

ASHUTOSH ROY

Gurugram, Haryana, India

+91-9810599837 ashu2003roy@gmail.com ashutosh-roy ashutoshroy02

SUMMARY

Experienced in Machine Learning, NLP, and LLM-based systems, with a focus on building scalable, production-ready AI solutions. Worked on OCR and Retrieval-Augmented Generation (RAG) pipelines at IIT Delhi for large-scale document understanding. Proficient in Python, PyTorch, Hugging Face, OpenCV, and FastAPI, with strong experience in developing, optimizing, and deploying end-to-end AI systems using vector databases like FAISS and Pinecone.

EDUCATION

CSVTU, Bhilai, Chhattisgarh Nov 2022 – May 2026
BTech (Hons) Computer Science and Engineering in Artificial Intelligence

St. Columbus School Apr 2020 – Mar 2021
12th (Science PCMB) - 87%

St. Columbus School Apr 2018 – Mar 2019
10th - 93.8%

EXPERIENCE

Research Intern | IIT Delhi Jan 2025 – Jun 2025

- Developed OCR-powered document understanding pipeline integrated with RAG for archaeological datasets.
- Implemented hybrid search using FAISS/Pinecone and Apache Solr (semantic + keyword retrieval).
- Benchmarked OCR systems (Tesseract, Marker, Gemini OCR), improving extraction accuracy.
- Built scalable APIs using FastAPI for search, retrieval, and LLM-based querying.

Software Engineer Intern | LeanTactics Solutions Pvt. Ltd. (Remote) Nov 2024 – Dec 2024

- Designed and deployed backend services using Python and FastAPI for AI-driven applications.
- Integrated ML inference pipelines into production systems.
- Optimized API performance, reducing response latency by 25–35%.

AI/ML Engineer Intern | MitoVoid AI (Remote) Oct 2024 – Nov 2024

- Fine-tuned transformer models (Mistral/LLaMA) for healthcare NLP applications.
- Built AI chatbot improving patient query resolution by 40%.
- Developed end-to-end ML pipelines (data → model → API deployment).

PROJECTS

Speech Emotion & Stress Detection Jan 2026 – Apr 2026

- Built ML/DL models using MFCC, pitch, and energy features.
- Applied SVM, Random Forest, CNN, and wav2vec2 embeddings.
- Improved performance using augmentation and feature engineering.

Archaeological Search Engine Jan 2025 – Jun 2025

- Built OCR + ML pipeline for extracting structured data from scanned documents.
- Enabled semantic search using embeddings and hybrid retrieval.

Exam-Helper RAG Dec 2024 - April 2025

- Built LLM-based study assistant using LLaMA 3 and RAG architecture.
- Integrated FAISS/Pinecone for efficient retrieval.
- Enabled multimodal querying (PDF, images, YouTube).

Medical Chatbot (LLM) Mar 2024 – Apr 2024

- Fine-tuned Mistral 7B using QLoRA for domain-specific NLP tasks.
- Integrated OCR pipeline for extracting insights from medical reports.

Music Identification App Sep 2024 - Dec 2024

- Built real-time audio recognition system using ACRCLOUD API.
- Designed pipeline: audio → feature extraction → matching → metadata retrieval.

PUBLICATIONS

- **SARCH: Multimodal Search for Archaeological Archives** – Co-authored with IIT Delhi. Developed multimodal retrieval system using OCR + image understanding + hybrid search. [↗](#)
- **Multitask Speech Emotion & Stress Detection (Wav2Vec2-BiLSTM)** – Accepted at RECAAP 2026 (IIT Palakkad). Achieved 85.9% accuracy for emotion + stress prediction from speech.

TECHNICAL SKILLS

Languages: Python, C++, MATLAB

ML/AI: Machine Learning, NLP, Computer Vision, Transformers, RAG, LLMs

Libraries & Frameworks: PyTorch, TensorFlow, Hugging Face, OpenCV, scikit-learn

Tools & Platforms: FastAPI, Docker, Streamlit, FAISS, Pinecone, Git, Power BI, Flutter

Developer Tools: GitHub, Google Colab, VS Code, Jupyter Notebook

ACHIEVEMENTS

- **First Runner-up** – Data Analytics Hackathon, IIM Nagpur
- **Strong problem-solving foundation** with 100+ DSA problems solved across core patterns
- **EF SET English Certificate** – 71/100 (C2 Proficient), Issued Jun 2025