

Karan Baid

Jaipur, Rajasthan · karanbaid1964@gmail.com · +91 8440019114 · GitHub · LinkedIn

EDUCATION

Vellore Institute of Technology
B.Tech in Computer Science

October 2022 - Present

WORK EXPERIENCE

Netgraph Networking Pvt. Ltd., *Machine Learning Intern*

On-site

- Built and deployed a personalized recommendation system for financial products using ML pipelines.
- Applied core machine learning concepts including data preprocessing, model training, and evaluation.
- Gained exposure to advanced AI/ML techniques and developed insights into fintech-specific use cases and data challenges.
- Collaborated with engineering team to align AI outputs with real-world fintech product structures.
- Gained hands-on experience in domain-driven ML solutioning, handling real user engagement data.

TECHNICAL PROFICIENCIES

Programming Languages:

Python,C++

Frameworks & APIs:

Flask, FastAPI , CrewAI , Langchain , Langgraph

MLOPS:

MLFlow, Docker , Airflow ,Dagshub , Grafana

Databases:

MySQL, PostgreSQL, Chroma , PineCone, FAISS

PROJECTS

Cyber-Saarthi - Fine-Tuned Legal AI Chatbot
GitHub Repository [↗](#)

- Fine-tuned TinyLlama-1.1B using QLoRA** on custom Indian cyber law dataset (544+ Q&A pairs), achieving **60.2%** accuracy and **93%** reduction in evaluation loss with 4-bit quantization.
- Built complete ML pipeline:** automated dataset generation with 10+ query variations per topic, parameter-efficient fine-tuning with LoRA adapters, and full-stack deployment.
- Implemented production features:** conversation memory, streaming responses, and domain-specialized chatbot achieving zero hallucinations on legal queries.

Tech Stack: Python, Hugging Face Transformers, PEFT, QLoRA, BitsAndBytes, PyTorch, Streamlit

News Analysis Crew
GitHub Repository [↗](#)

- Engineered a multi-agent AI system** using the CrewAI framework to automate the critical analysis of online news articles, simulating a professional research workflow.
- Developed a specialized "Web Content Extractor" agent** with a custom Python tool using BeautifulSoup to parse HTML and isolate core article text from web page clutter.
- Designed a second "AI Critical Thinking Partner" agent** tasked with performing nuanced analysis, including identifying core claims, tone, potential biases, and generating verification questions.
- Orchestrated the agent collaboration** in a sequential process where the extractor agent retrieves content and passes it to the analyst agent for in-depth reporting and final output generation.

Tech Stack: Python, CrewAI, Pydantic, BeautifulSoup, YAML, LangChain

RAG chatbot
GitHub Repository [↗](#)

- Developed an advanced RAG system** implementing hybrid retrieval combining BM25 and vector semantic search, enhanced with multi-query generation and Cohere reranking for superior document retrieval accuracy.
- Engineered multi-format document processing pipeline** leveraging Unstructured library to handle PDFs, Word documents, images, and text files with rich content extraction capabilities.
- Built production-ready conversational interface** with Reciprocal Rank Fusion, persistent vector storage, session management, and source citation tracking for context-aware multi-turn conversations.

Tech Stack: Python, LangChain, Streamlit, Groq API, Cohere, ChromaDB, Unstructured,

CERTIFICATIONS

Gen-AI Certification by IBM
Verification Link [↗](#)

Python for Data Science and Machine Learning Bootcamp by Jose Portilla
Verification Link [↗](#)

Complete Generative AI course with Langchain and Huggingface
Verification Link [↗](#)