WENYI (JULIA) XU

Fresh Graduate passionate about turning data into valuable insights. Proficient in analyzing data and implementing statistic and machine learning methods with Python to real-world business problems, including description, prediction and optimization problems
(646) 881-2612 | wenyi.xu@columbia.edu | Portfolio Website | Github | Linkedin

EDUCATION

Columbia University New York, NY
M.S. Business Analytics (GPA: 3.83 / 4.00) Aug 2019 - Dec 2020
- Coursework: Python, R, SQL, Statistics and Simulation, Optimization, Data Visualization, Machine Learning

Shanghai Jiao Tong University (SJTU) Shanghai, CN
M.E. Electronic Engineering (GPA: 3.50 / 4.00) Sep 2016 - Mar 2019

Nanjing University (NJU) Nanjing, CN
B.S. Electronic Engineering (GPA: 4.68 / 5.00, top 5%) Sep 2012 - Jun 2016

PROFESSIONAL SKILLS

- Proficient with Python (NumPy, pandas, Matplotlib, seaborn, scikit-learn, SciPy, Keras), R, SQL programming language
- Solid knowledge in statistics (probability distribution models and hypothesis testing) and ML/DL methods (regression, classification, clustering, dimension reduction, neural networks, ensemble methods)

WORK EXPERIENCE

Panasonic North America Newark, NJ
Data Scientist Intern / Business Intelligence Group Aug 2020 - Dec 2020
- Successfully predicted monthly revenue by forecasting transportation time using XGBoost regression model in scikit-learn
- Kept revenue prediction error under 15% by offsetting the error with constants optimized by RandomWalk Algorithm
- Recommended new products based on product specialists' insight and competitors' data scraped by professional tool import.io
- Conducted sentiment analysis on customer reviews from online distribution platforms using Azure Cognitive Service
- Built interactive reports using Power BI and Excel on an ad hoc basis to visualize analysis results

Balyasny Asset Management New York, NY
Quantitative Analyst Intern / Equity Group Jan 2020 - May 2020
- Forecasted short-term stock price movements based on historical time-series data with a Random Forest (RF) model in scikit-learn
- Identified big financial crises in the past 30 years by analyzing the historical S&P 500 index and VIX volatility index using Python
- Converted 20 classical trading strategies into quantitative factors; Combined these factors to form new 2-way & 3-way factors
- Trained and cross-validated the RF model on a rolling basis; Improved prediction accuracy by 9%
- Extracted a feasible trading strategy from the RF model and proved its superiority by return rate backtesting

Deloitte Shanghai, CN
Data Scientist Intern / Intelligence Group Dec 2019 - Jan 2020
- Automated content curation of a daily tech newsletter for internal consultants; Saved them over 1 hour's daily reading time
- Performed web scraping (BeautifulSoup4) and text mining (NLTK) to identify top 10 news topics and corresponding news articles
- Discussed with internal consultants regularly and served as the liaison between the consultant group and intelligence group
- Taught 2 new interns Deloitte's rules and regulations; Checked their work progresses daily and solved problems with them

Essence Securities Shanghai, CN
Industry Analyst Intern / Equity Group May 2018 - Jul 2018
- Built financial models using functions in Excel and automatically update financial data using Bloomberg Excel Add-In
- Analyzed and visualized public industry data using advanced Excel functions(e.g., Pivot Table, VLOOKUP) and graphs
- Wrote investment reports for the IC industry with unique insights, garnering views from over 1800 potential clients
- Collected information from over 30 industry experts and company executives through regular in-person conversations

PROJECTS

[Time Series, Clustering]	Small Business Growth Pattern in 2010 Winter Olympics	Citadel & Correlation One
[Django, SQL]	Squirrel Reporting Web App	CU
[Web Scraping]	Movies' Worldwide Box Office Revenue Prediction	CU
[Neural Network, SVM]	Forest Fire Detection Using Satellite Images	SJTU
[Least Square Regression]	Indoor 3-D Positioning of Wireless Communication Base Stations	Huawei

The details of all projects can be accessed here: https://wenyixu1994.github.io/