

PRESTON LAM

DATA SCIENTIST

✉ prestonlam53@gmail.com ☎ 16503939981 [in linkedin.com/preston-lam](https://www.linkedin.com/in/preston-lam) 🌐 prestonlam53

SKILLS

LANGUAGES: Python, SQL, PySpark, Git, Excel

LIBRARIES: Pandas, Matplotlib, Sklearn, BeautifulSoup, Selenium

MACHINE LEARNING TECHNIQUES: Random Forest, Naive Bayes, K-Means Clustering, TensorFlow, PyTorch, Dimensionality Reduction, XGBoost

EMPLOYMENT

Metis Data Science Bootcamp, Data Scientist, Chicago, IL June 2020 - Sept. 2020

Metis is a 12-week immersive data science program that builds skills in data analysis and machine learning through project development. Completed five-end-to end projects listed in project section.

Walmart ECommerce, Category Management Specialist, San Bruno, CA Sept. 2018 - Feb. 2020

Category Management Specialist September 2018 – August 2019

Manage P&L for 3 Hardware Categories – portfolio of \$8M – across Walmart.com, Jet.com, and Hayneedle.com.

- Grew overall category revenue by 30% and profit by 6%
- Negotiated \$400k+ in Marketing Funds from vendors
- Assortment Expansion Lead for Home Improvement Dept. – onboard 30+ key brands for the team
- Marketing Lead for Home Improvement Dept. – put together National Safety Month campaign for Walmart.com

Aramark, Fremont, CA May 2017 - Aug. 2017

• Developed process for ordering merchandising equipment that would cut costs by over \$1,000 per order – presented solution to GM, and branch is currently implementing this process

PROJECTS

Kaggle Competition - House Prices: Advanced Regression Techniques Apr. 2020 - Apr. 2020

• Placed in the top 40% using a stacked model of 4 components: Lasso, Ridge, XGBoost, and Random Forest

Predicting Stock Prices with Fundamental Ratios (Regression) June 2020 - June 2020

- Scraped financial data and valuation metrics from Morningstar and Yahoo Finance with Beautiful Soup
- Created a random forest regression model that predicts stock prices based on fundamental financial ratios

Predicting Loan Defaults (Classification) July 2020 - July 2020

- Developed a random forest classification model that predicted loan defaults with 84% accuracy and 81% recall
- Engineered 10 features using SQL that improved baseline model performance by 400%
- Managed a cloud-based database using SQL and AWS

Creating a Job Recommender (Natural Language Processing) Aug. 2020 - Aug. 2020

- Used TF-IDF vectorizer and Truncated SVD to topic model 150 job listings scraped from Glassdoor via Selenium
- Developed a web app that returns job recommendation, clustering charts, and keywords based on one's LinkedIn

Building a Forecasting Engine for 311 Service Calls (Time Series Analysis) Aug. 2020 - Sept. 2020

- Performed EDA on Chicago 311 Data - found that service times are 4 weeks avg and number of requests is highly correlated with completion time
- Created a tool using Facebook Prophet allowing Chicago to forecast service calls by neighborhood and department

EDUCATION

Washington University in St. Louis Aug. 2014 - May 2018

BA Psychology 2018

BS Marketing 2018