

# Ian Dong

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

: <https://github.com/ianhdong> : [www.linkedin.com/in/ianhdong](https://www.linkedin.com/in/ianhdong)

## Education

### University of California, Berkeley

Berkeley, CA

Computer Science & Data Science B.A. 3.98 GPA

Aug 2021 – May 2025

#### Relevant Coursework:

- Computer Programs, Data Structures, Machine Structures, Discrete Math & Prob Theory, Linear Algebra, Intro Data Science
- Computer Security, Operating Systems, Compilers, Internet Networks, Database Systems, Efficient Algorithms
- Intro to CV, Neural Nets, Intro to Graphics, Intro to AI, Intro to ML, Random Processes, Optim Models, Advanced DS

## Work Experience

### OpenAI

San Francisco, California

Member of Technical Staff

July 2025 – Present

- Incoming Applied Engineer at OpenAI

### Amazon

Seattle, Washington

Machine Learning Engineer Intern

May 2024 – Aug 2024

- Developed an end-to-end machine learning pipeline using PySpark, AWS CDK, and SageMaker to detect anomalous sign-ins
- Integrated Random Cut Forest (RCF) and SHAP algorithms to enhance anomaly detection for time-series data
- Reduced on-call response time for engineers from over 24 hours to approximately 3 hours by automating detection

### Amazon

New York City, New York

Software Development Engineer Intern

May 2023 – Aug 2023

- Built data processing workflows using AWS Glue and Andes to summarize Close-Match relationship data across RME jobs.
- Created library to automatically extract S3 files, update Glue tables, and partition data for Andes, streamlining analysis
- Saved the internal team over 12 hours weekly by automating business assessments and statistical analysis

## Teaching Experience

### UC Berkeley Electrical Engineering & Computer Sciences (EECS)

Berkeley, CA

CS 168: Internet Protocols Undergraduate Teaching Assistant

Aug 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)

Berkeley, CA

DATA 100: Principles of Data Science Course Staff

Aug 2023 – May 2024

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

### UC Berkeley Electrical Engineering & Computer Sciences (EECS)

Berkeley, CA

CS 70: Discrete Math and Probability Theory Course Staff

June 2022 – May 2023

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

## Projects

### Atlas - Open World Generation

Mar 2024 - May 2024

- Developed an OpenGL procedural terrain generator with Perlin noise and instancing for optimized rendering
- Implemented dynamic shading, fog models, and textured terrain with skybox and environment mapping for realistic visuals

### 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, Object Oriented Programming, Full Stack (Frontend, Backend)

**Machine Learning:** Deep Learning, Computer Vision, Anomaly Detection, NLP, Statistical Analysis, Predictive Modeling

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

**Libraries/Frameworks:** PyTorch, TensorFlow, PySpark, Keras, Node.js, React, Express, MongoDB

**AWS Cloud Computing:** CDK, EC2, Lambda, S3, Glue, Athena, Redshift, SageMaker, CloudWatch

## Awards

USACO Gold Division

Jan 2021

AMERICAN INVITATIONAL MATHEMATICS EXAMINATION (AIME) Qualifier

Mar 2019