

JOSEF SEEMAYER

DATA SCIENTIST

EXPERIENCE

Metis

Data Scientist · Oct. 2020 to Mar. 2021

Enrolled in 6-week introductory course and then intensive, 12-week data science bootcamp. Independently built five end-to-end data science projects that cover techniques including: exploratory data analysis, regression/classification, natural language processing, and neural networks. Please see project section.

June Life, Inc.

San Francisco, CA

Associate · June 2019 to Apr. 2021

Developed and deployed over 100 cook-programs directly to users ranging from everyday staples to frozen food to sous vide. Worked across many cross-functional groups to deliver product enhancements and changes:

- Devised and oversaw significant web design changes that led to hitting sales goals in consecutive quarters
- Developed new cook modes with embedded software team that shipped
- Established secondary warranty partnership and built Marketing Analytics Dashboard
- Directly worked with retail partners to ensure sales success

DuPont

Palo Alto, CA

Liquid Formulations Intern · May 2017 to Aug. 2017

- Increased throughput of automated system for analyzing protein stability by 200% using Biomek & KNIME software
- Investigated scalable filter plates
- Led meeting presenting the improvements and new challenges facing the system to full scale manufacturing team

PROJECTS

Predicting NFL Games

Using an adjusted Elo rating system, predict the winner of an NFL game. Interpreting probabilities as confidence levels, we can construct a betting recommendation and simulate expected profits. An interactive Tableau dashboard was built to explore team, individual, and season-long insights.

Factors Contributing to an NFL Total

Built an interpretive linear model of NFL scoring with 93% accuracy, or to within 1 field goal of the actual score, capturing 85% of variance

Exploring Mac Miller's Discography

Used NLP tools analyzed Mac Miller's discography for lyrical richness, sentiment analysis, and topic modeling. Comparing different albums over time, his personal struggles with addiction and depression were clearly illustrated with both topic modeling and sentiment analysis.

Classifying NBA Totals as Over/Under

Used time series data scraped online and built several classification models to predict the result of an NBA game's Vegas total line. Explored model including: logistic regression, decision trees, random forests, XGBoost, as well as, ensemble models and achieved 51% accuracy.

NYC MTA Street Team Deployment Plan

Developed multiple street team deployment schedules outside of subway stations in order to maximize sign ups & donations to the annual Women Tech Women Yes gala.

EDUCATION

Villanova University

BS Chemical Engineering 2018

Minor in Biochemical Engineering

CONTACT

✉ jseemayer2@gmail.com

☎ (650) 454-5622

in jseemayer2

🌐 jseemayer2

SKILLS

PYTHON

Pandas

Numpy

Scikit-learn

Spark

Keras

TensorFlow2

SUPERVISED LEARNING

Regression

Classification

Tree-based Methods

Ensemble Models

Neural Networks

UNSUPERVISED LEARNING

Dimensionality Reduction

K Means

DBSCAN

NATURAL LANGUAGE PROCESSING

Topic Modeling (NMF/LDA)

Sentiment Analysis

Word2Vec

TF-IDF Vectorization

VISUALIZATIONS

Tableau

Matplotlib

Seaborn

ggplot

DATABASES

PostgreSQL

NoSQL

Hadoop MapReduce

MongoDB

OTHER

Git

Google Collab

BeautifulSoup 4

HTML

Microsoft Office