Hao Zheng

Redrange.zheng@gmail.com | linkedin.com/in/haozheng98 | (+1) 3474012175 | Github: HaoZhengs

EDUCATION

Columbia University, United States   09/2021 – 05/2023
MS in Biostatistics, with a concentration of Public Health Data Science

University of Liverpool (UoL), United Kingdom   09/2019 – 07/2021
BSc in Applied Mathematics | Cum GPA: 3.72/4.0

Xi'an Jiaotong-Liverpool University (XJTLU), China   09/2017 – 06/2019
BSc in Applied Mathematics | 2017-18 XJTLU Academic Excellence Award (5%)

Core Courses: Statistical Inference / Data Science / Advanced Statistical Computing / Programming in Java / Numerical Methods

SKILLS

Programming Languages / Software: R | SQL | Python | Java | PostgreSQL | Power BI | SPSS | Excel | Latex
Libraries: Tidyverse | Tidytext | dplyr | ggplot | NumPy | SciPy | Pandas | Matplotlib | PyTorch | Keras

INTERNSHIP EXPERIENCES

Data Analyst Intern, Wunderman Thompson, Shanghai, China   06/2021–08/2021
- Used Power BI to help visualize the performance data for NIKE WMP platform, a platform which labels the customer when they access NIKE's WeChat official account. Presented the results for WMP platform using PowerPoint.
- Used MySQL and Excel to select and produce chart for customer desired information.
- Produced Word clouds to illustrate the information provided by Sina Public Opinion monitoring application, commented on and offered further opinions on the product or brand.

Data Analyst Intern, Ant Financial (the World's Largest Fintech Unicorn), Hangzhou, China   07/2020 – 09/2020
- Analyzed characteristics of foreigners' transactions on Alipay; determined significant, distinguishable, and predictive features, and wrote scripts for policy calls to improve the ability of Alipay risk control engine to identify and cover black incidents
- Used MySQL to complete the driver table and mark black and white events
- Associated the driver table with the device information table and the member information table and marked statistics of cases
- Analyzed characteristics of 20+ devices and assessed the differentiate ability (woe and IV values were calculated)
- Completed simple java script scripts with relevant features for online policy invocation (covering device language version, expense account and environment, high-risk recharge time period, binding the phone country and device language, and whether the input account is an international account, etc.)
- Mastered the model strategy, AB Testing, AARRR funnel model and Swap Set analysis

RESEARCH EXPERIENCES

Robust Indoor Localization Based on Deep Neural Network (DNN)
Research Assistant | Summer Undergraduate Research Fellowship, XJTLU | Advisor: Prof. K. Kim   06/2019 – 09/2019
- Analyzed and improved an existing multivariate dataset for DNN indoor localization, focusing on sensitivity evaluation and predictive accuracy improvement.
- Built a workflow for indoor localization based on Wi-Fi received signal strength (RSS) and geomagnetic field data, covering dimension reduction of input data using autoencoder (AE), and location estimation using convolutional neural network (CNN) and multilayer perceptron (MLP) methods; leveraged Python and PyTorch to implement the model.
- Systematically trained, validated, and compared the performance of CNN and MLP models.

COURSE PROJECTS

Newton-Rapson Optimization: Analyses of daily COVID-19 cases, hospitalization, death in NYC
Advanced Statistical Computing Course Group Project, CU   03/2022
- Joined a team of 5 to model the number of cumulative cases, hospitalization and death in a pandemic wave by the Richard growth curve.
  - Implemented a Newton-Rapson algorithm in four dimensions modified with the BFGS method.
  - Observed and fitted four curves for cumulative cases, hospitalization and death for each of the five boroughs with selected suitable initial values
  - Make comparison between the fitted curves of each pandemic wave and borough.

EXTRACURRICULAR ACTIVITIES

Team Leader | The Interdisciplinary Contest in Modeling   02/2020 – 03/2020
- Predicted and maximized Huskies' football team's performance, garnering the Meritorious Winner Award (Top 6.69%):
  - Built network model to describe ball passing process, allowing the mathematical representation of team collaboration from microscopic and macroscopic perspectives.
  - Built, trained, and validated a neural network (NN) performance model to evaluate the teamwork based on a set of performance indicators.
  - Built a Markov decision process (MDP) model with states and actions (pass, duel, foul, shot) defined to determine the optimal policy associated a simulated soccer match.
  - Integrated NN performance model and MDP model to evaluate the universal applicability of the optimal strategy.