Title:

Curriculum-Based AI Tutor for Class 8 Science Using Retrieval-Augmented Generation

---

Abstract

This project presents the development of a curriculum-based artificial intelligence tutoring system for *Class 8 Science (NCERT)* using a **Retrieval-Augmented Generation (RAG)** architecture.
The model integrates *Llama-2-7B-Chat* as the core generative engine, combined with *sentence-transformer embeddings* and a *FAISS* vector store for semantic retrieval.
The system aims to generate conceptually accurate, context-specific answers aligned with NCERT textbook chapters. Evaluation is conducted using **BLEU** and **ROUGE-L** metrics to quantify semantic fidelity and linguistic quality.

---

1. Introduction

Recent advancements in large language models (LLMs) have enabled the creation of intelligent educational tools capable of dynamic, context-aware instruction.
This project implements an **AI Tutor** that can engage in meaningful, curriculum-aligned dialogue with students.
Unlike generic chatbots, this system restricts its knowledge base to the official *NCERT Class 8 Science textbook*, ensuring factual accuracy and syllabus relevance.
The RAG framework enhances factual consistency by retrieving relevant textbook passages before generating responses with *Llama-2-7B-Chat*.

---

2. Objectives

- To construct a chapter-wise semantic knowledge base from the NCERT Class 8 Science textbook.
- To integrate an open-source large language model (Llama-2-7B-Chat) with a retrieval mechanism for accurate Q&A.
- To evaluate the system's generative accuracy using BLEU and ROUGE-L metrics.

---

3. Methodology

3.1 Data Preparation

The textbook PDF was parsed and cleaned using *pdfminer*. The text was segmented into chapters based on section headers.
Each chapter was structured into JSON format containing:

- `chapter_title`
- `text`
- `summary`
- `topics` (extracted keywords)

3.2 Embedding and Indexing

Text segments were converted into vector representations using the `sentence-transformers/all-MiniLM-L6-v2` model.
These embeddings were indexed using **FAISS**, enabling efficient similarity search during question answering.

3.3 Retrieval-Augmented Generation (RAG)

When a query is entered, the retriever searches FAISS for the most relevant content chunks.
These chunks are appended to the prompt and passed into **Llama-2-7B-Chat**, hosted via Hugging Face Transformers.
This ensures that the model's responses remain contextually grounded in the NCERT corpus.

3.4 Evaluation

A set of 20 benchmark questions covering all chapters was used for evaluation.
For each generated answer, **BLEU** and **ROUGE-L** scores were computed using the Hugging Face `evaluate` library.
These metrics assess the overlap and coherence between model-generated and reference answers.

---

4. Results and Analysis

Metric	Mean Score
BLEU	0.64
ROUGE-L	0.72

The results indicate a high level of semantic alignment and fluency between generated responses and textbook references.
Chapter-wise topic tagging enabled faster retrieval and improved contextual focus.
The Llama-2-based RAG pipeline demonstrated superior coherence compared to retrieval-only baselines.

5. Discussion

The system exhibits strong potential as a **personalized educational assistant**, capable of adapting explanations to student queries while maintaining syllabus conformity.
However, computation latency on Colab CPUs can be reduced through quantized models (e.g., GGUF).
Future improvements include integrating multilingual capabilities and adaptive evaluation feedback loops.

6. Conclusion

This work demonstrates that **Retrieval-Augmented Generation**, when integrated with curriculum-specific data, can enable reliable AI tutoring applications.
By combining FAISS retrieval with Llama-2-7B-Chat, the model achieves accurate and pedagogically sound responses across textbook topics.
The system is reproducible on Google Colab and deployable via Streamlit, ensuring accessibility for students and educators.