

# TIANYU (WILLIAM) ZHANG

(201) 238-9665 ■ [tianyuzhang@nyu.edu](mailto:tianyuzhang@nyu.edu) ■ [GitHub](#) ■ [Linkedin](#)

## EDUCATION

---

### NEW YORK UNIVERSITY

New York, NY

#### The Courant Institute of Mathematical Sciences

M.S. in Mathematics in Finance (Aug. 2018 – Dec. 2019)

- **Coursework:** Java and Python for computational finance, portfolio management, derivative pricing and hedging, machine learning, interest rate model, big data (Hadoop)\*, nonlinear options pricing\*, NLP\* (\* denotes current coursework)

### WUHAN UNIVERSITY

Wuhan, China

B.S. in Mathematics & B.A. in Finance, Ranked 1<sup>st</sup>, First-class Honor (Sep. 2014 – Jun. 2018)

- **Coursework:** Stochastic calculus, regression analysis, econometrics, algorithms, data structure, database (SQL), optimization, statistical inference, time series, deep learning (PyTorch & TensorFlow)

## EXPERIENCE

---

### BNP PARIBAS ASSET MANAGEMENT

New York, NY

*Machine Learning and Quant Research Project Leader* (Sep. 2019 – Present)

- **Hidden Markov Model:** Performed HMM to detect economic business cycle anomaly in equity, currency, inflation rate and economic growth
- **Asset Allocation:** Constructed macroeconomy trading signals by monthly index data and further developed a mix trading strategy with other signals

### FACTSET RESEARCH INC. (DEPT OF PORTFOLIO MANGT & TRADING)

New York, NY

*Data Scientist Intern (Quant Research)* (May. 2019 – Aug. 2019)

- **Overall performance:** Boosted the out-of-sample score from 0.08 to 0.26 in models that predict the illiquidity and urgency signals in a 5-minute execution model; rebuilt features and enhanced performance from 435/1257 to 25/1257 in mutual information ranking
- **Data cleansing:** Fetched 24GB large hedge funds' data and market data on a cloud server
- **Feature construction:** Changed features into dimensionless ratios to rule out the effect of various scale; implemented transformation functions on dimensionless ratios to improve the performance
- **Feature filtering:** Performed KBins-discretizer to draw partial dependency plot of on bins; implemented mutual information, Goodness-of-power Fit, coskew and cokurtosis to rank the non-linear dependency
- **Backtests and predictions:** Filtered out noisy features by multiple metrics and implemented Random Forest, XGBoost and Neural Network for prediction of illiquidity and urgency signals
- **Day-level mkt impact model:** Selected best impact model from guesses and do statistical test on assumptions
- **Bayesian method on news importance:** Performed Bayesian method to estimate the probability of jumps conditional on various news; ranked news by the conditional probability and replace the one-hot encoder

### FULKRUM (A REAL ESTATE INVESTMENT COMPANY)

New York, NY

*Data Scientist Intern (Part-time)* (Apr. 2019 – Aug. 2019)

- **Marco Data analysis:** Used MSA as cuts of regions and performed one-hot encoder on categorical variables; detect multicollinearity between the macroeconomy data
- **Rent prediction model:** Trained Ridge Regressions and Random Forest regressors to predict the rent price around the American; optimized the model by linear searching the hyper-parameters and cross-validation

### GUANGYUNQIANFAN ASSET MANAGEMENT

Shenzhen, China

*Quantitative Analyst & Financial Data Scientist* (Aug. 2017 – Jul. 2018)

- **Trading signal construction:** Trained Random Forest classifiers to classify wise trading intervals from the price-volume data of equities; constructed signals based on the positions and actions of sophisticated traders
- **Reinforcement learning trading:** Built a reinforcement learning trading framework and adopted signals as the state-vectors in Q-learning on portfolios (the total return reaches 14%, the max drawdown reaches 10%)
- **Emotion signals from alternative data:** Utilized news headlines and financial comments (10G) as an alternative data source to do natural language processing (LSTM) by transfer learning on existing models
- **Noise reduction of price data:** Adopted B-spline for reducing the noise of mid-price data
- **Data pre-processing:** Judged outliers with the local outlier factor; handled missing data with decision trees

### KPMG ADVISORY

Shenzhen, China

*Summer Intern* (Jul. 2016 – Aug. 2016)

- **Risk-warning system:** Updated the risk-warning system and constructed a KMV-based model to determine the threshold of EDF (14.9bp) and adopted logistic regression to predict default risks