Matthew Kang

Aspiring data engineer, passionate about creating robust data pipelines for an information driven world.

Berkeley, CA | 925-216-6150 | mjk095@ucsd.edu | linkedin.com/in/matthewjkang

EXPERIENCE

Coding Instructor

2021 - Present

Hackingtons I Pleasant Hill, CA

- Introduce children to coding through a visual programming language developed by MIT called Scratch.
- Teach static web development (HTML/CSS/JavaScript) to students through project-based learning.
- Teach web app development to students through Python (Django / Flask).

PROJECTS

Big Data Analysis of Chicago Taxi Cab Driver Data

In this project, my teammates and I analyzed the Chicago Taxi cab data set. This data set contains 25 million entries. (~ 30GB CSV file)

Machine Learning Algorithms Performance Testing

In this project I tested the performance of K-NN, SVM, and Logistic Regression on three different data sets. (9 trials total). Hyperparameter tuning was used to find the optimal configuration for K-fold cross validation.

ETL Pipeline #1

• A project ETL pipeline that I created. It intakes messy CSV or XYZ files, and outputs a clean CSV with extracted metadata and null values removed. The program is run via CLI.

All projects can be viewed at github.com/matthewjkang

EDUCATION

University of California San Diego | Bachelor of Science, Machine Learning

2017 - 2021

Eleanor Roosevelt Honors Program, Design Club, Triton Barbell

SKILLS

- Python, Pandas, Pytorch, Matplotlib, NumPy, Scikit-Learn, Django, Flask
- SQL, PostgreSQL, PgAdmin, MYSQL, SQLite
- HTML, CSS, JavaScript, Scratch
- Jupyter, GitHub, Visual Studio Code, Figma, Adobe Lightroom, Adobe Photoshop, Excel

CERTIFICATIONS

- Python3 Codecademy
- SQL Codecademy

INTERESTS

Weightlifting, Bay Area restaurants (Ask about my spreadsheet of local eats), Golden State Warriors, photography, farmers markets, repairing my 20-year-old Camry, and spending time with my girlfriend.

Personal website: mattkang.com