Natural Hazards
and Earth System
Sciences

Open Access

Discussions

EGU

NHESSD

**Interactive
comment**

Interactive comment on "Review article: Detection of informative tweets in crisis events" *by* Anna Kruspe et al.

Anonymous Referee #1

Received and published: 11 August 2020

General Comments This is a review paper summarizing at a high level currently available datasets and approaches that tackle crisis analytics on Twitter data, as well as highlights some challenges. The paper is overall well written and should be understandable by an in-expert audience. It may be valuable as an entry point for new researchers looking to work in the area of crisis informatics (although there are already good resources available in this area (e.g. the Big Crisis Data book). On the other hand, the paper is quite shallow in terms of detail in all aspects and so only acts as a guidepost for further reading on the subject. There are no individual experiments by the authors provided in the paper.

Detailed Comments The core limitation of this work from my reading is that it tries to summarize too many areas of the field of crisis informatics and so currently does not

provide enough detail on any one to provide significant insights that add value over the individual works. For example, Section 2 provides a brief summary of some of the definitions that different groups have used for analysing crisis content, but don't go into detail on what works are compatible with one another, or indeed provide information on the definition of 'Informative' regards to who is used in each work. As a result it is not clear to someone just entering the field what they should read. Similarly, Section 3 highlights some datasets used in crisis informatics, but Table 2 only lists tweet counts, not the volume labelled and what was labelled or for what task. Which datasets are complementary? Which are easy to work with? What datasets do the authors recommend researchers use?

The second question I have for this overview is whether COVID-19 datasets belong in this study or whether they should be considered separately. COVID is quite a different task compared to natural or manmade disasters, as there typically is not a strong timeliness component to related information needs. Hence, the definition of what is informative for pandemics and the associated target user groups are very different. I would recommend at least adding some discussion in Section 2 on this point.

Third, I would recommend structuring the discussion on the machine learning aspects along the lines of what task is being investigated, inputs, features and models. Indeed, it would be valuable to get some idea of how many works use each different approach, as well as get some data on the prevalence of different feature and text representation approaches used in the different works and critically, what patterns emerge on what works.

Other notes: - Table 2 should highlight the differences between labelled and un-labelled tweets - I believe the statistics for the TREC-IS data in particular is out-of-date, see the ISCRAM 2020 paper http://trecis.org/2020/ISCRAM_2020_TREC_IS.pdf

---