



## Learning from Free-Text Human Feedback in Task-Oriented Dialogues

### Motivation

Free-text human feedback, i.e., a user describing to the model what they did not like about the last generated utterance, is one of the most important sources of knowledge to iteratively improve language generation models, such as GPT or LLaMA, but datasets for research are scarce. For this reason, we used synthetic data generation methods to create a task-oriented dialogue dataset annotated with artificial free-text human feedback. However, since synthetic data is subject to various limitations, such as hallucinations or biases, we are looking for a thesis student to research and apply state-of-the-art data augmentation and cleansing methods to improve the quality of our data.

### Task Description

- Improvement and extension of a large internal dialogue dataset annotated with free-text feedback using state-of-the-art data augmentation and cleansing methods.
- Analyze and compare the quality of the new dataset with the original dataset.
- Training and evaluation of state-of-the-art language generation models using the new dataset.

### References

- Petrak, D., Moosavi, N., Tian, Y., Rozanov, N., and Gurevych, I. (2023). Learning From Free-Text Human Feedback – Collect New Datasets Or Extend Existing Ones?. EMNLP, 2023, Singapore.
- Kumar, S., Balachandran, V., Njoo, L., Anastasopoulos, A., Tsvetkov, Y (2023). Language Generation Models Can Cause Harm: So What Can We Do About It? An Actionable Survey. EACL , 2023, Dubrovnik, Croatia.
- Ouyang et al. (2022). Training language models to follow instructions with human feedback. NeurIPS, 2022, New Orleans, USA.

### Contact

Analysis



Programming



Literature



Prof. Dr. Iryna Gurevych

M.Sc. Dominic Petrak

[thesis@ukp.informatik.tu-darmstadt.de](mailto:thesis@ukp.informatik.tu-darmstadt.de)