650-291-4624
hirotaso92602@gmail.com

So Hirota

GitHub: soh09
LinkedIn: So Hirota
Website (soh09.github.io)

Education

Stanford University — **September 2025 -**
M.S. Computer Science — *Stanford, CA*

University of California, San Diego — **September 2021 - June 2025 (expected)**
B.S. in Data Science, Minor in Cognitive Science (GPA: 3.95) — *La Jolla, CA*

Research Experience

Research Assistant — **September 2024 - Current**
LLM Reasoning Lab by Prof. Zhiting Hu — *La Jolla, CA*

- Proposed methods to extend the Process Reward Modeling paradigm for LLM fine-tuning by incorporating natural language feedback instead of scalar feedback, and reviewed relevant literature to explore the feasibility and novelty
- Implemented a best-of-n sampling strategy for a browser-enabled LLM agent and benchmarked performance on Visual WebArena.
- Created presentations to share findings from relevant literature for group-wide brainstorming sessions

Work Experience

Machine Learning Intern — **July 2025 - Present**
NASA-JPL (Jet Propulsion Lab) — *Pasadena, CA*

- Developing scalable quality assurance tools for the ARIEL space telescope mission's EXCALIBUR data pipeline, automating detection of anomalies in various scientific data products
- Generated synthetic residual datasets by perturbing orbital parameters, enabling supervised training of model
- Training a convolutional neural networks to regress deviation from true orbital parameters (in sigma units) and identify data issues that manual inspection would miss.

Data Science Intern — **June 2024 - September 2024**
Franklin Templeton, Strategic Ventures — *San Mateo, CA*

- Developed a financial data pipeline from the ground up using DataBricks, Snowflake, and PowerBI, resulting in an 80% reduction in manual data processing time and enabling informed decision-making by stakeholders
- Leveraged SQL, Python, and Llama models with tailored prompts to parse emails, slide decks, spreadsheets, and text files
- Held frequent meetings with the users to understand their needs and prioritized ease of use for non-technical users
- Created extensive documentation of the system for future users and engineers

Machine Learning Engineer — **October 2023 - Present**
Argon Robotics — *La Jolla, CA*

- Rapidly prototyped and delivered a CNN for weld seam detection within 2 weeks to be used in a production environment
- Advised the founder on neural network architectures like autoencoders and U-Net and identified the optimal solution, based on requirements like core objectives, performance, and data availability

Data Science Intern — **July 2023 - September 2023**
menu, Inc — *Tokyo, JP*

- Menu is a major food delivery technology company, mainly competing with "Uber Eats" in Japan
- Constructed a regression model to predict daily online time per delivery crew by factoring in weather conditions, seasonality, and historical data
- Led the development of a time and order flow simulation to simulate order statuses, delivery driver locations, and order queue, which facilitated the testing of novel order-driver matching algorithms

Projects

Dinosaur Nugget Anomaly Detector (1st Place at DataHacks) — **April 2024**

- Prototyped and trained a convolutional autoencoder in PyTorch to detect anomalous shapes in dinosaur-shaped nuggets
- Collaborated with team members to collect data, load/transform data, and pitch our project to a panel of judges
- Winning project of the UCSD-hosted annual "DataHacks" hackathon

2048AI - Evolving Neural Networks — **December 2023 - Current**

- Implemented neural networks and the Neur-Evolution of Augmenting Topologies (NEAT) algorithm from scratch
- Designing fitness functions that balance human intuition and objective metrics to facilitate effective evolution
- Profiled the NN and NEAT implementation to identify inefficiencies in code, resulting in a 20% speedup in performance

American Sign Language Detection Glove (1st Place at HardHack) — **April 2023**

- Developed a glove that could measure the position of each finger to detect the ASL alphabet in real-time
- Trained a K-Nearest Neighbor algorithm that could detect signs with 85% accuracy using Scikit-Learn