Using Artificial Intelligence and Social Media Multimodal Data for Humanitarian Aid

Muhammad Imran, Ferda Ofli, Firoj Alam, and Umair Qazi

Qatar Computing Research Institute, HBKU, Doha, Qatar
{mimran,fofli,fialam,uqazi}@hbku.edu.qa

Extended Abstract

During disasters and emergencies, humanitarian organizations, government agencies, public health authorities, and military are tasked with responsibilities to help affected and vulnerable populations [4]. These formal response organizations rely on timely and credible information to make rapid decisions to launch relief operations. The information needs of these stakeholders vary depending on their role, responsibilities, and the circumstances that they deal with [10]. However, during time-critical situations, the importance of timely and factual information increases, especially when no information from other traditional sources such as TV or Radio are available [3]. The growing use of mobile technologies and social media platforms such as Twitter and Facebook has provided general public with practical and effective opportunities to disseminate and ingest information. Millions of people increasingly use social media to share information during natural or human-induced disasters [3, 5].

Making sense of social media data to help responders involves solving multifaceted challenges including parsing unstructured and brief content, filtering out irrelevant and noisy content, handling information overload, among others. Over the last six years, we have been working on a number of Artificial Intelligence (AI) techniques and computational methods to solve these big crisis data processing challenges. Specifically, we aim to build innovative technologies which can be used by decision-makers, NGOs, affected communities, and scholars to improve the effectiveness of humanitarian strategies such as preparedness, mitigation, and response during crises and emergencies.

Artificial Intelligence for Digital Response (AIDR) [7] is one of our AI platforms that collects and analyzes social media textual as well as imagery content in real time at the onset of a disaster to gain situational awareness and actionable information. We apply natural language processing, machine learning, and computer vision techniques to crisis information communicated via social media for humanitarian response. Most of these efforts are coordinated in collaboration with several regional and global partners including UN OCHA, and UNICEF.

On the analysis of textual data, we will present different techniques that were developed by the Crisis Computing team at QCRI to assist humanitarian organizations. These works include online Twitter stream classification using supervised machine learning techniques, active learning techniques to train ML models in real-time, and text summarization [9, 6]. Moreover, we will also talk about the applications of deep neural networks (DNNs) for the processing of crisis-related data. DNNs have shown great performance in classification tasks in NLP and data mining. However the success of DNNs on a task depends heavily on the availability of a large, labeled dataset which is not a feasible option in a crisis setting (i.e., no time to wait for labeling tweets at the onset of a disaster). We will present how we use a novel domain adaptation technique to overcome this labeled data scarcity issue [1].

In addition to textual data, images shared on social media at the onset of a disaster contain rich situational awareness information such as infrastructure damages, scenes of injured people, evacuations, shelter locations. In particular, damage assessment is one of the most challenging tasks that response organizations traditionally perform through field assessments, which are time consuming and costly. To tackle this challenge, we will present an image processing pipeline that mines images posted on Twitter in real-time to assess the overall severity of damage incurred by the disaster. This pipeline employs various DNNs to identify whether an image (i) is duplicate or not (i.e., to deal with the redundant content on Twitter), (ii) is relevant to the disaster or not (i.e., to eliminate junk content such as cartoons, banners, advertisements), and finally (iii) shows any infrastructure damage or not [2, 8]. Oftentimes, both textual and imagery data present complementary information which can be utilized further by a multimodal data analysis approach. We will present our ongoing work on multimodal learning for humanitarian aid.

References

1. Alam, F., Joty, S., Imran, M.: Domain adaptation with adversarial training and graph embeddings. In: ACL. pp. 120–128 (2018)
2. Alam, F., Ofli, F., Imran, M.: Processing social media images by combining human and machine computing during crises. IJHCI **34**(4), 311–327 (2018)
3. Castillo, C.: Big Crisis Data. Cambridge University Press (2016)
4. Gralla, E., Goentzel, J., Van de Walle, B.: Field-based decision makers information needs. Digital Humanitarian Network, Geneva (2013)
5. Hughes, A.L., Palen, L.: Twitter adoption and use in mass convergence and emergency events. IJEM **6**(3-4), 248–260 (2009)
6. Imran, M., Castillo, C., Lucas, J., Meier, P., Rogstadius, J.: Coordinating human and machine intelligence to classify microblog communications in crises. In: ISCRAM (2014)
7. Imran, M., Castillo, C., Lucas, J., Meier, P., Vieweg, S.: AIDR: Artificial intelligence for disaster response. In: WWW. pp. 159–162 (2014)
8. Nguyen, D.T., Ofli, F., Imran, M., Mitra, P.: Damage assessment from social media imagery data during disasters. In: ASONAM. pp. 569–576 (2017)
9. Rudra, K., Goyal, P., Ganguly, N., Mitra, P., Imran, M.: Identifying sub-events and summarizing disaster-related information from microblogs. In: SIGIR. pp. 265–274 (2018)
10. Vieweg, S., Castillo, C., Imran, M.: Integrating social media communications into the rapid assessment of sudden onset disasters. In: SocInfo. pp. 444–461 (2014)