

TECHNICAL REPORT

AI-POWERED INTERACTIVE LEARNING ASSISTANT FOR CLASSROOMS

TEAM NAME: CODEVERSE

DATE OF SUBMISSION: 12/07/2025

Abstract

The AI Learning Assistant is a smart, web-based classroom aid that is designed to empower teachers and assist students with the latest advancements in artificial intelligence. Built with NextJS and React for the responsive front end, and tailored with Tailwind CSS and ShadCN UI, the assistant enables students to ask questions through text, voice, or image input. Essentially, the system makes use of Google's Gemini AI models on the Genkit platform to provide real-time, context-aware explains and even draws diagrams or charts as and when needed. A major feature is this Engagement Monitor, which records the webcam (with the user's permission) and employs facial expression analysis powered by Gemini to measure student attention, confusion, and immediate attention. This feedback is received by instructors as visual feedback and recommendations, thereby enabling them to adjust their instruction methods in advance. Overall, the assistant generates a more personalized and responsive learning environment, bridging the gaps between traditional teaching and Existing AI technologies.

Introduction

In today's classrooms, learners are likely to miss timely individual assistance especially when they are reluctant to complain. Teachers, on the other hand, face challenges of establishing individual levels of engagement and adapting lessons on the fly. The AI Learning Assistant stands ready to assist with such issues by acting as a smart classroom guide that assists teachers as well as students in actual time. With images, voice, and text supported multimodal input it enables learning to be interactive and available to all learners.

Technically, the site has been developed with NextJS and React to offer a smooth, dynamic user interface. Clean Tailwind CSS and ShadCN UI provide a modern visual design. In the background is Google's Gemini AI, integrated through Genkit, enables the assistant to understand questions, process images, and provide reflective descriptions. One significant innovation is the Webcam image-based Engagement Monitor and facial expression-based Engagement Monitor analysis to recognize disengagement or confusion, to allow teachers to recognize classroom environment at a glance. Combined, these resources provide a better meaningful and effective learning experience.

Motivation Behind the Project

The primary impetus for this project is to break through the hurdle of maintaining one-to-one student interaction and providing individualized assistance in a class environment. Teachers usually have a hard time measuring the actual time-based levels of all of them at once. Traditional Q&A is dull and will fail to register the nuance of a pupil's confusion. This AI assistant came into being to act as an indefatigable, patient, and wise student and faculty assistant. For the learners, it offers a judgment-free way to ask questions and receive immediate, customized explanations. For educators, the Engagement Monitor provides a fresh, data-driven tool to understand classroom dynamics at a glance, allowing them to modify their instruction practices actively by expected levels of engagement, attention, and confusion.

Data Source

The program works with information submitted directly by the user in real-time. There is no static pre-existing dataset; rather, the system is dynamic and geared to react to live inputs. The data sources are:

- **Text Queries:** Students type questions directly into an input form.
- **Image Queries:** Students upload images (e.g., a picture of a textbook problem, a whiteboard diagram) to request explanations or analysis.
- **Voice Queries:** The app utilizes the browser's integrated Speech Recognition API to convert spoken questions into text, and send the text to the AI.
- **Live Camera Feed:** In the case of the Engagement Monitor, the app captures access to the user's webcam. It regularly captures still frames from the ongoing video stream.

These user-supplied data are passed to the set Google AI (Gemini) models via Genkit flows. The models take the input—text, images, or a mix—and use it to produce responses or do analysis. All data is processed ephemeral for a single query-response transaction and is not stored long-term by the app itself.

Work

AI Learning Assistant has two main parts:

1. Ask Anything – Multimodal Query System

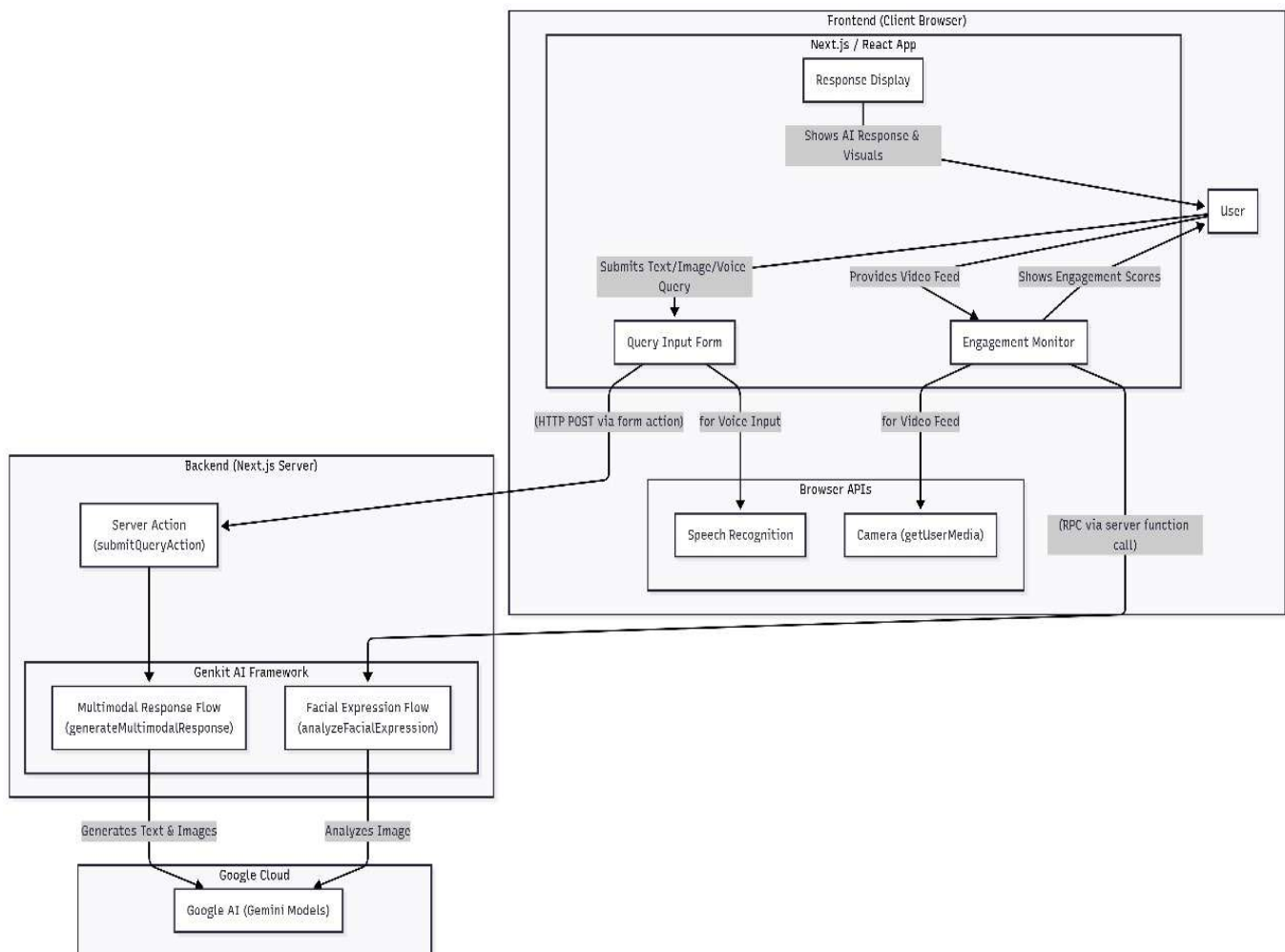
- Everybody is on the same screen — they can type, talk, or upload
- photos.
- It uses the smart AI flow to scan the question and generate:
- An explanation (in words)
- Helpful diagrams or illustrations (if required)
- Even with a voiceover that uses text-to-speech
- The goal is to give helpful answers — promptly, clearly, and in whatever manner.

2. Student Real-Time Feedback Engagement Monitor

- With permission access to cameras, the system tracks
- class.
- It takes photos and employs AI to understand
- Are students attentive?
- Are they functioning?
- Are they confused?
- It provides scores (out of 100) and even tips the teacher — like
- "Slow down and try" or "High engagement!"

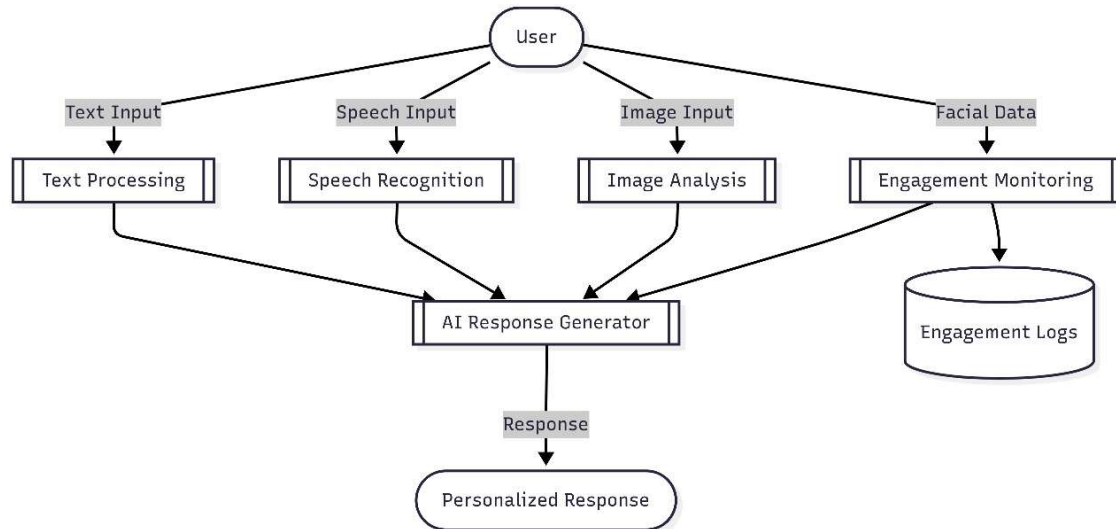
This means that teachers are provided with immediate, actionable feedback to adjust instruction in the moment.

System Architecture



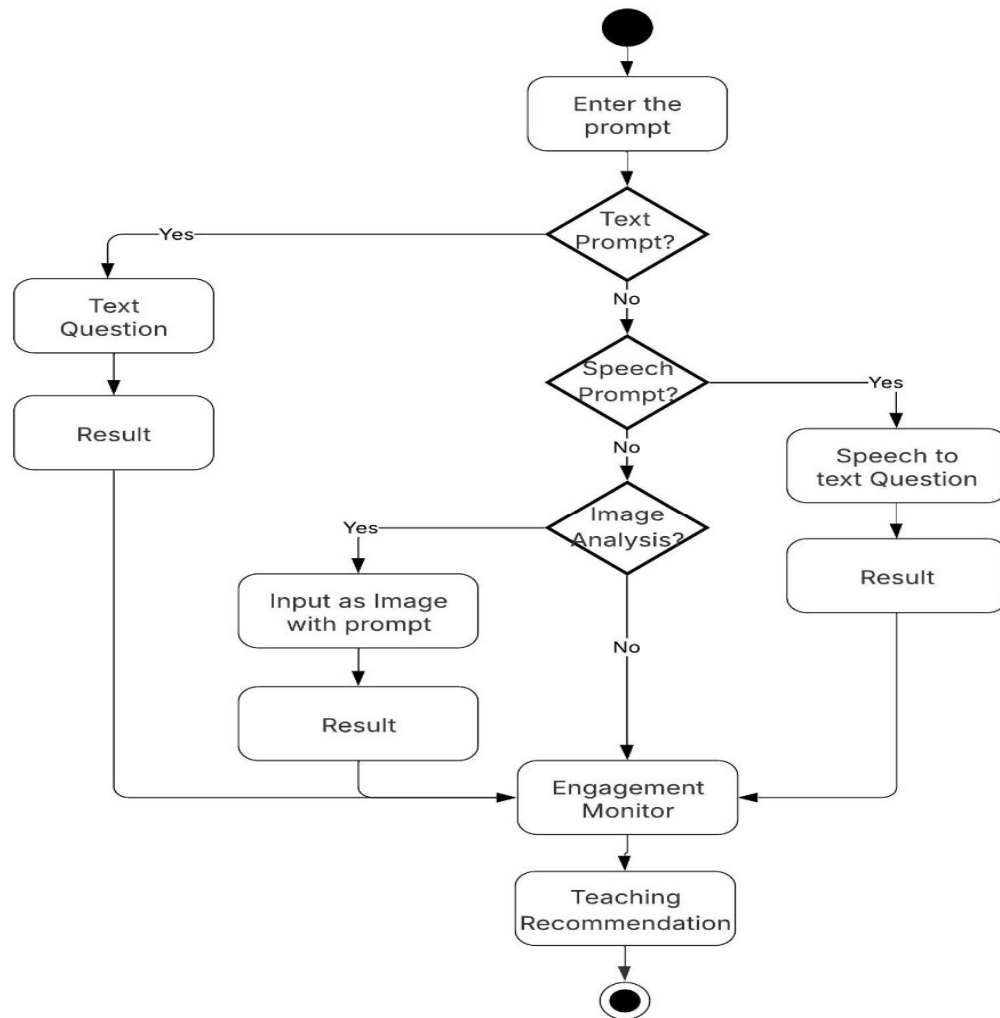
This architecture diagram illustrates how the AI Learning Assistant functions from the user's browser to the backend AI system. When a student asks a question with text, voice, or image, it passes through a frontend developed using Next.js and React. The voice input is processed using browser speech recognition, while webcam video is utilized for engagement tracking. All these inputs are passed to the backend where AI flows — driven by Genkit and Google Gemini models — produce responses or analyze expressions. These results, including personalized answers and engagement scores, are shown back to the user in real time, so the learning experience is smooth, smart, and interactive.

Data Flow Diagram



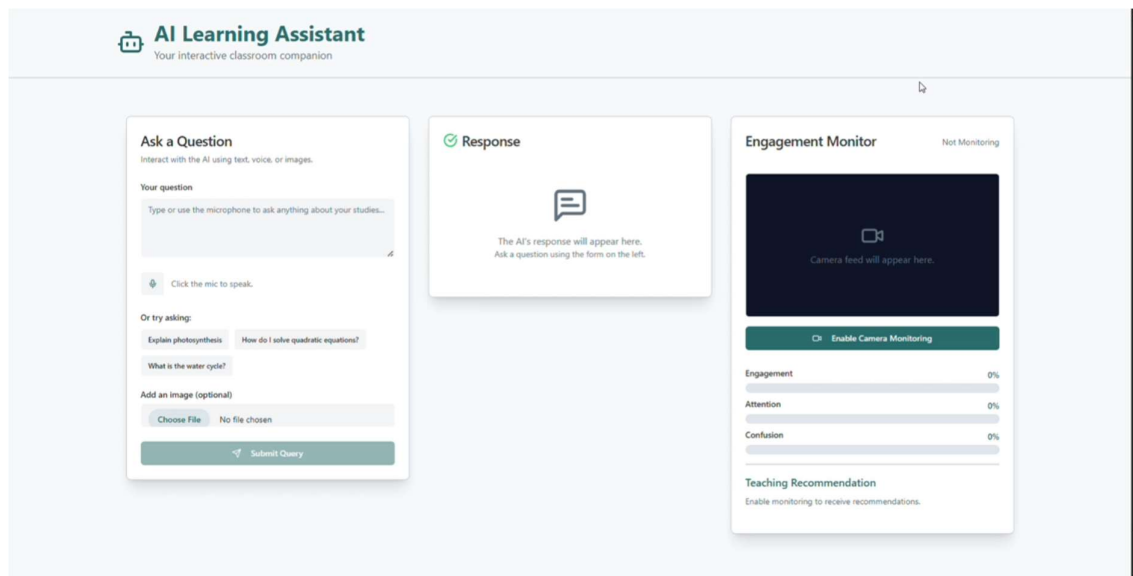
This flow of data shows how the AI Learning Assistant processes different kinds of user input. Whether a student types out a question, dictates it, or Uploads an image, the system runs it through the proper process — text processed, speech recognized, or image analyzed. All of these inputs are subsequently processed by AI Response Generator, creating a helpful and customized reply. At the same time, facial data from the webcam is used by the Engagement Track students' attention and store insights in engagement logs. It all together ensures intelligent responses to students and instant feedback for teachers.

Activity Diagram

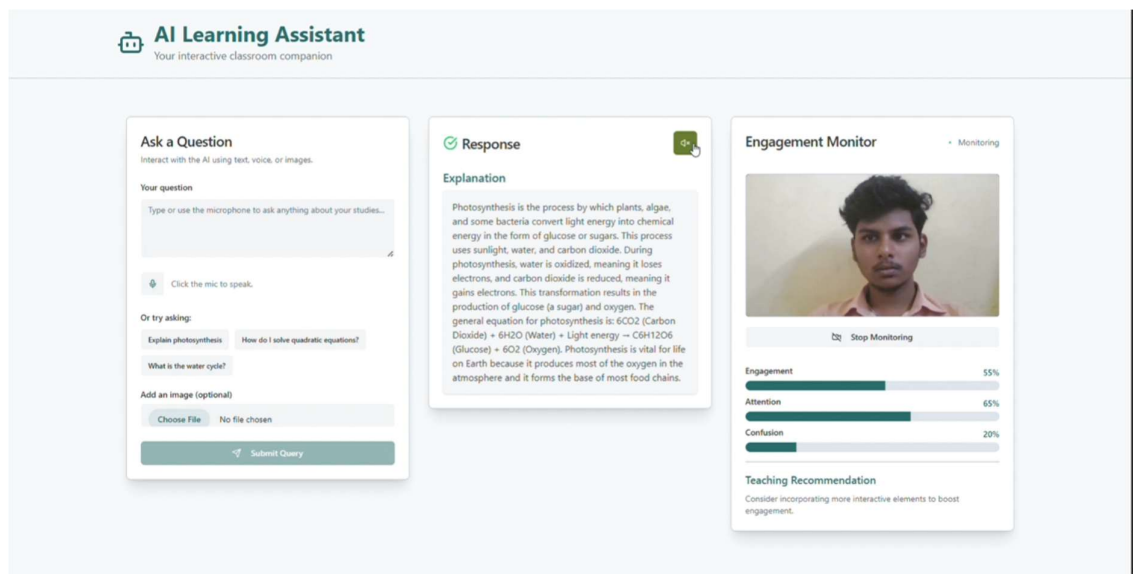


This activity diagram illustrates how the AI Assistant handles different types student questions. If a prompt is entered by the user, the system initially checks whether it's a text, speech, or a picture. Depending on what you enter, it interprets the question and answers. If an authentic question is not input, the system rather initiates the Engagement Monitor to monitor attention from students via the webcam. As reported by that, it offers an instructional recommendation to the instructor for enriching classroom engagement. It's a straightforward, step-by-step process to enable learning in any form.

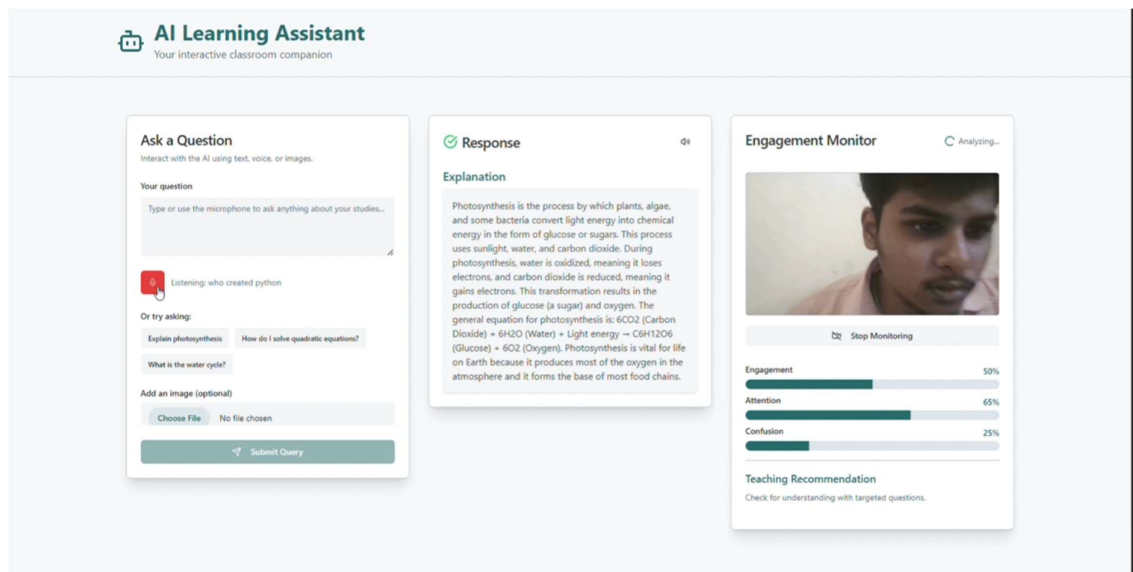
Result



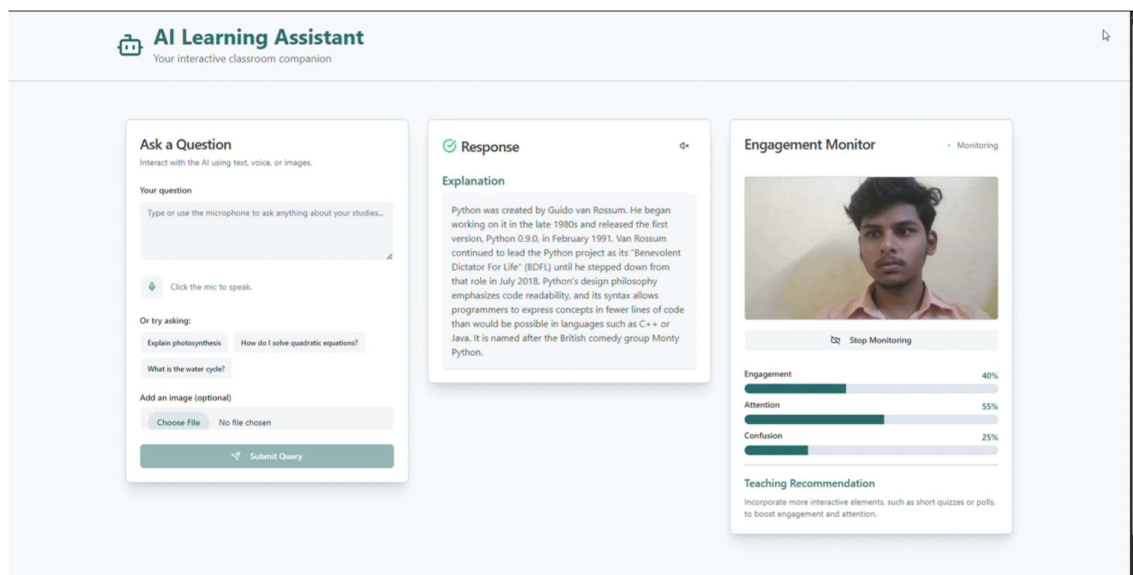
Result (1) Interface of the application



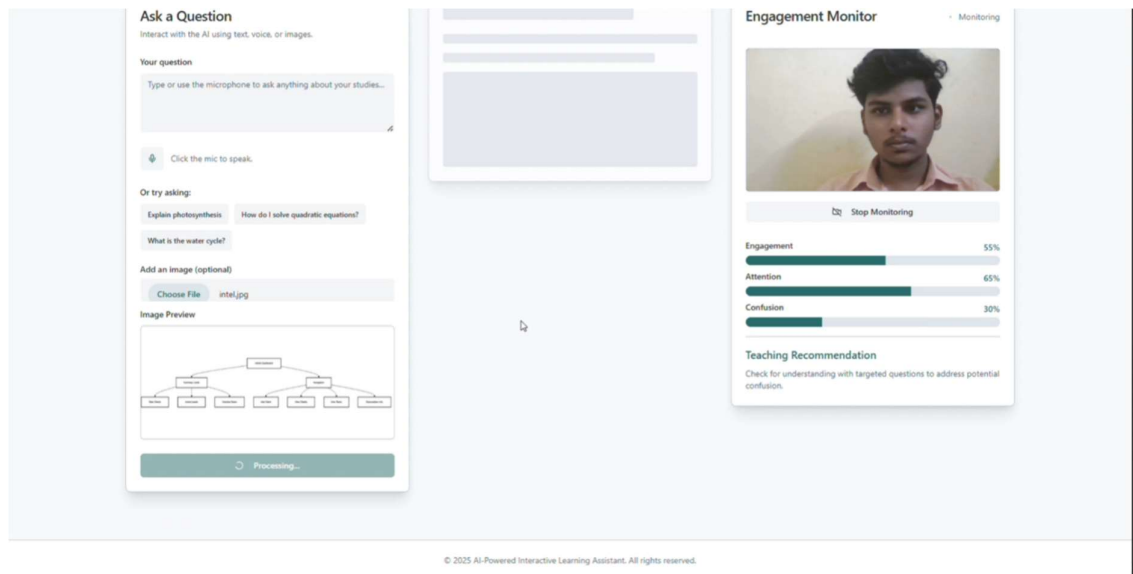
Result (2) User gives text prompt “Explain photosynthesis” and he gets the result on response field and engagement monitor analysing student interest and suggest teaching recommendation.



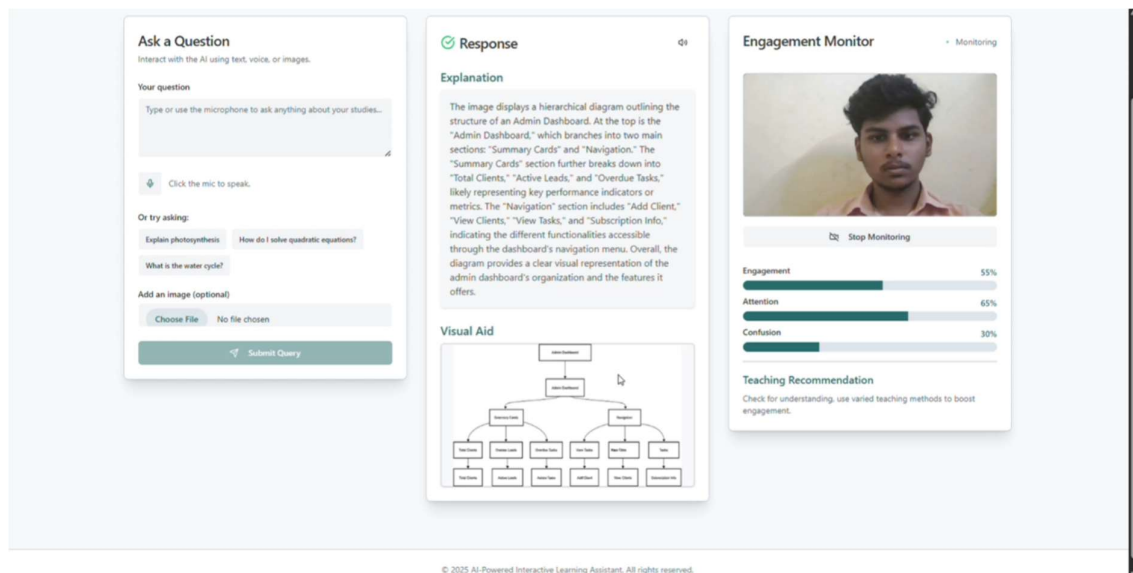
Result (3) User gives voice prompt “who created python”.



Result (4) User gets result for voice prompt “who created python” and he gets the result on response field and engagement monitor analysing student interest and suggest teaching recommendation.



Result (5) User gives a educational content image (a flow chat) and submits it.



Result (6) User gets result for image prompt and he gets the result on response field and engagement monitor analysing student interest and suggest teaching recommendation.

Links of the result:

Github link: <https://github.com/Nikitha6649/AI-Learning-Assistant.git>

Team members and contribution:

- **Dyapa Nikitha:** Integrated Genkit and Gemini AI, Handled Engagement Monitor, text-to-speech integration.
- **Dumpala Lokesh:** Build frontend components, connected input modules, Image analysis integration.
- **Bheemanapally Abhaya Sri:** Designed UI, Documentation, speech-to-text integration.