

YANQING LU

ylu62702@usc.edu | +1 (323) 397-0737 | yqlu1015.github.io

EDUCATION

University of Southern California
Master of Science in Computer Science

Jan 2024 – Dec 2025 (Expected)
Los Angeles, US

- GPA: 4.0/4.0
- Relevant Coursework: Machine Learning, Foundations of Artificial Intelligence

Southern University of Science and Technology
Bachelor of Science in Mathematics and Applied Mathematics

Aug 2019 – Jun 2023
Shenzhen, China

- GPA: 3.68/4.0 (Weighted Average Score: 89/100, top 30%)
- Awards: Excellent Freshman Scholarship, *Special Award*; Merit Student Scholarship, *3rd Class*; Excellent Graduation Thesis

PUBLICATIONS

Lu, Y., Zhang, M., & Tang, M. (2023). Caching for edge inference at scale: A mean field multi-agent reinforcement learning approach. *GLOBECOM 2023-2023 IEEE Global Communications Conference*, 332–337.

EXPERIENCES

WiDeS Group, University of Southern California
Research Assistant

Jan 2024 - Aug 2024
Los Angeles, US

- Contributed significantly to two key projects focusing on reinforcement learning and large language model, respectively; completed all the code from scratch, and conducted extensive simulation experiments to iterate the methods.
- Collaborated effectively with team members to achieve project goals and milestones; shared expertise in optimizing the workflow and helped with debugging.

Baixing.com
Software Development Engineer Intern

Sep 2023 – Dec 2023
Shanghai, China

- Developed a chatbot orchestration system by integrating a large language model (LLM) with a state machine for intent recognition, enabling the chatbot to follow predefined processes in conversations.
- Successfully resolved recurring LLM service outages through improvements to the API key distribution system.

PROJECTS

Context-aware Handover with Large Language Model

May 2024 – Aug 2024

- Developed a pipeline for sensing real-time network context data (signal strength, battery information, etc.), utilizing a locally deployed large language model (LLM) to make handover decisions, and executing the handovers.
- Fine-tuned Llama 3 and Phi-3 models with **QLora**, achieving **90%** accuracy in making optimal handover decision; implemented various **quantization** methods (GGML, GPTQ), and deployed the quantized fine-tuned model on a laptop.

Convolutional Neural Network for Bird Calls Identification

Mar 2024 - May 2024

- Trained a ResNet-18 for classifying bird species based on soundscapes, achieving an accuracy of **85%** across 264 species.
- Explored the relation between model structure and dataset features by comparing CNN, k -NN, and RNN for classification across varying numbers of classes.

Automatic Base Station Deployment via Deep Reinforcement Learning

Feb 2024 – May 2024

- Designed and implemented a Proximal Policy Optimization (PPO)-based algorithm to determine the locations of multiple base stations, aiming to maximize the overall signal coverage.
- Achieved over **95%** of the performance of the optimal base station locations, in terms of the channel capacity with less than **10%** of the time cost compared to exhaustive search.

Mean Field Multi-Agent Reinforcement Learning Based Edge Caching

Jul 2022 – May 2023

- Developed a mean field multi-agent reinforcement learning framework and implemented a mean field Actor-Critic (MFAC) approach to optimize neural network model caching in a large-scale edge intelligence system.
- Conducted extensive simulations to investigate the performance of MFAC, demonstrating the superiority of cooperative agents over competitive agents.

SKILLS

Programming Languages: Python, Java, C++, Swift, SQL, Matlab

Techniques: PyTorch, TensorFlow, Hugging Face, Git, Slurm, Redis, Docker, Spring Boot, \LaTeX

Languages: English(proficient), Mandarin(native), Japanese(basic)