

JIARUI XU



+86 181 1516 8409



Email



LinkedIn



XavierSpicy

Summary

- Pursuing a Master's in Data Science (Research) at the University of Sydney, specializing in machine learning.
- Proficient in common machine learning and deep learning algorithms, particularly focused on the application of LLMs.
- Strong programming and problem-solving skills, quick to adopt new technologies and apply domain knowledge.
- Active in the open-source community, exploring technical details through code & contributing to open-source projects.
- Able to stay updated with cutting-edge research, with strong communication, teamwork, and critical thinking abilities.

Technical Skills

Programming Languages: Python, C

Data Processing: NumPy, Pandas, SciPy, NetworkX

Query Language: MySQL

OS/Version Control/Containerization: Linux, Git, Docker

RAG & Agent Development: Langchain

ML Frameworks: SciKit Learn, PyTorch, HuggingFace

Education

University of Sydney, Sydney

Feb 2023 – Nov 2024

Master of Data Science

WAM: 81

Specialization: Machine Learning

Units: *Computational Statistical Methods (HD), Machine Learning and Data Mining (HD), Advanced Machine Learning (DI)*

Nanjing Tech University, Nanjing

Sep 2017 – Jun 2021

Bachelor of Industrial Engineering

GPA: 3.77/4

Specialization: Logistics and Supply Chain Management

Internship

Giant Network Technology Co., Ltd., Shanghai

Aug 2024 – Present

AI Lab - Large Language Model Algorithm - Algorithm Intern

- **Intelligent NPCs:** Developed LLM-driven NPCs with RAG using BM25 and embedding models, achieving 93% recall.
- **Alignment:** Explored innovations in alignment, built a lightweight framework, and fine-tuned with DPO, KTO, etc..

The University of Sydney, Sydney

Jun 2024 – Sep 2024

Faculty of Engineering - School of Computer Science - Vacation Research Intern

- **Segmentation:** Extracted named entities as text prompts, and combined CLIP with MedSAM to generate masks.
- **Captioning:** Cropped local images using masks, and captioned both global & local areas for multi-granularity info.
- **Question Answering:** Merged global and local captions as context for LLMs, using CoT prompting for QA.

Funplus Interactive Network Technology Co., Ltd., Shanghai

Dec 2023 – Feb 2024

Tech Middleware - AI Tech - AIGC Algorithm Intern

- **Product Demos:** Tracked SOTA models, quickly built demos, and explored feasibility and limitations in deployment.
- **Deployment:** Deployed models like GLM, MiniCPM, Qwen, integrated knowledge bases, and accelerated inference.
- **Development:** Developed advanced components for retrieval, including rewriting, routing, re-ranking, and stepback.

Research

Evolution of the Android App Eco-System through the Lens of Transformers

Aug 2023 – Jun 2024

- Evaluated 7 common topic models and 2 variants on coherence, diversity & similarity across million app descriptions.
- Proposed the category-conditioned summary to assign topic words to each app category with TopMEDAL framework.
- Analyzed app market trends through qualitative & quantitative analysis along with visualization techniques and LLMs.

Other

- **Open Source:** [Chinese Llama3](#), [NumPy MLP](#), [tabtransformers](#), [NumPy NMF](#), [Handwritten Character Recognition](#);
- **Language:** Mandarin (native), College English Test (CET-6) (541), TOEFL 97/120 (Writing 27/30, Reading 28/30).