

# Drooid Assignment

## Harsh Soni

Email :- soni.18@iitj.ac.in

Phone No :- +91 9450623469

Task Chosen:- Create a rudimentary Question Answering system that can answer questions related to the Israel Hamas war.

Link to the deployed T5 Model Method:-

<https://huggingface.co/spaces/Harsh2001/Israel-Hamas-War-Knowledge>

## Thought Process

Firstly, we need to understand that the task of making a rudimentary Question Answering System involves finding the relevant text from the dataset provided. Since Prompting the model directly with question and no context would lead to irrelevant answer being generated. Also passing the whole data at once to model would lead to a context window that would be larger than the model allows therefore it would again lead to irrelevant result. Also the data contains information that are not related to the Israel Hamas War.

Now in order to get relevant text out from the whole data (or Document), I divided the whole data into smaller chunks of text with length of 1000 tokens and 100 overlapping tokens between continuous text. This process of dividing the text into smaller chunks helps me calculate the sentence embeddings for them. These embeddings are then used to find the top k most similar chunks to the query embedding. These similar chunks are then used as the context for the LLM.

This helps us get an accurate context for the question. Now, In order to increase the quality of context for the follow-up questions I maintain a chat history that stores the previous questions and their respective answer and these are then passed along with the top k most relevant documents as the context to the query/question. The reason for this has been explained below:-

consider the question "What has been the international response?", this question could cause a problem while searching for the relevant documents since the question is not necessarily be in the context of Israel Hamas war it could easily bring up some other topic from the chunks.

## Models Used

For the given task i have used two models namely T5-large (770 Million Parameters) and LLama2(7 Billion Parameters). I have implemented them as two separate methods to which can be found in my branch.

For generating sentence embeddings for the chunks i have used the sentence transformer (all-mpnet-base-v2).

## Results

### T5-Large Model

This model has also been deployed on the link given above with which you can interact. The interaction results though might take some time (1-2 minutes) due to the absence of a gpu on the deployment server.

Query:- Describe about the Israel Hamas War.

```
"Israel's war with hamas that started on october 7 has now lasted for over two months,the conflict has claimed severa  
l thousands lives on both sides, with scores injured and displaced.the current conflict was triggered by the surprise  
hamas attack on israel on october 7, in which 1,200 israelis were gunned down in their homes, at a music festival and  
other public places."
```

Query:- What happened at the Al-Shifa Hospital?

```
'Tainment of the hospital has been a "bloodbath" hundreds of injured patients are crowded into the emergency departme  
nt of al-shifa hospital, who says.'
```

Query:- What has been the international response?

```
'The united nations general assembly voted overwhelmingly in favor of a humanitarian ceasefire in gaza, a resolution  
that was supported by 153 countries, opposed by 10, and seen abstention from 23. notably, the us vetoed a similar res  
olution in the security council, a move that stirred controversy on the global stage.'
```

Query:- What was the help provided to the victims of the war from Indian Govnerment?

'Naya: "india has been supporting palestine and palestinian refugees through significant contributions to unrwa between 2002 and 2023," he said. "india has been supporting palestine and palestinian refugees through significant contributions to unrwa between 2002 and 2023," he said.'

## Llama Model

Query:- Describe about the Israel Hamas War.

'The Israel Hamas war refers to the ongoing conflict between Israel and Hamas, a Palestinian militant group that rules the Gaza Strip. The war began on October 7, 2023, when Hamas launched a surprise attack on Israel, killing over 1,200 people and seizing some 240 hostages. Israel responded with a massive aerial and ground campaign, vowing to topple Hamas' regime in Gaza. The conflict has resulted in thousands of casualties, including civilians, soldiers, and militants, and has displaced approximately 80% of Gaza's population of 2.3 million people.'

Query:- What happened at the Al-Shifa Hospital?

'At the Al-Shifa Hospital, located in the Gaza Strip, there was a brutal and devastating attack by the Israeli military on November 11, 2023. The hospital, which was already overwhelmed with thousands of injured patients and displaced persons, was bombed and shelled by Israeli aircraft, causing significant damage and loss of life. The attack resulted in the deaths of at least two premature babies, and dozens of patients died in the Intensive Care Unit (ICU) over the following days. Medical staff and patients were also subjected to interrogation and detention by Israeli soldiers, and the hospital's state was described as "terrible" and "minimally functional". The situation at the hospital remains dire, with many people still trapped inside and in need of food, water, and medical attention.'

Query:- What has been the international response?

'The international community has responded swiftly and decisively to the ongoing conflict between Israel and Hamas. On October 7, 2023, the day of the initial attack, the United Nations General Assembly voted overwhelmingly in favor of a resolution calling for an immediate ceasefire in Gaza. The US and Israel were among the ten countries that voted against this proposition. Additionally, the UN Security Council passed a resolution calling for a temporary truce in Gaza to allow increased aid and the immediate release of hostages taken by Hamas. The US and Russia abstained from the vote, and the US blocked a Russian amendment demanding an urgent and sustainable cessation of hostilities. The international community has continued to monitor the situation closely, with many nations expressing concern about the escalating violence and humanitarian crisis in Gaza.'

Query:- What was the help provided to the victims of the war from Indian Government?

'The Indian Government provided aid to the victims of the war in Gaza, including essential life-saving medicines, surgical items, tents, sleeping bags, tarpaulins, and water purification tablets, among other items.'

## Conclusion

Upon examining the performance of two distinct models, T5-Large and LLama2, noticeable differences emerged. The T5-Large model, despite having fewer parameters, exhibited commendable performance on the task at hand. However, the LLama2 model, which boasts a larger number of parameters, demonstrated superior performance. It not only generated accurate responses but did so with a

remarkable depth of meaning. The exceptional efficacy of these models, particularly LLama2, underscores the potential of advanced machine learning tools in tackling complex tasks. This experiment, therefore, concludes that larger models may provide more meaningful results in the scope of question-answering systems.

## **File Contents**

DataProcessing.ipynb → Contains the code for generation of Data Chunks and their respective embeddings

T5-Model.ipynb → Contains the T5 Model based implementation of the task

LLama Model.ipynb → Contains the LLama Model based implementation of the task

requirements.txt → Contains all the necessary libraries required to run the code

documents.pkl → Already created chunks of data

embeddings.npy → Embedding vectors of the chunk of data created.

Report.pdf → A report file containing all the explanation and results of the task.

ContactMe.txt → A text file containing my personal contact information.