

Jiajing Chen

Tel: (617)-470-3445
203 Park Dr, Boston, 02215
jiajingchen@hsph.harvard.edu

Education

- **Harvard T.H. Chan School of Public Health** Boston, Massachusetts
M.S. Computational Biology and Quantitative Genetics Aug. 2017 – May. 2019 (expected)
 - Coursework: Applied Regression Analysis, Basics of Statistical Inference, Introduction to Data Science, Introduction to Social and Biological Network, Introductory Genomic & Bioinformatics for Health Research, Introduction to Epidemiology Methods I
- **China Agricultural University** Beijing, China
B.S. Biological Science of Honors Program, GPA: 3.85/4.0 Sept. 2013 – Jun. 2017
 - Coursework: Biostatistics, Bioinformatics, Algorithms Design and Analysis, C Language Programming, Advanced Mathematics (Analysis), Linear Algebra, Probability Theory and Statistics, Mathematical Analysis, Applied Statistical Methods (Multivariable Statistics), Numerical Analysis, Methods of Optimization, Data Mining (Machine Learning)

Skills

- **Programming:** Python, R, SAS, Java, C, SQL, Matlab
- Regression Methods, Statistics Inference, Multivariable Statistics, Object-Oriented Programming (Java, Python), Data Mining, Data Cleaning and Visualization, Optimization, Machine Learning (SVM, Random Forest, Naive Bayes, ...), Deep Learning Platform (TensorFlow), Database, Python Packages (Scikit-Learn, Numpy, Pandas)

Research Experience

- **Academic Research in Ziding Zhang's Lab** China Agricultural University
Machine Learning Application in Computational Biology Apr. 2015 – Jul. 2017
 - Used Support Vector Machine and Random Forest to build a machine-learning based predictors for Protein Ubiquitination sites through feature selection, optimization and cross-validation
 - Performance Evaluation: best-performed model with a AUC value of 0.869
 - Related Skills: Python, R, Database, MySQL, Linux, Machine Learning (SVM, Random Forest, Ensemble methods)
 - First-authored Paper Under Review
 - Built a Front-End web server: AraUbiSite for researchers to use
(Web server url: <http://systbio.cau.edu.cn/araubisite/index.php>)

Work Experience

- **Beijing QiHe Data Ctd. (Start-up in Data Science)** Beijing, China
Data Scientist Intern Jun. 2017 – Aug. 2017
 - Related Skills: Python (numpy, pandas, matplotlib), Tensorflow, R, Java
 - Applied deep learning platform (Tensorflow) in classification of highway images
 - Web scraping and data maintenance in relational database (Python, MySQL)
 - Large-scale longitudinal and geographical data modeling and analysis (R, Tableau)
 - Worked closely with software engineering team to provide front-end data pipeline

Award and Certificates

- National Scholarship, 2014
- SAS Base Certificate, 2016
- Honorable Mention, Mathematical Contest In Modeling, 2016

Online Courses Certificates on EdX and Coursera

- **PH525 Data Analysis for Life Sciences Series, HarvardX, By Prof. Rafael Irizarry, Department of Biostatistics**

PH525.1x: Statistics and R

PH525.2x: Introduction to Linear Models and Matrix Algebra

- **Fundamental of Computing Specialization, Coursera, Rice University**

An Introduction to Interactive Programming in Python

Principles of Computing

Algorithmic Thinking

The Fundamentals of Computing Capstone Exam

- **Python For everybody, Coursera, University of Michigan**

Programming for Everybody

Python Data Structures

Using Python to Access Web Data

Using Databases with Python

- **Data Science Specialization, Coursera, Johns Hopkins University**

The Data Scientists Toolbox

R Programming

Getting and Cleaning Data

Exploratory Data Analysis