

## EDUCATION

---

**University of California San Diego**  
*Bachelor of Science Data Science*

La Jolla  
2026

## EXPERIENCE

---

### **Student Intern - Innovating for National Security (i4NS)**

Sep 2024 – Dec 2024

*United States Department of Defense (DoD)*

*La Jolla, CA*

- Contributed to the development of a Venture Capital (VC) Dashboard to assist the Department of Defense (DoD) in identifying mature and emerging service-related technologies within investor portfolios.
- Applied the Lean LaunchPad methodology to help teams rapidly address complex challenges, iterate on solutions, and deliver customer-tailored minimum viable products.
- Fostered connections between military technology innovations and defense-focused venture capitalists, as well as entrepreneurs, to promote civilian-military technological advancements.
- Streamlined the transition of internal RD projects into market-ready solutions, enhancing operational efficiency and innovation within the DoD.

## PROJECTS

---

### **Language Model Development | May 2024**

- Developed a robust language modeling system consisting of Unigram and N-Gram language models in Python. This project involved constructing probabilistic models to analyze and generate sequences of tokens (words) based on their occurrences and contextual dependencies.
- Developed methods to compute the probability of a sequence of words and to generate random samples from the trained model.
- Created a recursive structure that allows the model to fall back to lower-order models (e.g., bigrams, unigrams) when needed, Employed advanced pandas operations to compute conditional probabilities of N-grams given the preceding (N-1)-grams.

### **Predicting Chronic Kidney Disease ML Model | July 2024**

[Link](#)

- Conducted comprehensive EDA to understand data distributions, correlations, and relationships.
- Addressed missing values by identifying their mechanism and applying appropriate imputation techniques.
- Created visualizations such as confusion matrices and classification reports to communicate model performance.
- Explored and engineered features to enhance model performance. Identified and combined relevant features to improve predictive accuracy.
- Employed hyper parameter optimization and rigorous cross-validation methodologies to refine model accuracy and robustness

### **Housing Evaluation Across the U.S | March 2025**

[Link](#)

- Designed and deployed a full-stack house valuation web application integrating a LightGBM-based regression model to predict property prices across the U.S. using over 2M listings and enriched geographic datasets.
- Performed robust data engineering and feature engineering including zip-level imputation, crime stats integration, proximity analysis, and home category classification using geodemographic segmentation.
- Built an interactive UI with React and Material-UI, featuring Google Maps integration, valuation scoring, and amenity-based neighborhood insights. Deployed using Vercel and Railway with scalable architecture.
- Outperformed baseline models (OLS, Ridge, Random Forest) with a LightGBM model achieving better RMSE, showcasing optimization via RandomSearchCV and effective use of engineered features.

## TECHNICAL SKILLS

---

**Languages:** Python, SQL, R, Java, JavaScript, HTML/CSS

**Frameworks:** React, Node.js, Django

**Developer Tools:** Git, Visual Studio, PyCharm, IntelliJ, Microsoft, Tableau, STATA, SQLite, PostgreSQL

**Libraries:** pandas, NumPy, Matplotlib, seaborn, Pytorch, sklearn