

EduGenius

A Multimodal Educational Assistant

Team Members:

PERERA S.A.I.M.	210471F
PRABASHWARA D.G.H.	219483T
PRANAVAN S.	210491P

Group ID: 21

Project ID: 22

Mentor: Dr. Chathuranga Hettiarachchi

Teaching Assistant: Mr. Nadun Kumarasinghe

1. Executive Summary:

- **Problem:** Building an interactive AI maths tutor that can interactively guide students through maths problems.
- **Data:** The marking schemes of the A/L pure mathematics and O/L mathematics exam papers.
- **Data science methods:** Data collection, Image Cropping, Optical Character Recognition, Vectorization, Contextual Information Retrieval, Large Language Model (LLM) Integration, Evaluation and Validation.
- **Expected Outcome:** Interactive guidance/ tutoring through A/L and O/L maths problems.

2. Problem Statement:

In modern education, personalised and efficient tutoring remains a challenge, particularly in mathematics. Traditional tutoring methods often lack flexibility and efficiency hindering the learning process. Our project aims to provide a personalised and efficient AI-powered tutoring system for A/Level and O/Level students in mathematics. Utilising pre-processed and pre-stored maths marking schemes in the vector database, and the latest advanced technologies like LangChain, the system generates interactive and personalised tutoring materials for students and evaluates their performance in real-time. The system is designed to provide instant, accurate feedback, and grades, enhancing the learning experience and supporting students in their learning journey.

3. Data Description:

The question-marking pair for the knowledge base (vector database) is cropped as images from the marking schemes available as PDFs in the following URLs:

- [Evaluation Reports - Department of Examinations - Sri Lanka \(doenets.lk\)](#)
- [MathsApi - Largest Online Mathematic Educational Website: Combined Maths Marking \(English Medium\)](#)
- <https://olmathematics.com/>

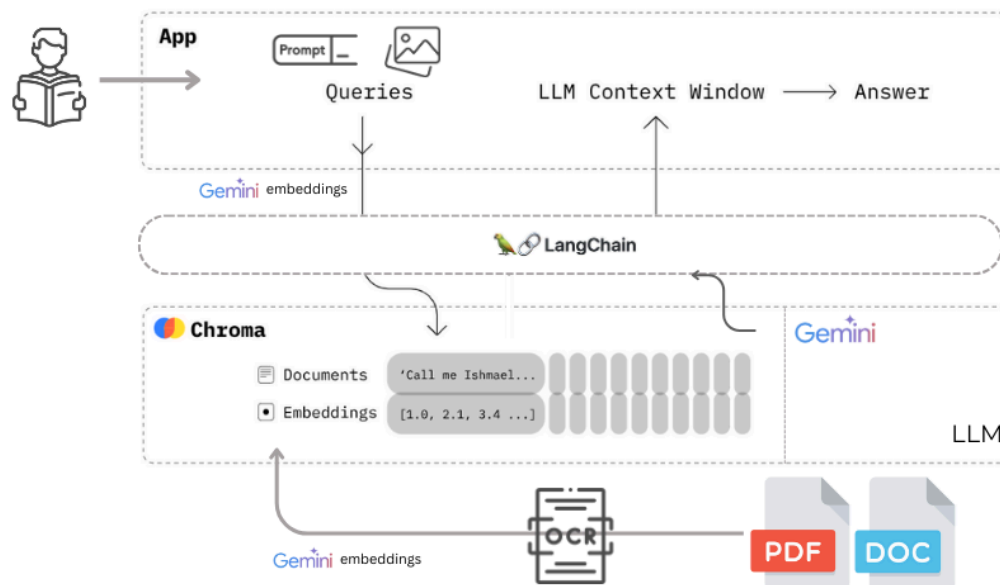
These PDFs and images are unstructured data types.

The project will utilise “**mathWriting**” dataset for recognizing the characters in user image inputs. The data set includes 253,000 human written expressions, 6000 isolated human written symbols and 396000 synthetic expressions. Those expressions are stored as a sequence of coordinates along with the Latex format of the expression. The dataset is structured.

4. Methods:

In this project, Natural Language Processing (NLP) is utilised to generate interactive tutorials for the users in order to guide them through the maths problems. Retrieval-augmented generation (RAG) enhances the relevance of this guidance by providing contextually accurate information. Additionally, Optical Character Recognition (OCR) is employed to extract text from uploaded images of question-marking pairs, ensuring comprehensive and accurate data processing.

1. First, the questions and respective markings are cropped from PDFs as images.
2. Those images are tested for clarity and then those images are doodled to indicate the allocated marks manually.
3. The textual contents in those images are extracted using OCR technology.
4. The textual content then will be embedded into the vector database.
5. Based on the question and user-written steps, interactive suggestions will be provided by obtaining context from the vector database and feeding it into LLM.
6. Evaluation.



- **User Interface:** Users can be interactively tutored via the user interface.
- **Chroma Vector database:** Vector database is used to store the textual information in vector format which can be easily compared for finding relevance.
- **Gemini:** Gemini API is integrated into the system to leverage the capability of LLM with the context obtained from vector databases to generate relevant responses.

5. Expected Outcomes and Success Criteria:

Expected Outcomes:

- Should provide comprehensive maths tutorials to the users
- Should provide correct marks to the users' answers
- Should give positive feedback to the user if users' answers are correct. And should describe what went wrong if the user's answers are incorrect.
- Should provide tips if users are taking too much time to answer.

Success Criteria:

- User feedback and usage statistics: Achieve a user satisfaction rating of at least 60%.
- Word Error Rate for the OCR component: Achieve a Word Error Rate of less than 10%.
- Human Evaluation for RAG: rate correctness, relevancy, fluency out of the scale from 0 to 10. Then take the average.

$$\text{HumanScore} = \sum(\text{correctness} + \text{fluency} + \text{relevancy}) / \text{no. of human observation}$$
 - HumanScore > 6

6. Preliminary Bibliography:

- Lang chain documentation: [Introduction | !\[\]\(13dd0e1ab3baa23f7c1ed52b3eec2756_img.jpg\) LangChain](#)
- Gemini API docs: <https://ai.google.dev/gemini-api/docs>
- [MathWriting: A Dataset For Handwritten Mathematical Expression Recognition](#)
- [RAG Evaluation | DeepEval - The Open-Source LLM Evaluation Framework \(confident-ai.com\)](#)
- <https://docs.smith.langchain.com/concepts/evaluation>