Weiwei Qi

Tel: 646-683-6801 | E-mail: wq2151@cumc.columbia.edu | Website: https://weiweiq.com

EDUCATION

Columbia University in the City of New York **New York, US**

Master of Science in Biostatistics, Theory and Methods track | GPA 3.67/4 Expected May. 2021

- **Course**: Data Science I & II (Statistical and Machine Learning methods), Deep Learning in Biomedical Engineering, Biostatistical Methods I & II, Statistical Inference, Probability, Epidemiology

China Pharmaceutical University **Nanjing, China**

Bachelor of Business Management, Pharmaceutical track | GPA 3.34/4 Sept. 2015–Jun. 2019

- **Course**: Management, Marketing, Linear Algebra, Micro & Macro Economics, Operation Research, Securities Investments, Pharmaceutics, Pharmacoeconomics, Statistics Principles, Mathematical Statistics

EXPERIENCE

Data Augmentation for Brain-Tumor Segmentation, Columbia Engineering **New York, US**

Research Assistant Mar. 2020-Present

- Write python code to build, train, validate, and test the Neural Network (U-Net) for brain-tumor segmentation using **PyTorch** and **TorchIO** on Google Cloud Platform with total dataset size of 335; Achieved a Dice Score of 0.81 on the test dataset;
- Implement traditional data augmentation methods such as flip, random bias field, and random noise to increase the dataset size to avoid the network from overfitting and memorizing the training set. After training with both original and augmented data, the model dice score is improved by 6.1%;
- Adopt state-of-the-art Generative Adversarial Network (GAN) architecture as a new data augmentation method and apply it into generating new high-quality data to further grow the dataset, which helps build a well generalized model for T1 neuroimage data.

CUG Golden Shield Environmental Technology **Wuhan, China**

Assistant to Chief Executive Officer Feb. 2019-Jun. 2019

- Assisted the CEO with daily administrative duties; Translated literatures for leadership decision-making in resources and technology;
- **Excel** and **Word** documents processing, such as **PivotTable** and **Formula Functions** to provide information support;
- Performed **desk research** regarding competitors, advanced soil remediation technology methods and drafted reports to executives.

School of Business, China Pharmaceutical University **Changzhou/Nanjing, China**

Summer Research Intern Jun. 2016-Sept. 2016

- Visited and conducted questionnaires with deans, attending doctors, pharmaceutical personnel and patients at five hospitals in Changzhou City, Jiangsu Province to provide data entry support as part of a multi-year study;
- Inquired about the basic medical facilities, drug reserve and usage, patient distribution, patients' understanding of the national drug subsidy policy, and the salary status of pharmaceutical personnel.

PROJECT

Data Analysis for Drug Abuse Related Death in Connecticut Oct. 2019-Dec. 2019

- Discovered the Number 1 drug of causing death (Fentanyl) and the association between age and illicit drug use by visualizing and analyzing the dataset. Built a project website with the team for final presentation;
- Matched the original county/city column with new geographic coordinates data downloaded from Google to expand the original dataset for visualizing the distribution of drug illicit use, age, race, location on state map. Built and deployed the interactive map to the project website with **ggplot2, Dashboard, Leaflet** and **R Shiny app**;
- Built and validated the model to predict the death rate associated with predictors such as age, type of drugs, race, and location.

Regression Analysis for Gender Discrimination in Setting Salaries Oct. 2019-Dec. 2019

- Conducted data exploration and visualization on existing data regarding two-year salaries of doctors at a college; Confirmed initial assumption that the discrimination exists, with hypothesis testing results computed by R;
- Built regression models with stepwise function and stratification; Validate gender as a main factor in faculty salaries among different title population; Conducted model diagnostics to identify outliers & influential points with Cook's distance and DFFITS as metrics.

Interpretation of China New Normal Economic Mar. 2017-Jun. 2017

- Outlined the main characteristics of domestic economy in the past three years with GDP data such as industrial structure, employment rate, and population data by fulfilling exploratory data analysis;
- Adopted subdivision analysis and comparative analysis to predict the new opportunities, changes, and the industry trend brought by the New Normal. Produced a final report and presentation.

AWARDS & ACTIVITIES

The First Prize of the Seventh Securities Investment Preliminary Contest, School of Business at CPU, Jun. 2018

Third-class Scholarship for Merited Students, China Pharmaceutical University, Mar. 2018

US-China Health Summit, Covid-19 News Report Group **New York, US**

R developer Mar. 2020-Present

- Develop R functions to auto merge and translate data input such as row and columns names to produce Chinese version chart output;
- Match countries with different colors to create tidy and consistent plots; Rewrite the R script with pipeline to optimize workflow.

SKILLS & CERTIFICATES

- Programming: Python, PyTorch, NumPy, Jupyter Lab, R, R shiny, SQL, SPSS, Visual Basic 6.0, Excel, GCP, GitHub, Linux
- Field: Data Analysis, Data Visualization, Statistical & Machine Learning, Deep Learning and Convolutional Neural Network
- Coursera: R programming; Database with Python, Applied Data Science in Python