Kaushal Sengupta

Uttarpara, West Bengal
(+91) 8132859752 | kaushalsubho2005@gmailcom | linkedin.com/in/kaushal-sengupta-2b4277256/ | github.com/kforkool11234

EDUCATION

Vellore Institute of Technology, Bhopal — **September 2022 – Ongoing**
B.Tech in Computer Science and Engineering — ***CGPA: 8.22/10***
Central Board of Secondary Education — **May 2022**
Senior Secondary Education — ***Percentage: 83.2%***
Central Board of Secondary Education — **May 2020**
Secondary Education — ***Percentage: 87.4%***

SKILLS

Python, C++, JavaScript, React.js, Node.js, Tailwind CSS, MySQL, MongoDB, AWS, Git/GitHub, Redis, Docker, PyTorch, OpenCV, TensorFlow, Generative AI

PROJECTS

RAG Assistant Platform | Python, Django, Sentence Transformers, OpenCLIP, ChromaDB

- Developed a Retrieval-Augmented Generation (RAG) system supporting PDF, DOCX, PPTX, and TXT for semantic search and Q&A.
- Integrated OpenCLIP for multimodal embeddings and the Gemini API for context-aware responses, improving response relevance by over 90%.
- Implemented a vectorization pipeline with OpenCLIP and ChromaDB for persistent, low-latency retrieval.
- **Role: Generative AI & Full-Stack Engineer**

Automated Blood Cell Analysis System | Python, TensorFlow, scikit-learn, OpenCV

- Developed a CNN that achieved **95.5%** accuracy for blood cell subtype classification, cutting manual review time by **80%**+.
- Prepared a dataset of **1,200+** microscopic images with preprocessing and augmentation to improve generalization and reduce overfitting.
- Built a Python pipeline with scikit-learn to extract morphological features, improving separation between visually similar classes.
- Deployed a lightweight model with **TensorFlow Lite** for near real-time inference on diagnostic devices.

Research-RAG | Python, Streamlit, Vector Embeddings , ChromaDB

- Developed a RAG pipeline to process PDFs: extracted sections (titles/abstracts), chunked content, and created embeddings (all-MiniLM-L6-v2) for semantic retrieval.
- Indexed vector embeddings and metadata in ChromaDB for low-latency similarity search and hybrid matching (query vector + keyword).
- Integrated Mistral LLM to answer queries grounded in retrieved papers or draft research sections from user input and examples.
- Bypassed token limits with chunking + overlap, top-k dynamic context windows, and iterative summaries to compress long inputs.
- Built a Streamlit UI for upload, search, and generation with a one-command setup (requirements.txt).
- **Role: Generative AI & Data Engineering**

EXTRA-CURRICULAR AND ACHIEVEMENTS

Achievements: Engineered an AI-powered data integration solution utilizing Informatica, securing a Top 20 placement out of 500+ global participants in the international AI/GenAI hackathon based on innovation criteria.
Responsibilities: Part of the VIT Bhopal Cyber-Security Club as a Web-lead, Design Team member for the Rajasthani club

CERTIFICATIONS

OCI Generative AI Professional - *Oracle* — *september 2025*
Oracle Cloud Infrastructure 2025 Certified Data Science Professional - *Oracle* — *september 2025*
Oracle AI Vector Search Associate - *Oracle* — *september 2025*