Lan (Ava) Yan

(412) 214-2245 | lany2@cs.cmu.edu | www.linkedin.com/in/lan-ava-yan-2baa51166

EDUCATION

Carnegie Mellon University   Pittsburgh, PA
Master of Science in Artificial Intelligence and Innovation   Aug 2023 – May 2025
GPA: 4.33/4.33
Selected Coursework: Introduction to Computer Systems, Introduction to Machine Learning, Web Application Development

Tsinghua University   Beijing, China
Bachelor of Architecture, Bachelor of Arts in Economics   Sep 2018 – June 2023
GPA: 3.85/4.00, **Rank**: 1/79
Selected Coursework: Advanced Python Programming, Java and Object-oriented Programming, Data Structure & Algorithms, Pattern Recognition, Search Engine, AI for Takeout Delivery, Mobile Big Data Mining, NLP Seminar, LLM Practicum

SKILLS

Programming Languages: Python, C/C++, Java, SQL, R, HTML/CSS, JavaScript
Machine Learning & Data Science: Large Language Models, Machine Learning Algorithms, Deep Learning (PyTorch), Natural Language Processing, Predictive Modeling, Data Analysis, Data Visualization, Pandas, Numpy, MySQL, Hadoop, Spark
Tools & Framework: Git, Linux, Docker, Visual Studio, Jupyter Notebook, AWS, Django, React.js, Vue.js, Power BI, Tableau, Sketch, Figma, Adobe Suite

WORK EXPERIENCE

ModelBest (10-Million-Dollar LLM Startup)   Beijing, China
Machine Learning Engineer Intern   May 2023 – Aug 2023
- Built an autonomous workflow that prompts ChatGPT to generate 200k+ instructions leveraging 16k+ real-world APIs.
- Integrated AutoGPT and LangChain frameworks to auto-prompt ChatGPT to select and invoke APIs for instructions.
- Fine-tuned LLaMA on ChatGPT-generated instruction response data, achieving tool-use ability comparable to ChatGPT.

THUNLP Lab (Tsinghua University Natural Language Processing Lab)   Beijing, China
Research Assistant   June 2022 – Jan 2023
- Collaboratively sourced and assessed interactive web search data for 5.5k long-form questions, and created a 4-stage pipeline.
- Fine-tuned Chinese LLMs for fact extraction, used data augmentation and hyperparameter tuning to improve model by 16%.
- Co-authored paper *WebCPM: Interactive Web Search for Chinese Long-form Question Answering*, accepted by ACL 2023.

Boston Consulting Group (Top-3 Consulting Firm)   Beijing, China
Part-time Assistant   June 2021 – Jan 2022
- Executed comprehensive competitor analysis to formulate a tailor-made strategic transformation plan for a retail enterprise.
- Conducted an in-depth survey of target demographics and developed marketing strategy for an island development initiative.
- Compiled and critically analyzed 200+ intelligent medical policies, independently delivered an exhaustive report for the client.

ACADEMIC PROJECTS

Racial Bias Mitigation in Crime Prediction (Mobile Big Data Mining Course Project)   Feb 2023 – July 2023
- Gathered demographic data of Chicago and New York, conducted EDA and correlation study of regions in the city.
- Proposed a GAN-based framework to de-bias crime prediction by filtering out racial information in regional embeddings.
- Tested the framework on 3 models: reduced AUC of race prediction to 0.5 from 1 without sacrificing crime prediction ability.

Takeout Delivery Time Prediction (AI for Takeout Delivery Course Project)   Feb 2023 – July 2023
- Conducted time series analysis on one week of real-world takeout delivery data and engineered 30+ features.
- Compared machine learning method (XGBoost) and deep learning method (MLP) for predicting the delivery time of order.
- Enhanced XGBoost via feature selection and optimized MLP with re-sampling and weighted loss, achieving RMSE of 270s.

Chinese Judiciary Search Engine (Search Engine Course Project)   Feb 2023 – July 2023
- Developed a 10k+ case Chinese judiciary search engine with multi-page UI, keyword search, similarity searches, etc.
- Constructed front-end user interface with Vue.js and Element UI, and back-end with Django, Whoosh, and SQLite3.
- Evaluated search engine performance: Avg. Accuracy 0.79, Precision@10 0.74, Reciprocal Rank 0.85, Success@10 1.