

EDUCATION

University of Pennsylvania
MA Mathematics

California State University, Northridge
BS Applied Mathematics
Cum Laude

SKILLS

PYTHON

Pandas
Numpy
BeautifulSoup
Selenium
Scikit-learn
NLTK

MACHINE LEARNING

Linear Regression
Classification
Natural Language Processing
Computer Vision
Cross Validation
Model Selection

DATABASES

PostgreSQL
MySQL

VISUALIZATION

Matplotlib
Seaborn
Tableau
Streamlit

EXPERIENCE

Data Scientist

Sept 2020 – Dec 2020

Metis

Completed Metis's 12-week accredited data science bootcamp focused on project-oriented skill-building in data wrangling, data visualization, statistical modeling, machine learning, problem solving, and communication of deliverables. Designed, implemented, and presented the following projects:

Classify Muay Thai

- Leveraged cloud computing, and YOLOv3 with darknet framework to train a convolutional neural network to classify images and videos of Muay Thai strikes.
- Labeled hundreds of images of strikes with labelling to use for model training.

Math Unsupervised

- Used natural language processing to explore topics and trends in the way people talk about math.
- Scraped 9,000 tweets with the keyword "math" using the Twitter API.
- Modeled topics using Non-Negative Matrix Factorization (NMF).
- Analyzed sentiment of each topic using VADER.

Predicting Contraception Usage

- Utilized supervised machine learning methods to explore important features contributing to whether or not women were using long term contraception.
- Compared performances of multiple classification models including random forest, gradient boosted, logistic regression, and KNN.
- Built a Streamlit app that makes predictions in real time based on user input.

Predicting the Lifespan of a TV Show

- Built a linear regression model that predicts the number of seasons a show will run for with a MAE of ~1.6 seasons.
- Scraped data on 1,000 tv shows from RatinGraph.com and IMDb.com using Selenium.

PM Lead Intern

June 2020 – July 2020

STEM-Away – Machine Learning Internship

- Developed project timeline including milestones, coordination of team meetings over 4+ different time zones, and delegation of tasks to build a recommendation model for open-source internet forum Discourse.
- Leveraged Python libraries Selenium and Pandas to scrape, store, and clean data.

Research Assistant

Dec 2016 – July 2017

Mathematical Research in Interdisciplinary Topics Group

- Read medical studies to guide development of a game theoretical model.
- Presented suggestions for further development of the model in meetings with peers and supervisor.

Research Fellow

Aug 2016 – April 2017

PUMP Undergraduate Research Group

- Utilized Python library Networkx to generate and statistically characterize scale-free, small-world, and random network models.
- Presented at the MAA Spring 2017 SoCal-Nev Section Meeting.