Aidan Nguyen

(408) 345-1470 | aidannguyen27@ucla.edu | linkedin.com/in/nguyenaidan | aidannguyen.netlify.app

EDUCATION

University of California, Los Angeles (UCLA) Expected June 2026
B.S. Statistics and Data Science, Minor in Data Engineering GPA: 3.97
Activities: Statistics Club (Social Chair), Data Resolutions (Consulting Project Lead)

EXPERIENCE

Improper Etiquette September 2024 - Present
Data Analyst Intern
- Modeled PETE cost structure using market research, public data, & regression analysis for 50% profit margin targets
- Analyzed 15+ LA-specific production variables to identify break-even points and optimize sales targets for hats
- Created data visualizations and detailed Excel reports with 10+ research papers sourced to support findings
- Presented regression-based insights on cost modeling and profit optimization to guide strategic client decisions

Signaling Systems Lab January 2024 – Present
Machine Learning Research Assistant
- Secured grant funding by creating visual cluster t-SNE plot that verified data from ongoing research initiatives
- Analyzed Single Cell RNA sequencing data (~6M rows) with R, caret, & Seurat to classify bone marrow stroma cells
- Deployed supervised learning models, including SVMs and coefficient classifiers, to evaluate data quality & scalability
- Fine-tuned SVM hyperparameters with R and e1071, improving model accuracy by 10% through cross-validation

Dune Road Capital June 2024 – September 2024
Data Science Intern
- Enhanced financial analysis using SQL queries for 10M+ rows of exchange-traded fund (ETF) performance metrics
- Predicted future 10+ ETF returns through Python time-series analysis, aiding portfolio strategy recommendations
- Developed data visualizations with Tableau to present key ETF metrics such as expense ratio and asset allocation

PROJECTS

DataFest at UCLA | *R, dplyr, ggplot2, caret, tidyverse*
- Achieved finalist out of 400+ participants in a 48-hour data hackathon, focused on analyzing complex datasets
- Analyzed 7 spreadsheets with 100K+ rows of user and site interaction data to improve a student textbook platform
- Identified key patterns in 1000+ users' behavior utilizing R and tidyverse to clean, transform, and merge datasets
- Presented visualizations recommending data-driven improvements for greater engagement & educational outcomes

Book Recommendation Web Application | *Python, numpy, pandas, tensorflow, scikit-learn, matplotlib*
- Developed a web app connecting 100+ readers, fostering community, & promoting youth engagement with literature
- Built a custom LLM-powered recommendation engine that increased personalized book suggestions by 75%
- Managed user data for 10,000+ libraries, interactions, and connections with MongoDB for 99.9% uptime
- Enhanced backend with RESTful API integration, boosting content delivery speed by 25%

YouTube Success Metrics Analysis | *PyTorch, transformers, pandas, numpy, googleapiclient, nltk, matplotlib*
- Analyzed 600,000+ YouTube comments using a fine-tuned DistilBERT model for multilingual sentiment analysis
- Created a gradient boosting regressor and predictive models, to assess how 1000+ titles & thumbnails impact views
- Implemented OpenCV to implement face detection in video thumbnails, correlating findings with viewership metrics

Programming Languages: Python (pandas, numpy, scikit-learn, plotly, flask, tensorflow), Java, C++, JavaScript, SQL, R
Tools: Git, GitHub, Tableau, AWS, Excel, MongoDB, Power BI
Interests: Volleyball, Hiking, Catan, Poker, Graphic Design