

Ian Dong

ihdong@berkeley.edu (925)-596-8235 Personal Website: <https://ianhdong.github.io>

: <https://github.com/ianhdong> : www.linkedin.com/in/ianhdong

Education

University of California, Berkeley

Computer Science & Data Science B.A. 3.97 GPA

Berkeley, CA

Aug 2021 – May 2025

Relevant Coursework:

- Computer Programs, Data Structures, Discrete Math & Prob Theory, Linear Algebra, Foundations of Data Science
- Machine Structures, Computer Security, Operating Systems, Database Systems, Efficient Algorithms
- Computer Graphics, Intro to AI, Intro to ML, Prob & Random Processes, Optim Models, Principle of Data Science

Work Experience

Amazon

Machine Learning Engineer Intern

Seattle, Washington

May 2024 – Aug 2024

- Created ML productionalized pipeline to detect anomalous sign-ins via PySpark, AWS CDK, and SageMaker
- Utilized Random Cut Forest (RCF) and SHAP algorithms to alert team on anomalous time-series metrics
- Reduced oncall engineer's tickets from 24+ hours to ~3 hours of detection/debugging

Amazon

Software Development Engineer Intern

New York City, New York

May 2023 – Aug 2023

- Utilized AWS Glue, Andes, and Cradle to summarize Close-Match relationship data between RME jobs
- Created publishing library to extract S3 files into Glue tables and crawl partitions into Andes tables
- Saved internal team 12+ hours for statistical analysis and business assessments

App Orchid Inc.

Software Engineer Intern

San Ramon, CA

May 2022 – Oct 2022

- Created AI services to parse through documents and extract metadata to reduce execution costs by 25%.
- Tested contract extraction solutions leveraging context-based reasoning and natural language processing

Teaching Experience

UC Berkeley Electrical Engineering & Computer Sciences (EECS)

CS 168: Internet Protocols Undergraduate Teaching Assistant

Berkeley, CA

August 2024 – Present

- Hosted weekly discussions to teach various topics on internet protocols
- Created rubrics, graded assignments, and hosted office hours for course-related questions

College of Computing, Data Science, and Society (CDSS)

DATA 100: Principles of Data Science Course Staff

Berkeley, CA

August 2023 – May 2024

- Created rubrics, graded assignments, and hosted office hours for course-related questions

UC Berkeley Electrical Engineering & Computer Sciences (EECS)

CS 70: Discrete Math and Probability Course Staff

Berkeley, CA

June 2022 – May 2023

- Created rubrics, graded assignments, and hosted office hours for course-related questions

Projects

Climate Synapse

Jan 2023 - May 2023

- Deployed an NLP model utilizing word2vec and PPMI algorithms that links researchers and startups and investors

CV for Nutrition

Jan 2023 - May 2023

- Utilized Keras and InceptionV3 to build a CNN model to identify foods from photos and produce nutritional facts
- Worked with transfer learning to fine-tune the model to increase classification accuracy

WordleWise

Mar 2023

- Developed a Python model to optimize word guesses for Wordle using entropy and mutual information principles
- Based model on uncertainty to determine most informative guesses to maximize word-guessing efficiency

Skills

Technical Skills: Agile Software Engineering, Systems Design, OOP, Full Stack, Frontend, Backend

Statistics/Machine Learning: Statistical Analysis and Modelling, CNN, Deep Learning, NLP, Computer Vision

Languages: Java, Python, C++, C, R, SQL, HTML, CSS, Javascript

Libraries and Frameworks: Git, Jira, NumPy, Pandas, Matplotlib, Seaborn, Node.js, React, Express, MongoDB

AWS Cloud Computing: S3, Glue, EC2, Lambda, Athena, RedShift, CloudWatch, SageMaker

Awards

USACO Gold Division

Jan 2021

AMERICAN INVITATIONAL MATHEMATICS EXAMINATION (AIME) Qualifier

Mar 2019