Sarabjot Singh

sarabjot.dev

+91 6200342799 sarabjot.singh799@gmail.com linkedin.com/in/sarabjot github.com/ricky-aufvaa

Summary

AI Developer with hands-on experience in building GenAI solutions using LLMs, vector databases, and full-stack tools. Delivered production-level applications at Ashok Leyland, optimizing sales and engineering operations using agentic workflows, OCR pipelines, and LLM orchestration.

Education

Chandigarh Engineering College, Landran, Punjab
B.E. in Electronics and Communication Engineering — Sep 2020 – Sep 2024

Experience

Ashok Leyland, Chennai — Jul 2024 – Present
Deputy Manager

- Built a suite of 3 GenAI applications for Engineers, Designers, and Functional Heads using LangChain, LlamaIndex, and AWS, improving document search and analytics.
- Enhanced retrieval workflows by fine-tuning LLaMA-3 8B model on proprietary automotive manuals, improving engineering knowledge access and technical query resolution.
- Developed an agentic sales assistant using LangGraph and Claude Sonnet via Bedrock to recommend vehicles based on customer requirements (e.g., sand transportation), automating brochure delivery, TCO calculations, and product matching—dramatically reducing manual lookup time for sales executives.
- Engineered OCR pipelines to extract structured data from scanned documents, improving accessibility and enabling AI-based document querying.
- Stack: Python, LangChain, LlamaIndex, Streamlit, ChromaDB, Neo4j, AWS, Flask, OpenAI API, Claude Sonnet, LLMSherpa, Bedrock.

Projects

WebCraft AI — Live Link
Next.js, React, Convex, Claude Sonnet 3.5, OAuth, PayPal, Sandpack

- Built a GenAI-powered developer platform that generates UI code from prompts using Claude Sonnet 3.5, with live preview/editing using Sandpack.
- Implemented Google OAuth authentication and PayPal integration for future monetization with tiered pricing plans.
- Created an interactive coding workspace with AI chat assistance and persistent project storage using Convex.

DSCR Loan Assistant — GitHub
LangGraph, FastAPI, Streamlit, Google Gemini, FAISS, marker, Sentence Transformers

- Developed an AI-agentic web application for real estate investors to validate loan applications and query DSCR loan guidelines using business rules and RAG.
- Built a LangGraph-powered agent that routes queries to either a finance Q&A or loan validation workflow based on input type (natural language vs JSON).
- Used `marker` for high-fidelity PDF text extraction and FAISS for efficient vector storage, improving retrieval accuracy for regulatory content.
- Designed structured prompting and rule-based post-processing to validate loan eligibility with explainable, auditable outputs.

Technical Skills

- **Languages:** Python, C++, JavaScript, SQL
- **Web/Frameworks:** React, Next.js, Flask, FastAPI, Streamlit
- **AI/ML:** Transformers, TensorFlow, Keras, LangChain, LlamaIndex, LangGraph, GenAI, NLP
- **Cloud/Tools:** AWS (EC2, S3), GCP, Azure Data Factory, Docker, Git
- **Databases:** MySQL, ChromaDB, Neo4j