

TARKESHWAR NARAYAN SHARMA

Senior AI/ML Engineer | Generative AI & LLM Specialist | RAG Systems Expert

📍 India ☎ +918539975201 📩 tns15june@gmail.com 🚩 linkedin.com/in/tarkeshwar-narayan-sharma

Professional Summary

Senior AI/ML Engineer with 8.5 years designing → training → shipping production models at enterprise scale. Key wins: +43% engagement boost, 7.5% conversion uplift, \$420K+ cost savings through intelligent automation. Expert in Generative AI, RAG systems, and multilingual NLP with proven track record leading cross-regional teams. Seeking Principal AI Engineer role to architect next-generation AI systems that drive \$50M+ business impact.

Technical Skills

- 🔥 Core AI/ML: Python, PyTorch, TensorFlow, Scikit-learn, MLOps, Deep Learning, Computer Vision, NLP
- 🚀 Generative AI: LLMs (GPT, Claude, LLaMA, Gemini), RAG, Prompt Engineering, Vector Databases, AI Agents
- ⚡ Emerging Stack: LangChain, LlamaIndex, Strands, MCP, Hugging Face, OpenAI APIs, Cohere v3, FAISS
- ☁️ Cloud & MLOps: AWS (SageMaker, Bedrock, Lambda), Databricks, Apache Spark, MLflow, Docker
- 🔧 Frameworks & Tools: FastAPI, YOLO, spaCy, Spark NLP, ChromaDB, PostgreSQL, MongoDB
- 👥 Leadership: AI Project Management, Stakeholder Management, ML System Design

Experience

Topsource Worldwide

November 2023 – Present

Senior Data Scientist

Remote, India

- Intelligent Logistics with AI Agents - Architected AI-powered logistics platform using MCP agents & Strands framework, implementing intelligent job assignment algorithms that automate delivery coordination and optimize driver efficiency. Integrated Mapbox APIs for real-time traffic analysis enabling dynamic route optimization & proactive prediction. Led 5-person development team building FastAPI backend, scaling system to handle 100+ concurrent jobs through agentic workflow architecture and intelligent scheduling automation.
- Multilingual Enterprise RAG Chatbot for DEWA - Led multilingual chatbot development for DEWA enterprise using Claude 3 & Cohere v3 embedding models. Integrated AWS SageMaker, Lambda & SharePoint for Arabic/English document processing. Pioneered advanced prompt engineering and implemented performance analytics dashboard, collaborating with Middle East stakeholders for enterprise-grade RAG architecture delivering optimized response accuracy across multilingual queries.
- Generative AI-Powered Airline Dynamic Pricing System - Directed cross-regional teams across North America, Europe & Asia, engineering AI-powered dynamic pricing system. Designed multi-agent architecture with autonomous agents for real-time price adjustments, market analysis & revenue forecasting. Integrated competitor pricing, demand patterns & seat inventory data, enabling automated price recommendations that significantly enhanced airline revenue optimization through intelligent algorithmic decision-making.
- Multilingual Spell Correction System for Quick Commerce - Architected RAG-enhanced spell correction system using Llama3-8B & Databricks, implementing semantic retrieval with Vector DB & ANN search. Engineered two-stage architecture with instruct fine-tuning, reducing token count by 40% while improving inference latency by 18s for 1000-query batches. Deployed production-scale solution handling vernacular phonetic inputs, achieving 7.5% conversion rate increase through intelligent multilingual query understanding.

Statusneo

June 2021 – November 2023

Consultant- Data Scientist

Gurugram, India

- AVA Tracking System - Computer Vision Application - Managed end-to-end development of sophisticated computer vision application deployed for global clients. Led multi-disciplinary team designing real-time object detection using YOLO model for logbook & starter pack detection. Ensured seamless deployment across different operational

environments, collaborating with diverse stakeholders to enhance operational efficiency through advanced computer vision and intelligent tracking capabilities.

- **Personalized News Recommendation System - News Platform** - Built personalized recommendation system using entity recognition & knowledge graphs for content relevance improvement. Integrated trending news summaries with real-time recommendation engine, boosting user engagement & interaction across News platform. Optimized recommendation algorithms for real-time updates, significantly increasing user retention metrics and content interaction through intelligent personalization and dynamic content delivery for enhanced user experience.
- **Custom Sentiment Analysis Model for Streaming Platform**- Developed custom sentiment analysis model supporting multiple languages for large streaming company. Trained deep learning models on diverse datasets for audience sentiment classification & content strategy optimization. Implemented real-time sentiment tracking and optimized content recommendation algorithms using sentiment insights, improving viewer retention & platform engagement through enhanced user personalization and intelligent content delivery across streaming platform.
- **AI-Powered HR System with Talent Management** - Designed internal HR system with skill extraction & JD-CV matching for talent database management. Built NLP-based resume parsing module enabling automated candidate-job matching & recruitment optimization. Improved talent acquisition efficiency through AI-driven analytics for skill-gap identification & workforce planning, streamlining recruitment workflows and enhancing hiring decision-making through intelligent candidate matching and automated talent assessment across organizational departments.

Ayurs Infotech Pvt Ltd

Senior Data Analyst

March 2017 – January 2021

New Delhi, India

- **Large-Scale Image Matching using Vector-Search** - Led team development of image matching system, enhancing retrieval accuracy by 30% using advanced vector-search algorithms. Implemented efficient vector search across massive datasets, creating scalable solution for image-based applications. Optimized search algorithms significantly reducing search times and improving user experience, delivering high-performance image retrieval through intelligent vector matching and scalable architecture for production deployment.
- Pioneered recommendation engine, boosting engagement by 43% and driving user retention and sales
- Developed NLP model for semiconductors, cutting costs by 42% through streamlined ticket management
- Built data science pipeline, streamlining processes and improving decision-making across the organization
- Developed NLP chatbots and led data-mining, forecasting retail sales with predictive analytics
- Spearheaded semantic search, enhancing product relevance by 32% and reducing query latency

Teaching & Mentorship Experience

- **Generative AI Mentor | Analytics Vidhya** (Remote, Gurugram) | Nov 2023 – Nov 2024 - Mentored students in GenAI Pinnacle & AI BlackBelt Plus Programs, implementing state-of-the-art AI models for real-world problems
- **Data Science Instructor | Skillovillia** (Remote, Bengaluru) | Feb 2023 – Apr 2023 - Led hands-on workshops on deep learning & ML, helping students apply data science concepts to industry problems
- **Engineering Mathematics and Statistics Tutor | Self-Employed** (Bhopal) | Jul 2014 – Feb 2017 - Provided tutoring in engineering mathematics & statistics, helping students build strong foundational problem-solving skills.

Academic & Co-Curricular Achievements

- Top 10 Winner of AWS Generative AI Innovation Center Partner Alliance Innovation MCP Hackathon | 2025
- Top 20 Rank Holder at RGPV University for excellent academic performance | 2014

Education

RGPV University

Bachelor of Science in Electrical and Electronics Engineering

June 2010 – June 2014

Bhopal, Madhya Pradesh