

2024 삼성 무선사업부 생성형 AI 실습 과제 가이드

과제 제출 형식

- 주피터 노트북 파일 제출 (.ipynb)
- 파일명 사번_이름 (0102998_박무선.ipynb)
- 파일내 최종 결과물 json 형태로 출력

```
# 아래 형식에 맞춰 결과를 출력해주세요
# key 값이 다음과 완전히 일치해야 채점이 진행됩니다 : name, team , ID_number,
ai_response1, ai_response2, ai_response3

final_result = {
    "name": "박무선",
    "team" : "MX 팀",
    "ID_number": "0102998_",
    "ai_response1": ai_response1,
    "ai_response2": ai_response2,
    "ai_response3": ai_response3
}

# JSON으로 출력
import json
print(json.dumps(final_result, ensure_ascii=False))
```

과제 공통 제약사항

- Python 3.10 환경의 Jupyter Notebook에서 작업한다.
- Langchain 버전 0.1.8 이상을 사용한다.
- LLM 모델은 AzureOpenAI GPT API를 활용한다.
- 언어 모델은 gpt-35-turbo, gpt-35-turbo-instruct 둘 중 하나를 사용하거나 두 개 모두 사용할 수 있다.
- Embedding 모델은 text-embedding-ada-002를 사용한다.
- Vector store는 ChromaDB 또는 FAISS를 사용한다.
- 최대 출력 토큰 수는 512 이하로 설정한다.
- 주어지는 사용자 질문(query)에 대한 결과를 ai_response로 반환한다.
- 주어지는 질문 개수 3개에 맞는 답 3개를 모두 반환해야한다. 3개 답이 모두 제출되지 않으면 무효 처리된다.
- 제공되는 데이터 외 다른 데이터는 사용할 수 없다.
- 필요한 패키지가 있는 경우 %pip를 이용해 설치한다.

- **AzureOpenAI key 연동 코드는 아래와 동일하게 작성한다**

```
from config import (
    AZURE_OPENAI_API_VERSION,
    AZURE_OPENAI_ENDPOINT,
    AZURE_OPENAI_KEY
)
os.environ["OPENAI_API_TYPE"] = "azure"
os.environ["AZURE_OPENAI_API_KEY"] = AZURE_OPENAI_KEY
os.environ["AZURE_OPENAI_ENDPOINT"] = AZURE_OPENAI_ENDPOINT
os.environ["OPENAI_API_VERSION"] = AZURE_OPENAI_API_VERSION
```

수행 과제 : PDF 문서 기반 질의응답

[최종 결과물]

- PDF 문서를 기반으로 질의응답 기능을 구현한다.

[기능 요구사항]

- 사용자 질문에 대해 짧고 간략한 답변을 제시한다.
- 문서에 나와있지 않은 내용에 대해서는 명시되어있지 않다고 답변해야한다.
- 모든 답변은 “질병관리청에서 알려드립니다.”로 시작한다.

[제공 문서]

- 2024GenAI_Mineral.pdf

[구현 사항] 아래 3개 질문에 대한 LLM 모델의 답 3개를 JSON 형태로 출력한다.

```
* query1 = "파일의 주요 내용을 1줄로 요약해줘"
* query2 = "2021년 기준 20대의 1일 평균 엽산 섭취량을 알려줘"
* query3 = "질병관리청에서 추천하는 비타민12가 많이 들어있는 음식 3가지를 알려줘"
```