Ziheng (Jack) Chen

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

Education

University of Illinois Urbana-Champaign (UIUC) Champaign, IL
Master of Science in Computer Engineering | GPA: 4.0/4.0 *Aug 2023 – Expected May 2025*

University of Illinois Urbana-Champaign (UIUC) Champaign, IL
Bachelor of Science in Computer Engineering | GPA: 3.9/4.0 *Aug 2019 – May 2023*

Leadership: ECE Graduate Student Advancement Committee, TA for CS/ECE 374: Intro to Algorithms (FA23, SP24, FA24)

Skills

Programming Languages:	Python, C++, Bash, JavaScript, TypeScript, Go, SQL, CUDA
Cloud & Infrastructure:	GCP, Azure DevOps, Octopus Deploy, Terraform, Kubernetes, Docker
ML & Data:	Databricks, PyTorch, Langchain, pandas, Apache Airflow, BigQuery, Splunk, Datadog, CARLA
Web Development:	React.js, Node.js, Flask, Axios, Firebase, MySQL, PostgreSQL, MongoDB, GraphQL

Experience

Site Reliability Engineer Intern, Talos Trading – New York, NY June 2024 – Aug 2024
- Improved system stability using Python by provisioning upgraded hardware in GCP based on past usage metrics, boosting performance by 50% and reducing costs by 25%, saving $100K monthly
- Designed an interactive TypeScript and Node.js webpage to visualize GCP infrastructure and market data connections
- Integrated Datadog API with Flask, enhancing system visibility and reducing troubleshooting time by 2 hours daily
- Optimized order reconciliation workflows by configuring Cloud Composer in GCP using Terraform and integrating BigQuery, centralizing proactive business alerts and uncovering a potential $1.5M annual cost increase
- Automated market data failover and remote log retrieval with Bash scripts in Octopus Deploy, reducing downtime by 30%
- Integrated validation checks into the release pipeline with Datadog alerts, reducing deployment time by 2 hours
- Implemented YAML linting with GitHub Actions in submodules, reducing configuration errors by 30%
- Traced logs across platform components in Linux using Postgres, conducting root cause analysis on cross-functional issues

Software Engineer Intern (co-op), StoneX Group – Chicago, IL Jan 2023 – June 2023
- Built a pipeline using Python and pandas to benchmark commodity futures indices with 200+ configurations
- Identified the optimal configuration, surpassing 10-year return performance targets by 23.3%
- Devised a system for version-controlling Databricks data processing tasks, reducing deployment errors by 25%
- Developed a CI/CD pipeline in Azure DevOps to migrate Databricks notebooks and jobs to production using REST APIs, validated through edge-case testing of workload status, properties, and parameters

Data Engineer Intern (co-op), StoneX Group – Chicago, IL Aug 2022 – Dec 2022
- Integrated Okta authentication into a Streamlit webpage via an Envoy Proxy microservice, using bearer tokens for user-specific access control, eliminating unauthorized access
- Designed a Splunk dashboard to track real-time usage by user group, identifying popular content for prioritization
- Streamlined the data curation ETL pipeline by migrating from Apache Airflow to a scheduled Databricks workflow, leveraging Azure Blob Storage for staging data and reducing load times from 150 seconds to 5 seconds

Data Engineer Intern, Ecolab – Saint Paul, MN May 2022 – July 2022
- Profiled Snowflake tables using JinjaSQL to identify key statistics, outliers, and trends, boosting data integrity
- Processed 19,300 hours of dishwasher service logs from Snowflake using Azure Cognitive Services, identified six root causes, and guided improvements for future product revisions

Projects

GPU Failure Analysis in Delta Supercomputer – High-Performance Computing, FMEA Mar 2024 – Present
- Pre-processed 6B+ Delta Supercomputer log lines, removing noise through error deduplication and keyword matching
- Boosted processing speed by 100x using Hyperscan and Python multiprocessing for error extraction
- Analyzed NVIDIA GPU failure modes for AI workloads, uncovered error propagation paths that impacted 361K user jobs

Traffic Risk Assessment and Mitigation – Autonomous Vehicles, Machine Learning, Safety Aug 2023 – Feb 2024
- Developed a novel traffic risk assessment method to enhance AV resiliency in unfamiliar, accident-prone scenarios
- Simulated 6,000 scenarios from NHTSA pre-crash typologies and trained lightweight Double DQN Reinforcement Learning Agents in PyTorch to preemptively brake using traffic risk as an indicator, reducing accidents by 72.7%