Skills

Technical

Machine Learning: Classification, Regression, Clustering, Feature engineering, Dimensional reduction techniques, Recommendation systems

Data Analytics: Cleaning, Manipulation, Scraping, Visualizations

Statistics: Inferential and Descriptive statistics, Bayesian methods, A/B testing

Software

Proficient: Python (NumPy, Pandas, Scikit-learn, Matplotlib, Plotly, Seaborn), SQL (PostgreSQL, MySQL), Git, Microsoft Office, MATLAB, LaTeX

Familiar with: C/C++, R, Win32, Linux

Work History

May 2018 - December 2019

Data Analytics & Market Research Intern

- Developed a character balancing model based on historical statistic performance and matchup characteristics, which increased the number of characters used in the 2018 season by 200%.
- Setup and managed a PostgreSQL database to store match statistics for extensive querying involving multiple joins between tables and schemas.
- Designed viewership feedback surveys that were completed by 50,000 users adhering to A/B testing design methodologies.
- Produced data visualization dashboard assets for the streaming platform in which viewership attained a 48% increase in max concurrent viewers.

Projects

Predicting Customer Churn in World of Warcraft **Spring 2020**

- Analyzed time-series user data during a one-year period applying exploratory data analysis techniques to identify pertinent features for predicting customer churn across a six-month period.
- Explored survival analysis techniques such as Kaplan-Meier estimator to gain insight into the relationship of features selected.
- Examined the performance of classification algorithms by applying Receiver Operating Characteristics (ROC) metrics with the highest scoring algorithm achieving a 96% ROC-AUC score.
- Deployed an application where users can input data for churn predictions and provides recommendations to reduce the risk of churning.

Backblaze: Modelling Hard Drive Failure **Spring 2020**

- Built a customized python-based data parser for data cleaning and manipulation on the 2019 Backblaze Hard Drive Reliability dataset.
- Investigated potential S.M.A.R.T Statistics (KPIs) documented by hard drive manufacturers for correlations to hard drive failure.
- Developed models for hard drive failure by employing classification algorithms, which led to a 3% increase in ROC-AUC score from Backblaze's baseline model.

PlayerUnknown's Battlegrounds Cheater Detection Tool **Winter 2019**

- Studied user match statistics scraped from PUBG.me to label features applicable to customer segmentation.
- Employed clustering algorithms and dimensional reduction techniques to flag users receiving the aid of illicit third-party programs.
- Deployed an application where users can input data to detect illicit users and inspect various 3-D cluster visualizations.

Education

May 2018

Masters of Science, **Electrical Engineering** Cumulative GPA: **3.80/4.00**

Relevant Coursework:

· Detection and Information Theory · Distributed Systems and Networks · Linear Systems · Machine Learning · Probability and Stochastic Processes

May 2016

Bachelors of Science, **Materials Science and Engineering** Cumulative GPA: **3.72/4.00**