Anthony Ko

github.com/springtea12 linkedin.com/in/koanthony12 koanthony12@gmail.com

EDUCATION

University of California, Los Angeles — Expected June 2025
Bachelor of Science in Statistics and Data Science

Pasadena City College — June 2021 - May 2023

RELEVANT COURSEWORK

Courses: Computation and Optimization for Statistics, Statistical Models and Data Mining, Data Management, Data Analysis and Regression, Computational Statistics with R, Linear Models, Design and Analysis of Experiment, Monte Carlo Methods, and Practice of Statistical Consulting, Intro to C++, Intro to Python, Object Oriented Programming, Data Structures and Algorithms

SKILLS

Languages: Python, SQL, R, C++, HTML/CSS, LaTeX
Databases: MySQL, PostgreSQL
Libraries: NumPy, Pandas, MatPlotLib, Seaborn, Tidyverse, Dplyr
Certificates: MIT IDSS Data Science and Machine Learning: Making Data-Driven Decisions

WORK EXPERIENCE

PV Circonomy | *Machine Learning Engineer Intern* — June 2024 – August 2024
- Trained a YOLOv7 model to accurately detect junction box dimensions for removal
- Employed A/B testing to evaluate the performance of training weights to increase model accuracy
- Created and annotated labels for items in images for object detection and classification
- Performed image augmentation using Albumentations to increase training and testing data

PROJECTS

Invoice PDF Scraper | *R, Excel* — July 2024
- Programmed an R script that scrapes contents of a semi-structured invoice and transfers contents onto an Excel sheet
- Utilized regular expressions (regex) to scrape targeted data such as name, items, total, and date

Predicting Potential Customers | *Python, NumPy, MatPlotLib, Scikit-Learn* — Aug 2023
- Built a Decision Tree and Random Forest model with Scikit-Learn to predict potential customers
- Conducted performance checks to ensure the model wasn't overfitting
- Performed analyses on the importance of features based on the two models to determine insights

FoodHub Order Analysis | *Python, NumPy, Seaborn, MatPlotLib* — July 2023
- Performed Exploratory Data Analysis on a FoodHub data set to determine correlations and characteristics.
- Utilized Seaborn and MatPlotLib to visualize food preparation time, cost of orders, and delivery times.
- Mutated the set for restaurants fulfilling rating count and score for promotional offers.

HACKATHONS

ASA DataFest | *R, Tableau* — April 2024
- Conducted data analysis on a dataset with over 5 million entries on interactions with eLearning for Statistics
- Designed dashboards using Tableau to present key findings and correlations within the dataset
- Developed a linear model to predict student chapter scores based on significant factors