

# SHULIN JI

+1 (201) 668-1150 | sj4016@nyu.edu | [LinkedIn](#)

## EDUCATION

### New York University, Center for Data Science

New York, NY

*Master of Science in Data Science*

Sept. 2022 – May. 2024 (Expected)

- Courses: Machine Learning, Big Data, Deep Learning, Fundamental Algorithms, Time Series, Causal Inference

### Shanghai Jiao Tong University, School of Mathematical Sciences

Shanghai, China

*Bachelor of Science in Mathematics & Applied Mathematics (Zhiyuan Honors Program)*

Sept. 2018 – Jun. 2022

- Honors: Outstanding Graduate (top 3%); Academic and Leadership Merit Scholarship (\$5000 awarded for the top 5%)

## SKILLS

- **Programming and Software:** Python (scikit-learn, numpy, pandas, matplotlib, etc.), C++, MATLAB, SQL, Tableau, LaTeX, SPSS, Weka, Origin, Wind, Microsoft Office (Excel, Word, PowerPoint, etc.), Linux operations
- **Data Science Methods:** A/B Testing, Inference, Machine Learning, DS pipeline in business (cleansing, wrangling, visualization, modeling, interpretation), Statistics, Stochastic Process, Dynamic Systems and Numerical Simulation

## PROFESSIONAL EXPERIENCE

### PayPal

Shanghai, China

*Data Analyst Intern* in Compliance Platforms

Sept. 2021 – Apr. 2022

- Constructed self-inquiry **database dashboards** and official reports to **automate tracking functionality** on business volume and solution efficacy with **SQL** and **Tableau**
- Identified 10+ effective **live issues** to alert potential **algorithmic flaws** with detailed data detection and **interactive data visualization**, and guided further java-based modifications with engineers
- Raised 4 applicative **metrics** for classifying algorithms on imbalanced data with F-beta score, ROC and PR curve
- Designed an automatic **data report email system** for team engineers, and a detailed Compliance **data analysis illustration platform** to enhance communication efficiency between departments

### Yinzhi Tech

Shanghai, China

*Data Engineer Intern* in Product Department

Jul. 2021 – Sept. 2021

- Developed an intelligent platform to automate real-time analysis for investment consulting, including **data dictionary** construction, knowledge graph analysis and **interface design** in **Mockplus**
- Researched the investment diversity of fund managers: constructed a **web-crawler system** in **Python** for data collection, conducted label definition and further **regression analysis**

### Lingjun Investment

Shanghai, China

*Quantitative Researcher Intern* in Futures Department

Apr. 2021 – Jul. 2021

- Generated 2 effective Supply/Demand **alpha factors** on commodities (methanol, rebar, etc.) in **Python** in **Linux**
- Formed a **data pre-processing** program with high generalization ability, later applied to all 48 futures commodities, **wrangling** all features in structure from unorganized details (diverse update frequency, delayed period, etc.)
- Customized **feature selection** with **product-based** knowledge, nearly tripling the raw volume, based on the former self-conducted Industry Analysis Reports for 48 commodities in Quantitative Futures Fundamentals

## PROJECTS EXPERIENCE

### Quantile Factor Model and Applications in Financial Technology

Jul. 2021 – Jun. 2022

- Developed **Quantile Regression** on **Factor Model** with high-dimensional applications in **MATLAB**
- Generalized the 2-D model to fit 3-D scenarios, and captured quantile-dependent factors (unobserved by PCA) on empirical cases with heavy-tailed errors and outliers
- Implemented Quantile Factor Model on individual **stock prediction** (return rate) with 10 quantitative factor indicators (volatility, momentum, etc.), outperforming traditional Approximate Factor Model and PCA by lowering 8~9% MSE

### Statistical Study on Typical Indicators of Credit Card Holders

Mar. 2021 – Jun. 2021

- Launched statistical research on 24 features with **normality test**, **correlation test** and **non-parametric test**
- Applied Machine Learning methods (**LR**, **SVM**, **NN**, **RF**, etc.) for default probability prediction in **Python**, with further statistical analysis in precision, recall and F-1 score, reaching general accuracy at over 80%