Jenny Rhee

jennyirhee@gmail.com | jennyrhee.github.io | jennyrhee | jenny-rhee-84ab5576

Skills

Languages/Tools Python (NumPy, pandas, matplotlib, SciPy, scikit-learn, etc.), AWS Redshift, PostgreSQL, Git, Tableau, Excel

Technical Skills Supervised and unsupervised machine learning, experimental design, A/B testing, web scraping, statistics, NLP

Marketing Tools Braze, Leanplum, Google Analytics

Experience

Marketing Analyst — Lafayette, LA

WAITR — *November 2019 to present*

- Tracked and analyzed the performance of marketing campaigns (push notifications, emails)
- Designed statistically rigorous A/B tests and analyzed results to optimize marketing campaigns
- RFM customer segmentation using k-means clustering and quartile bucketing
- Identified an approximately $1M per year stream of cannibalized revenue from coupon abuse
- Analyzed the order behavior effects of a loyalty program (+200% average orders per year) leading to program redesign with the goal of increasing customer penetration and engagement
- Created a marketing dashboard to track weekly KPIs and wrote scripts to automate complex data retrieval
- Worked with the sales and business intelligence teams to research restaurants not yet on the app and scrape data that was necessary for restaurant acquisition efforts
- Technologies used: Python, AWS Redshift, PostgreSQL, Tableau, Sisense, Braze, Leanplum

Research Associate — Baton Rouge, LA

LOUISIANA STATE UNIVERSITY — *June 2019 to November 2019*

- Compiled data from various sources to build an econometric model to analyze the effects of driving forces on a variety of environmental impacts in the U.S. and Germany over the past two centuries
- Developed a simulation model in Python to illustrate the relationship between embodied energy and fitness of a K-selected species
- Technologies used: Python, Excel

Data Analyst Intern — Lafayette, LA

ACADIAN AMBULANCE — *Aug 2018 to May 2019*

- Led the first data science projects to be conducted at the company
- Exploratory analysis (clustering via scikit-learn and topic modeling via Gensim and MALLET) of rejected medical records to categorize and understand the common reasons for rejection
- Analyzed the "virality" of medics with low medical documentation accuracy on their partners with the goal of changing training procedures
- Technologies used: Python, Microsoft SQL Server

Great Lakes Summer Fellow — Ann Arbor, MI

UNIVERSITY OF MICHIGAN — *May 2018 to Aug 2018*

- Developed a data processing script in Python to normalize historical time series data from 15 stations and buoys in the Great Lakes (2015-2017; 196 million observations)
- Proposed a data management solution for 15 stations and buoys in the Great Lakes. Data were in inconsistent formats and units over time and between stations, ranged over a decade, and stored as flat text files on a server.
- Designed and implemented a time-series database (TimescaleDB) prototype to manage historical and real-time streaming data
- Technologies used: Python, TimescaleDB

Projects

Predicting sentiment from tweets about U.S. airlines (Web App) (Github)

- Used a labeled data set of tweets about U.S. airlines to train a model to predict sentiment
- Technologies used: Python (pandas, NumPy, matplotlib, seaborn, nltk, scikit-learn, Flask, Plotly, Tweepy), Heroku

Predicting flood events in Louisiana (Github)

- Used a severe storm events and historical weather data from NOAA to train a random forest model to predict flood events in Louisiana with forecast data
- Technologies used: Python (pandas, NumPy, pandarallel, seaborn, matplotlib, BeautifulSoup, requests, SciPy, scikit-learn), SQLite

Education

Data Science Career Track, Certification — Online

SPRINGBOARD — *June 2020*

- Intensive course in data science, data visualization, machine learning, hypothesis testing, Python, SQL, and Spark

Bachelor of Science in Biology — Lafayette, LA

UNIVERSITY OF LOUISIANA AT LAFAYETTE — *May 2018*

- Relevant coursework: Advanced Data Structures and Software Engineering, Linear Algebra, Calculus I and II