Qimin Zhang

qz2392@cumc.columbia.edu· (+1) 646-525-8660

EDUCATION BACKGROUND

Columbia University, New York, NY 08/2019 – 05/2021
M.S. in Biostatistics, Theory & Method track
Coursework: Probability, Data Science, Statistical Inference, Biostatistics Methods, Experiment Design, Analysis of Longitudinal Data, Survival Analysis, Deep Learning in Biomedical Engineering, SQL Database, Statistical Learning and Data Mining

Sun Yat-sen University (SYSU), Guangdong, China 08/2015 – 06/2019
B.S. in Chemistry | GPA: 3.5 / 4.0 (Calculated by the university)
Coursework: Multivariable Calculus I and II (4.0/4.0), Linear Algebra (95/100), Probability (84/100), Artificial Intelligence and Neural Network (95/100), Mathematical Modeling Practice (92/100), Chemistry Information (4.0/4.0)

University of California, Los Angeles (UCLA), CA 08/2017 – 09/2017
Summer School | GPA: 4.0 / 4.0
Coursework: Intro to Statistical Reasoning (4.0/4.0), Economics (4.0/4.0)

Coursework (Online): Machine Learning, Python Programming, Applied Regression Analysis, Nonparametric Statistics, Stochastic Process, Survival Analysis, Data Structure and Algorithms

RESEARCH PROJECTS

Data Augmentation for brain tumor segmentation (Supervisor: Prof. Jia Guo)
Columbia University Biomedical Engineering 03/2020 –

- Implement traditional data augmentation methods like flip, rotation or translation with TorchIO, as well as Generative Adversarial Network (GAN) to significantly increase the size (and potentially representativeness) of training sets and further improves supervised tumor segmentation performance with deep learning.

Intelligent denoising of Tailored MR Fingerprinting data (Supervisor: Dr. Sairam Geethanath)
Columbia University Biomedical Engineering 03/2020 –

- To denoise Tailored MR Fingerprinting natural contrast reconstructed image across all contrasts using deep learning, specifically L1-regularized U-Net.

Independent Component Analysis (ICA) on brain images and signals (Supervisor: Prof. Seonjoo Lee)
Columbia University Department of Psychiatry 02/2020 –

- Modify 'colorICA' for larger scale data by employing mini-batch method and gradient descent;
- Paper review and presentation for 'Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker'.

WORK EXPERIENCES

Merck Co., Ltd., North Wales, PA
Trial Optimization Data Intern, Merck Research Lab 05/2020 – 08/2020

Yimian Data Co., Ltd., Guangdong, China
Data Analyst Intern, Data Science Lab 06/2019 – 08/2019

- Used Hive to streamline data storage and manipulation; used Tableau for data interpretation and visualization.
- Developed Python crawlers to automate the data collection from various online sources; conducted data cleansing to reconcile different data types and remove redundant attributes; applied various libraries in Python on daily work.
- Employed clustering algorithm on 100000+ social media users, with word segmentation tools used on the content they created; researched on frequency pattern on social media; Presented effective clustering method for incoming precise advertising.
- Assisted a client to create a market report of a giant tech company by analyzing the data from one of the biggest e-commerce platforms in China.

VOLUNTEER ACTIVITIES

US-China Health Summit
Volunteer 03/2020 –

- Data analysis and typesetting with Latex in R markdown for 'COVID-19 Morning Post', a newspaper for readers in China and USA.

SKILLS

Programming Languages / Software: Python, R, SQL, Tableau, SPSS, Microsoft Office, Spark/Hive, LaTex
Soft skills: Case study, Critical thinking, Team Collaboration, Communication, Public Speaking, Scientific Research