

Chandr Chuti Chauhan

📍 Bengaluru, India • ☎ +91-9876543210 • 📩 chandr.chut.chauhan@gmail.com

Objective

Data Scientist specializing in LLMs and Conversational AI with 3+ years of experience designing, fine-tuning, and deploying advanced NLP systems. Skilled in LangChain, LangGraph, and RAG pipelines using FAISS, Pinecone, and ChromaDB. Adept at building production-ready AI agents and integrating LLMs into enterprise environments.

Professional Experience

Senior AI Engineer – LLM & Conversational AI

CortexAI Technologies Pvt. Ltd. | Bengaluru | Jan 2023 – Present

- Developed and deployed multi-agent systems using LangChain, LangGraph, and AutoGen, powering virtual assistants for enterprise clients.
- Integrated LlamaIndex for structured and unstructured document retrieval, improving context relevance.
- Fine-tuned GPT-4, LLaMA-2, and Mistral-7B for customer service, HR automation, and legal summarization.
- Designed and optimized RAG pipelines using ChromaDB, FAISS, and Pinecone, reducing latency and boosting retrieval accuracy by 91%.
- Implemented prompt engineering and fine-tuning strategies for domain-specific model enhancement.
- Built RESTful APIs using FastAPI and deployed containerized LLM inference services on AWS using EC2, Lambda, and S3.
- Contributed to internal LLMOps pipelines using MLflow and LangServe for versioning, monitoring, and A/B testing.

Machine Learning Engineer – NLP & ML Systems

DataNova Solutions | Remote | Aug 2021 – Dec 2022

- Trained text classification and NLU models using Hugging Face Transformers, PyTorch, and spaCy.
- Created scalable document retrieval and matching systems using FAISS and sentence-transformers.
- Developed initial versions of customer support chatbots using fine-tuned transformer models and REST APIs.
- Led integration of NLP microservices with web dashboards and mobile apps.
- Contributed to continuous deployment of NLP models using MLflow and Docker.

Key Skills

- **Languages:** Python, SQL
- **LLM Ecosystem:** LangChain, LangGraph, AutoGen, LlamaIndex, LangServe, Hugging Face
- **Transformers:** GPT-4, LLaMA-2, Claude, Falcon, Mistral
- **Vector DBs:** FAISS, Pinecone, ChromaDB, Weaviate
- **ML/DL Frameworks:** PyTorch, TensorFlow, Scikit-learn
- **DevOps/Deployment:** FastAPI, Docker, AWS (EC2, Lambda, S3), GCP
- **MLOps/LLMops:** MLflow, Weights & Biases, GCP Vertex AI
- **Core Concepts:** Prompt Engineering, Fine-Tuning, RAG, Multi-agent Systems, Tool-use, Embeddings
- **Other Tools:** Git, Jupyter, LangSmith, Postman

Education

- **Bachelor of Technology in Computer Science**
- National Institute of Technology (NIT), Trichy | 2017 – 2021
GPA: 8.7/10
- Final Year Project: *LLM-Powered Legal Chatbot using LangChain and FAISS*
- **Projects**
 -  **RAG-based Legal Assistant (LangChain + ChromaDB)**
 - Built a conversational agent to assist lawyers with real-time statute and case retrieval using custom embeddings.
Achieved 92% relevance score in blind evaluations.
 -  **Resume Screening and Matching Bot**
 - Developed a LangChain-powered system to parse, match, and rank resumes against job descriptions.
Used Pinecone for real-time similarity search, reducing screening time by 60%.
Integrated fallback logic with AutoGen and deployed using FastAPI.
 -  **Multi-Agent Workflow Bot (LangGraph + AutoGen)**
 - Created a multi-agent task management assistant with tool-use capability and fallback mechanisms.
Integrated with Google Calendar, Notion API, and internal CRMs.
 -  **Domain-Specific LLM Evaluation Framework (LangChain + LangGraph)**
 - Built an auto-evaluation pipeline for LLM outputs using rubric-based grading, hallucination detection, and feedback generation.
Used LangSmith for trace logging and debugging.

Certifications & Awards

- AWS Certified Machine Learning – Specialty
- LangChain Developer Bootcamp – Built production-ready LangChain agents with integrations
- Hugging Face NLP Practitioner Certificate
- MLOps with GCP Vertex AI – Google Cloud
- Hands-on FAISS and Vector DBs – DeepLearning.ai / Coursera
- 1st Place – AI4Gov Hackathon (LLM Chatbot for Government Schemes)
- Research Paper: “*Fine-tuning LLMs for RAG Workflows*” – Presented at IndiaML 2024