American History Q&A with Wikipedia

University of Utah AI Bootcamp

Brian Nelson, Eric Caldwell, Jaime Portillo

Instructor- Jason Stewart, Teacher Assistant-Shay Schreurs

August 27th, 2024

Project Overview And Goals

Our project aims to…. develop an AI-driven system using that accurately answers questions about the American Revolution based on historical summaries

The goal of our tool is to improve the accessibility to historical information and enhance understanding through sentiment analysis

Data Collection Overview

- Collect articles from Wikipedia
  - Utilize Wikipedia's API
  - Query the API with categories and articles relevant to the “American Revolution”
  - Limit results to 100 articles
  - Extract Key Information (Title, Page ID, Description, and Full Article Content)
  - Organize the collected data into a Pandas DataFrame
- Created multiple datasets of relevant Wikipedia articles
- Key Objectives: Automated Retrieval, Scalable, and Comprehensive

Data Processing Overview

- Text Cleaning
  - Remove Digits and Special Characters
  - Normalize Text (converted all text to lowercase for uniformity)
- Tokenization
  - Word Tokenization
  - Stopwords Removed
- Data Structuring
  - Organized clean data into Pandas DataFrame
  - Prepared for Analysis with summarization
- Key Objectives: Data Quality, Consistency, and Enhanced Analysis for NLP Model

Model Training and Evaluation

BERT model fine tuned on the American Revolution data set

Evaluates potential answers by confidence score

Key findings-

-Effectiveness of Fine Tuning BERT

-Confidence Scores as a Useful Metric

-Importance of Specificity in Queries

Code Snippet

This code snippet shows how the model processes each summary, evaluates the confidence of the answers and selects the most accurate response

```
for _, row in df_articles.iterrows():
    context = row['Summary']

    # Ensure the context is a string and not empty
    if not isinstance(context, str) or context.strip() == "":
        continue

    # Use the QA model to get an answer
    try:
        result = qa_pipeline(question=question, context=context
    except Exception as e:
        print(f"Error processing context: {e}")
        continue

    # Compare the score with the best score so far
    if result['score'] > best_score:
        best_score = result['score']
        best_answer = result['answer']
        best_context = context  # Save the context in case it's
```

Project Approach

Data collection → Data processing → Model selection and fine-tuning → Integration with Gradio → Sentiment analysis with VADER

Challenges-

Multiple input arguments

Determining accuracy

Fine tuning

Reducing processing time with bigger datasets

Questions we are using.

Q: Who was involved in the American Revolution? Who else?

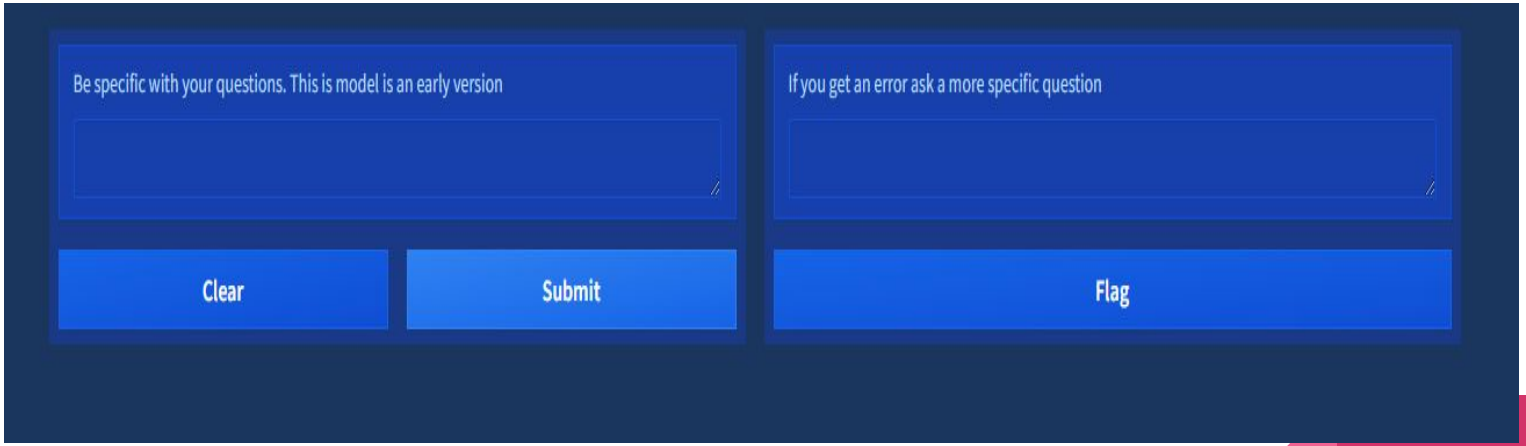
A: Paul Revere and Samuel Adams

Q:When did the American Revolution start?

A: 1772

Gradio Interface

Functions- gr.themes(dawood/microsoft_windows), gr.Textbox/placeholder, with gr.Blocks



Analysis and Conclusion

The project successfully demonstrated the feasibility of using AI to enhance historical education.

Benefits include allowing information to be retrieved faster, and with more relevant results, than traditional search methods.

The model allows for customized processes that can be added in the future to meet specific educational needs.

Questions?

Questions and Future Development

Questions that surfaced during the project development.

1. Would Gradio have the functions we need to create a customized message for the human to see?

2. How can we test the accuracy of the model?

Future steps if time allowed.

1. Add additional questions to the ones we already have in use.

2. Create a custom dataset with manually curated question-answer pairs specific to the American Revolution and fine-tune the model on this dataset. (Fine-tuning on a dataset specifically designed for the project's goals would likely improve the model's ability to generate accurate and contextually relevant answers)