WUCHEN (AUBREY) LI

New York, NY | (669) 302-7032 | wl758@cornell.edu
GitHub/ LinkedIn/ Website

EDUCATION

Cornell University, New York, NY   Aug. 2024 – May 2026
Master of Science in Information Systems | GPA: 3.98
Relevant Coursework: Machine Learning Engineering, Machine Learning for Health, Computer Vision, Data Science
Honors/Awards: Cornell Tech Merit Scholarship

ShanghaiTech University, Shanghai, CN   Aug. 2018 - Jun. 2022
Bachelor of Engineering in Computer Science and Engineering | GPA: 3.50
Relevant Coursework: Software Engineering, Database System, Algorithm & Data Structure, Optimization & Machine Learning

EXPERIENCE

Florens Asset Management Company Limited, Data Analyst, Shanghai, CN   Sept. 2022 - Jul. 2024
- **Developed Scalable Data Analysis Platform:** Designed and maintained data mid-platform to support business decisions.
  - Build ETL pipelines with HIVE, Greenplum, and PostgreSQL; created **40+** Tableau BI and FineBI dashboards.
- **Automated Asset Selection System:** Engineered a high-performance system to streamline portfolio asset selection.
  - Developed a linear optimization engine using Python, SQL, and CPLEX, supporting selection from **3+ million** assets.
  - Reduced selection time from several days to **5 minutes**, automating **90%** of the selection workflow.
- **Built AI-Powered Automation for Logistics:** Developed CV and LLM solutions to improve operational efficiency.
  - Developed a MaskRCNN-based CV model with **95+** accuracy to detect floor damages of returned shipping containers.
  - Built an LLM workflow using ChatGPT/Llama and LangChain for automated customer order booking email processing, integrated to Outlook via customized extensions.

Intel, Software Engineer Intern, Shanghai, CN   Nov. 2021 - Feb. 2022
- **Contribute to Open-Source Recommender Systems:** Contributed to **DeepRec**, an open-source recommendation engine developed in collaboration with Alibaba.
  - Conducted performance evaluations and testing of BST, DIEN, and DSSM models to identify optimization opportunities. Utilized Kubernetes and Alibaba Cloud for scalable training and testing environments.
  - Implemented BF16 precision and self-attention modules to improve model AUC performance.
- **Optimized LSTM Model Inference Speed:** Improved the inference performance of Intel's PyTorch LSTM Operator.
  - Integrated Intel dgemm matrix multiplication library using C++/C, achieving a **3.5x** speedup in reference.
  - Used Docker for reproducible benchmarking and streamlined testing environment.

TECHNICAL SKILLS

- **Coding Language:**   Python, C++/C, C#, SQL, Shell Scripting, JavaScript, HTML, CSS
- **Tools & Frameworks:**   Git, GitHub, Docker, Linux, PyTorch, TensorFlow, Unity, Blender, OpenCV, ITK-SNAP, SLURM
- **Professional Tools:**   PostgreSQL, HIVE, Greenplum, CPLEX, Pandas, JSON, Figma, Tableau BI, Excel, CI/CD
- **Other Relevant Course:** HCI, Building Startup Systems, Unity Game Development, Artificial Intelligence, Cryptography

PROJECTS

Cornell: MiniTorch Machine Learning Framework Project (Python, Numba, Pytest)   Aug. 2024 - Dec. 2024
Course Project: Developed a PyTorch-like ML framework based on Python with auto-differentiation and GPU acceleration.
- Implemented broadcasting, backpropagation, and auto-differentiation for neural network training.
- Integrated GPU acceleration using Numba and operator fusion, achieving **100x** speedup in training and inference.
- Established a Python modular architecture with Pytest unit tests, ensuring reliability and maintainability.

MICCAI: Semi-Supervised Tooth Segmentation (Python, nnUNet, ITK-SNAP)   Jul. 2023 - Sept. 2023
Research Project: Developed 3D CBCT tooth segmentation models and ranked **FIRST** in the MICCAI 2023 CBCT challenge.
- Enhanced nnUNet with additional encoding layers for improved feature extraction and generalization.
- Designed a two-stage training strategy incorporating maxilla-mandible position prediction, data smoothing, and pseudo-labeling.
- Addressed dataset limitations by generating synthetic data to mitigate metal artifacts, improving model robustness.