Richard So

347-281-3815 | richardso2021@gmail.com | github.com/richardso21 | in/richardso21 | sorichard.com

Education

Georgia Institute of Technology 08/2024 – 05/2025
M.S. Computer Science, Interactive Intelligence *Atlanta, GA*
- Coursework: Machine Learning, Computer Vision, NLP, Data-Centric ML, Databases, Algorithms Honors

Georgia Institute of Technology 08/2021 – 05/2024
B.S. Computer Science — GPA: 4.0 *Atlanta, GA*

Work Experience

Amazon Web Services 05/2024 — 08/2024
Software Engineering Intern (ML)
- Member of an investigative research team that quantitatively analyzes user and developer experience across AWS.
- Reduced runtime of an internal data pipeline by >85% leveraging a parallelized fork-join model with AWS Lambda functions.
- Automated activity labeling of user session screenshots using Amazon Rekognition, Textract, and Anthropic's Claude LLMs.

Georgia Tech Financial Services Innovation Lab 05/2024 — Present
Research Assistant
- Explored benchmarking strategies and metrics to evaluate against state-of-the-art LLMs in financial/economic contexts.
- Devised robust document parsers with BeautifulSoup, RegEx, and spaCy to compile immense datasets for LLM fine-tuning.

Tanium 06/2023 — 08/2023
Software Engineering Intern
- Implemented CRUD logging into an internal PostgreSQL database and RESTful API interface to elevate console visibility.
- Rapidly tackled 50+ Jira tickets within a 10-week internship maintaining a Knex.js and React TypeScript codebase.
- Exercised test-driven development and data validation using Jest, Jasmine, and Joi to ensure UI and API reliability.

Georgia Tech College of Computing 01/2023 — 05/2024
Senior Teaching Assistant
- Headed biweekly lectures on computer architecture foundations, the C language, and memory allocation principles.
- Developed unit testing suites, docker images for auto-grading, and course software servicing 1000+ students per semester.

Union Pacific 05/2022 — 08/2022
Technology Intern
- Designed explainable ML regression models to estimate rail shipment prices for customers using XGBoost and SHAP.
- Performed rigorous feature engineering to achieve a 31% RMSE decrease versus UP's existing pricing analytics solution.

Projects

Generative Data Augmentation for Image Classification | *PyTorch, Stable Diffusion, ControlNet* (GitHub) 04/2024
- Experimented with multiple image generative models to enhance image classification accuracy when data is scarce.
- Observed a 10% F1 increase for Resnet-50 on a compact dataset when augmented with ControlNet-generated images.

LC3Tools | *C++, Vue, Electron, LC-3 Assembly* (GitHub) 10/2023 — 05/2024
- Lead maintainer of the educational tooling suite to code, assemble, and simulate assembly programs for the LC-3.
- Added 20+ major quality-of-life improvements through student and instructor feedback as a fork from the original project.

Alaskan Wildlife Image Segmentation | *Python, PyTorch, Pillow* (GitHub) 09/2021
- Utilized and refined the FgSegNet segmentation model to predict and automatically annotate animal presence in image data.
- 1st Award Winner of 2021 Terra NYC STEM Fair and Milton Fisher Scholarship for Innovation and Creativity.

Skills

Programming Languages | Python, TypeScript/JavaScript, C/C++, Go, Java, Lua
Frameworks & Libraries | React, Vue, Express, Electron, Flask, NumPy, Pandas, SkLearn, PyTorch
Databases & Misc. | PostgreSQL, SQLite, RocksDB, MongoDB, Firebase, Git, Docker, AWS