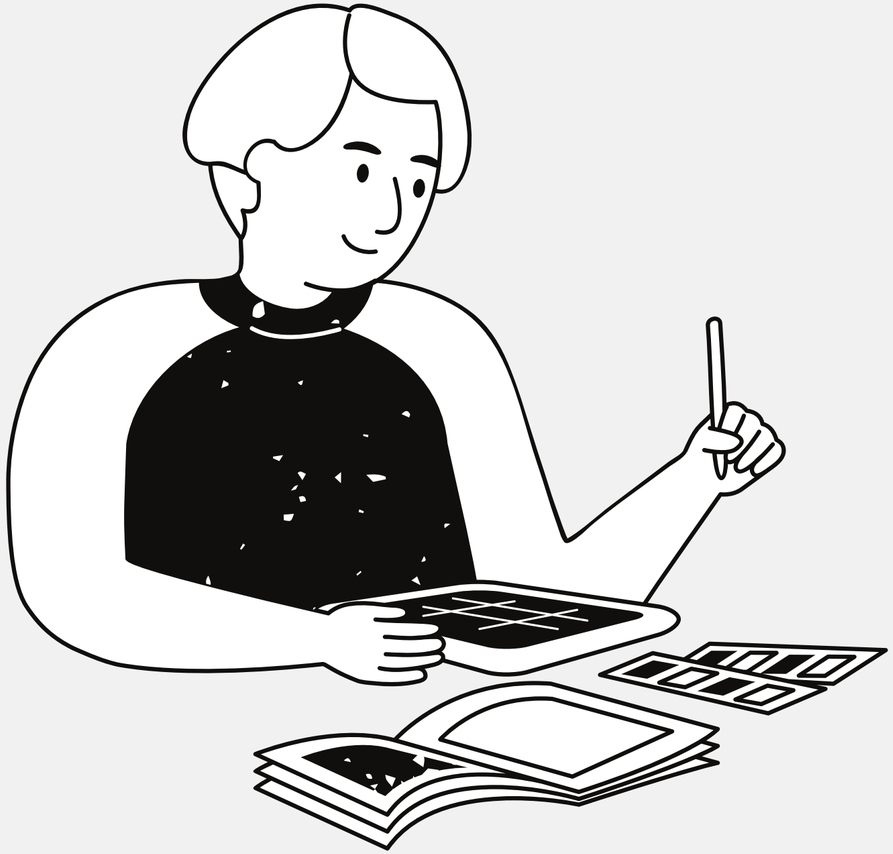
Visual & Hearing Assistance System

Prepared by Aditya Aman

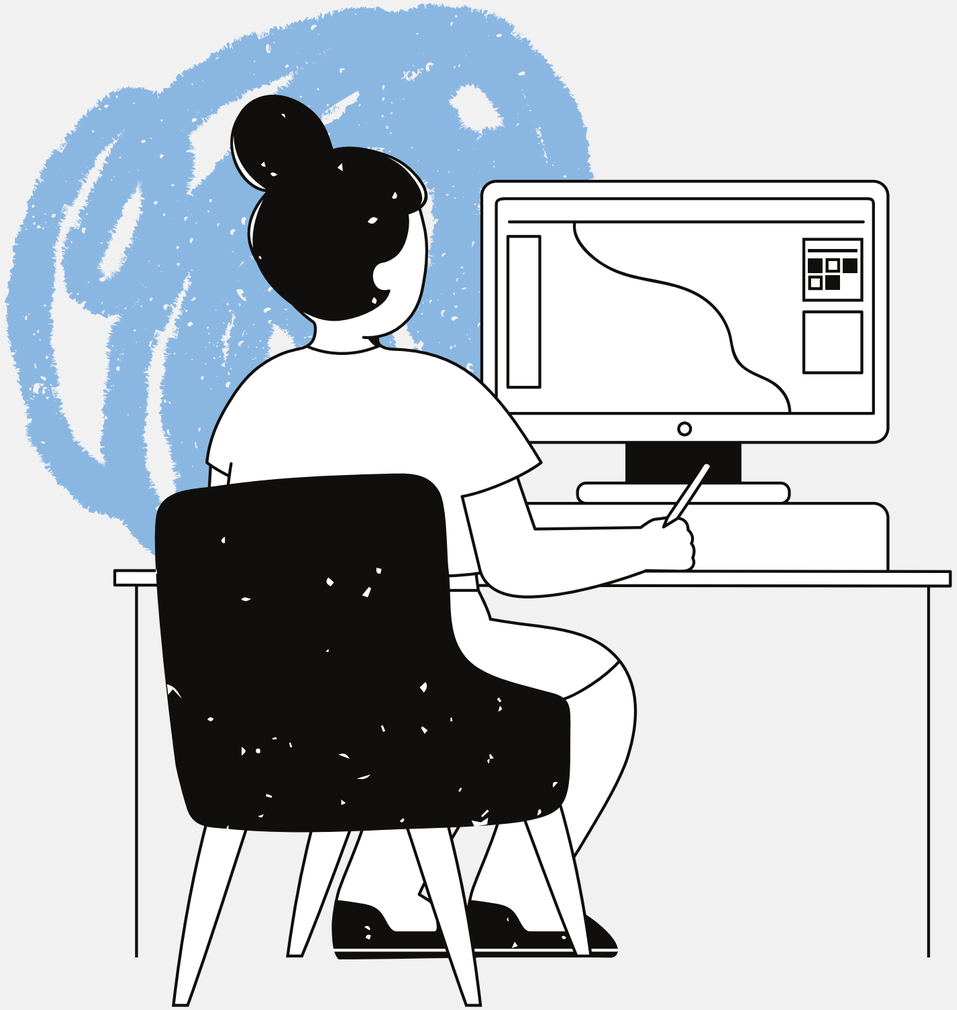
Overview

This project enhances learning for visually and hearing-impaired individuals with AI-powered tools. It includes a PDF Summarizer for concise study material, Scene Description with Voice Commands for better visual understanding, and Voice-Based Note-Making for effortless note-taking. Users can generate MCQs by Speaking Topics for quick revision, while a Hand Gesture Identifier helps deaf individuals communicate effectively. Designed for accessibility, this solution empowers inclusive education.



Problem Statement

Visually and hearing-impaired individuals face significant challenges in accessing and processing study materials. Traditional learning methods often lack inclusivity, making it difficult for them to read, take notes, and interact effectively. There is a need for an intelligent system that simplifies learning through voice commands, automated summaries, gesture recognition, and interactive study tools, ensuring equal educational opportunities for all.



Project vision and mission

This project aims to empower visually and hearing-impaired individuals by providing AI-driven tools that enhance accessibility and learning efficiency.

01. Firstly, it enables seamless study access through features like PDF summarization and scene description with voice commands, allowing visually impaired users to listen and understand study materials effortlessly.

02. Secondly, it simplifies note-taking and revision by introducing voice-based note-making and an MCQ generator, allowing users to create notes and practice questions just by speaking.

03. Lastly, it enhances communication for deaf individuals with a hand gesture identifier, bridging the gap between sign language and digital interactions, making learning more inclusive and interactive.

Tech Stack

01

Frontend: Built with Next.js, ensuring a responsive and dynamic user experience.

02

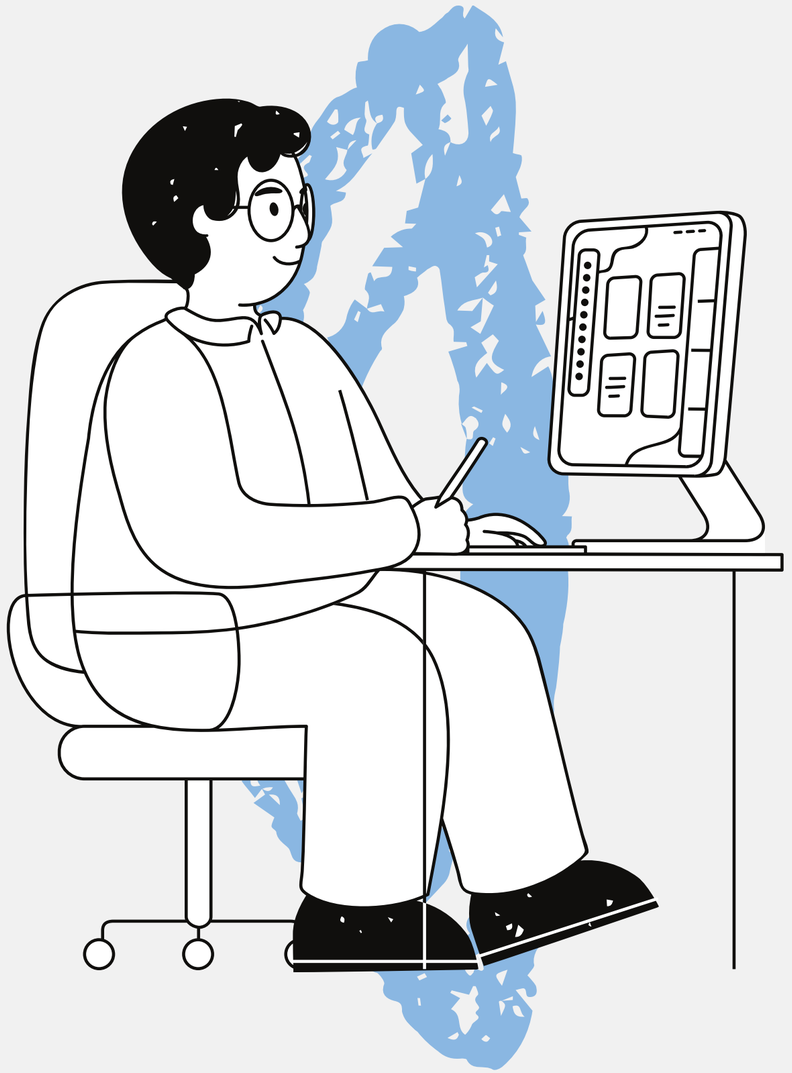
Backend & AI Processing: Powered by Flask, with LangChain, Hugging Face, FAISS, OpenCV, and Transformers for AI-driven features like PDF summarization, scene description, and gesture recognition.

03

APIs & Data Handling: Utilizes Google AI, Generative AI, YouTube Transcript API, PyPDF2, NumPy, and Scikit-Learn for efficient processing and data retrieval.

04

Security & Utilities: Incorporates Flask-CORS, SQLAlchemy, Requests, Pydantic, and Werkzeug, ensuring smooth API communication, authentication, and overall system robustness.

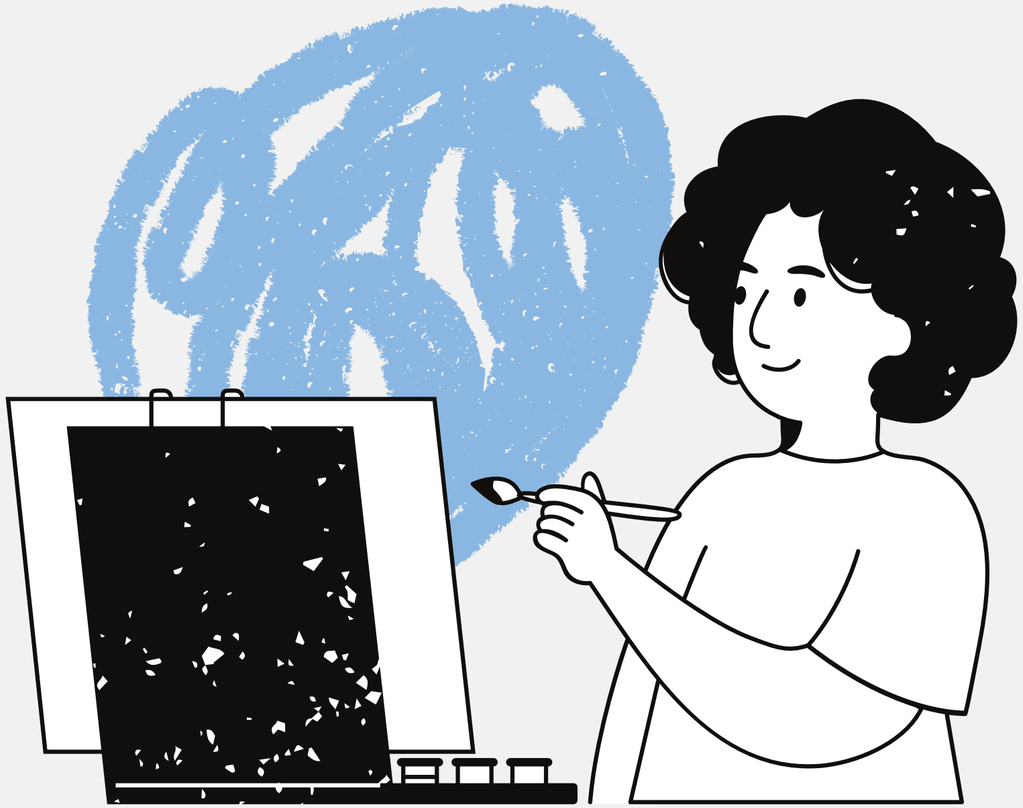


Challenges faced

Challenges included fine-tuning AI models for accuracy, ensuring seamless real-time voice commands, and optimizing data processing for speed. Managing speech-to-text, text-to-speech, and gesture recognition was complex, while ensuring accessibility and smooth integration in Next.js and Flask required extensive testing. Overcoming these hurdles led to a powerful and inclusive solution.

Final reflections and future steps

This project successfully enhances education for visually and hearing-impaired individuals through AI-powered tools like PDF summarization, voice commands, MCQ generation, and gesture recognition. Overcoming technical challenges led to a seamless, inclusive, and efficient solution. Moving forward, we plan to improve AI accuracy, expand multilingual support, integrate real-time sign language translation, and develop a mobile-friendly version to further enhance accessibility and user experience.



Thank you very much!

By: Aditya Aman