



Multimodal Combination of Text and Image Tweet for Disaster Response Assessment

Saideshwar Kotha, Smitha Haridasan, Ajita Rattani , Aaron Bowen, Glyn Rimmington, Atri Dutta,

Disaster Resilience Analysis Center

Wichita State University, USA.

Outline

Introduction

Problem Statement

Proposed Method

Dataset

Experimental Results

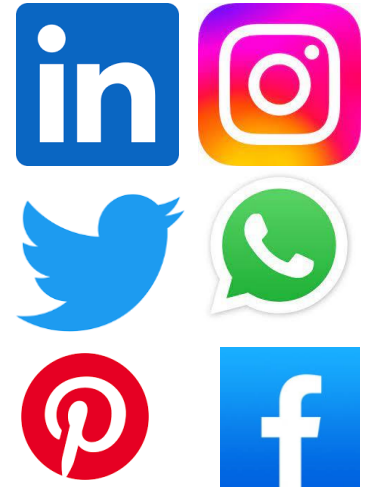
Conclusions and future work



Introduction

Information related to natural disaster is posted on social media

Sharing such information is useful for humanitarian support



Example keywords in tweets useful for Informative category

- Caution
- Help needed
- Warnings
- Rescue
- Donation request

Kathy @Katnow123 · Jul 1

Another **earthquake** in Dubai. Anyone feeling this?! It's still shaking!

 148

 54

 302



[Show this thread](#)

joshdoit8 @joshdoit8 · Jul 1

Another Saturday, another **#earthquake** in Dubai. Only this time, it was longer... almost a minute! Hopefully this quake has not caused any damage or casualty in the UAE and nearby countries. 🙏

Sample Non-Informative Tweet

Anyways, I drove through
the Redwood Forest then
next to the California
wildfires

<https://t.co/yiK5BQh7PF>

Rampant sharing of crisis related posts has led
to the need for categorizing data into
Informative and Non-Informative categories



Humanitarian Aid

- Humanitarian aid workers help people during crisis events to save lives, reduce suffering and rebuild affected communities.

Prior Work

- Ofli et al. evaluated VGG16 on tweet texts and tweet images from CrisisMMD dataset for two categories
 - Informative vs Non-Informative
 - Humanitarian categories

Modality	Informative vs Non-Informative categories (%)	Humanitarian categories (%)
Text only	80.80	70.40
Image only	83.30	76.80
Text + Image	84.40	78.40

Ofli, Ferda, Firoj Alam, and Muhammad Imran. "Analysis of social media data using multimodal deep learning for disaster response." *arXiv preprint arXiv:2004.11838* (2020).

Problem Statement

Aim of this research is to process multimedia data and propose a **fusion-based decision-making technique** to classify social media data into

- **Informative** and
- **Non-Informative** category

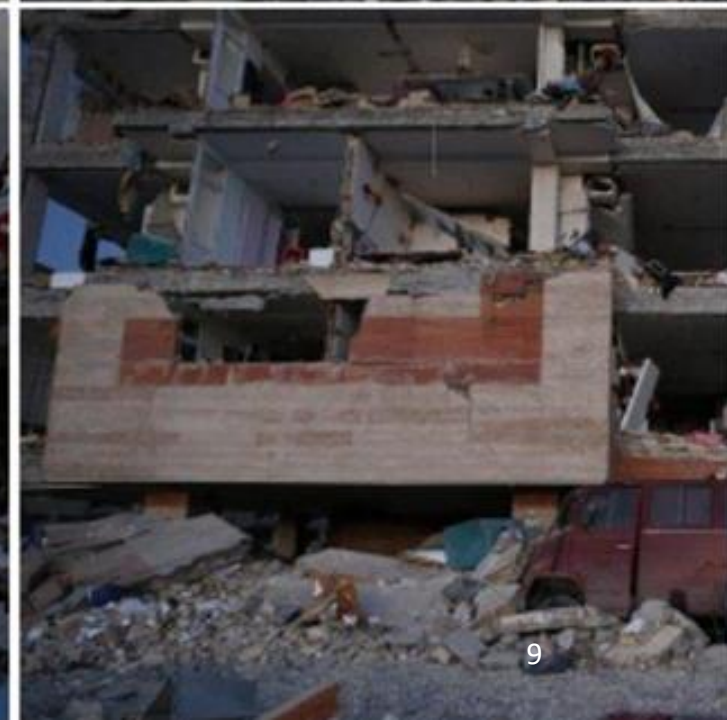
Informative tweets are then classified into various humanitarian categories such as

- Rescue volunteering or donation efforts,
- Infrastructure and utility damage,
- Affected individuals,
- Other relevant information
- Not-Humanitarian.

Dataset - CrisisMMD

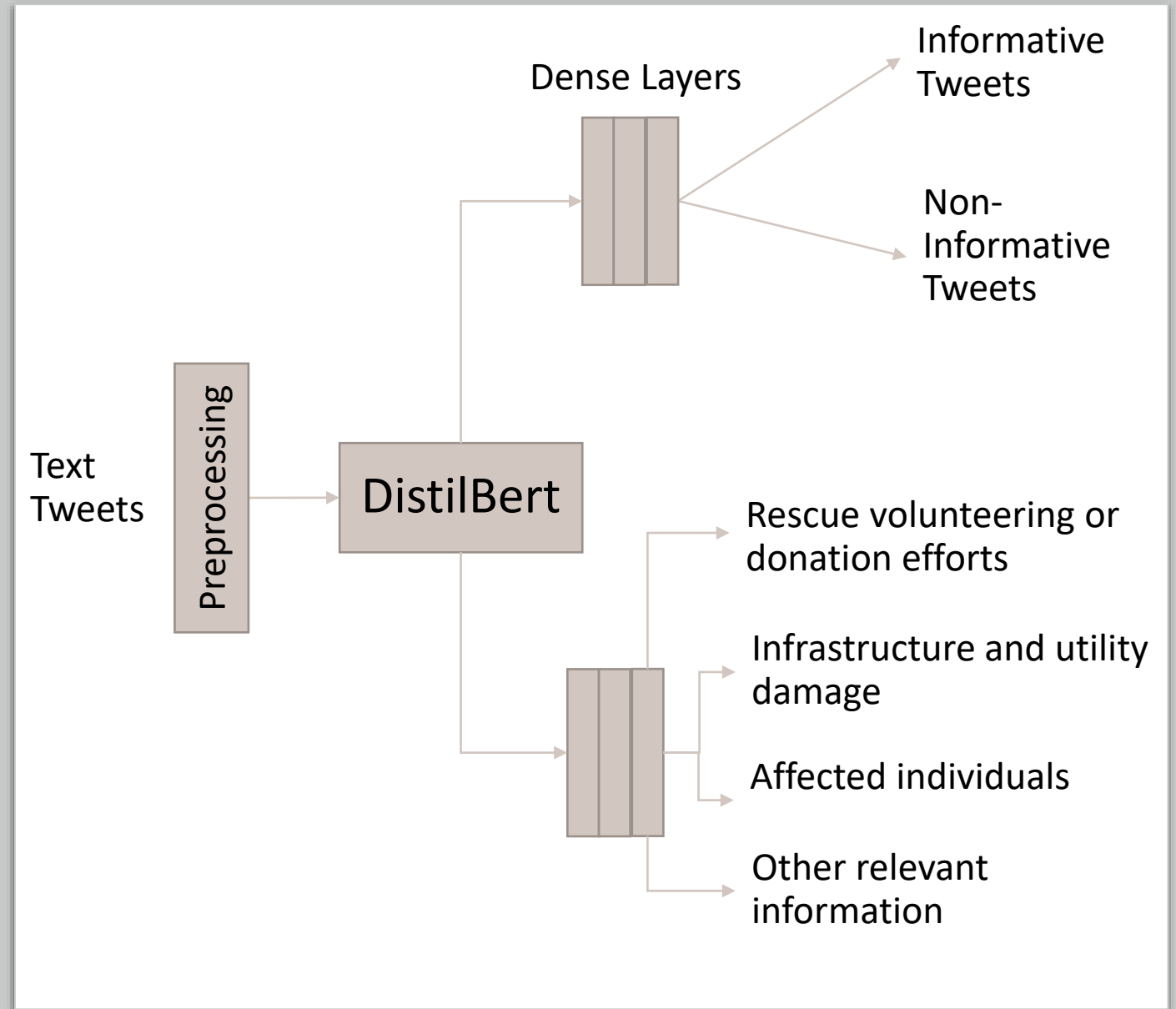
Multimodal Crisis (CrisisMMD) – consists of tweet texts and images collected during multiple major natural calamities including earthquakes, floods, hurricanes and wildfires

The abundance of social media data clearly indicates the possibility of image processing research mainly by assisting humanitarian aid workers.



Proposed Method

Experiment #1: Text Modality for Disaster Response Assessment



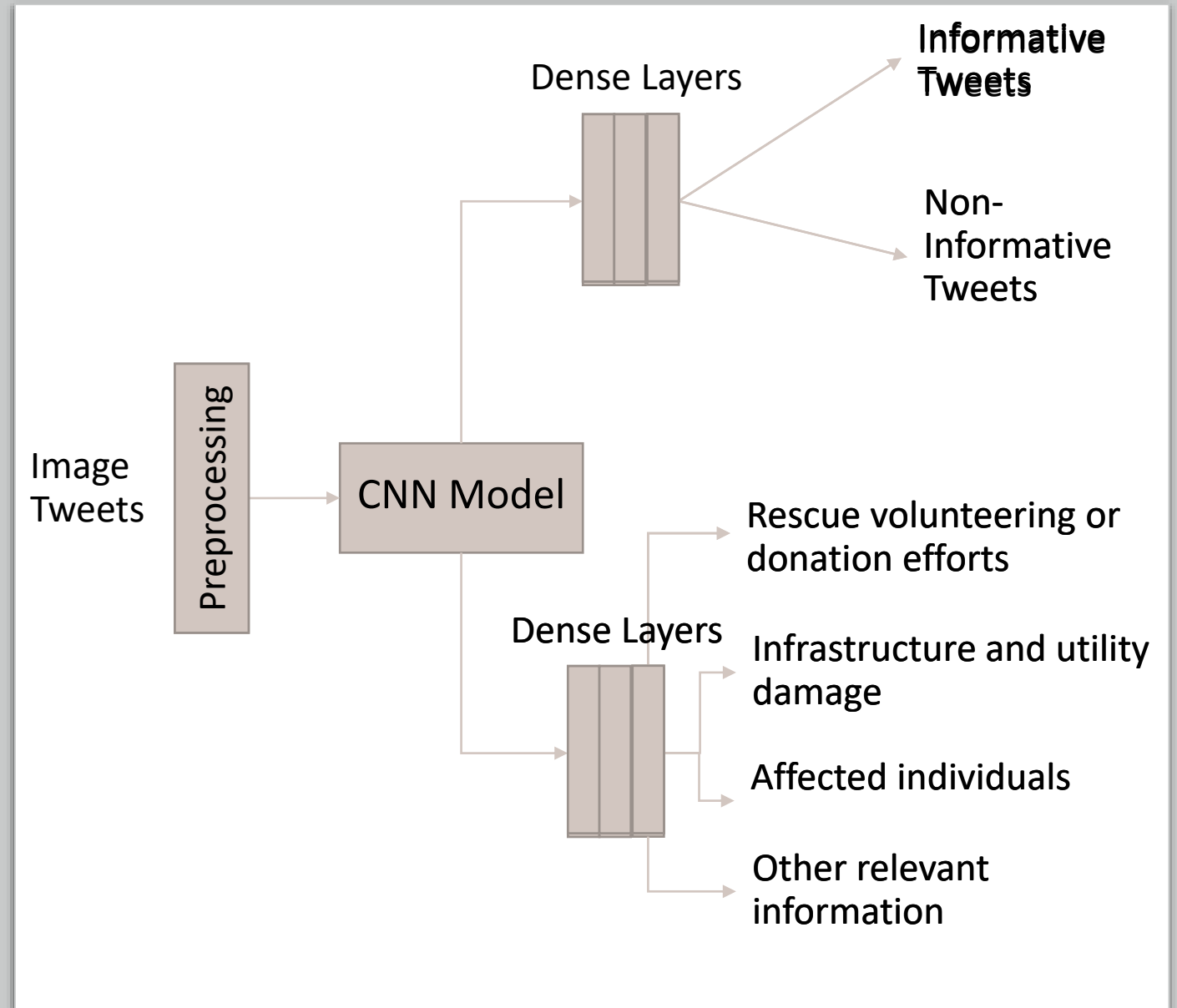
Evaluation of DistilBERT on Text Modality

Text Modality	Informative vs Non-Informative (%)	Humanitarian categories (%)
Test Accuracy	82.65	75.39

Experiment #2: Image Modality for Disaster Response Assessment

CNN models used:

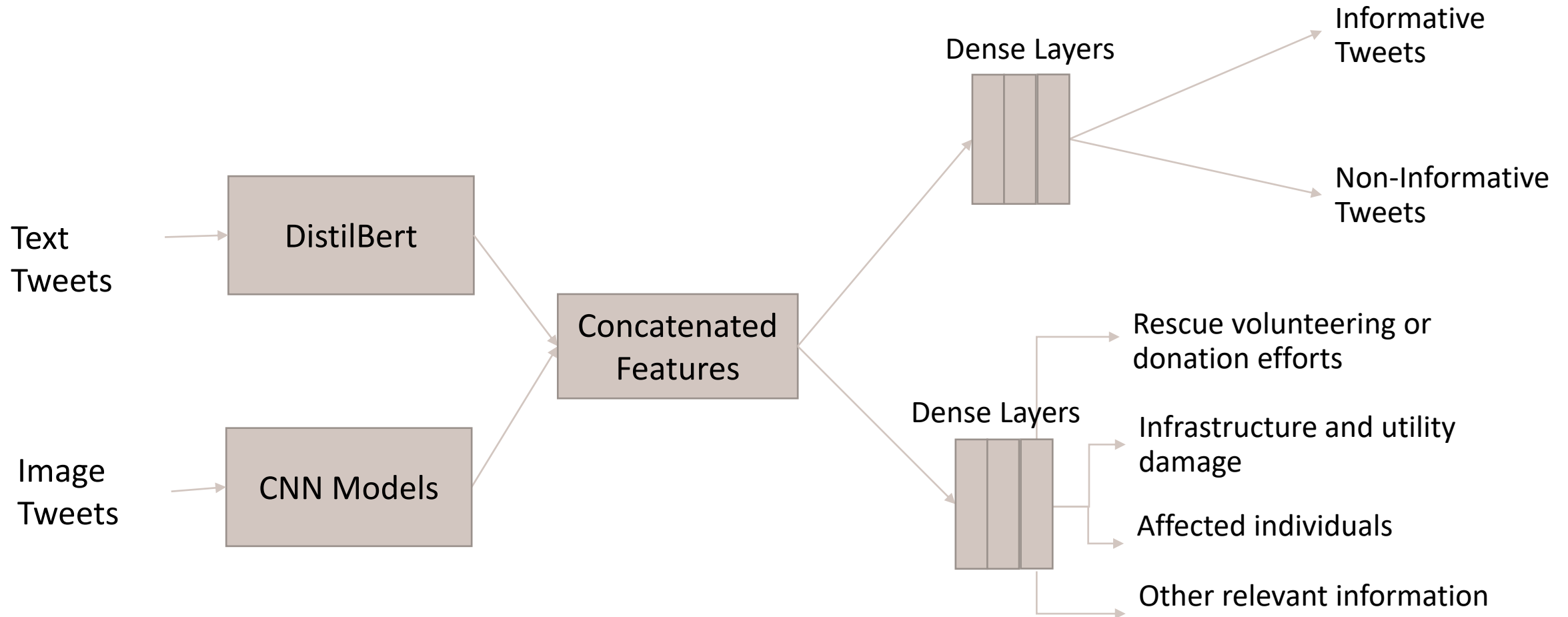
- VGG16
- VGG19
- ResNet50
- DenseNet121
- RegNetY320



Evaluation of CNN Models on Image Modality

Image Modality	Informative vs Non-Informative (%)	Humanitarian categories (%)
VGG16	83.11	73.5
VGG19	83.18	76.23
ResNet50	83.11	77.17
DenseNet121	83.63	76.12
RegNetY320	85.13	80.20

Experiment #3: Fusion of Text and Image Modality for Disaster Response Assessment



Evaluation of DistilBERT on Text + Image Modality

Image Modality	Informative vs Non-Informative (%)	Humanitarian categories (%)
Bert + VGG16	86.50	82.4
Bert + VGG19	86.83	81.67
Bert + ResNet50	88.00	84.71
Bert + DenseNet121	89.50	85.96
Bert + RegNetY320	89.63	86.59

Informative Vs Non-Informative Result Analysis

Modality	Informative vs Non-Informative categories (%)
Text only	82.65
Image only	85.13
Text + Image	89.63 (5% increase)

Humanitarian Result Analysis

Modality	Humanitarian categories (%)
Text only	75.39
Image only	80.20
Text + Image	86.59 (8.19% increase)

Conclusions

Experimental results suggest that fusion of text and image tweets using Multimodal deep learning model on CrisisMMD dataset performs better than either the single text or image modality.

Future Work

- As part of future work, we will explore advanced fusion-techniques for combining text and image modalities and advanced deep learning models to improve the classification performance.



Thank you

Questions