YNACC Annotation Guidelines

1 Expert Guidelines

Purpose

The goal of this test is to annotate and classify comments and comment threads. Advisory: Please be aware that the comments come from random, unfiltered samples. Some of the content of this test may be racist, bigoted, vulgar, or exceptionally mean.

Test Structure and Guidelines

Each "query" on the test is a comment thread from a Yahoo news article. We need to judge each comment in the thread, as well as the thread overall. The article title and URL are given for context. It won't always be necessary to read the article and in some cases a URL might not be available. Context can often be inferred from the article title. You will judge each comment based on sentiment, tone, (dis)agreement, topic, audience, and persuasiveness. You will then judge the overall subdialogue based on three categories, 1) constructiveness, 2) type classification and 3) agreement. Lastly, you provide a confidence score based on how confident you are in your judgement of the subdialogue as a whole and are given the opportunity to leave comments.

Per Comment Annotation

Judge each of the following factors for every comment. You will judge the sentiment, tone, agreement, topic, audience, and persuasiveness.

1. SENTIMENT SCORE: Strongly negative to strongly positive, neutral, or mixed. This is your opinion on the overall tone or sentiment of the comment. When considering sentiment, consider how the user feels with respect to what information they are trying to convey. Different sentiment types are:

   - A positive sentiment generally expresses feelings of positive emotion, for example "I love the Yankees! Gunna be a great year." When users express their opinion on something as being great or good, this is a positive sentiment. Attempts at making jokes or at being funny to "lift the mood" generally have a positive sentiment, unless they are mean-spirited.
   - A negative sentiment is more common and indicates the commenter is unhappy for some reason. Usually goes in hand with some complaint about the world or the issue. For example, "Donald Trump sucks. I can't believe he's allowed to be in the public eye." Additionally, mean or controversial comments are usually coming from a negative emotional place on the part of the commenter. Sarcasm, although funny, is usually used to express discontent or absurdity, is usually negative.

- A neutral sentiment has no emotion. Usually when a commenter is just stating fact, like "I think the season starts in July, actually." or "If you water your basil every day, it shouldn't die"

- A mixed sentiment is when a user seems to express both positive and negative emotion about something or several things. A good example is when a commenter expresses that they are in agreement with something one commenter said, but also is offended or takes issue by something else. Longer comments that contain many expressions of different emotions are usually "mixed."

- NA: Used for foreign or illegible comments.

2. TONE: Select all that apply to the overall tone of the comment. But don't force it. It's possible that none may apply.

- Informative (Constructive, Productive): This comment furthers the discussion by adding new information. Usually an attempt to be persuasive and convince others of the user's argument. It may be passionate, but it is not dismissive. NOT a personal story. Criteria for informativeness: 1. Mention of historical facts or evidence 2. Mention of statistics and other numbers 3. Quote or paraphrase of public discourse made by a popular figure 4. Mention of events, or "news" 5. Presenting a cogent or logical analysis or argument

- Controversial (Outspoken): This comment puts forth a strong opinion in a way that others will certainly strongly disagree with.

- Sarcastic: Not all sarcasm is funny, but this will often go hand in hand with the next option.

- Funny: This is an opinion. If you think it's kind of funny or intended to be funny, mark this option. Usually mark this when the comment shows a user is only trying to add humor, not necessarily new information.

- Sympathetic (Warm): The comment is nice, expresses positive emotion or sadness without being mean. Usually the user is expressing some type of agreement concurrently.

- Mean (Hateful): The comment is intended to be rude, mean, or hateful with no other intent. Be careful to not assign this to comments you personally disagree with. Note that mean, hateful comments/insults can still be on topic with the article or conversation.

- NA: To be used sparingly when the tone is neutral, the contribution is garbled, foreign or non-sensical, or tone is unable to be determined.

3. AGREEMENT: Check all that apply. An individual contribution can either agree or disagree with another commenter in the thread or explicitly express agreement or disagreement with the point or a series of points made in the article. If a user isn't explicitly expressing agreement with another commenter, they may be implicitly doing so by providing some "adjunct opinion" (where they express a personal belief that supplements some type of argument to be made in support of or against some point—further defined below).

- Agreement with other commenter: The comment indicates agreement with either another commenter explicitly, meaning, the user is clearly expressing they agree with another person or point of view in the thread. Can occur concurrently with other types of agreement/ disagreement.

- Disagreement with other commenter: The comment indicates disagreement with either another commenter. Can occur concurrently with other types of agreement/disagreement.

- Adjunct/personal opinion: The comment expresses a user's a NEW view or judgement about something—it has to be an opinion that has not yet been articulated in the context of the conversation. Usually starts with "I think," "I feel" or "I believe." BE CAREFUL not to use this if the comment is actually supporting or rejecting a point that has already been made. "Adjunct" opinion means that something is being added to by the user as a supplementary, rather than essential, part of the discussion.
- NA: To be used for foreign, garbled, nonsensical, or unclear comments where the contribution doesn't have any form of agreement or disagreement.

4. TOPIC: Specifically, is the comment off-topic in any way? Or, is the comment somehow related to the topic of other either the conversation or article but slightly less relevant because it is in the form of a personal story?

- Off topic with conversation: The comment has nothing to do with what others are saying in the context of the conversation. It may be relevant to the article in some way, however.
- Off topic with article: The conversation/contributions have begun to be irrelevant to the article (Also can be thought of as a digression). Each contribution that is off topic to the article once an off-topic conversation has begun should get an "off-topic" flag. Once a digression begins, every contribution within the digression is on topic with that conversation but OFF topic with the article.
- Personal story: The comment is intended to reflect a user's personal experience with the issue. Not usually terribly "informative," but can be used to express agreement or disagreement. This is contained within "topic" because often, personal stories can feel slightly off-topic, although the user's intention is often to share the story in an attempt to present anecdotal evidence for a point and be persuasive.
- NA: Used for foreign or illegible comments or those which are perfectly on topic in all ways.

5. AUDIENCE: Choose whether the comment is meant to be 1) a reply to a specific commenter; 2) broadcast message/general audience; 3) NA (used for foreign or garbled comments). *** Please note that a top-level comment with ALWAYS be a broadcast message with a general audience, by nature of being a top-level comment. ***

6. PERSUASIVENESS: Choose whether the comment is 1) persuasive; 2) not persuasive; 3) NA (used for foreign or garbled comments). This is an opinion based on whether you think the user means to be persuasive and how persuasive they seem to be. Generally, the commenter incorporates new information or a personal story, or uses specific language in attempt to convince other users of his or her point. In order for a comment to have true persuasiveness, they must present a well-reasoned argument.

Per Subdialogue Annotation

"Subdialogue" is the name we're giving to the comment threads. After judging each comment separately, judge the entire conversation on the factors (categories) below. Please try to look at the whole conversation in terms of its constructiveness, tone, topicality, and agreement.

1. Type classification (Please choose as many as apply, but try to stick to capturing the overall tone and topicality of the discussion)

- Argumentative ("Back-and-forth"): An argumentative discussion without much credit given by either side, but not particularly mean or hateful. "Lots of back and forth," without a necessary conclusion.

- Flamewar ("Insulting"): Nothing constructive; users are insulting and "yelling" at each other. Not informative.

- Snarky/humorous: The conversation is mostly snark or jokes that may be on- or off-topic. Users are choosing to engage one another in humor rather than argue or sympathize.

- Positive/respectful: Conversation is mostly users are all expressing opinions with respect and empathy.

- Off-Topic / Digression: Comments are completely irrelevant to article and/or each other; OR users start on topic but veer off in another direction.

- Personal stories: Users exchange personal anecdotes.

- NA: Used for foreign or illegible dialogs or for those which are too short to develop into a conclusive "type" of conversation.

2. Agreement (Please choose one)Judge whether the thread: 1) has agreement throughout; 2) has initial disagreement converging to agreement (convincing); 3) has continual disagreement throughout; 4) has agreement initially and devolves to disagreement, or 5) NA: the thread is too short or has too many digressions to rate the agreement level. Please also choose "NA" for foreign or illegible dialogs.

3. Constructiveness (Please choose one)

- Constructive: Constructive conversations do not require "conclusions." A conversation or argument does not have to have a winner or conclusion to be constructive, as long as there is a clear exchange of ideas, opinions, and information done so somewhat respectfully. A constructive conversation should contain one or multiple points of agreement and/or disagreement, all mostly on topic, and be relatively respectful. Comments should contain new information (informative) and/ or may also include intent to persuade. Comments may also seek to contribute humor, sarcasm, even a mean comment if in the context of a passionate attempt at persuasiveness. It will be up to you to decide how much "meanness" degrades the constructiveness—some people are more tolerant than others of disrespectful language when heated arguments occur.

- Non-constructive: Non-constructive conversations are those which are largely unproductive. Usually, the initial commenter's point does not get properly addressed (i.e. conversation(s) does/do not contain a clear communicative goal; conversation is disconnected), is comprised of few attempts at persuasiveness, and each speech act can be taken in isolation. A Subdialogue can also be deemed non-constructive if largely negative (i.e. just back and forth of insults) or if it is "all over the place" in terms of topic.

- NA: Used for foreign or illegible dialogs or for dialogs that are too short to determine constructiveness.

4. Confidence score (Choose between 1 and 5). On a scale of 1 to 5, 1 being not at all confident and 5 being fully confident, how confident are you in your overall judgement of your per-subdialogue annotation (the Type, Agreement, and Constructiveness of the thread as a whole)?

2   Crowdsourced Guidelines[1]

Read the linked article and the comment thread below and choose the best options for the conversation's Type, Agreement, and Constructiveness in the questions that follow.

1. Type (Please choose as many as apply, but try to stick to capturing the overall tone and topic of the discussion)

    - Argumentative ("Back-and-forth"): An argumentative discussion without much credit given by either side, but not particularly mean or hateful. ?Lots of back and forth,? without a necessary conclusion.
    - Flamewar ("Insulting"): Nothing constructive; users are insulting and "yelling" at each other. Not informative.
    - Snarky/humorous: The conversation is mostly snark or jokes that may be on- or off-topic. Users are choosing to engage one another in humor rather than argue or sympathize.
    - Positive/respectful: Conversation is mostly users who are all expressing opinions with respect and empathy.
    - Off-Topic / Digression: Comments are completely irrelevant to article and/or each other; OR users start on topic but veer off in another direction.
    - Personal stories: Users exchange personal anecdotes.
    - NA: Used for conversations that are too short to develop into a conclusive "type" of conversation.

2. Agreement (Please choose only one)
   Judge whether the thread:

    - has agreement throughout;
    - has initial disagreement converging to agreement (convincing);
    - has continual disagreement throughout;
    - has agreement initially and devolves to disagreement, or
    - NA: the thread is too short or has too many digressions to rate the agreement level.

3. Constructiveness (Please choose only one)

    - Constructive: Constructive conversations do not require "conclusions." A constructive conversation should contain one or multiple points of agreement and/or disagreement, all mostly on topic, and be relatively respectful. Comments should contain new information (informative) and/or may include intent to persuade. Comments may also seek to contribute humor, sarcasm, even a mean comment if in the context of a passionate attempt at persuasiveness.
    - Non-constructive: These are largely unproductive threads. Usually, the initial commenter's point does not get properly addressed, comments contains few attempts at persuasiveness, and each comment can be taken in isolation. A conversation can also be deemed non-constructive if largely negative (i.e. just back and forth of insults) or "all over the place" in terms of topic.
    - NA: Used for discussions that are too short to determine constructiveness.

---

[1]These guidelines were derived from the expert guidelines above but shortened and simplified for the platform.