JIAHANG LI

jiahangli0311@gmail.com | (857) 327-2261 | linkedin.com/in/jiahang-li-218811217/

EDUCATION

Boston University — **Boston, MA**
Master in Computer Science — **Sep 2022 - Expected Jan 2024**

Fudan University — **Shanghai, China**
Bachelor in Computer Science and Technology — **Sep 2018 - Jun 2022**

- President of Student Council, College of Computer Science. Directed a school-wide programming exchange platform.

SKILLS

- **Language**: C/C++, Python, HTML/CSS, Java, Javascript, Typescript, Rust, SQL
- **Framework**: Node.js, Django, React.js, Spring Boot, Spring, Express.js,
- **Tools & Platforms**: Git, Postman, Docker, Github, GitLab, Linux, MongoDB, Android, AWS, RocksDB, Redis, etcd

PROFESSIONAL EXPERIENCE

Open Source Developer | Google Summer of Code | United States — **Jun – Aug 2023**

- Developed the enhancement of cBioPortal web page by adding support for the analysis and visualization of categorical and binary data in Generic Assay format, allowing users to perform statistic tests and to view the results in a clear format.
- Built new backend models, services and controllers to supply necessary data with **Spring** in **JAVA**, effectively facilitating data communication and processing. Designed 2 new **RESTful API** endpoints for frontend and backend connection.
- Implemented frontend features on the comparison page and results tab with **React**, significantly improving data presentation and user experience by applying multiple charts and tables, showing 10 thousands of data samples.
- Ensured software quality and reliability by designing and executing end-to-end tests for the newly integrated feature, leading to robust and fault-tolerant application performance. Released the updated product with **Docker**.

Algorithm Engineer Intern | NewsBreak | China — **Jan – Apr 2021**

- Constructed a full stack local news platform, supplemented functions using **Sanic**, such as administrator's assignment of tasks and multi-level urban planning drop-down list.
- Devised front-end using **React**, upgraded back-end in **Python**, database with **MongoDB**, and tested with **postman**.
- Led version control, fixed out of sync trouble between server and client with auto timestamp.
- Implemented a machine learning pipeline to extract and label data for the labelling team, tuning and testing web page classification model for 0.4M web page, increasing performance accuracy from 89% to 95%.
- Utilized **Git** for version control and developed detailed program design and documentation

PROJECTS

Full-Stack BLOG APP: BLOGGER | *Node.js, Spring Boot, React.js, MySQL*

- Designed a social media application to display activities and interact with other users, created the front-end in React.js and connected the backend with **Node.js**, and developed back-end in **Java and Spring Boot** for consistent real-time data.
- Identified opportunities to reduce server pressure and network latency by implementing a proxy with http-proxy-middleware for high concurrency environments, reducing HTTP calls by 8%, and database queries by 15%.
- Designed and developed database schemas to store user information and blog content in **MySQL**
- Performed security protocols and management of sensitive data and personal information using **dotenv**.

Stream Processing System with State Disaggregation | *Flink, AWS, Docker, RocksDB*

- Developed an independent control plane that divided tasks and states, enhancing the state migration process within **Flink**.
- Employed **Java** and **gRPC** to design a distributed, event-driven framework with TaskManager in charge of managing operators. Leveraged **RocksDB** for TaskManager state storage and utilized etcd for routing table storage, thereby ensuring system fault-tolerance.
- Adopted watermarks to represent the logical ingestion time, facilitating the handling of late-arrival events in window operators. Introduced a consistent hashing mechanism with virtual nodes to reduce the state migration expense.
- Engineered a scalable deployment on **AWS EC2** using **Docker** Compose, with features such as auto-scaling and load balancing, demonstrating no downtime and only a 30% increase in latency.

PUBLICATIONS

- Yu Hong, **Jiahang Li**, Jianchuan Feng, Chenghua Huang, Zhixu Li, Jianfeng Qu, Yanghua Xiao, Wei Wang "Competition or Cooperation? Exploring Unlabeled Data via Challenging Minimax Game for Semi-Supervised Relation Extraction", Proceedings of the 37th Association for the Advancement of Artificial Intelligence. 2023.
- Chao Wang, Jingping Liu, Tianyi Zhuang, **Jiahang Li,** Juntao Liu, Yanghua Xiao, Wei Wang and Rui Xie. "A Sequence-to-Sequence Model for Large-scale Chinese Abbreviation Database Construction", Proceedings of the 14th ACM International Conference on Web Search and Data Mining. 2022.