

Bowen Deng bdeng2@stanford.edu • +1(650) 285-7174 • baolidakai (GitHub)

Objective

Full time Data Scientist/Quantitative Analyst/Software Engineer

Education Background

Stanford University, M.S. Statistics

GPA 4.16/4.3 Anticipated Graduation on March, 2016

Peking University, B.S. Math and Applied Math

Programming Languages

Primary

Python/Java/C++/R/Matlab/SQL/bash/Excel/LaTeX

Secondary

hadoop/HTML+CSS/JavaScript/sed/awk/ruby

Working Experience

Dealstruck, Carlsbad, CA, Quantitative Analyst Intern

06/2015-09/2015

- Built **SQL+python+R** machine which generates bank statement report for underwriting team using Locality Sensitive Hashing and Clustering Analysis.
 - Trained logistic regression and classification tree with delinquency data (10GB) to predict possible future delinquents.
 - Devised Expectation Maximization based model to analyze the trends for sales analytics.
-

Research Experience

Research Area: Machine Learning

Research Topic: Chinese Text Mining

Joint research work with Prof. Ke Deng, Tsinghua University

09/2013 - 02/2014

- Built text splitting tool using Expectation-Maximization algorithm (coded in **C++**).
- Proposed a graphical model of complex Chinese disease names and devised a widely applicable algorithm to do fast classifications, and analyzed the clusters with **python** package snap.py based on hierarchical clustering.
- Trained a neural network with Chinese Wikipedia data and built a search engine for similar Chinese phrase detection applied on medical data (**python**).

Research Area: Statistics

Research Topic: Parameter Estimation

Joint research work with Yanyang Kong

07/2015 - Present

- Developing an efficient Monte Carlo sampling algorithm for general parameter estimation problems with kernel method.
- Proving the consistency of the estimator with probability theory and large sample theory.
- Implementing the algorithm in an **R** package.

Research Area: Bioinformatics

Research Topic: Identification of Mutated Driver Pathways in Cancer

Senior thesis with Prof. Ruibin Xi, Peking University

03/2014 - 06/2014

- Invented and implemented (with **R**) Multiple Value Genetic Algorithm to maximize De Novo score for selected gene pathways in cancer genome.
- Generalized Genetic Algorithm solution to simultaneously identify multiple pathways and other combinatorial optimization problem including the knapsack problem and traveling salesman problem.
- Provided an automatic method using permutation test to choose parameters in the algorithm, including the target number of the cancer pathway and the expected number of pathways.

Research Area: Bioinformatics

Research Topic: Genome-Wide Association Study

Joint research work with Prof. Minping Qian, Peking University

07/2013 - 01/2014

- Studied the high dimensional data from yeast samples and presented a model that described the regulatory relationship among yeast genes.
- Designed an innovative algorithm for computing the matching score between two sequences via permutation method.
- Designed a graph clustering algorithm based on the Metropolis algorithm and identified well-known gene clusters (preprocessed in **SQL**, implemented in **R**).