

YUANZHAN GAO

Primary Email: yg8ch@virginia.edu

Primary Phone: (+1) 540-618-0766

Personal Website: <https://yuanzhangao.github.io/>

Education

University of Virginia, Charlottesville, VA

August 2020 – December 2023

Computer Science & Economics (Financial Concentration) Double Major

- GPA: 3.93/4.0
- Relevant Courses: Advanced Software Development, Machine Learning, Econometrics, Empirical Finance, Optimization

Work Experiences

Economics Department - University of Virginia, Charlottesville, VA

Data Science Research Assistant

December 2022 – Present

- Designed and implemented 2 [algorithms](#) using *Geopandas*, *Shapely*, *PyModis* libraries to automatically retrieve and process 10GB + satellite data from NASA database to reversely compute geographical coordinate based on pixel values, which were then used collect various regional satellite data from sub-regions of Pakistan
- Designed and implemented an [algorithm](#) using TF-IDF vectorization and fast sparse matrix multiplication libraries to perform efficient fuzzy matching tasks and record linkages; used the algorithm to perform FZ tasks between two 10GB+ loan datasets and significantly reduced processing time by 95%+
- Performed feature engineering tasks such as data cleansing, wrangling, dummy encoding, and regularization on 20 GB+ of financial dataset **in total** using *Pandas*, and run restricted 2-staged multivariate regressions on selected variables in STATA to discover underlying patterns
- Produced 200+ customized scatter plots, binned scatter plots, heatmaps using *Matplotlib* and *Seaborn* for multiple published and working Economics papers

Engineering Department – University of Virginia, Charlottesville, VA

June 2023 – Present

Software Development Research Assistant

- Developed a fully functional dashboard web application from scratch using *Dash* Framework, *Plotly.js*, *React.js* and *Flask*, which enables users to interactively visualize the affected regions of past 23 hurricanes in the US and track their movements within the given timeframes
- Developed a preprocessing pipeline using *Pandas* and *Polars* for parsing and feature engineering the existing Hurricane Database into acceptable format for *Dash* backend
- Designed client-side callbacks in JavaScript and Python that allow users to adjust the date parameter to modify the choropleth graph for visualizing hurricane-affected regions within browser only
- Containerized the web app using *Docker* for hosting and modified the existing nginx config file to allow larger net traffic for graph to render

Taiping Asset Management Limited, Shanghai, China (Remote)

Information Technology Research and Development Intern

June 2022 - August 2022

- Reviewed and studied 20+ papers into the history of the applications of ML in the asset management industry and classical models used in practices; drafted a 15 pages of studying reports to the manager
- Performed feature engineering tasks such as regularization and denoising on historical portfolio performance and macroeconomics data and EDA (correlation matrices, multicollinearity) on the aforementioned dataset to look for potential patterns and select relevant parameters
- Trained and fine-tuned a 3 memory-cells LSTM network using 20GB+ of aforementioned dataset to provide forecasts for upcoming market behaviors and adjust investment portfolios using *Tensorflow* and *Pytorch*; achieved an 83.7% accuracy level on the validation data

Research and Business Plan Development Intern

- Conducted business development research and language education market research by summarizing and synthesizing 50+ online sources and 25+ virtual interviews with potential clients
- Offered recommendations to 2 board members and CEO on strategic business development and pricing models by analyzed for-profit and nonprofit business models and operational strategies in a 45-page report
- Communicated and collaborated with 8 interns in 4 cross-functional teams using project management tools such as a Gantt Charts, project charter, project closeout report, and Google Workspace
- Drafted a 20 pages + business plan, a value proposition canvas, and a lean business canvas in a team setting

Projects

Study Buddy Finder

Class Project of CS 3240: Advanced Software Development

- As part of 5-developers team, developed a web application using Django framework to assist undergraduate students looking for relevant study buddies taking similar classes; includes communication functionalities.

Portfolio Optimization with Risk and Return Attribution Analysis

Class Project of Econ 4360: Empirical Finance

- Analyzed the historical returns of 5 pre-selected financial assets and performed portfolio optimizations by adjusting weights of each asset using past financial data using MATLAB

Relevant Interests

- Interested in the theoretical basis of mathematics in machine learning; current reading and taking notes on [*Mathematics of Machine Learning*](#) by Marc Peter Deisenroth, A.Aldo Faisal and Cheng Soon Ong
- Interested in quantitative finance and related mathematics background; current reading and taking notes on [*Asset Pricing*](#) by John Cochrane

Skills

- Proficiency in Data Processing and Analysis with Pandas/Polars, STATA, MATLAB
- Proficiency in Machine Learning with Scikit-Learn, Tensorflow, and Pytorch
- Proficiency in Full Stack Development with Django, Dash, Flask; Front-end Development with React.js, Vue.js and generic html/CSS/JavaScript
- Proficiency in Structured Query Language with MySQL and PostgreSQL