

SUMMARY

Data nerd who's always looking for new skills to learn. Skeptical of easy answers; loves analysis.

SKILLS

LANGUAGES

Python
R
SQL

MACHINE LEARNING

SKLearn
statsmodels
keras

DATABASES

PostgreSQL
pandas

VISUALIZATION

seaborn
Tableau
matplotlib.pyplot
folium
Streamlit

BIOINFORMATICS

BLAST
BioPython
ClustalW

EDUCATION

University of Washington •
Seattle, WA
PhD Microbiology
Jan 2015

EXPERIENCE

Metis Data Science Bootcamp, Remote

Sept-Dec 2020

Data Scientist

12-week data science intensive program focused on project-oriented learning. The core curriculum is centered around Python, statistics, supervised and unsupervised machine learning, exploratory data analysis, databases, and visualization techniques. Completed four self-designed data science projects from conception to presentation including data collection, data management, exploratory data analysis, and visualizations. Project highlights below.

Everett Community College, Everett, WA

Sept 2015-Mar 2020

Associate Faculty - Biology

Iteratively improving courses every quarter and using evidence-based teaching.

- Designed, updated, and presented material on a wide range of biology topics to a diverse audience of students
- Write and iteratively improve assessments of student learning to target and re-focus class lectures and activities as needed
- Created and delivered modules on professionalism, critical thinking, and collaboration in a work setting

Shoreline Community College, Shoreline, WA

Sept-Dec 2019

Associate Faculty - Biology

"Flipping" microbiology course to cover faculty member's sabbatical.

- Transitioned from lecture-dominant to project-based course design "just in time" during the quarter
- Helped students learn both textual material and lab skills for microbiology

Projects

Metis: using linear regression to predict house prices from scraped data

- Scraped various house data from Redfin (>3k records over 3 year span)
- Generated house sale price predictor using scraped and cleaned data; predictor is most useful for predicting approximate ROI of additions/improvements

Metis: topic modeling COVID tweets to steer users away from misinformation

- Use Twitter API to download 100k sample tweets over the previous month
- Topic modeling using LDA followed by PCA of per-tweet topic vectors to cluster Tweets into related groups

Metis: training a neural net to match sweater photos to Ravelry patterns

- Download (API and scraping), clean, and categorize sweater photos based on domain knowledge (4k photos usable)
- Use transfer learning to classify photos into one of 22 categories
- Cluster photos based on raw score and recommend "close" patterns when a user uploads a photo of a sweater they want

Independent: effect of local changes on 911 calls

2019-2020

- Built basic data cleaning pipeline for data from public SPD police calls database
- Wrote scripts to generate plots of how calls changed before and after a given event, including choropleth of city to show change by area