Tianxu Jiang

 in

Email: tianxuj@umich.edu
Mobile: +1-734-583-9692

EDUCATION

Emory University Atlanta, U.S.
Research Intern *May 2024 - Present*
Advisor: *Prof. Liang Zhao*
Theme: *Multimodal Models, Explainable AI, Medical AI*

University of Michigan Ann Arbor, U.S.
M.S. Data Science *Aug. 2023 - Present*
GPA: *4.0 / 4.0*
Courses: *Large Language Models, Information Retrieval, Machine Learning*

Fudan University Shanghai, China
B.S. Statistics *Sep. 2019 - June 2023*
Courses: *Data Structures and Introduction to Algorithms, Data Mining, Natural Language Processing*
Honor Program: *AI + X Micro-major Honor Program*

PUBLICATIONS

[Preprint] Yifei Zhang*, **Tianxu Jiang*** (Equal Contribution), Bo Pan, Jingyu Wang, Guangji Bai, Liang Zhao. *MEGL: Multimodal Explanation-Guided Learning.* Submitted to the Conference on Computer Vision and Pattern Recognition (CVPR) 2025

[Preprint] Yifei Zhang, James Song, Siyi Gu, Bo Pan, Guangji Bai, **Tianxu Jiang**, Liang Zhao. *XAI Benchmark for Visual Explanation.* Submitted to the Conference on Computer Vision and Pattern Recognition (CVPR) 2025

PROJECTS

Multimodal Explanation-Guided Learning with Large Language Models
Supervisor: Prof. Liang Zhao, Department of Computer Science, Emory University *May. 2024 - Oct. 2024*
- **MEGL Framework**: Developed MEGL framework to integrate visual and textual explanations, enhancing classification accuracy and model interpretability
- **Saliency-Driven Textual Grounding**: Proposed the Saliency-Driven Textual Grounding (SDTG) method to align visual and textual explanations through multimodal learning
- **Consistency**: Introduced the Visual Explanation Distribution Consistency loss to address incomplete annotations, leveraging multimodal large language models (MLLMs) for robus explanation generation

Self-Learning and Teacher-Guided Paradigms in Language Model Alignment *Feb. 2024 - May 2024*
- **Paradigm Exploration**: Explored fine-tuning frameworks to enhance performance of language models through self-learning and teacher-guiding
- **Learning Architecture**: Constructed self-reward preferences pairs to establish a self-learning architecture and utilized assessment of a larger model (Gemini 1.5 Pro) to establish a teacher-guided paradigm
- **Boosted AlpacaEval Performance**: Fine-tuned a 2.7B model (Phi-2) with proposed paradigms and achieved improvement in head-to-head evaluation and AlpacaEval 2.0 (19% progress on the pre-trained model), surpassing Falcon 13B and approaching Alpaca 7B and Davinci001

LLaVA-Recipe: Visual Instruction Tuning Enhanced Food Recipe VQA *Feb. 2024 - May 2024*
- **Culinary Assistant**: Augmented the capabilities of Large Language and Vision Assistant (LLaVA) for generating detailed cooking recipe from visual inputs
- **End-to-end fine-tuning**: Constructed a million level multi-turn dialogue dataset based on Recipe1M+ and conducted end-to-end fine-tuning with visual encoder weights frozen
- **Enhanced Performance in GPT Evaluation**: Performed knowledge injection to GPT 3.5 for better evaluation at a lower cost and enhanced the performance of model by up to 27% compared to baseline model

INTERNSHIP EXPERIENCE

Shanghai Consumer Big Data Lab (Fudan University) Shanghai, China
Research Intern *Mar. 2022 - Jul. 2022*
- **Prediction Framework**: Designed and implemented an LSTM-based framework , which effectively forecasted retail sales trend and reduced MAPE by 15% compared to the previous model
- **Evaluation Methodology**: Utilized Difference in Differences(DID) methodology to assess and quantify the impact of the "May 5th Shopping Festival" on consumer spending patterns and sales trend in Shanghai

China Pacific Insurance Group Data and AI Research Center Shanghai, China
Algorithm Engineer Intern *Jun. 2021 - Aug. 2021*
- **HRNet-Based Project**: Developed a project based on High-Resolution Net (HRNet) for facial analysis
- **Facial Attributes Detection**: Implemented HRNet-based facial expression recognition (e.g. smiles) and facial attributes detection (e.g. dark circles) from key facial coordinates, achieving an accuracy of 87.4%
- **Facial position estimation**: Established HRNet-based **facial position estimation** (e.g. facial tilts and frontal-profile status) from key facial coordinates, achieving an accuracy of 92.3%

SKILLS

Languages: Python, C, SQL, R
Frameworks: AWS, Google Cloud, Multiprocessing, PyTorch, Transformers, Deepspeed, XGBoost, LightGBM