

Richa Verma

richa.verma93@gmail.com — [LinkedIn](#) — [GitHub](#)

Summary

Senior Applied AI Scientist with 8+ years of experience building and deploying Generative AI and ML systems for enterprise applications. Strong expertise in LLMs, RAG pipelines and MLOps, with hands-on experience deploying AI solutions on AWS. Proven ability to translate research ideas into scalable, production-ready AI systems. Proven record of publications A* conference venues, patents, and large-scale ML systems. Current work focuses on post-training LLMs using preference optimization and RL-inspired objectives for safety-critical reasoning.

Education

IIT Madras	Aug 2020 – Present
Ph.D., Computer Science (Expected graduation: 07/26)	CGPA: 9.0
IIIT Delhi	2015 – 2017
M.Tech., Computer Science & Engineering	CGPA: 9.53
MAIT, GGSIPU, Delhi	2011 – 2015
B.Tech., Computer Science & Engineering	CGPA: 8.65

Work Experience

TATA Consultancy Services Ltd. - Applied AI Scientist	Aug 2017 – Present
<ul style="list-style-type: none">Designed and deployed end-to-end ML and Generative AI pipelines covering data ingestion, feature engineering, training, evaluation, and deployment.Built LLM-powered prototypes integrating structured enterprise data (SQL) with language models using retrieval-based and agentic approaches.Implemented preference-based and offline fine-tuning workflows with strong emphasis on evaluation, robustness, and safe deployment.Collaborated with cross-functional stakeholders to translate business requirements into AI-driven solutions.Designing post-training pipelines for LLMs using preference optimization and implicit reward modeling, evaluated on multi-step reasoning safety benchmarks.Designed RL algorithms for high-dimensional planning and combinatorial optimization problems.	

Skills

Programming: Python (expert), PyTorch (expert), SQL, Pandas, Numpy, Scikit-learn.

Generative AI & LLMs: RAG pipelines (embeddings, vector search), Prompt engineering, Prompt optimization, HuggingFace Transformers, PEFT/LoRA, Google ADK for agentic AI workflows.

ML Engineering & MLOps: ML pipelines, feature engineering, experiment tracking (WandB), model evaluation, Docker, Linux.

Deployment: AWS (S3, EC2, Bedrock), Docker, REST APIs, Git-based CI/CD.

Projects (Selected)

Enterprise RAG System for Structured & Unstructured Data (Dec 2025 - Present)

Built a RAG system combining SQL databases and document stores for enterprise Q&A. Implemented chunking, embedding, vector retrieval, and response synthesis. Exposed inference via REST API; containerized with Docker and deployed on AWS EC2/S3. Focused on retrieval evaluation and cost-aware design.

Agentic LLM for Data Analysis & Decision Support (Nov 2025 - Present)

Developed an LLM-driven agent capable of querying SQL databases and performing Pandas transformations. Implemented tool selection logic for data querying, transformation, and insight synthesis. Demonstrated enterprise use cases such as anomaly detection and trend analysis.

RL for Safety in Multi-step Language Model Reasoning (Oct 2025 – Present)

Designed post-training methods for large language models using preference-based optimization, inspired by offline RL and DPO. Evaluated safety and reasoning consistency on WildJailBreak using Qwen and DeepSeek-R1 models. **Direct Preference Optimization for Safe RL** (Aug 2024 – Oct 2025)

Casting the problem of safety in RL as a safety alignment problem for Mujoco continuous control.

Offline RL for Joint Optimization of EV & HVAC Systems (May 2025 – Aug 2025)

Offline RL on limited high-fidelity data for energy-efficient HVAC and EV scheduling.

Transfer Learning in Offline RL (May 2022 – Jun 2023)

Studied transfer learning in offline RL for safety-critical domains such as robotics.

Threshold-based Improvement Rewards for RL (Jan 2019 – Apr 2020)

Developed a reward-shaping technique to improve performance of existing RL algorithms.

RL for Robotic Package Loading (Jun 2018 – Apr 2019)

Built RL-based solution for parcel packing in transportation containers, maximizing volume utilization and stability.

Patents

System and Method For Autonomous Multi-Bin Parcel Loading System, A. Singhal, S. Thukral, V. Raju, **R. Verma**, H. Khadilkar, US 12,307,405 B2.

Method And System For Container Loading Sequence Generation Using Reinforcement Learning (RL), **R. Verma**, S. Saikia, A. Srinivasan, P. Agarwal, G. Shroff, IN Patent 492,460.

System and Method for Autonomous Multi-bin Parcel Loading System, A. Singhal, S. Thukral, V. Raju, **R. Verma**, H. Khadilkar, IN Patent 527,907.

Awards & Achievements

First place, WSAI Research Showcase at IIT Madras, 2025 (100+ participants).

Dean's list for academic excellence, 2015–2016.

Winner, Granular.AI Deep Learning Challenge on HackerEarth (2,681 teams).

Secured 99.4 percentile in GATE 2015.

Theses

Ph.D. Thesis: Safety in Offline RL

Advisor: Dr. Balaraman Ravindran

Making RL safer and deployable for continuous control and language model reasoning.

M.Tech. Thesis: Mobility Assistance for Visually Impaired (MAVI)

Advisors: Dr. Chetan Arora, IIT Delhi; Dr. M. Balakrishnan, IIT Delhi

Computer vision system to detect animals, road signs, pavement, and pedestrians.

Research Publications (Selected)

PREFINE: Preference-Based Implicit Reward and Cost Fine-Tuning for Safety Alignment, **R. Verma**, B. Kulur, S. Chawla, B. Ravindran, *AAMAS (A*)*, 2026.

Guiding Offline Reinforcement Learning using a Safety Expert, **R. Verma**, D. Kalwar, H. Khadilkar, B. Ravindran, *CODS-COMAD*, 2024.

Guiding Offline RL using Safety Expert, **R. Verma**, K. Bharadwaj, H. Khadilkar, B. Ravindran, *Offline RL Workshop, NeurIPS*, 2022.

SIBRE: Self Improvement Based Rewards for RL, S. Nath, **R. Verma**, A. Ray, H. Khadilkar, *AAMAS*, 2021.

A Reinforcement Learning Framework for Container Selection and Ship Load Sequencing, **R. Verma** et al., *AAMAS*, 2019.