

Creating crowd sourced sentiment analysis for the dataset would need to be relatively structured. We would not have freely typed analysis, and instead would have set sentiments to decide on. If we were able to change the given dataset, we would add some additional options than simply “boredom, enthusiasm and anger”, or at the very least add an ‘other’ option. Along with this, it may be a useful option to be able to select multiple options at once. Regardless of the choice, preset sentiments would need to be chosen to prevent multiple users from using different words to describe the same thing or deviating too much from the desired result set.

With a dataset such as this, we would aim to have a wide range of contributors and be unbiased to who annotates what. To accomplish this, a ‘push’ methodology would be given, one in which the workers would be given data to work with. This dataset is too large to expect one person to annotate the entire thing, and as such a random set of ~50 tweets would be given instead. This set could be randomly selected from the data with the lowest number of notations. This would allow us to gather information which is unbiased, as it is randomly distributed and sourced from multiple different people, giving results which would represent a more realistic distribution of how each tweet was perceived. If we allowed users to pick tweets to annotate themselves, certain controversial tweets could be ‘vandalised’ or otherwise explicitly targeted, and the number of given sentiments for each could vary greatly.

In order to incentivise getting information, it may be possible to provide a small monetary output to users who fill out the data, but it may also result in users who do not provide helpful annotations and who only join for the money. We want relatively less screening for contributors as that could introduce bias and we want a realistic/natural range of people, but we also need to ensure the quality remains high. We could include certain ‘control’ questions randomly distributed in each set of questions which have very obvious correct sentiments to ensure that the contributor is giving a genuine effort, and ignoring results which fail these questions. Ensuring that contributors can only provide one set of annotations and gathering data from large amounts of contributors should reduce the impact of some poor contributors.