DANNY SIU

dannysiu.me
(817) 320-7283
chat2danny@gmail.com

EDUCATION

University of California at Berkeley
Berkeley, CA
Aug. 2017 - May 2021
Bachelor of Arts in Data Science; Domain Emphasis: Linguistic Sciences
Minor in Computer Science

Coursework:

Principles and Techniques of Data Science	Concepts of Statistics
Foundations of Data Science	Discrete Math & Probability Theory
Efficient Algorithms & Intractable Problems	Intro to Artificial Intelligence
Designing Information Devices & Systems	Computer Architecture
Probability for Data Science	Data Structures

EXPERIENCE

University of California at Berkeley, Cohen Group - Research Intern
Berkeley, CA
Feb. 2019 - Present
○ Performing alignment of OMI (satellite) and WRF-CHEM model pixel data.
○ Implementing machine learning techniques for analysis of OMI NO2 using Python.
○ Enacting strategies for replacing an expensive WRF-CHEM run with ML model predictions by training on two years of model output and observations.

University of Texas at Austin - Summer Student Researcher
Austin, TX
Jun. - Aug. 2016
○ Optimized catalysts for efficient energy production through molecular modeling by utilizing VASP, Python libraries, modeling software, and Bash scripts.
○ Computationally simulated carbon monoxide binding on gold-palladium nanoparticles during oxidation-reduction reactions.

Jun. - Aug. 2015
○ Conducted neurological experiments on spiders, learned research lab techniques, and explored various scientific research studies with professors.

SKILLS

Programming: Python and libraries (Scikit-Learn, Matplotlib, Pandas, Seaborn, SciPy, NumPy), Java, SQL, Unix/Linux, Git, Bash, VASP
Data Science: Data analysis, Data visualization, Data cleaning, Data modeling, Feature engineering, Machine learning (ML), Artificial intelligence (AI), Exploratory data analysis (EDA)
Tools: Autodesk Inventor, Illustrator, Photoshop

PROJECTS

Python
(Pandas, Matplotlib, Seaborn, Scikit-Learn, NumPy)

Taxi Ride Duration Predictor
○ Utilized data visualizations in Matplotlib & Seaborn and conducted data manipulation & analysis in Pandas dataframes to assist in feature engineering.
○ Dramatically improved linear model predictions in Scikit-Learn by replacing outliers with mean values and implementing better feature selection.

Python
(Pandas, Matplotlib, Seaborn, Scikit-Learn, NumPy)

Spam Email Classifier
○ Performed exploratory data analysis (EDA) on text data to perform feature selection.
○ Implemented feature engineering for a logistic regression model and utilized carefully selected spam keywords to fit my spam classifier model.

Java

PokeMan
○ Developed a multi-player game in Java featuring randomized worlds.
○ Implemented saving of game versions using Git and a live heads-up display (HUD).