

Ravi Kumar

+91-6283867492 | ravibh5522@gmail.com | linkedin.com/in/ravibh5522 | github.com/ravibh5522 |

Bhawanipur, Supaul, Bihar 854339

Aspiring Backend & GenAI Developer with experience in building AI-powered applications, fine-tuning LLMs (LLaMA2, Gemini), and developing scalable REST APIs. Proficient in Python, Flask, FastAPI, and cloud services (AWS, GCP, Vercel).

Education

B.E. in Computer Science & Engineering — Chandigarh University 2022 – 2026

12th Govt. +2 High School, Birpur (2019 – 2021)

10th Govt. High School, Birpur (2017 – 2019)

Skills

Languages: Python, C++, Java, Dart, R

Technologies: Flask, FastAPI, REST APIs, Transformers, TensorFlow, Keras, Selenium

AI/ML: LLaMA2, Gemini API, Hugging Face, AI Fine-Tuning, Prompt Engineering, RAG Systems

Cloud/DevOps: AWS (Lambda, EC2, S3, SageMaker), Google Cloud Run, Vercel

Others: Bash, Linux, Shell Scripting

Projects

Multimodal RAG System

Developed a robust **Retrieval Augmented Generation (RAG)** system capable of handling over **20,000 pages of multimodal data** (text and images) extracted from PDFs. The system leverages advanced retrieval techniques (FAISS, BM25) to efficiently retrieve relevant information and generates both **textual** and **image-based responses**. Implemented using Flask and REST APIs for seamless data access and deployed on Google Cloud Run for scalability and reliability.

Tech Stack: Flask, REST APIs, Google Cloud Run, FAISS, BM25, Python

Automated News Research Tool

Created an automated tool to scrape news articles from Google and various other news websites using **Selenium**. Integrated a **fine-tuned LLaMA2** model to efficiently summarize the collected news, handling dynamic content and ensuring efficient data extraction and processing.

Tech Stack: Python, Selenium, LLaMA2

AI-Assisted Music Production Web App

Developed a web application that generates song lyrics based on user-defined parameters such as language, genre, and song description. Leveraged the Gemini API for lyric generation via a REST service. Deployed the application on Vercel for accessibility and scalability.

Tech Stack: Flutter, Gemini API, Vercel, REST APIs

AI Text-to-Voice Assistant API

Designed and implemented a complete pipeline for a text-to-voice assistant API using Flask. The system incorporates custom embeddings and fine-tuned models to enhance voice synthesis quality and accuracy.

Tech Stack: Python, Flask, AI Model Fine-Tuning, REST APIs

P2P File Sharing App

Description: Built a peer-to-peer (P2P) file-sharing application with a Python server backend hosted on Vercel. Each user is assigned a unique 8-digit ID to facilitate direct file transfers. The application incorporates robust error handling and deadlock prevention mechanisms.

Tech Stack: Python, Flutter, Vercel

Experience

Freelance GenAI & Backend Developer (*Remote / Jan 2024 – Present*)

- Delivered 10+ AI/Backend projects for 6+ clients.
- Built and deployed REST APIs, RAG systems, and automation tools using Flask, FastAPI, and cloud platforms.
- Specialized in integrating LLMs for document understanding, summarization, and user-facing apps.