

Yuxuan Zhang (Wayne)

http://www.yxzwayne.com

github.com/yxzwayne | linkedin.com/in/yxzwayne

yxzwayne@gmail.com | yuz226@ucsd.edu

EXPERIENCE

Student Trainee / Research Assistant

Aug. 2023 – Present

- Trained and tested VAE models on HIV protein sequences alignment data to learn epistasis that cause drug-resistance.
- Developed custom data analysis code for cryoDRGN experiments, enhancing team's capacity to interpret cryo-EM data and accelerating research outcomes.
- Work under the Lyumkis Lab at Salk Institute.

Teaching Assistant

Apr. – Aug. 2023

Hacıoğlu Data Science Institute, UC San Diego

San Diego, CA

- Head TA for DSC 30: Data Structures & Algorithms, helped design new programming assignments and unit tests in Java, led discussion sessions to clarify students' queries.

Software Engineer

Dec. 2021 – Jul. 2022

Awesung Inc.

Cranbury, NJ

- Worked on full-stacks with React.js and Django, shipped weekly new features to internal logistic management app. Optimized database queries, reducing API response times by **50%** for top five most-used queries.
- Revamped the relational database schema, achieving **more than 20% reduction** in average serialization time.
- Led the eBay Verification initiative for company warehouses; developed an ETL pipeline and integrated untapped operational data, resulting in a more than **15% reduction** in Item-Not-Received rate within 2 months.

Software Engineer Intern

Oct. 2020 – Dec. 2021

AILaw

Remote

- Shipped new features and unit tests in Vue.js for a legal workflow management app.
- Restored and optimized a defunct daily reporting script, reducing runtime by over 20% and enabling the recovery of a dozen clients within two months.

RESEARCH PROJECTS

Embedding based classification

Jan. 2024 – Present

- Ongoing project, researching novel ways to do multi-label classification in high dimensional text embedding space.

Medilora — Large language models, Fine-tuning

Oct. – Dec. 2023

- Led Medilora project, fine-tuned large language models with quantized low-rank adaptation.
- Set up model+data parallel training, inference and evaluation experiments in remote linux servers.
- Achieved comparable medical benchmark performance with 0.05% of the training data required by larger medical models, demonstrated effectiveness of small pretrained models.

Open-ended Document Classification

Apr. - Jun. 2023

- Adapted the recursive summarization method in 2109.10862 to generate class labels on various scientific papers.
- Conducted experiments using OpenAI ChatGPT API and local BERT models from HuggingFace Transformers.

Character-based LSTM RNN

Jan. – Mar. 2021

- Collected and cleaned text and conversation corpus focused on space missions in 5 languages; used PyTorch to train a LSTM recurrent neural network to make next-character predictions.

EDUCATION

University of California San Diego

La Jolla, CA

M.S. Data Science

Sep. 2022 - Present

- **Courses and Involvements** Attention/Diffusion Reading Group, RL/LLM seminar, Maths of Deep Learning, Learning Algorithms, Classical Statistical Modeling, Search and Optimization, Text Data Mining

University of Washington Seattle

Seattle, WA

B.S. Informatics

Sep. 2017 - Jun. 2021

TECHNICAL SKILLS

Programming: Python, Java, SQL (Postgres), JavaScript, R

Libraries: Keras, Jax, PyTorch, Pandas, NumPy, Matplotlib

Frameworks: React, Django, Node.js, Flask, JUnit