

2024 Nittany AI Challenge

Submission ID: 13

TEAM INFORMATION

1. Project Team Name:

Advisor

2. Team Members:

List each team member including Name, PSU Email, Campus, and College
(ex: Jane Doe, jxd123@psu.edu, Shenango, Engineering)

Vishnu Venugopal, vmv5177@psu.edu, University Park, College of Engineering
Rushil Kakkad, rkk5541@psu.edu, University Park, College of Engineering
Veer Shah, vas5455@psu.edu, University Park, College of Engineering
Krishna Pagrut, knp5451@psu.edu, University Park, College of Engineering

3. Primary Contact/Team Lead:

Vishnu Venugopal, vmv5177@psu.edu, 412-499-5076

PROJECT INFORMATION

4. Problem / Opportunity Statement

Documentation should include an overview of the problem and the method used to address that problem. This should demonstrate that you clearly understand the problem or opportunity you are addressing with AI. Response is limited to 300 words.

This year, Penn State has an undergraduate student population exceeding 42,000. With such a large student body, it should not come as a surprise when the advisor-to-student ratio falls significantly short of achieving an acceptable standard. Notably, within the College of Engineering, a field comprising over 5,000 students, there are merely 23 advisors. A substantial portion of these students fall under the 'Pre-Major' category, indicating that they have not yet selected a specific engineering field for their degree pursuit. This particular group is especially in need of comprehensive guidance as they navigate through the decision-making process regarding their academic path. This is a similar story throughout all the other colleges as well. Eberly College of Science has 3400 students for 11 advisors, a 312 to 1 student/advisor ratio. Liberal Arts is home to 5300 undergrads for 25 advisors, a 212 to 1 ratio.

Advisor serves as a comprehensive personal assistant, putting everything related to course and degree planning at students' fingertips for effortless access and convenience. Any question a student might have about their major and what classes to take, our product can accurately answer. This will save an incredible amount of time for students who do not read over the lengthy Penn State handbooks nor have to go through the hassle of scheduling an appointment with an advisor just for a relatively simple question.

The need for more advisors, especially in complex fields like engineering, sciences, and liberal arts, highlights the urgency for an improved advisory infrastructure. Our product aligns with President Neeli Bendapudi's Penn State visions of Transforming Internal Operations and Enhancing Student Success, offering 24/7 access to a user-friendly online portal for students seeking guidance on courses and degrees, streamlining the advice-giving process.

5. MVP Use Case

Provide a sample use case for the tool. Describe how someone will use the MVP functionality you intend to build and the benefits or impact the MVP will provide. Response is limited to 300 words.

Imagine a student struggling to navigate the college experience—from attending classes and study sessions to handling assignments and projects, the academic workload is undeniably demanding. In addition to these commitments, students often engage in extracurricular activities, attend social events, and may even take on part-time jobs. Within this challenging landscape students are expected to make difficult decisions regarding future courses and academic plans, with little to no help. Advisor is a streamlined solution for course planning, designed to save students time and mental energy. For instance, instead of consulting multiple sources spread out across different sites, students can use our product to receive succinct answers to any question about course planning.

Advisor also frees up advisors from administrative burdens, allowing them to prioritize deeper engagement with students needing guidance on academic progress, career exploration, and personal challenges.

Our product can also be used by advisors themselves, especially those who are unfamiliar with Penn State advising and might need some assistance generating course recommendations or suggesting degree programs.

In the future, we would like to implement a profile aspect to our app which allows Advisor to offer recommendations based off of past conversations as well as present ones. This will also let Advisor create a visualization of a student's academic journey, from completed courses to future aspirations, using clear and customizable flowcharts.

We would also like to expand Advisor's reach into offering possible clubs and research areas, enhancing all academic aspects of the college student experience.

As we are committed to data privacy, our product currently operates without retaining any information of the conversations between sessions, ensuring airtight security. To enhance the advice given by our LLM, we are brainstorming ideas that involve introducing user profiles that will keep track of any information that the user wants us to remember.

6. Data Availability

Detail the data sources leveraged within the prototype as well as the data sources necessary if this project moved to MVP. If available, please detail the location and availability of the data sources and/or the plan for collecting the necessary data. Remember that while we can provide some assistance with finding data sources, finding and gaining access to those sources is the team's responsibility. Response is limited to 300 words.

Data availability is a critical aspect of any AI project, as it determines the quality and reliability of the AI system. It is the backbone of our project that allows our product to specialize in college advising.

While developing our prototype, we used the course descriptions from the Penn State University bulletin as our data source (<https://bulletins.psu.edu/university-course-descriptions/undergraduate/>). This is an extensive list of every course offered at Penn State—a list far too long for any single student to read through entirely. Furthermore, we attempted to utilize the “Course Handbooks” offered by

different departments; however, a large issue we ran into was the disorganized and chaotic nature of each one. There was no similarity between different departments, and there was no common structure between them. By using the prior mentioned list, we were able to utilize the consistent formatting to extract data that fed our Large Language Model (LLM).

For our MVP, we plan to use additional data sources to make our software more robust and reliable. We hope to integrate program requirements, AP test information, and instructor ratings into our LLM.

These are the data sources that we leveraged for our prototype and our tentative plan for our MVP. We believe that these data sources are sufficient and necessary for our project, as they provide us with the information and insights that we need to create a high-quality AI Advisor that can provide academic guidance to students at Penn State University.

7. Technology

Provide a technical description of the approach the team used to achieve its proposed goal, including the ways in which the selected AI platforms are used within the prototype and how the team anticipates using those and other services in the MVP phase. Specifically, the documentation should include a list of the components of the selected AI platforms that are leveraged in the prototype, any additional components that may be leveraged in the development of the MVP, and additional services that may be necessary for continued development. Response is limited to 350 words.

Our software can be broken up into three parts: backend, pipeline, and frontend.

We are using ReactJS to develop our frontend as it provides a powerful framework for creating web applications. The user interface and experience is much easier to create when the styling and functionality are meshed into one powerful script. Moreover, utilizing React streamlines future mobile app development. ReactNative, a subset of React, specifically serves as the framework for creating native apps.

We use Flask to connect our frontend and backend. This lightweight Python microframework simplifies testing, making it easy to call LLM functions and pass data back to React using HTTP requests. As we scale up, we plan to transition to Django for handling increased traffic.

Our backend is the engine of our product and incorporates several different technologies that help us provide accurate responses to the questions the user asks. We first scraped information about every course and stored it in text files grouped by major. Using Open AI's embedding model, we then characterize the text data about each course into a vector and store it in a vector database. When the user submits a query, this same embedding model is again used to convert their text into a vector which is then compared to the other vectors in the database. It then pulls the most contextually relevant vectors and reranks them into a smaller subset, following the pipeline called Retrieval-Augmented Generation (RAG). This information is then fed into our LLM, guaranteeing an accurate and pertinent answer to the user's question.

For our MVP, we are thinking of making the shift to Azure OpenAI for faster responses and increased flexibility for our specific uses. Upgrading our LLM from GPT-3.5 to GPT-4 is certainly on our radar as well as it could provide more robust answers and give better advice. Finally, increasing the storage of our vector database and being able to better organize it through "Namespaces" and different pods is also a potential area that will help us grow.

8. Prototype Video Overview

All teams submitting a prototype for review are required to submit video demonstrations of their working prototypes. The videos must:

- *be no more than 5 minutes in length.*
- *explain the intent, goals, and potential impact of the solution.*
- *demonstrate the basic, working functionality of the prototype.*
- *be available through a YouTube link accessible for viewing by the Challenge reviewers.*

The production value of the videos will not be factored into the review, but they must clearly and accurately represent the prototype functionality. To help, Media Commons at Penn State provides free One Button Studio options throughout the Commonwealth.

****See Video Provided in the Folder**