

Jaisurya Sundarbabu

San Diego, California, United States | jsundarbabu@ucsd.edu | (858) 220-8033 | [LinkedIn](#) | [GitHub](#)

EDUCATION

University of California San Diego

Master's in Data Science. GPA: 3.77

San Diego, California

Sep 2025 – Dec 2026

Sri Venkateswara College of Engineering

B.Tech Information Technology. GPA: 8.82

Chennai, Tamil Nadu

Jul 2019 – May 2023

EXPERIENCE

Celebal Technologies

Associate Data Scientist

Jaipur, India

Mar 2024 – May 2025

- Designed, developed, and tested an **end-to-end LLM-based RAG** system with **semantic search** (embeddings + retrieval) for manufacturing issue resolution; achieved **95%** accuracy on unseen issues and reduced new Jira tickets by **70%** via structured evaluation and error analysis.
- Developed **predictive models** for retail **demand forecasting** on large historical datasets using time-series feature engineering and **statistics**-driven backtesting; delivered **15% MAPE** (1-year horizon) and supported planning decisions.
- Automated retail cataloging using **GPT-4 Vision + computer vision**: image preprocessing, view classification (front/back/side/detail), and prompt-based attribute extraction (color/pattern/collar/sleeves); stored structured outputs in **Databricks Delta Lake (Delta tables)** for e-commerce ingestion.
- Implemented an end-to-end **Databricks** pipeline using **MLflow** for experiment tracking, model deployment, and monitoring; enforced rigorous validation and logging, saving 10+ hours/week of manual oversight.

Samsung Digital Commerce

Software Development Engineer-1

Chennai, India

Jul 2023 – Jan 2024

- Owned **backend** services for the Samsung Shop app, integrating 6 APIs and scheduling recurring cron jobs to maintain data freshness and improve **availability** and **reliability**.
- Designed, developed, and tested backend changes for the global Samsung S24 launch; strengthened **observability** with logging/validation and implemented access controls aligned with internal **compliance** and production standards.

Actyv.ai

Data Science Intern

Bangalore, India

Feb 2023 – May 2023

- Delivered an **end-to-end** document classification pipeline using **deep learning + NLP**; reduced manual onboarding data entry by **80%** through robust preprocessing and quality checks.
- Created and evaluated an ensemble of DL/NLP models achieving **90%** accuracy; routed outputs to Amazon S3 with monitoring hooks and data handling aligned with onboarding compliance requirements.

KGiSL

Machine Learning Intern

Coimbatore, India

Aug 2021 – Nov 2021

- Designed an **NLP**-based resume filtering tool that automated **80%** of initial screening; iterated on **preprocessing and keyword/ranking** logic to improve precision and reduce manual review time.
- Built churn **predictive models** using Logistic Regression and **XGBoost**; achieved **93%** accuracy and presented data analytics insights with clear data visualization summaries for retention actions.

PROJECTS

Fitness Tracker

- Developed a computer vision workout-tracking app using MediaPipe to score form and provide real-time posture feedback; measured reps/accuracy, added a leaderboard, and won Best Solution Award at StepSetGo Technothon'22.

Fund-Level Mutual Fund AUM Forecasting (Tech, Energy, Healthcare)

- Implemented a Python time-series pipeline (OLS, VAR, SARIMAX, Kalman, Random Forest, XGBoost) to forecast monthly fund-level AUM using macro + market features; evaluated models with MAPE/RMSE and statistics tests (Mincer-Zarnowitz, Diebold-Mariano), and produced comparative data visualization reports.

CERTIFICATIONS

Machine Learning Professional - Databricks

Advanced machine learning tasks

Machine Learning - Coursera

Fundamentals of Machine Learning from Stanford

Neural Networks and Deep Learning

DeepLearning.AI

Improving Deep Neural Networks: Hyperparameter Tuning, and Optimization

DeepLearning.AI

Full Stack Web Development - Udemy

using HTML, CSS, JS, Python, and Django.

AWARDS & ACHIEVEMENTS

E-Mind by StepSetGo (May 2022): Best Solution Award for a real-time CV fitness tracker using MediaPipe pose landmarks to score exercise performance.

Technothon '22 (May 2021): Top 7/100 teams nationwide for a healthcare emergency platform connecting hospitals, blood banks, and donors (Python/Django).

HACKIT (May 2022): Award-winning full-stack recipe recommendation app providing personalized alternatives and nutrition insights via APIs.

Ease The Error 2.0 (May 2021): 3rd place in an inter-college hackathon for a yoga assistance app using MediaPipe for posture tracking and feedback.

SKILLS

Languages: Python, SQL, Java, JavaScript (JS), C/C++

ML/AI: Machine learning, Deep learning, LLM Inference, predictive models, regression, statistics, experimentation (hypothesis testing), NLP, computer vision, LLMs, RAG, semantic search, prompt engineering, Supervised Fine Tuning.

Data & Systems: Databricks, Spark/PySpark, Delta Lake, AWS, Azure, Git; monitoring, logging, observability

Frameworks: PyTorch, TensorFlow, Keras, Scikit-learn

Core CS: Object-oriented programming, data structures & algorithms, backend APIs, end-to-end pipelines