

USE CASE 1: Voice AI Agent for Telephonic Screening of Candidates

Scenario:

Design an automated calling agent that is capable of asking general screening questions (e.g., basic introduction, skillset, experience, current location, willingness to join, etc.).

Tasks:

- Make an automated call using the HR phone number for real-time conversation with a candidate.
- Transcribe audio (use Whisper / Hugging Face).
- Run sentiment & keyword analysis.

Agent decides:

- Are there tone issues?
- Is the candidate handling queries well?
- Should a better closing line be used or escalate to human HR?

Deliver insights as text (or simulated UI card)

Deliverables:

- Transcripts + NLP pipeline.
- Agent reasoning flow (how decisions are made).
- Sample feedback snippet.

Agentic Element:

The agent transcribes, analyzes, reasons, and screens in the candidates based on their responses.

OR

USE CASE 2: AI-Powered Video Knowledge Assistant

Scenario:

Design an AI assistant that processes a YouTube video (e.g., lecture, webinar, tutorial) and enables users to ask questions about its content. The assistant should transcribe the video, generate a concise summary, and provide citations with timestamps.

(*Knowledge base given below)

Tasks:

- Download video audio (via pytube or YouTube API).
- Transcribe using Whisper / Hugging Face ASR.
- Segment transcript into sections and the video frames into captions.
- Generate an overall video summary covering the main themes.
- Create embeddings with timestamp metadata.
- Build QA pipeline referencing timestamps.

Agent decides:

- Which transcript segments or frame caption are relevant to the query.
- When to give short vs detailed answers.
- Which timestamps to cite.
- When to suggest watching a full segment.

Deliverables:

- Transcript and fram captions with timestamps.
- Video summary.
- QA system with timestamped citations.

- Demo UI (Streamlit or similar).

Agentic Element:

The agent transcribes and captions the frame, embeds, retrieves, and answers user questions with relevant video segments and timestamp citations.

OR

USE CASE 3: Intelligent PDF Summarizer & QA Bot

Scenario:

Design an AI tool that ingests a PDF (e.g., report, research paper) and outputs a structured summary with the ability to answer detailed questions based on the document.

(*Knowledge base given below)

Tasks:

- Parse PDF text (PyMuPDF / pdfplumber).
- OCR scanned pages if needed / PDF page content extraction.
- Chunk content and generate embeddings.
- Produce a hierarchical summary.
- Build retrieval-augmented QA pipeline.
- Add page-number citations in answers.

Agent decides:

- Which sections to include in the summary.
- When to describe diagrams/tables.
- Which pages to cite.
- How much detail to give for different queries.

Deliverables:

- Extracted text & OCR output.
- Structured summary.
- QA interface (Streamlit or similar).
- Sample queries with page-cited answers.

Agentic Element:

The agent parses, summarizes, indexes, and retrieves from the PDF, giving context-grounded answers with page-level citations.

Final Deliverables (CHOOSE ANY ONE CASE)

Item

Description

Code

Collab link or GitHub repo with README

Agent Flow Diagram or Program Workflow flowchart

PNG or draw.io or Figma diagram

Explanation (PDF/Slides)

One-pager explaining agent use case & impact

Output Screenshots
Sample inputs/outputs of the agent
(Optional) Demo Video
Phone/Screen-recorded demo

****Tools Allowed but not limited feel free to choose others :**

Python (LangChain, ChromaDB, FAISS, OpenAI API, Anthropic Api, PyMuPDF, pdfplumber, Whisper, Hugging Face ASR, sentence-transformers)
Speech: Whisper / Google Speech-to-Text
SQL (for bonus tasks / embedded search layer)
Streamlit / Gradio / FastAPI for simple frontends
AutoGPT/ CrewAI/ LangGraph (for advanced submissions)
Visualization: Notion, Figma, Canva, Mermaid, Figma, Canva, Notion, Tesseract OCR, Google Cloud Vision OCR

How to Submit:

Google Drive /GitHub link with: [While Sharing pls make it accessible]
Code
Architecture
Demo/video if any

Google Drive Link for Knowledge base:

<https://drive.google.com/drive/folders/1dUwu2HFej594NUiWvFBcUWDmvi1qmAig?usp=sharing>