Feng Qing

210 S Fourth Street, Champaign IL61820 | 217-979-6704 | fqing2@illinois.edu

EDUCATION

University of Illinois at Urbana-Champaign August 2019-Expected June 2021
Master of Science, Statistics GPA: 4.0/4.0
Relevant Coursework: Data Structures, Database Systems, Statistics and Probability II, Applied Regression and Design, Basics of Statistical Learning, Computational Statistics

Sun Yat-Sen University August 2019-June 2019
Bachelor of Science, Statistics GPA: 3.5/4.0
Relevant Coursework: Advanced Computer Language Programming, Numerical Analysis, Calculus, Linear Algebra, Mathematical Statistics, Multivariate Statistical Analysis and Its Application

SKILLS & CERTIFICATES

Programming Language: Proficient at Python, C++ and R, beginner in SQL and Cypher
Technical Skillset: Proficient at Django, Scikit-learn; beginner in PyTorch, MongoDB and Neo4j
Online Course Certificate: Object-Oriented Data Structures in C++ (UIUC), Applied Data Science with Python (University of Michigan), The Complete SQL Bootcamp (Udemy)

RELEVENT EXPERIENCE

Shenzhen ChengQi Capital (Python, SQL) Shenzhen, China
Quantitative Research Analyst Intern July 2018-September 2018

- Developed an event-driven quantitative strategy using naive Bayes, which outperformed 13% annualized net return, 1.69 Sharp Ratio on backtesting data
- Constructed a MySQL database of statistics like event frequency and price tendency, based on 23 financial events
- Conducted vectorization and optimized code framework, to reduce the run time from 30 to 3 minutes after several iterations of improvement on the codes

RESEARCH PROJECTS

City University of Hong Kong – Management of financial data in empirical study (Python, SQL) May 2019
Research assistant for Prof. Gavin Feng

- Measured firm characteristics and risk factors in the empirical analysis of asset-sector allocation, retrieving data through MERCURY server from WRDS and CSMAR database
- Replicated and updated data to verify some classical models, like the Fama-French 5-factor model

Kesci Competition – Data Visualization and Machine Learning (Python, Seaborn) January 2019

- Predicted quantities of hotel booking with an R^2 of 0.52, by optimizing LightGBM with grid search
- Analyzed variable distributions to implement feature engineering. For instance, discovered a 7-day periodic pattern of "AverageOutput", based on which we used an external API to convert "CheckinDate" into categorized variables as workdays, weekends and holidays

Kaggle Competition – Data Mining and Machine Learning (Python) July 2018

- Predicted the total ride duration of taxi trips in New York city based on over 1.4 million samples, ranked top 10% with an RMSLE of 0.374
- Improved accuracy of the models by adding supplementary datasets from external datasets, like weather data of New York
- Scrutinized the effectiveness and accuracy of 7 algorithms: Xgboost, CatBoost, Gradient Boosting, Linear Regression, Random Forest and LightGBM to determine the best option