Kelvin Nguyen

714-657-6269 | kelvinnguyen514@gmail.com | Github: kelvinnguyenn | Linkedin: linkedin.com/kelvinnguyenn

EDUCATION

University of California, San Diego **Expected Graduation:** June 2024
B. S., Data Science *La Jolla, CA*
GPA: 3.59

Relevant Coursework: Linear Algebra, Foundations of Data Science, Statistical Methods, Data Structures & Algorithms, Multivariable Calculus, Practice & Application of Data Science
Planned Coursework: Data Management, Systems for Scalable Analytics, Data Visualization, Recommender Systems

EXPERIENCE

Project Lead Jan 2022 - June 2022
Data Science Student Society (DS3) at UCSD *La Jolla, CA*

- Organized and facilitated weekly meetings consisting of progress updates, improvements to model, etc.
- Coordinated with Projects Committee mentor to determine best course of action in moving forward with each step of the project
- Developed structure & timeline of project from formation of hypotheses & dataset sourcing to completion of various ML models

PROJECTS

Internet Outage Analysis in SEA & MENA | *Python, sk-learn, Pandas, BS4* Jan 2022 - June 2022

- Sourced **100,000+** news articles about Southeast Asia & the Middle East regions via NYT API and Guardian API in Python
- Created regression/classification model with sk-learn predicting severity of an outage/ predicting an outage in a given country with an accuracy of **77%** & an f1-score of **9%**
- Developed interactive geographic dashboard that allows users to interact with available regions using Streamlit

House of Representatives Stock Action Classifier | *sk-learn, Pandas, matplotlib* May 2022 - June 2022

- Prepared dataset for analysis & usage in model with industry-best practices (data cleaning, imputation, etc.)
- Utilized cross-validation techniques to optimize model parameters to increase final accuracy by **11%**
- Explored relationships between political affiliation and stock trades by creating visualizations with univariate & bivariate features using matplotlib
- Trained classifier using a SVM that predicts whether a stock trade was a buy/sell with an accuracy of **65%**

San Francisco Parking Citations EDA | *Seaborn, geopandas, plot.ly* August 2022

- Cleaned & analyzed over **18 million rows** of data points to gather insights on citation trends in San Francisco, ranging from citation-level to location-based using pandas
- Utilized various statistical aggregate functions to determine the most common types of citations & the most common locations
- Created visualizations of citation trends that highlighted important features using Plot.ly and Seaborn

SKILLS

- Coding Languages: Python, SQL, Java, HTML, CSS, Javascript, R, Matlab
- Computer Programs/Software: Git, Anaconda, NumPy, pandas, scikit-learn, matplotlib, Plot.ly
- Languages: English (native), Vietnamese (elementary)