

### **Some Questions**

- 1) What will be the things to look out for when collecting the appropriate tweets for a certain event?
- 2) Will there be any hardware or software issues that will arise when collecting these types of data?
- 3) How will I initially pull data using an API so that I will run into the least amount of resistance?
- 4) In the event that the data provider runs into a conflict, what will be my role in making sure the dataset is pure?
- 5) Given a certain amount of data at one point in time, what is the best way to make sure it is safe from potentially harmful changes?
- 6) What measures must be taken in order for my API key to be kept safe? In the event that it is compromised, how shall I proceed with the data already collected?
- 7) If any part of the data shall be omitted for the purpose of running a less but more efficient model, then how should the foregone data be interpreted? If context is lost, then what should be the excuse?
- 8) What are the ethical ways to avoid possible consequences such as harming Twitter users or empowering a misleading trend.?
- 9) What will be the most common trait such that it will be prone to missing data, and how should that be dealt?
- 10) What is the order of importance for this type of data? If the lesser important data is found to be lacking, then should this be mostly ignored? If more important data is lacking, then how should this proceed?
- 11) When would be the practical way to retrieve data? Should we wait a certain amount of days after an event happens? Should this depend on the severity of the event, or the heat of the moment?

- 12) When it is determined to retrieve data, how should it be collected? Should data via tweets be piling in, or do we wait a certain amount of time before that happens?
- 13) Is it possible to deploy such a model in which all reactions are analyzed correctly to different degrees of events?
- 14) For each event and their reactions, should there be an expectation that some events should trigger bigger reactions than others? What would be the consequences if the dataset were to take this into account.
- 15) Would using other API's be worth considering? There are other social media platforms like Reddit, Instagram, and Facebook, and it might be worth noting their responses.