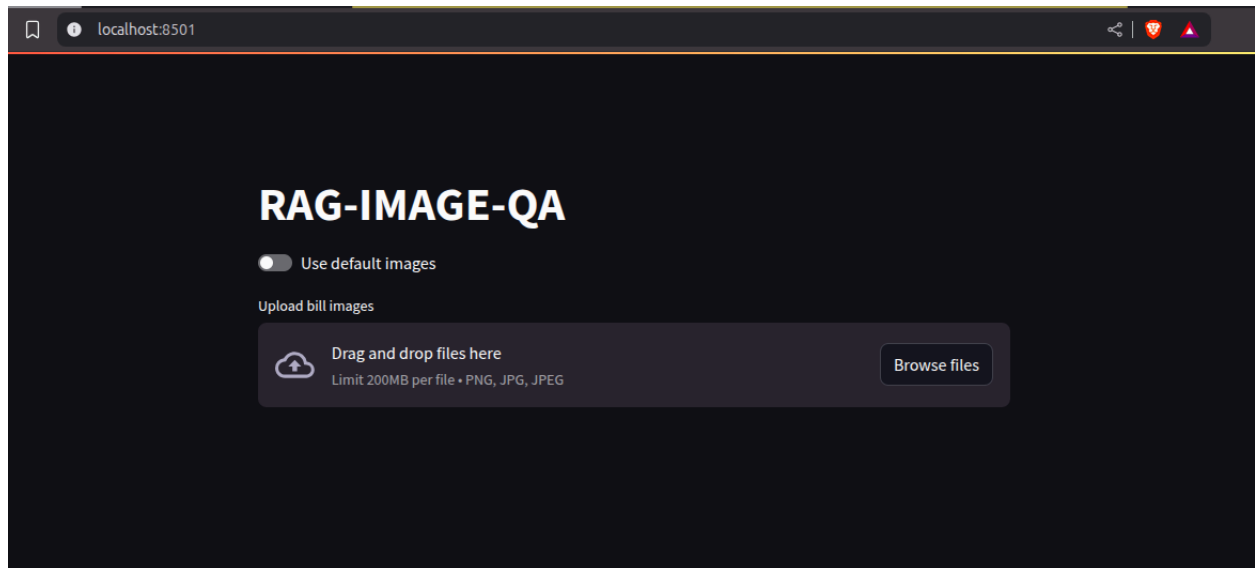


UI Specification

Application: RAG-IMAGE-QA
Author: Yash Meshram
Date: 22/05/2025

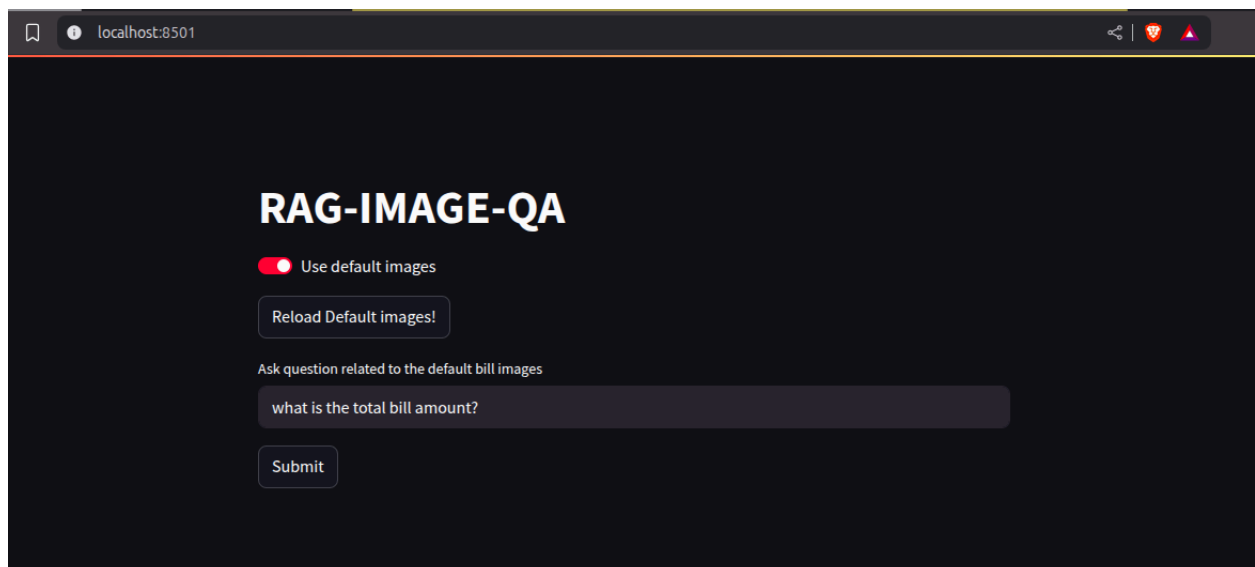
Home Screen:

This will be the screen users will see when open the application.



Functionality:

- Users can upload their own images (one or more) for analysis
- Users can use the default images present in the application by enabling the toggle at the top (below the title).



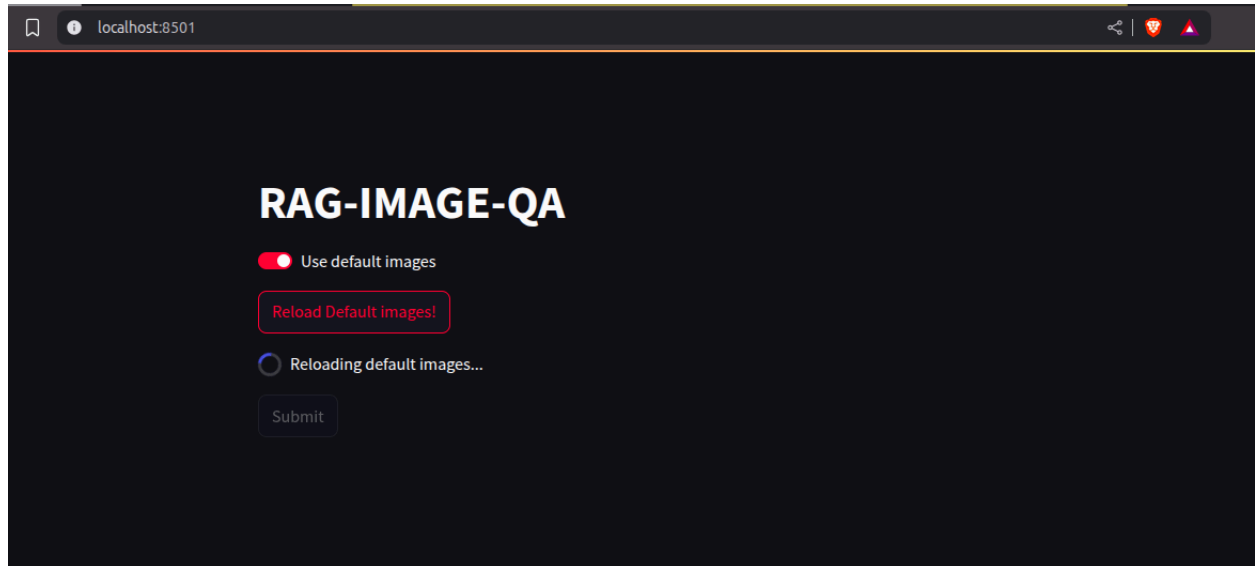
Functionality:

- Reload Default Images: Will re-parse the bill images (present in the data/default/Images) folder. This will be use if any images are been added/removed/edited.

- Or user can directly interact with the image data by asking the questions (eg: What is the total bill amount)

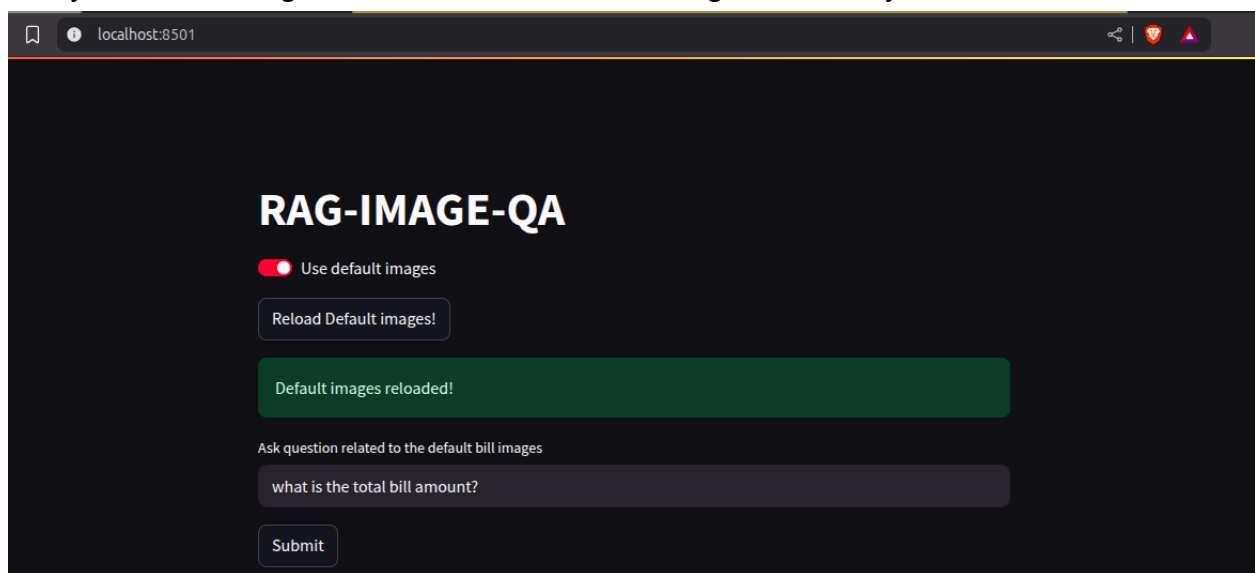
Reloading Default Images:

While reloading the question text box and submit button will be disabled.



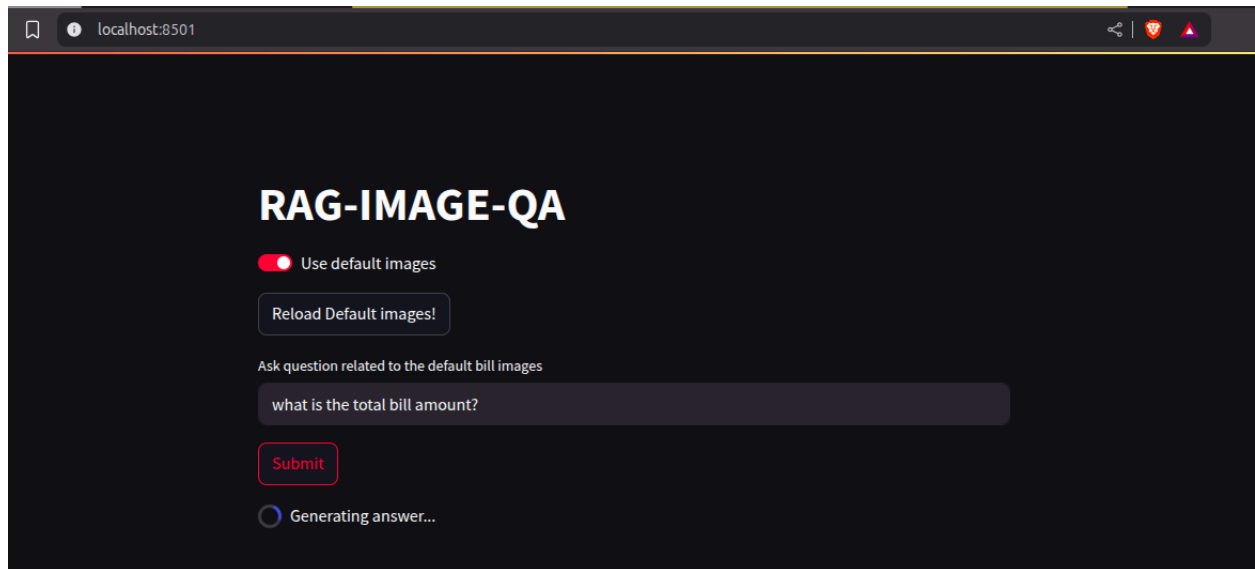
Reloading Default Images completed:

After the reloading is completed the message will be displayed “Default images reloaded!”. Now whatever question the user will ask the answer will be based on the newly reloaded images from the “data/default/Images” directory.

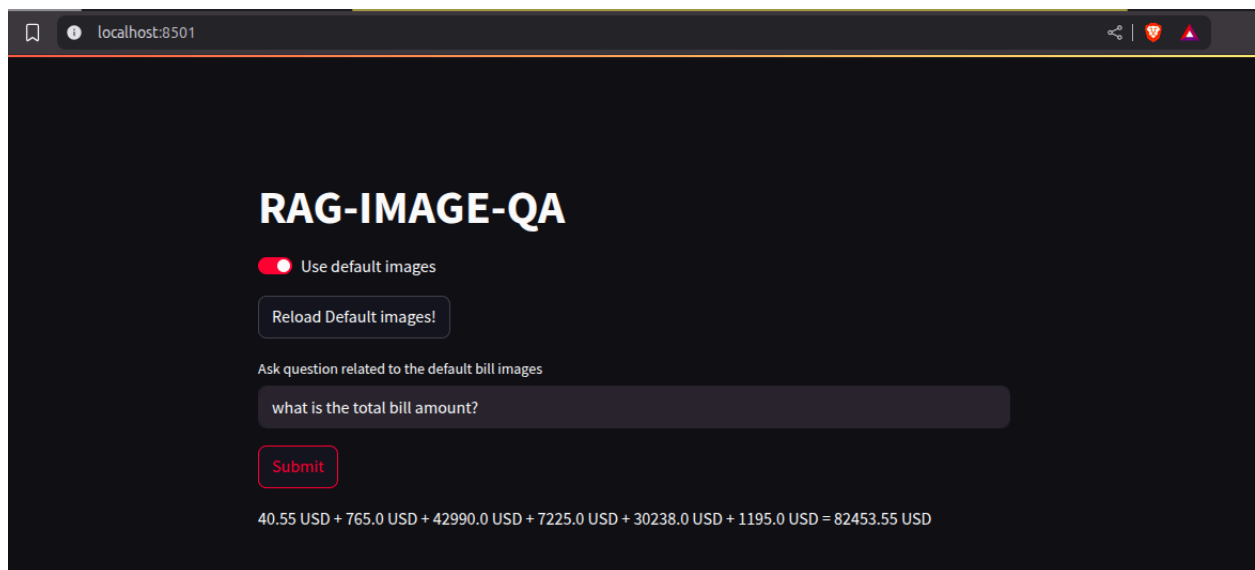


Submit:

Once the user will submit the question the model will take some time to respond. Till then the user will see the loading icon with text 'Generating answer...'

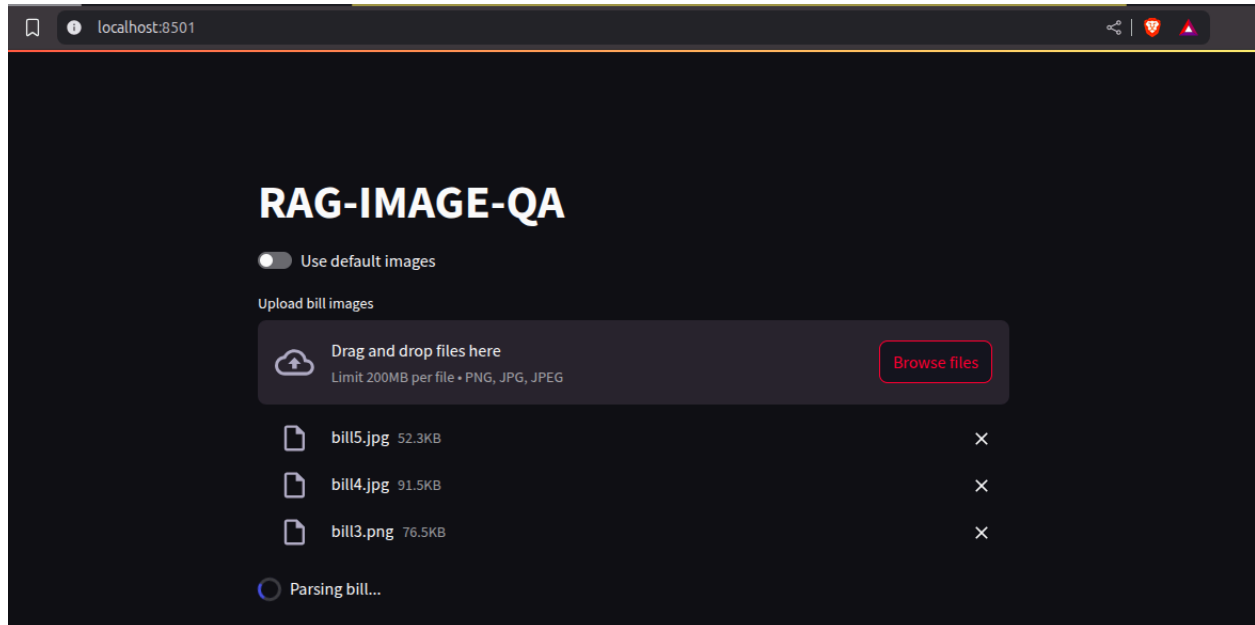


The answer will be displayed just below the Submit button.

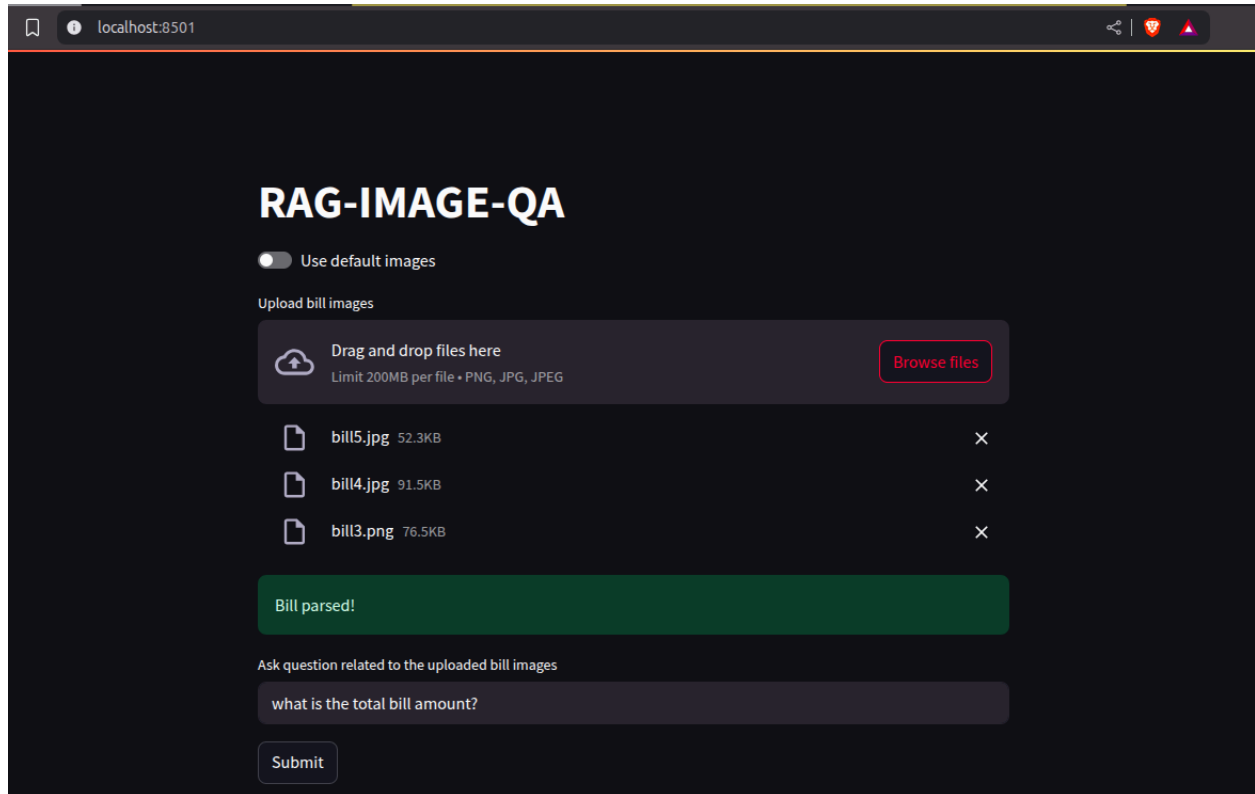


Uploading the images:

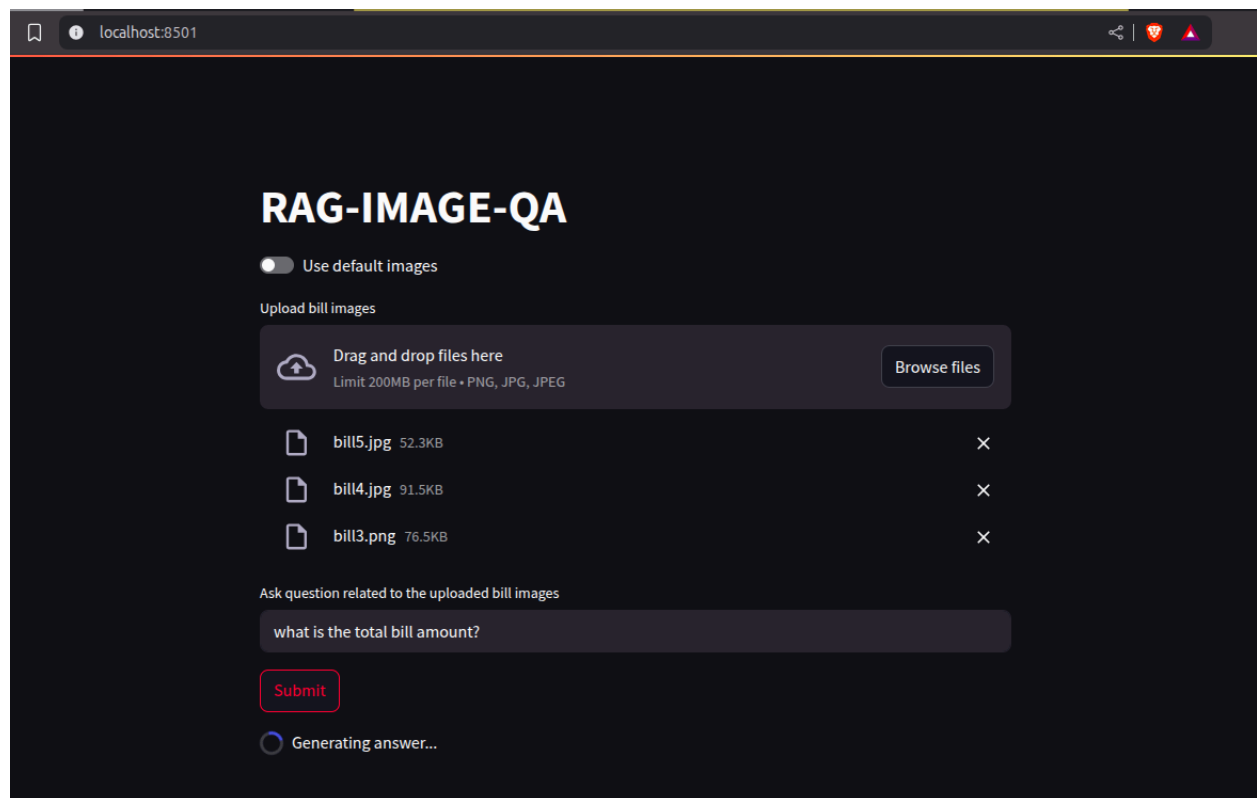
Once the user uploads the images application will take some time to analyze the bill images and extract relevant information from it. Till then the user will see the loading icon with text 'Parsing bill...'



Once the application completely analyzes and extracts relevant information from the text the success message will be displayed 'Bill Parsed!'. After that the question text box and submit button will also be displayed. Now user can type any question related to the uploaded bill images.



Once users type its question and hit submit the model will take some time to answer. Till then the user will see the loading icon with text 'Generating answer...'



The answer will be displayed just below the Submit button.

The screenshot shows a web browser at localhost:8501 displaying the 'RAG-IMAGE-QA' application. At the top, there is a toggle switch labeled 'Use default images' which is currently turned off. Below this is a section titled 'Upload bill images' containing a file upload area with the text 'Drag and drop files here' and 'Limit 200MB per file • PNG, JPG, JPEG', along with a 'Browse files' button. Three files are listed: 'bill5.jpg' (52.3KB), 'bill4.jpg' (91.5KB), and 'bill3.png' (76.5KB), each with a delete icon. Below the file list is a text input field with the question 'what is the total bill amount?' and a red 'Submit' button. The answer displayed at the bottom is '40.55 USD + 2765.0 USD + 30238.0 USD = 33043.55 USD The total bill amount is 33043.55 USD.'

localhost:8501

RAG-IMAGE-QA

☐ Use default images

Upload bill images

Drag and drop files here
Limit 200MB per file • PNG, JPG, JPEG

Browse files

bill5.jpg 52.3KB ×

bill4.jpg 91.5KB ×

bill3.png 76.5KB ×

Ask question related to the uploaded bill images

what is the total bill amount?

Submit

40.55 USD + 2765.0 USD + 30238.0 USD = 33043.55 USD The total bill amount is 33043.55 USD.

If the question is not related the the bill images like shown in the below image. The response will be “No information in the provided images”

This screenshot shows the same 'RAG-IMAGE-QA' application interface. The 'Use default images' toggle is still off. The file upload section and the list of three bill images are identical to the previous screenshot. However, the text input field now contains the question 'What is relu activation function'. The red 'Submit' button is present, and the response displayed at the bottom is 'No information in the provided Images.'

localhost:8501

RAG-IMAGE-QA

☐ Use default images

Upload bill images

Drag and drop files here
Limit 200MB per file • PNG, JPG, JPEG

Browse files

bill5.jpg 52.3KB ×

bill4.jpg 91.5KB ×

bill3.png 76.5KB ×

Ask question related to the uploaded bill images

What is relu activation function

Submit

No information in the provided Images.