TONG GUAN

541-914-3580 | yuanyuanbingzi@gmail.com | linkedin.com/in/tongg | https://yuanyuanbingzi.github.io/

EDUCATION

B.S. in Computer Science and B.A. in Comparative Literature
University of Oregon, Eugene, OR | Expected Graduation: December 2024
Dean's Lists, CAS Hands-on Scholarship
Relevant Courses: Data Structures and Algorithms, Database, Software Engineering, Operating System, Computer Architecture, Compiler Design, Artificial Intelligence, Machine Learning, GPU Programming

WORK EXPERIENCE

Amazon, Seattle, WA: Software Dev Engineer Intern **06/2024 - 09/2024**
GenAI Full-Stack Application: Serverless Architecture with AWS

- Built a **serverless backend** using **AWS CDK** with **REST APIs** managed by **AWS API Gateway** and two **Lambda functions** for data ingestion and query.
- Integrated **RAG pattern** with **Amazon S3**, **OpenSearch Serverless**, and **Bedrock**, utilizing the **Claude 3.5 Sonnet model** to improve data retrieval accuracy and reduce hallucination.
- Implemented a **robust pipeline** with **alpha** and **beta** stages, incorporating **unit tests** and **end-to-end integration tests** using **Hydra Test** to ensure system reliability.
- Developed a **user-friendly front end** with **React** and **TypeScript**, implementing **AWS Amplify** for user authentication and token management, enhancing user experience.
- Enhanced **application security** with **AWS Cognito** and **Federated Authentication**, restricting access to internal APIs and safeguarding sensitive data.
- **Achieved a 40% increase in information access efficiency** and **improved understanding of internal tools by 30%**, driving significant user productivity gains.

University of Oregon Machine Learning Lab: Undergraduate Researcher **11/2023 - 03/2024**

- Identified **high-frequency features** in **target and non-target categories** within datasets, contributing to the enhancement of **data pattern recognition** and **classification accuracy**.
- Applied advanced **data manipulation** techniques using **Large Language Models** to selectively **rephrase** datasets, generating **adversarial (poisonous) data** for experimental analysis.
- Collaborated with the research team to **optimize data workflows**, leveraging **Python scripts** to **automate** the generation and evaluation of modified data sets.

IpserLab, Mountain View, CA: Full Stack Engineer Intern **01/2023 – 06/2023**

- Designed an **interactive web page** using **React/Redux** and **Ant Design**
- Built and integrated backend services with **Java** and **Spring Boot**, supporting critical features like data upload, deletion, advanced search, and order management.
- Implemented **secure data handling** with **MySQL** and **OAuth** protocols, ensuring compliance with authentication and authorization standards.
- Optimized **application scalability and performance** by deploying server-side solutions to **Docker** and **Kubernetes** on **AWS EKS**, reducing downtime and enhancing deployment efficiency.

TECHNICAL SKILL

Programming Languages: Java, Python, C/C++, Go, JavaScript, TypeScript, SQL, CUDA
Databases and Tools: MySQL, PostgreSQL, MongoDB, Unix/Linux, Git, Docker, Kubernetes
Web Development: React/Redux, Node.js, Spring MVC/Boot, Hibernate, Flask, Django, HTML/CSS
Cloud and DevOps: AWS (S3, Lambda, API Gateway, IAM, Cognito, EC2), CDK, CI/CD, Jenkins
GenAI Applications: Amazon Bedrock, Kendra, OpenAI, Claude models, fine-tuning and deployment