

Course Assistant Bot

Using Retrieval Augmented Generation (RAG) to build effective course resources

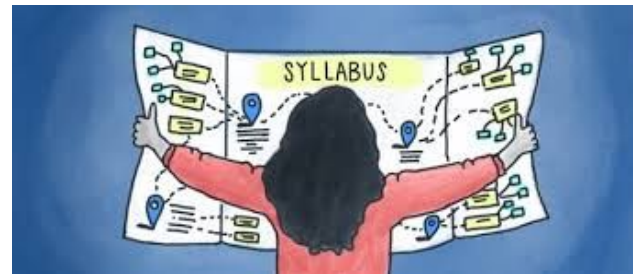
Developed by Reggie Bain



Background + Motivation

Designing Effective Syllabi

- Syllabi plays a key role in building course
 - Sets course policies (academic integrity, late policies)
 - Defines grading scheme
 - Explains structure of course (labs, homework, studio)



+

Goals

- Build RAG tool that assesses syllabi
- How well does a syllabus answer important questions?
- Our Targets → **Correctness of responses**
- Our Features → **Question and answer pairs**



Datasets

Data Sources

- Used real syllabi from different institutions as proof-of-concept
- Examples (all college-level courses):
 - Algebra-based Intro Physics Course for advanced high-school students
 - Intro Python-based CompSci Course for non-majors
 - Intro Racket-based CompSci Course for CompSci Majors



SOUTH CAROLINA GOVERNOR'S SCHOOL FOR
SCIENCE+
MATHEMATICS



NORTHWESTERN
UNIVERSITY

KPIs + Stakeholders

Key Performance Indicators

1. Build RAG pipeline for asking questions of syllabi
2. Successful evaluation of pipeline using synthetically generated Q&A set
3. Construct basic app that effectively allows users to query syllabi in multiple formats.

Stakeholders

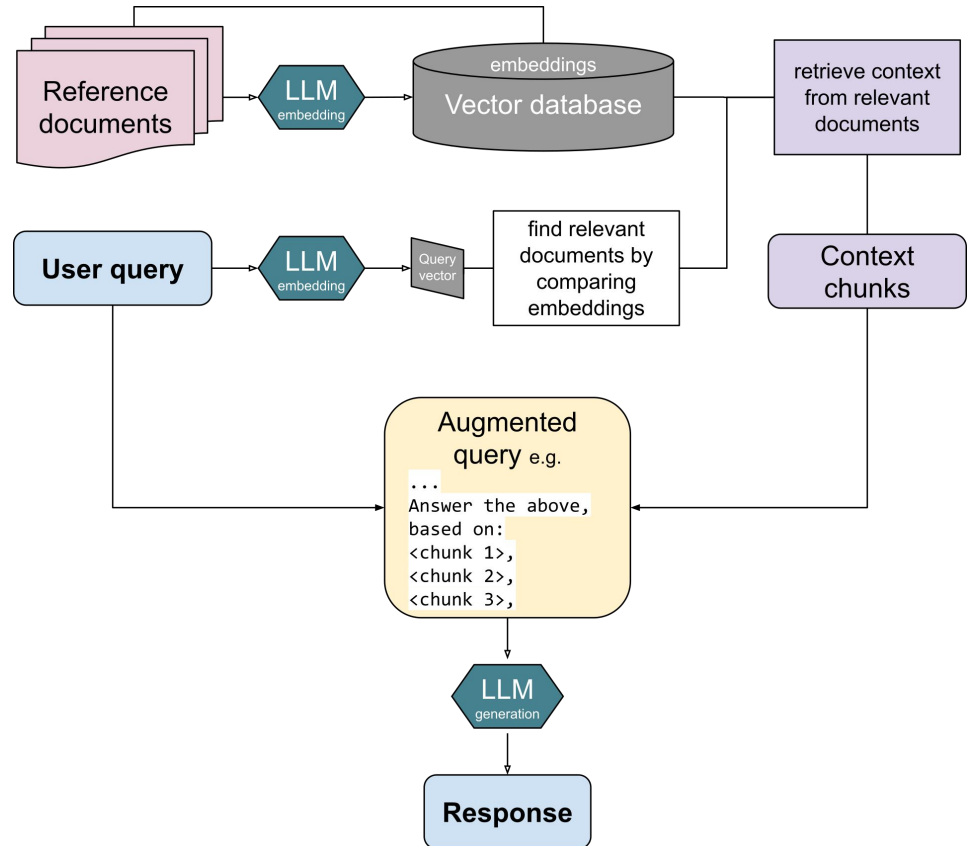
- Instructors, departments, and colleges for creating effective syllabi
- Students, to better able to understand course expectations
- Administrators/legal teams assessing syllabi as “contracts”



Retrieval Augmented Generation (RAG)

Outline of RAG Process

1. Build vector database of source documents
2. Create embeddings of user query and source documents
3. Retrieval of semantically relevant context using search
4. Provide query + context to LLM trained for Q&A

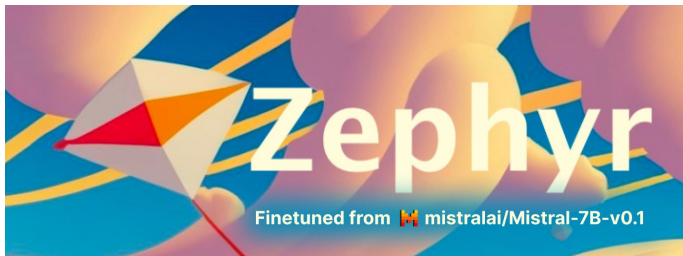


Models

Choice of LLMs + Databases

- Embeddings: [General Text Embeddings - Small](#)
- Synthetic Question Generation: [Mixtral-8x7B-Instruct](#)
- Q&A with Context: [Zephyr 7B Beta](#)
- Q&A with Context: [Llama 3.1-8B](#)
- LLM as Judge Agent: [GPT-4-1106-Preview](#)

LLaMA
by  **Meta**



 **OpenAI**
GPT-4

 **MISTRAL**
AI_

Methods

LLM Evaluation of Q&A Pairs

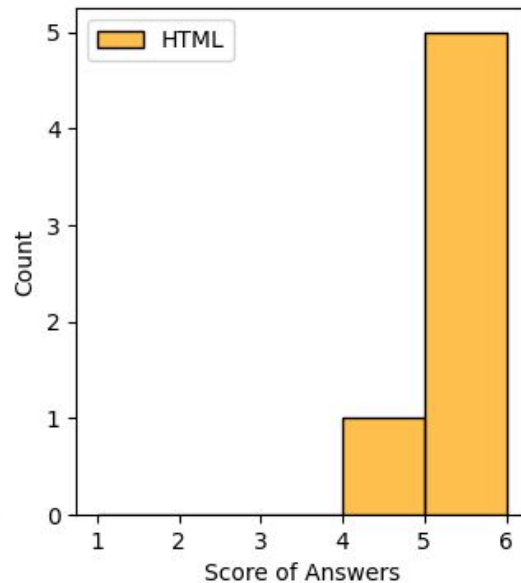
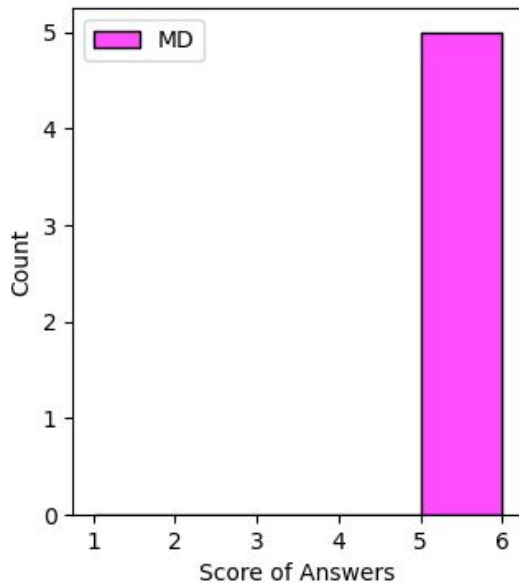
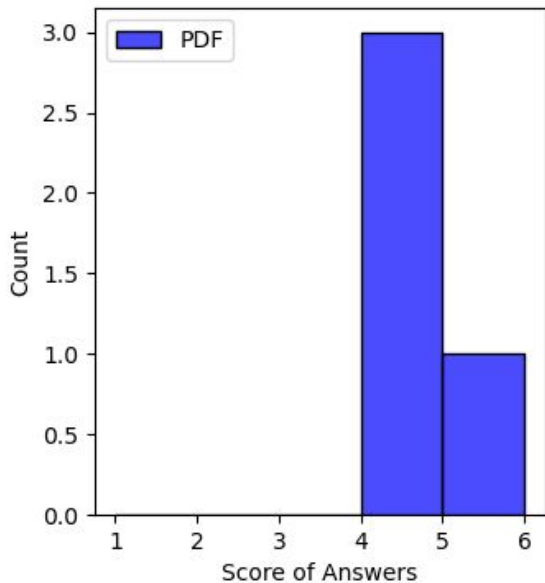
1. Generate synthetic Q&A dataset based on source documents
2. Feed synthetic dataset to LLM to get answers for each question
3. Use LLM to evaluate several metrics: "groundedness", "relevance", and "standalone"
4. Evaluate answers based on the questions and source documents using a more powerful LLM as judge agent



Results

Evaluating Q&A on Synthetic Dataset w/ RAG

GPT4 Evaluation of Q&A Pairs using RAG



Results

Evaluating Q&A Dataset without RAG

| Question | Generated Answer | Score |
|--|--|-------|
| What is the grading range for an F grade in the Graduate course? | Cannot answer. | 0 |
| How many extensions can a student request for assignments during the quarter? | Cannot answer. | 0 |
| What is programming used for? | Programming is used to instruct computers to perform specific tasks, solve problems, and automate processes. | 5 |
| Who are the people that can be found in the course staff? | Cannot answer. | 0 |
| What is the primary textbook for the course? | Cannot answer. | 0 |
| What should students bring to the exam according to the instructor? | Cannot answer. | 0 |
| How many college credit hours can students earn in the General Physics I course? | Cannot answer. | 0 |
| What is the purpose of edSTEM? | Cannot answer. | 0 |

Conclusions

Achieving KPIs

- Build RAG pipeline for syllabus analysis
- Developed basic syllabus analyzer app
- Establish pipeline evaluation scheme

Beating the Baseline

- Used baseline of LLM query without RAG
- RAG pipeline (somewhat obviously) beat queries without context

Future Work

Proprietary Models + APIs

- Use proprietary models with more parameters (8B -> 200+B parameters)
- Pro level accounts to Hugging Face, OpenAI to fix rate-limiting
- Try state-of-the-art models for Q&A, judge agent, synthetic Q&A generation tasks

Compute + Storage

- Use AWS (Redshift + S3 + Glue scripts) to store and deploy high-parameter models
- Host UI and databases for higher user volume.

Additional Use Cases

- More syllabus formats
- Integration with Canvas, Moodle, Blackboard
- **MORE TOOLS** - Querying course notes, lecture audio, etc.

Web Application

Can your syllabus answer basic questions?

- Try your at: <https://course-assistant-bot.streamlit.app/>
- Enter your own questions and download answers in JSON



Syllabus Evaluation App

Upload your syllabus (PDF or Markdown):

Drag and drop file here
Limit 200MB per file • PDF, MD, HTML, HTM

Browse files

Choose an LLM for questioning your syllabus:

Zephyr-7b

View and/or edit questions (one per line):

What topics are covered in this course?
What are the prerequisites for this course?
What is the grading policy?
What is the attendance policy?
What are the lab requirements?
What is the policy on late assignments?

Save Questions

Results: 50.00% of questions answered

Q: What topics are covered in this course?

A: The course, COMP_SCI 110: Introduction to Computer Programming, covers the key ideas of computer programming and assumes no prior programming knowledge. The course uses the programming language Python with the goal of introducing students to the fundamental constructs of computer programming and how to apply them creatively and usefully. The course consists of interactive lectures on Mondays, tutorial days on Wednesdays, and practice days on Fridays. Students are expected to review course materials, including recorded lecture videos, slides, and notes, before tutorial days and work in small groups to complete assignments during class time. The course aims to be an inclusive learning community and encourages students to correct instructors on their name or preferred pronouns if a mistake is made. The course also emphasizes the importance of dedicating enough time to practicing the basics in order to progressively move to more complex programming practice. (Source: Documents 0, 1, 3, 4, and 6)

Q: What are the prerequisites for this course?

A: The prerequisites for this course, CS 110: Introduction to Computer Programming, are listed in Document 1 as none. This information can be found under the section titled "Term Spring 2024 Course" and the course is identified as COMP_SCI 110.