SHUN XIE

sx2337@cumc.columbia.edu • 917-8536527

EDUCATION

Columbia University, New York, NY   Sep 2022 - present
MS in Biostatistic, Public Health Data Science
Imperial College London, London, UK   Oct 2018 - Jun 2022
MSci in Mathematics (Bachelor's and Master's degree equivalent)
  · **Undergraduate GPA**: 3.7/4.0, **Graduate GPA**: 3.83/4.0, **First-Class Honours Degree** (UK)

RESEARCH EXPERIENCE

Correlation between Unemployment and Wages   Oct 2021 - Jun 2022
Master Thesis, Supervisor: Prof. Mauricio Barahona, Dr. Jonathan Clarke, Imperial College   London, UK
  · Applied Distance Correlation in a new field and captured additional 10% non-linear correlations
  · Implemented spatial regression with approximate profile-likelihood estimator (APLE) to solve the problem of nonzero spatial correlation
  · Iterated over confounders and concluded that the correlation arises from confounders under different time lags

Depression Status Predicted by COVID-19 Associated Behavior Change   Jun 2021 - Sep 2021
Group Leader, Supervisor: Prof. Shuangge Ma, Yale University   Remote
  · Increased accuracy by an average of 6% after replacing selected models (Logistic Regression, Mixed Logistic Regression, Random Forest, k-NN) by Linear SVM
  · Improved sensitivity by 0.6 in k-NN and LR using Under-sampling, Lower-sampling to tackle unbalanced data.
  · Identified Random Forest achieved the optimal F-score (0.787) and sensitivity (0.750) using 4-fold cross-validation

Evaluation of Supervised Classification Models for Image Recognition   Jan 2021 - Mar 2021
Academic, Instructor: Prof. Mauricio Barahona, Prof. Kevin Webster, Imperial College   London, UK
  · Built and optimized the Multilayer Perceptron model to prevent overtraining based on the performance on CIFAR-10 dataset
  · Compared results between MLP and CNN model; Concluded that CNN model has a 20% higher accuracy at epoch 40 and saved 7 megabytes of data

Comparison of Different Word Embedding Models based on IMDb Reviews   Jun 2020
Group Leader, Supervisor: Prof. Kelvin Webster, Imperial College   London, UK
  · Compared LSA, GloVe, and Word2vec models via theoretical analysis using intrinsic and extrinsic evaluation and discovered no global difference among models after applying sentiment analysis on IMDB movie reviews

WORK EXPERIENCE

Caitong Security Asset Management Co. Ltd.,   Jun 2020 - Aug 2020
Data Analyst Intern, Marketing Department   Shanghai, China
  · Improved accuracy by imputing missing value using developed formula.
  · Analyze data extracted from Wind and generate analysis report on the feature of public equity fund

PUBLICATIONS

Shun Xie, **Comparison of Models' Performance for Predicting Depression under COVID-19**, accepted by the 2021 International Conference on Statistics, Applied Mathematics and Computing Science (**CSAMCS 2021**)

SKILLS

Python, Matlab, R, STATA, SPSS, Latex, Scikit-learn