Amandeep Singh

Punjab, India
+91 7814986066 amandeepsaroa13@gmail.com Portfolio Website LinkedIn

Education

Lovely Professional University **2021 – 2024**
Bachelor of Computer Application *Jalandhar, India*

Professional Experience

ResoluteAI Software **May 2025 – Present**
AI Engineer - Full Time *Bengaluru (Remote)*

- Built and deployed client-facing **GenAI and Agentic AI** apps using FastAPI, **LangGraph, CrewAI, Azure**, MongoDB, and Docker.
- Fine-tuned LLMs and developed predictive models for domain-specific use cases integrated with **ERP systems like SAP**.

ResoluteAI Software **Oct 2024 – Apr 2025**
AI Solution Engineer - Internship *Bengaluru (Remote)*

- Contributed to **client-focused AI projects** optimizing image classification and object detection models.
- Fine-tuned **computer vision** models like **YOLO** on custom datasets to improve KPI performance.

Projects

Scalable Multi-Model GenAI System for Structured and Unstructured Data with Agentic Workflow

- Built a **production-grade** multi-model RAG pipeline handling **PDFs, images, tables** with modular agent workflows.
- Integrated **LlamaIndex, Pinecone, PostgreSQL, MongoDB,** for hybrid semantic retrieval across structured and unstructured data.
- **Outcome:** Processed over **50K documents** with **sub-6-second response time** and high retrieval accuracy.

Database Migration Automation using Agentic AI | *SQLAlchemy, PostgreSQL, MongoDB, Docker, LangChain*

- Built a **multi-agent system** to automate **database migration from MongoDB to PostgreSQL**.
- Built specialized agents to analyze **schemas, plan migrations**, validate data, and monitor performance.
- **Outcome:** Migrated **50+ datasets with 98% data integrity** across PostgreSQL and MongoDB.

Agentic AI CRM Assistant with Predictive Analysis

- Built an Agentic AI CRM with multi-agent workflows for lead management, predictive follow-ups, and customer insights.
- Integrated XGBoost, SHAP, and RAG for explainable predictions; deployed via Dockerized FastAPI on Azure.
- **Outcome:** Increased lead prediction accuracy by 85% and reduced task time by 70%.

Mixture-of-Experts (MoE) Based Efficient Fine-Tuning for Domain Adaptation | *PyTorch, LoRA*

- Used sparse MoE layers on top of LLMs like **LLaMA** to adapt models for specific domains such as **finance and healthcare.**
- Compared this approach with normal fine-tuning and LoRA, cutting training parameters by 40% with similar accuracy.
- **Outcome:** Improved model **speed by 30%** through optimized expert selection.

Technical Skills

Deep Learning & AI: Generative AI, Computer Vision, NLP, RAG, MoE

Programming: Python, C/C++

Frameworks & Tools: LangChain, CrewAI, PyTorch, Docker, llamaindex, Transformers

Databases & Cloud: MongoDB, PostgreSQL, SAP HANA & ARIBA

Certificates

AWS SAA-C03 (Udemy)
Introduction to Software Engineering (Coursera)
AI/ML for GeoData (ISRO)
Neural Networks and Deep Learning (Coursera)