

MRITUYANJAY SHARMA

AI Automation Engineer · LLM Systems Engineer · Agentic AI Developer

Vadodara, India · +91 9313127246 · Sharmamritujanjay@gmail.com · github.com/Misharma1234

PROFESSIONAL SUMMARY

Production-focused AI Automation & LLM Systems Engineer with hands-on experience architecting end-to-end agentic pipelines, RAG systems, and autonomous AI agents that deliver measurable business impact. Proven track record of deploying LLM-powered workflows using N8N, LangChain, and multi-agent orchestration frameworks — reducing manual effort by up to 45% and cutting lead qualification overhead by 40%. Pursuing M.Sc. in AI & Data Science at BITS Pilani (2027). Strong ML foundation across deep learning (CNNs), vector databases, inference optimization, and cloud-native AI deployment on GCP and AWS. Seeking to bring production-grade AI engineering depth to a fast-moving GenAI or AI Automation team.

TECHNICAL SKILLS

LLM / Agentic AI	LangChain, LangGraph, N8N, Agentic Workflows, Multi-Agent Orchestration, RAG (Retrieval-Augmented Generation), Prompt Engineering, LLM Fine-Tuning, Vector Databases, Context Management, AI Observability
ML / Deep Learning	TensorFlow, Keras, Scikit-learn, CNNs, Random Forest, OpenCV, Feature Engineering, Model Evaluation, Inference Optimization, NumPy, Pandas, Matplotlib, Seaborn
MLOps & Cloud	GCP, AWS (Elastic Beanstalk, EC2), Docker, CI/CD Pipelines, REST APIs, Webhook Integration, AI Pipelines, Software Architecture, Git
Languages & Data	Python, TypeScript, JavaScript, Java, PHP, SQL, MongoDB, PostgreSQL, Supabase
Web & Tools	Next.js, Airtable, Pygame

WORK EXPERIENCE

AI / LLM Solutions Engineer · Makeshift Solutions

Sep 2025 – Present

- ▶ LeadXpress Architected and deployed
- ▶ Architected and deployed Lead Xpress, an N8N-based multi-agent AI system for autonomous lead qualification from LinkedIn and Google; eliminated 40% of manual vetting time and integrated LLM-based scoring logic with structured vector database outputs.
- ▶ Designed and shipped a production-grade GenAI Reddit automation agent leveraging LLM inference pipelines for contextual brand engagement; achieved ~25% uplift in post engagement across client campaigns through prompt-optimized reply generation.
- ▶ Engineered end-to-end AI pipelines integrating LLM APIs, vector databases, and structured data outputs; owned architecture decisions covering context management, API orchestration, and inference reliability for 3+ concurrent client automation workflows.
- ▶ Implemented AI observability hooks and validation layers across agentic pipelines to ensure output quality, enabling consistent scalable deployment and reducing debugging turnaround by ~35%.

Teaching Assistant · Parul University

Sep 2023 – Jan 2024

- ▶ Mentored 200+ students in core software engineering across Python, Java, and PHP; conducted weekly code reviews, debugging sessions, and 1:1 consultation that measurably improved assignment pass rates.
- ▶ Designed and deployed a centralized software system for tracking student mentorship records and session logs, improving team-wide reporting efficiency by 30%.

EDUCATION

M.Sc. — Artificial Intelligence & Data Science

Sep 2025 – Present (Expected 2027)

SELECTED PROJECTS

Adaptive RAG — Agentic AI Chatbot | *LangChain · LangGraph · FastAPI · Qdrant · MongoDB · Streamlit* [[GitHub ↗](#)]

- ▶ Built an end-to-end Adaptive RAG system with intelligent query routing that dynamically classifies user queries into index-based retrieval, general knowledge, or real-time web search pipelines — optimizing response accuracy across diverse query types.
- ▶ Designed a multi-node LangGraph workflow with 7 specialized nodes (query analysis, retrieval, relevance grading, query rewriting, response generation, web search, general LLM) orchestrated via a stateful directed graph architecture.
- ▶ Implemented vector similarity search using Qdrant with configurable document chunking and OpenAI embeddings for semantic indexing; integrated an automatic relevance grading and query rewriting system to improve retrieval precision.
- ▶ Developed a production-grade async FastAPI backend with RESTful endpoints, MongoDB-backed persistent chat history for multi-session tracking, and an interactive Streamlit frontend with user authentication and real-time document upload.

Lead Generation AI Agent | *N8N · LLMs · LangChain · Vector DB · REST APIs* [[GitHub ↗](#)]

- ▶ Built an autonomous agentic lead generation system with multi-step reasoning: prospect discovery, contact enrichment, deduplication, LLM-based quality scoring, and structured database logging — reducing manual prospecting effort by ~40%.
- ▶ Implemented context management and validation logic across agent nodes to maintain pipeline reliability and data fidelity at scale.

AI-Powered HR CV Screening Pipeline | *N8N · LangChain · LLMs · RAG · Vector Databases* [[GitHub ↗](#)]

- ▶ Designed a production-grade automated screening pipeline ingesting PDF resumes, extracting qualifications via LLM parsing, and evaluating candidates against role-specific criteria using retrieval-augmented generation (RAG).
- ▶ Replaced manual resume screening with structured AI outputs; reduced screening time by ~60% and improved consistency of candidate shortlisting at scale.

Deep Research Multi-Agent System | *Agentic AI · Tavily Search · LLMs · Structured Outputs* [[GitHub ↗](#)]

- ▶ Engineered a multi-step autonomous research agent with recursive intent clarification, high-signal web sourcing, relevance ranking, and citation-aware report synthesis — delivering structured technical outputs with ~45% reduction in research effort.
- ▶ Implemented AI observability and output validation layers ensuring structured, actionable report generation with minimal hallucination risk.

RISE Research HR Portal | *Next.js · Supabase · Airtable · TypeScript · REST APIs* [[GitHub ↗](#)]

- ▶ Architected a full-stack internal HR management system with role-based access control, real-time Supabase ↔ Airtable data synchronization, and automated data pipelines — cutting HR administrative overhead by over 50%.
- ▶ Implemented secure authentication flows, dynamic dashboards, and automated reporting — deployed to production and actively used by the organization.

Handwritten Digit Classifier (Real-Time CNN) | *Keras · TensorFlow · CNN · Pygame · OpenCV* [[GitHub ↗](#)]

- ▶ Designed and trained a multi-layer CNN achieving ~98% test accuracy on MNIST; integrated real-time inference into a Pygame GUI supporting live drawing, image upload, and immediate classification with user authentication.

FinGenius — Real Estate Price Prediction | *TensorFlow · Scikit-learn · Python · Feature Engineering* [[GitHub ↗](#)]

- ▶ Engineered a regression model with intensive feature selection and EDA achieving RMSE of 0.2 on held-out test data; documented model architecture and evaluation methodology for reproducibility.

Additional Projects

- ▶ Career Prediction Software (Freelance): Integrated a Random Forest ML model into a resume generator SaaS platform for data-driven career path recommendations.
- ▶ Online Railway Reservation System: Full-stack platform (PHP, Tailwind CSS, JavaScript) with complete CRUD operations, seat booking logic, and session-based authentication.
- ▶ GroceryGo: Mobile delivery app developed with Advanced Java and XML featuring real-time inventory and order management.

CERTIFICATIONS & TRAINING

- ▶ Designed a scalable, fault-tolerant hosting architecture using AWS Elastic Beanstalk for a high-traffic client; produced stakeholder-ready technical documentation covering system design and cost structure.