Miriam Ali

Data Scientist

SF Bay Area • 408.636.6675 • miriamaali@outlook.com • Github • LinkedIn • Website

Python SQL Docker FastAPI Git/hub AWS

Projects

Incident Tracker Dashboard, *Human Rights First,* Project Link, current
- Built the data science backend, from data mining and API creation to data visualization, for a website tracking incidents of police use of force as interactive map of localized data
- Sourced data from numerous news outlets, organizations, and social media posts to create a relational Postgres database and API of incidents and evidence using Python, hosted in AWS

Music Recommendation Engine, Project Link, May 2020
- Collaborated in a team of 8 members, as the sole machine learning engineer
- Sourced and processed audio feature data on over 120k unique songs to train a predictive k-nearest neighbors model, which was pickled and deployed with Flask, improving upon previous models' accuracy by over 30% and overall response time by 50%

MediCabinet, Project Link, April 2020
- Deployed a Flask app and developed a RESTful API to support the project database and NLP predictive algorithm pipeline as the data engineer for a medical marijuana strain recommendation application

Education

Data Science Fellow, Lambda School, 2020
Program Description: Lambda School is a 9+ month full-time CS & Software Engineering academy that provides an immersive hands-on curriculum
Coursework Includes: Computer Science, Descriptive & Predictive Statistics, Machine Learning, Data Engineering
- Gained applied experience with machine learning; able to determine the most appropriate methods, implement, and tune NLP models and neural networks.
- Served as a data scientist in cross functional teams on multiple ready-to-deploy projects

B.S. in Biology, University of Michigan, 2015
Honors: Dean's List (2013-2015)
Coursework Includes: Stats, Calc I/II, Linear Algebra, Ethics, graduate-level research & seminars

Technical Skills

Programming Expertise: Python (Pandas, NumPy, Scikit-learn, Matplotlib), SQL, MATLAB
Frameworks: TensorFlow, Keras, NLTK, FastAPI, Flask, Plotly Dash
Tools: AWS, Docker, Jupyter, PostgreSQL, MongoDB, MySQL
Skills: Data sourcing and cleaning, Data analysis, Linear and multivariate regressions, Predictive analytics, Data visualization, Machine learning, Natural language processing, Neural Networks, Git and version control, Linux and Windows OS