

4. How many annotated examples do you think you will be able to collect?

Provide a justification for this number (e.g. note how many instances of unlabeled data are in your source repository and guess how many you will have after cleaning and filtering. Then, if your team members will do the annotation, calculate how many items they will need to annotate each week in order to hit your target).

Youtube has around 300 hours of data uploaded every minute. So the size of the source will not be an issue even after cleaning and filtering the data. Manual annotation of the data will be the bottleneck, and a conservative estimate would place the video count that we could annotate to 1000 considering every video is 4 minute long on average. This is 4000 minutes of video-viewing over 4 people which is around 17 hours spent viewing videos.

5. What are the names and usc email addresses of your team members?

Kaladhar Mummadi - mummadi@usc.edu

Akshay Bhobe - abhobe@usc.edu

Tanay Shankar - tshankar@usc.edu

Nikhil Mago - mago@usc.edu