

Jialiang Xu

Email xjl@stanford.edu | Homepage liamjxu.github.io | LinkedIn [linked.in/xjl](https://www.linkedin.com/in/xjl)
Google Scholar scholar.google.com/citations?user=S_mgVngAAAAJ

EDUCATION

Stanford University

Master of Science, Computer Science, AI Track

2023-2025, GPA 3.94 / 4.00

University of Illinois at Urbana Champaign

Bachelor of Science, Computer Engineering, Highest Honor

2018-2022, GPA 3.98 / 4.00

Minor, Computer Science

2019-2021, GPA 4.00 / 4.00

PROFESSIONAL EXPERIENCE

Stanford University

Dec 2023 – Present

Research Assistant

Stanford, CA

- Working with Professor Percy Liang at the Center for Research on Foundation Models (CRFM) on LLM Evaluation.
- Designed, implemented, and maintained key features of the Holistic Evaluation of Language Models (HELM) framework, adding support for 10+ cutting-edge datasets and 10+ state-of-the-art LLMs, including features for chain-of-thought prompting, LLM-as-a-judge, and encrypted frontend workflows.
- Curating a novel benchmark leveraging HELM to provide a generalist capability evaluation for the latest LLMs.

Amazon

Jun 2024 – Sep 2024

Applied Scientist Intern

Seattle, WA

- Led research project on User Response Prediction via LLM-based User Persona Modeling.
- Designed and implemented a framework with Python and Amazon Bedrock that mines scenario-specific user persona ontologies and user personas from interaction history, leveraging Claude-3-sonnet and causal inference. The pipeline uses the inferred user personas in Retrieval Augmented Generation for user response prediction.
- Evaluated model on a recent user response prediction benchmark, got results significantly outperforming the previous state of the art (PersonaDB) on Accuracy (12% better) and 1-Wasserstein Distance (11% better relatively).

Amazon

Jun 2023 – Sep 2023

Applied Scientist Intern

Seattle, WA

- Led research project on Large Language Models for Recommendation Systems.
- Created a framework with PyTorch and HuggingFace libraries including Transformers and Datasets 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.
- Evaluated model on internal product environment dataset, final model outperformed the product baseline model (LightGCN, BGCF) by a large margin (10% better) 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 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

- Led research on the topic of numerical reasoning and semi-structured data understanding. Produced one preprint and 2 papers published on ACL and EMNLP.
- Collaborated with product teams on transferring research output into product features in Bing, Azure Synapse Notebook, and Excel. Delivered tools that assist Bing in identifying table fields for web table visualization, Azure in automating pivot table generation, and Excel in helping users generate spreadsheet visualization and analysis.

PUBLICATIONS

Google Scholar: https://scholar.google.com/citations?user=S_mgVngAAAAJ

“*” denotes equal contribution.

Peer-reviewed Conference Publications

[P7] Do LLMs Know to Respect Copyright Notice?

Jialiang Xu, Shenglan Li, Zhaozhuo Xu, Denghui Zhang
EMNLP 2024

[P6] SPINACH: SPARQL-Based Information Navigation for Challenging Real-World Questions

Shicheng Liu, Sina Semnani, Harold Triedman, Jialiang Xu, Isaac Dan Zhao, Monica Lam
EMNLP 2024 findings

[P5] 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, 0.72%]

[P4] 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

[P3] 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

[P2] 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

[P1] 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

[M3] Reverse Image Retrieval Cues Parametric Memory in Multimodal LLMs

Jialiang Xu*, Michael Moor*, and Jure Leskovec
ArXiv, 2024

[M2] InfoPattern: Unveiling Information Propagation Patterns in Social Media

Chi Han*, Jialiang Xu*, Manling Li*, Hanning Zhang*, Tarek Abdelzaher, Heng Ji
ArXiv, 2023

[M1] 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

Patents

[PT2] Detection device and detection method for tripping mechanism of molded case circuit breaker

Guichu Wu, Hongwei Pan, Jialiang Xu, Bumeng Liang
No. CN107870299A

[PT1] An automated synthesis test system and method for realizing rapid formulation optimization

Jing Jiang, Hongzhe Ai, Yifan Long, Xubin Zheng, Zhaojian Tang, Jiani He, Jialiang Xu
No. CN116046764A

PROFESSIONAL SERVICE

NeurIPS: Reviewer (2024, 2025).
ICLR: Reviewer (2025).
ICML: Reviewer (2025)
ACL: Reviewer (2023).
EMNLP: Reviewer (2022, 2023), Conference Volunteer (2022).
NAACL: Reviewer (2024).
ARR: Reviewer (Feb 2024, Apr 2024, Jun 2024, Aug 2024, Oct 2024, Dec 2024, Feb 2025).

HONORS AND AWARDS

ACL 2024 Outstanding Paper Award	2024
Brown Institute 2024 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
Edmund J. James Scholarship	2018

MEDIA COVERAGE

SPINACH paper featured in the *Wikipedia Signpost*'s Recent Research section, November 18, 2024.
https://en.wikipedia.org/wiki/Wikipedia:Wikipedia_Signpost/2024-11-18/Recent_research
Brown Institute Magic Grant featured on the Stanford School of Engineering website, 2024–2025 cohort.
<https://engineering.stanford.edu/news/brown-institute-announces-its-2024-2025-magic-grant-recipients>