Menglin Wang

Email: mw976@cornell.edu | Phone: 607-262-5553 | Location: New York, NY | LinkedIn: linkedin.com/in/melvinw

ABOUT

I am a recent graduate from Cornell University with a strong passion for software development. At Cornell, I work on developing a platform that helps researchers better leverage machine learning techniques for their research. I have finished core courses in software development, and I am skilled in distributed systems, cloud computing and NLP.

EDUCATION

Cornell University, College of Engineering   Aug 2021 – Aug 2023
• Master of Science (research-based), GPA: 3.8/4.0
• Concentrate on **Materials Engineering** and **Computer Science**

Anhui University, China   Aug 2016 – June 2020
• Bachelor of Engineering (Grade Rank: **1**/ 45), GPA: 87.5/100, **Merit Graduate**

Coursework: OOP & Data Structures (Java), Design and Programming for the Web, App Design and Prototype, Natural Language Processing, Machine Learning, Cloud Computing, System Programming, Advanced Database Systems

EXPERIENCE

• **Cornell University**   Oct 2021 – Aug 2023
Graduate Research Assistant (Software Engineering Track)
  o Designed, built, and managed a machine learning platform that provides dataset management, training, and testing.
  o Engineered a user-friendly web page using React.js, enhanced data visualization using Dash, and implemented efficient data pagination resulting in a **20 % reduction in page load time**.
  o Developed the backend features using Django, including APIs, datasets versioning, and workflow manager to automate the data transformation. Integrated PostgreSQL to manage datasets containing over **100K** data points.
  o Implemented ML models with gRPC APIs using Go and significantly improved the **feature generation speed by 65%** by the concurrent data processing workflow.
  o Communicate backend using Axios and **reduced the response latency by 50%** by Redis job queue and Memcached.

• **Westlake University, China**   May 2021 – Aug 2021
NLP Research Intern
  o Performed tokenization and masking by NLTK. Implemented a multi-head attention classifier to fine-tune a large language model (BERT) to enable automatic summarization. Achieved a Rouge-2 precision of 0.22.
  o Performed word embeddings with NLTK. Implemented a seq2seq model with a bidirectional LSTM encoder and a unidirectional LSTM with an attention decoder for semantic labeling. Achieved a low f1 score of 0.25.

PROJECT

• **Distributed Database, Distributed PostgreSQL managed on Cloud** *(C, AWS, PostgreSQL)*   *Instructor: Prof. Immanuel*
  o Distributed PostgreSQL on AWS EC2. Implemented distributed tables, query planning and execution via **UDFs** and **hooks**. Configured CloudWatch, Grafana, AWS Lambda, and AWS SDKs to enable **automatic node management**.
  o Benchmarked with pgbench in a **multi-tenant work pattern**. Achieved latency from 150ms to 50ms by automatic node addition and reduced scaling time from 10min to 1min compared to Azure for Postgres.

• **Distributed System, Sharded Key/Value Storage Service with Raft** *(Golang, Raft, RPC)*
  o Implemented **Raft** protocol, including log replication, leader election, and persistence state to ensure fault tolerance.
  o Constructed a sharded KV store service using Raft for replication and a shard controller to **manage the assignment of shards** to replica groups, enabling load balancing, scalability, and seamless reconfiguration of the system.

• **Cloud Computing, Intelligent City Safety Records Tracking Platform** *(Python, Azure, Flask)*   *Instructor: Prof. Ken Berman*
  o Constructed and developed a real-time analytic platform using **Python**/**Flask** based on **Microsoft Azure** and **MVC.**
  o Implemented data storage (over 1 million) by **CosmosDB/NoSQL.** Showed streaming record on an interactive map by **Leaflet.js** and trend by **Plotly**. Provided travel advice with statistical model, and notification by **Azure Function.**
  o Evaluated system by **Apache Bench** and achieved low latency of 10-20ms. Structured controller layer to handle asynchronous requests using **REST APIs**. Created and deployed containers on **Azure Kubernetes.**

TECHNICAL SKILLS

• Python, Java, Golang, Vue, Flask, Django, RESTful/RPC, Docker, AWS, Azure, Postgres