ZHEWEN LYU

(+1)346-542-0735 | xzavierlv@outlook.com | https://github.com/ZavierLyu

EDUCATION

Rice University December 2023
Master of Computer Science Houston, US
The Chinese University of Hong Kong, Shenzhen July 2021
Bachelor of Engineering in Computer Science and Engineering Shenzhen, CN

- GPA: 3.589 / 4.000, Major GPA: 3.633 / 4.000
- Rank: School of Data Science: 36 / 268, Computer Science and Engineering Major: 22 / 123
- Relevant Coursework: Distributed and Parallel Computing (A), Cloud Computing (A), Data Structures (A-)

PROFESSIONAL EXPERIENCE

Pinduoduo Inc. Jul. 2021 – Jun. 2022
Software Development Engineer Shanghai, CN

- Maintained a data transfer service system for Pinduoduo's **search engine** based on **Kafka**, **Hive**, **HDFS**, **Hbase** and **Flink** with an administration web application based on **SpringBoot** and **MySQL;** data transfer system is able to deliver streaming data **1TB/day** and load offline data **100TB/day**
- Implemented the load balance distribution of **Kafka** consumer and processor nodes in consistent hash algorithm, resulting the reduction of the variance of load by 25%.
- Devised a monitor system for monitoring memory usage, CPU usage of Kafka client nodes based on CAT (Central Application Tracking) and **InfluxDB**

ONES Technology Company Jun. 2020 – Aug. 2020
Software Engineering Internship Shenzhen, CN

- Built SDK migrating APIs in Golang to Java by AST tools in Golang
- Applied the text template render in Golang to convert the Iris web framework's router interface to the Java syntax's classes and functions
- Contributed to the project which establishes a plugins platform on which developers can build their own plugins

Shenzhen Research Institute of Big Data Sep. 2018 – May. 2019
Research Assistant Shenzhen, CN

- Analyzed students' Wi-Fi records data using **Pandas** and implemented an algorithm for calculating real-time course attendance to characterize students' behaviors
- Deployed a course attendance system website based on the **Flask** backend
- Presented in department's symposium and recognized as a crucial research achievement

PROJECT EXPERIENCE

Distributed and Parallel Programming Practices Sep. 2020 – Nov. 2020

- Programmed an odd-even transposition sorting algorithm with MPI (Message Passing Interface)
- Used Pthread (POSIX Thread) and MPI to conduct the computation of the Mandelbrot set in static and dynamic architecture, and built a GUI to plot the graph with X11 library
- Simulated an astronomical N-body system and a heating system using **Pthread**, **OpenMP**, and **MPI**
- Summarized the experiment result in different configurations and implementations for each assignment above, and analyzed Amdahl's law and Gustafson's law for parallel performance

Real-time Box Office Predictor Sep. 2019 – Dec. 2019

- Led a four-person team to conduct a data-driven project predicting box office based on the sentiment of tweets as calculated by NLP toolset TextBlob
- Applied the Random Forest model in **Scikit-learn** to predict box office returns and deployed the model on a website with AWS Elastic Beanstalk service
- Delegated machine learning model establishment, web crawler, data collection to group members and coordinated projcet schedule

Operating System Simulations Sep. 2019 – Dec. 2019

- Practiced process scheduling in Linux using fork and implemented a multi-thread game using POSIX Threads
- Operated GPU memory and CUDA API to simulate a memory management strategy in the least recently used (LRU) caching algorithm and implemented in-memory file systems with sequential and linked structures
- Simulated blocking and non-blocking operation in the IO system

HONORS & AWARDS

- Dean's List (Top 10% schoolwide) 2017-2020
- Provincial Third Prize, China Undergraduate Mathematical Contest in Modeling 2018

SKILLS & CERTIFICATIONS

- Programming Languages: Java, Python, C/C++, JavaScript, R, SQL
- Toolsets & Frameworks: SpringBoot, Kafka, Hadoop, Flink, Pandas, Numpy, OpenCV, LaTex, Scikit-learn, Flask