JUSTIN GEE

justing33193@gmail.com ◇ Portfolio ◇ LinkedIn ◇ Github

EXPERIENCE

Statistical Consultant *Sep 2019 - Dec 2019*
University of Massachusetts Amherst

- Worked closely with clients providing statistical and technical assistance on experimental and analytical projects. Consultations included recommending statistical methods, and experimental design.
- Developed and implemented a variety of statistical methods used for departmental student survey. Topics include survey design, methods for data collection, and hypothesis testing.

TECHNICAL SKILLS

Programming	Java, Python, R, Scala, SQL
Libraries	NumPy, Pandas, Scikit-Learn, TensorFlow
Visualization Tools	GGplot, Matplotlib, Seaborn, Tableau
Other Tools	Git, Kafka, MySQL, Spark, SQL Server

RELEVANT COURSES

Probability & Statistical Inference, Statistical Computing, Regression Analysis, Applied Mathematics, Data Analytics & Visualization, Corporate Finance

PROJECTS

Graduate Admissions Prediction & Analysis

- Used Gradient Boosting algorithm in Python to predict the probability of acceptance to graduate.
- Used Grid-Search and cross-validation for hyperparameter optimization.
- Improved R^2 score from 0.896 to 0.963 after hyperparameter optimization.
- Results found GPA and GRE Score to be the most influential when determining graduate acceptance.

Melbourne Housing Market

- Used Decision Tree and Random Forest algorithms in R; predicted the price of a home given its features.
- Conducted data analysis and feature engineering to reveal variable interactions associated with housing prices.
- Achieved R^2 score of 0.86 using Random Forest algorithm.
- Results found home location, type, age and size to be strong features to predict housing prices.

Student Performance Analysis

- Performed data analysis on a public dataset about student performance to identify variable interactions associated with final grades.
- Used Python's Seaborn to constructed data visualization and Python's SciPy to use statistical hypothesis tests to gain insight into relationships associated with final grades.

Micro Center Scraper

- Used BeautifulSoup framework for web automation; extracted product data from HTML for data measurement.
- Developed ETL pipelines to build a database hosted on MySQL using Python database interfaces.
- Performed data analysis on product data to evaluate their on-shelf presence along with sales knowledge regarding locations across the country.

EDUCATION

University of Massachusetts Amherst, Amherst *May 2020*
B.S. Mathematics; Concentration in Statistics