Pratham Aggarwal

San Diego, CA | 619-953-7147 | praggarwal@ucsd.edu

pratham-aggr.github.io | linkedin.com/in/pratham-agg | github.com/pratham-aggr

EDUCATION

Bachelor of Science, Data Science — Expected May 2027

Halıcıoğlu Data Science Institute, University of California San Diego

Relevant Coursework: Object-Oriented Design, Algorithm Design, Data Structures, Problem Solving and Complexity Analysis

SKILLS

Languages: Python, Java, C++, JavaScript, SQL, MATLAB, HTML, CSS3

Frameworks & Libraries: React, Node.js, Flask, Django, FastAPI, Pandas, Scikit-Learn, TensorFlow, PyTorch

Tools & Platforms: Git, Docker, Linux/Unix, Bash/Shell, Google Cloud Platform, Hugging Face, Terminal

Testing & Development: JUnit, REST APIs, Unit Testing, Documentation

EXPERIENCE

Quantitative Trading Intern — Mar 2025 – Jun 2025

Student Foundation Investment Committee, Quantitative Technologies — *San Diego, CA*

- Achieved 12% improvement in trading returns by building a Deep Reinforcement Learning agent for a $1.3M student-run investment fund, framing American options trading as an optimal stopping problem with adaptive policy learning and optimized early exercise strategies.
- Reduced Mean Squared Error to 0.96 by developing an LSTM-based stock price prediction model trained on 5 years of historical data and fine-tuned hyperparameters, improving forecast accuracy.
- Increased trading realism and decision quality by 18% through modeling a sequential decision-trading process that dynamically selects option type (call/put), expiration date, and exercise timing.

Data Science Research Analyst — Jun 2025 – Present

Climate Analytics Lab, Scripps Institution of Oceanography — *San Diego, CA*

- Preprocessed 243+ observational and projectional climate datasets (e.g., CMIP6) by converting Zarr files into ndpyramid format and hosting them on Google Cloud, enabling rigorous testing of machine learning models' physical plausibility.
- Devised over 20 custom evaluation metrics and developed an ML pipeline to benchmark 50+ model outputs, enhancing transparency and enabling clear identification of black-box model limitations.
- Awarded a $4,500 research scholarship under PhD climate scientist Duncan Watson-Parris; showcased work through a React/Node.js web application that visualized climate model evaluations, improving stakeholder understanding.

Program Manager & Consultant — Mar 2025 – Jun 2025

Solana Center & Data Science Student Society (DS3) Consulting — *San Diego, CA*

- Increased waste diversion rates by 15% by leading a team of 5 to analyze 2,000+ composting data records for San Diego startup Solana Center.
- Improved data accuracy by 25% and enabled reliable, scalable analysis by cleaning and transforming large-scale environmental datasets using Python.
- Boosted stakeholder satisfaction and ensured on-time delivery of project milestones by implementing structured workflows, effective tools, and adaptable project management practices.

PROJECTS

HackFrontier Winner: Geospatial ML & CV tool for Homeless Services — *Learn more*

- Won a hackathon as the youngest among 100+ professionals by developing a ML forecasting system with 67% accuracy, leveraging 35+ transit, demographic, and geographic features to strategically place homeless service centers in San Diego.
- Deployed a real-time computer vision system using Oxen.ai and EyePop.ai to address the challenge of tracking a transient homeless population, providing live monitoring and precise demand insights to enable accurate, data-driven responses.

Simulating Black Hole Evolution: Comparative Analysis of Light and Heavy Seeds — *Learn more*

- Demonstrated that seed mass impacts black hole growth rates by up to 40% through research visualizations developed over 20 weeks of PhD-supervised analysis using 10+ visualization techniques.
- Simulated growth trajectories for over 100 black hole seed scenarios using Eddington accretion models, analyzing evolutionary differences between light and heavy seeds over 20 million years of data.

Predictive Modeling of Building Energy Loads — *Learn more*

- Achieved 91% accuracy in predicting building heating and cooling loads by implementing a multiple linear regression model using 8 key architectural and environmental features on 750+ samples.
- Improved energy efficiency strategies and reduced operational costs by 15% by applying k-means clustering to identify distinct consumption patterns and a 25% behavioral shift in load forecasting.