

# Shanshan Li

New York – NY 10001

☎ (631) 377-6277 • ✉ [shaniavina@gmail.com](mailto:shaniavina@gmail.com) • 📄 [shaniavina.github.io](https://shaniavina.github.io)

## QUALIFICATIONS

---

- Over 5 years of research and programming experience in statistical modeling and data mining
- Hands-on experience of large data sets analysis and deep understanding of algorithm
- Excellent in communication; Goal and detail oriented; Self learner and quick starter

## SKILLS

---

Python, R, SQL, C++, Unix/Linux, Shell, Hadoop, Spark

## EDUCATION

---

### **Stony Brook University**

*Ph.D in Applied Mathematics and Statistics*

**Stony Brook, NY**

2012 - 2017

### **Nankai University**

*Dual degree of B.S. in Applied Mathematics and B.S. in Economics*

**Tianjin, China**

2008 - 2012

## RESEARCH EXPERIENCE

---

### **Stony Brook University**

*Research Assistant*

**Stony Brook, NY**

September 2012 - June 2017

- Constructed large-scale network structures to estimate stochastic time series of attributes, and implemented machine learning methods with two approaches: Partial Correlation network and VAR adjacency network.
- Designed Multi-Active-Shooting(MAS), a computationally efficient algorithm which applied regularized kernel smoother for solving lasso-type problems, and implemented the algorithm in Python.
- Processed high frequency data (1 minute) of 233 stocks in SP 500 from year 2002 to 2013. Identified patterns of cross-sectional interconnections and created animation for visualizing results. Managed and analyzed more than 10 GB of Data.
- Developed 4 anomaly detection methods applied in multivariate time series: SPC chart, Generalized Likelihood Ratio(GLR), Bayesian single change-point model and Bounded Complexity Mixture(BCMIX).
- Improved the weighted k-means clustering at 81% accuracy. Designed a iterative algorithm to obtain the globally optimal weights based on KKT conditions.
- Led a group to build a innovative self-service system "Intelligent Life Insurance" which is applied in collecting the feedback from potential guests through decision tree computation.

### **Cold Spring Harbor Laboratory**

*Research Assistant*

**Cold Spring Harbor, NY**

April 2014 - May 2016

- Implemented Python modules to process and parse millions of genomic sequences. Constructed SQL relational databases from un-structured files, allowing more flexible and faster exploration for the variant information.
- Developed Bayesian method to identify the genetic mutations from patients. Performed hierarchical clustering method for breast cancer gene expression data.

## WORK EXPERIENCE

---

### **Mckinley Capital Management LLC**

*Research Intern*

**Stony Brook, NY**

June 2013 - August 2013

Won the first prize of HorseRace Portfolio Competition. Processed the massive historic data of 100 stocks for 12 years, constructed a better fitted mathematical model (MNTS-ARMA-GARCH), and obtained a large-scale portfolio with 15% growth rate

## PUBLICATIONS

---

- S. Zhang, S. Li and W. Zhu, "An iterative algorithm for optimal variable weighting in K-means clustering", *Submitted - Communications in Statistics* (2016)
- S. Li and H. Xing, "Estimation and detection of network variation in intraday stock market", *Submitted - Journal of Network Theory in Finance* (2017)