

The Flood-Related Multimedia Task at MediaEval 2020

Stelios Andreadis¹, Ilias Gialampoukidis¹, Anastasios Karakostas¹, Stefanos Vrochidis¹,
Ioannis Kompatsiaris¹, Roberto Fiorin², Daniele Norbiato², Michele Ferri²

¹Information Technologies Institute - Centre of Research and Technology Hellas, Greece

²Eastern Alps River Basin District, Italy

{andreadisst,heliasgj,akarakos,stefanos,ikom}@iti.gr

{roberto.fiorin,daniele.norbiato,michele.ferri}@distrettoalpiorientali.it

ABSTRACT

This paper provides a description of the Flood-related Multimedia Task at MediaEval 2020. The primary goal of the task is to analyse and combine textual and visual content from social media data that reflect real-world events. The focus is on natural disasters and especially on flooding incidents, which are frequent around the globe and have large social consequences for communities and individuals. In particular, the task requires participants to identify Twitter posts that are relevant to flood events in a specific area of interest, based on their text and images. The automatic classification of posts as relevant or not relevant will essentially improve the quality of retrieved social media data, so that they can play a more valuable role in the emergency management.

1 INTRODUCTION

The worldwide dominance of social media in the daily life of modern people has led to a vast amount of available crowd-sourced information. The large streams of public social media data have created new directions in the research community and their analysis has the potential to affect positively several domains. One domain that can significantly benefit is natural disaster management where the exploitation of social media data is able to assist in every stage of a hazard event: a) they can notify about a possible disaster in the pre-emergency phase; b) they can provide insights on the evolution of the incident and detect regions in danger during the disaster; and c) they can assist in the damage control in the post-emergency phase. To this end, the Flood-related Multimedia Task at MediaEval 2020 focuses on improving situational awareness for flooding incidents in specified areas of interest.

However, the acquisition of social media data about a particular natural disaster, such as floods in our scope, raises a number of challenges. First, keyword-based search can result to social media posts that contain keywords related to floods, but their content is in fact irrelevant; for example, when a word is used metaphorically. Another challenge is that posts often consist of both textual and visual information (i.e. message and attached image) that may have different relevance to the disaster. Finally, a third challenge is to filter out posts that refer to past events or incidents occurring outside the area of interest.

Towards addressing these challenges, several works aim at estimating the relevance of social media content to natural disaster

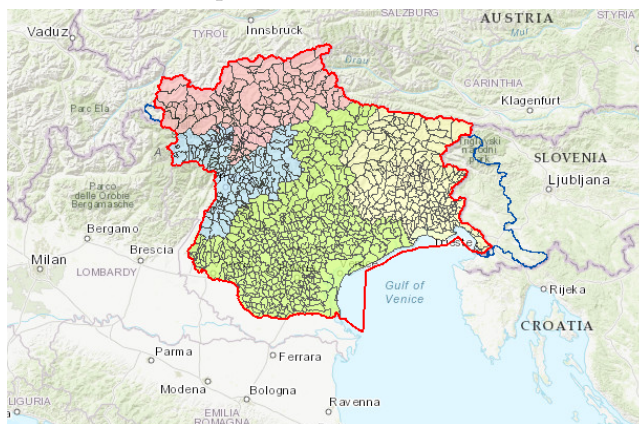


Figure 1: The area of interest - Eastern Alps river basin district in North-East Italy

incidents. The work in [7] proves that the proximity of disaster-related georeferenced social media images to a disaster event can be a significant indicator of relevance. The authors of [4] apply machine learning techniques to classify informative versus non-informative tweets posted during an earthquake, while [6] proposes an approach to decide whether a tweet is relevant to flooding by combining text classification of the posted message and visual classification of the attached image. Moreover, [5] investigates the usefulness of annotated social media data from a prior disaster together with unlabeled data from a current disaster so as to learn domain adaptation classifiers.

Following the Multimedia Satellite Task that was introduced in 2017 [3] and continued in years 2018 [1] and 2019 [2], the Flood-related Multimedia Task remains focused on the high-impact natural disaster of floods, but this year solely through the perspective of social media data. The overall goal of the task is to tackle the aforementioned challenges and combine text and images from social media streams in order to identify posts about floods in a predefined area. Moreover, the task poses an additional challenge by involving Italian social media posts in order to encourage researchers to move away from a focus on English. Better ability to separate relevant and not relevant tweets will contribute to improving the quality of the incoming information available to support first responders and civil protection authorities.



#allertameteo in tutto il paese, a #Venezia
nuovo picco di acqua alta
askanews.it/cronaca/2019/1...



3:16 PM · Nov 17, 2019 · Twitter Web App

Figure 2: An example of a tweet that is relevant to floods in NE Italy

2 TASK DESCRIPTION

The Flood-related Multimedia Task tackles the analysis of social multimedia from Twitter for flooding events. In this task, the participants receive a set of Twitter posts (tweets) and their associated images, which contain keywords related to floods in a specific area of interest, specifically, the Eastern Alps district in Northeastern (NE) Italy (Fig. 1). However, the relevance of the tweets to actual flooding incidents in that area is ambiguous.

The objective of the task is to build an information retrieval system or a classifier that is able to distinguish whether or not a tweet is relevant to a flooding event in the examined area. An example of a relevant tweet can be seen in Fig. 2. The dataset of the task consists of Italian-language tweets, motivated by the common flood events in the cities of Eastern Alps (e.g., Venice, Vicenza, Trieste, Padua, Pordenone) and surrounding areas. Participants can tackle the task using text features, image features, and a combination of both, and are allowed to submit 5 runs:

- Required run 1: automated, using a fusion of textual and visual data
- Optional run 2: automated, using textual information only
- Optional run 3: automated, using visual information only
- General run 4,5: everything automated allowed, including using data from external sources

3 DATASET DESCRIPTION

The dataset of the task consists of 5,419 Tweet IDs (development-set) and 2,288 Tweet IDs (test-set) that have been collected from Twitter between 2017 and 2019, by searching for Italian flood-related

Table 1: Keywords used in the acquisition of tweets

Italian keywords	Translation
alluvione	flood
alluvione vicenza	flood Vicenza
allagamento	flooding
bacchiglione	Bacchiglione (river)
fiume piena	full river
allerta meteo	weather alert
sottopasso allagato	underpass flooded
allerta meteo vicenza	weather alert Vicenza
esonazione	flood
livello fiume	river level

keywords inside the tweet text. The complete list of keywords can be seen in Table 1. All tweets contain an attached image and had to be still online at the time of releasing the dataset. In order to be compliant with the Twitter Developer Policy, only the IDs of the tweets are distributed to the participants.

The ground truth data of the dataset consists of one class label for the relevancy of each Tweet ID and has been collected with human annotation. Each tweet has been annotated by a single person and the annotators work for the Eastern Alps River Basin District, who are experts on flood risk management in the Eastern Alps district of NE Italy. Apart from the fact that tweets were in their native language, their expertise allowed them to annotate as relevant tweets that referred to specific events, known a priori to them as real, in their district and to other phenomena that may be indirectly related to flooding.

Ground truth is provided to participants only for the development-set as key-value pairs of Tweet ID and ground truth label for the relevancy (0=not relevant/ 1=relevant), where out of the 5,419 tweets 1,140 (21%) are relevant and 4,279 (79%) are not relevant. On the other hand, ground truth for the test-set is not provided to the participants, since it is used in the evaluation.

4 EVALUATION

The official metric for evaluating the correctness of retrieved tweets for the two classes relevant (1) and not relevant (0) is the F1-Score metric on the test set. F1-Score has been selected, because it is defined as the harmonic mean between precision and recall.

ACKNOWLEDGMENTS

This work has been supported by the EU's Horizon 2020 research and innovation programme under grant agreements H2020-776019 EOPEN and H2020-832876 aqua3S.

REFERENCES

- [1] Bischke Benjamin, Helber Patrick, Zhao Zhengyu, Borth Damian, and others. 2018. The Multimedia Satellite Task at MediaEval 2018: Emergency response for flooding events. (2018).
- [2] Benjamin Bischke, Patrick Helber, Simon Brugman, Erkan Basar, Zhengyu Zhao, Martha Larson, and Konstantin Pogorelov. The Multimedia Satellite Task at MediaEval 2019: Estimation of Flood Severity.

- In *Proc. of the MediaEval 2019 Workshop* (Oct. 27-29, 2019). Sophia Antipolis, France.
- [3] Benjamin Bischke, Patrick Helber, Christian Schulze, Venkat Srinivasan, Andreas Dengel, and Damian Borth. 2017. The Multimedia Satellite Task at MediaEval 2017.. In *MediaEval*.
 - [4] Muhammad Imran, Carlos Castillo, Ji Lucas, Patrick Meier, and Sarah Vieweg. 2014. AIDR: Artificial intelligence for disaster response. In *Proceedings of the 23rd International Conference on World Wide Web*. 159–162.
 - [5] Hongmin Li, Nicolais Guevara, Nic Herndon, Doina Caragea, Kishore Neppalli, Cornelia Caragea, Anna Cinzia Squicciarini, and Andrea H Tapia. 2015. Twitter Mining for Disaster Response: A Domain Adaptation Approach.. In *ISCRAM*.
 - [6] Anastasia Moutzidou, Stelios Andreadis, Ilias Gialampoukidis, Anastasios Karakostas, Stefanos Vrochidis, and Ioannis Kompatsiaris. 2018. Flood relevance estimation from visual and textual content in social media streams. In *Companion Proceedings of the The Web Conference 2018*. 1621–1627.
 - [7] Robin Peters and João Porto de Albuquerque. 2015. Investigating images as indicators for relevant social media messages in disaster management.. In *ISCRAM*.