LLM QA Chatbot Builder

Instructions:

In this page you can prepare data for LLM fine-tuning, testing and embedding model finetuning your model. The data can be provided through Excel file or CSV file or directly via web interface. Additionally, data can be parsed from the target website (Data parsing for RAG) to further enhance the model performance.

1. If you want to provide data in Excel file or CSV file for model fine-tuning and testing.

- Create an Excel or CSV file in the data folder and name it `finetune_data.xlsx` or `finetune_data.csv` for finetuning the model.
- Create an Excel or CSV file in the data folder and name it `testing_data.xlsx` or `testing_data.csv` for generating answers using the fine-tuned model.
- `finetune_data.xlsx` or `finetune_data.csv` has two columns: `question` and `answer`. `testing_data.xlsx` or `testing_data.csv` has three columns: `question`, `ground_truth`, `context`.

`finetune_data.xlsx` | `finetune_data.csv`

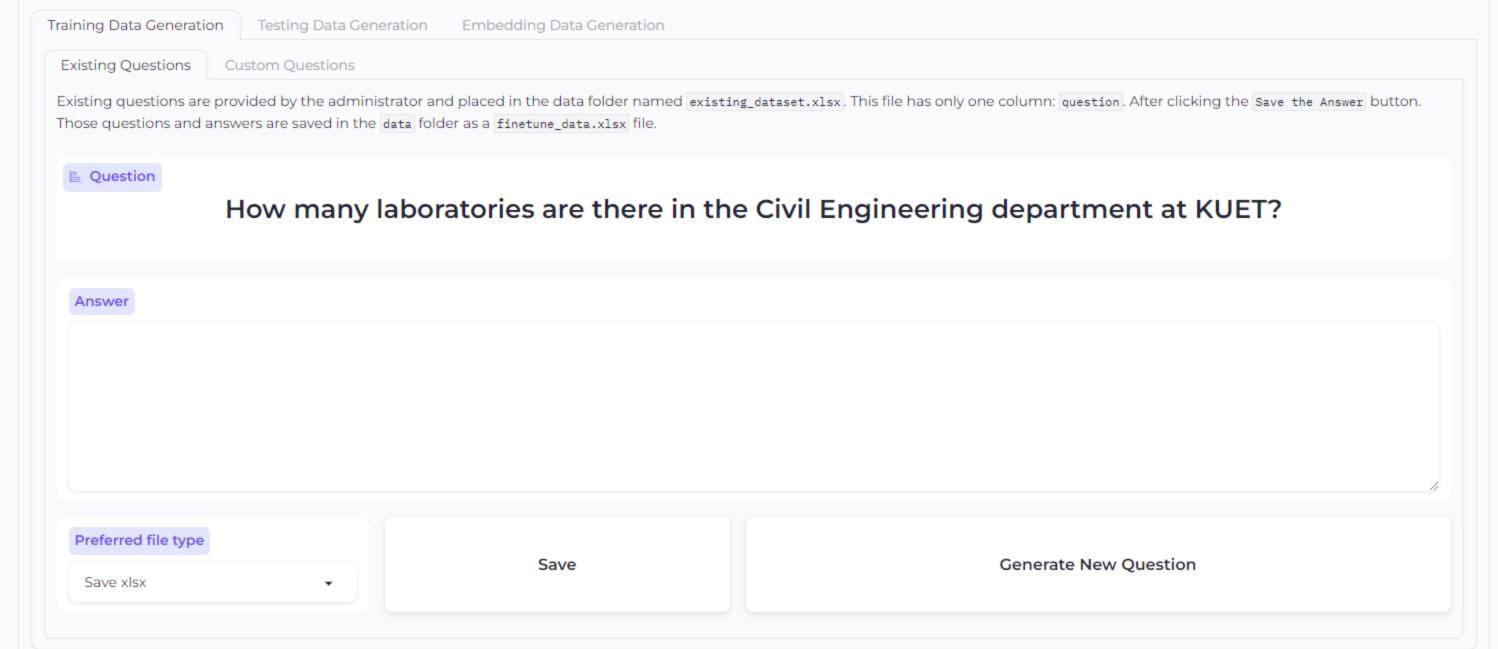| question	answer
What is the history and evolution of KUET?	Khulna University of Engineering & Technology (KUET) originated as Khulna Engineering College in 1974. It transformed into the Bangladesh Institute of Technology (BIT), Khulna in 1986. Later, in 2003, it was upgraded and renamed as Khulna University of Engineering & Technology with the aim of enhancing higher education and research in engineering and technology. The campus spans 101 acres, offering a picturesque setting for academic pursuits. KUET aspires to be a leading engineering university with a focus on teaching, research, and intellectual development.
Where is the KUET campus located, and what surrounds it?	The KUET campus is situated at Fulbarigate, approximately 13 km north of Khulna City along the Khulna-Dhaka highway. The campus is surrounded by various institutions, such as Govt. B. L. College to the south, Teacher's Training College and Technical Training Center to the west, and Jahanabad Cantonment to the north. The Khulna-Dhaka highway runs along the eastern side of the campus.

`testing_data.xlsx` | `testing_data.csv`

question	ground_truth	context
What are the requirements for postgraduate programs at KUET?	Applicants for master's programs need a B.Sc. Eng degree or equivalent in the relevant field, with good academic records. Admissions include written and/or viva-voce tests.	Applicants for the masters programs must have B.Sc. Eng. degree or equivalent in the relevant field from a recognized University/Institute with good academic records. Students who have higher research aptitude are welcome to the program. KUET invites application twice in a year (January Term and July Term). The respective departments arrange an interview at a suitable time to select candidates for this program. The selected candidates have to take admission by depositing prescribed amount of money to the bank. Full-time and part-time students are there in this program. For full time meritorious students financial help can be provided. A part-time student must have consent from his employer to pursue postgraduate studies.
How are earned credits and CGPA improvement handled at KUET?	Courses with a 'D' grade or higher are counted as earned credits. 'F' grades are not counted for GPA but remain on transcripts. Students can repeat a course for improvement but can't earn above a 'B' grade.	The courses in which a student has obtained 'D' or a higher grade will be counted as credits earned by him/her. Any course in which a student has obtained 'F' grade will not be counted towards his/her earned credits calculation. A student who obtains an 'F' grade in any core course in any Term, he/she will have to repeat the course. If a student obtains an 'F' in an optional course he/she may choose to repeat the course or take a substitute course, if available. F grades will not be counted for GPA calculation but will stay permanently on the grade sheet and transcript. When a student will repeat a Backlog course in which he/she previously obtained 'F' grade, he/she will not be eligible to get a grade better than B+ (B plus) in such a course. A student obtaining D grade in a course, will be allowed to repeat the course for the purpose of grade improvement if CGPA of the student falls below 2.20 In such case he/she will be awarded the new grade thus he/she obtains or retains his/her previous grade if he/she fails. A student obtaining 'C' or a better grade in a course will not be allowed to repeat the course for the purpose of grade improvement if CGPA of the student falls below 2.20. Absence in Term final examination will result 'F' grade unless he/she has withdrawn from the Term as per Article 11.8.

2. You can use the below interface to create the dataset for training and testing models.

Training Data Generation | Testing Data Generation | Embedding Data Generation

Existing Questions | Custom Questions

Existing questions are provided by the administrator and placed in the data folder named `existing_dataset.xlsx`. This file has only one column: `question`. After clicking the `Save the Answer` button. Those questions and answers are saved in the `data` folder as a `finetune_data.xlsx` file.

📄 Question

How many laboratories are there in the Civil Engineering department at KUET?

Answer

Preferred file type

Save xlsx ▾

Save	Generate New Question

3. Data parsing for RAG

Enter Link to Parse Data for RAG

To provide the link for parsing the data from the website, this link can help create RAG data for the model.

Parse Data