Jiayi Wang

Address: New York, NY || Cell: (646)891-8987 || w.jiayi@columbia.edu || LinkedIn: www.linkedin.com/in/JiayiWang-Skylar

SUMMARY

Solid master of applied analytics with strong programming, analytical and mathematical skills. Have a wide range of data analysis, modeling, machine learning, and visualization experience. Strong knowledge of SQL, Python, Spark and R. Detail-oriented and significant ability to work in team environments

EDUCATION

Columbia University *New York, NY* Expected December 2022
Master of Science in Applied Analytics *GPA: 4.0/4.3*
Coursework: Managing Data, Machine Learning, Python, SQL, Statistics

The University of Nottingham *Ningbo, China* June 2020
Bachelor of Science in International Business Economics in the First Class Scholarships in 2016, 2017, 2018 *GPA: 3.7/4.0 (10%)*

PROFESSIONAL EXPERIENCES

Co-founder, Hi-Tech Beauty Instrument Project, Ningbo Intelligent Technology Research Institute Jan 2021 - Present

- Raised $80K from investors for start-up focused on developing innovative facial cleanser that detects make-up residue
- Develop marketing strategy based on targeted consumer segmentation, consumer needs through survey and focus groups; benchmark products with industry trends and competitors' strategies to optimize offer
- Build the potential customers' database, and leverage database to iterate on product design, and automate the data collection process to fetch competitors' information; the prototype of the first generation has been launched [Python]

Consulting Intern, PwC Strategy&, Shenzhen, China Sept 2020 - Dec 2020

- Developed strategic plan for Smart Campus design for a tourist district in Shenzhen by researching industry trends, growth drivers in frontier technology for tourism and transportation sectors [Tableau]
- Gathered insights from 50+ client interviews, and summarized the key industry findings and identified 400+ key Smart Campus offer attributes in a database; presented 30-page in-depth report adopted in the project [Python, Excel]
- Modeled "static" & "chronic" smart transportation scenarios and built information management system architecture

Strategy Analysis Department Intern, NetEase, Inc., Shanghai, China Jul 2019 - Aug 2019

- Identified attractive industries for investment and built investment theses based on research of foreign capital markets
- Conducted due diligence on 3 representative companies in different industries including analysis of market landscape, business model, consumers segmentation, consumer needs and pain points, and growth strategy [Excel, Tableau]
- Crawled complex data on 256 US Sequoia portfolios and shortlisted to 17 priority companies by efficient data cleaning, exploratory data analysis and regression analysis [EDA, ML, Python]

PROJECT & ACTIVITY

Bank Customer Churn Analysis and Prediction Dec 2021-Feb 2022

- Developed algorithms for bank to predict customer churn probability based on 10,000 labeled data via Python programming
- Preprocessed dataset by data cleaning, categorical feature transformation and standardization, etc. [Python]
- Trained Logistic Regression, Random Forest and K-Nearest Neighbors, and applied regularization with optimal parameters to overcome overfitting; Evaluated model performance (accuracy 0.86) via k-fold cross-validation technique [ML]

Airbnb Rental Price Prediction [Kaggle] Sep 2021- Dec 2021

- Developed machine learning models in R to predict 9210 NYC Airbnb rentals' prices
- Preprocessed 400,000 rentals data along 90 variables by removing duplicates, encoding categorical features, and imputing missing values; performed exploratory data analysis and utilized Lasso for feature selection [EDA, R]
- Built linear regression, boost tuned gbm and random forest models to predict rental prices and found optimal model via comparing the root mean square error (RMSE) with 59 and ranked top 5% in Kaggle competition [ML]

Bookstore Database Design and Implementation [Distinction Project] Sep 2019 -Dec 2019

- Created a database solution for a bookseller to reduce business costs and develop customer relationships
- Designed flowchart to manage various entity relations and maximize the allocation of book resources; analyzed the age and gender distribution of customers and ratio of their respective spending on the overall sales [SQL]
- Extracted key elements from customer feedbacks to analyze their satisfaction; built multivariate regression analysis to predict the future trend of book consumption [STATA, Python, Sentiment Analysis]

SKILLS

Programming: Python (sklearn, pandas, numpy), R, SQL, Tableau, Advanced Excel, PySpark, MongoDB, Google Analytics
Machine Learning: Classical & Penalized Regression Methods (Lasso, Ridge), Decision Tree, Random Forest, K Nearest Neighbors, Clustering, K-means, Text Mining, Regularization, Principal Component Analysis (PCA), Model Evaluation
Statistics Analytics: Hypothesis Testing, A/B testing, Time Analysis
Credentials: CFA Level II Candidate, AMAC Certificate Holder, 4th Place 2020 Roland Berger Campus Talents Competition, 2nd Place 2021 Smart C-Terminal Science and Technology Bootcamp ($80,000 initial investment)