

Next-Gen Answers: Application using RAG with LLMs for Medical Q&A on PubMed Data

Capstone-Project | Computational Data Science | Cohort:6 | Group:3

Annexure

| | |
|---------------------|---|
| <u>RAG</u> | Overview of Why RAG? |
| <u>Approach 1:</u> | Implementing the solution using Google Vertex AI |
| <u>Approach 2:</u> | Implementing the solution using open source workspace |
| <u>Comparison</u> | Approach 1 vs Approach 2 |
| <u>Future scope</u> | Further applications |
| <u>Appendix</u> | |

Contributors:



Why RAG? Enhancing LLM Accuracy with Contextual Awareness

Starting Point: Challenges with Open Source LLMs

- Problem: LLMs can "hallucinate," generating incorrect but plausible answers.
- Example: For a pharma company open source model might inaccurately /outdated information in explaining drug interactions based only on open source data

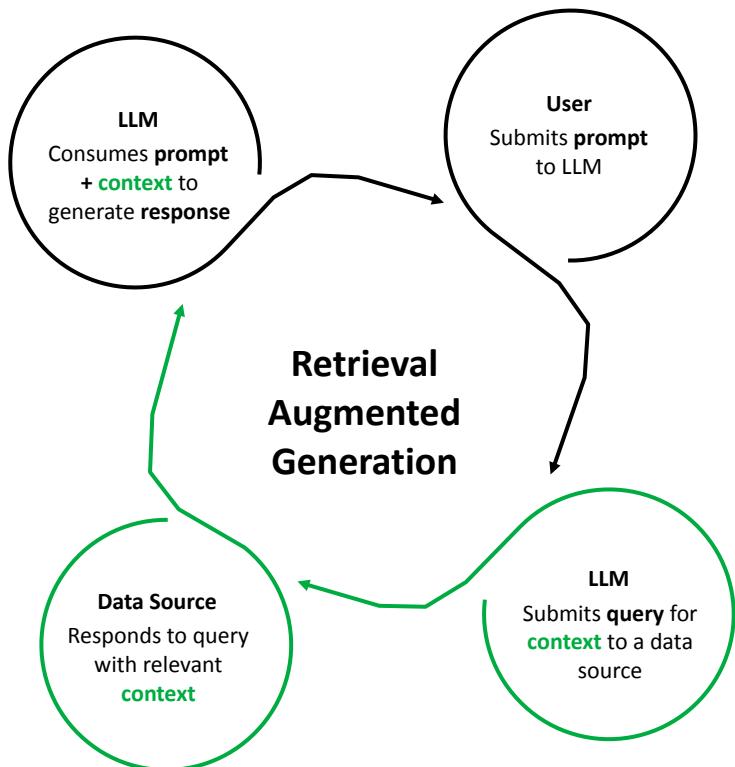
RAG as a Solution

- Approach: Combines LLMs with a retrieval system to use real-time, relevant documents from internal company databases
- Benefit: Reduces errors and enhances response relevance and accuracy.

Our Exploration: Results with PubMed Data

- Impact: Ensures answers are supported by the pubmed research journals increasing reliability.
- Use Case: Improves accuracy in responses about drug and alcohol interaction studies, referencing specific journals

HIGHLIGHT ON RAG



Benefits

-
- Enhanced Accuracy**
Incorporates a layer of fact-checking via relevant information from external databases or sources → responses are grounded in source
 - Tailored Responses**
Outputs leverage external documents to add relevant context to responses, particularly in complex or specialized domains
 - Efficiency and Scalability**
Information retrieved from databases as opposed to generated from scratch, with ability to update database to incorporate new information without retraining entire model

RAG Process Pipeline Overview



Data Source: PubMed/URLs
Topic: Alcohol and Drug Interactions
(PDF Format)

Data Cleaning: Removal of header and footer

Text splitting: Content is split into chunks using Recursive Text Splitter to keep related content intact

Data Storage: BigQuery Table
Centralized storage for PubMed metadata.
(Journal Title, content)

Data Cleaning: Removal of header and footer

Text splitting: Content is split into chunks using Recursive Text Splitter to keep related content intact

Data Storage: BigQuery Table
Centralized storage for PubMed metadata.
(Journal Title, content)

Text Embedding: Converting text into numerical form for analysis.

Options Explored:

- mpnet-base-v2*: Efficient embedding using MPNet architecture.
- gpt-4all*: Extensive language understanding for embedding.
- Gecko-text-embedding*: Google vertex AI advanced text embedding.

Vector storage:

- FAISS DB*: Efficient storage and retrieval of embeddings.
- Vertex Vector Store*: Scalable storage solution for embeddings.

Retrieval

Retrieval : The process of retrieving relevant splits from a large dataset based on a given query.

Options Explored:

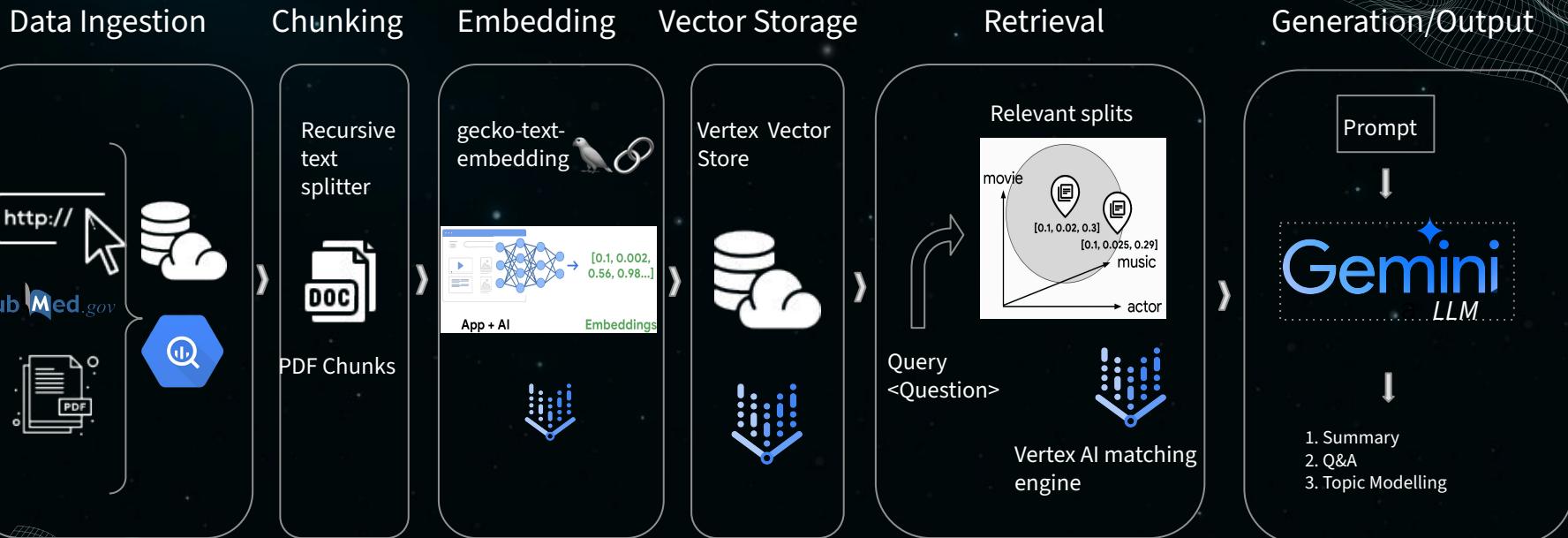
- FAISS*: Utilizing FAISS for efficient similarity search and retrieval.

Vertex AI Matching Engine: Leveraging Google's Vertex AI for advanced matching capabilities.

Generation Process:

- Summary:** Condensing information from retrieved documents.
- Q&A:** Extracting answers from documents based on questions.
- Topic Modeling:** Identifying themes in the retrieved data.
- LLM Options Explored:** Gemini, PaLM, Text-Bison, GPT-2
- Parameters:** #Tokens, temperature**

Approach 1: Utilising Google Vertex AI Architecture



Approach 2: Utilising Open Source Architecture

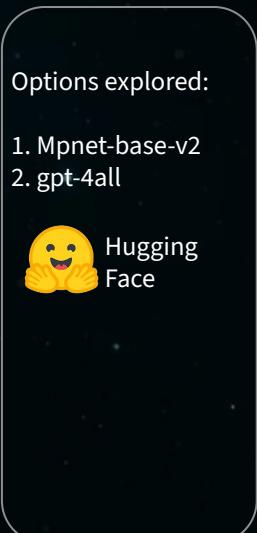
Data Ingestion



Chunking



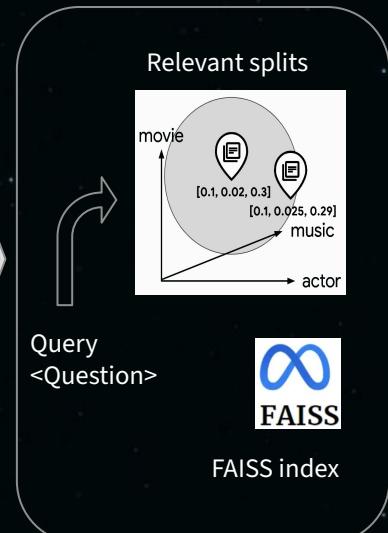
Embedding



Vector Storage



Retrieval



Generation/Output



Output: Query 'Prolong alcohol intake impact human health?' colab link: [Gemini](#)

```
ask("prolong alcohol intake impact human health?")
```

REFERENCE #0

Document Source: Title: Effects of Alcohol Consumption on Various Systems of the Human BodyA Systematic Review.pdf,page:1

Content:

content: Open Access Review Article DOI: 10.7759/cureus.30057 Effects of Alcohol Consumption on Various Systems of the Human Body: A Systematic Review Jerin Varghese 1 , Sarika Dakhode 2 Received 09/16/2022 Review began 09/20/2022 Review ended 09/30/2022 1. Medical School, Jawaharlal Nehru Medical College, Datta Meghe Institute of Medical Sciences, Wardha, IND 2. Published 10/08/2022 Community Medicine, Jawaharlal Nehru Medical College, Datta Meghe Institute of Medical Sciences, Wardha, IND Varghese et al. This is an open access Corresponding author: Jerin Varghese, jerinazero@gmail.com article distributed under the terms of the Creative Commons Attribution License CC- BY 4.0., which permits unrestricted use, distribution, and reproduction in any Abstract medium, provided the original author and source are credited. Prolonged alcohol intake for many years has been known to cause serious ailments in human beings since time memorial

REFERENCE #1

Document Source: Title: Impact of Alcohol Consumption on Young People.pdf,page:24

Content:

content: Does alcohol cause brain changes? For ethical reasons, it is not possible to conduct prospective research on the subject of alcohol use and brain effects in children and young people. However, the available observational evidence suggests that adolescents are likely to be more vulnerable than adults to both subtle brain damage and long lasting cognitive deficits following significant alcohol exposure [25]. Cognitive and learning impairments from chronic heavy exposure to alcohol can arise from the neuro-toxic effects of ethanol on vulnerable hippocampus and limbic structures.

Output: Query 'Prolong alcohol intake impact human health?' colab link: [Gemini](#)

Q/A Generation

```
print(model.generate_content(prompt, generation_config=generation_config).text)

## 10 Questions to Test Student Knowledge about Prolonged Alcohol Intake and Its Effects on Human Health:

**Multiple Choice:**
```

1. Which of the following is NOT a negative impact of prolonged alcohol intake on the brain?
 - a) Reduced grey and white matter volume
 - b) Improved cognitive function
 - c) Increased risk of dementia
 - d) Higher vulnerability in adolescents
2. Which of the following is the most common type of liver damage caused by prolonged alcohol intake?
 - a) Fatty liver disease
 - b) Alcoholic hepatitis
 - c) Cirrhosis
 - d) Liver cancer
3. Which of the following is NOT a cardiovascular complication associated with prolonged alcohol intake?
 - a) High blood pressure
 - b) Cardiomyopathy
 - c) Osteoporosis
 - d) Arrhythmia
4. Which of the following cancers is NOT associated with increased risk due to prolonged alcohol intake?
 - a) Mouth cancer
 - b) Lung cancer
 - c) Esophageal cancer
 - d) Breast cancer

Output: Query 'Prolong alcohol intake impact human health?' colab link: [Gemini](#)

Topic Modelling

```
[ ] generation_config = GenerationConfig(  
    temperature=0.2, max_output_tokens=1024, top_k=40, top_p=0.8  
)  
  
print(model.generate_content(modified_prompt, generation_config=generation_config).text)
```

5 Questions about Prolonged Alcohol Intake and Its Effects on Human Health:

1. **What are the specific mechanisms by which prolonged alcohol intake leads to brain damage and cognitive decline?**
2. **How does the severity of negative health impacts vary depending on the amount and duration of alcohol consumption?**
3. **Is there a safe limit for alcohol consumption that minimizes health risks while still offering potential benefits?**
4. **What are the most effective strategies for preventing and treating alcohol-related health problems?**
5. **How can we better educate individuals and communities about the risks and potential benefits of alcohol consumption?**

Summary Output

Embedding Type
MPTNET

Search Type
Similarity Search

Model Output Type
Summarize

Query
Is it safe for adults below 40 to drink moderately?

Search

Result

No amount of alcohol is "safe" or beneficial for your Alcohol use disorder (AUD) is a chronic brain disorder health. AUD can be mild, moderate, or severe one drink per day. Some people should not drink. Alcohol misuse at an early age increases the risk of all, including: developing AUD.

Embedding Type
MPTNET

Search Type
Similarity Search with Score

Model Output Type
Summarize

Query
How does alcohol consumption contribute to cancer development?

Search

Result

Alcohol consumption may give rise to malignant tumours via the toxic metabolite acetaldehyde. acetaldehyde is mutagenic and can cause cancer by damaging DNA and preventing it from being repaired. Drinking alcohol greatly increases the level of acetaldehyde found in saliva. All types of alcoholic drinks, including wine, beer and spirits, can increase the risk of cancer.

[view api](#) • built with gradio 

Embedding Type
GPT4

Search Type
Similarity Search

Model Output Type
Summarize

Query
prolong alcohol intake impact human health?

Search

Result

As of 2021, 29.5 million people health. Alcohol is a part of cultural traditions all ages 12 and older had an alcohol around the world...and It's also a drug that chemically use disorder in the past year. People of all ages need to understand these effects.

Embedding Type
MPTNET

Search Type
Similarity Search

Model Output Type
Summarize

Query
what is the relationship between alcohol and type2 diabetes?

Search

Result

Alcohol consumption is considered a significant risk the metabolic syndrome and type II diabetes mellitus factor for the development of haemorrhagic stroke (Mohan et al, 2009) Alcohol supplies only calories (Zodpey and Tiwari, 2005) and has been considered which compromises nutrition.

Summary Output

Embedding Type
MPTNET

Search Type
Similarity Search

Model Output Type
Summarize

Query
What is the World Health Organization's stance or guidance on the relationship between alcohol consumption and cardiovascular disease?

Search

Result
This is an official publication of the World Heart Federation. The WHO has called for a 10% and several other non-communicable diseases, relative reduction in the harmful use of alcohol between 2013-2025. Lack of universal advice and stringent policy deaths due to infectious diseases, intentional measures have contributed towards increased uptake and and unintentional injuries.

Embedding Type
GPT4

Search Type
Similarity Search with Score

Model Output Type
Generate

Query
what is the relationship between alcohol and type2 diabetes?

Search

Result
content: 7.3. Stroke/Metabolic: Studies in the last decade have shown that alcohol consumption is significantly associated with Alcohol intake is considered a significant risk the metabolic syndrome and type II diabetes mellitus factor for the development of haemorrhagic stroke (Mohan et al., 2009). Alcohol intake only calories (Zoddy and Twiss, 2005) and has been connected with increased triglycerides, reduced HDL cholesterol and increases blood as a modifiable risk factor for the occurrence sugar. Apart from this, it also contributes to the increase in the risk of stroke and heart attack. Alcohol can increase the risk of stroke by increasing the risk of hypertension and heart attack. Alcohol can increase the risk of stroke by increasing the risk of hypertension and heart attack. In 2004, acute and chronic pancreatitis is often the largest population based neuro-associate with and aggregates diabetes, but if it rises too steeply, the resulting overproduction of insulin can actually lead to low blood sugar, a condition called hypoglycemia. This is especially dangerous for diabetics, especially those taking certain drugs to lower their blood sugar. Alcohol also affects the endocrine system by interfering with how the body absorbs calcium, a chemical necessary for bone strength. Alcohol can also contribute to the development of osteoporosis, a condition where the bones become brittle and fragile. Chronic drinkers are more likely to develop osteoporosis than non-drinkers. Alcohol can also contribute to numerous, heavy or hazardous drinking on a single occasion slows your body's ability to fight off infections, even up to 24 hours after getting drunk. Over time, chronic drinkers are more liable to contract diseases like pneumonia and tuberculosis than those who do not drink above the recommended guidelines. Apart from congenital disorders of the cardiovascular system, it indeed is a very well-known fact, which could be understood from the history of most of the patients diagnosed with pancreatic cancer, that alcohol consumption is a major risk factor for the development of pancreatic cancer. Chronic alcohol intake impacts the repair ability of the structure of the exocrine pancreas, thereby leading to pancreatic dysfunctioning [14]. Most of the patients diagnosed with pancreatic have a strong history of chronic intake of alcohol. Liver diseases related to alcohol are known to humankind from the very beginning and probably one of the oldest known forms of injury to the liver [15]. In liver diseases linked to alcohol, liver cirrhosis is a major concern. Among alcohol-related diseases, liver cirrhosis is the most common and is associated with a high mortality rate. The risk of death from alcohol consumption or for Alcohol consumption is estimated to cause 1.1 million deaths worldwide (WHO 2008). Deaths risk of mortality and morbidity due to physical (et al., 2006). The impact of alcohol on disease and illness (Thun et al., 1997) is well-known. The studies of alcohol and health have been conducted in many countries around the world, but there is lack of power or validity. The studies of alcoholism and the cause of disease have been criticized as lacking power or validity. In conclusion, the findings of the first comprehensive epidemiological study of alcohol consumption and incidence of chronic and chronic diseases, in the world of the Americas and Africa, in the last decade, are convincing. The findings of these epidemiological studies are well documented and have been published in the journal of the American Journal of Public Health (AJPH) and the International Journal of Alcohol and Alcoholism (IJAA). An international team of researchers (A.A. Sajad and J.S.P. Shikha, 2013) conducted a comprehensive analysis to determine the risk of chronic and chronic diseases among alcoholics in India and Central Asia. They found that the use of alcohol increased by 21% (OR, 2.6 and increased by 10%, IOR)

Embedding Type
GPT4

Search Type
Similarity Search with Score

Model Output Type
Generate

Query
prolong alcohol intake impact human health?

Search

Result
content: 36 Volume 18, Issue 3 NHM MedicinePlus Magazine - STIDERC Alcohol's health effects: What you need to know risking alcohol is so common that people may not question how even one beer, cocktail, or glass of wine could impact their As of 2021, 28.5% of the global population of all ages (15 and older) had an alcohol use disorder worldwide, and it's also a drug that chemically changes the brain in the past year alters the body. People of all ages and genders are affected. SOURCE: GLOBAL SURVEY ON DISEASE AND HEALTH IN 2019. The International Agency for Research on Cancer has resources to help water. Because women tend to have less water in their bodies than men, if a woman and a man of the same size looking to change their drinking habits. Expert evidence provided for the paper showed that the only group with potential to have an overall significant reduction in risk of death in the UK is women over the age of 55 (especially if drinking around five units a week or less), but that heart disease was a significant risk factor for heavy drinking in the short-term and regular drinking in the long-term.5 A meta-analysis of over 90 studies that link drinking and cardiovascular disease for over 260,000 people found that reducing the amount of alcohol beverages consumed, even for light-to-moderate drinkers, improves cardiovascular health, including a reduction in heart disease risk. This has been followed by a number of other studies, including the Global Burden of Disease, Injury and Risk Studies, which has shown that previous research have estimated the possible health benefits of alcohol and under-estimated its health risks: content: nutrients Article The Global Impact of Alcohol Consumption on Premature Mortality and Health in 2016. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5313911/>. [Anfrankin1, BéthanyChen1top12,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,4,33,34,35,36,37,38,39,40,41,42,43,44,45,46,49,50,52,54,55,56,57,58,59,60,62,64,67,69,70,71/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5313911/) content: THE IMPACT OF ALCOHOL CONSUMPTION ON CARDIOVASCULAR HEALTH MYTHS AND MEASURES Over the past several decades, the most for cardiovascular diseases, including hypertension prevalence of cardiovascular disease (CVD) heart disease, cardiomyopathy, atrial fibrillation and flutter, and stroke. The evidence for the link between alcohol and cardiovascular disease is clear, but the mechanism of action is not fully understood. Alcohol has been distributed in [Topic]. Lack of universal advice and stringent policy deaths due to infection, disease, intentional measures have contributed towards increased uptake and and unintentional injuries, digestive diseases, diabetic diseases, especially availability of alcohol. The WHO has called for a 10% and several other non-communicable diseases, relative reduction in the harmful use of alcohol between 2013-2025. There is still no effective policy for the prevention and control of chronic pain in children and adolescents.7,8,9 Health costs, the cost of treatment and the burden of hospitalizations increase greatly the more alcohol is consumed.10,11,12 The most common cause of death in a child is not alcohol but alcohol-related death. The main cause of death is unintentional. The problem with the myth that alcohol increases death is that, as it has become more widely accepted, it has become more easily recognized as a factor that contributes to the increasing incidence of cardiovascular disease.13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,4,33,34,35,36,37,38,39,40,41,42,43,44,45,46,49,50,52,54,55,56,57,58,59,60,62,64,67,69,70,71/ content: THE IMPACT OF ALCOHOL CONSUMPTION ON

Embedding Type
MPTNET

Search Type
Similarity Search with Score

Model Output Type
Summarize

Query
prolong alcohol intake impact human health?

Search

Result
5.3% of all deaths occurred worldwide in the year 2016 were because of harmful alcohol use. It drastically increases the severity of diseases and also makes the treatments less effective. Alcohol not only affects the person physiologically, but it has many adverse effects psychologically and socially too.

Q&A Output

Embedding Type
MPTNET

Search Type
Similarity Search with Score

Model Output Type
Q&A

Query
How does alcohol consumption contribute to cancer development?

Search

Result
Chronic intake of alcohol may promote the genesis of cancer

Embedding Type
MPTNET

Search Type
Similarity Search with Score

Model Output Type
Q&A

Query
How does alcohol consumption contribute to cancer development?

Search

Result
Chronic intake of alcohol may promote the genesis of cancer

Embedding Type
GPT4

Search Type
Similarity Search

Model Output Type
Q&A

Query
what is the relationship between alcohol and type2 diabetes?

Search

Result
Alcohol use has also been liver which in turn predisposes to type ii diabetes

Top retrieved splits: Gecko (Vertex-AI) vs GPT-4ALL

Gecko-text-embedding

```
me.similarity_search("prolong alcohol intake impact human health?", k=5)
```

[Document(page_content='content: Open Access Review Article DOI: 10.7759/cureus.30057 Effects of Alcohol Consumption on Various Systems of the Human Body: A Systematic Review Jerin Varghese 1 , Sarika Dakhode 2 Received 09/16/2022 Review began 09/20/2022 Review ended 09/30/2022 1. Medical School, Jawaharlal Nehru Medical College, Datta Meghe Institute of Medical Sciences, Wardha, IND 2. Published 10/08/2022 Community Medicine, Jawaharlal Nehru Medical College, Datta Meghe Institute of Medical Sciences, Wardha, IND Varghese et al. This is an open access Corresponding author: Jerin Varghese, jerinazero@gmail.com article distributed under the terms of the Creative Commons Attribution License CC- BY 4.0., which permits unrestricted use, distribution, and reproduction in any Abstract medium, provided the original author and source are credited. Prolonged alcohol intake for many years has been known to cause serious ailments in human beings since time memorial', metadata={'source': 'Title: Effects of Alcohol Consumption on Various Systems of the Human BodyA Systematic Review.pdf,page:1', 'chunk': 2630}),
Document(page_content='content: Does alcohol cause brain changes? For ethical reasons, it is not possible to conduct prospective research on the subject of alcohol use and brain effects in children and young people. However, the available observational evidence suggests that adolescents are likely to be more vulnerable than adults to both subtle brain damage and long lasting cognitive deficits following significant alcohol exposure [25]. Cognitive and learning impairments from chronic heavy exposure to alcohol can arise from the neuro-toxic effects of ethanol on vulnerable hippocampus and limbic structures. Research findings from animal models and observational human studies have shown that persistent alcohol abuse can result in a decrease in the overall volume of the hippocampus, the area responsible for long term learning and memory [25]', metadata={'source': 'Title: Impact of Alcohol Consumption on Young People.pdf,page:24', 'chunk': 1171}),
Document(page_content='content: Does alcohol cause brain changes? For ethical reasons, it is not possible to conduct prospective research on the subject of alcohol use and brain effects in children and young people. However, the available observational evidence suggests that adolescents are likely to be more vulnerable than adults to both subtle brain damage and long lasting cognitive deficits following significant alcohol exposure [25]. Cognitive and learning impairments from chronic heavy exposure to alcohol can arise from the neuro-toxic effects of ethanol on vulnerable hippocampus and limbic structures. Research findings from animal models and observational human studies have shown that persistent alcohol abuse can result in a decrease in the overall volume of the hippocampus, the area responsible for long term learning and memory [25]', metadata={'source': 'Title: Impact of Alcohol Consumption on Young People.pdf,page:24', 'chunk': 1171}),
Document(page_content='content: Does alcohol cause brain changes? For ethical reasons, it is not possible to conduct prospective research on the subject of alcohol use and brain effects in children and young people. However, the available observational evidence suggests that adolescents are likely to be more vulnerable than adults to both subtle brain damage and long lasting cognitive deficits following significant alcohol exposure [25]. Cognitive and learning impairments from chronic heavy exposure to alcohol can arise from the neuro-toxic effects of ethanol on vulnerable hippocampus and limbic structures. Research findings from animal models and observational human studies have shown that persistent alcohol abuse can result in a decrease in the overall volume of the hippocampus, the area responsible for long term learning and memory [25]', metadata=

Top retrieved splits: Gecko (Vertex-AI) vs GPT-4ALL

GPT-4ALL embedding

```
new_db.similarity_search_with_score(query)

[(Document(page_content='content: 36 Volume 18, Issue 3 NIH MedlinePlus Magazine :STIDERC Alcohol's health effects: What you need to know rinking alcohol is so common that people may not question how even one beer, cocktail, or glass of wine could impact their As of 2021, 29.5 million people health. Alcohol is a part of cultural traditions all ages 12 and older had an alcohol around the world..and it's also a drug that chemically use disorder in the past year. alters the body. People of all ages need to understand these effects. SOURCE: NATIONAL SURVEY ON DRUG USE AND HEALTH The National Institute on Alcohol Abuse and Alcoholism (NIAAA) has information on how alcohol The alcohol you consume resides mostly in the body's impacts your health. It also has resources to help water. Because women tend to have less water in their bodies than men, if a woman and a man of the same those looking to change their drinking habits', metadata={'source': 'Title: alcohol_health_overview.PDF_.final_.080123.pdf,page:1', 'chunk': 3399}, 0.57196635), (Document(page_content='. Expert evidence provided for the paper showed that the only group with potential to have an overall significant reduction in risk of death in the UK is women over the age of 55 (especially if drinking around five units a week or less), but that heart disease was a significant risk factor for heavy drinking in the short-term and regular drinking in the long-term.5 A meta-analysis of evidence from more than 50 studies that linked drinking habits and cardiovascular health for over 260,000 people found that 'reducing the amount of alcoholic beverages consumed, even for light-to-moderate drinkers, may improve cardiovascular health, including a reduced risk of coronary heart disease'.6 This has since been followed by an international review of nearly 4 million respondents from 87 studies, which has shown that previous research 'over-estimated the possible health benefits of alcohol and under- estimated its health risks', metadata={'source': 'Title: The-physical-and-mental-health-effects-of-alcohol.pdf,page:5', 'chunk': 3471}, 0.5762557), (Document(page_content='content: nutrients Article The Global Impact of Alcohol Consumption on Premature Mortality and Health in 2016 IvneetSohi1,*;AriFranklin1,BethanyChrystoja1,2,3,AshleyWettlaufer1,JürgenRehm1,2,3,4,5,6,7,8 andKevinShield1,2,4 1 CentreforAddictionandMentalHealth,InstituteforMentalHealthPolicyResearch,33UrsulaFranklin Street,Toronto,ONM5S2S1,Canada;Ari.Franklin@camh.ca(A.F.);bethany.chrystoja@ucalgary.ca(B.C.);Ashley.Wettlaufer@camh.ca(A.W.);jtrehm@gmail.com(J.R.);kevin.shield@camh.ca(K.S), metadata={'source': 'Title: The Impact of Alcoholic Beverages on Human Health .pdf,page:24', 'chunk': 4314}, 0.58662367), (Document(page_content='content: THE IMPACT OF ALCOHOL CONSUMPTION ON CARDIOVASCULAR HEALTH: MYTHS AND MEASURES Over the past several decades, the for most cardiovascular diseases, including hypertensive prevalence of cardiovascular disease (CVD) heart disease, cardiomyopathy, atrial fibrillation and flutter, and stroke. The widespread message for over 30 has nearly doubled, and alcohol has played years has been to promote the myth that alcohol prolongs a major role in the incidence of much life. chiefly by reducing the risk of coronary heart disease of it. Alcohol has
```

Limitation while executing Capstone Project:

- Llama3 taking more GPU resources than available, resulting in crashes due to insufficient memory allocation.
- Llama3 encounters conflicts with the GPU configurations in Google Colab, leading to instability and frequent crashes during execution.
- API associated with the open-source model may have usage limits tied to the provided API key, resulting in errors when these limits are exceeded.
- At the proof of concept stage, it was not feasible to allocate resources towards investing in an API due to the project being in a learning stage.

Future Scope:

Evaluating contextual model: Evaluating LLMs differs significantly from traditional machine learning algorithms.

While traditional models output discrete values like 0/1 for classification or numerical values for regression, LLMs generate variable-length text, images, or videos. For instance, minor changes in wording can alter the meaning of LLM-generated content drastically, unlike in traditional models. This distinction complicates evaluation processes, emphasizing the need for nuanced assessment methods

Methods: ROUGE, NIST, BLEU etc. can be used depending on the usage of the LLM.

ROUGE (Recall Oriented Understudy for Gisting Evaluation):

- Rouge compares computer-generated summaries with human references.
- Used for text summarization tasks.
- Scores range from 0 to 1, with higher scores indicating greater similarity.
- Rouge-1, Rouge-2, and Rouge-L assess similarity at different levels.

Performance Improvement:

Hyperparameter: Temperature, Token Size, similarity mechanism

LLM: Different LLM model based on use case Q&A

Embedding: chunking Type, Overlap, different embedding mechanism

Search Mechanism: Graph network, elastic search

Reference: [ROUGE](#)



Thanks!

Do you have any questions?

References:

1. <https://docs.smith.langchain.com/cookbook/hub-examples/retrieval-qa-chain>
2. https://github.com/Adi8885/gcp-generative-ai/tree/update_vector_search_tutorial/gemini/prompts
3. <https://medium.com/@vikastiwari708409/how-to-use-gpt4all-langs-with-langchain-to-work-with-pdf-files-f0f0becadcb6>
4. https://python.langchain.com/docs/use_cases/question_answering/quickstart/
5. [https://cloud.google.com/blog/products/ai-machine-learning/how-to-use-grounding-for-your-langs-with-text-embeddingg6.](https://cloud.google.com/blog/products/ai-machine-learning/how-to-use-grounding-for-your-langs-with-text-embeddingg6)

6. https://github.com/GoogleCloudPlatform/vertex-ai-samples/blob/main/notebooks/official/matching_engine/sdk_matching_engine_create_stack_overflow_embeddings_vertex.ipynb
7. <https://cloud.google.com/vertex-ai/generative-ai/docs/embeddings/get-text-embeddings>
8. <https://cloud.google.com/vertex-ai/docs/vector-search/overview>
9. https://colab.research.google.com/drive/1R3ddBhX6Fdqm39W_D9kmvamOBitkn5p0#scrollTo=kKDw5VXMkXb3
10. <https://cloud.google.com/blog/products/ai-machine-learning/generative-ai-applications-with-vertex-ai-palm-2-models-and-langchain>
11. https://github.com/FullStackRetrieval-com/RetrievalTutorials/blob/main/tutorials/LevelsOfTextSplitting/5_Levels_Of_Text_Splitting.ipynb