

Team members' information:

- 1- Fatma Mohammed Al-Farsi (Sultan Qaboos University) > A. ID: 2
- 2- Asmahan Juma Al-Mamari (Sultan Qaboos University) > A. ID: 1

(d) Our Bias guidelines

1.1 Define the Objective:

This shared task aims to shed light on the bias and double standards in media coverage regarding the Gaza-Israel 2023-2024 war. It seeks to uncover the layers of bias and propaganda found in news articles and whether or not factors such as language and culture play a role in the level of bias these news outlets contain. The main objective of this collaborative task is to build a shared corpus for the process of a comprehensive annotation, including creating guidelines tailored to overcome conflicting views regarding sensitive topics such as this current task. Furthermore, this task helps in fostering a collaborative community in which team members overcome challenges together through discussion and data analysis. In addition, one of the main aims of this task is to raise a generation of Natural Language Processing (NLP) researchers, with proper skills which can help them deal with raw data sources. It also equips participants with the skills needed to propose the guidelines and methods to analyze a complex set of data in addition to allowing the space for participants to experiment with different tools that can facilitate the process for them.

1.2 Describe the Task:

FIGNEWS 2024 Datathon Shared Task has provided two subtasks to participate in; bias subtask and propaganda subtask and the participants can choose either both or one of them. As a team, we have decided to focus on the 'Bias Subtask' as it is an ongoing concern on social media platforms nowadays, especially after what happened on the 7th of October. A document was given to the participants that consisted of 6 types of sheets. The first two sheets are the annotation sheets; the main and the IAA (inter annotations agreement). These two sheets consist of the dataset to be annotated and each includes 15 batches in which each batch has 5 different

languages (English, Arabic, French, Hebrew, and Hindi). English and Arabic translations were provided for all of the datasets to facilitate annotation for participants who are not familiar with other languages. The annotations aim to identify if the text is unbiased, biased against Palestine, biased against Israel, biased against both Palestine and Israel, biased against others, unclear, or not applicable and the labeling is based on guidelines which are set by team members. The main sheet is annotated by all members of the team while IAA is annotated individually where each member has her own sheet to annotate (e.g. IAA-1, IAA-2 etc.). Then there is the ‘status’ sheet where participants have to do nothing as everything is going to be updated automatically based on members’ annotations in the main and IAA sheets. In addition, an ‘annotation team’ sheet is used to be filled by each of the team members to provide some information regarding the team name, type of subtask, ID, source of annotation language, native language, gender, country of origin, and education level. As for the last two sheets, there are ‘our bias guidelines’ sheet and ‘our propaganda guidelines’ sheet. In our case, the focus is going to be only on the ‘our bias guidelines’ sheet which is considered to be one of the most important things in the shared task as it identifies the main purpose and the process of the subtask in detail.

1.3 Establish Categories:

1. Unbiased: A text is unbiased when the information presented within it is objective and does not favor a side or a certain perspective involved in the topic or conflict being discussed. Unbiased texts aim to provide fair representations of different points of view without including personal opinions.
2. Biased against Palestine: A text is considered biased against Palestine if it contains prejudice against Palestine or Palestinians. The information, language, or tone in the text shows a negative or unfavorable view of everything related to Palestine. It also includes bias that manifests itself in the selection and presentation of facts, the omission of relevant information, and the use of loaded language.
3. Biased against Israel: This term describes a text that includes bias or discrimination against Israel. The information, language, or tone in the text shows a negative or unfavorable view of everything related to Israel. It also includes bias that manifests itself

in the selection and presentation of facts, the omission of relevant information, and the use of loaded language.

4. Biased against both Palestine and Israel: This label refers to texts that show unfairness or prejudice towards both sides of the conflict. This type of bias can be seen through the way facts are presented, the use of inflammatory language, or the lack of balance in discussing the perspectives of both sides.
5. Biased against others: This phrase indicates the type of bias found in texts targeted towards entities other than Palestine or Israel. It refers to texts that exhibit prejudice against neighboring countries, international figures, or any other relevant bodies involved in the conflict. Again, the bias may manifest itself in the form of negative portrayals, unfair generalizations, or the exclusion of relevant information that can provide a balanced understanding of other entities.
6. Unclear: When a text contains elements that appear to be biased but the direction of the bias is not evident or easily described then it can be labeled as “unclear”. It can be a result of vague language use or lack of clear bias indicators.
7. Not Applicable: This term is used when the text reviewed is not related to the Israel-Palestine conflict or does not contain relevant bias that requires annotation. It indicates that the text does not exhibit any bias targeted toward any side or entity involved in the conflict.

1.4 Detailed Category Guidelines:

The application criteria that the SQUad team has used when annotating were divided into four stages:

1- Biased vs Unbiased:

At first, the team members have to decide whether a text is biased or unbiased and this is going to be based on critically analyzing the following:

- Author's Perspective and Background:

Where we have to investigate the author's expertise, potential biases, and background as we have to take into account whether they have any interests or ideologies that could influence their perspective when reporting about the conflict. It is important to note that there is no information provided about the author but the source which is based on the languages.

- The Use of Language and Tone:

Analyzing the use of terms (e.g. harsh or biased adjectives) that could raise specific emotions and beliefs toward things or persuasive techniques that might have a sense of bias. For example, sometimes they deny the Palestinians' Identity referring to them as “ Gaza People” and “residents” Instead of Palestinians. In addition, when reporting, sometimes they are show more emphatic toward Israel and the opposite.

- Facts vs Opinion:

The members have to determine whether the text is a fact or it is only an opinion as the biased text might rely heavily on opinions that are not based on or supported by evidence. As for facts, they must be unbiased because they can either be known logically based on evidence or something which is well-known.

For example, it is a fact that Israel is an occupier so when it comes to reporting it sounds biased (opinion) but it's not that's why it might be considered to be unbiased as it is a fact.

Also, the number of killed people might not be accurate when reporting which is not considered to be fact but opinion.

- Balance and Fairness:

Where we have to assess if a text presents a balanced reporting for both affected parties or not. Usually, unbiased texts consider views from both sides.

Sometimes they show more empathy for Israel. Moreover, when reporting they focus on one side rather than mentioning both affected parties.

- Source and citation:

Examining what sources are cited in each text. Biased texts might choose sources that suit their point of view while ignoring the other perspectives.

In the case of this kind of dataset, there is no mention of the specific resources but the languages and it appears that some of the languages are more biased than others.

Also when reporting they mention things like “ the Israel government said that...” without looking at the other side's source of information to have a complete picture.

2- Biased against ...

If any of the team members decided that a text is biased, then they have to choose whether it is biased against:

- Palestine:

- Ignoring the Palestinian identity:

- 1- Referring to them as Gaza people and other names without Palestinians.

- The use of some adjectives.

- The way they describe the reporter:

- 1- the Israeli government said that

- 2- The Hamas group said ... > they deny the fact that Palestine is a country!> It doesn't exist for them!

- Denying the fact that Hamas group are Palestinians:

- > Hamas group are Palestinians >fact > but they deny again!

- Israel:

- The use of some terms and adjectives.

- Hate

- Dislike because of what they have done:
 “I need the West to understand something. Bombs kill people.” Cenk Uygur, the founder of the American news organization The Young Turks, appeared on Piers Morgan Uncensored on Monday where he spoke about the ongoing Israel-Palestine war. During the interview, Uygur, who announced he would run for president earlier this month, criticized Israel over its ongoing bombing campaign in Gaza and the United States for supporting it.

- Palestine and Israel:
 - Someone affected by this war
 - Biased against the conflict in general




- Others:
 - > Russia
 - > Iran
 - >.. etc

3- Unclear:

Text that has some sort of ambiguity and is difficult to determine what a text is talking about or biased against whom if it is not motioned anything related to the intended side.




- “Seven hostages who returned, but who left parents, children, siblings and friends behind in Gaza, describe what they endured” (20, B01, IAA). Here it is not clear who the hostages are. Are they Palestiniens? Or Israelis?

- “She is being prosecuted for advocating terrorism.” (54, B01, IAA). How is she? And is she Israeli? Or Palestinien?

- “  Israel's rapid attacks on Gaza. Watch LIVE  India's most wanted terrorist Shahid Latif killed in Pakistan  Afghanistan's first batting against India #ZeeLive #IsraelUnderAttack #IsraelWar #Israel #Hamas” (88, B01, IAA). It has more than one idea. Talking about Israel then Pakistan, It cannot be biased as a whole text against a specific side.

4- Not Applicable:

Texts that are not related to the annotation purpose at all. They might have about one word or more. It seems that some texts are generated by AI so it might choose texts that have certain words that are used when reporting about the conflict but in fact, they don't have the same meaning. Consider the following examples:

- “ Look | Great enthusiasm in the Jordanian march to support the starving people in Souq Waqif #Come_Asia | #AsiaCup2023 #Qatar2023 | #AsianCup2023 #Asian_Cup_Cup_Channels #See_the_Cup” (124, B02, IAA). This text is from the Arabic language where there was the word “ Hamas” which means “enthusiasm” in English and this context. The reason might be because the data was collected by AI that considered the word “Hamas” in Arabic as the Hamas group but it is clear also in the English translation that it is not about Hamas group that is why it considered to be not applicable.
- “PUBG Mobile explodes a surprise with exciting details that excite players. More details in the first comment    ” (40, B0, IAA). This is also another example of the same thing but the word “exciting” which means “Hamas” in Arabic it is not related to the conflict at all.
- “She suffered terrible trauma.” (143, B02, IAA). Who is she? And why? Who is responsible?

1.5 Include Examples:

1- Biased examples and common mistakes:

O “ Arabs! Residents of #Gaza, wake up!! # Hamas leaders are in heaven, enjoying the luxuries of #life, while you are in #hell, sending you to #death. Khaled Meshal and others like him are exploiting you so that their families can live in luxury and comfort #Because_of_Hamas” (224, B03, IAA) > Some think that we have to consider Hamas as Other or Palestinians (others vs Palestine) because they are Palestinians fighting for the sake of Palestine.

2- Biased against Israel vs. unbiased:

- Some of the things that Israel has done and some refer to them as crimes because referring to them as crimes might be biased but it is a fact that everyone knows.

Other Mistakes: Sometimes we could be affected by our own perspectives and beliefs.

1.6 Outline the Process:

For this task, the data analysis process was divided into two sections: Manual data annotation and technical data annotation (using tools).

A- Manual data annotation process:

The team members began this task by reading the aims and components of the task to have a clear vision of how the work is supposed to be done. The first step was agreeing on the criteria we will follow to indicate whether the texts provided are "Unbiased," "Biased against Palestine," "Unbiased Biased against Israel," "Biased against both Palestine and Israel," "Biased against others," "Unclear," or "Not Applicable." The team agreed on doing the inter-annotator agreement (IAA) manually, where each member will annotate four batches from B1 to B4).

Since the batches included sources from five different languages which are Arabic, English, Hebrew, French, and Hindi, we relied on the Arabic translations of the texts as it is the first language of both team members. Once the agreements were done, each team member read the guidelines and criteria on their own to make sure that they had a clear understanding of every label they would be using to decide the type of bias exhibited in the texts if any. The annotators made sure to look out for any word or phrase that indicates direct bias as well as paying attention to the tone, source of origin, level of logic, and credibility each text contained to help in the process of annotating. Throughout the whole process of annotation, the members consulted each other and constantly reviewed the criteria to avoid any personal views getting in the way.

B- Technical data annotation process:

As for the the posts under the “Main”, they were labeled using zero-shot classification with the Facebook/BART-large-MNLI model. The selection of this model was based on its robust architecture and proven performance in natural language processing tasks. Candidate labels were defined to cover a spectrum of potential classifications within the dataset, including "Unbiased," "Biased against Palestine," "Unbiased Biased against Israel," "Biased against both Palestine and Israel," "Biased against others," "Unclear," and "Not Applicable." To manage the large dataset efficiently, batch processing was adopted with a batch size of 100 samples per iteration. This approach facilitated computational efficiency and reduced memory usage. The zero-shot classification process involved tokenization, input encoding, model inference, and label assignment. Hyperparameters were tuned to optimize the model's performance in assigning labels to the data accurately. Labeled data was saved after each batch processing step, utilizing formats such as Excel files for easy accessibility and compatibility with downstream tasks. In this project, we initiated the text classification pipeline using the specified pre-trained model, namely "facebook/bart-large-mnli". The pipeline was instantiated with the model using the `pipeline()` function, specifying the task as "zero-shot-classification" and the pre-trained model to be used. This approach leveraged the convenience and efficiency of the pipeline architecture for seamless integration of the model into the data labeling workflow. By utilizing the pipeline, we were able to streamline the classification process, from tokenization to inference, while

benefiting from the model's capabilities in zero-shot classification. This ensured a consistent and efficient methodology for labeling the dataset batch-wise, contributing to the overall success of the research endeavor.

1.7 Set Quality Standards:

1. Set and follow proper guidelines: Once the guidelines are set, annotators should make sure to thoroughly understand them by familiarizing themselves with the definitions, criteria, and examples of the different types of biases before annotating the posts.
2. Clear communication: Annotators should maintain regular communication with one another as well as the project organizers. Any challenges or concerns should be discussed in order to reach clarification and accuracy. Open communication helps address ambiguities and ensures consistency.
3. Regular Quality Checks: Annotators should seek quality checks from a supervisor or an experienced annotator if possible to assess the annotations' consistency and quality. Reviewing all annotated texts might be impossible, so selecting a random sample to be reviewed to verify if the guidelines were followed correctly. Based on the quality check, the annotators can be given feedback and suggestions for improvements.
4. Collaborative feedback: Annotators should have constant discussions to correct each other's mistakes and learn from one another. Doing so helps annotators share insights, discuss challenges, and receive guidance to enhance their annotation process.

Following the steps mentioned above can contribute to the overall quality, consistency, and accuracy of the annotation process.

1.8 Handle Ambiguities:

Whenever any of the team members face ambiguous texts they do the following steps:

- 1- Review the guidelines again individually and search (e.g. what some words might indicate if it is not clear enough).
- 2- Check with other team members to share views collaboratively.
- 3- Additional help from our professors at SQU to check.

1.9 Ensure Consistency:

- Annotators will undergo calibration sessions to align their interpretations and annotations.
 - Consistency checks will be performed periodically to monitor annotator agreement.
-
1. **Consistency and Accuracy:** Check for consistency and accuracy in the information presented. Biased texts may contain inconsistencies or inaccuracies, especially if they prioritize promoting a particular viewpoint over factual correctness.

1.10 Ethical Considerations:

Regardless of the annotators' backgrounds, they should maintain objectivity throughout the process. Here are some of the considerations that the annotators followed to maintain unbiased annotation practices when handling sensitive data:

1. Impartiality: The task should be approached without any personal prejudice or biases.

Annotators should focus on identifying instances of bias exhibited in the texts solely based on the defined criteria and guidelines, regardless of personal beliefs and opinions.

1. Fair representation: Annotators should provide a balanced representation of both parties involved in the conflict within the texts. This includes considering different sources, avoiding focusing on specific information, and analyzing different points of view.
2. Sensitivity: Annotators have to deal with sensitive data with care, especially when it comes to portraying individuals or groups that are directly involved in the conflict. They should consider how their annotations might impact the rights and dignity of those involved.
3. Transparency: Annotators should discuss the reasoning behind their annotations to ensure that the process is accountable and allows for review if needed.
4. Fact-based approach: Annotators should prioritize factual information and evidence while carrying out their annotation process. Doing so helps maintain objectivity and avoid any subjective and biased opinions.

It is important to note that regardless of the annotators' personal beliefs on a sensitive topic like this, they should maintain an unbiased and professional act while carrying out the task.

1.11 Training and Support:

There were no training procedures that the annotators took but they sought support from the following resources:

1. The team members searched for resources and sought guidance from experienced individuals, including their professors and professionals in the field of NLP.

2. The Annotators supported each other and shared their experiences within the team.
3. Annotated sample texts were used to assist annotators.
4. The team members had constant communication to motivate one another.
5. The guidelines were always accessible to annotators to ensure consistency and accuracy.
6. The collaboration and knowledge shared among team members were encouraged to enhance understanding and performance.
7. External resources, such as research papers and online forums, were utilized to further expand knowledge and address specific challenges.

1.12 Review and Update:

Taking into account feedback and new insights, we schedule guideline reviews by doing the following:

1. We check on each other's work once every week by holding meetings either online or face-to-face to provide feedback and discuss any updates.
2. Send each other any update emails / contact each other during the week.
3. Ask the organizers for clarification.

1.13 Feedback Mechanism:

Feedback for this task was gained from more than one party, and it is further discussed in the following points:

1. Annotators' Feedback:

- The annotators provided each other feedback at every chance possible. They did regular check-ins on each other's work and suggested ways to improve certain issues whenever needed.

2. Stakeholder Engagement:

-The annotators contacted project managers, researchers, and professors to obtain feedback on the effectiveness and relevance of annotation guidelines.

Once feedback was obtained, the team members moved on to the next step which was:

- Documentation and Actionable Insights: The feedback from both annotators and stakeholders was documented and analyzed and based on that action was taken and improvements were made to the process of setting guidelines and annotation.

Overall, the annotators focused on implementing the robust review and updating the process, along with applying effective feedback mechanisms and supporting continuous improvement and optimization of annotation guidelines and processes.