DANNY SIU

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

EDUCATION

Berkeley, CA
Aug. 2017 - May 2021

University of California at Berkeley
Bachelor of Arts in Data Science; Domain Emphasis in Linguistic Science
Minor in Computer Science

Coursework:

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

EXPERIENCE

Berkeley, CA
Feb. 2019 - Present

University of California at Berkeley, Cohen Group - Research Intern
- Implementing machine learning techniques in Python for analysis of NASA satellite measurements of NO2.
- Enacting strategies for replacing an expensive WRF-CHEM climate model run with ML model predictions by training on two years of model output and observations.

Austin, TX
Jun. - Aug. 2016

University of Texas at Austin - Summer Student Researcher
- Optimized catalysts for improved efficiency in energy production by utilizing VASP, Python libraries, modeling software, and Bash scripts.
- Computationally simulated carbon monoxide binding on alloys of gold-palladium nanoparticles during oxidation-reduction reactions.

Jun. - Aug. 2015
- Investigated the effects of brainwave entrainment on orb weaver spiders by microscopically analyzing their webs.

SKILLS

Programming: Python and its libraries (Scikit-Learn, Matplotlib, Pandas, Seaborn, SciPy, NumPy), Java, SQL, C, Unix/Linux, Git, HTML, CSS, 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 (Scikit-Learn, Matplotlib, Pandas, Seaborn, 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 (Scikit-Learn, Matplotlib, Pandas, Seaborn, 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).