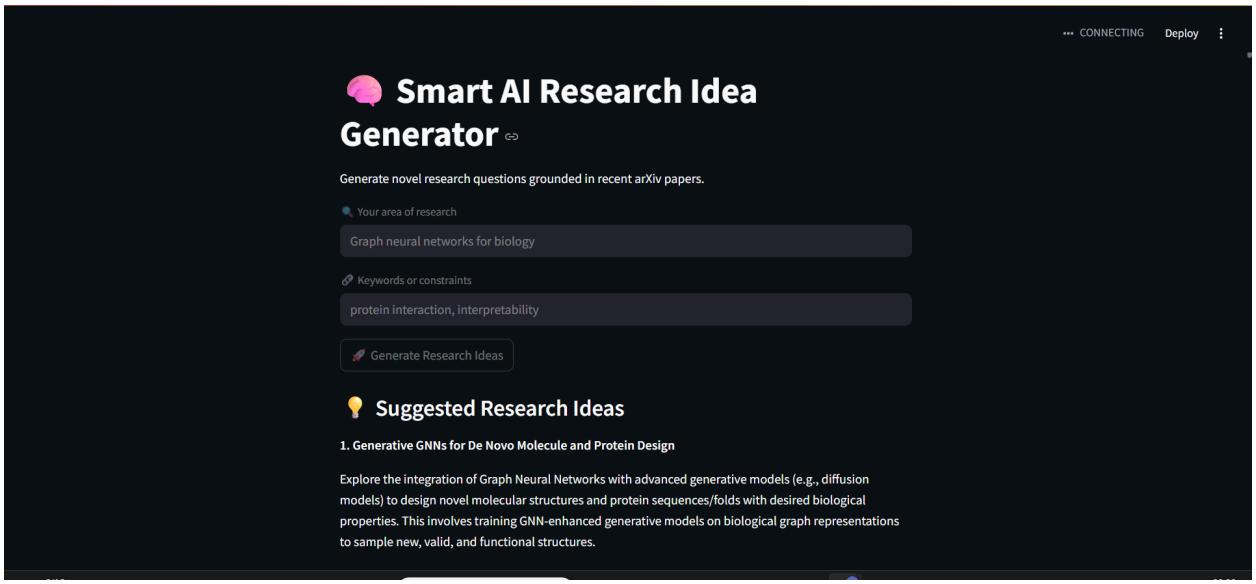
AI Agent for scientific Research

Running the app:

An AI agent has been developed for the purpose of aiding researchers in idea generation. A streamlit frontend is used and Gemini API is used for embedding. The model specifically used is `gemini-embedding-exp-03-07`
A screenshot of the UI/UX is pasted below.



The simple webapp can be run by first installing all requirements in a virtual environment by running the command:
`pip install -r /path/to/requirements.txt`

Then simply run the following command to start the streamlit app:
`streamlit run app.py`

Working:

The app takes in input of area of research of user. Based on this area of research, the app uses the arxiv API to fetch all research papers with a simple pattern matching. The text of the papers is then embedded and the embeddings stored in a vector database of FAISS and indexed accordingly. The query of user is also embedded and based on a similarity search, the top K most relevant papers are retrieved. These papers are then passed as context into the Gemini `gemini-2.5-flash-preview-04-1` model and the model then outputs a novel idea based on the papers and prompt.

The model then calculates the novelty score of the idea by performing cosine similarity between the new idea proposed and the ideas of the papers. Based on a threshold value, the idea can be classified as truly novel or similar to existing work.

Evaluation:

The evaluation of the model is based on novelty of the ideas generated with respect to existing literature in the field. The cosine similarity scores of all ideas are calculated.