

ATHARVA YEOLA

+1 8582203508 | atharvayeola12@gmail.com | [in/atharvayeola](https://in.linkedin.com/in/atharvayeola) | [G](https://www.github.com/atharvayeola) /atharvayeola | [I](https://www.instagram.com/atharvayeola) /atharvayeola

EDUCATION

University of California San Diego (UCSD)
Master of Science (M.S) – Data Science & Machine Learning

Sep 2023 – Jun 2025

San Diego, USA

Pune Institute of Computer Technology (PICT)
Bachelor of Engineering (B.E.) – Electronics & Telecommunication

Aug 2019 – Jun 2023

Pune, India

WORK EXPERIENCE

Applied Data Scientist

Qualcomm Institute

Jul 2024 – Jun 2025

San Diego, USA

- Built ‘Waldo’, an open-sourced LLM tool to flag adverse events in unstructured text; fine-tuned RoBERTa in PyTorch on 360K+ Reddit posts, built custom evaluation scripts for rare-class metrics and conducted failure mode analysis.
- Conducted a causal impact study using synthetic control and ARIMA time series forecasting to quantify a 23% rise in gambling-related risk post U.S. sports betting legalization; insights shaped policy recommendations and outreach.

Research Scientist

Indian Institute of Technology (IIT) Patna

Dec 2022 – Aug 2023

Remote

- Generated a 100K-image (30% increase) synthetic dataset using image augmentation, improving computer vision model robustness for occluded traffic sign detection in autonomous driving.
- Designed and evaluated a deep ensemble (ViT + EfficientNet), improving F1 score by 9% over experimental baselines and achieving over 85% precision on long-tail edge cases.

Senior Data Scientist, Ranking

RhythmFlows Solutions Pvt. Ltd

Aug 2022 – Apr 2023

Pune, India

- Devised a personalized context-aware music recommender for a mental-health app, curated 75 seed tracks collaborating with psychiatrists, mitigating cold-start; boosted NDCG@10 by 18% and hit-rate@5 by 14%
- Productionized the model on AWS Sagemaker at scale with continuous training and rollout. Designed an A/B test; treatment group showed a 33% increased click-through rate (CTR), and increased session duration by 20%.

Machine Learning Engineer

Atomic Loops

Jan 2022 – Jun 2022

Pune, India

- Spearheaded a team of three in building a YOLOv5-based real-time object detection model for for conveyor-based food delivery, achieving 98% detection precision and boosting footfall by 30%.
- Optimized real-time inference by aligning model latency with ops & hardware teams to, reducing false negatives by 15%; later reused for inventory tracking, reducing food waste by 5x.

PUBLICATIONS & CONTRIBUTIONS

- Authored a statistical investigation in JAMA Internal Medicine (acceptance rate: 4%) on gambling addiction trends.
- Led and authored a novel research study in IEEE Xplore (CVMI 2024) on occluded traffic sign detection.

TECHNICAL PROJECTS

e-CommerceAgent | FastAPI, Next.js, Transformers, Docker

- Engineered a multi-modal commerce agent (text + image) with deterministic fallbacks; delivered sub-1.5s p95 chat responses across a 50-product, 10-category catalog in local benchmark runs.
- Hardened web enrichment by adding allowlisted OG/JSON-LD parsing and a server-side Browse.ai adapter, cutting scrape/extractor failures from 45% to less than 10%.

MediSyn-Vision | Diffusion, LoRA, dataset-creation

- Productized the diffusion-LoRA and StyleGAN2 pipeline into six CLI stages that automate data prep, training, sampling, evaluation, reporting, and demo launch for SynMedVision.
- Streamlined generation of 10K-image batches with manifest logging supporting tracking of KID and TSTR metrics.

TECHNICAL SKILLS

Programming – Python, PyTorch, SQL, C++, Java, TensorFlow, Pandas, NumPy, SciPy, scikit-learn

Data & Platforms – AWS (SageMaker, S3, EC2/ECR, Batch), Docker, Git/GitHub Actions, Spark, dbt, ETL

Science & Modeling – Search Ranking, LLMs, Computer Vision, Causal Inference & A/B testing

Visualization – Tableau, PowerBI, Plotly, Matplotlib, Dash

Other – Probability, Statistics, Agentic AI, Metric design, Failure analysis