

Course Project Proposals

**1. What is the name of your task?**

*WatchTower* – Classification of disaster related tweets as ‘*Actionable*’ or ‘*Non Actionable*’.

**2. Why this is an interesting data set? What would be learned from building a good classifier of this data?**

The disaster dataset has a collection of tweets which was pulled from twitter by querying on a few keywords. The tweets range from topic of discussions related to disasters, calamity, accidents, asking for help under threat or headlines about disasters. The dataset also consists of irrelevant tweets which do not discuss about the disaster topic but has those keywords.

We hope to add to this dataset, a label detailing whether the information in the tweet is ‘*Actionable*’ or ‘*Non Actionable*’.

There is no dataset which draws this insight from the text. This classification tool can be used to do real time threat detection of tweets and possibly provide help for those in need.

**3. How will the data be collected and labelled?**

The Disaster on Social Media data is collected from Kaggle. The dataset currently has labels ‘*Relevant*’ and ‘*Not Relevant*’. We hope to add a new annotation to this dataset based on the utility of the information in the tweets.

We are also planning to pull tweets from Twitter using its search API to ensure we have a statistically significant number of ‘actionable’ tweets.

The two labels we are considering while annotating the dataset are:

- ‘*Actionable*’ – Disaster related tweets which claims for help under emergency situation or threat.  
Example: there is a forest fire at spot pond, geese are fleeing across the street, I cannot save them all
- ‘*Non Actionable*’ – Tweets discussing about disaster post event, or irrelevant tweets.  
Example: Sinking ships burning buildings & Falling objects are what reminds me of the old us.

The rules the team is looking to annotate/label the dataset as ‘*Actionable*’ are:

- Loss of life
- Risk of serious injury
- Danger to general population
- Risk to public necessities

- Evacuation alert / Natural disasters
- Area quarantine

In essence we are looking for tweets that imply a threat to an individual or group of people. A tweet that meets one or more of these conditions is not necessarily *actionable*. But for a tweet to be *actionable* it must meet at least one of these conditions.