

EDUCATION

Program	Institute	CGPA/%	Year
Dual Degree (B Tech+M Tech) in Engineering Design	Indian Institute of Technology Madras	8.60/10.0	2026
Class XII (CBSE)	Velammal Bodhi Campus, Chennai	98.2%	2021
Class X (CBSE)	Velammal Bodhi Campus, Chennai	95.8%	2019

SCHOLASTIC ACHIEVEMENTS

- JEE Advanced 2021: **Top 1.5%**(Among 2 Lakhs); JEE Mains 2021: **99.78%**; National Entrance Screening Test 2021: **Rank 36**
- National Standard Examination in Physics 2021: **State Topper**; International Olympiad of Mathematics 2021: **State rank 25**
- Codeforces - **Specialist**; CodeChef - **3 Stars (Global Rank 35 in CodeChef Starters 202)**; Arithmetic Zetamac - **78 in 2 mins**

PROFESSIONAL EXPERIENCE

- **Featurely AI** *June 2025 - July 2025 (2 mos)*
Applied AI Intern San Francisco Bay Area Based
 - Featurely is a US AI startup backed by **Bling Capital**, whose partners have backed companies such as **Palantir** and **Lyft**.
 - Built an explainable **human-behavior simulation model** under data constraints using **LLMs, Probabilistic Modelling, graph-based memory system** simulating user behavior, for faster, user-focused product development and feature prioritization.
 - Engineered **episodic memory system** (Neo4j, Pinecone) with recursive retrieval (**vector RAG and graph RAG**).
 - Deployed the system (**FastAPI + Docker on AWS EC2** backend; **Vercel** frontend), enabling remote testing by stakeholders.
- **Caterpillar** *Jan 2025 - May 2025 (5 mos)*
Machine Learning Intern, Global Data Science Team (Optimisation Pod) Chennai, India
 - Developed **Multi-Agent Reinforcement Learning** dispatch system for mining fleets using **Proximal Policy Optimisation** algo.
 - Engineered **multi-objective Reward Framework**: Productivity **8-10%**↑, truck congestions **30%**↓, power efficiency **10%**↑.
 - Automated mine layout generation, cutting setup time from **hours to minutes** for rapid training (across 50+ site layouts).

RESEARCH EXPERIENCE

- **ADOPT Lab, IIT Madras**
 - (**Patent Applied**) Proposed and evaluated a **topology-preserving synthetic tabular data generation** approach for scarce and non-convex datasets, demonstrating improved stability, downstream task utility and fidelity over standard baselines.

PROJECTS

- **Fraud Ring Investigator Arena: RL Benchmark for Sequential Anti Money Laundering Investigation**
 - Built an **OpenEnv benchmark** for fraud triage with observability, action costs, delayed effects, and clear/escalate decisions.
 - Designed typed investigation/intervention actions with deterministic baselines on easy/medium/hard tracks for RL agents.
- **End-to-End RL Fine-Tuning Pipeline for DeepSeek-7B using Custom Reward Modeling**
 - Engineered a **Reinforcement Learning (GRPO) pipeline** to **fine-tune** a 7B LLM, implementing a custom reward function.
 - Optimized training efficiency using **gradient checkpointing** and **8-bit quantization** within the Hugging Face trl ecosystem.
- **Efficient LLM Adaptation: QLoRA-based Supervised Fine-Tuning of DeepSeek-7B**
 - Implemented **SFT²** using **LoRA³** adapters and 4-bit quantization (**QLoRA⁴**) to specialize LLMs for mathematical reasoning.
 - Integrated **gradient accumulation and mixed-precision training** to enable **efficient large-model tuning on consumer-grade GPUs**.
- **Productionizing LLMs: Quantization and Performance Profiling for Efficient Inference**
 - Executed a model optimization pipeline reducing **memory footprint by 70%** via 4-bit NF4 and Int8 **quantization** techniques.
 - Benchmarked **inference latency and throughput** across precision formats to identify optimal configurations for **edge deployment**.
- **Implementation of Custom Deep Learning Architectures for Computer Vision**
 - Implemented **YOLOv2 for Car Detection**, using IoU⁵/NMS⁶ for real-time bounding boxes in **autonomous driving**.
 - Designed **U-Net for Semantic Segmentation**, achieving 80% pixel accuracy on **object outlining tasks** (noisy data).
- **Clinically RAG: Chat-First Healthcare AI Research Assistant**
 - Developed a **chat-first healthcare assistant** using Next.js with a **7-stage LangGraph** pipeline on a 21-article corpus.
 - Implemented **hybrid retrieval**, citation-linked streaming, and unsupported-query refusal to keep answers **corpus-grounded**.
- **Foundational Machine Learning and Statistical Modeling**
 - Supervised Learning Models**
 - Implemented regression (**Ridge, Lasso**), on high-dim 2D data, achieving 99.5% MSE reduction and 86% feature sparsity.
 - Derived **Bayesian LDA/QDA⁷** classifiers with asymmetric loss on 2D (binary/multiclass) data, reducing errors by 72%.
 - Unsupervised Learning Models**
 - Implemented **clustering (DBSCAN, spectral)** for non-convex data, achieving 100% visual purity via Laplacians.
 - Developed **Expectation Maximization for Gaussian Mixture Models** (ELBO-monitored cluster alignment) on 2D data.

SKILLS

- **Programming:** C++, Python, SQL; **Tools:** Git, Docker, Redis, AWS, AirFlow, FastAPI, Neo4j, Pinecone
- **Utility Libraries:** statsmodels, PyTorch, Tensorflow, Scikit-learn, NumPy, Pandas, Matplotlib, Seaborn, Keras, SciPy

EXTRACURRICULAR

- **Tech Contingent in Inter IIT Tech Meet 13.0 (Represented IITM competing against 23 IITs; Organised by ISRO):**
- Generated a lunar map at sub-pixel resolution by analyzing X-ray fluorescence data from Chandrayaan-2's CLASS instrument.

RELEVANT COURSEWORKS

- Probability, Statistics and Stochastic Processes; Linear Algebra; Differential Equations; Series and Matrices
- Machine Learning and Deep Learning Specialisation*; Pattern Recognition and Machine Learning
- Fine-tuning & RL for LLMs: Intro to Post-training**; First Principles Tools; Data Analysis & Visualization in R/Python/SQL
- Operations Research; Operations Management; Principles of Economics; Econometrics; Business Analytics

POSITIONS OF RESPONSIBILITY

- **Core Team Member**, Wellness Team, IIT Madras *Aug 2022 - May 2025 (3 yrs)*
Managed a team of **50** coordinators; Led numerous wellness events; Offered peer counselling support to students.
- **Department Coordinator**, Academic Mentorship Cell, IIT Madras *Sep 2023 - May 2024 (1 yr)*
Facilitated academic journeys for **20** mentees by matching them with suitable mentors and monitoring mentees' progress.

*- Stanford Online, **-DeepLearning.AI, 1-Root Mean Squared Error, 2-Supervised Fine Tuning, 3-Low-Rank Adaptation, 4-Quantized Low-Rank Adaptation, 5-Intersection Over Union, 6-Non-Max Suppression, 7-Linear Discriminant Analysis/Quadratic Discriminant Analysis