

准备工作

a. 对于标注员1

- batch1_sampled(用于原始确定性标注)
- batch2_sampled(用于重新标注)
- batch3_sampled(用于重新标注)

b. 对于标注员2

- batch1_sampled(用于重新标注)
- batch2_sampled(用于原始确定性标注)
- batch3_sampled(用于重新标注)

c. 对于标注员3

- batch1_sampled(用于重新标注)
- batch2_sampled(用于重新标注)
- batch3_sampled(用于原始确定性标注)

步骤1: 重新审视讽刺的定义(data_collection)

步骤2: 重新标注batch1_sampled(66个话语)

(<15分钟)

操作步骤

1. 标注员打开“batch1_sampled_a*”文件夹中的文件, 获取batch 1视频的访问权限。
2. 每位标注员根据标注表上列出的ID编号逐个观看“batch1视频”。
3. 观看完每个视频后, 在“标签”栏中给出标签(“s”或“ns”), 并对标记为“s”的话语给出确定性分数。
4. 对于不确定的标记为“a”(表示模糊)。
5. 标注结束后, 通过讨论会解决“a”标签(可以通过多数票达成一致或直接丢弃)。

步骤3: 重新标注batch2_sampled(116个话语)

(<20分钟)

操作步骤

1. 标注员打开“batch2_sampled_a*”文件夹中的文件，获取batch 2视频的访问权限。
2. 每位标注员根据指定标注表上的ID编号逐个观看“batch2视频”。
3. 观看完每个视频后，在“标签”栏中给出标签(“s”或“ns”)，并对标记为“s”的话语给出确定性分数。
4. 对于不确定的标记为“a”(表示模糊)。
5. 标注结束后，通过讨论会解决“a”标签(可以通过多数票达成一致或直接丢弃)。

步骤4: 重新标注batch3_sampled(130个话语)

(<30分钟)

操作步骤

1. 标注员打开“batch3_sampled_a*”文件夹中的文件，获取batch 3视频的访问权限。
2. 每位标注员根据指定标注表上的ID编号逐个观看“batch3视频”。
3. 观看完每个视频后，在“标签”栏中给出标签(“s”或“ns”)，并对标记为“s”的话语给出确定性分数。
4. 对于不确定的标记为“a”(表示模糊)。
5. 标注结束后，通过讨论会解决“a”标签(可以通过多数票达成一致或直接丢弃)。