

# Himanshu Parihar

+91-9319264679    himanshuparihar853@gmail.com    LinkedIn    Github    Leetcode

## Experience

### HiVoco Content-Tech Studios

Nov 2023 – Present

ML Engineer – OpenAI, LangChain, Vector DBs, AWS

Delhi, India

- Developed **RAG-based applications**, including a **saffron-related chatbot** and a **research paper assistant**, leveraging **LangChain, OpenAI, and vector databases** to improve query resolution and content understanding.
- Built a voice-to-voice AI assistant for customer support, integrating STT, TTS, and RAG using OpenAI and open-source LLMs, reducing LLM costs by **50%** while maintaining high response accuracy.
- Engineered a **scalable voice-based quiz campaign** serving **500,000+ users**, optimizing performance with **Groq API, Flask, and FassiDB**.
- Led the development of **"Closeup Love Tune"**, a **personalized video generation platform** for Valentine's Day, utilizing **audio and video synthesis techniques**.
- Currently working on **"Razer Sharp"**, a **video and audio analytics platform** for QSRs and retail, leveraging **computer vision and NLP** to extract business insights.
- Designed **CI/CD pipelines** using **GitHub Actions**, deployed applications on **AWS (EC2, ECR, S3)**, and managed scalable databases (**MySQL, MongoDB**).

## Projects

### Voice-to-Voice AI Assistant (STT → RAG → TTS) | Python (Flask), OpenAI, LangChain, Vector DB, Groq API

Github

- Developed an interactive voice assistant that processes speech input, retrieves relevant responses via **RAG**, and generates natural voice replies using TTS.
- Integrated **OpenAI, Groq API, and vector databases** for accurate contextual understanding.
- Enabled customizable backends, allowing seamless integration with different speech APIs and LLMs.
- Deployed on AWS EC2**, ensuring scalability and availability for real-time interactions

### Chatbot for Webpage | Python, Flask, FAISS, HuggingFace, Chrome Extension

Github

- Developed a Chrome extension that enables users to interact with any webpage via a chatbot.
- Implemented automated web scraping to extract and preprocess text from URLs.
- Used HuggingFace embeddings and FAISS for efficient vector-based search.
- Built a REST API for querying webpage content in natural language.

## Technical Skills

**Programming Languages:** Python, SQL

**Data Validation and Processing:** Pydantic, Pandas, NumPy

**Databases:** SQL, NoSQL, Vector DBs (FAISS, Pinecone, ChromaDB, FassiDB)

**Version Control:** Git, GitHub

**Cloud and Deployment:** AWS (EC2, S3, ECR, RDS), Docker, GitHub Actions

**Frameworks and Libraries:** LangChain, OpenCV, PyTorch, Scikit-learn, flask

**LLMs and RAG:** OpenAI API, Groq API, Custom Retrieval Augmented Generation Pipelines

## Certifications

- HackerRank Advanced SQL Certification
- Data Science Certification

## Education

Delhi University B.Sc (Hons.) Mathematics with Computer Science

July 2020 - June 2023

CGPA: 7.73