Rohan Bali

Machine Learning Researcher

rohanbaliwork@gmail.com   +1 (774) 524-0505   Boston, MA

GitHub   Portfolio   LinkedIn   Twitter   Medium

Summary

Graduate student in Data Science at UMass Dartmouth focusing on spatio-temporal machine learning, network science, and computational social systems. My research examines how spatial structure, temporal dynamics, and network interactions shape the behavior of cities, elections, mobility networks, and scientific communities. I build predictive and interpretable models that operate under distribution shift, limited supervision, and heterogeneous real-world constraints. Previously at Capgemini and Upcred.ai.

Education

M.S. in Data Science, University of Massachusetts Dartmouth   2024–2026
Thesis: *Learning from Minimal Data under Distribution Shift (BrainLab OS)*
Relevant Coursework: Machine Learning, Deep Learning, Advanced Statistics, Network Science, Optimization, Big Data Analytics, Market Research

B.Tech in Computer and Communication Engineering, Manipal University Jaipur   2017–2021

Research Projects

Neural Time Capsule: Urban Growth Prediction   2025–2026

- Developed a ConvLSTM-based forecasting model for long-horizon urban expansion using GHSL built-up rasters and OpenStreetMap road networks.
- Trained on 2,313 CONUS tiles (1975–2000); achieved MSE = 0.000218 and RMSE = 0.00641.
- Demonstrated 67% improvement over U-Net baseline with stable generalization across regions.

Academic Collaboration Networks   2024

- Analyzed 180K+ co-authorship edges across 12 academic domains using small-world statistics.
- Found that 72% of high-impact researchers cluster within a 3-hop neighborhood.
- Developed reproducible pipelines for citation graph construction and author-impact inference.

Disaster Risk Monitoring Using Satellite Imagery – NVIDIA DLI   2025

- Trained U-Net on Sentinel-1 radar imagery for flood segmentation (Dice = 0.82, IoU = 0.78).
- Designed preprocessing and augmentation pipeline optimized for noisy radar data.
- Integrated predictions into Google Earth Engine for real-time risk assessment.

Selected Engineering Projects

Tripsy: AI Travel Assistant Using RAG   2025

- Built retrieval-augmented assistant with geographic embeddings and OpenRoute integration.
- Reduced hallucinations by 35% and improved route relevance by 22% over GPT-only models.

Local MapReduce Using LevelDB   2024

- Implemented local MapReduce architecture with LevelDB storage and multiprocessing.
- Achieved near-linear scaling on 8 cores ($R^2 = 0.94$) and 1.8× improvement over naive baselines.

Certifications

NVIDIA DLI Certified – Disaster Risk Monitoring with Satellite Imagery   2025
Docker Certified – Foundations Professional   2025

Technical Skills

Machine Learning: Spatio-temporal models, ConvLSTM, GNNs, forecasting, distribution shift
Mathematics: Generalization theory, robustness, PAC learning, optimization
Languages & Libraries: Python, R, SQL, PyTorch, TensorFlow, NumPy, Pandas
Systems: Docker, Git, Airflow, LevelDB
Frameworks: LangChain, Earth Engine, scikit-learn
Cloud: AWS, Azure, GCP, Milvus Vector DB

Publications & Writing

- "Understanding Academic Collaboration Networks through Small World Theory."
  swn465.home.blog (2024)
- "Neural Time Capsule: Urban Growth Prediction."
  Bali2025_NeuralTimeCapsule_UrbanGrowthPrediction.pdf (2025)

Open Source & Portfolio

- GitHub: github.com/rohanbalixz
- Portfolio: rohanbalixz.github.io
- Medium: medium.com/@bali2rohan
- Google Scholar: scholar.google.com/citations