Yu (Lana) Jin

yj465@cornell.edu 110 Veterans Place, Ithaca, NY 14850 (607) 262-3539

EDUCATION

Cornell University, College of Engineering, Ithaca, NY
Master of Engineering in Financial Engineering, **GPA: 4.0/4.3** **Expected December 2022**

Peking University, Beijing, CHN
Bachelor of Economics in Finance, **GPA: 3.7/4.0** **July 2020**

Selected Coursework: Monte Carlo Simulation, Machine Learning with Big Data, Python for Finance, Stochastic Processes, Numerical Methods and Principles, Financial Derivatives, Fixed-Income Securities, Financial Risk Management

SKILLS

Technical: C++, MATLAB, Stata, Python, MS Office, Tableau
Languages: Chinese(native), English(fluent), French(advanced), Italian(intermediate)
Certificate: CFA Level II Candidate

EXPERIENCE

Forex Specialist (full-time), *BOE*, Beijing, CHN **(Top New Employee of the year)** **July 2020 to July 2021**

- Designed an automatic system that predicts Forex trend with BeautifulSoup, based on daily Forex comments scraped from the financial news website 'HuiTongWeb' and used Python to perform linear regression analysis based on exchange rates within our transaction cycle.
- Analyzed the Forex trends (EUR/CHN, USD/CHN, USD/JPY) with the two methods above, advised foreign exchange transaction of over 15 subsidiaries biweekly and reduced book loss on Forex transaction for 1 million RMB within 6 months.

Quantitative Research Intern, *China Securities*, Beijing, CHN **May to Sep. 2019**

- Assisted in research on 51 financial quality factors for building multi-factor stock-selection model, including profitability, income quality, cashflow, capital structure, selected and used Python to test the performance of a series of ROE factors under 9 different calculation methods accordingly.
- Constructed WLS regression model on factor exposure in period T and the stock return in period T + 1 based on size-neutral and industry-neutral assumption, and conducted IC analysis with cross-section data and time series data to discover its performance among different industry and time-span.

Quantitative Investment Intern, *SWS MU Fund Management*, Beijing, CHN **Feb. to Mar. 2019**

- Collected, cleaned and visualized data on the prices of 3078 stocks in Chinese A share market within past 245 weeks and performed linear regression on detecting correlation among stock returns.
- Utilized pre-selected factors to conduct linear regression analysis and assisted in selecting significant factors with high factor return based on t statistics and the information coefficients obtained.

PROJECTS

Wine Quality Analysis (team of 3), *Cornell University*, Ithaca, US **Sep. to Nov. 2021**

- Employed ridge and Lasso regression, decision tree, random forest, SVM and KNN to discover the relationship between physicochemical characteristics including acidity, residual sugar, density, pH and alcohol of white wine 'Vino Verde' and its sensory performance, which is an ordinal score from 1 to 10 evaluated by wine experts.
- Imputed missing values with Gaussian Copula model, standardized features with feature scaling, applied 5-folds cross validation and grid search for hyperparameter tuning.

Airbnb Price Prediction (team of 3), *Cornell University*, Ithaca, US **Sep. to Nov. 2021**

- Based on real-valued, Boolean, categorical and text data input, employed linear regression, random tree and ensemble models including bagging, Ada boosting, gradient boosting and xgboosting to discover the pattern of Airbnb pricing.
- Experimented with Universal Sentence Encoder and neural network architecture to pre-train text features from long-form texture descriptions and further incorporate the text features into the training models.

Chaping CTR Research (team of 10), *Peking University & The University of Chicago*, Beijing, CHN **Sep. to Dec. 2019**

- Conducted word segmentation with five different lexicons based on different word attributes, manually traversed over 2000-word segmentation results, modified unreasonable word segmentation, established and constantly updated the word database.
- Used our word database to segment over 300 article titles of the platform, trained LSTM model with MSE, hinge and binary cross-entropy loss to avoid overfitting, performed sentiment analysis and further explored how sentiments reflected in the title affected the click through rate of the articles on the platform 'ChapingWeb'.