Jianing (Julia) Chen

jchen801@usc.edu · LinkedIn · (765)337-0250

Education

University of Southern California, Los Angeles, CA — Jan. 2021 – Dec. 2022
Purse Master of Science in *Applied Data Science*.

University of California, Santa Barbara (UCSB), Santa Barbara, CA — Sep. 2019 – Dec. 2020
Earned Bachelor of Science in *Statistics and Data Science*, *GPA*: 3.60/4.0.

Purdue University, West Lafayette, IN — Aug. 2016 – May. 2018
Intended Bachelor of Science in *Applied Statistics*, *GPA*: 3.60/4.0.

Related Project

Heart Failure Survival Analysis — Nov. 2020 – Dec. 2020

- Applied Kaplan-Meier estimate, Cox Proportional Hazard model and Likelihood test, along with Forward Selection with comparing AIC/BIC value to find the variables that are statistically significant in leading to a heart failure medical record for the patient.
- Improved the model using Stratification; utilized KNN and Random Forest to build classifiers and obtained 82.67% accuracy based on age, ejection fraction, high blood pressure, serum creatinine.

Literature Review about Interventions of COVID-19 — May. 2020 – Jun. 2020

- Studied the COVID-19 SEIR-based Transmission Dynamical modeling and Bayesian Epidemiological modeling and quantified the impact of the intervention.
- Concluded that the intervention strategy, such as physical distancing in both Wuhan and British Columbia are helpful, but over-relaxation on intervention will lead to an exponential increasing on the infection.

Building a Classifier for Credit Default Risk — May. 2020 – Jun. 2020

- Extracted useful insights by visualizing relationships between predictive features using Python Altair.
- Applied Principal Component Analysis to reduce redundancy variables among bill amounts and payment amounts, successfully reduced 12 variables into 7 PCs and achieved 78 % of the accuracy.
- Built precision and recall logistic models for oversample, undersample, simple logistic model and undersample with feature selection. Proved that under-sampled model with feature selection will be the best fit, which gives 44.23 % precision and 56.72% recall.

Time Series Analysis Project — Feb. 2020 – Mar. 2020

- Analyzed the behavior of monthly reported chickenpox breakout and identified a best SARIMA model fitting the breakout in New York City from 1957 to 1971.
- Utilized SARIMA model to delineate disease patterns and accurately forecasted 12 months into the future, and learned how time series is being practical in other related contagious disease analyses.

Courseworks & Skills

Relevant Courses:

Statistics Machine Learning, Bayesian Analysis, Regression Analysis, Probability & Statistics, Time Series, Design of Experiment, Statistical Process, Bayesian Analysis, Linear Algebra, Multivariate Calculus, Ordinary Differential Equation, Discrete Math

Programming Skills:

Python, R, SAS, SQL, C, C++, Java, Android Studio, LaTex, etc.