YINHONG LIU

+44 7923592505 ⋄ yl535@cam.ac.uk

Homepage: http://www.yinhongliu.com

EDUCATION EXPERIENCE

University of Cambridge, Language Technology Lab; PhD **2021.10 - present**

- Supervisor: Prof. Nigel Collier
- Research area: Large Language Model (LLM) Evaluation and Alignment. Text and Multimodal Evaluation Techniques. Controllable Text Generation. Data Synthesis for Alignment and Down-stream applications. Advanced Dialog Systems. Discourse Structure Analysis for Long-form Text.
- Scholarship: Cambridge Trust

University of Cambridge, Computer Lab; Mphil **2020 - 2021**

- Subject: **Advanced Computer Science, in Natural Language Processing**.
- Grade: **Distinction**.
- Research area: Learning Functional Distributional Semantics from Multi-modal data.

University of Cambridge, Engineering Department; BA, MEng **2014 - 2018**

- Subject: **Information Engineering and Computer Science**.
- Grade in Master of Engineering: **Merit degree** with honours.
- Grade in Bachelor of Arts: **1st Class**.
- Final year project: **Statistical enhancement of multi-channel audio data**.

WORK, RESEARCH EXPERIENCE

Mircrosoft Research; Cambridge. *Research Intern* **Jun. 2024 - Sep. 2024**

· Developed an advanced Automatic Prompt Optimization pipeline to enhance Model Migration efficiency and improve Model Routing for Microsoft Copilot teams.

AI/ML, Apple; Cambridge. *Research Intern* **Jun. 2023 - Oct. 2023**

· Developed an innovative automatic Task-Oriented Dialog generation pipeline, incorporating advanced analysis of future virtual assistant response styles and simulating complex apps interactions. This work benefits both internal Apple Intelligence development and external NLP community.

· Authored a research paper detailing the dialog generation methodology, which was accepted by ACL 2024.

Gambit Research; London. *Full-time Quant-analyst* **Oct. 2018 - Jul. 2020**

· Working on various statistical football teams & games modelings, market price & volume predictions, strategies developing and back-testing.

· Extensive experience in various programming environments, including Python, SQL and BASH.

Jigsaw Unintended Bias in Toxicity Classification, Kaggle competition Jun. 2019

· Got a 225th place out of 3165 teams (top 8%), a bronze medal, as a solo participator.

- Comments that mention certain minority groups, like black people, women and Asians, have higher chance to be toxic. Therefore, challenge exists in classify those comments' toxicity correctly regardless of their bias in the minority groups.
- Applied the latest language models, including LSTM, BERT and GPT-2. Designed a custom loss function based on the evaluation metrics. Model implemented in PyTorch and trained across multiple cloud GPUs.

PUBLICATIONS

- **Yinhong Liu**, Han Zhou, Zhijiang Guo, Ehsan Shareghi, Ivan Vulic, Anna Korhonen and Nigel Collier. Aligning with Human Judgement: The Role of Pairwise Preference in Large Language Model Evaluators. COLM 2024.

- **Yinhong Liu**, Yimai Fang, David Vandyke and Nigel Collier. TOAD: Task-Oriented Automatic Dialog generation pipeline with diverse system response style. ACL 2024.

- **Yinhong Liu**, Yixuan Su, Ehsan Shareghi and Nigel Collier. Unlocking Structure Measuring: Introducing PDD, an Automatic Metric for Positional Discourse Coherence. NAACL 2024, main conference.

- Han Zhou, Xingchen Wan, **Yinhong Liu**, Nigel Collier, Ivan Vulić, Anna Korhonen. Fairer Preferences Elicit Improved Human-Aligned Large Language Model Judgments. Under review.

- Zhongshen Zeng, **Yinhong Liu**, Yingjia Wan, Jingyao Li, et al. MR-BEN: A Comprehensive Meta-Reasoning Benchmark for Large Language Models. Under review.

- **Yinhong Liu**, Yixuan Su, Ehsan Shareghi and Nigel Collier. Plug-and-play Recipe Generation with Content Planning. EMNLP 2022, Generation, Evaluation & Metrics Workshop.

- **Yinhong Liu** and Guy Emerson. Learning Functional Distributional Semantics with Visual Data. ACL 2022, main conference.

- **Yinhong Liu** and Simon Godsill. Multi-channel audio statistical restoration. IEEE-ICSPCC, 2020.

PRIZE & EXTRA-CURRICULUM ACTIVITIES

Best Reviewer Award, EMNLP 2023, Natural Language Generation track.

PhD scholarship: Cambridge International Scholarship by Cambridge Trust.

Academic scholarship: Caldwell Scholarship by Corpus Christi College, University of Cambridge

Cambridge Chinese Drama Society: Protagonist in a university play, *Our Assassin*; Producer of two university plays *The Robin Murders* and *The Village*.

LANGUAGE & SKILLS

Languages: Chinese Mandarin (native), English (fluent), Chinese Cantonese (fluent).

Programming Languages:

- Inter-mediate Level: Python, C++, SQL, Matlab, Latex.
- Beginner Level: Javascripts, HTML, Objective-C.