Jacob Merrell

303-960-5156 • merrell.jacob@gmail.com • data science website

EXPERIENCE

Data Analyst/Data Scientist *Mercer*
February 2019 – Present

- Used discriminant classification/clustering model and geospatial data in SAS to impute missing data (with 89% accuracy on test data) on a 2 TB Medicaid dataset
- Risk adjustment project leader for California's Medicaid program. Oversee running of regression models in SAS to create cost weights program. Results are implemented state wide
- Predicted dispensing fee costs for pharmacies using R and linear regression. Preprocessed the data and imputed values where there were errors
- Automated a process in Python saving 30 minutes to an hour of time for each project
- Mange trainings for new employees and provide mentorship

Data Science Analyst *Brigham Young University*
September 2017 – November 2018

- Linked individuals (using random forests and logistic regression in R) with 90% accuracy from a dataset of 60,000 European immigrants to a pioneer dataset recorded in the Utah Valley
- Created quadratic discriminant clustering model in R to identify malignancy of cancerous tumors
- Led projects and managed two other junior statistics analysts.

Actuarial Analyst *Milliman*
December 2015 - August 2017

- Used R to simulate large claims experience and based accrual recommendations on results
- Developed a process to streamline the creation of pro forma scenarios which showed high and low estimates for the gain/loss

Statistics/Property Valuation Intern *The Church of Jesus Christ of Latter-day Saints*
May 2014 - Dec 2014

- Saved $72 million by building a VBA model to predict membership utilization of meeting houses
- Created a multiple regression model in R to predict land values

PASSION PROJECTS AND SKILLS

- Skills and Languages: Python, R, SAS, SQL,VBA, Spark, pandas, scikit learn, TensorFlow,
- To improve my own YouTube videos I accessed YouTube's API and XGBoost in Python to predict success of YouTube videos with 86% accuracy. Built deep learning model using TensorFlow, Google Vision and OpenCV for facial and text recognition (See More)
- Scraped box office data using Python's Beautiful Soup package for nearly 13,000 movies. Trained random forest model which explained 80% of the variation in box office revenue (See More)
- Used spatial regression and geospatial data to predict air pollution for most of the USA. The adjusted R-squared was 0.6129, with a normalized RSME of 0.174 (See More)
- Estimated the expected savings ($86.69) for using solar energy using a time series AR(1) model. The model had an adjusted R-squared of 0.9376 (See More)
- Fluent in Spanish (speaking, writing, and reading)

EDUCATION

Brigham Young University, Provo, UT
Bachelor of Science, Actuarial Science August 2015

- Cumulative GPA 3.91/4.0
- Certificates: SAS Programming, and Advanced Statistics
- Actuarial Exams: Financial Mathematics (FM), Probability (P), and Models for Financial Economics (MFE)