

# Detecting the Need for Resources and their Availability

Nelleke Oostdijk\*

Centre for Language Studies (CLS), Radboud University  
Erasmusplein 1  
Nijmegen 6525 HT, The Netherlands  
n.oostdijk@let.ru.nl

Ali Hürriyetoglu

CLS, Radboud University, Nijmegen and Statistics  
Netherlands  
CBS-Weg 11  
Heerlen, The Netherlands  
a.hurriyetoglu@cbs.nl

## ABSTRACT

In this working note we describe our submission to the FIRE2017 IRMiDiS track. We participated in both sub-tasks, the first of which was directed at identifying the need or availability of specific resources and the second at matching tweets expressing the need for a resource with tweets mentioning their availability. Our linguistically motivated approach using pattern matching of word n-grams achieved an overall average MAP score of 0.2458 for sub-task (1), outperforming our machine-learning approach (MAP 0.1739) while being surpassed by two other (automatic) systems. The linguistic approach was also used in sub-task (2). There it was the best-performing approach with an f-score of 0.3793.

## CCS CONCEPTS

•**Computing methodologies** → **Information extraction**; *Natural language processing*; *Support vector machines*; •**Information systems** → *Content ranking*; *Social tagging*; *Social networks*; *Content analysis and feature selection*; *Information extraction*; *Expert search*; *Clustering and classification*; *Web and social media search*;

## 1 INTRODUCTION

The FIRE2017 IRMiDiS track[1] was directed at microblogs and the aim was to identify actionable information such as what resources are needed or available during a disaster. In this track there were two sub-tasks: (1) identifying need tweets and availability tweets and (2) matching need tweets and availability tweets. The dataset consisted of tweets posted during the Nepal 2015 earthquake.

Our submission consisted of the results of two runs each using a different semi-automatic approach addressing sub-task (1), and one run with one of our approaches addressing sub-task (2).

The structure of the text below is as follows: The data are described in more detail in Section 2. In Section 3 we describe the approaches used to address subtask (1) and the results obtained. In Section 4 the approach used for subtask 2 is described as well as the results obtained. Section 5 summarizes the main findings and includes suggestions for future work.

## 2 DATA

The data for this track were tweets posted during the 2015 Nepal Earthquake and include tweets in English as well as tweets in local languages such as Hindi and Nepali and code-mixed tweets. The data were provided by the organizers in the form of tweet IDs, together with a script for downloading them.

\*This is the corresponding author

Table 1: Composition Test Dataset

Language tag	Tweet count	Tweet %
English	35,682	79.72
Hindi	3,060	6.84
Nepali	4,180	9.34
Other <sup>a</sup>	1,837	4.10
All	44,759	100.00

<sup>a</sup>Other is used to refer to a wide range of languages/language tags as well as the UND tag which occurred with 976 tweets.

Participants were given two datasets which they could use to develop their methods. The first set (development data) included roughly 18,000 tweets without any class labels. The second set (train data) contained some 800 tweets with class labels identifying them as need tweet (211) or availability tweet (718). The test data contained 46,920 tweets of which we managed to download 44,759 tweets. The composition of the dataset we downloaded is shown in Table 1.

## 3 SUB-TASK (1): IDENTIFYING NEED AND AVAILABILITY

For sub-task (1) participants were expected to develop methodologies for identifying need tweets and availability tweets. The following descriptions were provided: <sup>1</sup>

**Need-tweets:** Tweets which inform about the need or requirement of some specific resource such as food, water, medical aid, shelter, mobile or internet connectivity, etc.

**Availability-tweets:** Tweets which inform about the availability of some specific resources. This class includes both tweets which inform about potential availability such as resources being transported or dispatched to the disaster-struck area, as well as tweets informing about the actual availability in the disaster-struck area, such as food being distributed, etc.

### 3.1 The Linguistic Approach

The approach we used here was based on that applied on a previous occasion.[3] In this approach a (monolingual English) lexicon and a set of hand-crafted rules are used to tag the relevant n-grams.

<sup>1</sup><https://sites.google.com/site/irmidisfire2017/>

The tagged output is then used for assigning class labels. As for the current track approaches were expected to be able to handle also non-English data, a pre-processing step was introduced for dealing with these data. In order to produce the results in a ranked fashion, where this is not part of the approach as such, we used some heuristics. More details on each of the steps outlined above are given below.

*Pre-processing.* As a pre-processing step we translated part of the non-English tweets to English by means of Google Translate (<https://translate.google.com>), that is, only tweets that carried the (Twitter) language tag HI or NE were translated. All other non-English tweets (again, according to the language tag provided) were left untouched, as were all tweets that the language tag identified as English where in fact they were either mixed tweets (English and some other language, possibly Hindi or Nepali) or a completely different language.

An example of a tweet where Twitter identified the language as Hindi is given below (Example 1). Its translation as produced by Google Translate is shown in Example 1a. As neither of the authors has any knowledge of Hindi or Nepali, we had no way of knowing to what extent the language tag and the translation were reliable and we simply had to take the translation for what it is worth. However, based on what we saw with some other tweets in languages that we do know, we could tell that the language tag was sometimes completely off. Yet, we decided that we were not going to spend any time on this issue within the scope of the present task.

[Example 1]

- नेपाल में भीषण भूकम्प से व्यापक तबाही के लगभग 72 घंटे बाद भी लोग भोजन, पानी, दवाईयों और बिजली की कमी से जूझ रहे...  
<http://t.co/VFIELVXFZL>

[Example 1a]

- About 72 hours after the devastating earthquake in Nepal, people are still struggling with food, water, medicine and lack of electricity ...  
<http://t.co/VFIELVXFZL>

*Tagging relevant n-grams.* A lexicon was constructed containing some 1,400 items. The lexicon contains mostly unigrams and bigrams that are considered relevant for the task at hand. All lexical entries are typed. Main types are V(erb)-like or N(oun)-like. V-like items typically are verb forms (e.g. *distribute*) and nominalizations (e.g. *distribution*) which express some action. With each V-like item a tag is associated which indicates whether the items expresses need or availability (tags act-n and act-a respectively). N-like items are typically nouns (e.g. *food*). With each N-like item a tag is associated which indicates whether the item is a specific (e.g. *water*) or more general resource (e.g. *aid*). Tags here are res-a and res. In addition to the lexicon there is a small rule set (currently comprising 10 rules) which specifies how lexical items may combine to form multiword n-grams. The lexicon and the rules are used by a software module we have developed in another project. The module assigns the appropriate tags to the relevant word n-grams.

**Table 2: Evaluation Results Sub-Task 1 for Availability Tweets**

Submission Run ID	Availability Tweets		
	Prec.@100	Recall@100	MAP
Radboud_CLS_task1_1	.7300	.3153	.2062
Radboud_CLS_task1_2	.6100	.2224	.1660

**Table 3: Evaluation Results Sub-Task 1 for Need Tweets**

Submission Run ID	Need Tweets		
	Prec.@100	Recall@100	MAP
Radboud_CLS_task1_1	.7500	.4309	.2853
Radboud_CLS_task1_2	.4900	.3934	.1812

Example 2 shows the result of the tagging of the (translated) tweet shown in Example 1a.

[Example 2]

food [res-a] water [res-a] medicine [res-a] lack of  
[act-n] electricity [res-a]

*Assigning class labels.* The labeling module we developed automatically assigns a class label ('Nepal-Need' or 'Nepal-Avail') to each tweet. Input for this module is the output of the tagging described above. For each tweet, the tags assigned are unique, that is a single instance of each tag type is maintained. Example 3 shows the result of this step for our example tweet. The module makes use of the label pattern list which specifies which labels are to be associated with the (combinations of) tags that can occur in a tweet (the tag patterns). For our example this means that given the tag pattern [act-n][res-a] the label 'Nepal-Need' is assigned (Example 3a). A typical pattern for a tweet expressing the availability of a resource is [act-a] [res-a].

[Example 3]

[act-n] [res-a]

[Example 3a]

[act-n] [res-a] Nepal-Need

*Ranking the output.* As this method does not yield any confidence or likelihood scores, a ranking was obtained in the following manner. The output was first ranked based on a human-estimated confidence of specific class label + tag pattern combinations. This resulted in an initial ranking of the sets of tweets that showed a particular tag pattern. The final ranking was obtained by ordering the tweets within these ranked sets according to their tweet ID.

**Results.** The results as evaluated by the organizers are shown in Tables 2 and 3 as run ID Radboud\_CLS\_task1\_1. The results were the best of all submissions using a semi-automatic approach on all counts (Precision@100, Recall@100, and MAP, both for the need tweets and the availability tweets). The average MAP of the Availability and Necessity tweets are 0.2458 and 0.1736 for Radboud\_CLS\_task1\_1 and Radboud\_CLS\_task1\_2 respectively.

### 3.2 The Relevancer Approach

For sub-task (1) we also applied the Relevancer approach[2]. Relevancer was used to generate 194 clusters for the tweets tagged as English or Hindi. English clusters, which are one third of the clusters, were annotated and used as training data for the support vector machines (SVM) based classifier. The cluster annotation yielded 272 availability and 38 need tweets. The training data was extended with additional data from [3], the gold annotations released by the organization team, and the development data released in the scope of this shared task. The final classifier was used to predict label of the test tweets. The classifier confidence was used to rank the results.

**Results.** The results as evaluated by the organizers are shown in Table 2 as run ID Radboud\_CLS\_task1\_2.

### 4 SUB-TASK (2): MATCHING NEED AND AVAILABILITY

In sub-task (2) participants were required to develop methodologies for matching need tweets with appropriate availability tweets. For this task we used the tagging output we had obtained in the processing the tweets for sub-task 1 using the linguistic approach. For every need tweet all the word n-grams that had been tagged as identifying a resource, we would attempt to find an exact match in the availability tweets. In both cases (need and availability tweets), the ranked list was used and the software program would work its way down. Since the task was to find up to 5 availability tweets with each need tweet and the algorithm would always start at the top of the list of ranked availability tweets, only a small portion of the availability tweets actually appears in the matching results. Only when no exact match could be found, would the software attempt to find near-matches. This was typically the case for tweets where at best tweets could be found that yielded a partial match (e.g. matching 2 out of 4 requested resources).

To illustrate the approach described above, let us get back to the example we have been using throughout this paper. Our example tweet was identified as a need tweet. From the tagging we obtained (shown in Example 2) we would only keep those word n-grams and their tags that identified a resource. The result of this is shown in Example 2b.

[Example 2b]

food [res-a] water [res-a] medicine [res-a] electricity [res-a]

In order to find matching availability tweets we would look for tweets where the same word n-grams could be found, regardless of the order in which they occurred. The five matching tweets found amongst the highest ranking availability tweets in the case of our example were those shown in Examples 4-8.<sup>2</sup>

[Example 4]

RT @aolnepal: Anybody interested in donating food, medicine, water and sanitation materials for earthquake victims, contact Art... <http://t...>

<sup>2</sup>In each case the word n-grams identified and tagged as resource are shown below the tweet text.

food [res-a] medicine [res-a] water [res-a] sanitation [res-a] sanitation materials [res-a] materials [res]

[Example 5]

Pakistan distributing Beef, Missionaries distributing Bible @RSSorg @bst\_official @ArtofLiving distributing Food+Water #Earthquake

food [res-a] water [res-a]

[Example 6]<sup>3</sup>

Soldiers of Jagadal Battalion of NA distributing food and water to earthquake victims. नेपाली सेनाको जगदल गणका... <http://t.co/HJLs1MpS9N>

food [res-a] water [res-a]

[Example 7]<sup>4</sup>

RT @abpnewshindi: विमान में खाना, पानी और कंबल नेपाल के लिए भेजे गए हैं . एस. जयशंकर \#NepalEarthquake लाइव देखें- <http://t.co/MG3hLqR5bO>

[Example 7a]

RT @abpnewshindi: Food, water and blankets are sent to Nepal in the aircraft. s. See Jaishankar #NepalEarthquake Live- <http://t.co/MG3hLqR5bO>

food [res-a] water [res-a] blankets [res-a]

[Example 8]<sup>5</sup>

• About 72 hours after the devastating earthquake in Nepal, people are still struggling with food, water, medicine and lack of electricity ... <http://t.co/VFIELVXFZL>

[Example 8a]

Under the service of religion, Christian missionaries are distributing the Bible in food, water, and clothing in Nepal. <Http://t.co/4E6IHcEqM4> via @the-lapine

clothing [res-a] food [res-a] water [res-a]

**Results.**

The results were submitted under run ID Radboud\_CLS\_task2\_1 and evaluated by the organizers as follows: precision@5 0.3305, recall 0.4450 and f-score 0.3793. Thus this approach was found to outperform all other approaches.

## 5 DISCUSSION AND CONCLUSIONS

Our participation in the IRMiDiS track was rather successful: we achieved a third and sixth place in the overall (MAP) ranking for

<sup>3</sup>The language tag for this tweet was 'und', therefore no preprocessing was applied.

<sup>4</sup>The language tag here was for Hindi. Example 7a shows the English translation obtained from Google Translate.

<sup>5</sup>The language tag here was for Hindi. Example 8a shows the English translation obtained from Google Translate.

**Table 4: Distribution across Languages**

Language Tag	Test set	Availability		Need	
	Count	Count	%	Count	%
English	35,682	2,913	8.16	1,235	3.46
Hindi	3,060	252	8.24	55	1.80
Nepali	4,180	254	6.08	78	1.87
Other	1,837	39	2.12	8	0.44
All	44,759	3,458	7.73	1,376	3.07

sub-task (1), and scored best of all approaches on sub-task (2).<sup>6</sup> Our success can in part be attributed to the experience gained through our participation in previous shared tasks.

However, there are some issues that we struggled with and where we expect we might do better on a future occasion. Thus we found that for the current task the definitions of what constitutes a need tweet and what an availability tweet were somewhat unclear, more particularly in specifying what exactly was meant by ‘a specific resource’. The examples in the task description were all clear-cut cases, including food, water, medicine, electricity, and blood donors. But what about donations or support which is what we also encounter in the data.

Looking at the results obtained through our linguistic approach for sub-task (1), we note that the distribution of the need and availability tweets over the various languages is rather uneven. While on average 7.73% of the data is classified as availability tweet and 3.07% as need tweet, the ratios especially for other language tweets are much lower (Table 4). We speculate that the fact that the data comprised multiple languages has affected our recall obtained for sub-task (1). While the use of Google Translate for tweets tagged as Hindi or Nepali was reasonably effective, we expect that better results could be achieved if we put more effort into the preprocessing of the tweets. This would involve both improving the language identification and finding a way to handle code-mixed tweets.

The results obtained with the linguistic approach for sub-task (1) might have been better if we had allowed for multiple class labels to be associated with a given tweet. However, as we expected sub-task (2) to be easier if a tweet carried only a single label, we opted for a forced choice for one of the two classes and we ignored tweets where we could not decide for either class.

In matching need and availability of resources for sub-task (2) we restricted ourselves to only exact (literal) matches which means that we fail to match instances such as the need for shelter and the availability of tents, or the need for food and the availability of packed meals. In future work we might include synonyms, hypernyms, and hyponyms.

The Relevancer approach suffered from the fact that the training set of labelled data was rather small (929 tweets) and quite unbalanced (211 need vs 718 availability tweets). However, we think that a combination of the linguistic approach and Relevancer approach has the potential to overcome such limitations. We are currently conducting experiments in which we aim to combine the strengths of the two approaches. So stay tuned!

<sup>6</sup>The organizers distinguish between automatic and semi-automatic approaches without specifying how to discriminate between them. We consider our approaches semi-automatic but actually in their execution they operate automatically.

## ACKNOWLEDGMENTS

The authors are grateful to Peter Beinema for his help with the software used in sub-task (2).

## REFERENCES

- [1] Moumita Basu, Saptarshi Ghosh, Kripabandhu Ghosh, and Monojit Choudhury. 2017. Overview of the FIRE 2017 track: Information Retrieval from Microblogs during Disasters (IRMiDis). In *Working notes of FIRE 2017 - Forum for Information Retrieval Evaluation (CEUR Workshop Proceedings)*. CEUR-WS.org.
- [2] Ali Hürriyetoğlu, Nelleke Oostdijk, Mustafa Erkan Başar, and Antal van den Bosch. 2017. *Supporting Experts to Handle Tweet Collections About Significant Events*. Springer International Publishing, Cham, 138–141. [https://doi.org/10.1007/978-3-319-59569-6\\_14](https://doi.org/10.1007/978-3-319-59569-6_14)
- [3] Ali Hürriyetoğlu and Nelleke Oostdijk. 2017. Extracting Humanitarian Information from Tweets. In *Proceedings of the First International Workshop on Exploitation of Social Media for Emergency Relief and Preparedness*. Aberdeen, United Kingdom. <http://ceur-ws.org/Vol-1832/SMERP-2017-DC-RU-Retrieval.pdf>