Zhenqing (ZQ) Li

0	lizhenqing.fudan@gmail.com https://github.com/digitalpig 614-271-5528	http://lizhenqing.net https://www.linkedin.com/in/zhenqingli
	— Experiences ——	

09/2016 - Present

Principal Data Scientist, Capital One, McLean, VA

- Validating behavior, fraud and recommendation models in card customer acquisition and management at model risk office
- Extensively using AWS cloud based computing for modeling and building data engineering pipeline
- Investigating new tools to provide opinions to 1st line card data scientist team
- Maintained/modernized three legacy SAS based model-monitoring tools
- Contributing to various internal open source projects

06/2012 - 09/2016

Research Chemist, Lubrizol, Wickliffe, OH

- Developed a statistical model to predict chemical toxicity, saving registration cost (\$1M)
- Wrote an image processing software to assess hair straightening after shampoo treatment
- Applying Design of Experiment concepts consistently in new product development
- Developing rheology modifier for lubrication and supporting pilot scale-up/manufacturing
- Supervise technicians

— Work Authorization —

Permanent Resident

—— Skills ——

 $Python \, (numpy, scipy, matplotlib, pandas), R(data.table, ggplot2), H2O \, (www.h2o.ai), Spark, \\$

Core Skills SQL, AWS, EMR, Ansible, Terraform

Unix/Linux Experiences Extensively use Linux/Unix system for 10 years

Statistics Statistical data analysis, Design of Experiment (DoE), Six Sigma Green Belt

Language Chinese Mandarin (native), English (fluent)

— Open Source Projects —

Smart Underwriter https://smartunderwriter.herokuapp.com

I made a stochastic gradient descent trained SVM model, using Fannie Mae's single family housing data from 2000 to 2015. My project models the automated underwriting system (AUS) credit decision. For detail on data acquisition, processing and machine learning, please check the data acquisition or training models. You can also checkout the GitHub repo.

NYC Bus Track https://nyc-bus.herokuapp.com.

Analyze NYC's bus status using both historical and real time data. Source: Github

Trulia Scrapping Study of Trulia Traffic

Using Python to scrape Trulia housing data in different US states via its API. Analyze and plot those housing data by using R with ggplot2 and ggmaps. Source: Github

—— Education ——

03/2016 - 05/2016 Fellow at The DataIncubator Inc.

2016 - Present Machine Learning Track Neuron Networks and Deep Learning; Improving Deep Neural

Networks: Hyperparameter tuning, Regularization and Optimization

2016 - Present Functional Programming Track Functional Programming Principles in Scala

2015 - Present Data Science Track by John Hopkins University (Coursera)

The Data Science's Toolbox; R Programming; Getting and Cleaning Data; Exploratory Data Analysis

2015 - Present Fundamentals of Computing by Rice University (Coursera)

An introduction to interactive programming in Python: Part I; An introduction to interactive programming in Python: Part II

05/2008 - 06/2012 PhD, Materials Science ♂ Engineering

The Ohio State University, Columbus, OH

o8/2006 - o5/2008 MS, Materials Science & Engineering

Clemson University, Clemson, SC

09/2002 - 07/2006 BS, Macromolecular Science & Engineering

Fudan University, Shanghai, China

—— Selected Publications ——

Full publication list is at Google Scholar

- Li Z, Guo X, Guan J. An oxygen release system to augment cardiac progenitor cell survival and differentiation under hypoxic condition. Biomaterials. 33(25): 5914-5923
- Li Z, Guo X, Guan J. A Thermosensitive Hydrogel Capable of Releasing bFGF for Enhanced Differentiation of Mesenchymal Stem Cell into Cardiomyocyte-like Cells under Ischemic Conditions. Biomacromolecules 13(6): 1956-1964.
- 2011 Li Z, Guan J. Thermosensitive hydrogels for drug delivery. Expert Opin Drug Deliv. 8(8): 991-1007.
- 2011 Li Z, Guan J. Hydrogels for cardiac tissue engineering. Polymers. 3(2): 740-761; [DOI]