

Allison Lynn

Los Angeles, CA | allienlynn@gmail.com | 469-401-1939 | Portfolio | LinkedIn | GitHub

Education

BS in Statistics and Data Science, The University of California Los Angeles Sept 2022 – June 2026

- Cumulative GPA: 3.77/4.00 | Dean's Honors List: Winter 2023, Fall 2024
- **Coursework:** Data Analysis and Regression, Probability Theory, Design and Analysis of Experiments, Mathematical Statistics, Calculus of Several Variables, Linear Algebra, Programming in R, Programming in C++, Programming in Python, Statistical Models and Data Mining, Linear Models, Computational Statistics with R

Experience

Data Strategy Intern, Ryan, LLC – Los Angeles, CA June 2024 – Aug 2024

- Worked with machine learning, deep learning, and NLP pipelines using PyTorch, TensorFlow, and BERT.
- Developed Python scripts to automate and optimize large-scale data analysis workflows.
- Wrote efficient SQL queries to manage Databricks databases and supported strategy for data lake storage implementation.
- Designed dashboards in Tableau and Power BI to visualize storage patterns and data inflow into the data lake.

Calculus Tutor, Unlimited Tutoring – San Diego, CA April 2023 – June 2024

- Conducted tutoring sessions for high school and university calculus courses.
- Leveraged extensive mathematics knowledge and strong communication skills to provide one-on-one tutoring.

Projects

Modeling Urban Flight in U.S. Commercial Real Estate Real-Estate-Analysis

- Presented at ASA DataFest 2025 at UCLA; awarded the Judges' Choice Award out of 50+ participating teams.
- Analyzed a proprietary dataset from Savills to uncover trends in commercial lease migration from central business districts to suburban markets over the past seven years.
- Engineered features to build a logistic regression model that classified lease location type with **90%** accuracy.

Predicting ICU Length of Stay ICU-Stay-Prediction

- Applied Natural Language Processing (NLP) to predict ICU length of stay, utilizing BioBERT embeddings and training Random Forest models to capture complex patterns, and applied LASSO to reduce overfitting.

NBA Analysis: Predicting Game Outcomes and Player Salaries NBA-Analysis

- Developed a classification model in Python to predict NBA game outcomes, achieving **93.87%** accuracy through feature engineering (rolling averages, weighted stats) and feature selection with Linear SVC.
- Trained Random Forest and Gradient Boosting models on engineered and selected features.
- Applied linear regression in R to model and predict NBA player salaries, identifying key performance metrics.

Predicting Heart Failure with LDA Heart-Failure

- Applied Linear Discriminant Analysis (LDA) with **83.33%** accuracy, validating model assumptions, and demonstrating a primarily linear relationship between significant predictors and heart failure.
- Utilized LASSO regression for feature selection, simplifying the model while maintaining predictive accuracy.

Skills

Statistical Modeling & Machine Learning: Logistic Regression, Linear Regression, LDA, QDA, Random Forest, Gradient Boosting, Lasso, Ridge, Linear SVC, Kernel SVM, KNN, PCA, Neural Networks

Data Analysis & Visualization: EDA, Experimental Design, Data Mining, Data Visualization (Tableau, Power BI)

Programming & Tools: Python (pandas, matplotlib, PyTorch, scikit-learn), R, SQL, Git, Excel

Mathematics: Calculus, Linear Algebra, Probability Theory, Optimization

Data Engineering: ETL, Databricks