Yuxuan Zhang

(206-557-8944) | yuz226@ucsd.edu

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

RESEARCH PROJECTS

Epistasis Modeling   August 2023 – Present
- Developing, training and testing deep generative models on HIV protein sequence alingments using Keras and PyTorch, with a focus on variational autoencoder.
- Researching novel approaches to interpret epistasis in HIV virus, with ongoing efforts to bridge performance from physics-based models and mechanistic interpretability in transformer language models.

Medilora   October 2023 - December 2023
- Led Medilora project to fine-tune large language models with quantized low-rank adaptation.
- Achieved comparable medical benchmark performance with 0.05% of the training data required by larger models, demonstrated effectiveness of small pretrained models.
- Implemented quantization, inference, and benchmark code, providing insights into model performance scaling.

Open-ended Document Classification   April 2023 – June 2023
- Adapted the recursive summarization method in 2109.10862 to generate class labels on various scientific papers.
- Conducted comparative experiments of OpenAI ChatGPT API and local language models, demonstrated a marginally higher preference (10%) for ChatGPT approach over text encoding methods in both summarization and label discovery tasks.
- Employed k-nearest neighbors to model label output variance and stability.

WORK EXPERIENCE

Teaching Assistant   April 2023 – August 2023
Halıcıoğlu 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   December 2021 – July 2022
Awesung Inc.   *Cranbury, NJ*
- Developed the internal logistics management dashboard with React.js and Django. Optimized database queries, reducing API response times by 50% for top five most-used queries.
- 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.
- Revamped the relational database schema, achieving more than 20% reduction in average serialization time.

Software Engineer Intern   October 2020 – December 2021
AILaw   *Remote*
- Developed new features and unit tests in Vue.js for NEUTRINET, a legal SaaS app for law firms workflow.
- Restored and optimized a defunct daily reporting script, reducing runtime by over 20% and enabling the recovery of a dozen clients within two months.

EDUCATION

University of California San Diego   La Jolla, CA
M.S. Data Science   *Sep. 2022 - Present*

University of Washington Seattle   Seattle, WA
B.S. Informatics   *Sep. 2017 - Jun. 2021*

TECHNICAL SKILLS

Languages: Python, Java, SQL (Postgres), JavaScript, HTML/CSS, R
Libraries: Tensorflow, PyTorch, Jax (Equinox), Pandas, NumPy, Matplotlib
Frameworks: React, Node.js, Flask, JUnit