

Zhongyou Shi

zhoyoushi@gmail.com

New York, NY 10002 ▪ (917) 704-6876

EDUCATION

Baruch College, MS in statistics, Data Science Track, **GPA 3.917**; New York, NY
Dean's List

Expected Dec 2021

Stony Brook University, BE in Electrical Engineering; Stony Brook, NY
Dean's List

Dec 2018

PROJECTS

New York City Housing Price Prediction

Gathered 1.6 Million records, stored in the cloud, produced a giant dataset, applied 9 filters, and predicted the prices.

- Generated a single data frame with 1.6 million housing records from the NYC Department of Finance in python
- Cleaned and manipulated the redundant dataset to improve the accuracy of future prediction on housing
- Visualized the dataset into various trending plots to present the correlation between features and the price
- Performed multiple Scikit-learn machine-learning techniques and obtained optimal model for future prediction

Optimal stock selection simulator

Obtained stock tickers and information from yahoo's API, and built a stock selection model to compete with S&P 500

- Requested stock information from yahoo's API and combined the valid stocks into a single data frame
- Utilized median price as references and converted the daily price into daily returns with 10 years transaction
- Built stock selection model based on returns, monthly downside variances, and betas obtained by Scikit-learn
- Simulated stock selection with the model based on top 1, top 3, and top 10 performance which beat the S&P500

Chinese MNIST Classifier

Gathered Chinese digit images from Kaggle and built the neural network model by using Pytorch to classify these digits

- Loaded 15000 Chinese digit images into Kaggle directory and split them into training and test set
- Utilized Pytorch package and converted image data into the tensor matrix and loaded them into batches
- Implemented a feedforward neural network model with 4 hidden layers and SGD as the optimizer
- Performed 45 epochs and achieved 88% accuracy on the unseen test image data with the fitted model

Tutoring Center Database

Designed a database for a language tutoring center with MS Access to simulate a management system

- Designed and drew Entity-Relationship diagrams of a tutoring center and converted them into Relational Model
- Implemented database tables from the normalized Relational Model in SQL with MS Access
- Wrote queries based on several business scenarios as a simulation to monitor the business system

CERTIFICATIONS

- Python Data Analyst, Python Data Scientist, & SQL Data Analyst (DataCamp)
- R & SQL Data Analytics Professional Certificate (Google)
- Python & R Machine Learning A-Z In Data Science (Udemy)

PROFESSIONAL EXPERIENCE

L2 Computer, Inc. New York, NY - *Technician*

Oct 2019 – August 2020

- Diagnosed defective mac books with troubleshooting and problem-solving skills
- Updated information on the service form for further follow up with customers
- Replaced defective hardware parts and reassembling mac book
- Performed final established test procedure before marking it as done on the service form.

RELEVANT COURSES

Principles of Database Management Systems, Programming for Analytics, Data Mining for Business Analytics
Modern Regression Analysis, Multivariate Statistical Methods, Big Data Technologies, Data Visualization

SKILLS AND INTERESTS

Technical Skills: Python, Excel, SQL, R, SAS, Tableau, MS Access, C/C++, Mat-lab, Scikit-learn, Pytorch

Interest: Traveling, Snowboarding, PC building, Coding