

# CIS 5300 Final Project Proposal: Multimodal Approaches for Disaster Tweets Classification

Gokul Nair, Rohan Saraogi, Shivani Prasad Bondapalli, Yash Agrawal

October 30, 2023

## 1 Motivation

Social media platforms are vital for disseminating information during disaster events. Through platforms like Twitter, individuals share multimedia content on various aspects of the unfolding situation, such as updates on casualties, infrastructure damage, the status of missing or found individuals etc. Numerous studies have highlighted the value of this online information for humanitarian organizations, as when processed swiftly and efficiently, it significantly enhances their situational awareness, enabling them to plan and execute relief operations more effectively. To this end, our goal in this project is to build a multi-modal classifier to determine if a given tweet and/or image is useful for humanitarian aid or not.

## 2 Problem, Dataset & Plan Description

The authors in [1] introduced the CrisisMMD: Multimodal Crisis Dataset and several associated tasks. The dataset consists of several thousands of manually annotated tweets and images collected during seven major natural disasters including earthquakes, hurricanes, wildfires, and floods that happened in the year 2017 across different parts of the world. We plan to use this dataset and to work on Task 1. In particular, this is a binary classification task to determine whether a given tweet or image is useful for humanitarian aid purposes (“Informative”) or not (“Not informative”). The authors consider a tweet/image as “Informative” if it reports/shows one or more of the following: cautions, advice, and warnings, injured, dead, or affected people, rescue, volunteering, or donation request or effort, damaged houses, damaged roads, damaged buildings; flooded houses, flooded streets; blocked roads, blocked bridges, blocked pathways; any built structure affected by earthquake, fire, heavy rain, strong winds, gust, etc., disaster area maps. Images showing banners, logos, and cartoons are not considered as “Informative”.

For our milestone 2 simple baseline we are planning to use a simple majority classifier. For our strong baseline, we are tentatively planning to use a text-only model like LSTM/GRU etc. with pre-trained text embeddings eg. GloVe embeddings. For our milestone 3 extension we are planning to train separate text and image models and feed the output into a final classification model (eg. a logistic regression model) for final prediction. And for our milestone 4 extension we are planning to jointly train text and image models. The architecture that we had in mind is to something similar to that used by the authors in [2], where they concatenate the layer outputs of text and image models and pass it through a series of feed-forward layers for the final prediction.

## References

- [1] Firoj Alam, Ferda Ofli, and Muhammad Imran. “CrisisMMD: Multimodal Twitter Datasets from Natural Disasters”. In: *Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM)*. USA, June 2018.
- [2] Ferda Ofli, Firoj Alam, and Muhammad Imran. “Analysis of Social Media Data using Multimodal Deep Learning for Disaster Response”. In: *17th International Conference on Information Systems for Crisis Response and Management*. ISCRAM. ISCRAM, May 2020.