Jialiang Xu

Email: xjl@stanford.edu | Website: liamjxu.github.io | LinkedIn: www.linkedin.com/in/xjl

EDUCATION

Stanford University

Master of Science, Computer Science   *2023-2025, GPA 3.9 / 4.0*

University of Illinois at Urbana Champaign

Bachelor of Science, Computer Engineering, Highest Honor   *2018-2022, GPA 3.98 / 4.0*

Minor, Computer Science   *2019-2021, GPA 4.0 / 4.0*

PROFESSIONAL EXPERIENCE

Amazon   *Jun 2024 – Sep 2024*
Applied Scientist Intern   *Seattle, WA*

- Working on LLMs for User Modeling.

Amazon   *Jun 2023 – Sep 2023*
Applied Scientist Intern   *Seattle, WA*

- Worked on Large Language Models for Recommendation Systems. Developed a framework that uses a GNN-based recommendation model as a candidate item short-lister and a prompt-augmented, domain-adapted, and instruction-tuned LLaMA-2 13B model as a candidate re-ranker.
- The final model outperformed the baseline model that is used in the current product by a large margin on the metrics of recall@k and NDCG@k (k=5,20).

Bosch Research   *Mar 2023 – Jun 2023*
Research Intern   *Sunnyvale, CA*

- Adapted open-sourced Large Language Models (LLM) to domain-specific chat applications. The motivation behind this is to create open-sourced alternatives for popular close-sourced chat LLMs, therefore alleviating concerns for private data security observed in the common API-based approach.
- Finetuned a series of Flan-T5 models that outperform the vanilla version and popular API-based chat systems on corporate-owned data. Applied techniques of domain adaptation and prompt tuning to inject domain knowledge into LLM while avoiding catastrophic forgetting of the general language capabilities.

Microsoft Research   *Jul 2021 – Jul 2022*
Research Intern   *Beijing, China*

- Research: Focused on understanding semi-structured data (e.g., tables, forms, logs) with Natural Language Processing techniques. Produced 3 research papers submitted to ACL, KDD, and EMNLP. The PDF files, slides, posters, and code repo can be found at https://liamjxu.github.io/publications/.
- Production: Cooperated with product teams from Bing, Azure, and Excel on transferring research output into features for Microsoft products including Edge, Synapse Notebook, and Excel. Provided fundamental tools that allow 1) Bing to identify table fields for best visualization, 2) Azure to automate pivot table generation, and 3) Excel to intelligently assist users in generating analysis and visualizations for their spreadsheets.

Discovery Partners Institute   *Aug 2022 – Dec 2022*
Machine Learning Engineer   *Chicago, IL*

- Added Semantic Search functionalities to a multi-source biomedical searching platform 1-Search.
- Finetuned a Biomedical-domain-specific pre-trained language model on a Learning-To-Rank dataset and evaluated the model on both public datasets and an internal dataset collected from 1-Search. Implemented pipelines for two downstream functionalities. Improved the inference latency from over 3 seconds to less than 2 seconds by utilizing dynamic model quantization. Served the model on an Azure virtual machine with TorchServe.

COMMUNITY SERVICE

NeurIPS: reviewer (2024)

ARR: reviewer (Feb, Apr, Jun, Aug 2024).

EMNLP: reviewer (2022, 2023), conference volunteer (2022).

ACL: reviewer (2023).

PUBLICATIONS

"*" denotes equal contribution.

Peer-reviewed Conference and Journal Publications

[P8] Word Embeddings Are Steers for Language Models

Chi Han, **Jialiang Xu**, Manling Li, Yi Fung, Chenkai Sun, Nan Jiang, Tarek Abdelzaher, Heng Ji

ACL 2024, **Outstanding Paper Award**

[P7] SPAGHETTI: Open-Domain Question Answering from Heterogeneous Data Sources with Retrieval and Semantic Parsing

Heidi C. Zhang, Sina J. Semnani, Farhad Ghassemi, **Jialiang Xu**, Shicheng Liu, and Monica S. Lam

ACL 2024 findings

[P6] SUQL: Conversational Search over Structured and Unstructured Data with Large Language Models

Shicheng Liu, **Jialiang Xu**, Wesley Tjangnaka, Sina J Semnani, Chen Jie Yu, Monica S Lam

NAACL 2024 findings

[P5] Inferring Tabular Analysis Metadata by Infusing Distribution and Knowledge Information

Xinyi He, Mengyu Zhou, **Jialiang Xu**, Xiao Lv, Tianle Li, Yijia Shao, Shi Han, Zejian Yuan, Dongmei Zhang

ACL 2023 findings

[P4] Towards Robust Numerical Question Answering: Diagnosing Numerical Capabilities of NLP Systems

Jialiang Xu, Mengyu Zhou, Xinyi He, Shi Han, Dongmei Zhang

EMNLP 2022

Manuscripts and Pre-prints

[P3] Reverse Image Retrieval Cues Parametric Memory in Multimodal LLMs

Jialiang Xu*, Michael Moor*, and Jure Leskovec

Submitted to NeurIPS 2024

[P2] InfoPattern: Unveiling Information Propagation Patterns in Social Media

Chi Han*, **Jialiang Xu***, Manling Li*, Hanning Zhang*, Tarek Abdelzaher, Heng Ji

ArXiv, 2023

[P1] LUNA: Language Understanding with Number Augmentations on Transformers via Number Plugins and Pre-training

Hongwei Han*, **Jialiang Xu***, Mengyu Zhou, Yijia Shao, Shi Han, Dongmei Zhang

ArXiv, 2022

HONORS AND AWARDS

ACL 2024 Outstanding Paper Award	2024
Brown Institute Magic Grant	2024
Microsoft Stars of Tomorrow Award	2022
Horace and Kate Wu Scholarship	2022
Daniel W. and Carol A. Dobberpuhl Scholarship	2022
ECE Visionary Scholarship	2022
Yunni and Maxine Pao Memorial Scholarship	2021
ECE Alumni Association Scholarship	2021
First Place, DARPA SocialSim Final Evaluation	2021
First Place, UIUC EOH Original Undergraduate Research Award	2021
Omron Scholarship	2020
UIUC, Dean's list	2018
UIUC, Edmund J. James Scholarship	2018