YICHEN ZHAO

https://shenshenzyc.github.io/yichen-zhao-portfolio/ | alexyczhao@gmail.com | linkedin.com/in/yichenzhao1

EDUCATION

University of Texas at Dallas	Ph.D. in Computer Science	*Exp. May 2028*
University of San Francisco	M. S. in Data Science	*Jun 2023*
University of Pennsylvania	M. A. in Applied Mathematics and Computational Science	*May 2022*
Georgia Institute of Technology	B. S. in Mathematics, Applied Mathematics	*May 2020*

WORK EXPERIENCE

Data Science Intern, Dagshub Inc. – Developer Relations   *Nov 2022 – Jun 2023*

Next Word Prediction
- Build a neural network model implementing bidirectional LSTM for next word prediction. Use MLflow APIs to log experiments with parameters, metrics, and model artifact. (Dagshub repo)
- Write two blog posts, one introducing the theory behind RNN and LSTM, and the other explaining the workflow of building this next-word-prediction model with MLOps integration on Dagshub. (Blog post)

Snowflake x Dagshub
- Integrate Dagshub tracking with a Snowflake database, using Git to log SQL queries and DVC to log result tables. Optimize SQL queries to retrieve data for optimizing a Lasso regression model that predicts Lego set prices. (Dagshub repo)
- Write a blog post to explain to Dagshub users how they could take advantage of Dagshub tracking features when working with cloud-based database on Dagshub. (Blog post)

Annotation Transfer for YOLO-formatted Annotations
- Create a Colab notebook facilitating the transfer of YOLO-formatted annotations to enable rendering of annotations in Dagshub repositories for explicit display of images with annotations. (Colab notebook)

PROJECTS

Neural Style Transfer on Videos and Images
- Develop a program using PyTorch and Python OpenCV library to perform neural style transfer, transferring the artistic style of one image to another image or video by leveraging the hierarchical feature extraction capabilities of the pre-trained VGG19 model.
- Implement an iterative process to update outcome by minimizing a loss function that combines content loss and style loss, resulting in synthesized outcome with the desired artistic style. (Github repo)

Hand Gesture to Sign Language Alphabet Translator
- Develop an interactive program using Python OpenCV that captures live video from the webcam and translates users' hand gestures into sign language alphabet in real-time.
- Leverage Google's Mediapipe package to identify hand landmarks and fine-tune a random forest model to classify their coordinates as sign language letters, achieving an accuracy of 92.52%. (Github repo)

ShenShenPL – A New Programming Language
- Develop ShenShenPL, a new programming language focused on readability and flexibility in Python.
- Design an intuitive syntax that enable programmers to write maintainable code. Create an interactive shell for code execution, providing a user-friendly environment for prototyping and projects. (Github repo)

SKILLS

Python-related: Statsmodels, NumPy, Pandas, Plotly, Scikit-learn, OpenCV, PyTorch, Gensim, SpaCy, Flask, Streamlit, BeautifulSoup, Airflow, PySpark, MLflow, Docker, Kubernetes, Metaflow, Evidently
Non-Python: MySQL, PostgreSQL, MongoDB, JavaScript, HTML, CSS, Java, Git, DVC, C/C++, Unix
ML-related: A/B Testing, Time Series Analysis, Deep Learning, Reinforcement Learning, Optimization Theory