

Problem:

Create a timeline summarization list, that shows the major events of the Israel Hamas war that had occurred during the time-period mentioned in the Json file

Steps followed Thought process

- Loading and preprocessing the json data.
- Taking the input from the user like Israel-Hamas War and Generating the keywords
- Extract the date and the news article description
- Summarize the description using the transformer model (here, used the facebook/bart-large-cnn for text summary generation)
- generate the vector embeddings for each article (i.e. description) and then form clusters of the similar articles (i.e. description) by means of some clustering algorithm (here, used DBSCAN)
- Using the keywords filter the articles based on the description and then get a central vector which represent all the filtered articles via keywords, now based on the central vector find a cluster that it belongs to based on all the data and corresponding to the vector get all the articles
- get the date corresponding to these articles and that will your results
- further summarize the articles by forming the cluster and using title or description to generate new title or description for a enhanced summary like mentioned in the problem document
"An example of a time timeline summarization list is as follows (excerpt from Wikipedia):".

Steps Followed in the code

Step 1: Loading and Preprocessing the JSON Data

Simple Explanation: Preprocess and clean news data to ensure it's accurate, relevant, and easy to analyze. This process removes errors, duplicates, and irrelevant information, making the data consistent and suitable for generating meaningful insights and summaries.

Step 2: Taking Input and Generating Keywords

Simple Explanation:

To ask the user which event he/ she would like view like "Israel Hamas War" and generate the keywords so that we can retrieve the more relevant information related to the given prompt or user input.

Example: You tell me you want to know about the "Israel-Hamas War". I think of important words like "Israel", "Hamas", "war", "conflict", etc.

Step 3: Extracting Dates and Descriptions

Simple Explanation:

We look at the stories we found and write down when they happened and a short description of what happened.

For Example: We find a story about a big event on October 7, 2023. We write down "October 7, 2023: A big event happened."

Step 4: Summarizing Descriptions

Simple Explanation:

The stories might be very long, so we use a special tool to make them shorter but still keep the important parts.

For Example: We have a long story about a speech. The tool shortens it to "A leader gave an important speech."

Step 5: Generating Vector Embeddings and Clustering

Simple Explanation:

We turn the short descriptions into special codes that a computer can understand and group similar stories together.

For Example: We make a code for each short story. Stories about speeches get one code, and stories about fights get another. Then, we put stories with similar codes together.

Step 6: Filtering Articles Using Keywords

Simple Explanation:

We use our important words to find the stories that match what we're looking for.

For Example: We look at our grouped stories and see which ones mention "Israel", "Hamas", or "war". We keep those stories.

Step 7: Finding Relevant Clusters

Simple Explanation:

We find the group of stories that best match the important words we picked earlier.

For Example: We have groups of stories. One group talks a lot about the Israel-Hamas war. We pick that group.

Step 8: Summarizing the Articles Again with timelines

Simple Explanation:

We make a final short summary of the most important stories in our chosen group with timelines.

For Example:

Final code output

22 December 2023 - 25 December 2023: 'Israel increases strikes in central Gaza, killing scores Khawaja denied permission to have peace symbol on bat : Reports Pope Francis says Jesus' message of peace is being drowned out by 'futile logic of war' Shipping giant Maersk prepares to resume operations in Red Sea. 50 elections globally amid wars cast doubt on 2024 economic outlook.

10 September 2014 - 10 September 2014: 'Israeli soldier killed in rocket attack from Lebanon. Soldier was killed by a rocket fired from Lebanon, Israel says. Israeli military says it is investigating the attack and has launched an investigation into the source of the rocket attack. Israel says the rocket came from Lebanon and was fired from inside the country.

18 December 2023 - 22 December 2023: World celebrates Christmas in shadow of war. UN Security Council adopts resolution on aid to Gaza watered down by US pressure. Syrians cancel Christmas festivities in solidarity with Gaza. Pope laments war in Holy Land on solemn Christmas Eve. Pope Presides Over Christmas Eve Mass: 'Our Hearts Are in Bethlehem'