

Ansh Thukral

athukral007@gmail.com | github.com/thukralansh | [linkedin.com/in/ansh-thukral/](https://www.linkedin.com/in/ansh-thukral/) | 3-year OPT eligibility

PROFESSIONAL SUMMARY

Machine Learning Researcher with 2 years of experience developing deep learning models and scalable ML pipelines across scientific computing, high-energy physics, and time-series financial analysis. Trained in computational physics, I enjoy breaking down complex problems from first principles, designing rigorous experiments, and building models that are both scientifically grounded and practically useful. Passionate about advancing ML research while translating novel ideas into robust, production-ready systems.

EDUCATION

University of California, San Diego

Jun 2026

B.S. Physics w/ Specialization in Computational Physics; Minor: Speculative Design

WORK EXPERIENCE

Duarte Lab - Particle Physics & Machine Learning

Feb 2025 - Jun 2026

Undergraduate Research Assistant

- Developed end-to-end machine learning pipelines on the NRP Nautilus HPC cluster for large-scale collider-physics datasets, integrating preprocessing, feature engineering, event weighting, distributed GPU training, and model evaluation.
- Researched, developed, and rapidly iterated supervised and unsupervised models for anomaly detection and jet tagging, using hyperparameter optimization, calibration, and controlled benchmarking to improve predictive performance.
- Collaborated with researchers to formulate hypotheses, design experiments, interpret model behavior, and communicate findings that informed architecture selection and future research directions.

PROJECTS

Deep Learning for Scientific Time-Series Denoising (LIGO)

- Designed, trained, and evaluated a PyTorch CNN for time-series regression, reproducing the published DeepClean architecture to predict environmental noise in LIGO strain data from auxiliary witness channels.
- Built a reproducible benchmarking pipeline with independent train datasets, enabling rigorous comparison of denoising performance, model generalization, and robustness across diverse noise conditions.
- Optimized a hybrid MSE and ASD loss, demonstrating improved spectral reconstruction and generalization while preserving astrophysically relevant waveform characteristics.

Machine Learning for Higgs Boson Classification (Duarte Lab)

- Built a research pipeline for Higgs Boson classification on ~ 250,000 simulated ATLAS collision events, performing preprocessing, feature scaling, and orchestrating distributed training-validation jobs via Kubernetes.
- Developed and benchmarked a PyTorch multilayer perceptron, random forest, and gradient-boosted decision tree, comparing AMS, class-weighted accuracy, calibration, and computational efficiency to identify robust classification strategies.

ML-Driven Time-Series Stock Portfolio Forecasting <https://ansh-stock-forecasting.streamlit.app/>

- Developed and deployed an end-to-end machine learning system for financial time-series forecasting and portfolio optimization using Prophet, constrained Markowitz optimization, and cloud-based automation.
- Productionized the workflow through CircleCI, automated testing, scheduled execution, Supabase persistence, and cloud deployment, enabling reliable daily inference and decision support.

SKILLS

Programming & Frameworks: Python, SQL, C, PyTorch, scikit-learn, SciPy, Claude Code, Cursor, Codex, FastAPI

Machine Learning: Deep Learning, Classification, Anomaly Detection, Time-Series Modeling, Signal Processing, Feature Engineering, Hyperparameter Optimization, Model Evaluation, Benchmarking, Optimization

Research & Infrastructure: Distributed GPU Computing (Nautilus HPC), Docker, Kubernetes, CI/CD, LangGraph, RAG

Domain Knowledge: Embedded systems, Sensor Integration, Signal Processing, CAD