SHAN LU

257 Thayer St, Apt 120C, Providence RI 02906

(401)339-7001 • shan_lu@brown.edu

Homepage: baizhima.github.io • GitHub Account: baizhima

Education

Brown University — **Providence, Rhode Island**

Master of Science (Sc.M) in Computer Science — September 2015 – May 2017 (expected)

Courses undertaken: Database System Management, Computer Networks

Renmin University of China — **Beijing, China**

Bachelor of Science in Applied Mathematics — June 2015

Bachelor of Management in Agricultural Economics and Management — June 2014

University of California, Davis — **Davis, California**

Exchange Student (cumulative GPA 3.97/4.0) — January 2012 - December 2012

Experiences

Citadel (Hong Kong) Securities, LLC — **Hong Kong**

Intern quantitative researcher — June 2015 – August 2015

- Parsed FIX Adapted for STreaming(FAST) encoded market data stream from Shanghai and Shenzhen Stock Exchanges
- Implemented a Python/C++ extension that wraps data stream into Python objects by message types (StockStatus, Snapshot, Index, Trade, Order), supporting cross-language function callbacks
- Applied perfect hashing on attribute names to achieve O(1) time field accessing without explicit declarations

Multimedia Computing Laboratory, School of Information, Renmin University of China — **Beijing, China**

Undergraduate Research Assistant (Advisor: Prof. Xirong Li) — September 2013 – March 2015

- Extracted the 2048-dimensional DSIFT descriptors from raw images by using Bag-of-Words and Clustering
- Reassembled the SVM classifier with a Histogram Intersection Kernel to improve the overall top-5 accuracy by 8%
- Multimedia information retrieval research in annotating public source pictures from Flickr by their tag features

Latest Projects

Snowcast (Internet Radio Station) — September 2015

- Network programming based on Berkeley socket API, sending messages between server and clients under TCP/UDP
- Multithreaded programming using POSIX threads on server-side to support non-blocking I/O intercommunication
- Finely tuned streaming rate with respect to each thread in order to play mp3 formatted songs smoothly

New York Times Blogs Popularity Prediction — March 2015

- MOOC course project originated from MIT Analytics Edge, competition held on Kaggle, final ranking 102nd/2923
- Built an ensemble learning model on logistic regression and random forest in R, test set ROC metric: 0.90672

MOOC Certificates

- Analytics Edge (MIT, 90%)
- Bioinformatics Algorithms I (UCSD, 97.2%)
- Coding the Matrix (Brown, 97.5%)
- Computing for Data Analysis (John Hopkins, 99%)
- Computational Thinking and Data Science (MIT, 93%)
- Computational Investing (Georgia Tech, 100%)
- Introduction to Databases (Stanford, 90%)
- Interactive Programming in Python (Rice, 90.9%)
- Machine Learning (Stanford, 95.6%)
- Mining Massive Datasets (Stanford)
- Statistical Learning (Stanford)
- Functional Programming in Scala (EPFL)

Skills

- Languages: Native in Chinese, full-proficiency in English
- Programming Skills: C/C++, Python, Java, MATLAB, R