Shorthills

Kushagra Wadhwa

SUMMARY

AI Engineer with a strong focus on building real-world, scalable Generative AI solutions. Experienced in developing and deploying production-grade LLM applications, RAG pipelines using tools like LangChain, LlamaIndex, Pinecone, and workflow automation systems using tools like n8n. Adept at translating cutting-edge AI research into practical applications that drive business value, improve information retrieval, and enhance user experiences. Proven track record across NLP, document intelligence, automations workflows and RAG based systems, with hands-on experience delivering AI systems in startup environment.

SKILLS

- Python
- Flask
- FastAPI
- REST APIs
- Hugging Face
- Langchain
- LlamaIndex
- LLMs
- RAG Pipelines
- Vector Databases (Pinecone, ChromaDB)
- NLP
- Machine Learning
- SQL
- MongoDb
- HTML
- CSS
- Streamlit
- Agile Development
- Team Collaboration
- Problem Solving
- Adaptability
- Project Management

EDUCATION

- B. Tech Artificial Intelligence and Machine Learning , **Guru Gobind Singh Indraprastha University**

CERTIFICATIONS

- GATE DA 2025
- Diploma in Programming and Data Science, **IIT Madras**
- Diploma in Data Science
- Introduction to Generative AI, **Google**
- Postman API Fundamental Student Expert, **Postman**

KEY RESPONSIBILITIES

Project 1: Job Description (JD) based Resume Evaluation Retailoring Pipeline

- Automated HR Workflow (n8n + LLMs).
- Built an end-to-end automation pipeline using n8n to score candidate resumes against job descriptions and dynamically tailor profiles to align with role requirements.
- Resulted in significant reduction of manual intervention by 90% and improved personalization in outbound applications.

Project 2: Project X — RAG-Based Multimodal Content Retriever

- Built a topic-aware content retrieval engine that intelligently matched user queries with relevant videos and articles using a dual-namespace RAG pipeline.
- Achieved fast and accurate delivery of top 3 videos and 1 article per query, significantly improving user content discovery across formats.

Project 3: Sales-Assist AI — AI-Powered Conversational Assistant

- Developed a production-grade LLM-based virtual assistant using Retrieval-Augmented Generation (RAG) to improve real-time sales support with accurate, context-aware answers.
- Integrated OpenAI Whisper and Gemini LLM APIs for dynamic conversational capabilities.