Automatic Multiple Correct Answer Question Generation

Project Overview

The goal of this project is to develop a solution that can automatically generate multiple-choice questions (MCQs) with multiple correct answers based on a given chapter from a subject. The generated questions should test the reader's understanding of the chapter and have more than one possible correct answer to increase the complexity and challenge of the questions. The objective is to assist educators in creating engaging and challenging assessments for their students.

Prerequisites

This project requires several libraries for text processing and natural language understanding. You can install these libraries using `pip`:

pip install PyMuPDF

You will also need to download NLTK data for tokenization:

nltk.download('punkt')

Usage

Here's an overview of how this project works:

1. Extract Text from PDF: The project begins by opening a PDF file and extracting text content. Images and non-text content are ignored.

2. Text Preprocessing: Special characters and punctuation are removed from the extracted text to prepare it for question generation.

3. Question Generation: The cleaned text is processed to identify sentences. For each MCQ, a sentence is randomly selected, and words from that sentence are chosen as both correct and incorrect answers. The correct answers are then shuffled among the choices.

4. Display MCQs: The generated MCQs are displayed, including the question, choices, and correct answers.

Sample Code

Here's a sample of how the code is used:

```
# Open the PDF file and extract text
pdf_file_path = 'chapter-2.pdf'
pdf_document = fitz.open(pdf_file_path)

# Initialize an empty string to store the extracted text
extracted_text = ""

# Iterate through each page and extract text for page_num in range(len(pdf_document)):
    page = pdf_document.load_page(page_num)
    page_text = page.get_text()
    extracted_text += page_text

# Close the PDF document
pdf_document.close()

# Remove special characters
cleaned_text = remove_special_characters(extracted_text)

# Generate MCQs
doc = cleaned_text[:2000]
generated_mcqs = generate_mcqs(doc, num_mcqs=5, num_correct_answers=2)

# Display the generated MCQs
for i, mcq in enumerate(generated_mcqs):
    print(f"MCQ {i + 1}: {mcq['question']}")
    print("Choices:")
    for j, choice in enumerate(mcq['choices']):
        print(f"{chr(65 + j)}. {choice}")

    # Display correct answers as sorted multiple letters (e.g., ['A', 'C', 'D'])
    correct_answers = [chr(65 + mcq['choices'].index(ans)) for ans in mcq['correct_answers']]
    print("Correct Answers:", sorted(correct_answers))
    print()
```

Sample MCQs

MCQ 1: In fact when a massive rebellion against British rule broke out in 1857 Bahadur Shah Zafar the Mughal emperor at the time was seen as the natural leader
Choices:
A. the
B. time
C. rule
D. Shah
E. In
Correct Answers: ['A', 'C']

MCQ 2: As powerful regional kingdoms emerged in various parts of India Delhi could no longer function as an effective centre By the second half of the eighteenth century however a new power was emerging on the political horizon the British Did you know that the British originally came as a small trading company and were reluctant to acquire territories How then did they come to be masters of a vast empire
Choices:
A. reluctant
B. Delhi
C. As
D. a
E. kingdoms
Correct Answers: ['C', 'E']

MCQ 3: From Trade to Territory The Company Establishes Power 2 Aurangzeb was the last of the powerful Mughal rulers He established control over a very large part of the territory that is now known as India After his death in 1707 many Mughal governors subadars and big zamindars began asserting their authority and establishing regional kingdoms
Choices:
A. rulers
B. Power
C. is D. in
E. From
Correct Answers: ['A', 'C']

MCQ 4: As powerful regional kingdoms emerged in various parts of India Delhi could no longer function as an effective centre By the second half of the eighteenth century however a new power was emerging on the political horizon the British Did you know

that the British originally came as a small trading company and were reluctant to acquire territories How then did they come to be masters of a vast empire
Choices:
A. second
B. emerging
C. in
D. come
E. that
Correct Answers: ['C', 'E']

MCQ 5: In this chapter you will see how this came about Fig 1 Bahadur Shah Zafar and his sons being arrested by Captain Hodson After Aurangzeb there was no powerful Mughal ruler but Mughal emperors continued to be symbolically important
Choices:
A. was
B. Captain
C. Mughal
D. his
E. this
Correct Answers: ['B', 'E']

Conclusion

This project demonstrates how to extract text from a PDF, preprocess it, and generate multiple-choice questions with multiple correct answers. The generated questions can be used to create engaging and challenging assessments for educational purposes.