Ashutosh Roy

AI/ML Engineer

+91-9810599837 | ashu2003roy@gmail.com | linkedin.com/in/ashutosh-roy | github.com/ashutosh-roy | Gurugram, Haryana

Summary

AI/ML Engineer with hands-on experience building end-to-end LLM, RAG, and speech AI systems. Published researcher (RECCAP 2026, IIT Palakkad) with a co-authored paper from IIT Delhi. Skilled in fine-tuning transformer models, designing hybrid retrieval pipelines, and deploying production ML systems via FastAPI. Focused on building scalable, high-accuracy AI solutions from research prototype to production.

Experience

Research Intern **Jan 2025 – Jun 2025**
Indian Institute of Technology (IIT) Delhi *New Delhi, India*

- Built an OCR-powered document understanding pipeline for archaeological datasets; benchmarked Tesseract, Marker, and Gemini OCR — selecting optimal strategy per document class to improve extraction accuracy.
- Designed hybrid retrieval (dense: FAISS/Pinecone + sparse: Apache Solr) for high-precision semantic + lexical document search over thousands of historical records.
- Deployed scalable **FastAPI** services for real-time LLM-based querying; integrated with frontend for sub-second retrieval responses.
- Co-authored **SARCH: Multimodal Search for Archaeological Archives** with IIT Delhi collaborators — submitted for peer review. *arxiv.org/abs/2511.05667*

AI/ML Engineer Intern **Oct 2024 – Nov 2024**
MitoVoid AI *Remote*

- Fine-tuned transformer-based NLP models (Mistral/LLaMA variants) for healthcare applications — medical record summarization and symptom-driven conversational AI.
- Reduced patient query resolution time by **˜40%** by deploying an AI chatbot for telemedicine with optimized intent understanding and response generation.
- Built end-to-end ML pipelines: data preprocessing → model inference → API deployment for real-time healthcare applications.

Projects

Speech Emotion & Stress Detection | *PyTorch, Wav2Vec2, BiLSTM, Hugging Face* | ***RECCAP 2026*** Aug 2024 – Present

- Proposed a **self-supervised** Wav2Vec2 + BiLSTM multitask architecture for simultaneous 5-class emotion classification and continuous stress regression; achieved **85.9% accuracy**, **0.856 F1**, RMSE **0.1268**.
- Paper **accepted at RECCAP 2026, IIT Palakkad**; evaluated across 3 corpora (RAVDESS, TESS, SAVEE) with unified preprocessing (VAD, silence removal, log-Mel spectrograms).

Exam-Helper RAG System | *Python, LLaMA 3, FAISS, Pinecone, LangChain, Streamlit* Dec 2024 – Apr 2025

- Built an LLM-powered contextual QA system on **LLaMA 3** with RAG; multimodal ingestion pipeline supports PDFs, images, and YouTube transcripts.
- Implemented **dense retrieval** via FAISS/Pinecone for sub-second similarity search, reducing hallucination and improving context precision.

Medical Chatbot (LLM Fine-Tuning) | *Python, Mistral 7B, QLoRA, Tesseract OCR* Mar 2024 – Apr 2024

- Fine-tuned **Mistral 7B** via **QLoRA** (4-bit quantization) for domain-specific medical NLU; reduced compute cost vs. full fine-tuning while preserving task accuracy.
- Built OCR pipeline using Tesseract to extract and structure information from unstructured clinical reports for downstream LLM inference.

Music Identification App | *Python, ACRCloud API, Streamlit, Flutter* Sep 2024 – Dec 2024

- Built real-time audio recognition pipeline: audio capture → feature extraction → ACRCloud API fingerprinting → metadata retrieval; cross-platform UI via Streamlit and Flutter.

BulkyMail | *Python, Streamlit, SMTP, Jinja2* Dec 2024

- Built a Streamlit-based bulk email automation tool with dynamic templating for personalized large-scale outreach (name, role, company fields); integrated SMTP APIs for reliable delivery of 100+ emails per run.

Technical Skills

AI/ML & Research: LLMs, RAG, Transformer Fine-Tuning (QLoRA/LoRA), Self-Supervised Learning, Multimodal AI, Speech Processing, NLP, Evaluation & Benchmarking, Dense Retrieval, Foundation Models
Libraries & Frameworks: PyTorch, Hugging Face, LangChain, TensorFlow, Scikit-learn, Pandas, NumPy, FastAPI, Streamlit
Databases & Search: FAISS, Pinecone, Apache Solr, MySQL
Languages: Python, C++, SQL
Tools: Git, Docker, Linux, Jupyter Notebook, Google Colab, VS Code

Education

Chhattisgarh Swami Vivekananda Technical University (CSVTU) **2022 – 2026**
B.Tech (Hons) in Computer Science and Engineering (Artificial Intelligence) *Bhilai, Chhattisgarh*

St. Columbus School **2020 – 2021**
Higher Secondary (Class XII - Science PCMB) - 87% *Faridabad, Haryana*