



# MedAdvisor KR

The MedAdvisor aims to provide users with accurate and prompt information about medications, including their effects.

자연어처리 01분반 김예원 이지원 조윤영

# Project Overview

## Goal

Develop a chatbot that allows users to easily access medication information.

## Key Features

Natural language understanding and data-driven provision of medication details.

### Chatbot (Demo)

You:

식욕감퇴할 때 먹는 약과 그 약의 주의사항은?

Send

 Bot:

식욕부진 증상이 있을 때 먹는 약은 '트레스오릭스워트정'입니다. 이 약의 주의사항은 과민증 환자, 녹내장, 요저류 환자, 고령자, 쇠약자, 신생아, 갈락토오스 불내성, Lapp 유당분해효소 결핍증 또는 포도당-갈락토오스 흡수장애 등의 유전적인 문제가 있는 환자는 이 약을 복용하지 말아야 합니다. 또한, 복용 후 졸음이 나타날 수 있으므로 운전 및 기계조작 시 주의해야 합니다.

# Project Overview

## Chatbot (Demo)

You:

소화불



Press Enter to submit form

Send

# Project Background

1

## **Difficulty Accessing Information**

Users often struggle to find reliable and up-to-date medication information from a variety of scattered online and offline sources. This can lead to confusion, uncertainty, and potentially unsafe medication usage.

2

## **Information Complexity**

Medication details can be highly technical, including complex names, dosage instructions, and detailed side effect information. This technical jargon can be challenging for many users to fully understand, especially in urgent situations.

3

## **Lack of Emergency Response**

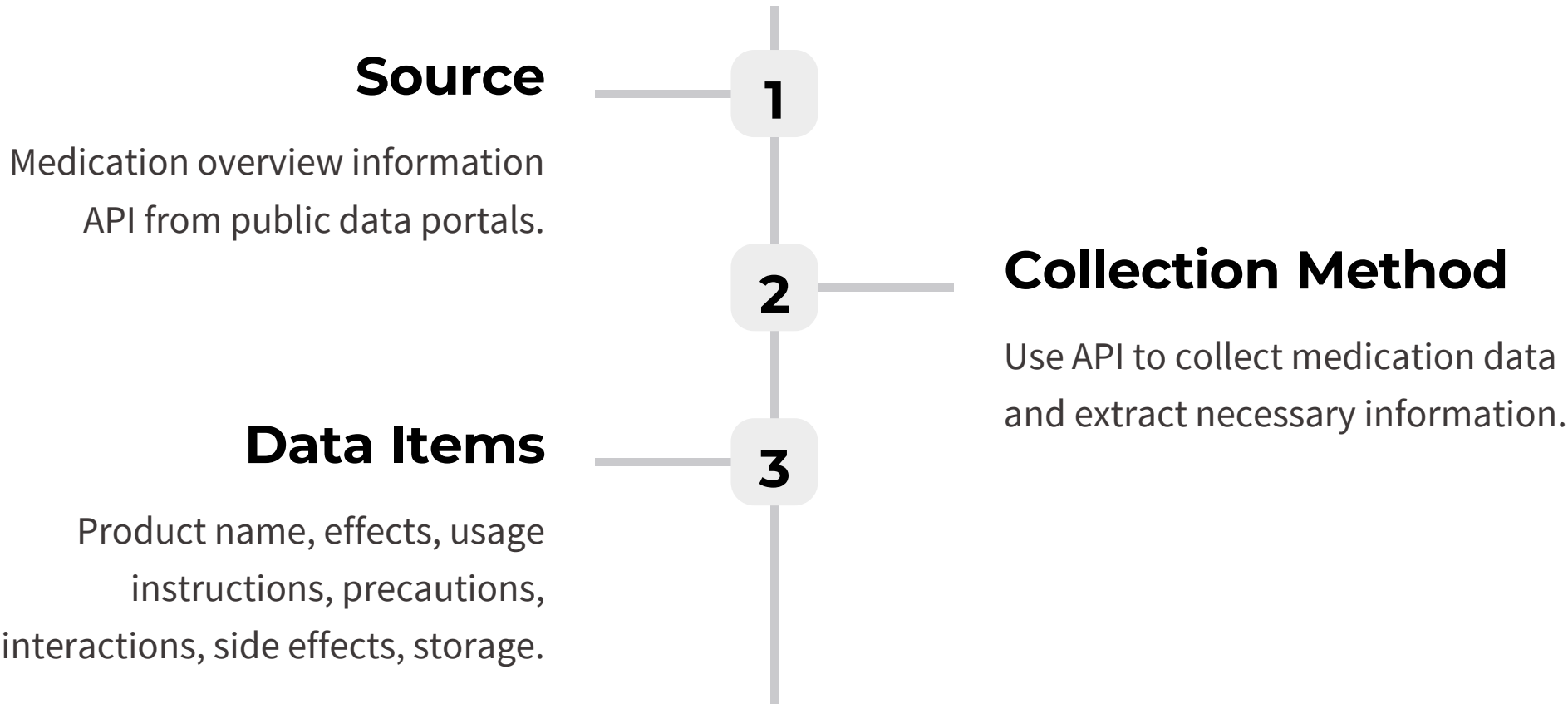
When faced with an urgent medication-related question or concern, it can be extremely difficult for users to quickly obtain accurate and reliable information. This can be a significant problem in time-sensitive emergency situations.

# Core Idea

## **“Symptom-based prescription system”**

A smart medication prescription system that recommends appropriate drugs based on user-entered symptom information.

# Data Collection



# Data Collection

공공데이터포털 >  
“식품의약품안전처\_의약품정보개요”

업체명	제품명	품목기준코드	효능	사용법	주의사항경고	주의사항	상호작용	부작용	보관법	공개일자	수정일자	날알이미지	bizrno
-----	-----	--------	----	-----	--------	------	------	-----	-----	------	------	-------	--------

DATA 공공데이터포털  
. GO . KR

데이터찾기

국가데이터맵

데이터요청

데이터활용

정보공유

이용안내

## 오픈API 상세



XML

JSON

식품의약품안전처\_의약품개요정보(e약은요)

활용신청

공급실적이 있는 일반 의약품 주요정보(업체명, 제품명, 품목기준코드, 효능, 사용법, 주의사항, 상호작용, 부작용, 보관법) 등을 목록으로 조회하는 서비스

👍 26

👎 2

🔖 관심

# Data Collection

We used ...

업 체 명	제 품 명	품목기준 코드	효능	사용법	주의 사항 경고	주의사항	상호 작용	부작 용	보관법	공개일자	수정 일자	날알 이미지	bizrno
0	활명수	195700020	이 약은 식욕감퇴(식욕부진), 위부 팽만감, 소화불량, 과식, 체함, 구역, 구토에 사용됩니다.	만 15세 이상 및 성인은 1회 1병(75 mL), 만 11세이상~만 15세미만은 1회 2/3병(50 mL), 만 8세 이상~만 11세 미만은 1회 1/2병(37.5 mL), 만 5세 이상~만 8세 미만은 1회 1/3병(25 mL), 만 3세 이상~만 5세 미만은 1회 1/4병(18.75 mL), 만 1세 이상~만 3세 미만은 1회 1/5병(15 mL), 1일 3회 식후에 복용합니다. 복용간격은 4시간 이상으로 합니다.	NaN	만 3개월 미만의 젖먹이는 이 약을 복용하지 마십시오.\n\n이 약을 복용하기 전에 만 1세 미만의 젖먹이, 임부 또는 임신하고 있을 가능성이 있는 여성, 카라멜에 과민증 환자 또는 경험자, 나트륨 제한 식이를 하는 사람은 의사 또는 약사와 상의 하십시오.\n\n정해진 용법과 용량을 잘 지키십시오.\n\n어린이에게 투여할 경우 보호자의 지도 감독하에 투여하십시오.\n\n1개월 정도 복용하여도 증상의 개선이 없을 경우 복용을 즉각 중지하고 의사 또는 약사와 상의하십시오.	NaN	NaN	습기와 빛을 피해 실온에서 보관하십시오.\n\n어린이의 손이 닿지 않는 곳에 보관하십시오.	2021-01-29 00:00:00	2024-05-09	NaN	1108100102



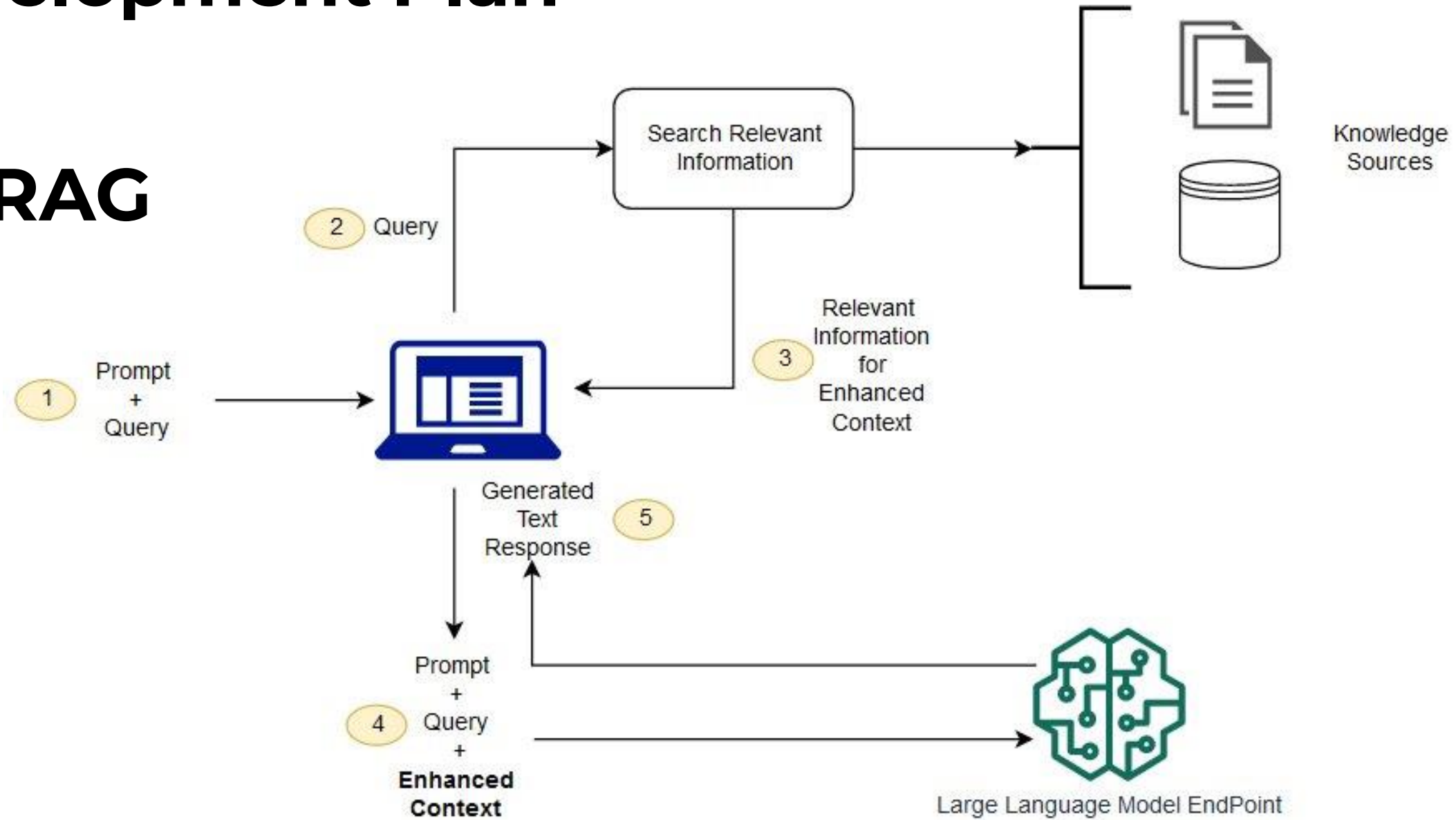
# Development Plan

## 1. In-Context Learning

symptom	count	구내염	361	부기	163
가래	330	구역	270	부종	160
가려움	308	구토	263	불안	55
가슴답답함	2	굳은살	10	비염	155
가슴쓰림	2	기관지염	130	빈혈	108
각막	13	긴장	3	사마귀	6
간질환	75	다래끼	14	상처	309
감기	579	두근거림	5	색소침착	78
건조	254	두드러기	143	소화불량	242
결막염	91	두통	647	속쓰림	268
경련	171	목결림	177	수족냉증	191
고름	19	발진	20		
고혈압	56	배아픔	33		
관절통	893	변비	197		
		복통	94		

# Development Plan

## 2. RAG



# In-Context Learning # Prompt

```
{'role': 'system', 'content': '너는 지금부터 어떤 증상이 있을 때 약을 알맞게 추천해주는 일을 할거야. 내가 증상을 알려주면 그에 맞는 약과 그 약의 효능을 알려줘.'}
```

```
{'role': 'user', 'content': '콧물 증상에 복용할 수 있는 약을 추천해줘. 그리고 그 약의 효능에 대해 알려줄래?'}
```

```
{'role': 'assistant', 'content': '콧물 증상에 복용할 수 있는 약으로는 페니라민정(클로르페니라민말레산염), 판콜에이내복액, 액티피드정이 있습니다.\n\n페니라민정(클로르페니라민말레산염)의 효능은 고초열(꽃가루 알레르기비염), 두드러기, 가려움성 피부질환(습진, 피부염, 피부가려움증, 약물발진), 알레르기 비염, 혈관운동성 코염, 코감기에 의한 재채기, 콧물, 기침, 혈관운동성 부기에 사용합니다.\n\n판콜에이내복액의 효능은 감기의 여러 증상(콧물, 코막힘, 재채기, 인후(목구멍)통, 기침, 가래, 오한(춡고 떨리는 증상), 발열, 두통, 관절통, 근육통)의 완화에 사용합니다.\n\n액티피드정의 효능은 코감기, 알레르기 및 혈관운동성 코염에 의한 재채기, 콧물, 코막힘, 눈물의 완화에 사용합니다.\n\n'}
```

```
{'role': 'system', 'content': '너는 지금부터 어떤 증상이 있을 때 약을 알맞게 추천해주는 일을 할거야. 내가 증상을 알려주면 그에 맞는 약과 그 약의 효능을 알려줘.'}
```

```
{'role': 'user', 'content': '소화불량 증상에 복용할 수 있는 약을 알고싶어. 그리고 그 약의 효능은 뭐야?'}
```

```
{'role': 'assistant', 'content': '소화불량 증상에 복용할 수 있는 약으로는 활명수, 겔마현탁액, 보화소합원(대환,소환)이 있습니다.\n\n활명수의 효능은 식욕감퇴(식욕부진), 위부팽만감, 소화불량, 과식, 체함, 구역, 구토에 사용합니다.\n\n겔마현탁액의 효능은 위염, 식도염, 소화불량, 가슴쓰림, 열공(구멍)헤르니아(탈장), 복부(배부분)팽만감에 사용합니다.\n\n보화소합원(대환,소환)의 효능은 식욕감퇴(식욕부진), 위부팽만감, 소화불량, 과식, 체함, 구역, 구토에 사용합니다.\n\n'}
```

```
{'role': 'system', 'content': '너는 지금부터 어떤 증상이 있을 때 약을 알맞게 추천해주는 일을 할거야. 내가 증상을 알려주면 그에 맞는 약과 그 약의 효능을 알려줘.'}
```

```
{'role': 'user', 'content': '위염 증상이 있어 ㅍㅍ . 어떤 약을 먹는지 알려주고 그 약의 효능은?'}
```

```
{'role': 'assistant', 'content': '위염 증상에 복용할 수 있는 약으로는 겔포스현탁액(인산알루미늄겔), 일양노이겔현탁액(규산알루미늄산마그네슘), 일양노이시린에이정(규산알루미늄산마그네슘)이 있습니다.\n\n겔포스현탁액(인산알루미늄겔)의 효능은 위·십이지장 궤양, 위염, 위산과다(속쓰림, 위통, 구역, 구토)의 제산작용 및 증상의 개선에 사용합니다.\n\n일양노이겔현탁액(규산알루미늄산마그네슘)의 효능은 위산과다 위·십이지장염 위염의 제산작용 및 증상의 개선에 사용합니다.\n\n일양노이시린에이정(규산알루미늄산마그네슘)의 효능은 위산과다 위·십이지장염 위염의 제산작용 및 증상의 개선에 사용합니다.\n\n'}
```

```
{'role': 'system', 'content': '너는 어떤 증상이 있을 때 약을 추천해주는 최고의 약 추천 챗봇이야.'}
```

```
{'role': 'user', 'content': '속쓰림 증상에 복용할 수 있는 약을 추천해줘. 그리고 그 약의 효능을 알려줄래?'}
```

# In-Context Learning # inference

“소화불량 증상에 복용할 수 있는 약을 추천해줘. 그리고 그 약의 효능을 알려줄래?”

```
def get_prediction(prompt):  
    response = openai.ChatCompletion.create(  
        model="gpt-3.5-turbo",  
        messages=prompt,  
        max_tokens = 500 ,  
        temperature=0,  
    )  
    return response.choices[0].message["content"]
```

**base model  
: GPT-3.5**

# 소화불량



Non-existent



Non-existent

소화불량 증상에 복용할 수 있는 약으로는 메틸메틸포스포닐콜린염산염, 메틸메틸포스포닐콜린염산염이 있습니다.

메틸메틸포스포닐콜린염산염의 효능은 소화불량, 위장궤양, 위산과다, 위염, 식도염, 위장관 운동장애, 구역, 구토, 속쓰림, 소화불량, 위장궤양, 위산과다, 위염, 식도염, 위장관 운동장애, 구역, 구토, 속쓰림에 사용됩니다.

# In-Context Learning # inference

“가래 증상에 복용할 수 있는 약을 추천해줘. 그리고 그 약의 효능을 알려줄래?”

```
def get_prediction(prompt):  
    response = openai.ChatCompletion.create(  
        model="gpt-3.5-turbo",  
        messages=prompt,  
        max_tokens = 1000 ,  
        temperature=0,  
    )  
    return response.choices[0].message["content"]
```

**base model  
: GPT-3.5**

# 가래

Similar  
Composition ▲

Similar  
Composition ▲

✗ Non-existent

가래 증상에 복용할 수 있는 약으로는 **류코틴정**, **브롬헥신시럽**, **레보세티리진정**이 있습니다.

류코틴정의 효능은 기관지염, 기관지천식, 만성기관지염, 기관지확장증, 기관지점막염, 기관지점막비염, 기관지점막비염에 따른 기침, 가래, 호흡곤란, 기관지점막비염에 따른 기침, 가래, 호흡곤란에 사용합니다.

브롬헥신시럽의 효능은 기침, 가래, 기관지염, 기관지천식, 기관지점막비염, 기관지점막비염에 따른 기침, 가래에 사용합니다.

레보세티리진정의 효능은 기관지염, 기관지천식, 만성기관지염, 기관지확장증, 기관지점막염, 기관지점막비염, 기관지점막비염에 따른 기침, 가래, 호흡곤란, 기관지점막비염에 따른 기침, 가래, 호흡곤란에 사용합니다.

# RAG + LLM

Since **in-context learning** can generate  
**“non – existent drug names”** as responses,

Let's utilize a **Retrieval Augmented Generation (RAG)** model  
that can leverage external knowledge

# RAG Data Preprocessing

# original data : 4704 drugs info

We used ...

업체명	제품명	품목기준코드	효능	사용법	주의사항 경고	주의사항	상호작용	부작용	보관법	공개일자	수정일자	날알 이미지	bizrno
0	동화약품(주) 활명수	195700020	이 약은 식욕감퇴(식욕부진), 위부팽만감, 소화불량, 과식, 체함, 구역, 구토에 사용됩니다.	만 15세 이상 및 성인은 1회 1병(75 mL), 만 11세 이상~만 15세미만은 1회 2/3병(50 mL), 만 8세 이상~만 11세 미만은 1회 1/2병(37.5 mL), 만 5세 이상~만 8세 미만은 1회 1/3병(25 mL), 만 3세 이상~만 5세 미만은 1회 1/4병(18.75 mL), 만 1세 이상~만 3세 미만은 1회 1/5병(15 mL), 1일 3회 식후에 복용합니다. 복용간격은 4시간 이상으로 합니다.	NaN	만 3개월 미만의 젖먹이는 이 약을 복용하지 마십시오.\n\n이 약을 복용하기 전에 만 1세 미만의 젖먹이, 임부 또는 임신하고 있을 가능성이 있는 여성, 카라멜에 과민증 환자 또는 경험자, 나트륨 제한 식이를 하는 사람은 의사 또는 약사와 상의하십시오.\n\n정해진 용법과 용량을 잘 지켜십시오.\n\n어린이에게 투여할 경우 보호자의 지도 감독하에 투여하십시오.\n\n1개월 정도 복용하여도 증상의 개선이 없을 경우 복용을 즉각 중지하고 의사 또는 약사와 상의하십시오.	NaN	NaN	습기와 빛을 피해 실온에서 보관하십시오.\n\n어린이의 손이 닿지 않는 곳에 보관하십시오.	2021-01-29 00:00:00	2024-05-09	NaN	1108100102

# RAG - Data Preprocessing

```
df = pd.read_csv('/content/drive/MyDrive/NLP/약품_data_nadrop.csv')

idx = df[~(df['효능'].str.startswith('이 '))]['효능'].index
df.loc[idx, '효능'] = df.loc[idx, '효능'].apply(lambda x: '이 약은 '+x)

사용_idx = df[~(df['효능'].str.contains('사용합니다'))]['효능'].index
df.loc[사용_idx[0], '효능'] = df.loc[사용_idx[0], '효능'][:-9]+'합니다'
df.loc[사용_idx[1], '효능'] = df.loc[사용_idx[1], '효능'][:-4] +'에 사용합니다'

def custom_split(text):
    text = text.strip()
    # 숫자와 숫자 사이의 점을 제외하고 분리
    parts = re.split(r'(?!\Wd)\W.(?!Wd)', text)
    parts = list(filter(None, parts))
    return [i.strip() for i in parts]

#### 사용법 ####
pd.set_option('display.max_rows', None)
df['사용법'].fillna("", inplace=True)
df[df['사용법'].str.contains(',')]['사용법']
df['사용법'] = df['사용법'].apply(lambda x: x.replace(',', '@'))
q = df.apply(lambda x: [x['제품명'] + '의 복용법은 ' + i for i in custom_split(x['사용법'])], axis = 1).to_frame()
df['사용법_concat'] = q[0].apply(lambda x: ('. '.join(x)).replace('@', ',')+'.')

#### 주의사항 ####
df['주의사항'].fillna("", inplace=True)
df[df['주의사항'].str.contains(',')]['주의사항']
df['주의사항'] = df['주의사항'].apply(lambda x: x.replace(',', '@'))
q = df.apply(lambda x: [x['제품명'] + '의 주의사항은 ' + i for i in custom_split(x['주의사항'])], axis = 1).to_frame()
df['주의사항_concat'] = q[0].apply(lambda x: ('. '.join(x)).replace('@', ',')+'.').to_frame()

full_data = df.apply(lambda x: (f"다음은 약 {x['제품명']}을 어떤 증상이 있을 때 먹어야 하는지 {x['제품명']}의 효능과 {x['제품명']}의 복용법",
                                x['사용법_concat'], x['주의사항_concat']), axis = 1).to_frame()

# 저장할 디렉토리 생성
output_dir = '약품_full_data_별도의txt'
if not os.path.exists(output_dir):
    os.makedirs(output_dir)

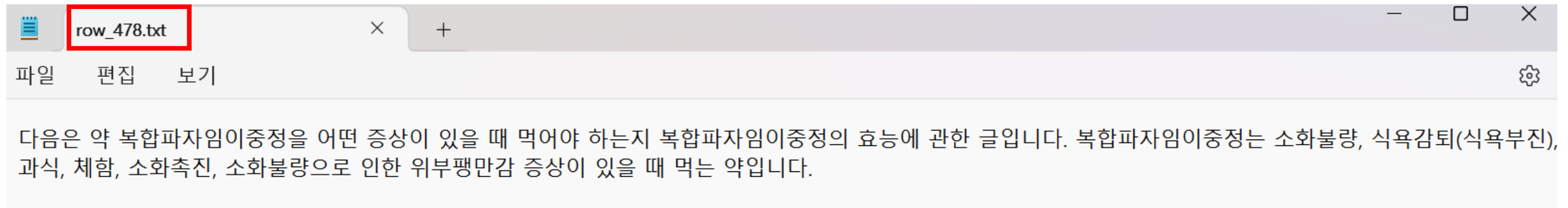
# 데이터프레임의 각 행을 텍스트 파일로 저장
for idx, row in enumerate(full_data[0]):
    file_path = os.path.join(output_dir, f'row_{idx}.txt')
    with open(file_path, 'w', encoding='utf-8') as file:
        file.write(row)
```



# RAG - Construct dataset



For evaluation, generate separate text files containing descriptions that include ['제품명', '효능'] for each drug.



# GPT-3.5 using RAG

## HuggingFaceEmbeddings object using the 'ko-sroberta-multitask' Korean language model

```
model_huggingface = HuggingFaceEmbeddings(model_name = 'jhgan/ko-sroberta-multitask'
                                           , model_kwargs = {'device':device}
                                           , encode_kwargs = {'normalize_embeddings' : True}))
```

## Splits the loaded Document objects into smaller text chunks

```
loader = DirectoryLoader('약품_full_data_별도의.txt', glob="*.txt", loader_cls=TextLoader)
documents = loader.load()
```

```
text_splitter = RecursiveCharacterTextSplitter(chunk_size=1000, chunk_overlap=200)
texts = text_splitter.split_documents(documents)
```

## Creates a Chroma vector database object

```
chroma = Chroma.from_documents(texts, model_huggingface, persist_directory="./chroma_")
chroma.persist()
```

```
retriever = chroma.as_retriever()
```

## Creates a retriever object from the Chroma vector database, which can be used to retrieve similar documents

```
chatopenai = ChatOpenAI(model_name = 'gpt-3.5-turbo'
                        , api_key = 'sk-6Lg1W'
                        , temperature = 0 )
```

## provide answers to questions based on the retrieved relevant documents

```
qa_chatopenai = RetrievalQA.from_chain_type(llm=chatopenai
                                           , retriever = retriever
                                           , retriever = retriever)
result_chatopenai = qa_chatopenai(q)
```

**GPT-3.5**

**GPT-3.5**

+

**RAG**

**RAG**

# RAG inference

## # 소화불량

“소화불량 증상에 복용할 수 있는 약을 추천해줘. 그리고 그 약의 효능을 알려줄래?”

```
chatopenai = ChatOpenAI(model_name = 'gpt-3.5-turbo',
                        , api_key = 'sk-6LgIWY[REDACTED]',
                        , temperature = 0 )

qa_chatopenai = RetrievalQA.from_chain_type(llm=chatopenai
                                           ##### 여기에 disk 에서 읽은 chroma 넣기
                                           , retriever = retriever)

result_chatopenai = qa_chatopenai(q)
```

### < RAG output>

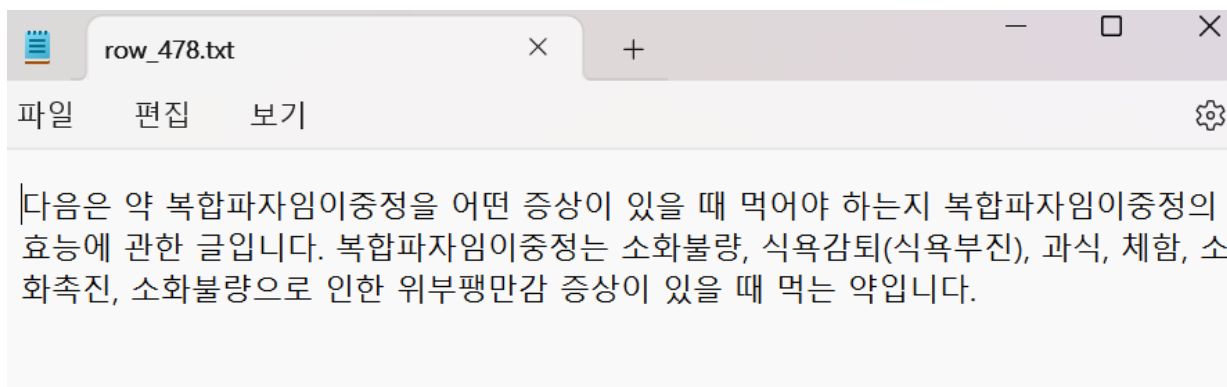


**prediction :** 복합파자임이중정은 소화불량 증상에 복용할 수 있는 약으로, 소화불량, 식욕감퇴, 과식, 체함, 소화촉진, 소화불량으로 인한 위부팽만감 등의 증상을 개선하는데 도움을 줍니다. 이 약은 성인 및 15세 이상은 1일 3회 식후에 복용하며, 만 7세 이하의 어린이나 특정 유전적 문제가 있는 환자는 복용하지 않아야 합니다.

### context

**documents:** context docs: ['row\_1605.txt', 'row\_32.txt', 'row\_2326.txt', 'row\_1843.txt']

### < train document (ground\_truth) >



### RAG model provides ...

1. **predicted answer**

2. **context documents** it referred to when generating response  
-> it can be used for “evaluation”

# RAG inference

“가래 증상에 복용할 수 있는 약을 추천해줘. 그리고 그 약의 효능을 알려줄래?”

# 가래

```
chatopenai = ChatOpenAI(model_name = 'gpt-3.5-turbo',
                        , api_key = 'sk-6LgIWY[REDACTED]',
                        , temperature = 0 )

qa_chatopenai = RetrievalQA.from_chain_type(llm=chatopenai
                                           ##### 여기에 disk 에서 읽은 chroma 넣기
                                           , retriever = retriever)

result_chatopenai = qa_chatopenai(q)
```

## < RAG output>

### prediction :

가래 증상에 복용할 수 있는 약으로는 "아스노캡셀"과 "두캅스에이연질캡슐"이 있습니다.

- 아스노캡셀의 효능: 기침과 가래 증상을 완화시키는 약입니다.
- 두캅스에이연질캡슐의 효능: 기침과 가래 증상을 완화시키는 약입니다.

위 두 가지 약 중 하나를 선택하여 가래 증상을 완화시킬 수 있습니다.

### context documents:

context docs: ['row\_2502.txt', 'row\_2128.txt', 'row\_3725.txt', 'row\_2894.txt']

## < train document (ground\_truth) >

row\_1023.txt

파일 편집 보기

다음은 약 아스노캡셀을 어떤 증상이 있을 때 먹어야 하는지 아스노캡셀의 효능에 관한 글입니다. 아스노캡셀은 기침, 가래 증상이 있을 때 먹는 약입니다.

row\_1311.txt

파일 편집 보기

다음은 약 두캅스에이연질캡슐을 어떤 증상이 있을 때 먹어야 하는지 두캅스에이연질캡슐의 효능에 관한 글입니다. 두캅스에이연질캡슐은 기침, 가래 증상이 있을 때 먹는 약입니다.

# Performance Evaluation

## Evaluation Metrics

Accuracy, precision, recall, BLEU, and ROUGE

Commonly used to assess the performance of natural language processing models.

## RAGAS

RAGAS(Retrieval Augmented Generation Assessment) is a framework designed to evaluate Retrieval Augmented Generation (RAG) pipelines. It measures the quality of generated text by comparing it to reference outputs, accounting for semantic similarity.

## Relevance Score

The relevance score measures the relevance of predicted document indices by comparing them to true relevant indices.

# Evaluation - RAGAS

**ragas score**

generation

**faithfulness**

how factually accurate is  
the generated answer

**answer relevancy**

how relevant is the generated  
answer to the question

retrieval

**context precision**

the signal to noise ratio of retrieved  
context

**context recall**

can it retrieve all the relevant information  
required to answer the question

**Question**

**Ground\_truth**

**Contexts**

**Answer**

**Faithfulness**

**Answer\_correctness**

# Evaluation - RAGAS

Evaluating: 100%

10/10 [00:28<00:00, 3.22s/it]

Filter



index	question	ground_truth	contexts	answer	faithfulness	answer_correctness
2	소화불량 증상에 복용할 수 있는 약을 추천해줘	소화불량 증상에 복용하는 약으로는 '복합파자임이중정'이 있습니다. 이 약은 소화불량, 식욕감퇴(식욕부진), 과식, 체함, 소화촉진, 소화불량으로 인한 위부팽만감에 사용합니다. 하지만 정확한 처방을 위해서는 의사나 약사와 상담하여야 합니다.	<p>[‘다음은 약 복합파자임이중정을 어떤 증상이 있을 때 먹어야 하는지 복합파자임이중정의 효능과 복합파자임이중정의 복용법과 복합파자임이중정을 복용할 때 지켜야 할 복합파자임이중정의 주의사항에 관한 글입니다. 복합파자임이중정은 소화불량, 식욕감퇴(식욕부진), 과식, 체함, 소화촉진, 소화불량으로 인한 위부팽만감 증상이 있을 때 먹는 약입니다. 복합파자임이중정의 복용법은 성인 및 15세 이상은 1회 1~2정 1일 3회 식후에 복용합니다. 복합파자임이중정의 주의사항은 만 7세 이하의 어린이,\xa0갈락토오스 불내성, Lapp 유당분해효소 결핍증 또는 포도당-갈락토오스 흡수장애 등의 유전적인 문제가 있는 환자는 이 약을 복용하지 마십시오. 복합파자임이중정의 주의사항은 이 약을 복용하기 전에 알레르기 체질인 환자,\xa0임부 또는 임신하고 있을 가능성이 있는 여성은 의사 또는 약사와 상의하십시오. 복합파자임이중정의 주의사항은 정해진 용법과 용량을 잘 지키십시오.2주 정도 복용하여도 증상의 개선이 없을 경우\xa0복용을 즉각 중지하고 의사 또는 약사와 상의하십시오.’</p> <p>‘다음은 약 뉴이라센액을 어떤 증상이 있을 때 먹어야 하는지 뉴이라센액의 효능과 뉴이라센액의 복용법과 뉴이라센액을 복용할 때 지켜야 할 뉴이라센액의 주의사항에 관한 글입니다. 뉴이라센액은 자양강장, 허약체질, 육체피로, 병후의 체력저하, 식욕부진, 영양장애, 발열성·소모성 질환 증상이 있을 때 먹는 약입니다. 뉴이라센액의 복용법은 성인은 1일 1회, 1회 1병을 복용하십시오. 뉴이라센액의 주의사항은 이 약에 과민증, 레보도파 성분의 약을 복용하는 환자는 이 약을 복용하지 마십시오. 뉴이라센액의 주의사항은 이 약을 복용하기 전에 의사의 치료를 받고 있는 환자, 임부, 수유부는 의사 또는 약사와 상의하십시오. 뉴이라센액의 주의사항은 정해진 용법과 용량을 잘 지키십시오. 뉴이라센액의 주의사항은 일정기간 복용하여도 증상의 개선이 없거나 악화될 경우에는 의사 또는 약사와 상의하십시오. 뉴이라센액의 주의사항은 경구용으로만 사용하고 주사용으로는 사용하지 마십시오. 뉴이라센액의 주의사항은 요를 황색으로 변하게 하여 임상검사치에 영향을 줄 수 있습니다.’</p> <p>‘다음은 약 위제로정을 어떤 증상이 있을 때 먹어야 하는지 위제로정의 효능과 위제로정의 복용법과 위제로정을 복용할 때 지켜야 할 위제로정의 주의사항에 관한 글입니다. 위제로정은 소화불량, 식욕감퇴(식욕부진), 과식, 소화촉진, 소화불량으로 인한 위부팽만감, 위산과다, 속쓰림, 위부불쾌감, 체함, 구역, 구토, 위통, 신트림 증상이 있을 때 먹는 약입니다. 위제로정의 복용법은 만 15세 이상 소아 및 성인은 1회 3정 1일 3회 식후 또는 식간에 복용합니다. 위제로정의 주의사항은 투석요법을 받고 있는 환자, 만 15세 미만 소아, 갈락토오스 불내성, Lapp 유당분해효소결핍증 또는 포도당-갈락토오스 흡수장애 등의 유전적인 문제가 있는 환자는 이 약을 복용하지 마십시오. 위제로정의 주의사항은 이 약을 복용하기 전에 신장(콩팥)장애, 알레르기 체질, 임부 또는 임신하고 있을 가능성이 있는 여성, 황색4호 과민증 환자 또는 경험자, 나트륨 제한 식이를 하는 사람은 의사 또는 약사와 상의하십시오. 위제로정의 주의사항은 정해진 용법과 용량을 잘 지키십시오. 위제로정의 주의사항은 만 15세 이상 소아에게 투여할 경우 보호자의 지도 감독하에 투여하십시오. 위제로정의 주의사항은 2주 정도 복용하여도 증상의 개선이 없을 경우 복용을 즉각 중지하고 의사 또는 약사와 상의하십시오.’</p> <p>‘다음은 약 다이락스센정을 어떤 증상이 있을 때 먹어야 하는지 다이락스센정의 효능과 다이락스센정의 복용법과 다이락스센정을 복용할 때 지켜야 할 다이락스센정의 주의사항에 관한 글입니다. 다이락스센정은 변비와 변비에 따른 식욕부진(식욕감퇴), 복부(배부분)팽만, 장내이상발효, 치질의 완화 증상이 있을 때 먹는 약입니다. 다이락스센정의 복용법은 만 15세 이상은 1회 1~2정, 만 11세 이상~만 15세 미만 1회 2/3정~1과 1/3정, 만 7세 이상~만 11세 미만 1회 1/2정~1정을 1일 1회, 취침 시(공복 시) 복용합니다. 다이락스센정의 복용법은 다만, 초회(처음)는 최소량을 복용하고 변의 모양과 상태를 보면서 조금씩 증량(양을 늘림) 또는 감량(줄임)하십시오. 다이락스센정의 주의사항은 급성 복부(배부분)질환(충수(맹장고리)염, 장출혈, 궤양성 결장염 등) 환자, 장폐색(창자막힘) 환자, 만 7세 이하의 어린이, 갈락토스 불내성, Lapp 유당(젖당)분해효소 결핍증 또는 포도당-갈락토오스 흡수장애 등의 유전적인 문제가 있는 환자는 이 약을 복용하지 마십시오. 다이락스센정의 주의사항은 이 약을 복용하기 전에 심한 복통(배 아픔) 또는 구역, 구토 환자, 임부 또는 임신하고 있을 가능성이 있는 여성, 다른 약물을 복용하고 있는 사람, 직장의 출혈 혹은 장부전의 병력이 있는 사람은 의사 또는 약사와 상의하십시오. 다이락스센정의 주의사항은 정해진 용법과 용량을 잘 지키십시오. 다이락스센정의 주의사항은 어린이에게 사용할 경우 보호자의 지도 감독하에 투여하십시오. 다이락스센정의 주의사항은 반드시 필요한 경우 이외에는 만 10세 이하의 어린이는 복용을 피하십시오. 다이락스센정의 주의사항은 복용 시 정제를 씹지 마십시오.]</p>	소화불량 증상에 대해 복용할 수 있는 약으로는 "복합파자임이중정"을 추천해드릴 수 있습니다. 해당 약은 소화불량, 식욕감퇴, 과식, 체함, 소화촉진, 소화불량으로 인한 위부팽만감 등의 증상이 있을 때 복용할 수 있는 약으로 알려져 있습니다. 하지만 의사 또는 약사와 상의한 후에 안전하게 복용하시는 것이 좋습니다.	0.3333333333333333	0.8088523907906952



# RAG Evaluation

## # test samples about 10 symptoms !

symptoms

['소화불량', '가래', '습진', '설사', '빈혈', '속쓰림', '비염', '두통', '결막염', '가려움']

	user input
소화불량	소화가 잘 안 되는 것 같아 ㅠㅠ 이때 복용할 수 있는 약을 추천해줘.
가래	가래가 자주 나오네.. 가래 많이 나올 때는 무슨 약 먹어?
습진	또 습진이 났어!! 습진 났을 때는 무슨 약 먹어야 하는지 말해줘.
설사	설사함....ㅠㅠ 이때 무슨 약 먹어야해?
빈혈	빈혈이 심하다. 빈혈 심할 때 약 뭐 먹어야해?
속쓰림	지금 속쓰림이 너무 심한데 무슨 약 먹어야해?
비염	환절기에 비염 때문에 너무 고생해.. 나한테 약 추천해줘.
두통	나 갑자기 머리가 너무 아파.. 약 뭐 먹으면 좋아?
결막염	나 결막염 걸린 것 같은데 무슨 약 먹어야하지?
가려움	피부가 너무 가려워.. 이땐 무슨 약 복용하는 게 좋아?

# 소화불량  
# 가래  
# 습진  
# 설사  
# 빈혈  
# 속쓰림  
# 비염  
# 두통  
# 결막염  
# 가려움



# RAG Evaluation

## (1) Document Precision

Evaluates how accurate the documents referred to by the model are in generating the answer. It measures the quality of the retrieved relevant documents.

```
# 전체 데이터셋에서 일치하는 데이터 찾기
def find_matching_rows(symptom, data):
    keyword_index = []
    keyword_index += data[data['효능'].str.contains(symptom)].index.tolist()
    for i in data[data['제품명'].str.contains(symptom)].index.tolist():
        if i not in keyword_index:
            keyword_index.append(i)
    return keyword_index

def extract_row_index(doc_filenames):
    return [int(fname.split('_')[1].split('.')[0]) for fname in doc_filenames]

true_indexes_list = []
symptom_list = []

for symptom in symptoms:
    # 전체 데이터셋에서 일치하는 데이터 찾기
    true_indexes = find_matching_rows(symptom, df)
    true_indexes_list.append(true_indexes)
```

```
def evaluate_relevance(true_indexes_list, predicted_indexes_list):
    relevance_scores = []
    for true_indexes, predicted_indexes in zip(true_indexes_list, predicted_indexes_list):
        match_count = sum([1 for idx in predicted_indexes if idx in true_indexes])
        relevance_scores.append(match_count / len(predicted_indexes) if predicted_indexes else 0)
    return sum(relevance_scores) / len(relevance_scores)

relevance_score = evaluate_relevance(true_indexes_list, predicted_doc_indices)
print('참고한 정확한 문서 수 / 참고한 문서 수')
print("Document Precision:", relevance_score)
```

참고한 정확한 문서 수 / 참고한 문서 수  
Document Precision: 0.825



**True Document Indices:** [8, 90, 256, 277, 502, 503, 512, 513, 579, 582, 583, 586, 590, 596, 607, 614, 616, 619, 625, 631, 819, 888, 894, 895, 896, 903, 904, 920, 925, 927, 938, 941, 0, 1308, 1353, 1441, 1469, 1522, 1530, 1566, 1605, 1709, 17489, 1998, 2000, 2009, 2102, 2248, 2326, 2352, 2368, 2370, 23764, 2781, 2798, 2850, 2865, 2868, 2889, 2890, 2913, 2946, 23495, 3514, 3524, 3545, 3564, 3565, 3568, 3679, 3710, 3720, 4543, 4659]  
**Predicted Document Indices:** [896, 681, 1605, 2798]

# Document Precision: 0.825

# RAG Evaluation

## (2) Answer Precision

```
답변에있는약품이정답 = []
답변에있는약품=[]
for i,j,t in zip(predicted_answers, predicted_doc_indices, true_indexes_list):
    aa = df['제품명']
    s = [] # 존재하는 약품 인덱스
    q = [] # 답변에있는 약품이 정답
    for k in aa.values :
        if k in i:
            idx = df[df['제품명'] == k].index[0]
            s.append(idx)
    답변에있는약품.extend(s)
    for f in s:
        if f in t:
            q.append(f)
    답변에있는약품이정답.extend(q)

print('답변에 있는 정확한 약품 수 / 답변에 있는 약품 수')
print('Answer Precision:', round(len(답변에있는약품이정답) / len(답변에있는약품), 3))
```

Evaluates the accuracy of the final answer generated by the model itself, regardless of the reference documents. It measures the correctness of the model's answer, independent of the documents it referred to.

답변에 있는 정확한 약품 수 / 답변에 있는 약품 수  
Answer Precision: 0.862

# Answer Precision: 0.862

# Conclusion and Future Work

1

## Conclusion

The MedAdvisor has proven to be a effective and accessible solution for users. By leveraging advanced natural language processing, the system is able to provide accurate and prompt responses to a wide range of medication-related queries. This has helped to bridge the gap between technical medical information and the needs of everyday users, particularly in urgent situations.

2

## Future Work

Looking ahead, there are several key areas of development that will be explored to further enhance the capabilities of the chatbot. These include adding multi-language support to reach a broader audience, integrating voice recognition technology for hands-free accessibility, implementing data-driven improvements based on user feedback and usage patterns, and incorporating robust security measures to protect sensitive user information.

**Thank you**