DRAGON: a Dedicated RAG for October 7th conflict News

Framing of Israel-Gaza NEWS shared task Guideline

Note: It is worth noting we annotated all Main and IAA sheets completely. However, the auto-fill column in the Google sheet did not update for our annotation. We must inform you of this so that our work will be reviewed completely.

Define the Objective

Bias

The task objective is to establish a shared corpus to facilitate the study of potential biases in news associated with the Israel-Gaza war.

Propaganda

The task objective is to prepare a corpus to facilitate the study of propaganda and its variations in the news associated with the war mentioned above.

Describe the Task

Both tasks can be considered annotation tasks accompanied by the provisioning of a guideline.

Establish Categories

Bias

The corpus has seven different categories reflecting the biases against or in favor of the sides in the war mentioned above.

Propaganda

The corpus has four different categories reflecting the presence of propaganda in the news content of the war.

Detailed Category Guidelines and Include Examples

Bias

Label	Definition	Example

Unbiased	The label is associated with the news presenting information neutrally without favoring any particular side.	In the ongoing Israel-Palestine conflict, recent events have escalated tensions. Yesterday, Israeli forces conducted operations in response to rocket attacks from Gaza. Both sides have reported casualties. International leaders are calling for restraint and a return to peace talks.
Biased against Palestine	The label corresponds to news that blames the Palestinian side for breaking and interrupting peace talks and ignoring international laws preventing attacks against non-militant citizens.	Once again, Palestinian aggression has disrupted peace in the region. Palestinian extremists, ignoring efforts for peace, launched unprovoked attacks on innocent Israeli civilians. Israel's response, though portrayed as harsh by some, is a justified measure to protect its citizens.
Biased against Israel	The label corresponds to news that criticizes activities of the Israeli side for harsh attacks against non-militant citizens such as women and children and for committing war crimes, especially genocide.	In a typical display of excessive force, Israeli troops have yet again targeted Palestinian areas, causing numerous civilian casualties. This aggression, under the guise of self-defense, highlights the ongoing oppressive tactics Israel employs against Palestinians.
Biased against both Palestine and Israel	This label is related to news that questions the activities of both sides and expresses opinions that the loss of peace is caused by the irresponsibility of both sides involved in the war.	In the latest chapter of their endless and futile conflict, Israeli and Palestinian forces have once again engaged in senseless violence. Both sides continue to commit atrocities, showing a complete disregard for peace or human life.
Biased against others	The label is associated with news that contains complaints	In the shadow of the Israel-Palestine conflict, external

	against other countries, people, and organizations.	actors, particularly Iran, are exacerbating tensions. Iran's covert support for extremist groups shows its intent to destabilize the region, disregarding the catastrophic impact on both Israeli and Palestinian civilians.
Unclear	The label corresponds to the news indicating its stance non-clearly or exhibiting ambiguity in its presentation.	Recent developments in the Middle East have seen an increase in hostilities. The situation in the region is complex, with various factors contributing to the current state of affairs. The international community remains divided on the issue.
Not applicable	This label can be considered as the label of any topics that are not directly related to the war conflicts or the bias annotations.	In other news, the annual technology conference in Tel Aviv has unveiled groundbreaking advancements in cybersecurity. Industry leaders from around the globe gathered to showcase innovations that promise to shape the future of digital security.

Propaganda

Label	Definition	Example
Propaganda	Information, especially biased or misleading, used to promote a particular political cause or point of view.	In a display of unmatched heroism, our troops have once again safeguarded our nation from the brink of destruction, heroically neutralizing the threat from Gaza, which aims to undermine our very existence.

Not Propaganda	Information or communication that is factual, neutral, or unbiased.	Once again, Palestinian aggression has disrupted peace in the region. Palestinian extremists, ignoring efforts for peace, launched unprovoked attacks on innocent Israeli civilians. Israel's response, though portrayed as harsh by some, is a justified measure to protect its citizens.
Unclear	Information that is ambiguous, vague, or difficult to determine its intended purpose or bias.	The situation in Gaza remains tense, with reports of civilian distress and military movements. While some sources claim the military actions are defensive, others argue they are provocative, leaving the true nature of the situation open to interpretation.
Not Applicable	Information that does not fall under the category of propaganda or its opposite; it is unrelated to the promotion or criticism of a political cause or viewpoint.	A feature on Gaza's cultural scene highlights the resilience of its art community, showcasing how local artists use their craft to express hope and endurance amid challenging circumstances, without delving into the political context.

Outline the Process

We delve into the problem with a classification approach to automate and compare the annotation procedures. Having this kind of view of the task, we have adopted a process with the following steps;

1. We first prepare a subset dataset, sampled from “Main” data, to use it in a supervised configuration as train and test data splits.
2. The sample subset data is annotated manually by the team members.
3. Different models then use the annotated subset data for training, testing, and prompting, and afterward for annotation of “Main” and “IAA” data.

Having different algorithms experimented with, we annotated the data sheets as expressed below in the table:

sheet	IAA-1	IAA-2	IAA-3	IAA-4	Main
algorithm	BGE M3 embeddings + KNN	RAG	E5 Mistral 7B instruct embeddings + KNN	Multilingual E5 Large embeddings + KNN	RAG

Set Quality Standards

Since we model the task as a multiclass classification problem, we use standard metrics of classification problems including accuracy, precision, recall, and F1 score.

Training and Support

Having prepared train and test data, we used a combination of the “embedding model” and “classifier model” to establish our classification approach. For the classifier model, we experimented with our embeddings with “KNN”, “SVC”, and “XGBoost” and the results demonstrated that KKN outperforms other classifiers in our case.

Embedding model

To create embeddings from our data, we use four embedding models:

BGE M3 + KNN

BGE M3 embedding, distinguished for its versatility in Multi-Linguality, Multi-Functionality, and Multi-Granularity, is the first solution we employed to extract embeddings from the subset data.

Multilingual E5 large + KNN

E5 text embedding models, released in 2023 in three different sizes (small, base, and large), offer a balance between inference efficiency and embedding quality.

E5 Mistral 7b instruct + KNN

With the rise of large language models, their usage in different tasks, text embedding generation in our case, is broadly employed. E5-Mistral-7b-instruct is one of the efforts of researchers to use the power of large language models in the production of embedding.

Nomic Embed + KNN

nomic-embed-text-v1, is the first fully reproducible, open-source, open weights, open-data, 8192 context length English text embedding model.

RAG

The method used to classify the news in the IAA-2 sheet is based on Retrieval Augmented Generation (RAG). In this method, we employ a dynamic approach to prompting rather than using traditional ways of prompting. That is, we change the prompt for each news article that we want to classify. If we consider each row of the sheet as a query, we search for the most similar documents to this query from the

annotated subset data using a retriever. We then select some of these documents, along with their corresponding labels, based on a specific condition, and include them as examples in the prompt
.
To elaborate further, we have a subset of data annotated by our team members. For each news article in the IAA-2 sheet, we identify a number of the most similar documents and their corresponding labels from the annotated subset. After experimenting with different numbers of documents to retrieve, we discovered that retrieving 7 documents provides satisfactory results. These 7 documents have various labels. The condition for determining the number of documents to include as examples in the prompt is based on the count of unique labels among these 7 documents. To clarify, we include examples based on the number of unique labels found among the retrieved documents. We ensure that we include only one example per label. Additionally, besides including examples in the prompt, we also used definitions of labels based on the unique labels of the retrieved similar documents and finally we prompted ChatGPT to classify the news.

We have used the “gpt-3.5-turbo” as the text generator model.

Feedback Mechanism

Our feedback mechanism is based on an “iterative direct on-demand QA” between the annotators and other team members. We applied a majority voting approach as a calibration session to neutralize potential personal biases and utilize collaborative wisdom during the annotation process for any ambiguous or controversial data sample.