

Education

Carnegie Mellon University

Master of Science in Artificial Intelligence Engineering

Expected Dec. 2024

Pittsburgh, PA

University of Liverpool

Bachelor of Science in Computer Science (First Class Honor) | GPA: 3.96 / 4.0

Jul. 2023

Liverpool, UK

Skills

- **Programming Languages:** Python, Java, PHP, HTML/CSS, Svelte, R, SQL (PostgreSQL), C#
- **Frameworks:** Linux, Git, ssh, Docker, Dynamo DB, Django, PyTorch, OpenCV, Stable Baseline, Spark, pandera
- **Relevant Courses:** Software Engineering, Database Development, Data Structure and Algorithm, Computer System, Distributed System, Artificial Intelligence and Machine Learning, Data Mining, Computer Vision

Professional Experience

Data Engineer Intern | *Python, Stable Baseline, gymnasium, pandera*

Sept. 2022 – Aug. 2023

Department of Computer Science, University of Liverpool

Liverpool, UK

- **Trading System Design:** Designed a reward system based on returns and Sharpe ratio, and a daily trading environment for ten anonymized financial time series, involving market and limit order execution.
- **Custom Gym Development:** Developed "mbt-gym", a gym environment tailored for training high-frequency trading reinforcement learning agents, including performance evaluation on the combination of different models.
- **Model Training:** Trained deep reinforcement learning agents (A2C, PPO, DDPG) using an ensemble approach.
- **Achievement:** Achieved a **30%** accumulated return and a **1.78** Sharpe Ratio over a two-year testing dataset.

Data Engineer Intern | *Python, PostgreSQL, Dynamo DB, Faiss, Spark*

Jun. 2022 – Aug. 2022

Shanghai Action Information Technology Co., Ltd.

Shanghai, China

- **Data Retrieval Optimization:** Created a multi-stage product quantization algorithm to reduce vector data retrieval latency by **20%**, leading to a speed enhancement in the personalized recommendation system. The algorithm was successfully used in the company's financial customer service.
- **TensorDB Development:** Developed a search algorithm capable of using a single query statement to fetch both traditional attribute fields and vector data, resulting in a **20%** reduction in the number of search statements required compared to the conventional multi-level filtering approach based on heterogeneous database queries.
- **Testing:** Tested whether different Faiss structures could execute successfully on the system. Executed a performance analysis to determine the most effective indexing structures, resulting in a **30%** enhancement in search speed.

Data Engineer Intern | *Java, PostgreSQL*

Jun. 2020 – Aug. 2020

Industrial and Commercial Bank of China

Shanghai, China

- **Data Mining:** Implemented SQL query algorithms to identify valuable customers across multiple bank branches based on customer statements and credit history for various cards. This process reduced manual efforts by **80%**.
- **Data Processing Automation:** Employed a SQL-based data pipeline to process over one million original customer profiles in the past 20 years, resulting in a **30%** reduction in data ambiguity caused by variations in record standards.

Research Experience

Neuropricing in Immersive Environment | *C#, Unity*

Sept. 2021 – Jun. 2022

Individual Research, Supervisor: Georg Meyer

University of Liverpool

- **Experiment Setup:** Investigated human willingness to pay in virtual and physical environments. Developed VR sales environments in Unity, and conducted cross-over studies with 30 participants.
- **Findings:** Discovered no statistically significant difference between VR and physical environments, while identifying significant influences of product type and luxury vs. normal settings in both contexts (**p<0.001**).
- **Accomplishment:** Presented findings as a poster at *IMRF 2022*.

Machine Listening | *Python, Pytorch, Librosa*

Jun. 2021 – Sept. 2021

Research Assistant, Supervisor: Shengchen Li

Xi'an Jiaotong Liverpool University

- **Preprocessing:** Extracted acoustic features from DCASE 2021 into spectrograms to prepare for network training.
- **Model Development:** Designed a model that included self-attention and convolutional layers to generate audio embeddings and fine-tuned the ResNet 50 to generate visual embeddings. Deployed a fusion model that incorporated both audio and visual data for sound event detection.
- **Accomplishment:** Achieved **90%** accuracy in development dataset and **77%** in the baseline.