

RAG Chatbot using Together API (Python)

This is a simple **Retrieval-Augmented Generation (RAG)** chatbot that:

- Loads knowledge from a `.txt` file
- Accepts user input via terminal (CLI)
- Uses the **Together API** to generate responses
- Runs on the **Mixtral-8x7B-Instruct** model

Features

- Loads a local knowledge base (`knowledge.txt`)
- Context-aware GPT answers only based on that content
- Lightweight, pure Python solution (no vector database needed)
- Easily expandable to PDF/embeddings

File Structure

```
rag_chat_project/  
├─ rag_chat.py          # Main chatbot script  
├─ knowledge.txt        # Your custom knowledge base  
└─ .env                 # API key (ignored by git)
```

.env File Format


Create a file called `.env` in the root directory:


```
TOGETHER_API_KEY=sk-abc123456...
```


Example `knowledge.txt`

```
Gravity is a force that pulls objects toward each other.  
The Moon's gravity is about 1/6th as strong as Earth's gravity.  
The Sun's mass gives it the strongest gravity in our solar system.
```

Example Conversation

 RAG Chat. Ask anything related to the knowledge.txt file!

 You: How strong is Moon's gravity?

 GPT: The Moon's gravity is about 1/6th as strong as Earth's gravity.

Setup Instructions

1. Install dependencies

```
pip install python-dotenv requests
```

2. Run the bot

```
python3 rag_chat.py
```

Notes

- This is a simple proof of concept (RAG without retrieval)
- It can be extended to use FAISS/Sentence Transformers for smarter context lookup
- Also tested via CLI and Telegram integration as part of a full AI tooling case

Tech Used

- Python 3.10+
- Together API (<https://api.together.xyz>)
- Mixtral-8x7B-Instruct
- dotenv + requests

Author

Built by VikNo as part of AI toolchain demonstration.