Junjie Li

San Francisco
(415) 244-5428
jli5064@gmail.com

Education

Sep 2018 - June 2023
University of California, San Diego | B.S Data Science | 3.0 GPA

Relevant Courses

- Data Structures & Algorithms • Object Oriented Programming • Statistical Methods
- Machine Learning
- Data Management

Skills

- Python • SQL • NumPy • Pandas • SciKit-Learn • Java
- Jupyter • Microsoft Suite • Machine Learning

PROJECTS

Jan 2021 - Mar 2021
Video Game Sales

- Cleaned messy dataset, carried out EDA, and used linear regression on a Kaggle video games dataset to find what attributes affect sales the most.

May 2021 - June 2021
Predictive Model for Stroke Occurrence

- Used forward-stepwise selection to determine the best predictors for stroke in a logistic regression model in this dataset.

May 2021 - June 2021
Power Outages in the U.S.

- Cleaned data, performed permutation test to measure whether the missingness of a column is missing at random or missing completely at random.
- Used Grid Search to find that Decision Tree Classifier as the best model to predict the causes of power outages.
- Performed permutation test at the end to test for fairness of model.