

MMR Guidelines - Supplementary Materials

Overview

Imagine a search bar in your gallery app that allows you to search memories to recall a specific moment. Users can search memories by concept, visual content, or life moments, the system will then display the most relevant. To ensure quality results, we need to make sure that the videos selected by the assistant fit the memories the user is referencing.

Task

In order to evaluate the relevance of video search results you will be shown a natural language query, paired with a video from an imagined user's video album. You will review the pair and determine the relevance score for that query, based on how likely you think it is that the query will select the correct result, without selecting a lot of other, loosely related videos

QUERY RELEVANCE

A query is considered **relevant** if all the search criteria in the query are evident in the video content. If any or all of the criteria in the query are not present in the video, it is considered **not relevant**. A query cannot be partially relevant, if the video does not meet *all* of the criteria, then it must be considered not relevant. The video may also contain elements that are not referenced in the query, this is still considered relevant as long as all the referenced criteria *are* present.

Consider the following examples of possible queries:

Query:

- Show me a video of dogs playing in the backyard
- Show me a video of the sunset at the beach
- Show me videos of singing with kids
- Show me videos baking with Amanda
- Show me videos of birthday

Must Contain:

- Dogs, Backyard, Playing
- Sunset, Beach
- Singing, Kids
- Baking, Amanda
- Birthday

More detailed queries will have stricter criteria to be met. If the query contains adjectives, qualifiers, or quantifiers, the accompanying image must meet a reasonable interpretation of those criteria, as in the following:

Query:

- Show me a video of some people in orange jackets dancing
- Show me a video of 3 kids singing
- Show me videos of a tall guy biking

Must Contain:

- More than one person, Orange jackets, Dancing
- Singing, Three Kids
- Biking, Guy, Tall

Some descriptions may seem relative or subjective, e.g. *tall, orange*. If the description seems close enough that the average person could reasonably consider it correct, it doesn't need to be exact, e.g. a patterned jacket with mostly orange in it, or a shade of peach close to orange could both be considered *orange*. If the interpretation seems likely and plausible you may rate them as relevant, **if you aren't sure, err on the side of caution and consider them not relevant.**

DEALING WITH AMBIGUOUS OR VAGUE QUERIES

Some queries may have more than one reasonable interpretation, for example, *Show me videos of Mountain View* could refer to videos that show views of mountains, or to the city of Mountain View. Similarly, *Show me videos of the big apple* could refer either to New York city or a really large fruit. Some queries might seem vague, or it is difficult to tell from the video whether the specific criteria is present or not, for example, the query *Show me videos of walking in Paris*, might display a video of a cityscape with French signs and cobble streets, but without a specific landmark that identifies it as Paris'. This could still be considered **relevant**. Another query might be *Show me videos of birds on the beach at sunset*, the birds and the beach are present, but it isn't possible to tell from the duration of the video whether the sun is rising or setting, this could also be considered **relevant**. Use your best judgement with ambiguous or vague queries. As above, if the interpretation seems likely and plausible (even without concrete specific evidence) you may rate them as **relevant**.

If you aren't sure, err on the side of caution and consider them not relevant. For example, considering the examples above: if the scene looks kind of Parisian, but the street signs are in Spanish, then it is much less likely that this is Paris, and the query should be considered **irrelevant**. Likewise, the sun was neither setting or rising (ambiguous) but unambiguously high in the sky without a hint of pink or orange, that too should be considered **irrelevant**.

Sensitive videos (Violating vs Explicit Content)

The videos you are reviewing may not be suitable for all viewers. To that end, we will ask you to curate the data set as you label <query, video> pairs and flag sensitive or inappropriate videos/queries. When you reject a job, you will be asked to select a reason from a list of options. Note that there is a difference between **Violating Content** and **Explicit Content**. If you are not sure whether a video is violating or explicit, please select 'Violating.'

Violating content is specifically outlined in the following document and **must** be reported for safety reasons

Explicit Content are videos that do not meet any of the criteria for violating content but still may be considered offensive, inappropriate, or triggering to some people, as in the list below:

1. PII
 - a. First Name & Last Name (just one name is not PII)
 - i. if the first and last name seem to be used in a slot that would indicate a public figure, such as musical artist, fictional character, or political figure, please **do not mark it as containing PII**. If you are unsure check if the name has a Wikipedia page. If so, **do not mark as containing PII**.
 - b. phone numbers, credit card numbers, or social security numbers
 - c. email
 - d. Addresses are NOT considered PII unless they are accompanied by another piece of PII (i.e. name), in which the combined information would allow you to identify the user.
2. Offensive, racist, biased and non-tolerant behavior
 - a. Profanity, slurs, language that is offensive to any cultural, racial, and religious groups.
 - b. Bias towards or unequal treatment to any cultural, racial, and religious groups.
 - c. Anything inconsistent with the values of tolerance and respect for diversity.
3. Sexual or flirtatious behavior
 - a. Any reference to sexual behavior or materials, legal or illegal.
 - b. Sexual or flirtatious expressions or innuendo.
 - c. Explicit or sexual language or physical descriptions.

4. Controversial and Polarizing Topics
 - a. Disputed regions or events
 - b. Sexuality

Rejection Reasons

1. Query not Loading

- a. Please use this rejection reason if you are unable to annotate the video because the text for the query is not loading

2. Video not Loading

- a. Please use this rejection reason if you are unable to annotate the video because the video is not loading

3. Violating Content

- a. Using the guidance above in the Sensitive Videos section, if any video meets the criteria for violating content, please use this rejection reason and follow the escalation protocol

4. Explicit Content

- a. Using the guidance above in the Sensitive Videos section, if any video meets the criteria for explicit content, please use this rejection reason and follow the outlined protocol

Avoiding Bias

There may be strong personal feelings when viewing query/videos pairs, but the labeling process should be as relevant to the task as possible. Your own political and social views should not influence the way you label the relevance of the query/video pair as it may not be aligned with the user's perspective and because of the implications it imposes. If a query/video pair is problematic, reject the job and provide the reason for rejection.