Ziheng (Jack) Chen

zihengchen2000@gmail.com | zihengjackchen.com | github.com/zihengjackchen | US Permanent Resident

Education

University of Illinois Urbana-Champaign (UIUC) — Aug 2023 – Expected May 2025
Master of Science in Computer Engineering — GPA: 4.00/4.00
University of Illinois Urbana-Champaign (UIUC) — Aug 2019 – May 2023
Bachelor of Science in Computer Engineering (Highest Honor) — GPA: 3.99/4.00

Leadership: ECE Graduate Student Advancement Committee Member, CS/ECE 374: Introduction to Algorithms
Achievements: Ackmann Family Scholarship, I-Promise Scholarship, Top 5% in LeetCode Contests

Skills

Programming Languages: Python, C++, C, Go, SQL, JavaScript, TypeScript, Bash, Assembly, CUDA, Scala
Web Development: React.js, Node.js, MySQL, MongoDB, Mongoose, Neo4j, Google Analytics, Firebase, Axios, Streamlit
Big Data and Cloud: Databricks, Splunk, Envoy Proxy, Airflow, Kafka, Spark, Redis, Flink, Protocol Buffers, gRPC
Technologies and Tools: Git, Linux, Docker, Kubernetes, Rancher, Azure DevOps, pandas, PyTorch, pytest, CARLA, TCP/IP

Experience

Machine Learning Engineer (co-op), StoneX Group – Chicago, IL — Jan 2023 – June 2023
- Created a proof of concept pipeline to benchmark prototypical commodity indices, enabling their potential product release
- Analyzed the profitability of commodity indices corresponding to 200+ expert-provided configurations based on historical prices, surpassing performance targets by 23.3% in 10-year total return with the highest-performing index
- Accelerated benchmarking by 3000% from code optimization in Python using pandas and parallel execution in Databricks
- Collaborated closely with marketing associates to refine evaluation criteria by simplifying technological concepts
- Devised a procedure to version track Databricks workflows in the repository and developed a Python script for CI/CD in Azure DevOps to automate workflow migration from production environments using APIs, eliminating manual processes
- Verified migration of Delta Live Tables and Spark SQL commands in complex scenarios through unit and integration testing

Data Engineer Intern, StoneX Group – Chicago, IL — Aug 2022 – Dec 2022
- Enhanced and optimized a Streamlit web application through security integration and ETL pipeline improvements
- Researched for solutions to integrate Okta authentication into the web application, deploying Envoy Proxy microservice, achieving user access control using bearer tokens and MS SQL Server, enhancing data confidentiality
- Developed a dynamic usage analysis dashboard in Splunk, delivering associate usage and category-specific insights
- Migrated the data curation ETL pipeline from Apache Airflow to a continuous Databricks workflow, efficiently managing staging data with Azure Blob Storage, and updated the application using Docker to reduce load times by 30x to 5 seconds
- Managed microservices using Rancher for Kubernetes management and documented the overall system design for clarity
- Utilized Agile methodologies to ensure project flexibility and timely completion, adapting to evolving project needs.

Data Engineer Intern, Ecolab – Saint Paul, MN — May 2022 – July 2022
- Profiled tables in Snowflake using SQL to examine key statistics, showing outliers and trends, enhancing data integrity
- Identified 201M invalid rows and created an interactive Snowsight dashboard to support business decisions
- Automated query generation with JinjaSQL in Python, utilizing Alation catalog, accelerating data quality evaluation
- Analyzed 19300 hours of Service Requests logs of dishmachines, cleansed using Python from Snowflake, processed using Azure Cognitive Service API to identify 6 common issues and their locations, informing potential refresh strategies

Projects

Vocabulary Builder Application – Web Programming, Full-Stack, Development — Jan 2024 – Apr 2024
- Developed a full-stack vocabulary builder app using React.js for the front-end and Node.js with MySQL for the back-end
- Designed a data schema using a UML diagram and deployed a MySQL instance on GCP for scalable storage solutions
- Optimized database queries and implemented indexing strategies, achieving a 60% reduction in operational costs
- Added user authentication and supported social networking features to promote daily engagement and progress tracking

Workout Logging Application – Web Programming, Full-Stack Development — Oct 2022 – Dec 2022
- Created a responsive workout logging web application in React.js and TypeScript, utilizing Node.js, deployed on Heroku
- Supported user authentication via Firebase to facilitate progress tracking and deliver personalized recommendations
- Connected the JavaScript back-end to MongoDB using Mongoose and employed Axios for front-end integration