

Yiwen (Chloe) Shen

(626)567-1566 | yiwen.shen@tufts.edu | <https://www.linkedin.com/in/yiwen-chloe-shen-0ab538220/>

Data Scientist with Experience in Financial Analysis

EDUCATION

Tufts University

M.S. in Computer Science, GPA 3.69

May 2023

Relevant Coursework: Machine Learning, Software Engineering, Bayesian Deep Learning, Algorithms, Database Systems, Big Data, Networks & Protocols, Mobile App Development iOS&Swift, Probability, Statistics, Customer Discovery and Solution Design

Nanjing Audit University China

B.S. in Mathematical Finance, GPA 3.78

June 2020

Relevant Coursework: Mathematical Statistics, Applied Stochastic Process, Time Series Analysis, Financial Calculations, Macroeconomics and Microeconomics, Accounting, Algebra Analytic Geometry

SKILLS & COMPETENCIES

- **Programming Language:** Python, R, MATLAB, Java, Swift, Stata, SPSS, Tableau
- **Tools and frameworks:** Excel, GCP, BigQuery, Hadoop, Spark, Pandas, NumPy, Matplotlib, TensorFlow, PyTorch, JAX, Seaborn
- **Databases:** SQL (PostgreSQL), NOSQL (MongoDB)

PROJECTS

Neural Network Zero-Inflated Gaussian Process for Spatiotemporal Predictions · Python

November 2022

- Leveraged the number of deaths for each census tract and quarter between 2000 and 2018, utilized Zero-inflated Gaussian Process (ZiGP) to predict the death count in future tract-quarters. Transformed basic ZiGP model by replacing the Gaussian Process which models the non-sparsity of the output with a Neural Network.
- Adapted stochastic variational inference with inducing points to learn models in ZiGP.
- Calculated mean percentage of best-possible reach (%BPR) to analyse model forecast quality. Average %BPR-100 of basic ZiGP is 34.5% and that of the modified ZiGP is 32.3%.

Research on Airbnb Price Prediction Based on Deep Learning · Python (Senior Honors Thesis)

May 2021

- Collected and pre-processed a dataset of 67000 samples using Pandas and NumPy.
- Implemented KNN, random forest, XGBoost and Bayesian Neural Network to predict future Airbnb prices. For Bayesian Neural Network, predicted price via posterior predictive density estimated by Monte Carlo method.
- Got R^2 of 72.3%, and MSE of 0.128 from Bayesian Neural Network.

Simulating Prices for European options · R

March 2019

- Utilized historical data of European options in 10 years to produce future theoretical prices for these options based on Black-Scholes, Binary Model and Stochastic Volatility Model. Compared estimated prices with actual prices to test these three models.
- Achieved the best result from Stochastic Volatility Model, with RMSE 0.981, 1.002, 1.043 in one, two and three months respectively.

EXPERIENCE

Dongxing Securities, Shanghai, China

Investment Banking Analyst

February 2021 to May 2021

- Supported the investment team on complex investigative projects to discover anomalies and unusual behaviors to measure risks in terms of forensic regulation, compliance, and litigation work.
- Analyzed the visualized data using Excel and Python to develop presentation materials for clients and senior management.

Everbright Securities, Shanghai, China

Research Assistant

November 2019 to August 2020

- Developed an Excel system to automatically update company financial data and third-party information. Evaluated new data sources to update our estate database.
- Interpreted latest events and financial data in Real Estate industry to help mentors produce in-depth research and predict future trends. Successfully predicted the trend of Real Estate market in June 2020, doubling the stock price of Everbright in two months.
- Analyzed China Overseas Landing & Investment based on their financial statements, as well as information of comparable companies and the industry, and evaluated the company to help investors make reasonable decisions.