

Assignment #09

Hope to Skills

Free Artificial Intelligence Advance Course

Instructor: Irfan Malik, Dr. Sheraz

Submission:

- Create a new **python file** and create a Streamlit application.
- Submit a .zip file which contains the **.py** and **requirements.txt** files. No other format will be accepted.
- Submission file should be named as **Assignment_09_StudentName.zip**
- Deadline for this Assignment is **Friday 20-09-2024**.
- Strictly follow the submission deadline.
- Make Submission in the **Assignment-09** Google Form and press the submit button.
- Click [here](#) to submit the Assignment

What you will learn

- How to build a chatbot application from scratch using Streamlit.
- How to integrate and use any Large Language Model (LLM) API keys with Langchain.
- How to use retrieval-augmented generation (RAG) to generate more accurate and contextually relevant chatbot responses.
- Implement audio input/output handling in Streamlit and use text-to-speech (TTS) and speech-to-text (STT) functionalities for Urdu.
- Basics of conversational design using modern AI tools.
- Use of vector database

Solve the Following Task

Question 1: Create an intelligent chatbot using Streamlit and Langchain (RAG), where the chatbot can receive voice input in Urdu, process it, and return both text and audio responses in Urdu. The chatbot should be able to interact with users fluently, allowing for seamless audio-to-text and text-to-audio communication in the Urdu language.

Workflow:

- Build the Streamlit interface for real-time Urdu audio input and output.
- Integrate **Langchain** (RAG) with an LLM (Language Model) API to generate dynamic responses based on the user's input. (use PDF files only)
- Ensure the chatbot responds not only with a text-based answer in Urdu but also converts that response back to audio and plays it for the user.

Instructions:

In this assignment, you are required to use one of the following models for building your chatbot:

- **Claude (Anthropic):** You can use the Claude Haiku model, which is lightweight and optimized for fast responses. Claude provides \$5 worth of free credit for new users to explore their API. Visit [Claude AI](#) for more details.
- **AI21 Labs (Jamba 1.5 Mini):** This model is designed for efficient, lightweight inference tasks. AI21 Labs offers \$10 free credits for three months, allowing you to experiment without a credit card. You can learn more at [AI21 Labs](#).
- **Google Gemini API:** Google provides the **Gemini 1.5 Flash** model, which is highly optimized for multimodal tasks such as text, images, and audio. For details, visit [Google Gemini API](#).
- **OpenAI GPT (OpenAI API):** OpenAI offers **GPT-4o mini**, a smaller and more cost-effective model that is ideal for chatbot tasks. Check out [OpenAI's pricing](#) for more details.
-

You can choose either model for this project to handle the chatbot's generative tasks effectively.

Note:

This assignment was actually announced by Sir Irfan Malik in the **38th lecture of Free AI Advance Course (on 5th September, 2024)**. Now we are sharing the form with you to submit the assignment.

You can find the lecture here: <https://www.youtube.com/live/CzeLannZJ-E?si=hvg-xilyN8iZPeuv>