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 tool designed to support students and empower educators using the latest advancements in artificial intelligence. Built with **NextJS** and **React** for a responsive frontend, and styled with **Tailwind CSS** and **ShadCN UI**, the assistant enables students to ask questions using text, voice, or image input. At its core, the system uses **Google's Gemini AI models** via the **Genkit** framework to provide instant, context-aware explanations and even generate diagrams or charts where needed. A key feature is the **Engagement Monitor**, which accesses the webcam (with user permission) and uses facial expression analysis powered by Gemini to evaluate student focus, confusion, and attention in real time. This feedback is presented to teachers as visual scores and recommendations, allowing them to adapt their teaching methods proactively. Overall, the assistant fosters a more personalized and responsive learning environment, bridging gaps between traditional teaching and modern AI capabilities.

Introduction

In today's classrooms, students often struggle to get timely, personalized help especially when they feel hesitant to speak up. Teachers, on the other hand, face challenges in identifying individual engagement levels and adapting lessons on the fly. The **AI Learning Assistant** is built to address these issues by acting as an intelligent classroom companion that supports both students and educators in real time. With support for multimodal input including text, voice, and images it makes learning interactive and accessible for all types of learners.

Technologically, the platform is developed using **NextJS** and **React** to ensure a smooth, dynamic user interface. **Tailwind CSS** and **ShadCN UI** provide clean and modern visual design. Under the hood, **Google's Gemini AI**, integrated through **Genkit**, powers the assistant's ability to understand questions, analyze images, and offer meaningful explanations. A notable innovation is the **Engagement Monitor**, which uses webcam snapshots and facial expression analysis to detect confusion or disengagement, helping teachers understand classroom dynamics at a glance. Together, these technologies enable a more connected and effective learning experience.

Motivation Behind the Project

The primary motivation for this project is to address the challenge of maintaining individual student engagement and providing personalized support in a classroom setting. Educators often struggle to gauge the real-time comprehension levels of every student simultaneously. Traditional Q&A can be time-consuming and may not capture the nuances of a student's confusion. This AI assistant was conceived to act as a tireless, patient, and knowledgeable aide for both students and teachers.

For students, it offers a non-judgmental way to ask questions and receive immediate, tailored explanations. For teachers, the Engagement Monitor provides a novel, data-driven tool to understand classroom dynamics at a glance, allowing them to adjust their teaching strategies proactively based on estimated levels of engagement, attention, and confusion.

Data Source

The application operates on data provided directly by the user in real-time. There is no pre-existing static dataset; instead, the system is designed to be dynamic and responsive to live inputs. The data sources include:

- **Text Queries:** Students can type questions directly into an input form.
- **Image Queries:** Students can upload images (e.g., a photo of a textbook problem, a diagram from a whiteboard) to ask for explanations or analysis.
- **Voice Queries:** The application uses the browser's built-in Speech Recognition API to transcribe spoken questions into text, which is then sent to the AI.
- **Live Camera Feed:** For the Engagement Monitor, the application accesses the user's webcam. It captures still frames periodically from the live video stream.

This user-provided data is sent to the configured Google AI (Gemini) models through Genkit flows. The models process the input—be it text, images, or a combination—to generate responses or perform analysis. All data is handled ephemerally for the purpose of a single query-response interaction and is not stored long-term by the application itself.

Work

AI Learning Assistant has two main parts:

1. Ask Anything – Multimodal Query System

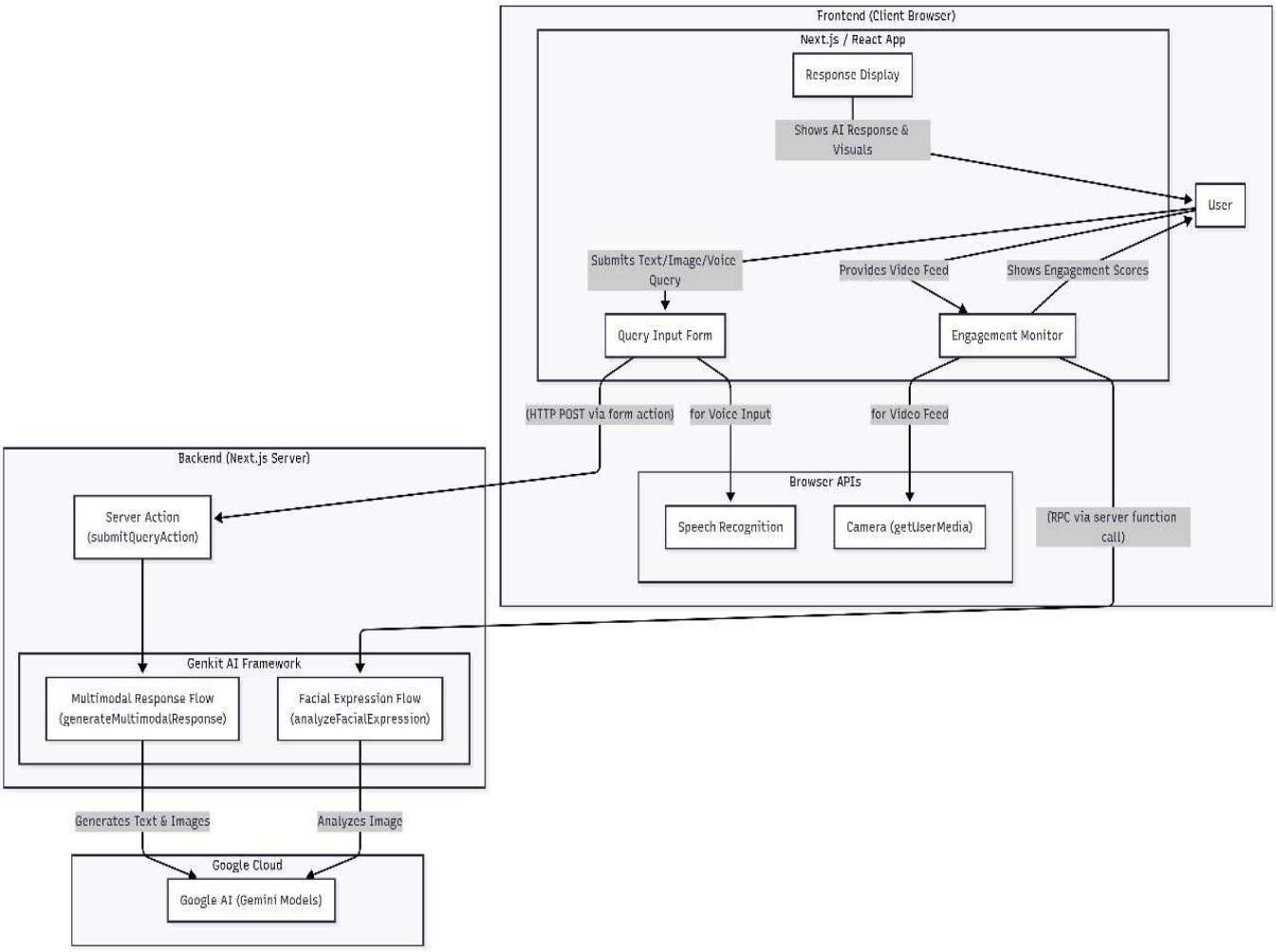
- Students interact through a single screen — they can type, talk, or upload images.
- The system uses a smart AI flow to understand the question and generate:
    - A clear explanation (in text)
    - Helpful diagrams or visuals (if needed)
    - Even a spoken version using text-to-speech
- The goal is to give helpful answers — fast, clear, and in any format.

2. Engagement Monitor – Real-Time Student Feedback

- With camera access (always with permission), the system watches the class.
- It takes snapshots and uses AI to understand:
    - Are students focused?
    - Are they engaged?
    - Are they confused?
- It gives scores (out of 100) and even shares tips with the teacher — like "Try slowing down" or "Engagement is high!"

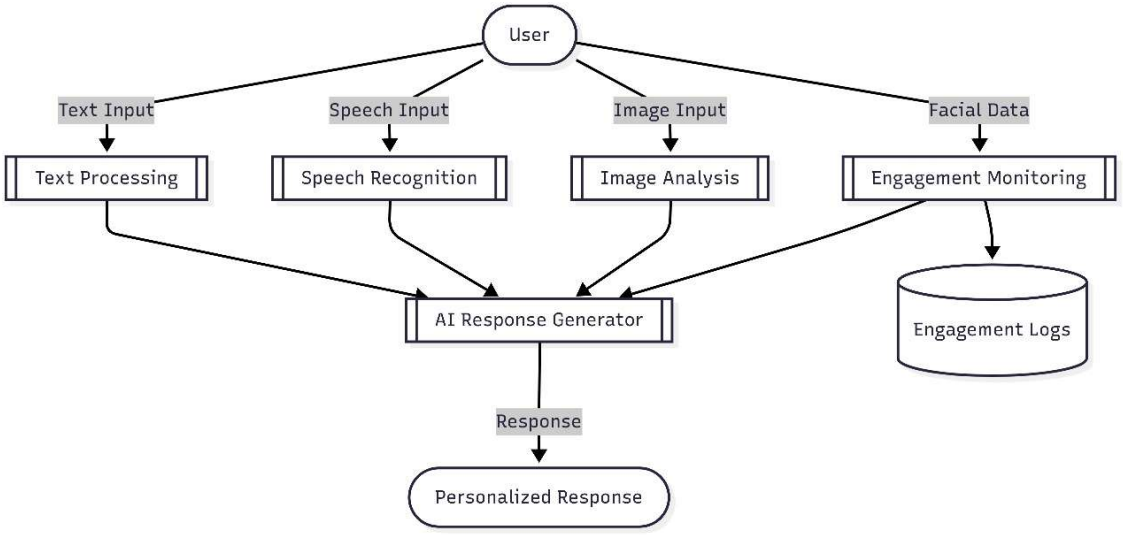
This means teachers get live, actionable insights to adapt their lessons on the go.

System Architecture



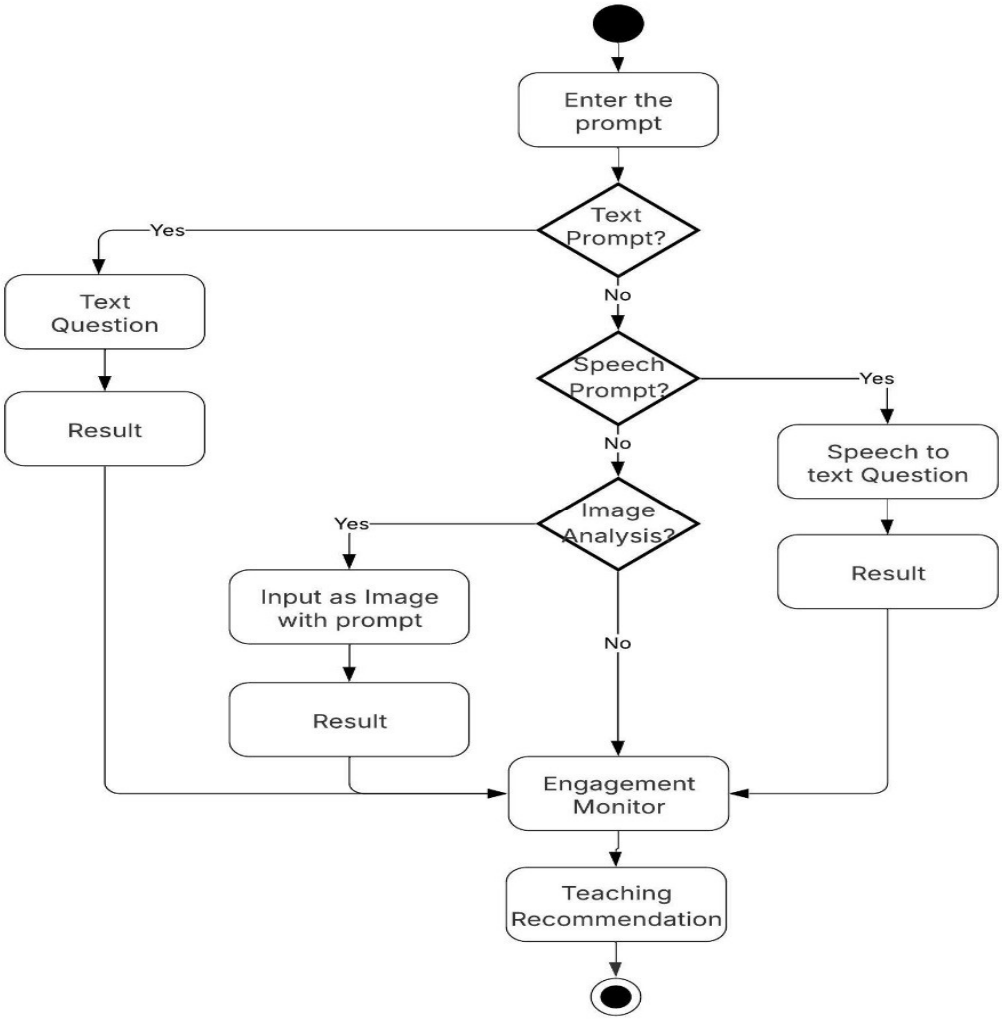
This architecture diagram shows how the AI Learning Assistant works from the user's browser to the backend AI system. When a student submits a question using text, voice, or an image, it goes through a frontend built with Next.js and React. Voice input is handled through browser speech recognition, while webcam video is used for engagement tracking. These inputs are sent to the backend where AI flows — powered by Genkit and Google Gemini models — generate answers or analyze facial expressions. The results, including personalized responses and engagement scores, are displayed back to the user in real time, making the learning experience smooth, smart, and interactive.

Data Flow Diagram



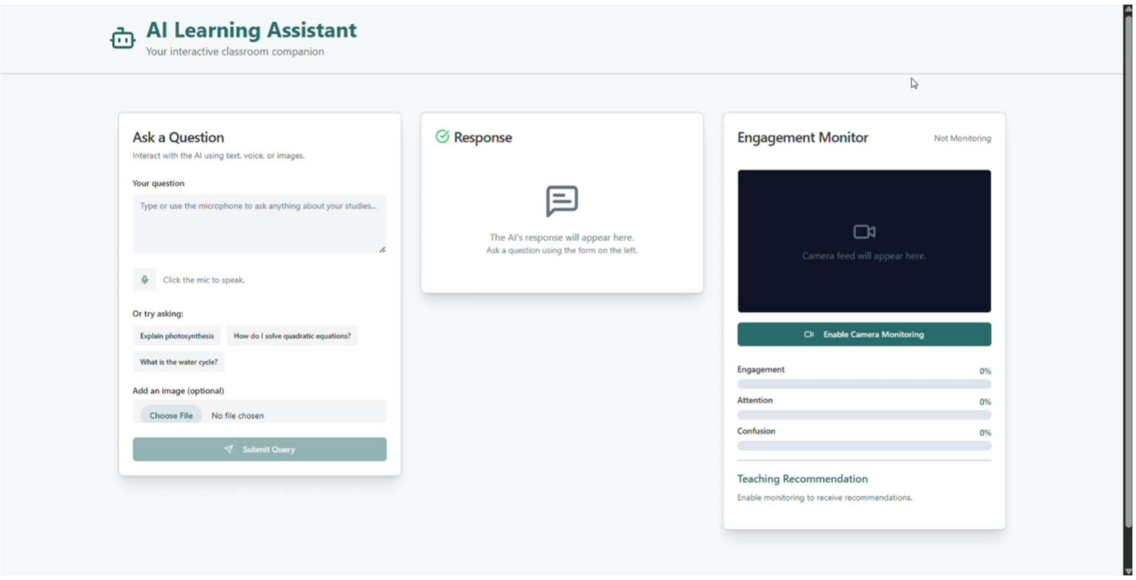
This data flow diagram shows how the AI Learning Assistant processes different types of input from the user. Whether a student types a question, speaks it, or uploads an image, the system routes it through the right process — text processing, speech recognition, or image analysis. All these inputs are then handled by the AI Response Generator, which creates a helpful and personalized reply. At the same time, facial data from the webcam is used by the Engagement Monitor to track student focus and store insights in engagement logs. Together, it ensures both smart answers for students and real-time feedback for teachers.

Activity Diagram

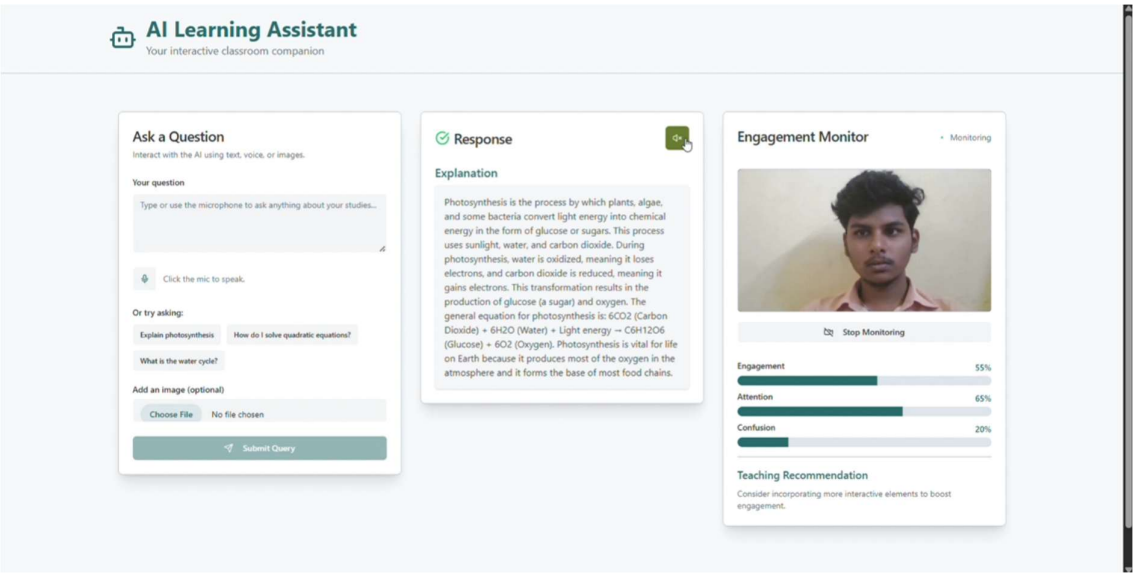


This activity diagram explains how the AI Assistant handles different types of student questions. When a user enters a prompt, the system first checks whether it’s text, speech, or an image. Depending on the input, it processes the question and provides a result. If no valid question is given, the system instead activates the Engagement Monitor to analyze student focus through the webcam. Based on that, it gives the teacher a teaching recommendation to help improve classroom engagement. It's a smooth, step-by-step system designed to support learning in any form.

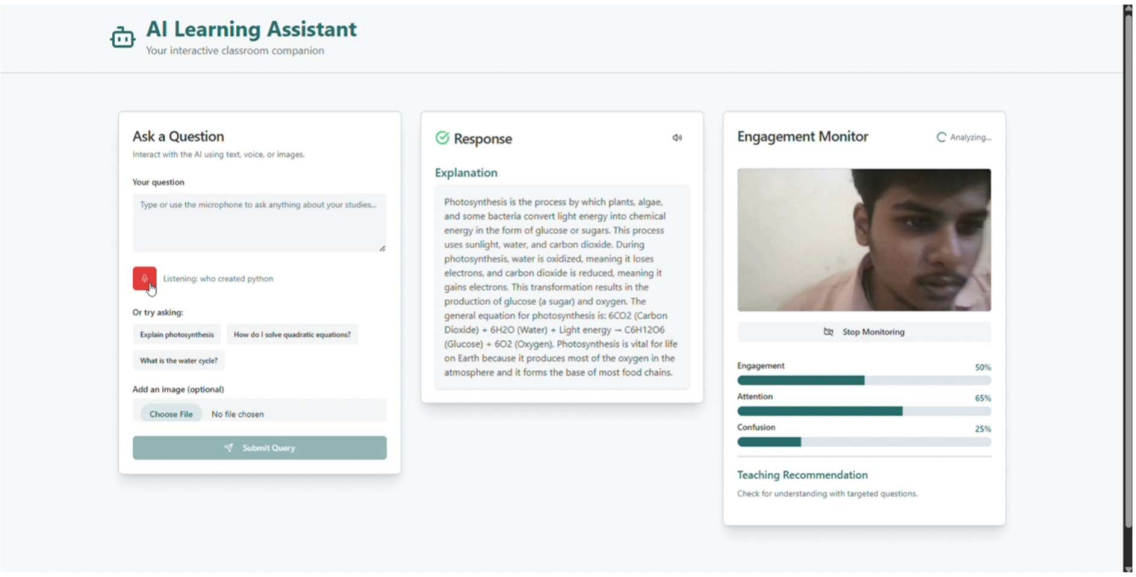
Result



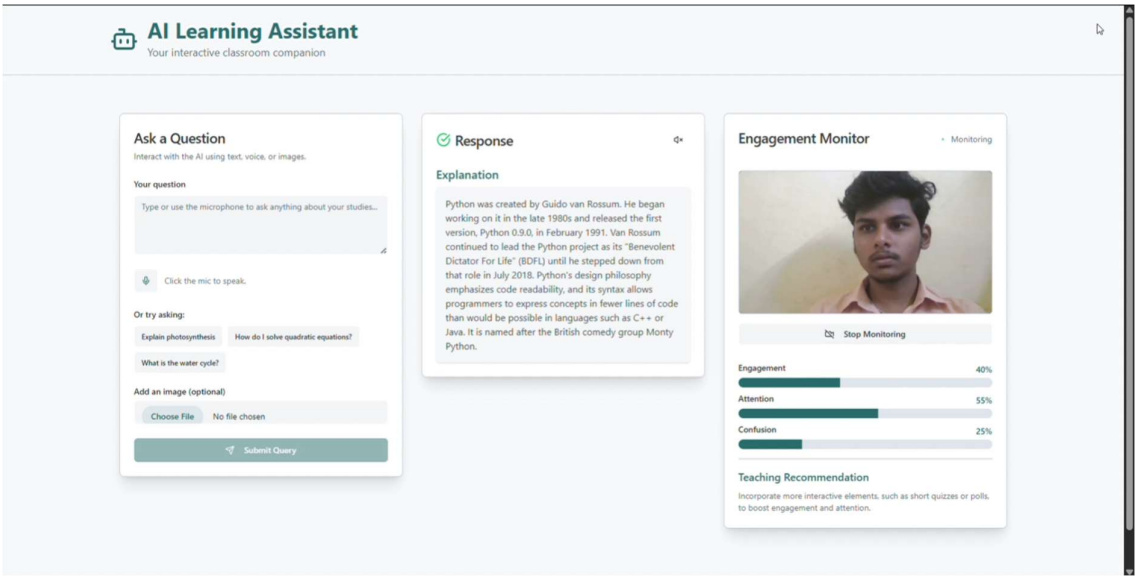
Result (1) Interface of the application



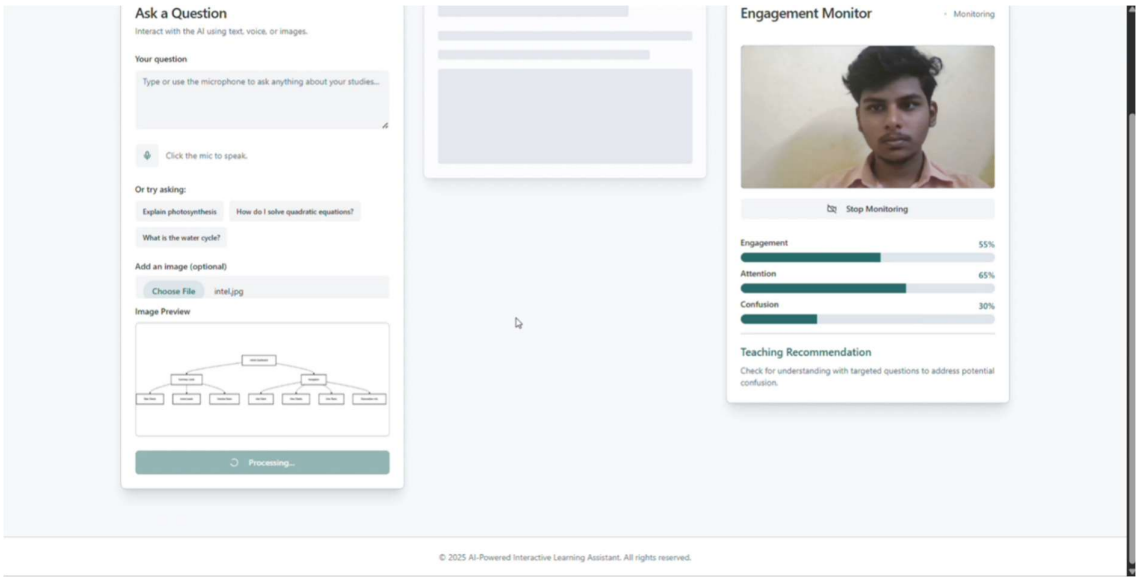
Result (2) User gives text prompt “Explain photosynthesis” and he gets the result on response field and engagement monitor analysing student interest and suggest teaching recommendation.



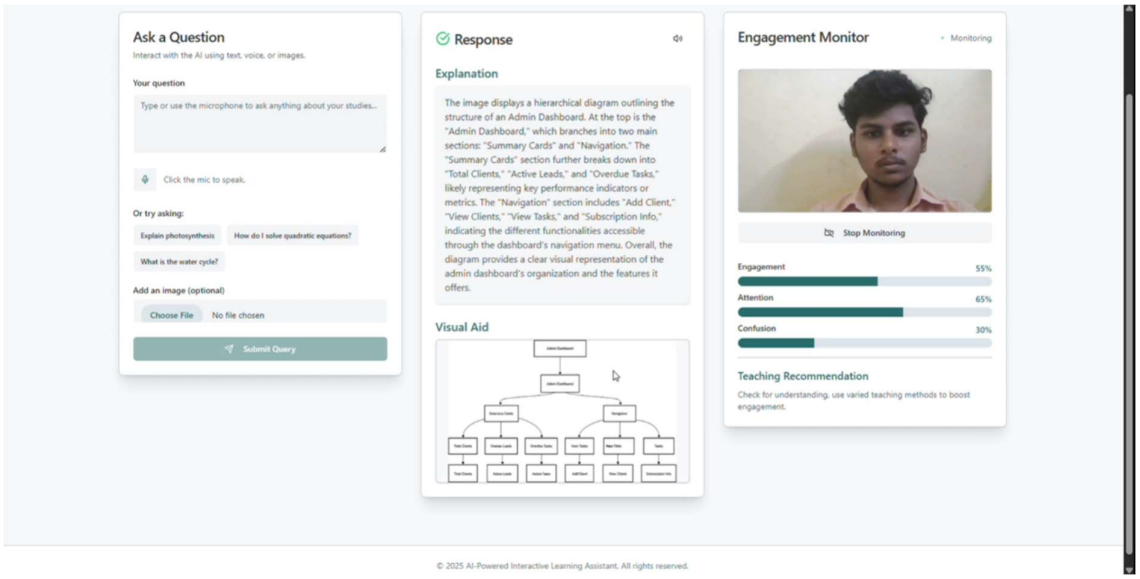
Result (3) User gives voice prompt “who created python”.



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.