James Zhao

jjzhao2@cs.cmu.edu • 415-259-8170 • in/jamesjzhao • jameszhao01.github.io

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
Masters in Machine Learning — December 2024
- Relevant Coursework: Machine Learning, Probability and Statistics, Advanced Natural Language Processing

University of California, San Diego — La Jolla, CA
Bachelors of Science in Computer Science, Minor in Math — June 2023
- Cumulative GPA: 3.985 / 4.0, Major GPA: 4.0 / 4.0
- Relevant Coursework: Deep Learning, Recommender Systems, Reinforcement Learning, NLP, Computer Vision, Computer Graphics, Image Processing, Probabilistic Reasoning, Computer Security, Networks, Databases
- Honors/Awards: Summa Cum Laude, Engineering Honors (Tau Beta Pi, Eta Kappa Nu)

SKILLS

Languages: Python, PostgreSQL, C++, R, Typescript, Java, Kotlin, C#/.NET
Libraries: PyTorch, Numpy, Scikit-Learn, SciPy, Matplotlib, Pandas, React, gRPC, Bootstrap
Tools: Kubernetes, Git, Latex, Bash, Powershell, Regex

EXPERIENCE

Doordash — San Francisco, CA
Software Engineer Intern — June 2023 - September 2023
- Architected identity verification deduplication system for DoorDash's internal Fraud Workstation Platform
- Serviced **240,000** dashers with **100 write QPS** and **50ms** latency, decreasing redundant handling time by **3 minutes**
- Constructed UI components in **React**, BFF endpoints in **GraphQL**, and Kafka event consumers & gRPC endpoints in **Kotlin** to dynamically persist user verification statuses in a **Postgres** database, backfilling with **200,000 identities**

The Cottrell Lab — La Jolla, CA
Research Assistant — January 2021 - June 2023
- Validated deep neural network architectures and methods on improving molecular fingerprint classification accuracy and F1-score from Nuclear Magnetic Resonance and Mass spectra data using **PyTorch Lightning**
- Improved previous CNN-based architectures through **transformers**, feature engineering, and positional encodings in terms of cosine similarity and informational-retrieval metrics, obtaining a **Recall@01 of 0.41**
- Prototyped architectures using multimodal (**CLIP**) and language modeling (**BART**) techniques to generatively predict molecular SMILES strings, dereplicate existing spectra, and perform vector lookups

Doordash — San Francisco, CA
Software Engineer Intern — June 2022 - September 2022
- Migrated internal Django monolithic entity tooling to microservice architecture, serving **200,000** daily requests
- Developed GraphQL BFF endpoints in **Typescript**, and built **React** front-end components to dynamically bootstrap and validate form schemas for flexibly searching, editing, creating, and deleting arbitrary database entities
- Expanded backend authorization for CRUD endpoints for **14,000** support agents and integrated Elasticsearch indexing

Scripps Institution of Oceanography — La Jolla, CA
Machine Learning Programmer — January 2022 - June 2022
- Experimented with image processing techniques and Vision Transformers to enhance multi-class plankton image classification, achieving an accuracy of **87.66%**
- Constructed and deployed web applications in **Plotly** and **Dash** to visualize experiment results, inspect mis-classified images and confusion matrices, and efficiently compare **100,000** individual images within plankton datasets

EXTRA-CURRICULARS

University of California, San Diego – Computer Science Tutor — September 2022 - June 2023
- Taught Deep Learning to two classes of 200 students by holding office hours and answering online forum questions
- Created exams and graded programming assignments related to Perceptron's, CNN's, RNN's, and Transformers
- Generated exploratory data analysis (**EDA**) tools for 70 student groups in class-wide Kaggle competition

Triton UAS — October 2019 - January 2021
- Implemented CNN-based image segmentation for shapes and characters through a U-Net, surpassing previous morphological methods and attaining an Intersection-Over-Union of **0.99** on generated images