ISCRAM 2022

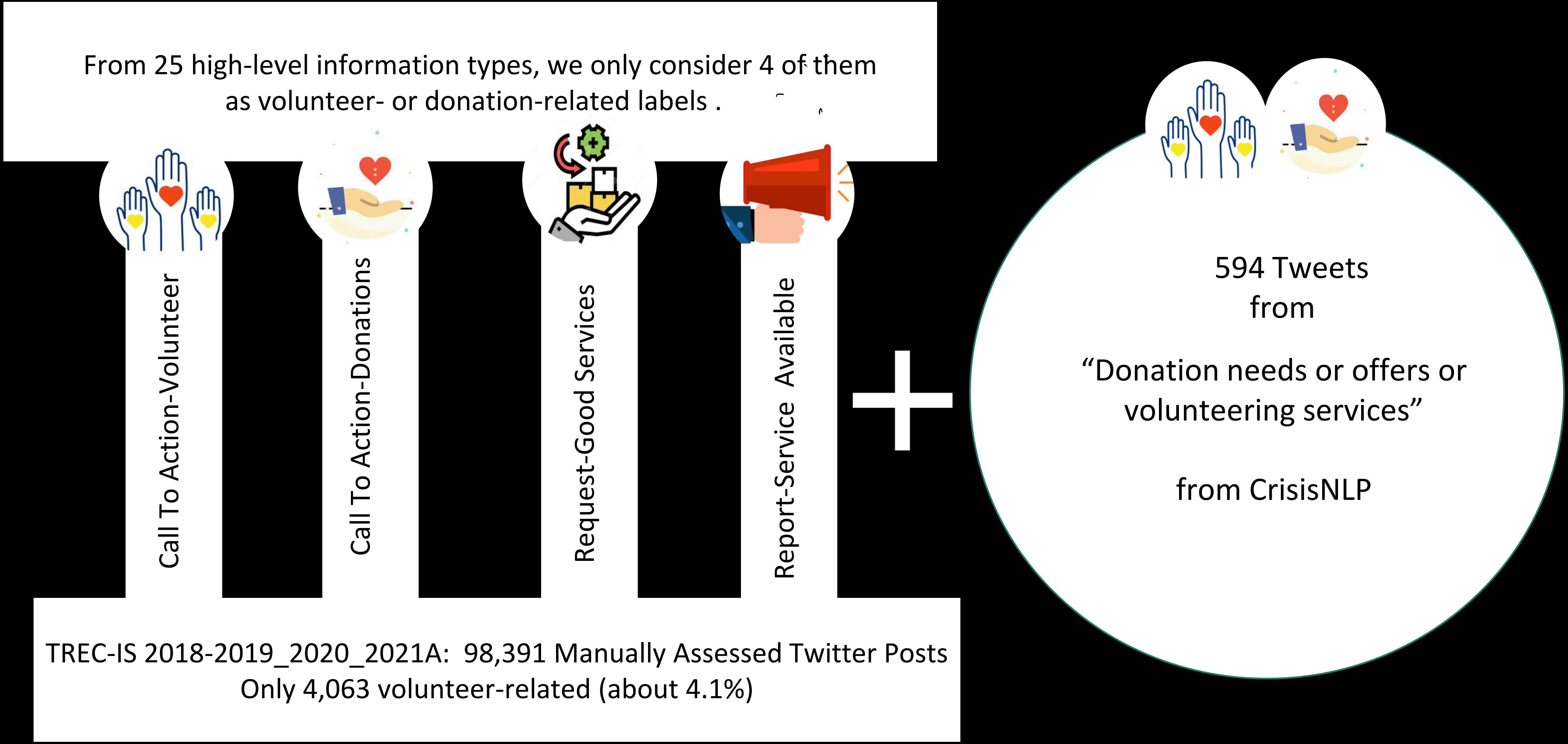
Tarbes, France

"Please Donate for the Affected": Supporting Emergency Managers in Finding Volunteers and Donations in Twitter Across Disasters

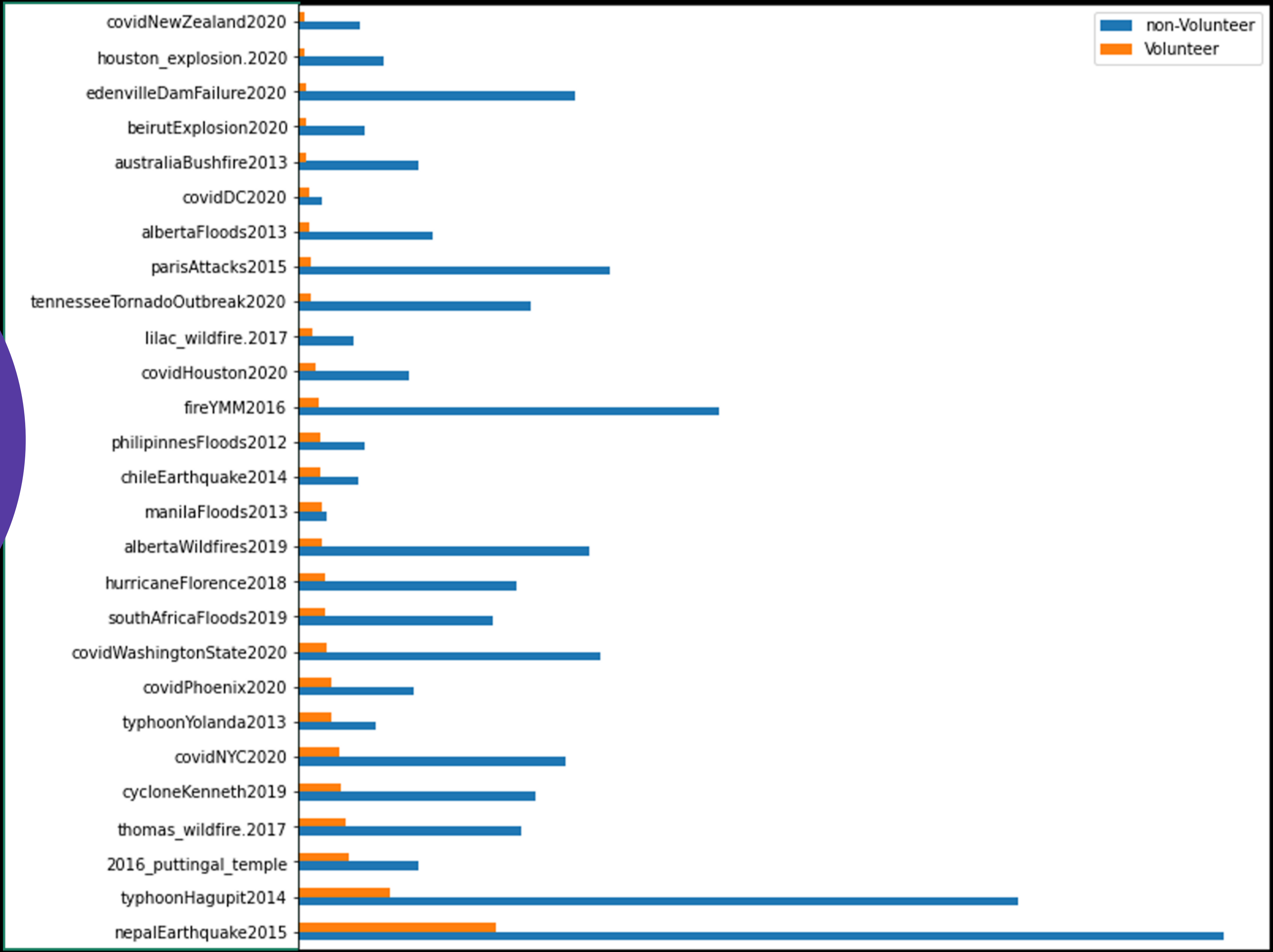
Pooneh Mousavi, University of Maryland, College Park

Cody Buntain, University of Maryland, College Park

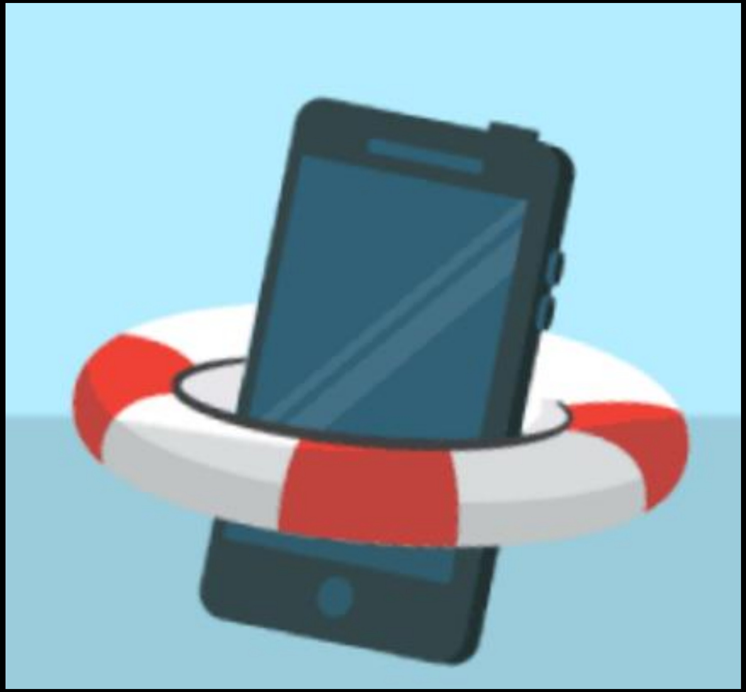
Despite the outpouring of social support posted to social media channels in the aftermath of disaster, finding and managing content that can translate into community relief, donations, volunteering, or other recovery support is difficult due to the lack of sufficient annotated data around volunteerism. This paper addresses these challenges by constructing a general machine learning model that is transferable from one crisis to the other and by introducing a method for integrating domain expertise into language classification to improve the classification accuracy.


From 25 high-level information types, we only consider 4 of them
as volunteer- or donation-related labels .
Call To Action-Volunteer
Call To Action-Donations
Request-Good Services
Report-Service Available
TREC-IS 2018-2019_2020_2021A: 98,391 Manually Assessed Twitter Posts
Only 4,063 volunteer-related (about 4.1%)
+
594 Tweets
from
“Donation needs or offers or volunteering services”
from CrisisNLP

A clear imbalance exists between recovery- and non-recovery-related data.



Analyses of online volunteerism efforts are primarily facilitated by qualitative methods among few crises. This leaves opportunities to study volunteerism and local community-based volunteer groups from a quantitative perspective and evaluate consistencies across disaster events especially in presence of sparsity of volunteerism content.



How similar is volunteerism and recovery language across different events and event-types?

Which strategies are better for re-weighting crisis events of similar types when learning the volunteer/donation information?

What source of weakly supervised volunteerism data leads to the largest performance improvements in our models?

How similar is volunteerism and recovery language across different events and event-types?

Which strategies are better for re-weighting crisis events of similar types when learning the volunteer/donation information?

What source of weakly supervised volunteerism data leads to the largest performance improvements in our models?

Islamic Relief UK
@IslamicReliefUK

This infographic from @decappeal shows the staggering needs in the #Philippines - pls give generously #TyphoonAid twitpic.com/dl0vp0

Oxfam Australia
@OxfamAustralia

Oxfam now preparing to respond to #CyclonePam after it unexpectedly veered west placing number of Vanuatu's islands directly in eye of storm

---

Anna Masera
@annamasera

Donate to #Nepal Earthquake Relief Fund goto.gg/20304 via @GlobalGiving

---

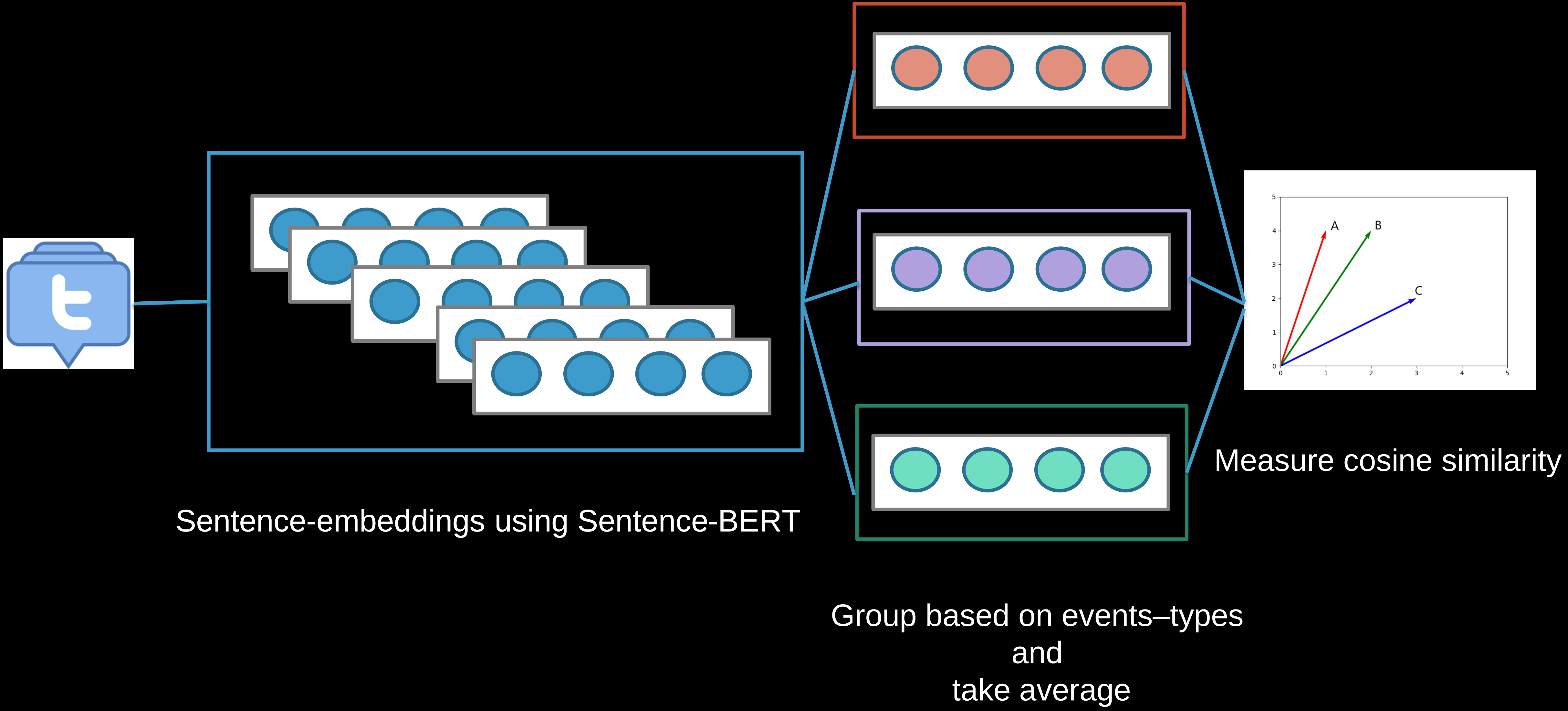
#PrayforthePhilippines

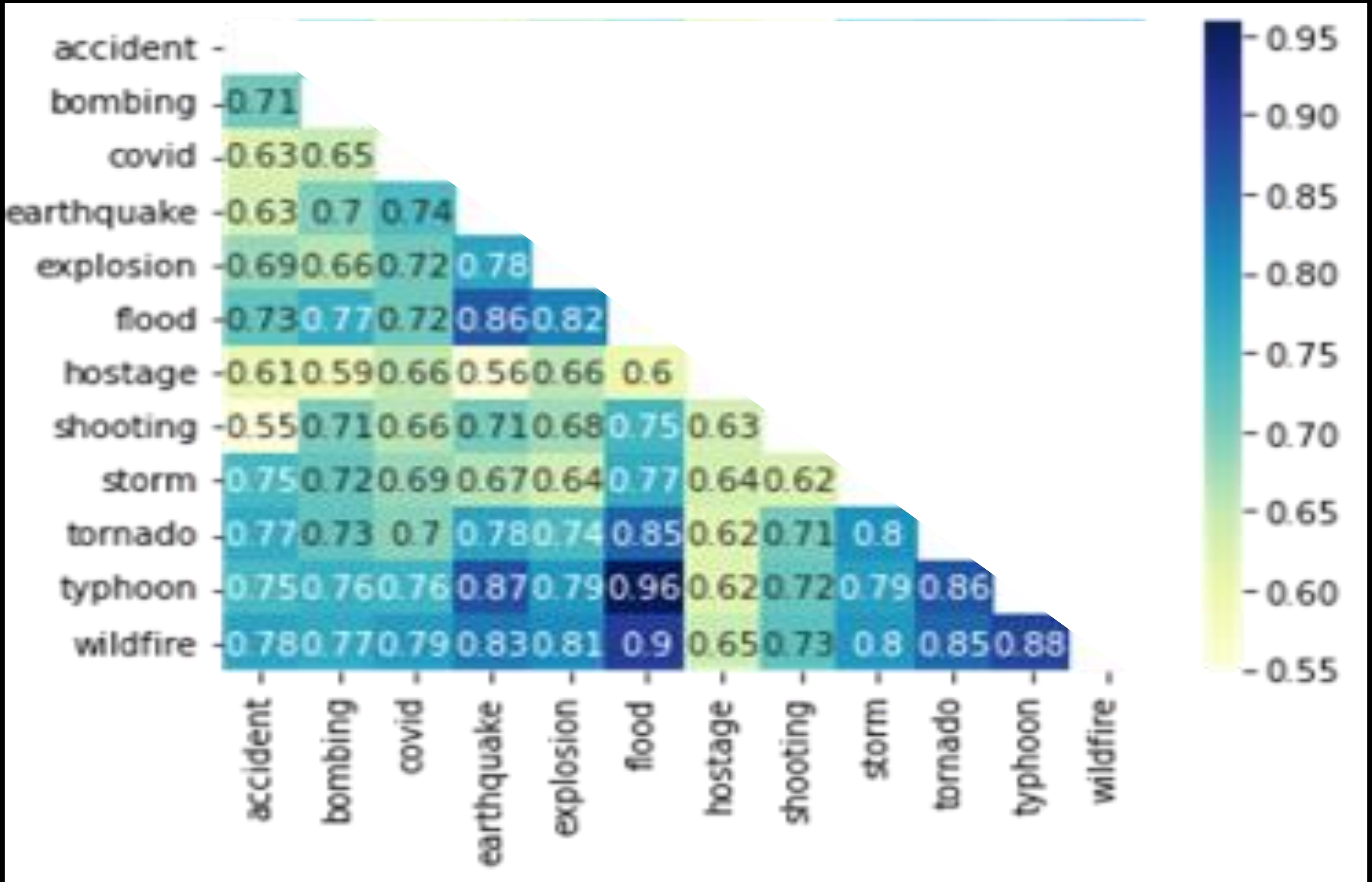
Canada is Ready to Help! Read more here: on.fb.me/1u5KQWn fb.me/34aeCkESD

Prabal Gurung
@prabalgurung

Pls. help my #fundraiser: #Nepal #Earthquake Fund. #Donate & spread the word. Every #dollar counts-PG. fw.to/Yyd30Ga #HELPNEPAL

Top story: Here's how you can help victims of the Paris terror attacks mashable.com/2015/11/14/par..., see more tweetedtimes.com/media/mashable...


Sentence-embeddings using Sentence-BERT
Group based on events–types
and
take average
Measure cosine similarity
A
B
C
0
1
2
3
4
5



CONSISTENCY IN VOLUNTEERISM ACROSS CRISES

Information type	mean	std	# tweets
CallToAction-Donations	0.7425	0.1193	1099
CallToAction-Volunteer	0.6632	0.1519	290
CallToAction-MovePeople	0.6236	0.0983	882
Other-Advice	0.6703	0.1115	3781
Other-Sentiment	0.6552	0.1036	11627
Other-ContextualInformation	0.6422	0.1036	4884
Other-Discussion	0.6297	0.1036	5364
Report-Weather	0.6671	0.1236	8473
Report-MultimediaShare	0.6455	0.1076	24105
Report-News	0.6386	0.1095	19404
Report-Location	0.6332	0.1140	26120
Report-ServiceAvailable	0.6281	0.2179	2533
Report-EmergingThreats	0.6279	0.0970	7367
Report-Official	0.6276	0.1159	3106
Report-ThirdPartyObservation	0.6223	0.1095	19060
Report-Hashtags	0.6018	0.1286	17208
Report-Factoid	0.5992	0.1215	11006
Report-FirstPartyObservation	0.5981	0.1217	5496
Report-NewSubEvent	0.5823	0.1134	2919
Report-CleanUp	0.5806	0.1690	516
Report-OriginalEvent	0.5402	0.1175	5046
Request-GoodsServices	0.6037	0.2144	361
Request-InformationWanted	0.5414	0.1475	509
Request-SearchAndRescue	0.4876	0.1854	308
All-Volunteer	0.7270	0.0855	4063
All-Non-Volunteer	0.6417	0.1083	50927

CONSISTENCY IN VOLUNTEERISM ACROSS CRISES

Result 1

How similar is volunteerism and recovery language across different events and event types?

Volunteer- and donation-related social media content appears similar across disaster events and disaster types to warrant transferring models across disasters.

How similar is volunteerism ad recovery language across different events and event-types?

Which strategies are better for re-weighting crisis events of similar types when learning the volunteer/donation information?

What source of weakly supervised volunteerism data leads to the largest performance improvements in our models?

		Weighting			
		No Weighting	Up Weighting	Event-type weighting	Heuristic Weighting
SAMPLING	No Sampling	None	NONE-UPW	-	-
	Down-Sampling	DOWN	DOWN-UPW	-	-
	Up-Sampling	UP	UP-UPW	UP-EV	UP-HEURISTIC

TRANSFERABILITY OF VOLUNTEERISM DATA ACROSS CRISES

method	precision	recall	f1_score
UP-EV	0.4636	0.6484	**0.5196**
UP	0.4448	0.6695	0.5151
UP-HEURISTIC	0.4552	0.6489	0.5131
DOWN	0.3854	**0.7133**	0.4781
UP-UPW	0.4149	0.5103	0.4385
NONE-UPW	0.4795	0.4036	0.4208
DOWN-UPW	0.3175	0.6834	0.4120
NONE	**0.7210**	0.2034	0.2892

Result 2

Which strategies are better for re-weighting crisis events of similar types when learning the volunteer/donation information?

There is no clear winner for sampling and weighting strategies. One could decide which strategies to use based on which metric is more important for their problem.

“UP” strategies outperform other weighting and sampling strategies on F1 while preserving relatively high precision and recall.

How similar is volunteerism ad recovery language across different events and event-types?

Which strategies are better for re-weighting crisis events of similar types when learning the volunteer/donation information?

What source of weakly supervised volunteerism data leads to the largest performance improvements in our models?

What source of external data leads to the largest performance improvements in our models?

NGO

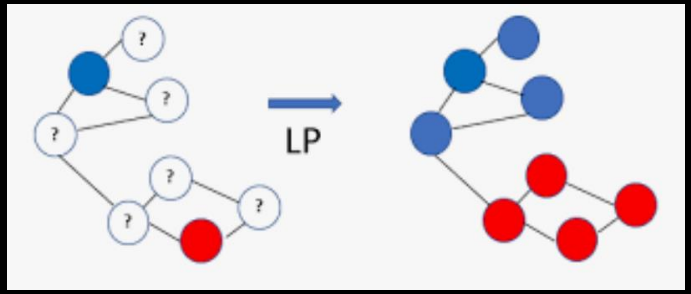
Unlabeled Crisis

59,734 tweets from unlabeled TREC-IS

Random

60k randomly sampled English tweets

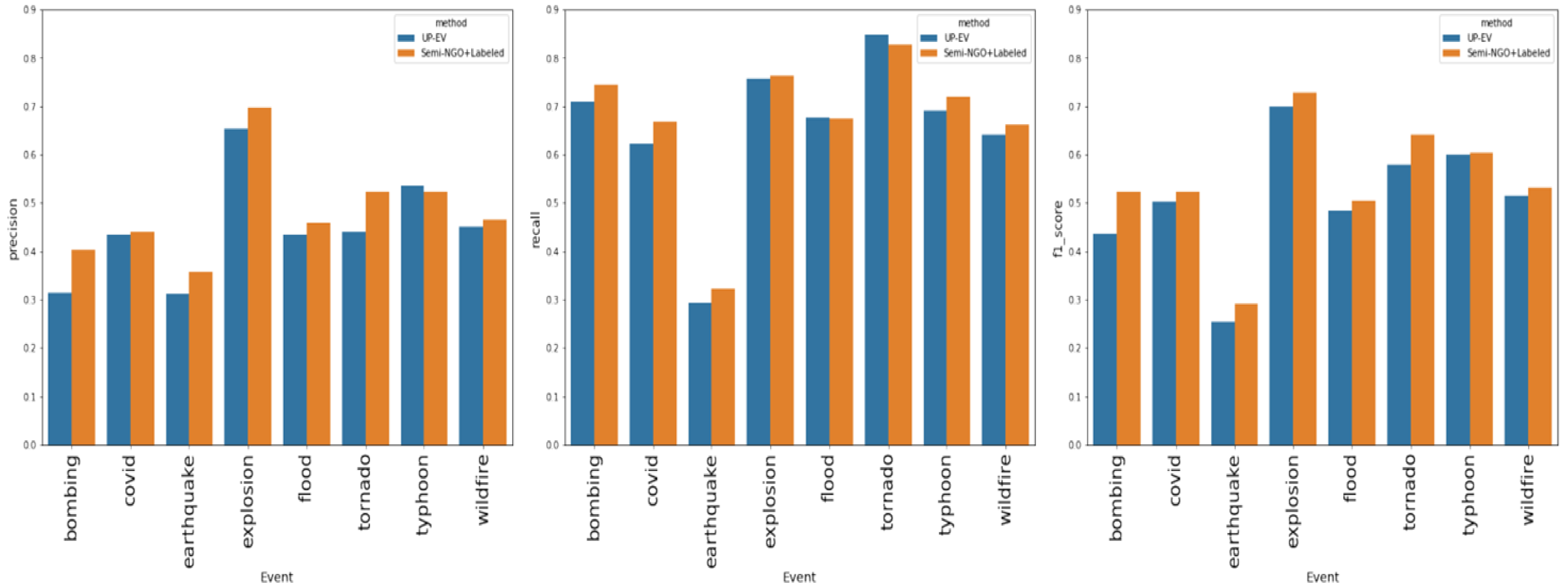
Weak-Supervision Methods

Labeled all as "volunteer"	Label-spreading	Semi-supervision
		

WEAK SUPERVISION FOR INTEGRATING DOMAIN EXPERTISE

method	precision	recall	f1_score
Semi-NGO+Labeled	**0.4847**	0.6700	**0.5430**
Semi-NGO+UnlabelledCrisis+Labeled	0.4762	0.6694	0.5365
Semi-UnlabelledCrisis+Labeled	0.4557	0.6779	0.5246
Semi_Random+Lableled	0.4454	0.6899	0.5206
UP-EV	0.4636	0.6484	0.5196
LS-NGO+Labeled	0.4358	0.69980	0.5168
LS-NGO+UnlabelledCrisis+Labeled	0.3897	**0.7312**	0.4848
LS-UnlabelledCrisis+Labeled	0.3893	0.7202	0.4806
All-volunteer-NGO+Labeled	0.3742	0.7073	0.4615

WEAK SUPERVISION FOR INTEGRATING DOMAIN EXPERTISE



(a) Precision (b) Recall (c) F1 Score

Result 3

What source of weakly supervised volunteerism data leads to the largest performance improvements in our models?

First, augmenting our training data with weakly supervised data increases overall performance.

Second, the study sheds light on the importance of integrating domain knowledge(e.g. NGOs) to identify additional sources of data to augment the initial dataset.

Despite the outpouring of social support posted to social media channels in the aftermath of disaster, finding and managing content that can translate into community relief, donations, volunteering, or other recovery support is difficult due to the lack of sufficient annotated data around volunteerism. This paper addresses these challenges by constructing a general machine learning model that is transferable from one crisis to the other and by introducing a method for integrating domain expertise into language classification to improve the classification accuracy.


Thank you! Questions?