

# Huy Ho

huyhdho@gmail.com | (858) 275-4189 | linkedin.com/in/huyhdho | huy-ho.github.io

## EDUCATION

### Georgia Institute of Technology

Atlanta, GA

*Master of Science: Computer Science*

Expected: May 2027

- **Relevant Coursework:** Robotics AI, Software Development Process, Information Security

### University of California, Berkeley

Berkeley, CA

*Bachelor of Arts: Data Science*

Dec 2022

- **Relevant Coursework:** Data Structures, Algorithms, Principles of Data Science, Database Systems, Artificial Intelligence, Machine Learning, Data Engineering, Data Inferences and Decisions, Probability, Web Design

## EXPERIENCE

### Cisco Systems

Atlanta, GA

*Software Engineer*

Jan 2023 - Present

- Designed and implemented ESG audit automation across 20 databases, 10 transformations, and 10 ETL pipelines with Snowflake, DBT, Dagster, and API ingestion, rolled out across Supply Chain IT, cutting manual effort by 80%
- Engineered core AI infrastructure for supply chain teams, building internal LLM platforms with agent-based RAG workflows over 10,000+ pages of unstructured documents using OpenAI API, LangGraph, Weaviate, and CortexLLM.
- Lead engineer for a Data Observability Platform at Cisco, managing deployment, integration, and platform operations supporting 10 Supply Chain IT teams and driving enterprise-wide data observability.
- Built scalable FastAPI APIs handling 500+ daily calls and optimized database queries with caching, reducing retrieval times by 44% and lowering API load.
- Deployed apps with Docker and AWS (ECR, ECS), cutting EC2 image size by 25% and improving efficiency.
- Migrated 50+ legacy Control-M jobs to DBT Cloud and orchestrated ETL pipelines using Dagster, increasing pipeline reliability by 40% and reducing query runtimes by 30%.
- Built hybrid rule-based and ML recommendation engine, reducing upgrade support tickets by 35%.
- Reduced energy consumption by 29% by applying LSTM and ARIMA models for forecasting and anomaly detection.
- Mentored three interns, providing technical guidance and supporting successful project delivery.

### DataGood at UC Berkeley

Berkeley, CA

*Project Lead*

Sep 2020 - Dec 2022

- Led a team of 5 students to deliver data-driven insights for clients, including UC Berkeley, CIEE, and CalSafe.
- Taught fundamental computer science skills such as Git, SQL, and relevant Python libraries.

### IDEXX Laboratories

San Diego, CA

*Data Science Intern*

May 2022 - Aug 2022

- Investigated a data leakage issue with the company's machine learning pipeline using Spark and SQL.
- Built a classifier with an accuracy of 88% to classify transaction items using feature engineering, natural language processing, H2O AI, and deployed it into the company's latest ontology version.
- Automated Avro and Parquet data wrangling, saving 1.5 hours per week previously spent on manual tasks.

### UC Berkeley Research Laboratory

Berkeley, CA

*Undergraduate Research Intern*

Sep 2021 - May 2022

- Conducted research in machine learning (DE-TR) in collaboration with the US Army Research Laboratory.

### Nozomi Networks

San Francisco, CA

*Software Engineer Intern*

May 2021 - Aug 2021

- Developed neural network models to forecast time-series data using Keras and TensorFlow; Presented to the CTO.

## SKILLS

**Languages:** Python, Java, C, R, SQL, HTML, CSS, JavaScript, TypeScript

**Libraries/Frameworks:** FastAPI, LangChain, LangGraph, Dagster, NumPy, Pandas, Tensorflow, Keras, React

**Others:** AWS, Docker, OpenAI, Kubernetes, Dbt, Spark, Hadoop, Weaviate, JUnit, Jenkins