NLP

Project Assessment - 1

Open Domain Question Answering
based on NCERT

Motivation

The need to create an open domain question answering (ODQA) platform based on NCERT textbooks arises from the necessity to bridge educational gaps and enhance learning efficiency.

Clear and immediate answers to students' questions are often lacking, holding back understanding and progress.

We, through this platform aim to provide instant, accurate responses, ensuring that students have continuous access to reliable information , thereby supporting their academic growth and fostering a more effective learning environment.

Problem Statement

The aim is to enhance access to standardized, curriculum-aligned knowledge for students across India.

ODQA intends to provide scalable, precise, and contextually relevant answers, therefore enhancing personalized learning and improving study efficiency. By addressing the limitations of traditional learning methods, this initiative will serve as a reliable educational tool to support students' academic progress.

Proposed Pipeline

Objective 1: Data Collection

Tasks:

1. Collect digital versions of all relevant NCERT textbooks.
2. Include additional educational resources.

Objective 2: Data Preprocessing

Tasks:

1. Text Normalization: Standardize the text.
2. Tokenization: Break down the text for easier processing.
3. Content Filtering: Remove irrelevant content

Objective 3: NLP Model Development

Tasks:

1. Model Selection: Choose appropriate NLP models (e.g., BERT, GPT)
2. Training: Understand and process questions effectively.
3. Fine-Tuning: Improve accuracy and relevance of the answers.

Objective 4: Question Answering System Implementation

Tasks:

1. System Design: Design the architecture, including components for question processing, answer retrieval, and response generation.
2. Integration: Integrate the trained NLP models into the system to handle user queries.
3. Optimization: Optimize the system for speed and accuracy.

Objective 5: User Interface Design

Tasks:

1. Interface Development: Develop a user-friendly interface.
2. Accessibility Features: Ensure the interface is accessible to all users.
3. User Experience Testing: Gather feedback and make necessary improvements.

Objective 6: Testing and Evaluation

Tasks:

1. Extensive System Testing: Ensure the system works correctly under various conditions.
2. Accuracy Evaluation: Evaluate the accuracy using a set of benchmark questions.
3. User Feedback: Identify areas for improvement.
4. Iterative Refinement: Continuously refine the system based on feedback and testing results to enhance performance and user satisfaction.

Timeline

October 1
Data Collection & Preprocessing

- Collect NCERT textbooks and supplementary materials.
- Standardize text, tokenize content, and remove irrelevant information.

October 15
NLP Model Development

- Select appropriate NLP models (e.g., BERT, GPT).
- Train and fine-tune the models on the preprocessed dataset.

November 1
System Implementation

- Design the system architecture.
- Integrate trained models and ensure system functionality.

November 15
UI Development & Testing

- Develop a user-friendly, accessible interface.
- Conduct user testing and refine the system based on feedback.

Expected Outcome

We plan to develop an NLP based open domain question answering system using NCERT textbooks. The application will assist learners, teachers, and students in accessing precise answers in a variety of disciplines. The system will ease information access and improve educational engagement by accurately interpreting and answering complicated inquiries.

This tool has the potential to transform learning experiences by making academic content more accessible and interactive.

THANK YOU

TEAM 20

Chandana goud V-SE22UARI217

Kshitij Karmuse-SE22UARI134

Mithil Josyula-SE23LARI001

Yuvan reddy M-SE22UARI091