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, TA for CS/ECE 374: Intro to Algorithms (FA23, SP24, FA24)
Achievements: Ackmann Family Scholarship, I-Promise Scholarship, Olesen Award Nomination, Top 5% in Coding Contests

Skills

Programming Languages: Python, C++, C, Go, SQL, JavaScript, TypeScript, Bash, Assembly, CUDA, SystemVerilog
Technologies and Tools: Git, Linux, Docker, Octopus Deploy, Terraform, Azure DevOps, Kubernetes, pandas, PyTorch, CARLA
Web Development: React.js, Node.js, MySQL, PostgreSQL, MongoDB, Neo4j, Firebase, Axios, Google Analytics
Distributed Systems and Cloud: Databricks, Splunk, Datadog, GCP, Cloud Composer, Airflow, BigQuery, Blockchain

Experience

Site Reliability Engineer Intern, Talos Trading – New York, NY June 2024 – Aug 2024
- Configured a Cloud Composer instance on GCP using Terraform and implemented order reconciliation DAGs using BigQuery, enabling actionable business-level alerts, analyzed costs to avoid a potential $1.5M annual increase
- Optimized system performance and cost by developing pipelines to project hardware specifications from current metrics, meeting target usage, achieving a 50% performance boost and $100K monthly savings
- Mapped VM connections in GCP using TypeScript to visualize complex market data flows, integrated real-time Datadog metrics for each VM using Flask, enhancing system health visibility and accelerating initial troubleshooting by 30%
- Automated market data failover process and integrated post-deployment validation into the Octopus Deploy pipeline with Datadog alerting, increasing efficiency, enhancing system reliability, and security by minimizing direct access
- Integrated commit message checks on pull requests with GitHub Actions, enabling traceability from Jira tickets
- Added automatic YAML/JSON validators using submodules to ensure code quality, leading to faster code reviews
- Analyzed logs in Linux to debug trading platform issues using Postgres database, demonstrating cross-functional expertise

Machine Learning Engineer (co-op), StoneX Group – Chicago, IL Jan 2023 – June 2023
- Created a pipeline to benchmark commodity futures indices, backtested profitability across 200+ configurations, and exceeded targets by 23.3% in 10-year return with the top-performing index, paving the way for potential product release
- Developed a Python script for CI/CD in Azure DevOps to automate the migration of workflows to production environments using REST APIs, verified correctness through extensive testing, eliminating manual processes
- Devised a procedure for version controlling Databricks workflows in the repository, enabling traceability and rollback

Data Engineer Intern (co-op), StoneX Group – Chicago, IL Aug 2022 – Dec 2022
- Researched solutions to integrate Okta authentication into an existing web application, deploying an Envoy Proxy microservice, and achieved user access control using bearer tokens and MS SQL Server, enhancing data confidentiality
- Developed a dynamic usage analysis dashboard in Splunk, providing insights into associate and category usage
- 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 improve load times by 30x to 5s

Data Engineer Intern, Ecolab – Saint Paul, MN May 2022 – July 2022
- Profiled tables in Snowflake using SQL to examine key statistics, identifying outliers and trends, enhancing data integrity
- Analyzed 19300 hours of Service Requests logs of dishmachines, cleansed using Python from Snowflake, processed using Azure Cognitive Service REST APIs to identify 6 common issues and their locations, informing potential refresh strategies

Projects

Traffic Risk Assessment and Mitigation – Autonomous Vehicles, Machine Learning, Safety Aug 2023 – Feb 2024
- Enhanced the resiliency of AVs in unfamiliar and accident-prone scenarios with a novel traffic risk assessment method
- Engineered a ResNet variant for computer vision tasks, reducing the footprint by 50% while achieving over 95% accuracy
- Constructed 6000 scenarios from NHTSA pre-crash typologies and trained lightweight Double DQN Reinforcement Learning Agents in PyTorch to preemptively brake using the traffic risk as an indicator, reducing accidents by 72.7%
- Created multi-threaded data generation and testing pipelines and boosted efficiency by 200% in Python