**Research Project Objective**

We aim to use supervised machine learning techniques to build an image classifier that can distinguish the difference between images of disaster-related damages and images that do not contain damage related content. The classifier may further label the kind of damages as either structural, natural, fire, flood, or human. Once classified, these social media images, paired with linked text and global positioning information, would support situational awareness in the aftermath of a disaster and provide valuable information for a coordinated response.

**Problem Statement**

In the past few years, the flooding from hurricanes has broken state records in North Carolina, leading to loss of homes, buildings, town infrastructure, and lives. These powerful storms are an example of how a natural disaster may bring chaos and destruction, disabling support and communication systems that residents rely on. During a state of emergency, timeliness of information is vital to help emergency responders, government officials, and disaster response teams reach the hardest hit areas and assess damages for rebuilding.

Social media provides a unique resource for real-time documentation of damages after a natural disaster. However, as with any big data resource, it can be difficult to make sense of the plethora of text and images that are shared. An image classifier that can locate images of major damages from social media posts and determine which kind of disaster response to deploy, would provide an innovative, low-cost, and efficient solution to this data problem.

**The Data**

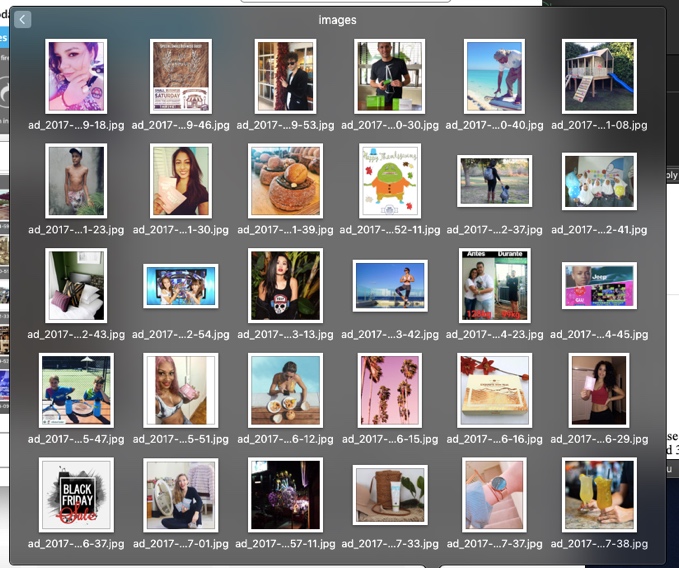
The dataset from which we will build this classifier contains 5,879 social media images, and was obtained from the machine learning repository from UCI Irving. Of these images, 2,906 (49.4%) are labeled as containing damages, and the remaining 2,973 (50.6%) are labeled as non-damages.

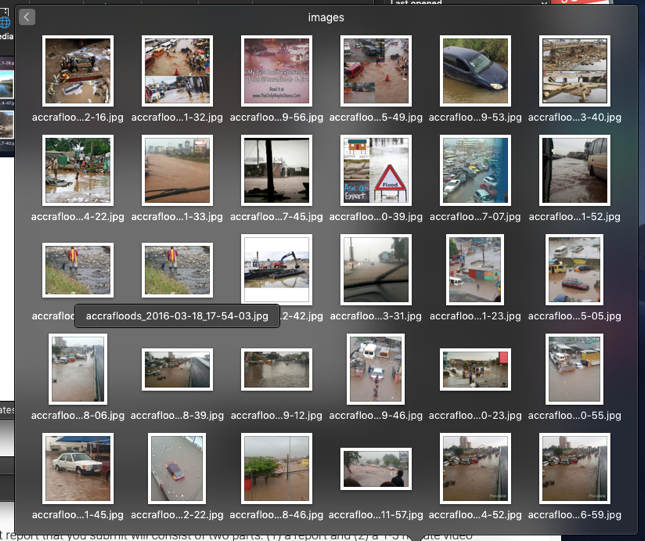
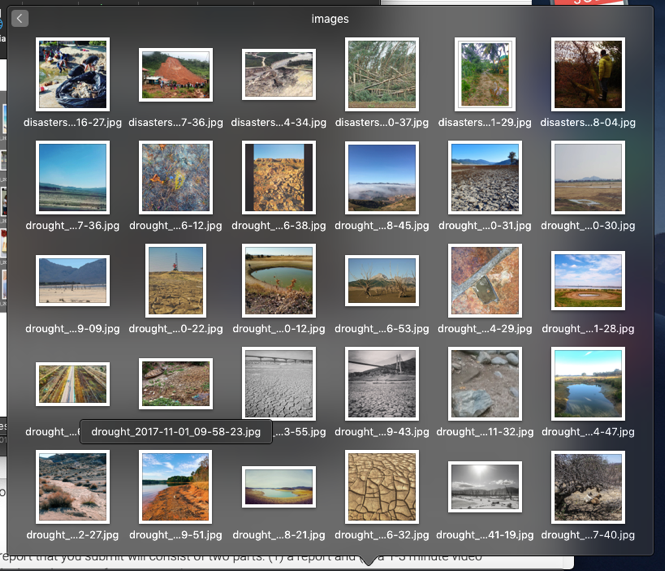
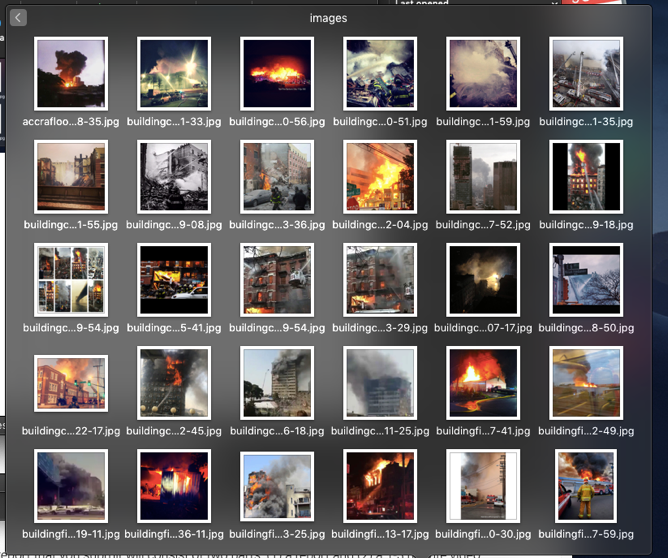
**Proposed Machine Learning Approach**

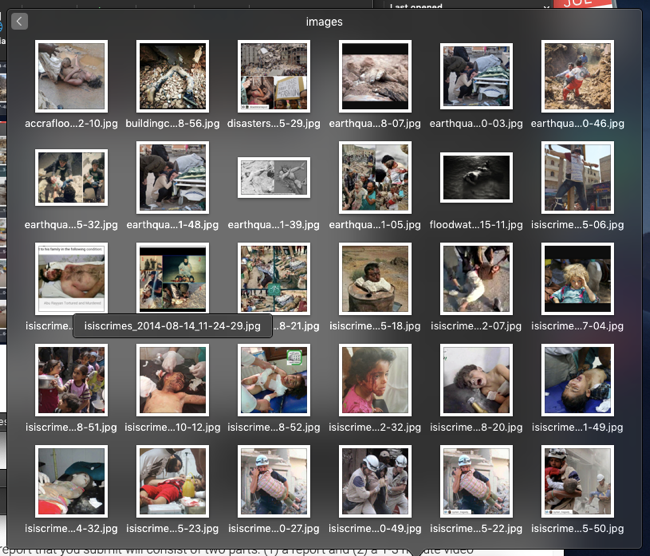
**Evaluation Strategies**

**Real-world Application Test**

Once we have trained and refined our model, we hope to use Twitter’s API to create an additional test dataset to look at images posted during the timeframe of a natural disaster, such as Hurricane Matthew in North Carolina. Bringing the model to a real-life, close to home scenario, brings both meaning and proof of concept to its usefulness outside of an academic setting.

Non Damage Photos (n=2,973) Structural Damage (n=1,419)

Natural Damage (n=516) Fire Damage (n=350)

 Flood Damage (n=386) Human Damage (n=242)