

Allen Chen

Mobile: (754)-715-6018 | alchen518@gmail.com

Education

Nova High School

Davie, Florida | GPA: Weighted: 5.263 / Unweighted: 4.0

August 2019 - June 2023

AP Coursework: Statistics, Calculus, Computer Science, Environmental Science, Human Geography, Microeconomics

University of California, Los Angeles (UCLA)

Bachelor of Science in Statistics and Data Science

Expected Graduation: September 2023 - June 2026

Los Angeles, California | GPA: 3.89

Coursework: Machine Learning, Linear Models, Data Analysis, GIS, R/C++ Programming, Optimization

Experience

Data Analytics Intern | Supermap GIS

July 2024 - September 2024

- Deployed and evaluated the performance of binary classification models to differentiate residential buildings from bodies of water and vegetation in satellite imagery with 90% accuracy
- Developed an LSTM deep learning model using Docker on JupyterLab to determine which factors more significantly impact Supermap stock

Project Lead | DataRes at UCLA

October 2024 - Present

- Led a team of 5 members to create a random forest model using scikit-learn to determine the most influential feature variables in predicting art prices at auction. Presented results to 150+ members
- Created data visualizations to showcase the volume traded of the top 5 tech stocks and analyzed the factors that influenced the trends in the graph. Presented results to 150+ members

Project Lead | National Student Data Corps (NSDC)

October 2023 - Present

- Collaborated with a group to perform data collection, smoothing, and visualization with Python to implement a decision tree classifier to detect fraudulent IP addresses based on GPS coordinates
- Improved performance of the model to an average 95% accuracy rate for a dataset containing 1.6 million observations

Co-Founder | Power Scholars Online Tutoring

September 2020 - November 2022

- Oversaw channels to ensure smooth communication among the team, interviewed new potential tutor applicants, and led weekly meetings discussing long-term progress/goals with tutors
- Worked to grow the organization from a starting group of 10 students to over 300
- Cooperated with Hack+ to obtain 501(c)(3) nonprofit status and initial grant funding for marketing
- Created structured study plans and leadership positions, increasing tutor productivity by 15%

Projects

Reddit Post Popularity Analysis

January 2025

- Devised a strategy to efficiently optimize Reddit postings to garner more upvotes using Vader sentiment analysis with 87% accuracy, and NLP to perform text mining to identify the most frequent words in posts

Used Car Price Analysis

June 2024

- Developed a multiple linear regression model with R on a dataset of used cars from 10 car brands from India to determine which factors are more significant in determining used car prices

House Price Prediction

June 2024

- Trained and deployed an XGBoost model with scikit-learn to predict housing prices, with a low mean absolute error of approximately \$200,000
- Used matplotlib to visualize the geographic distribution and clustering of houses by price with color coding

Skills

Languages: Python (Numpy, Pandas, matplotlib, scikit-learn), R (tidyverse), SQL, Java, C++

Tools: Docker, Tableau, Microsoft Office Specialist (Word, PowerPoint, Excel), QGIS, ArcGIS