650-291-4624
shirota@ucsd.edu

So Hirota

GitHub: soh09
LinkedIn: So Hirota

Education

University of California, San Diego   **September 2021 - June 2025 (expected)**
BS, Data Science   *La Jolla, CA*
Coursework: Data Structures and Algorithms, Machine Learning, Data Visualization, Databases
GPA: 3.91

Work Experience

Computer Vision Engineer   **October 2023 - Present**
Neuronflo   *La Jolla, CA*
- Developing computer vision-based quality assurance tools for manufacturing plants to detect faulty parts
- Utilizing pattern detection methods like template matching, the OpenCV library, and models such as YOLO

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
- Conducted KPI analysis and communicated insights to decision makers and a multi-disciplinary team
- Constructed a regression model using geospatial data that predicted daily online time per delivery crew within an average 20 minute difference by factoring in weather conditions, seasonality, and historic data
- Lead the development of a time and order flow simulation algorithm to simulate order statues, delivery driver location, and order queue, which facilitated the testing of novel order-driver matching algorithms

Student Researcher, Cool Stars Lab   **April 2023 - Present**
UCSD Center for Astrophysics and Space Sciences   *La Jolla, CA*
- Exploring the Generative Adversarial Network framework to address low data availability in exoplanet science by generating synthetic data indistinguishable from real exoplanet brightness data
- Researching methods to automate exoplanet discovery using the Lightkurves Python library and machine learning methods

Projects

American Sign Language Detection Glove   **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
- Designed and developed Python scripts which streamlined the workflow of data generation, data ingestion, data processing, modeling and inference
- Winning project of the UCSD-hosted annual "HardHack" hackathon

Shazam 2.0   **January 2023 - May 2023**
- Created a music genre classifier that works on less than 10 seconds of a song, similar to Shazam
- Employed PyTorch for training and optimizing hyperparameters of CNN models, resulting in 99% accuracy
- Utilized the Million Song Dataset, a dataset of over 300gb of music data which included MFCC-like features
- Developed Python scripts to ingest the dataset from an AWS server to a local server, and process it into multiple CSV files to streamline training of machine learning models
- Utilized Fourier transforms and dimentionality reduction techniques like PCA to process raw audio data

FIRST Robotics Competition Data Viz Website   **April 2023**
- Developed and designed an interactive data visualization website for FIRST Robotics Compeitition data using Javascript, the D3.js library, HTML, and CSS
- Implemented search and highlighting functionality to facilitate geospatial data exploration by the user
- Interfaced with the FIRST API to pull competition data and utilized the Pandas Python library and D3.js to clean and aggregate data

Skills

Data Science	Pandas, SKLearn, PyTorch, OpenCV, Machine Learning, Databases, Dask
Programming	Python, Github, SQL, Java, Javascript, C++, Object-Oriented Programming
Soft Skills	Cross-functional Communication, Detail Oriented, Organization, Collaboration