
- This information can be used to personalize and suggest advanced information in various systems, although it's not well used.
- This study suggests **a novel recommendation task based on open-domain chat dialogues**.
- We proposes a **SumRec** framework using speaker **Sum**mary and **Rec**ommendation information to manage our task.
- To facilitate analysis of the recommendation task, we construct **a dataset called ChatRec to recommend tourist spots based on chat dialogues.**

Proposed Task and Data Example

- Predict speaker's preference score using open-domain dialogue and tourist spot description.

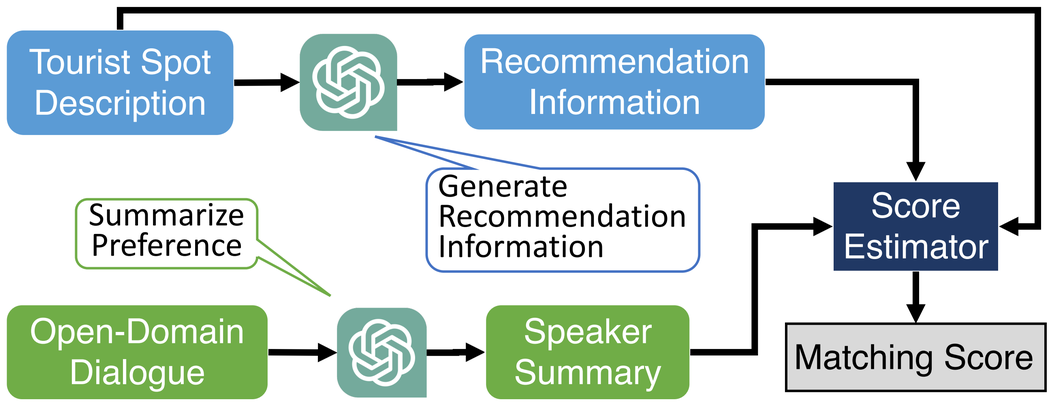


Challenges in Our Task

1. **Open-domain dialogue has unnecessary information for recommendation**
   - EX) Greeting, Filler, Interjection.
   - It might be noise for the score estimator.
2. **The description of a tourist spot usually focuses on its features**,
   - The description doesn't have enough information about the types of visitors.
   - It makes linking the user's preference and the feature of tourist spots difficult.

Proposed Method(SumRec)

- Speaker Summary
  - Extracting and summarizing speaker's preference, it eliminates some noises.
- Recommendation Information
  - ChatGPT generates a sentence what kind of person you should recommend.
  - This promotes the model link speaker and tourist spot information.
- Score Estimator
  - We use speaker summary, tourist spot description, and recommendation information as inputs for the score estimator, which then provides a predicted evaluation score.
  - The score estimator is available for any language model regardless of its properties.
  - We adopted **RoBERTa** and **ChatGPT(5-shot)** as score estimator in this experiment.



Speaker Summary

She states that train travel is a relaxing way to travel. She likes to take pictures and finds mountain scenery calming. She also says that a trip to ...

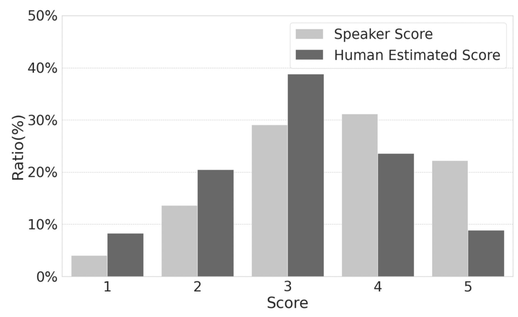
Recommendation Information

... Especially when you visit during the seasons of fresh greenery and autumn leaves, you can enjoy an exceptional view. Perfect for photographer and those who love to feel nature. It is also recommended for those who want to enjoy hiking, and there are courses such as Ofukazawa Promenade.

Data Collection

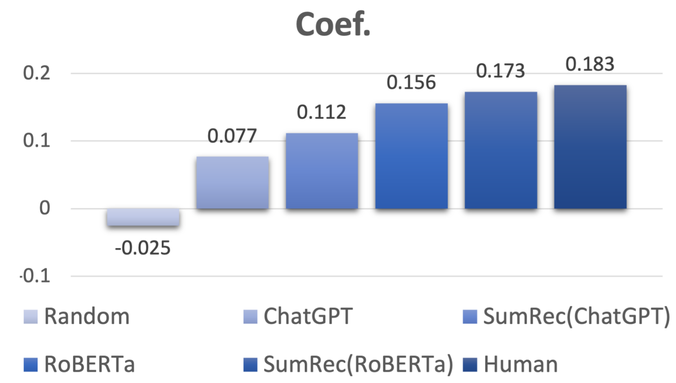
- We created a dataset for our task using 3,290 tourist spots, 1,005 two speaker's dialogue, and speaker's preference scores.
- For collecting the ground truth score, speakers evaluated each tourist spot after conversation.
  - We assigned 10-20 tourist spots to each speaker.
- This dataset has human estimated score to compare with our method.

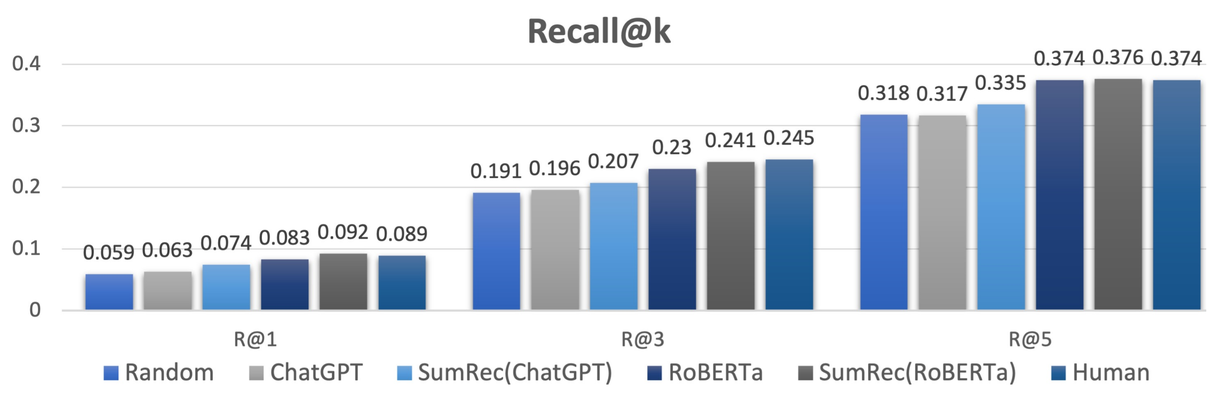
Dialogues	10005
Utterances	21982
Words per Utterances	23.44
Spot per Dialogues	15.73



Experiment and Results

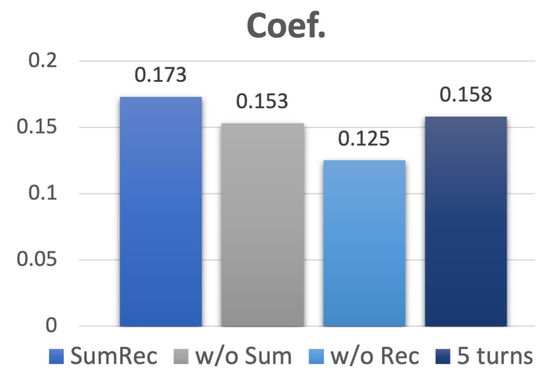
- We create a ranking using the speaker's preference score for a tourist spot as predicted by the score estimator.
- In this experiment, we evaluate this ranking using the following evaluation metrics:
- Recall@k
  This metric calculates the proportion of top-k items correctly ranked.
- Coef.(Spearman's Coefficient)
  This metric measures the relation between a model's predicted ranking order and the actual order of correct responses.

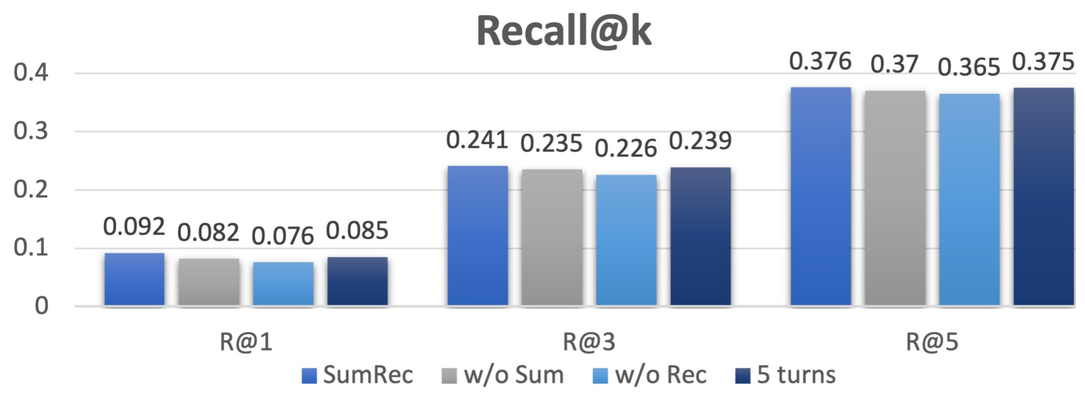




Ablation Study

- w/o Sum : This model does not use speaker summary but it uses dialogue directly.
- w/o Rec : We remove recommendation information.
- 5 turns : This model use speaker summary consists of the initial 5 turns (10 utterances) in dialogue.





Conclusion

- We proposed the framework SumRec for recommendation using open-domain dialogue.
- Experimental results indicated that our framework exhibits excellent performance on all the considered evaluation metrics.
- Nevertheless, travel-related dialogues continue to show considerable potential for improvement.
- We plan to explore more efficient methods by incorporating prior knowledge about the trip and integrating external knowledge into the model.

Our code is available on GitHub github.com/Ryutaro-A/SumRec

Inaba Lab. The University of Electro-Communications
The University of Electro-Communications