Yuxiang Xiao

EMAIL: yuxiangxiao02@gmail.com

Education

National University of Singapore Singapore
M.Sc. in Robotics *Aug.2024 – Expected May.2025*

South China University of Technology(SCUT) China
B.Eng. in Robotic Engineering (GPA: 3.60/4.00) *Sept.2020 – Jun.2024*
Bachelor's Thesis: "Noise-robust Zero-shot Event Detection Method"
Relevant Coursework: Robotics Technology, Autonomous Driving System, Natural Language Processing

Publication

Zhang, T., Zeng, Z.*, **Xiao, Y.**, Zhuang, H., Chen, C., Foulds, J., & Pan, S. (2024). GenderAlign: An Alignment Dataset for Mitigating Gender Bias in Large Language Models. arXiv preprint arXiv:2406.13925. ***(Under Review)***

Zeng, Z., Wu, R., **Xiao, Y.**, Zhong, X., Wang, H., Lu, Z., & Zhuang, H. (2024, May). Zero-shot Event Detection Using a Textual Entailment Model as an Enhanced Annotator. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation ***(LREC-COLING 2024)*** (pp. 17851-17857).

Li, F., Chen, J., Zhou, Z., Xie, J., Gao, Z., **Xiao, Y.**, ... & Zhou, Y. (2023). Lightweight soft robotic glove with whole-hand finger motion tracking for hand rehabilitation in virtual reality. ***Biomimetics(SCI, JCR-Q1)***, 8(5), 425.

Research Experience

Alignment Dataset for Mitigating Gender Bias in LLMs *Sept.2023 – Feb.2024*
Core Member | Advisors: Prof. Zeng of SCUT(China) and Prof. Foulds of UMBC(USA)

- Proposed an automated annotation scheme to generate an alignment dataset named *GenderAlign*to mitigate a comprehensive set of gender biases, the dataset consists of 8k single-turn dialogues.
- Collected seed texts that can provide relevant topics to initiate gender-related dialogues and potentially trigger unaligned LLM to generate gender-biased responses.
- Aligned the LLMs using various algorithms and evaluated the effectiveness of our datasets

Zero-shot Event Detection via TE Model as an Enhanced Annotator *Mar.2022 – Oct.2023*
Core Member | Advisors: Prof. Zeng of SCUT

- Built a zero-shot text event extraction system using Text-Entailment(TE) model.
- Annotated sentences via TE model to generate an event detection dataset with 100,000+ instances.
- Responsible for fine-tuning the TE model and reproducing baseline results.
- The experimental results show that our method can outperform SOTA by 15% on the ACE05 dataset.

Instruction Finetuning of ChatGLM-6B with LORA *Jun.2023 – Jul.2023*
Team leader | Course Project of Natural Language Processing

- Used effective prompts for instruction Finetuning to improve the robustness of ChatGLM-6B
- Achieve the diversity of prompts by designing and collecting 5,000 + human-made prompts with various domains
- Finetuned the model with LoRA and analyzed the impact of different instructions

Academic Experience

Attended the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation(*LREC-COLING 2024*), presented a paper with poster (*May 20 – 25, 2024*)

Honors and Awards

Second Prize in APMCM Mathematical Modeling Competition *2021*
First Prize in China Undergraduate Mathematical Contest in Modeling *2021*

Skills

Computer Skills: Python(Pytorch), Matlab, C/C++, Shell Script, LaTeX
Languages: Mandarin Chinese(Native), English(IELTS:6.5)