

# Bingyu (Bonnie) Zhang

(+86) 1582-183-9107 | bingyuz77@gmail.com

## EDUCATION

### Fudan University (FDU)

*M.S in Applied Statistics, School of Mathematical Sciences*

**Shanghai, China**

*Sept. 2019 – June 2021*

- Rank: 4/20
- Awards: Outstanding Student (10%)
- Coursework: *Data Mining (A), Time Series Analysis (A-), Machine Learning (A-), Neural Networks and Deep Learning (A-), Game Theory (A)*

### Shanghai Jiao Tong University (SJTU)

*B.S in Mathematics & Applied Mathematics, School of Mathematical Sciences*

**Shanghai, China**

*Sept. 2015 – June 2019*

- Cumulative GPA: **89.18/100**; Rank: **3/57**; **Graduate with Honors**
- Awards: Kwang-Hua Scholarship (5%), SFI Scholarship (5%)
- Coursework: *Mathematical Analysis (91), Advanced Algebra (94), Probability, Real Analysis (98), Scientific Computation (100), Stochastic Process (93), Statistics, C++ Programming (92)*

### University of Oxford

*Mathematics Programme, Hertford College*

**Oxford, United Kingdom**

*July 2017 – Aug. 2017*

- Coursework: *Manifolds (A), Group Theory (A)*

## PROFESSIONAL EXPERIENCE

### Amazon

*Solutions Architect Intern, AWS GCR Public Sector*

**Shanghai, China**

*May 2020 – Oct. 2020*

- Built LSTM model related to English Automatic Speech Recognition of Chinese teenagers for primary schools with 0.134 of MAE and applied it to corresponding English teaching circumstances
- Established churn prediction model for a media company via EMR cluster; provided high-coverage (100%) and high-precision (increased by 8%) forward-looking prediction
- Developed invoice identification solution for public hospitals through Chinese OCR model; supported template reuse with average accuracy >92% and cost savings up to 80%
- Applied regression model to prenatal diagnosis for a high-throughput sequencing company; provided double diagnosis and reduced false positive rate

### YanAnTang

*Data Analyst Intern, Research Institute*

**Shanghai, China**

*May 2019 – Aug. 2019*

- Developed product recommendation system based on customer similarity and taboo matrix; reduced customer service training time from 4–5 days to 3 days
- Expanded company's product database by Python-based web crawler
- Established KNN, SVM and XGBoost classification models on customer satisfaction data, and acquired precision of 0.72 for prediction; contributed to final decisions to target the market

## RESEARCH EXPERIENCE

### Cell Atlas Annotation Based on Joint Scoring and Supervised Learning

*Supervised by Prof. Shuqin Zhang, FDU*

*Sept. 2020 – Present*

- Defined a hierarchical markup language for specifying cell types using specifically expressed genes, in that a cell type can have subtypes
- Assigned overall marker score assessing the value of marker genes; applied TMM transformation to identify representative cells bearing markers
- Generated clusters using Louvain community detection on cell-cell map of jaccard coefficients; trained representative cells with GLMnet model and classified cells across datasets

## **Application of Bayesian Stochastic Volatility Model in Finance**

*Nov. 2018 – May 2019*

*Supervised by Prof. Dewen Xiong, SJTU*

- Established the SV model in the mixed Gaussian space based on Bayesian statistics
- Estimated parameters for standard SV model, linear space model and mixed Gaussian space model by Markov Chain Monte Carlo Bayesian estimation
- Selected stock price of SSE(Shanghai Stock Exchange) Composite Index for empirical analysis; SV-T model (one extended model) performed best with 9.17 of MAE

## **A Study for Regularized Method in Image Restoration**

*Sept. 2018 – May 2019*

*Supervised by Prof. Ling Pi, SJTU*

- Considered total variation, data fidelity and regularity criterion for optimization; used mirror symmetry and Dirichlet boundary for boundary condition processing
- Proved convergence of alternating iteration algorithm; calculated gradients, Lipschitz moduli and proximal operator
- Implemented PALM algorithm and achieved peak signal-to-noise ration (PSNR) of 32.16, a better restoration for thin structures such as thin blood vessels

## **A Study on the Efficiency of Metro Networks**

*Oct. 2016 – Apr. 2017*

*Supervised by Prof. Mijia Lai, SJTU*

- Established topology networks based on small-world model and took transfer effects into account
- Applied indicators such as betweenness centrality to measure significance of specific stations
- Proposed a method for city planners to assess the efficiency of any new lines; helped decide the best route by comparing the proposed parameters and changes of transfer stations

## **SELECTED PROJECTS**

---

### **Learning Hierarchical Features from Generative Models**

*June 2020 – July 2020*

- Used a ladder structure under the training framework of single-layer Variational Auto-encoders, which avoided conflict between optimization of evidence lower bound and efficient representation
- Learned disentangled hierarchical features on image datasets MNIST (stroke width, tilt, and digit), SVHN (color schemes, shape, and digit) and CelebA (skin and hair color, and face identity)

### **Community Mining and Link Prediction**

*Oct. 2019 – Nov. 2019*

- Detected community characteristics of 2 datasets and visualized by Gephi; applied fast-unfolding algorithm and clique penetration algorithm to community mining
- Used network characteristic indexes to carry out link prediction of the 2 community networks; achieved 0.92 of AUC with Adamic Adar algorithm
- Calculated node similarity by 3 definitions and carried out k-means and DBSCAN to identify special nodes (such as celebrities)

## **LEADERSHIP & ACTIVITIES**

---

Class Council of the Alumni Association, SJTU

*June 2019 – Present*

Founder of the Joint Undergraduate Committee, SJTU

*Apr. 2018 – June 2019*

Minister of School of Mathematical Sciences Student Union, SJTU

*June 2016 – Nov. 2017*

Vice Minister of Exploration and Outdoor Life Society Association, SJTU

*June 2016 – Nov. 2017*

Volunteer of Shanghai International Marathon

*Oct. 2016*

## **SKILLS**

---

Programming: C++, Python, R, Matlab

Tools: LaTeX, Linux, Git, SPSS, SQL

Languages: Japanese (JPLT N2); English (TOEFL 104); Mandarin (native)