

# Ty Alba

[Personal Portfolio](#) | 951-526-7249 | Ty Alba | [GitHub](#) | [LinkedIn](#)

## EDUCATION

### **University of California, San Diego - B.S. in Data Science, minor in Mathematics**

September 2023 – Present

- 3.79 GPA through two years
- Relevant courses: Modeling and ML, Representation Learning, Data Management, Data Visualization, Probability & Statistics, Data Structures & Algorithms

## EXPERIENCE

### **UCSD NCAA D1 Baseball Team, La Jolla, CA - Data Analyst**

October 2025 – Present

- Collect, clean, analyze, visualize data from Trackman
- Create models and scouting reports for coaches and players as needed

### **Halıcıoğlu Data Science Institute, La Jolla, CA - Tutor**

September 2024 – Present

- Tutor for DSC 20, Programming and Data Structures for Data Science
- Write tests, grade exams, assist students with deeper understanding in office hours and online Q&A

### **Data Science Alliance, San Diego, CA - Data Science Intern**

July 2024 – November 2024

- Collaborate with coworkers and work individually on projects that help the community of San Diego County
- Research and clean large datasets, create detailed customized interactive maps in Tableau and using GeoJSON files

## PROJECTS

### **MLB Automatic Ball-Strike (ABS) Helper - Individual Project**

- Built and validated XGBoost models to estimate challenge success probabilities from pitch, game-state, and umpire data
- Feature engineered 100,000 Statcast pitches to maximize win expectancy, visualized model insights through heatmaps and an interactive application to inform strategic decision-making for MLB teams

### **SMT Data Challenge - Individual Project**

- 1 of 4 finalists out of 114 contestants
- Cleaned and manipulated raw player-tracking data to derive crucial game-state information
- Built and validated XGBoost model to estimate baserunning advancement success probabilities on infield ground balls to grade teams' and players' quick decision making skills on the basepaths

### **League of Legends Data Analysis and Classification - Partner Project**

- Performed data cleaning, hypothesis testing, and EDA
- Built and validated Random Forest Classification model to predict position type to achieve 94% test accuracy

## REFERENCES

### **Marina Langlois - DSC 20 Professor**

malanglois@ucsd.edu

### **Dr. Meredith Wills - SMT-U Lead & Advisor**

m.wills@smt.com, (720) 289-0228