Xiran Wang

(803) 567-0771 | xiranwangwork@gmail.com | https://www.linkedin.com/in/xiran-wang-4b6715158/ | Phoenix, AZ

EDUCATION

M.S. in Statistics | University of South Carolina - Columbia, SC USA | May 2020
Full Scholarship | *Graduate Assistantship* | *GPA: 3.8/4.0*

B.S. in Statistics | University of South Carolina - Columbia, SC USA | May 2018
Minor: Risk Management and Insurance
Cum Laude | *Mu Sigma Rho* | *Dean's Honor List* | *GPA: 3.7/4.0*

TECHINICAL EXPERIENCES

RNN-LSTM Learning for S&P500 Stock Market Prediction from Quantitative Data | April 2020
- Created highly flexible stacked LSTM algorithm capturing dynamic stock data every minute from API & Wikipedia and making stock price prediction.
- Tuned the stacked LSTM model with optimal hyper-parameters among 100+ candidate models with TensorFlow.
- Minimized test loss to 0.002, which is 250% less than the loss of the baseline LSTM model.

Predictions of Smoking-Related Disease Cases and Medical Expenditures | Dec 2019
- Invented a "bridge" model transferring the prediction problem from "smoking⇒$cost" to "smoking⇒chance of getting disease⇒$cost" for solving the inherent causality snags in cost prediction.
- Selected and validated optimal multivariate linear regression model and logistic regression model to classify disease cases and medical costs by minimizing collinearity, influential points, and normality departures problems.
- Produced expenditure predictions for medical treatment of disease cases as a function of various data clusters identified by selected model features.

Time Series Analysis on U.S. Suicide Rate | May 2019
- Inferred1940-2016 suicide rate in U.S. using time series ARIMA model.
- Maintained the model's accuracy prediction for extremely unusual events, such as The Great Depression.
- Minimized average prediction error down to 0.72, secured 100% of predictions within 95% confidence interval.

Cost Differences Between Patient Clusters in the U.S. | Nov 2018
- Performed model and feature selection through an implementation of regression model diagnostics and collaboration with subject-matter experts.
- Increased R^2 by 280% from the baseline multivariant linear regression model.
- Solved client request for an uncertainty metric on predicted differences in average expenditures by patient cluster through the use of delta method- and Bootstrap-based standard errors and confidence intervals.

PROFESSIONAL EXPERIENCES

Data Scientist Intern | June 2020 – Present
Business Analytics & Data Science, Apollo Education Group – University of Phoenix | Phoenix, AZ
- Developed and applied analytical, statistical, and machine learning methods that are related to project work.
- Provided sound perspective on modeling approach, technique and tools in resolving the business problem.
- Engaged with various teams to perform research and provide actionable recommendations in work-related areas.

Lecturer and Teaching Assistant | Aug 2018 – May 2020
Department of Statistics, University of South Carolina | Columbia, SC
- Assumed responsibility for lecturing and coaching to 70+ undergraduate students in STAT 201 Elementary Statistics.
- Delivered streamlined lectures every week focusing on fundamentals of modern statistical methods, including probability, random sampling, hypothesis testing, estimation methods, and linear regression.
- Led students through end-to-end data science pipeline: process of data manipulation, data analysis, and result interpretation to solve real-life problems.

TECHNICAL SKILLS

R, Python R Markdown, SQL, High Performance Computing, MATLAB, SAS, Weka, LaTeX, Tableau, Microsoft Office.

CERTIFICATIONS

• SAS Certified Base/Advanced Programmer for SAS9 • Python Structure • SQL for Data Science • Data Wrangling and AB Testing • Neural Networks and Deep Learning

AFFILIATIONS

• American Statistical Association • US National Statistics Honor Society • Chinese Youth Volunteers Organization
• Student Body Government