Yufan Zhang

yz605@duke.edu | (+86) 138 5318 9701 | github.com/iamyufan | linkedin.com/in/yufanbruce

Education

Duke Kunshan University (DKU) / Duke University Dual Degree Aug 2019 – May 2023
B.S. in Data Science (by DKU) Kunshan, China
B.S. in Interdisciplinary Studies (Subplan: Data Science, by Duke)
- Cumulative GPA: 3.76/4.0 | Major GPA: 3.91/4.0
- Dean's List: Spring 2021, Spring 2022 (with Distinction)
- Courses: Calculus, Linear Algebra, Probability, Algorithms and Databases, Principles of Machine Learning, Numerical Analysis

Professional Experience

Saint Laurent Consulting (on Amazon.com account) | *Business Intelligence Engineer Intern* Sep 2022 – Present
- Designed and built an end-to-end data solution covering front end (Flask), data transformation (RESTful APIs and Python based ETL), and backend database (PostgreSQL).
- Automated daily process to fetch daily COVID-19 data from CDC and calculate the risk exposure of all Amazon fulfillment centers.
- Integrated an order tracking web application with an existing inventory management system with Flask and MySQL.

SciEcon CIC (an NPO for blockchain research) | *Research Affiliate & Chair of Creativity and Innovation* Oct 2021 – Present
- Queried over 2 million transaction records of DeFi protocols with Google BigQuery, and performed statistical analysis (regression, distribution analysis) and social network analysis (core-periphery structure analysis, network feature analysis) on them.
- Built static websites leveraging BootStrap and MkDocs and deployed the websites on Inter Planetary File System (IPFS).
- Publication: Chen, Z., Liu, Y., **Zhang, Y.**, & Zhang, L. (2022). Blockchain Network Analysis: A Comparative Studies of Decentralized Banks. [*In Progress*].

COMPSCI 201 Teaching Team, DKU | *Teaching Assistant* Jan 2021 – Mar 2021
- Developed 5 Java programming exercises covering topics including object-oriented programming and data structures.
- Hold office hours to troubleshoot bugs in 30+ students' programs and lectured 120 min-lab sessions.

Research Experience

HCI Blockchain, Data Science Research Center at DKU Mar 2022 – Sep 2022
Supervisors: Prof. Xin Tong (Asst. Prof. of Computation and Design) and Prof. Luyao Zhang (Asst. Prof. of Economics)
- Performed SQL to query 16k+ transaction records and architected a database of CryptoPunks NFT using MySQL.
- Conducted data analysis (distribution analysis, sentiment analysis) with Python on 13+ million tweets on NFT's ethical issues.
- Created a dashboard leveraging Plotly Dash to visualize the price difference of CryptoPunks with different genders and skin tones.
- Publication: **Zhang, Y.**, Chen, Z., Zhang. L., & Tong, X. (2022). Visualizing Non-Fungible Token Ethics: A Case Study On CryptoPunks. *ChinaVis 2022*. doi:10.48550/arXiv.2206.12922

Interdisciplinary Data Analysis Project, DKU Jan 2022 – Jul 2022
Supervisor: Prof. Peng Sun (Asst. Prof. of Data Science)
- Built a end-to-end GAN-based model with PyTorch to achieve arbitrary font style transfer in a few-shot learning fashion, improving the generating images quality by 22% of SSIM and 19% of user evaluation scores compared to the state-of-the-art method.
- Designed a novel skip connection to improve the generating ability of letters from untrained languages by 25% of SSIM.
- Publication: **Zhang, Y.**, Man, J., & Sun, P. (2022). MF-Net: A Novel Few-Shot Stylized Multilingual Font Generation Method. *ACM Multimedia 2022*. doi:10.1145/3503161.3548414

AlphaFold Research Group, DKU Sep 2021 – Jun 2022
Supervisor: Prof. Huansheng Cao (Asst. Prof. of Environmental Science)
- Developed a Graph Convolution Network (GCN)-based model using PyTorch to predict catalytic turnover numbers in Bacillus subtilis with the intermediate outputs from AlphaFold, which achieved an increase in R^2 of 14% than the state-of-the-art model.
- Conducted dimensionality reduction of 1.7TB of output data from AlphaFold by implementing Principle Component Analysis (PCA).

Institute of Computing Technology, Chinese Academy of Sciences Jul 2021 – Aug 2021
Supervisor: Prof. Shuhui Wang (Assoc. Prof. with Institute of Computing Technology, Chinese Academy of Sciences)
- Enhanced CycleGAN's robustness of the precision of the generating images by 67% by developing a new category loss with PyTorch.
- Presented the work to the program supervisor with an oral presentation and a technical written report with LaTeX.

Skills

- **Programming Skills:** Python, SQL, Java, C++, MATLAB, Markdown, LaTeX
- **Data Science & Machine Learning:** MySQL, PostgreSQL, Hadoop, Spark, PyTorch, Scikit-Learn, XGBoost, Pandas, Numpy
- **Software & Web Development:** JavaScript, HTML, CSS, Flask, Django, ReactJS, BootStrap
- **Miscellaneous Technologies:** Data scraping, Data visualization (D3.js, Plotly), Linux, Git, Conda, Slurm, Docker, ArgoCD

Additional Information

Languages: English (TOEFL iBT 110), Chinese Mandarin (native)
Honors and Awards:
- Meritorious Winner (MCM/ICM Contest in Modeling 2021)