

HAMZA JAMAL KHAN

- New Delhi, India, PIN: 110019 • hjskhan47@gmail.com • +91 9456615644
- Portfolio: hjskhan.github.io • LinkedIn: linkedin.com/in/hjskhan • GitHub: github.com/hjskhan

ABOUT

AI Engineer specializing in GenAI, GANs, LLMs, Chat bots, NLP, Computer Vision, Machine Learning, Deep Learning, MLOps and Data Analytics. Experience in developing RAG model chatbots. Key project Legal GPT chatbot based on OpenAI API. Proficient in deploying models through CICD pipelines, Azure, AWS, Docker and Flask. Skilled in Python, LLM model development, and deployment, with a commitment to continuous learning and growth

EDUCATION

MSC. DATA SCIENCE <i>Aligarh Muslim University (AMU)</i>	Aligarh, India 2022-2024
B.TECH MECHANICAL ENGINEERING <i>Aligarh Muslim University (AMU)</i>	Aligarh, India 2017-2021

PROFESSIONAL EXPERIENCE

TECHENHANCE IT SERVICES PVT LTD Artificial Intelligence Engineer	Bengaluru, India May-2024 –Present
<ul style="list-style-type: none">• Developed a 'Legal DST' LLM chatbot designed for legal professionals to efficiently search court judgments and related legal documents.<ul style="list-style-type: none">○ Tech Stack: OpenAI API, Assistant API, Azure Services, API development, Node JS○ Enhanced the initial architecture proposed to ensure optimal performance and usability which lead to reduction in running costs of application up to 25%.• Developed OCR based application for automating invoice's details extraction using Multimodal LLM<ul style="list-style-type: none">○ Tech Stack: OpenAI GPT 4, Gemini Vision Pro, Azure, Flask○ Enabled efficient data extraction, reducing manual effort and errors by leveraging advanced AI for precise and rapid processing of large volumes of invoices	
Artificial Intelligence Intern	Feb-2024 –May-2024
<ul style="list-style-type: none">• Developed a chatbot with RAG model using OpenAI API and Gemini API.<ul style="list-style-type: none">○ Tech Stack: Azure AI Search, OpenAI API, Gemini API, Astra DB (Vector DB), Flask, HTML, CSS, JS.○ Executed seamless CICD deployment using Azure Web App for efficient integration.• Fine-tuned Open Source Models with minimal parameter (2B-8B) on custom data and built Flask application.<ul style="list-style-type: none">○ Tech Stack: HuggingFace, Llama, Azure, AWS, QLoRA, Unsloth, Gemma, Mixtral, Falcon, Flask.○ Fine-tuned Open Source models for less compute resource utilization and low running cost.	
INZANE LABS Machine Learning Intern	Mumbai, India Sep-2023 – Dec-2023
<ul style="list-style-type: none">• Designed and used various trading strategies (HFT), making predictions more accurate.• Made data work at least 20% faster using Pandas and Numpy.• Benefitted the company by improving algorithmic trading capabilities.	

SKILLS

AI AND MACHINE LEARNING: LLM, GenAI, Audio Transcription, Computer Vision, Chat Bot, RAG, GANs, Deep Learning, NLP, Neural Networks, Time Series Analysis, MLOps, CICD Pipelines, Data Analytics.

TOOLS: Azure, AWS, Heroku, Docker, Flask, Stream Lit, FastAPI, LangChain, Hugging Face, OpenAI, MySQL, Astra DB, GitHub Actions, Keras, Tensor Flow, PyTorch, OpenCV, Power BI, Postman

PROGRAMMING LANGUAGES: Python, R, HTML, CSS, Node JS