

Annotation Guidelines

Annotator Training Process

Before annotating the dataset, the two annotators underwent annotation training to prepare themselves for the task of annotating propaganda techniques on the code-switched dataset. We considered 20 different propaganda techniques whose definitions can be found in the document:

Both annotators were experienced AI researchers with Master's degrees, fluent in English, Urdu, and Roman Urdu. Annotating propaganda techniques presents a greater challenge compared to tasks such as image classification or sentiment analysis. This is due to the annotators' requirement to comprehend each propaganda technique, recall and identify suitable techniques for specific texts, and accurately label the text spans associated with each propaganda technique. The training process consists of three stages.

For Stage 1 and Stage 2, both annotators were provided with English text examples containing propaganda. In **Stage 1**, both annotators were given a subset of these examples (Set A) and were asked to independently annotate them. The annotators had to identify and list all possible labels. Once they completed their annotations, both annotators compared their results and cross-question each other to resolve any conflicts that may arise. After reconciling any discrepancies, they prepared a final document.

Stage 2 is a repeat of Stage 1 but with a different set of examples. The annotators annotated Set B independently and prepared the document the same way they did for Set A.

In **Stage 3**, the annotators discussed their annotated documents with a domain expert who provided valuable feedback to refine their understanding of propaganda techniques. After a discussion, both annotators updated their knowledge state and repeated the entire process, this time focusing on annotating the text spans and corresponding propaganda labels.

This thorough training program ensured that the annotators had a comprehensive understanding of propaganda techniques and were well-prepared to annotate our code-switched dataset.

Definitions

1. Presenting Irrelevant Data (Red Herring)

Introducing irrelevant material to the issue being discussed, so that everyone's attention is diverted away from the points made.

Example 1: In politics, defending one's own policies regarding public safety - "I have worked hard to help eliminate criminal activity. What we need is economic growth that can only come from the hands of leadership."

Example 2: "You may claim that the death penalty is an ineffective deterrent against crime -- but what about the victims of crime? How do you think surviving family members feel when they see the man who murdered their son kept in prison at their expense? Is it right that they should pay for their son's murderer to be fed and housed?"

2. Misrepresentation of Someone's Position (Straw Man)

When an opponent's proposition is substituted with a similar one which is then refuted in place of the original proposition.

Example: Zebedee: What is your view on the Christian God?

Mike: I don't believe in any gods, including the Christian one.

Zebedee: So you think that we are here by accident, and all this design in nature is pure chance, and the universe just created itself?

Mike: You got all that from me stating that I just don't believe in any gods?

Explanation: Mike made one claim: that he does not believe in any gods. From that, we can deduce a few things, like he is not a theist, he is not a practicing Christian, Catholic, Jew, or a member of any other religion that requires the belief in a god, but we cannot deduce that he believes we are all here by accident, nature is chance, and the universe created itself.

3. Whataboutism

A technique that attempts to discredit an opponent's position by charging them with hypocrisy without directly disproving their argument.

Example 1: a nation deflects criticism of its recent human rights violations by pointing to the history of slavery in the United States.

Example 2: "Qatar spending profusely on Neymar, not fighting terrorism"

4. Causal Oversimplification

Assuming a single cause or reason when there are actually multiple causes for an issue.

It includes transferring blame to one person or group of people without investigating the complexities of the issue

Example 1: "President Trump has been in office for a month and gas prices have been skyrocketing. The rise in gas prices is because of President Trump."

Example 2: The reason New Orleans was hit so hard with the hurricane was because of all the immoral people who live there.

Explanation: This was an actual argument seen in the months that followed hurricane Katrina.

Ignoring the validity of the claims being made, the arguer is blaming a natural disaster on a group of people.

Example 3: if France had not have declared war on Germany then world war two would have never happened.

5. Obfuscation, Intentional vagueness, Confusion

Using words which are deliberately not clear so that the audience may have its own interpretations.

For example when an unclear phrase with multiple definitions is used within the argument and, therefore, it does not support the conclusion.

Example: It is a good idea to listen to victims of theft. Therefore if the victims say to have the thief shot, then you should do that.

Explanation: the definition for "listen to" is equivocated here. In the first case it means listen to their personal account of the experience of being a victim of theft. Empathize with them. In the second case "listen to" means carry out a punishment of their choice.

6. Appeal to authority

Stating that a claim is true simply because a valid authority or expert on the issue said it was true, without any other supporting evidence offered. We consider the special case in which the reference is not an authority or an expert in this technique, although it is referred to as Testimonial in literature.

Example: Richard Dawkins, an evolutionary biologist and perhaps the foremost expert in the field, says that evolution is true. Therefore, it's true.

Explanation: Richard Dawkins certainly knows about evolution, and he can confidently tell us that it is true, but that doesn't make it true. What makes it true is the preponderance of evidence for the theory.

Example 2: "According to Serena Williams, our foreign policy is the best on Earth. So we are in the right direction."

Details: since there is a chance that any authority can be wrong, it is reasonable to defer to an authority to support a claim, but the authority should not be the only justification to accept the claim, otherwise the Appeal-to-Authority fallacy is committed.

7. Black-and-white Fallacy

Presenting two alternative options as the only possibilities, when in fact more possibilities exist. We include dictatorship, which happens when we leave only one possible option, i.e. when we tell the audience exactly what actions to take, eliminating any other possible choices.

Example 1: You must be a Republican or Democrat. You are not a Democrat. Therefore, you must be a Republican

Example 2: I thought you were a good person, but you weren't at church today.

Explanation: The assumption here is that if one doesn't attend church, one must be bad. Of course, good people exist who don't go to church, and good church-going people could have had a really good reason not to be in

church.

Example 3: There is no alternative to war

8. Name Calling or Labeling

Labeling the object of the propaganda campaign as either something the target audience fears, hates, finds undesirable or loves, praises.

Examples: "Republican congressweasels", "Bush the Lesser" (note that lesser does not refer to "the second", but it is pejorative)

Guidelines for span selection: include the object of the labeling if possible

9. Loaded Language

Using specific words and phrases with strong emotional implications (either positive or negative) to influence an audience.

Example 1: "[...] a lone lawmaker's childish shouting. "

Example 2: "how stupid and petty things have become in Washington"

Guidelines for span selection: focus on the adjectives only, even if the context seems to be needed to convey the technique, es. Truly ill in how he manages his financial life

10. Exaggeration or Minimisation

Either representing something in an excessive manner: making things larger, better, worse (e.g., "the best of the best", "quality guaranteed") or making something seem less important or smaller than it really is (e.g., saying that an insult was just a joke).

Example 1: "Democrats bolted as soon as Trump's speech ended in an apparent effort to signal they can't even stomach being in the same room as the president "

Example 2: "We're going to have unbelievable intelligence" Example 3: I was not fighting with her; we were just playing.

11. Flag-waving

Playing on strong national feeling (or to any group; e.g., race, gender, political preference) to justify or promote an action or idea

Example 1: "patriotism mean no questions" (this is also a slogan)

Example 2: "entering this war will make us have a better future in our country."

12. Doubt

Questioning the credibility of someone or something.

Example: A candidate talks about his opponent and says: Is he ready to be the Mayor?

13. Appeal to fear/prejudice

Seeking to build support for an idea by instilling anxiety and/or panic in the population towards an alternative.

In some cases the support is built based on preconceived judgements.

Example 1: "either we go to war or we will perish" (this is also a Black and White fallacy))

Example 2: "we must stop those refugees as they are terrorists"

14. Slogans

A brief and striking phrase that may include labeling and stereotyping. Slogans tend to act as emotional appeals.

Example 1: "The more women at war . . . the sooner we win."

Example 2: "Make America great again!"

15. Thought-terminating cliché

Words or phrases that discourage critical thought and meaningful discussion about a given topic. They are typically short, generic sentences that offer seemingly simple answers to complex questions or that distract attention away from other lines of thought.

Examples: It is what it is; It's just common sense; You gotta do what you gotta do; Nothing is permanent except change; Better late than never; Mind your own business; Nobody's perfect; It doesn't matter; You can't change human nature.

16. Bandwagon

Attempting to persuade the target audience to join in and take the course of action because "everyone else is taking the same action".

Example 1: Would you vote for Clinton as president? 57% say yes

Example 2: 90% of citizens support our initiative. You should.

17. Reductio ad hitlerum

Persuading an audience to disapprove an action or idea by suggesting that the idea is popular with groups hated in contempt by the target audience. It can refer to any person or concept with a negative connotation.

Example 1: "Do you know who else was doing that ? Hitler!"

Example 2: "Only one kind of person can think in that way: a communist."

18. Repetition

Repeating the same message over and over again so that the audience will eventually accept it.

19. Smears

A smear is an effort to damage or call into question someone's reputation, by propounding negative propaganda. It can be applied to individuals or groups.

20. Glittering Generalities (Virtue)

These are words or symbols in the value system of the target audience that produce a positive image when attached to a person or issue. Peace, hope, happiness, security, wise leadership, freedom, "The Truth", etc. are virtue words. Virtue can be also expressed in images, where a person or an object is depicted