

Xiang LIU

Email: xiang.liu@connect.hku.hk

Mobile: +852-97479004

EDUCATION

- **The University of Hong Kong** HK, CHINA
Master of Science(MSc) in Computer Science Sep 2022 - Aug 2023
- **George Mason University** VA, USA
Bachelor of Science in Computer Science; GPA: 3.71/4.0 Aug 2018 - Aug 2022
Honors/Awards: Dean's List (2018-2020)
Courses: Natural Language Processing, Visual Computing, Data Structures, Database Concepts
- **Henan University** Henan, CHINA
Bachelor of Engineering in Networking Engineering; GPA: 88/100; Rank: 4/292 Sep 2017 - Jun 2022

EXPERIENCE

- **HKUST Statistics and Machine Learning Research Group** HK, CHINA
Research Intern Dec 2022 - present
 - Contributed to the LMFlow Benchmark, a new benchmark which provides a cheap and easy-to-use evaluation framework that can help reflect different aspects of LLMs.
 - Contributed to the LMFlow, a framework that allows fine-tuning and deploying personalized LLMs with minimal cost and effort. And focused on fine-tuning Large Language Models, including LLaMa, Bloom and Vicuna.
 - Conducted research on the Chain-of-thought (COT) method to enhance Large Language Models' logical ability.
 - Supervised by Professor Tong Zhang.
- **Baidu Research Cognitive Computing Lab** Beijing, CHINA
Research Intern Dec 2021 - June 2022
 - Worked on dependency parsing using the Open Information Annotation (OIA) method to convert sentences into directed acyclic graphs (DAGs).
 - Enhanced the performance of the OIA method for Chinese sentences by integrating node type and edge type specific to Chinese OIA, achieving parity with English OIA.

COMPETITIONS

- **Kaggle — Feedback Prize - Predicting Effective Arguments Competition** Jun 2022 - Aug 2022
Team Leader — Silver Medal (Top 2%)
 - Developed baseline code, and designed various data preprocessing strategies, and model structures.
 - Implemented token classification instead of sequence classification, boosting rank and saving time on training and inference. Acquired proficiency in using Transformers training API.
- **Kaggle — U.S. Patent Phrase to Phrase Matching** Mar 2022 - Jun 2022
Solo — Bronze Medal (Top 9%)
 - Developed the entire code pipeline with various combinations of pre-trained language models and model structures.
 - Implemented multiple fine-tuning methods such as mean pooling, multi-dropout, and layer normalization.
- **CCF Collegiate Computer System & Programming Contest** Sep 2021 - Nov 2021
Team Member — Top 5%
 - Utilized data science and machine learning pipeline to predict consumer default vs non-default.
 - Employed LightGBM and XGBoost models and model stacking method. Mastered real-world data science challenges, machine learning pipeline, and team collaboration.

PROJECTS

- **Rumor Prediction Model for Microblog Epidemic** Mar 2021 - May 2021
Solo Project
 - Built a Covid-19 rumor prediction model based on the BERT pre-trained model, completing the deep learning model process.
 - Gained a profound understanding of the NLP project process by going through idea generation, data search, reference papers, code practice, model building, and article writing.

PROFESSIONAL SKILLS

- **Languages:** English, Mandarin
- **Computing Skills:** PyTorch, Git, Linux, NLTK