Shenyi Xie

shenyix@andrew.cmu.edu | (412) 889-4905 | GitHub | LinkedIn

EDUCATION

Carnegie Mellon University **Pittsburgh, PA**

Bachelor of Science in Statistics and Machine Learning, additional major in Computer Science May 2024

- Relevant Coursework: Data Structures (C); Software Design, Construction and Concurrency (Java); Functional Programming (SML); Machine Learning (Python); Computer Systems (C); Computer Vision (Python); Deep Learning(Python); Natural Language Processing (Python); Parallel and Sequential Data Structures and Algorithms (SML); Probability Theory; Statistical Inference; Designing Human-Centered Software; Distributed Systems (expected);

SKILLS

Programming Languages: Python, Java, JavaScript, C, C++, SML, SQL, R, HTML & CSS
Frameworks & Tools: RStudio, PyTorch, Pandas, NumPy, Figma, Axure, SpringBoot, Django, Bootstrap
Technical Skills: algorithm development and optimization, data analysis and visualization

WORK EXPERIENCE

CMU Human-Computer Interaction Institute Laboratory **Pittsburgh, PA**

Research Assistant Sep. 2023 - Dec. 2023

- Design, prototype, and implement an AI agent using **Python** to assist the elderly with memory problems in daily tasks such as medication reminders, appointment management, emergency response, etc.
- Craft interviews to research and analyze how an agent's affiliation impacts participants' acceptance and use of it

QTT Software Inc. **Shanghai, China**

Backend Developer Intern July 2023 - Aug. 2023

- Utilized **Java** with **MyBatis-Plus** and **SpringBoot** to design and implement an administrative system backend for *AI Universe*, a WeChat Mini Program, to efficiently perform CRUD operations on user accounts, orders, and model tags
- Implemented 10+ RESTful APIs with user authentication, interception of unauthorized access, and encryption for sensitive data, enabling seamless and secure communication between the backend and the frontend
- Conducted research to gather more than 100 Lora models and installed these on current checkpoint models for experimentation to select the ones with better performance to optimize the diversity and accuracy of the models that the AI universe uses

Kingsoft Office Software Inc. **Beijing, China**

Data Analyst Intern Jan. 2023 - June 2023

- Developed an automatic pipeline using **Python** for order system data retrieval, calculations, performance metrics validation, and upstream sales system upload, reducing the daily sales calculation latency from the original 30 minutes to under 1 minute
- Conducted comprehensive analysis through cleaning 700,000+ historical sales data and exploring aspects such as customer churn, geographic growth, and business performance seasonality to build 10+ BI dashboards for business use
- Optimized an order system using Axure to prototype redesigns and crafted PRDs with 20+ pages for scoping the changes.

PROJECTS

Social Networking Web Application - *Individual* May 2023 - June 2023

- Used Python with **Django, Bootstrap,** and **AJAX** to build a web app that supports posting, commenting, file uploading, and following/unfollowing users with complete modern authentication such as sign-up, log-in, password reset, secure cookies, etc.
- Deployed on Heroku with Amazon S3

Automatic Speech Recognition - *Individual* Sep. 2022 - Nov. 2022

- Built a convolutional, long short-term memory, fully connected **deep neural networks** to output a vector of probabilities corresponding to each phoneme
- Trained the model on 360 hours of 16kHz read English speech dataset, optimizing the model with CTC loss, and decoded the output probabilities into a sequence of phonemes using Beam Search, achieving a Levenshtein distance of 5.4 on the test dataset

Concurrent Caching Web Proxy - *Individual* Dec. 2021

- Implemented a web proxy in **C** that simultaneously accepts multiple incoming connections, read and parse requests, forward requests to web servers, read the servers' responses, and forward those responses to the corresponding clients
- Employed a least-recently-used (LRU) eviction policy and Pthreads readers-writers locks to ensure cache access is thread-safe