YICHEN ZHAO

(972)-375-7525 | alexyczhao@gmail.com | LinkedIn | Github | Personal Website

EDUCATION

University of Texas at Dallas Ph.D. in Computer Science *May 2026 (Exp.)*

University of San Francisco M. S. in Data Science *Jun 2023*

- GPA: 3.96, top 5% of the class.

University of Pennsylvania M. A. in Applied Mathematics *May 2022*

- GPA: 3.40.

Georgia Institute of Technology B. S. in Mathematics, Applied Mathematics *May 2020*

- GPA: 3.71; major GPA: 3.90; graduated summa cum laude.

WORK EXPERIENCE

Data Science Intern *Nov 2022 – Jun 2023*

Dagshub Inc., Remote

(Project) Next Word Prediction

- Boosted next word prediction performance by 57.6% utilizing MLOps tools to systematically track experiments, identify optimal model architectures, and promote reproducibility. Operationalized top performing LSTM model into production through CI/CD pipelines, delivering low-latency API responses.
- Coached Dagshub users to adopt MLOps best practices by authoring tutorial with 500+ views on incorporating model tracking, automation, and monitoring into machine learning workflows.

(Project) Snowflake x Dagshub

- Forged integration between Dagshub tracking and Snowflake databases enabling seamless tracking of SQL queries and table artifacts using Git and DVC. Authored a blog post read by 700+ Dagshub users detailing how to leverage Dagshub tracking features with cloud-hosted databases on Dagshub.

(Project) YOLO-formatted Annotation Transfer

- Streamlined rendering of image annotations in Dagshub repositories by building an automated pipeline to convert YOLO-formatted object detection into visual overlays, reducing processing time by 80% compared to existing tools.

SIDE PROJECTS

Neural Style Transfer on Videos and Images

- Built PyTorch solution harnessing VGG19 network and transfer learning to enable rapid style transition on images and videos**.** Slashed generation time of neural style transfer by 75% by parameter tuning.

Hand Gesture to Sign Language Alphabet Translator

- Built real-time sign language alphabet translator, achieving 92.52% prediction accuracy. Orchestrated computer vision pipeline with optimized random forest classifier, leveraging OpenCV and MediaPipe.

ShenShenPL – A New Programming Language

- Architected a new programming language, ShenShenPL, enabling reusable and scalable code authoring and optimizing for readability and adaptability.

SKILLS

Python-related: Statsmodels, Pandas, Plotly, OpenCV, PyTorch, NLTK, SpaCy, Flask, Streamlit, BeautifulSoup, Airflow, PySpark, MLflow, Docker, Kubernetes, Metaflow, Evidently
Non-Python: MySQL, PostgreSQL, MongoDB, JavaScript, HTML/CSS, C/C++, Git, DVC
ML-related: Computer Vision, NLP, Deep Learning, Data Mining, MLOps, A/B Testing, Time Series Analysis