Database Project (Fall 2023)

Homework #10 (50pts, Due date: Dec 6)

Student ID: 2020315798

Student Name: Choi Jin Woo

Instruction: In this homework, we provide a jupyter notebook file (DBP_Homework10.ipynb). You should follow the instructions in these documents. Only the provided code's '*EDIT HERE*' sections must be edited. You must *DELETE* your *GOOGLE_API_KEY* before submitting.

Submission Guide: Submit two files as follows:

-DBP_Homwork10_StudentID.zip
- DBP_Homwork10_StudentID.ipynb
- DBP_Homwork10_StudentID.pdf

1. [20pts] You want to provide a Question Answering service based on Retrieval-Augmented Generation (RAG) for a given webpage. To raise the satisfaction of users, you have to do prompt engineering. Prompt engineering is to make an optimal prompt for a given task.

So, you should implement the RAG pipeline as given parameters.

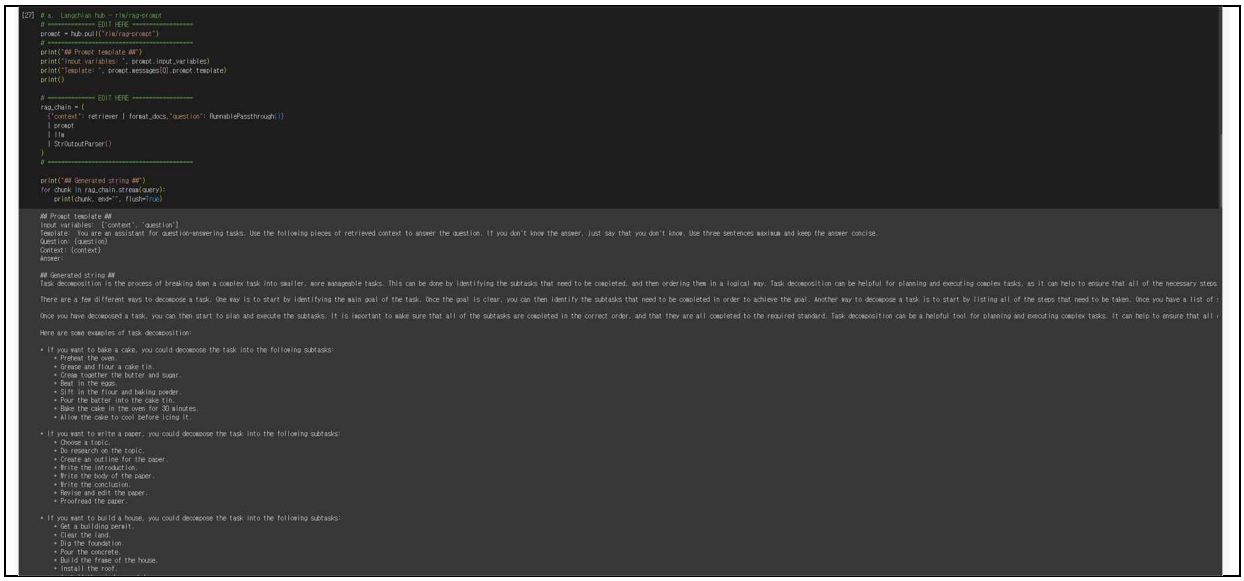
Parameters
Splitter: RecursiveCharacterTextSplitter
Chunk size for splitter: 800
Chunk overlap size for splitter: 100
Vector store: Chromadb
Embedding for vector store: GooglePalmEmbeddings
Retriever: Same as vector store
Retriever search type: similarity
Retriever search kwargs k: 6

Next, you should do prompt engineering for the given query string *"What is Task Decomposition?"*. **You should apply three prompts to the RAG pipeline: two prompts from Langchain hub (https://smith.langchain.com/hub) and one prompt on your own.**

a. Langchian hub - rlm/rag-prompt

[Answer]

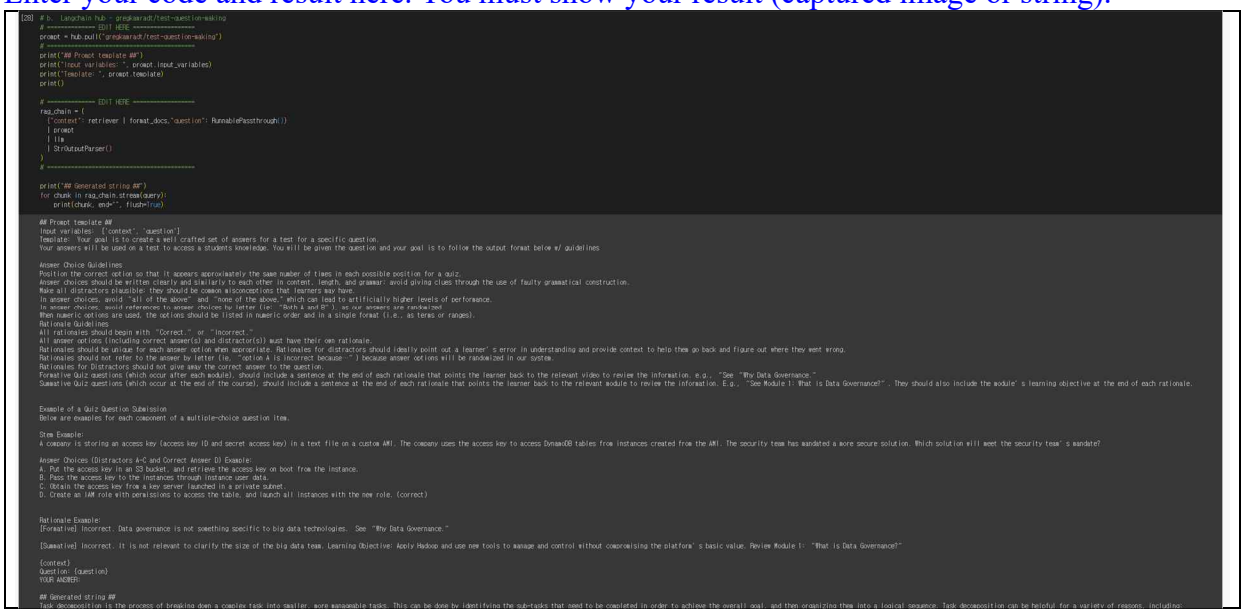
Enter your code and result here. You must show your result (captured image or string).



b. Langchain hub – gregkamradt/test-question-making

[Answer]

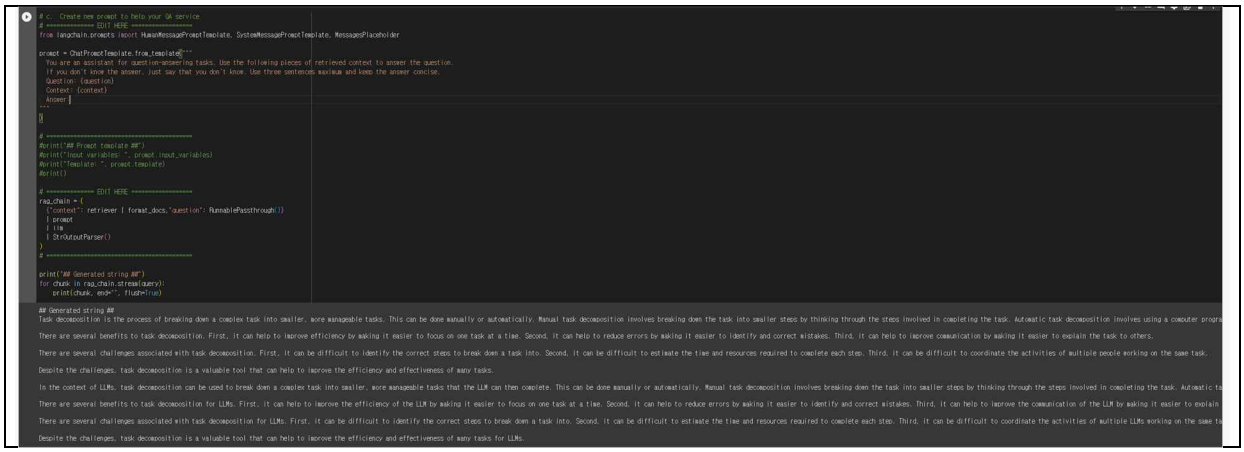
Enter your code and result here. You must show your result (captured image or string).



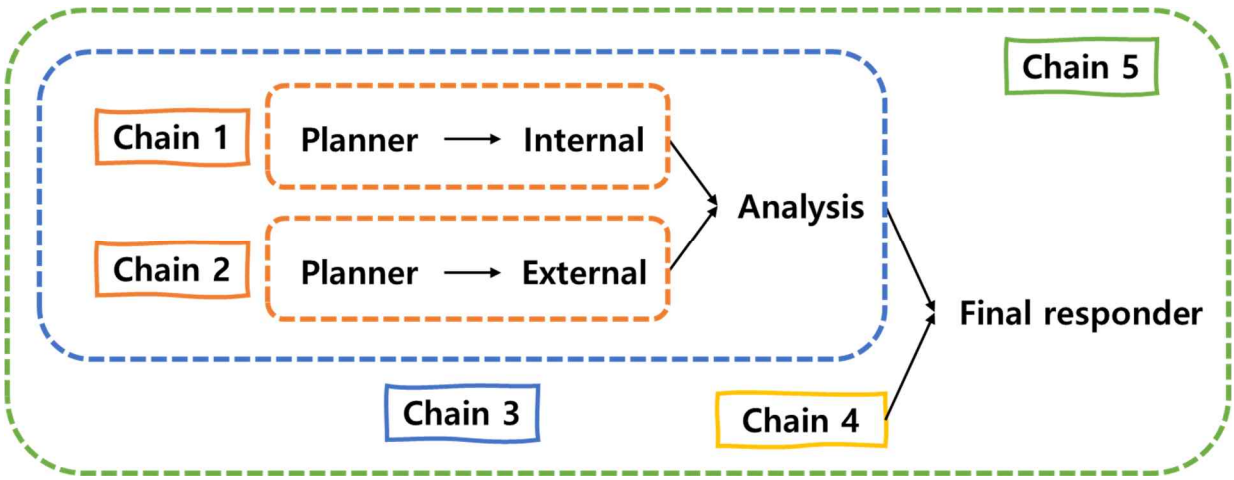
c. Create a new prompt to help your QA service.

[Answer]

Enter your code and result here. You must show your result (captured image or string).



2. [30pts] Implement a RAG multiple chain pipeline. The purpose of this pipeline is to find information about stakeholders for a given project. **The pipeline contains five chains.** The details of the chains are shown in the below figure.



The implementation parameters are same as question 1 in `text_splitter`, `vectorstore`, `retriever`, `llm` variables. The prompts for `planner`, `internal` and `external` are in the given ipynb file. You can use `from_tempate()` method for them. But you have to use `from_messages()` method for `analysis` and `final_responder`. The specific messages are shown in the table below.

analysis	("system", "Generate a stakeholder analysis map for a given project."), ("human", "The details of a given project is following: {result3}"), ("ai", "Internal stakeholders:\n{result1}\n\nExternal stakeholders:\n{result2}")
final_responder	("system", "Generate a final response given the information."), ("ai", "{a_response}"), ("human", "Common pitfalls:\n{a_pitfalls}")

Read the given ipynb file carefully and write all your codes and results here.
[Answer]
Enter your code and result here. You must show your result (captured image or string).

```
planner_prompt_template = "I want to do the project management for {input}"
planner_prompt = ChatPromptTemplate.from_template(planner_prompt_template)

# ============= EDIT HERE =================
# Planner
# Use `from_template` method
planner = (
    planner_prompt
    | llm
    | StrOutputParser()
)
# ========================================
```

```
internal_prompt_template = """Answer the question based only on the following context:
{context}

Question: {question}

Answer in the following: {base}
"""
internal_prompt = ChatPromptTemplate.from_template(internal_prompt_template)

# ============= EDIT HERE =================
# Internal
# Use `from_template` method
# Use `question`, `retriever`, `format_docs` for `context`
internal = (
    {"context": itemgetter("question1") | retriever | format_docs, "question": itemgetter("question1"), "base": planner}
    | internal_prompt
    | llm
    | StrOutputParser()
)
# ========================================
```

```
external_prompt_template = """Answer the question based only on the following context:
{context}

Question: {question}

Answer in the following: {base}
"""
external_prompt = ChatPromptTemplate.from_template(internal_prompt_template)

# ============= EDIT HERE =================
# External, it is similar to `internal`
# Use `from_template` method
# Use `question`, `retriever`, `format_docs` for `context`
external = (
    {"context": itemgetter("question2") | retriever | format_docs, "question": itemgetter("question2"), "base": planner}
    | internal_prompt
    | llm
    | StrOutputParser()
)
# ========================================
```

```
# ============= EDIT HERE =================
# Analysis
# Use `from_messages` method
analysis = (
    ChatPromptTemplate.from_messages(
        [
            ("system", "Generate a stakeholder analysis map for a given project."),
            ("human", "The details of a given project is following: {result3}"),
            ("ai", "Internal stakeholders:\n{result1}\n\nExternal stakeholders:\n{result2}")
        ]
    )
    | llm
    | StrOutputParser()
)
# ========================================
```

```
# ============= EDIT HERE =================
# Final responder
# Use `from_messages` method
final_responder = (
    ChatPromptTemplate.from_messages(
        [
            ("system", "Generate a final response given the information."),
            ("ai", "{a_response}"),
            ("human", "Common pitfalls:\n{a_pitfalls}")
        ]
    )
    | llm
    | StrOutputParser()
)
# ========================================
```

```
# You can check that you implement `planner` correctly using this cell
planner.invoke({"input": "scrum"})
```

```
'Scrum is a framework for developing and managing complex products. It is a lightweight framework with a focus on iterative development, team collaboration, and accountability. Scrum teams are cross-functional and self-organizing, and they work in sprints of two weeks or less.\n\nThe Scrum Master is responsible for ensuring that the Scrum process is followed and that the team is able to deliver the product backlog. The Product Owner is responsible for prioritizing the product backlog and ensuring that the team is working on the most important features. The Development Team is responsible for developing the product according to the Scrum process.\n\nTo do project management for Scrum, you need to understand the Scrum framework and the roles and responsibilities of each team member. You also need to be able to manage the Scrum process, including sprint planning, sprint execution, and sprint retrospectives.\n\nHere are some tips for project management for Scrum:\n\n* **Understand the Scrum…'
```

```
# ============= EDIT HERE =================
# Chain 1
chain1 = internal
# ========================================

# You can check that you implement `chain1`, `internal` correctly using this cell
chain1.invoke({"input": "scrum", "question1": "List the internal stakeholders."})
```

```
'Internal stakeholders are people within an organization who are affected by or have an interest in the outcome of a project. They can include:\n\n* Project manager\n* Project team members\n* Project sponsor\n* Executive leaders\n* Other cross-functional internal teams\n\nIt is important to identify and manage internal stakeholders effectively in order to ensure the success of a project. This can be done by:\n\n* Communicating regularly with stakeholders\n* Keeping stakeholders informed of progress and changes\n* Involving stakeholders in decision-making\n* Managing stakeholder expectations\n\nBy taking the time to manage internal stakeholders effectively, project managers can help to ensure that projects are completed on time, within budget, and to the satisfaction of all stakeholders.\n\nHere are some additional tips for managing internal stakeholders:\n\n* **Identify your stakeholders early on.** The first step in managing internal stakeholders is to identify who they are. This inclu…'
```

```
# ============= EDIT HERE =================
# Chain 2
chain2 = external
# ========================================

# You can check that you implement `chain2`, `external` correctly using this cell
chain2.invoke({"input": "scrum", "question2": "List the external stakeholders."})
```

```
'External stakeholders are people or groups outside of the organization that are affected by the project or have an interest in its success. They can include customers, suppliers, partners, regulators, and the public.\n\nExternal stakeholders can play a significant role in the success of a project. They can provide feedback on the project, help to identify risks and opportunities, and advocate for the project within their own organizations. It is important to identify and manage external stakeholders early in the project lifecycle. This will help to ensure that their needs are met and that they are supportive of the project.\n\nThere are a number of ways to manage external stakeholders. One way is to create a stakeholder register. This is a document that lists all of the project's external stakeholders and their contact information. The stakeholder register should also include information on each stakeholder's interest in the project, their level of influence, and any potential risks or…'
```

```
# ============= EDIT HERE =================
# Chain 3
# For `result3`, use the following line
# "result3": itemgetter("question3") | retriever | format_docs
chain3 = (
    {
        "result3": itemgetter("question3") | retriever | format_docs,
        "result2": chain2,
        "result1": chain1
    }
    | analysis
)
# ========================================

# You can check that you implement `chain3`, `analysis` correctly using this cell
chain3.invoke({
    "input": "scrum",
    "question1": "List the internal stakeholders.",
    "question2": "List the external stakeholders.",
    "question3": "How to do a stockholder analysis?"
    })
```

```
'Stakeholder analysis is a process of identifying and assessing the needs, interests, and potential impact of all the people or groups who may be affected by a project. It is an essential part of project management, as it helps to ensure that all stakeholders are considered and that their needs are met.\n\nStakeholders can be internal or external to the organization. Internal stakeholders are those who are directly involved in the project, such as the project manager, project team members, and project sponsor. External stakeholders are those who are not directly involved in the project, but who may be affected by its outcome, such as customers, suppliers, partners, and the public.\n\nThere are a number of benefits to conducting a stakeholder analysis. First, it helps to identify all of the stakeholders who may be affected by the project. This is important, as it ensures that all stakeholders are considered and that their needs are met. Second, it helps to assess the needs and interests …'
```

[134]
```
# Chain 4 (Don't modify)
chain4 = (
    {"question": itemgetter("question4"),
     "context": itemgetter("question4") | retriever | format_docs}
    | ChatPromptTemplate.from_template("{context} {question}")
    | llm
    | StrOutputParser()
)

chain4.invoke({"question4": "List the common stakeholder mapping pitfalls."})
```

```
'Here are the common stakeholder mapping pitfalls:\n\n* **Lack of boundaries:** Overeager project stakeholders can cause scope creep. To avoid this, implement a change control process.\n* **Not including stakeholders:** During the initial stakeholder identification phase, you forgot about a stakeholder who has a high interest in this initiative. To avoid this, involve your project team during the stakeholder identification process.\n* **Bringing stakeholders in too late:** You already kicked off your project, and then started your stakeholder analysis. To avoid this, create a stakeholder analysis map up front.\n* **Not communicating effectively:** You're not communicating effectively with your stakeholders, which is leading to misunderstandings and frustration. To avoid this, create a communication plan and stick to it.\n* **Not managing expectations:** You're not managing your stakeholders' expectations, which is leading to disappointment and frustration. To avoid this, set clear expec...'
```

[135]
```
# ============ EDIT HERE ================
# Chain 5
chain5 = (
    {
        "a_response": chain3,
        "a_pitfalls": chain4
    }
    | final_responder
)
# =======================================

# You can check that you implement `chain5`, `final_responder` correctly using this cell
chain5.invoke({
    "input": "scrum",
    "question1": "List the internal stakeholders.",
    "question2": "List the external stakeholders.",
    "question3": "How to do a stockholder analysis map?",
    "question4": "List the common stakeholder mapping pitfalls."
    })
```

```
'Stakeholder analysis is a critical part of project management. By understanding the needs and expectations of stakeholders, project managers can develop strategies to manage them effectively. However, there are a number of common pitfalls that can derail stakeholder management efforts.\n\nOne common pitfall is **not including stakeholders**. This can happen if stakeholders are not identified early in the project or if they are not given an opportunity to provide input. When stakeholders are not included, they may feel like they are not being heard and may become resistant to the project.\n\nAnother common pitfall is **not communicating effectively**. Stakeholders need to know what is going on with the project and they need to know it in a timely manner. If communication is not effective, stakeholders may become frustrated and may start to spread rumors or negative information about the project.\n\n**Not managing expectations** is another common pitfall. Stakeholders need to know what t...'
```