

# YINGZHI WANG

257 Gold Street, New York, NY, 11201 | C: (929) 213-4519 | yw2744@nyu.edu

## EDUCATION

**New York University, Tandon School of Engineering** (GPA: 3.5/4.0) Sep. 2016 - May 2018  
Master of Science in Finance Engineering, Computational Finance

**Sun Yat-Sen University** (GPA: 3.7/4.0) Sep. 2012 - Jun. 2016  
Bachelor of Science in Mathematics and Statistics

## SKILLS & COURSEWORK

- Programming Skills: **Python, R, MATLAB, SQL, Tableau, Excel VBA, Bash, STATA, SAS, C++, LaTeX**
- Modeling and Analysis: Financial Econometrics, Algorithm and Data Structure, Data Visualization, Feature Engineering
- Math and Statistics: Machine Learning, Probability and Statistics, Stochastic Calculus, Optimization and Linear Algebra

## PROFESSIONAL EXPERIENCE

**Data Assistant, NYU Wagner Graduate School of Public Service, NY** Sep. 2017 – May 2018

- Architected an effective way using AWS Lambda and SNS storing webform data into S3 for backup and security improvement.
- Apply machine learning classification on program applicant web survey data to focus on how different environment and background factors of candidate impact the application result based on Python sklearn.
- Salesforce: Create new contact information objects and improve efficiency of data management more than 50%; queried, cleaned and analyzed students' academic data and visualize degree enrollment distribution of students in Python for Program Team to make management decisions; build solution on duplicate accounts with Python fuzzywuzzy.
- Collaborate with colleagues over inspecting Eventbrite registration errors and provide faculties assistance through ServiceLink.

**Financial Analyst, PrimeAlpha, NY** Jun. 2017 - Oct. 2017

- Construct a Excel VBA model to accurately collect and update daily and weekly financial data (net returns and AUMs) efficiently with accuracy improved 90% from clients' financial reports.
- Conducted quantitative and qualitative research through about 30 investment funds in the real estate space under Bloomberg terminal; conclude key characteristics of the funds and main factor attribution of stock indexes using Machine Learning PCA method.
- Maintained well organized, extensive, and up-to-date due diligence documentation on hedge funds with average AUM of 500MM, testing, and verification/quality control documents and programs in compliance with company and clients standards.
- Evaluated investment funds based on strategy, opportunity set, investment process, risk and portfolio management, and fee structure.

**Data Researcher, Teaching Assistant, Dataguru, CN** Jun. 2015 - May 2016

- Developed a prediction model for course opening situation and student involvement in R using Logistic Regression and SVM method; improved the questionnaire for investigating potential learner's interests on the opening online courses.
- Created my own online courses on the Dataguru website relative to the application of basic statistics knowledge using STATA.
- Evaluated and instructed online candidate performance and provided solutions in data analysis courses (SPSS, SAS, Excel VBA, etc.).
- Organized the development and implementation of daily, weekly or long-range strategic preventive, predictive and scheduled curriculum plans.

## PROJECTS

**Prediction of Credit Default Risk (Python) - Feature Engineering** May 2018

- Predict applicants' repayment abilities using variety of alternative data, including historical telco and transactional information.
- EDA: examine the distribution, handle missing value or anomalies and encoding categorical variables; explore the correlations and make pairs plot between different factors and discuss relationships.
- After building polynomial and main-attribution features, we use Logistic Regression as a baseline and Random Forest as an improved model from Scikit-Learn. Finally, the Gradient Boosting Machine using the LightGBM library gives us the best prediction score.

**NBA Player Type Redefine (Python, R) - Unsupervised Learning** Jun. 2018 - Jul. 2018

- Scraping NBA Players general and team data from Basketball Reference website using module BeautifulSoup and requests.
- Data Dimensionality Reduction with Linear Discriminant Analysis to prevent the increasing volume of space resulting to the available data to become sparse; perform K-Means clustering into 8 groups based on the best silhouette score.

**Course Projects (Python) - Applied Machine Learning, NYU** Feb. 2018 - May 2018

- Investigated breast cancer data to classify between malignant or benign by contributing variables using Decision Trees and measured conditional probabilities; visualized and compared the the results before and after prone the trees.
- Built a Bayesian Network in order to analyze the probability of student suicide based on different factors.
- In order to predict middle price of given time lag, we use Linear Regression as a baseline and Trees, Neural Network and SVM as improved models with rolling dataset. SVM give us the highest predict accuracy as a result.