GAITA: Using AI to Craft Personalized Learning Pathways for Computer Science

By Lois Wong
September 2024

Agenda

- Introduction
- Problem Statement
- Demo
- Project Overview
- Q & A

Introduction

A bit about myself

- BA in Linguistics from UC Berkeley (Dec 2021)
- Current 2nd-year Computer Science MS student at Johns Hopkins University
- Research in augmenting Generative AI using Information Retrieval systems (RAG) and personalized education
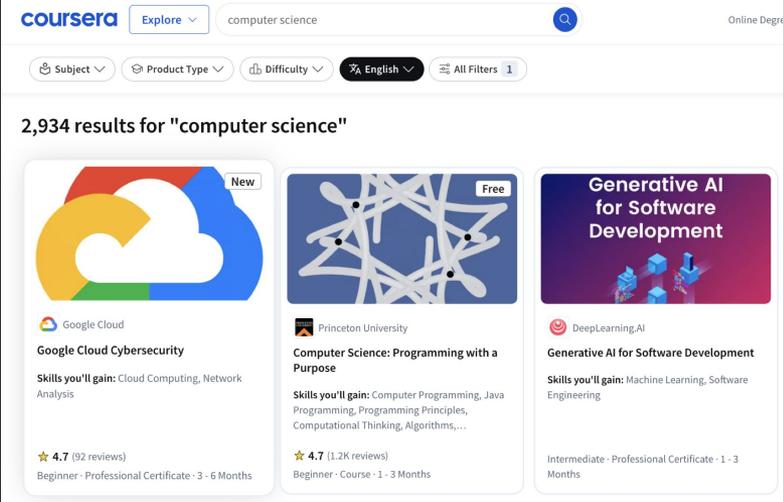- AI Education TA since August 2023

Motivation

- Inspired by my experience of transitioning into CS from a non-technical background
- Found an abundance of high quality online learning resources
- Difficult to keep goals in mind while handling prerequisites

Current Work in the Space

- **Course aggregators** compile numerous courses
    - Lack of personalization: doesn't tailor to individual needs/goals
    - Choice overload: difficult to choose among many options
   - Generic recommendations: lead to broad, non-specialized learning outcomes

Current Work in the Space

- **Personalized courses**
  - Standardization issue: personalized courses can lead to inconsistent skill assessments, making it hard to verify if claimed skills meet desired standards

Problem Statement

Problem Statement

Users cannot find courses(s) suited to their needs using existing learning platforms
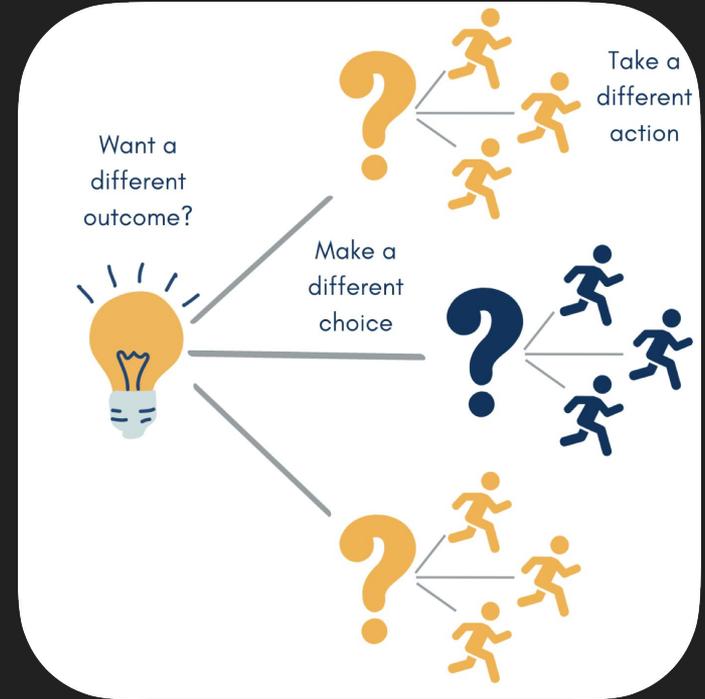
Excerpt from README

GAITA: Your AI-Powered Computer Science Learning Companion

Navigating the changing landscape of Computer Science is more difficult than ever, for both newcomers to the field and seasoned professionals aiming to stay informed and up-to-date on trends.

With a plethora of online learning resources, self-learners are often left to haphazardly browse the web and enroll in the first or most affordable program they find—a method that not only impedes their success and overall satisfaction, but also neglects to consider their unique perspectives and potential contributions to the field. It's extremely difficult to figure out which course(s) are right for you, and even harder to keep your goals in mind when worrying about prerequisites you might need. This can lead to enrolling in overly general introductory courses and ending up with a generic background in the field, ultimately making it harder to stand out in a competitive job market or make impactful contributions in your field.

We aim to solve this problem by developing GAITA, a Retrieval Augmented Generation (RAG) chatbot that generates personalized learning pathways from our database of over 1,200 open-access Computer Science courses from Coursera and MIT OpenCourseWare tailored to individuals' specific backgrounds, needs, and ambitions.

Evidence of Problem

- **Study**: Iyengar, s.s., & Lepper, M.R; (2000). "When choice is demotivating: Can one desire too much of a good thing?" *Journal of Personality and Social Psychology*
    - **Finding**: users faced with too many options are less satisfied and more likely to delay decisions
- **Study**: Chen, X., & Lo, J. (2020): "Personalized recommendation in education: A review and future directions." *Educational Technology Research and Development.*
    - **Finding**: Personalized learning tools show a lot of promise and is largely motivated by technology integration

Demo

# Project Demo

Architecture Diagram
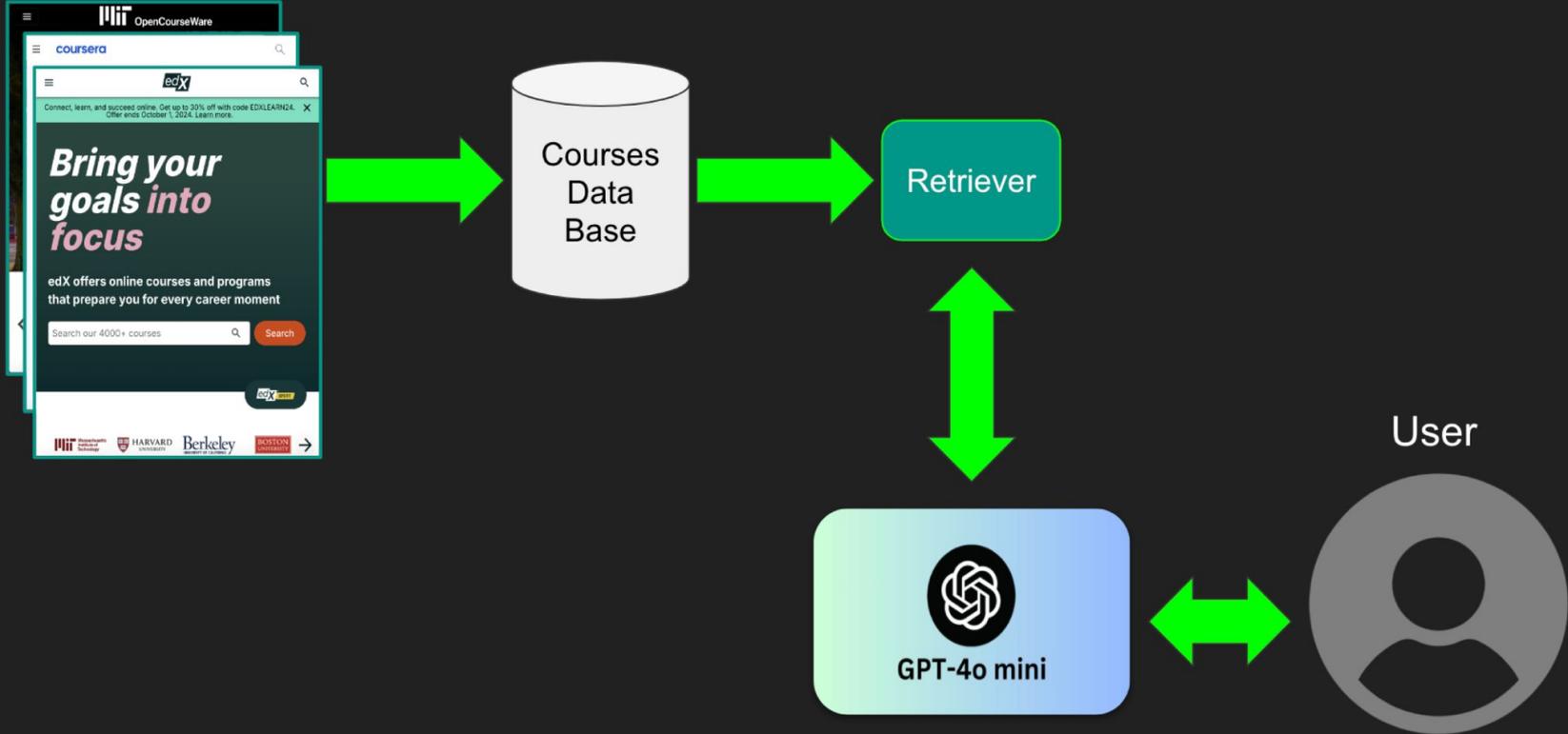
Part 1: Scraping and Database Construction

- Web crawling to create a database of 1,200+ open-access CS courses
- Database consists of course title, description, and URL
- Vectorized the title and description using OpenAI's text-embedding-3-small model
- Intermediary project: used Hugging Face's SentenceTransformers to recommend a single course from user's plaintext requests
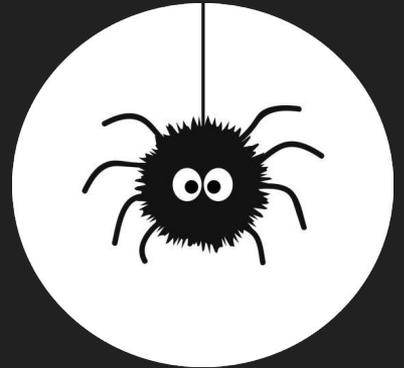
Part 2: RAG-Based Recommendation System

- Begin by prompting the user for their CS learning goals
- Cosine similarity to retrieve the top 3 relevant courses
- GPT 4-o mini to serve the recommendations with brief descriptions of how they fit the user prompt
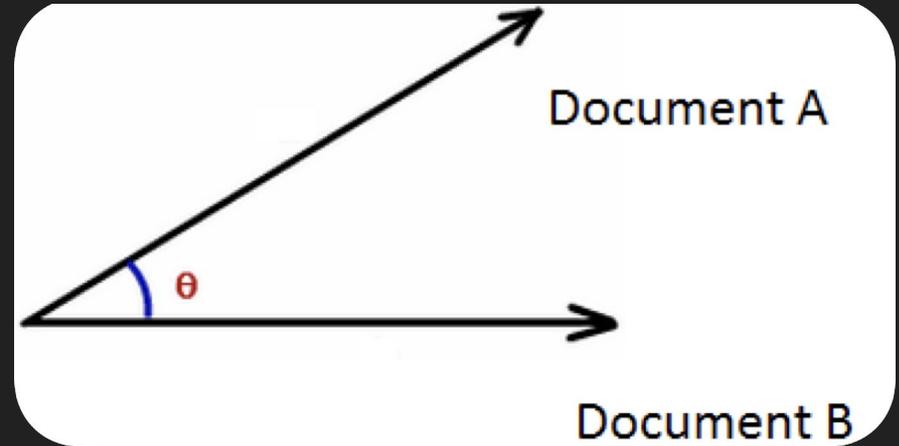
Part 3: Turning Recommendations into Pathways via Iterative Prompting

- Implement a multi-stage recommendation system via Iterative Prompting to craft a learning pathway that matches the user's current knowledge level
- Prompt the LM to generate follow-up questions (comfort level with prerequisites)
- User response is used to generate a new query for the retrieval module to recommend prerequisite courses (same process using LLM to serve recommendations)

TL;DR

- **Motivation**: transitioning in CS highlighted the problem of choice overload and unclear learning paths
- **Problem**: users cannot find the course(s) suited to their individual needs using current learning platforms
- **Solution**: an AI teaching assistant that builds personalized learning pathways based on user goals and backgrounds

Future Extensions

- Improving recommendation/iterative prompting
- Introduce caching (option to create an account and store users' course history)
- Write thesis on personalized education

Q & A

Thank You!

Contact & Links

- Gaita: https://gaita.onrender.com/
- GitHub: https://github.com/Lois-Wong/gaita
- Contact
  - lois_wong@apple.edu
  - lois@berkeley.edu