

# Ethan Zhou

xethanhzhou@gmail.com | 650-455-6366 | LinkedIn: [@ethannhzhouu](#) | ethannhzhouu.vercel.app | Los Angeles, CA

---

Data science professional bridging analytics and engineering to deliver impactful solutions. Skilled in Python, SQL, and JavaScript, with a strong background in CI/CD, ETL pipelines, dashboards, and workflow automation. Committed to leading projects and building systems that are reliable, scalable, user-friendly, and meaningful.

---

## SKILLS

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

**Frameworks & Libraries:** React.js, Next.js, Node.js, Tailwind, Flask, Django, PyTorch, TensorFlow, NumPy, scikit-learn, pandas

**Industry Tools:** Git, Tableau, PowerBI, AWS, PostgreSQL, Docker, Kubernetes, Apache Spark, Excel, Google Workspace

## WORK EXPERIENCE

### AI/ML Project Manager @ Apple, Data Operations

July 2024 - July 2025

INSPYR Solutions | Cupertino, CA

- Managed terabytes of audio, text, and video data in **S3** for ML evaluation and user accessibility improvements on Apple products.
- Spearheaded Name Detection data collection of **500,000+ assets** aimed to develop features for deaf and impaired users to facilitate attention and communication.
- Developed **Python** automation for file search and detection processes, **reducing manual tracker updates by ~80%**.
- Leveraged **pandas** to analyze metric data in **Excel**, generating reports to enhance project decisions and track status for clients.

### Data Collection Specialist @ Google, Augmented Reality

November 2023 - June 2024

Cognizant | Mountain View, CA

- Led **300+** data collection sessions for UX feedback and data, driving ML model improvements for Google AR glasses and headsets.
- Streamlined AR glasses collection workflows by troubleshooting operational issues, validating data, and updating documentation.
- Implemented metric automation in **Google Sheets** to record data and monitor key performance indicators (KPIs) for clients.

### Data Science & Machine Learning Team Lead

May 2023 - September 2023

AICamp | Palo Alto, CA

- Developed a robust scraping pipeline to collect **15,000+ contact data** using **LLMs, OpenAI API, and Google Search API**.
- Utilized **Langchain** and **GPT-3.5** to extract dynamic/static HTML contact data, **expanding customer outreach by ~37.5%**.
- Managed team of 5 interns to create a data science and machine learning curriculum to deliver to **100+ students**.
- Oversaw the development of 10+ language model projects made by students and achieved an average mentor rating of 9/10.

### Data Science & Machine Learning Intern

May 2022 - September 2022

AICamp | Palo Alto, CA

- Implemented a product review text generation model by training my own **GPT-2 model** with **Python** and **HuggingFace**.
- Deployed model onto a website built with **HTML, CSS, and Flask** for **100+ users to demo**.
- Developed a survey administration mobile app (similar to Google Forms) in **React Native** for students to demo.

## PROJECTS

### macOS Personal Portfolio | Next.js, React.js, TypeScript, Tailwind

August 2025 - September 2025

- Implemented clean and responsive UI components such as menu bar, control center, spotlight, and dock.
- Engineered interface with light/dark mode, brightness slider, and system lifecycle features (boot, sleep, shutdown).
- Created **13** different apps to showcase my skills, experiences, projects, and creativity.

### TigerGraph: Fraud Detection in Ethereum Networks | Python, G-SQL, ML

January 2023 - March 2023

- Implemented a machine learning model called Topology Adaptive Graph Convolutional Network (TA-GCN) with **Python** and **G-SQL**, with an average classification accuracy of **~82.2%**, to predict if an Ethereum wallet in the transaction graph is fraudulent or not.
- Documented the effectiveness of graph-based algorithms and traditional ML algorithms to detect fraud in Ethereum networks.

### Carbon Reduction with RunBuggy @ UCSD Innovation Sprints | Python, SQL, Tableau

February 2022 - May 2022

- Wrote SQL queries and Python scripts to process company logistics data, reducing truck usage and mileage for vehicle deliveries.
- Built a k-nearest neighbors model to classify CO<sub>2</sub> emissions with **~96% accuracy**, optimizing RunBuggy's transporter network.
- Presented project results to **30+ participants** on **showcase day** with dashboards built in **Python** and **Tableau**.

## EDUCATION

### University of California-San Diego (UCSD)

September 2019 - June 2023

B.S. in Data Science, Minor in Cognitive Science. GPA: 3.6

*Relevant Coursework:* Advanced Data Structures and Algorithms, Data Mining & Web Scraping, Data Visualization, Machine Learning Algorithms, Object Oriented Algorithms and Design, SQL/NoSQL Databases, Systems for Scalable Analytics, Web Development

Coursera - Google Advanced Data Analytics Professional

Udemy - Python Data Science and Machine Learning Bootcamp