



Automatic detection of informative tweets during disasters

Anna Kruspe, Jens Kersten, and Friederike Klan
German Aerospace Center (DLR)

Messages on social media can be an important source of information during a disaster. They can frequently provide details about developments much faster than traditional sources (e.g. official news) and can offer personal perspectives on events, such as opinions or specific needs. In the future, these messages can also serve to assess disaster risks.

One challenge for utilizing social media in disaster situations is the detection of informative messages in a flood of data. Researchers have started to look into this problem in recent years, beginning with crowd-sourced methods. Lately, approaches have shifted towards an automatic analysis of messages.

In this study, we present methods for the automatic detection of crisis-related messages (tweets) on Twitter. We start by showing the varying definitions of importance and relevance relating to disasters, as they can serve very different purposes. This is followed by an overview of existing data sets.

We then compare approaches for solving this problem based on Machine Learning techniques with regard to their focus, their data requirements, their technical prerequisites, their efficiency and accuracy, and their time scales. These factors determine the suitability of the approaches for different expectations, but also their limitations. We identify which aspects each of them can contribute to the detection of informative tweets, and which areas can be improved upon in the future.

We point out particular challenges, such as the linguistic issues concerning this kind of data. Finally, we suggest future avenues of research, and show connections to related tasks, such as the subsequent semantic classification of tweets.