

Pratham Aggarwal

San Diego, CA | 619-953-7147 | pragarwal@ucsd.edu
pratham-aggr.github.io | linkedin.com/in/pratham-aggr | github.com/pratham-aggr

EDUCATION

Bachelor of Science in Data Science, Minor in Probability & Statistics Expected Dec 2026
University of California, San Diego (School of Computing, Information and Data Sciences) GPA 3.6
Relevant Coursework: Machine Learning, Data Structures & Algorithms, Object-Oriented Programming, Databases
Probability & Stochastic Processes, Monte Carlo Simulation, Time Series Analysis, Linear Algebra, Optimization, A/B testing

SKILLS

Programming Languages: Python, SQL, C++, MATLAB, Java, JavaScript, HTML/CSS
Data Analysis & Visualization: Excel, ETL/Data Pipelines, Tableau, Matplotlib, Seaborn
Frameworks & Libraries: Pandas, NumPy, Scikit-Learn, PyTorch, TensorFlow, React, Node.js, FastAPI
Tools & Platforms: Google Cloud Platform (GCP), Docker, Git/GitHub, Linux/Unix, Bash/Shell, CI/CD

EXPERIENCE

Quantitative Trader Intern Mar 2025 – Jun 2025
SFIC Quantitative Technologies San Diego, CA

- Increased options trading returns by 12% for a \$1.3M student-run investment fund by developing a Deep Reinforcement Learning agent with adaptive policy learning and optimized early exercise strategies.
- Improved trading decision quality by 18% by modeling a sequential decision-making framework that dynamically selects option type, expiration, and exercise timing from historical time-series patterns.
- Reduced prediction error to 0.96 MSE using an LSTM-based stock price forecasting model trained on 5 years of historical market data with extensive hyperparameter tuning.

Data Science Researcher Jun 2025 – Present
Climate Analytics Lab, Scripps Institution of Oceanography San Diego, CA

- Awarded a \$4,500 research scholarship under PhD Duncan Watson-Parris; processed 243+ observational and projected climate datasets (e.g., CMIP6) on Google Cloud to assess machine learning models for physical plausibility.
- Developed 20+ custom evaluation metrics and an ML-based ETL pipeline to benchmark 50+ climate model outputs, creating interactive time series visualizations that reveal model limitations and enhance interpretability.
- Designed an interactive React/Node.js web application to visualize model benchmarks and trends, enhancing stakeholder understanding and accessibility of climate insights.

Program Manager & Consultant Mar 2025 – Jun 2025
Solana Center San Diego, CA

- Evaluated program success by analyzing 2,000+ composting records, revealing a 6.7% increase in waste diversion.
- Improved environmental data accuracy by 9% by imputing missing values, validating entries, and performing statistical analysis on large-scale datasets to enable precise, scalable insights.
- Improved stakeholder decision-making and ensured on-time project delivery by creating structured, data-informed workflows and presenting actionable insights clearly.

PROJECTS

QuantFlow: Full-Stack AI-Powered Portfolio Risk Management Platform Aug 2025

- Built a production-grade platform processing real-time data for 50+ stocks through an intelligent data pipeline with caching, reducing API costs by 80% while delivering sub-second portfolio valuations.
- Developed an AI-driven risk engine leveraging Monte Carlo simulations and volatility modeling, achieving 92% predictive accuracy and providing actionable portfolio rebalancing insights that outperformed benchmarks by 15%.

HackFrontier Winner: Geospatial ML & CV Tool for Homeless Services Jun 2025

- Developed an ML forecasting system achieving 67% accuracy, optimizing placement of homeless service centers in San Diego using 35+ transit, demographic, and geographic features.
- Delivered live monitoring and demand insights for a transient homeless population by deploying a real-time computer vision system with Oxen.ai and EyePop.ai, enabling data-driven service allocation.

Predictive Modeling of Building Energy Loads with Regression & Clustering Jan 2025

- 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 load forecasting accuracy by 12.3% using k-means clustering to identify distinct energy consumption patterns, reducing operational costs by 9%.