

Atharv Khare

✉ atharvkhare18@gmail.com | 📞 8817916681 | 🌐 atharvk.me | [📄 LinkedIn](#) | [📄 GitHub](#)

Education

Indian Institute of Technology Madras

Bachelor of Science: Data Science & Applications

CGPA: 9.02

Expected 2027

Skills

Languages: Python, SQL, JavaScript, TypeScript

Machine Learning: Classical ML, Ensemble Learning, Time-Series, Statistical Inference

AI Systems: LLMs, RAG, LangGraph, LangChain, Vector Databases, Prompt Engineering, Multi-Agent Systems

Frameworks & Tools: FastAPI, Flask, PostgreSQL, pgvector, Scikit-Learn, CatBoost, LightGBM, Pandas, NumPy, Git

Coursework: Machine Learning, Probability & Statistics, Data Structures & Algorithms, Business Analytics, DBMS

Experience

Learning Assistant: Tools in Data Science, IIT Madras

Jan 2026 – May 2026

- Selected as a Learning Assistant for IIT Madras Tools in Data Science, supporting 1,000+ students through technical mentoring, doubt resolution and community initiatives.
- Built TDS-Collab, a student platform for assignment tracking, submission discovery and leaderboard-based engagement, improving peer collaboration across the course community.
- Developed TDS-Manual, a searchable knowledge base consolidating course content, solutions and learning resources into a centralized reference used by students throughout the term.

Data Analyst: Shri Melan FF (QSR), Bhopal

Sep 2025 - Feb 2026

- Engineered Apriori Association Rules on a 153-item menu, driving a **14% ticket-size increase**; built **SARIMAX + Bayesian demand forecasting** for 90-day SKU projections and a BCG-adapted Menu Matrix, improving gross margin by **8.5%** and cutting stockouts by **35%**.

Projects

Sentio: NLP-Powered Cognitive Bias Awareness Platform [🔗 Live](#)

2025–26

- Built an AI-assisted journaling platform combining Claude Haiku, VADER sentiment analysis and a custom 15-class cognitive-bias taxonomy; processed journal entries into longitudinal behavioural profiles and personalized cognitive archetypes.
- Designed a two-stage retrieval pipeline using all-MiniLM-L6-v2 embeddings, pgvector semantic search and Cohere reranking; retrieved top-k historical reflections for context-aware insight generation with SSE response streaming.
- Implemented adaptive coaching workflows using Continuous Bayesian Knowledge Tracing, a 7-state Socratic dialogue engine and SM-2 spaced repetition to personalize reflection depth and follow-up recommendations.

GitSyntropy: Developer Compatibility Intelligence Platform [📄 GitHub](#) [🔗 Live](#)

2026

- Developed a software-team compatibility engine using GitHub behavioural telemetry and adaptive psychometric profiling; analysed 46 open-source developers across 10,886 commits to derive pairwise compatibility scores.
- Built a circular-statistics chronotype classifier from commit timestamps, achieving Macro-F1 of 0.86 and 75% cross-platform agreement between GitHub and Stack Overflow activity patterns.
- Built a recommendation engine that simulates candidate-team interactions using Monte Carlo search and compatibility scoring, generating optimal hiring profiles while reducing psychometric assessment length by 37.5% ($r=0.965$ score fidelity).

Flipkart Gridlock: Spatio-Temporal Traffic Demand Predictor [📄 GitHub](#) [🔗 Live](#)

2025

- Built a CatBoost-LightGBM ensemble for next-day traffic demand forecasting across 1,249 geohash regions, achieving a competition score of 92.397 on 77,299 labelled observations.
- Engineered 15+ spatial and temporal features including K-fold target encodings, cyclic time representations and inter-day demand anchors; target-encoded geohash-slot interactions contributed nearly 30% of overall feature importance.
- Investigated demand-distribution behaviour using log-transformed targets and spatial baselines ($R^2=0.695$), improving forecast stability while preventing cross-day leakage during training.

Achievements

Awards: Best Capstone Project in Business Data Management, IIT Madras BS Degree Program.

Certifications: NPTEL Distributed Systems (Elite), HackerRank Advanced SQL, DataCamp Data Science Associate.

Research: Author of three research preprints spanning AI systems, educational analytics and remote sensing: GitSyntropy, TerraHeal and CS-LLN.

Competitions: Achieved Top-10 placements in three Kaggle competitions; Winner – Anthropic CBC Spring Challenge, Inter House Hackathon IIT Madras, Anukriti IIT Madras.