DAIICT-Hildesheim @ Information Retrieval from Microblogs during Disasters (IRMiDis 2018)

Noushin Fadaei[1], Chanjong Im[1], Sandip Modha[2], and Thomas Mandl[1]

[1] Hildesheim University, Hildesheim, Germany
{fadaei,imchan, mandl}@uni-hildesheim.de
[2] Dhirubhai Ambani- Institue of Information and Communication Technology, Gandhinagar, India
sjmodha@gmail.com

1 Introduction

Twitter as one of the most used sources for quickly spreading news covers a range of necessary information particularly in emergency cases. However these situations are inevitably come along with tweets expressing the sympathy of people, rumours and claims that are not very clearly referenced. Operating the disaster relief properly and rapidly the information processing and trustworthiness of the tweets are required to be examined. The purpose of Information Retrieval from Microblogs during Disasters (IRMiDis) [2] track of FIRE 2018 is to distinguish between fact-checkable and non fact-checkable tweets to ease this process. Our group has participated in this track with three runs for task 1.

2 Methodology

2.1 run 1: DAIICT-Hildesheim-mod1-sif and run 2: DAIICT-Hildesheim-mod1-nosif

Run 1 and 2 are fully automatic. Both runs are applied on the semantically *neutral* tweets. The semantic label of the tweets is obtained through Stanford semantic analysis library (based on *Recursive Neural Network*)[3].

The following steps are then carried out on semantically *neutral* tweets assuming they potentially cover less emotional and more informative content in terms of fact-checkability:

1. Elimination of the tweets containing sentiment factors
2. Text Preprocessing
   (a) Tweets that are not part of utf-8 encoding are removed from the text.
   (b) **@string_value**, **RT**, **(,)** and **Urls** are removed
   (c) Pattern with more than two . or ! is replaced with ... or !!! respectively.
3. Word embedding creation

   (a) Train the model with the Nepal-Earthquake tweets. (dimension of 300)
   (b) If the term co-occur in both google-news data[3] and our trained model, replace the respective word vector in our trained model to google-news word vector.
4. Sentence vector creation. (This part is where it differs between run1 and run2)
   (a) For run 1:
      i. average (sum and divide by sentence length) all the word vectors appearing in the sentence.
      ii. Multiply with the first principle component with each sentence vector.
   (b) For run 2:
      i. Follows the methods from [1] for making the sentence vectors.
5. Similarity computation
   (a) We compute cosine similarity between fact-checkable tweets (83 tweets) and the rest.
   (b) Each sentence vectors from the set which does not include fact-checkable tweets will have 83 score values. Among 83 score values, maximum value is selected to represent the fact score. Based on the list of fact score the ranking is decided.

2.2 Run 3: DAIICT-Hildesheim-mod3-conv

This is a weakly supervised approach for objectivity detection. In this run, we anticipated this problem as an objectivity detection problem. We approached this problem as a 2-class classification problem Objective and subjective. There is only 83 Fact checkable tweets given by organizers. Total no of tweets in the dataset is more than 50000.

Preparation of Training Data We randomly choose 100 tweets from the dataset and labeled as a subjective or non-fact-checkable tweet and 83 fact-checkable tweets provided by track organizer labeled as objective. We train our Convolution neural network on these training data and tested the model on remaining 50000 tweets. At this stage we are not interested in the class but, we have sorted all the tweets based upon the predicted probability of the objective-class. Then after, we selected top 2000 tweets with highest predicted probability. We have randomly selected tweets and gave relevance judgment based upon objectivity on 1000 tweets and manually extracted 300 tweets as subjective which may minimize the false positives. Remaining 1700 tweets labeled as objective or fact-checkable tweets. We selected last 1700 tweets with the least probability of the class objective and labeled them as subjective. So our Training corpus has 1700 objective and 2000 subjective tweets.

[3] Available at: https://drive.google.com/file/d/0B7XkCwpI5KDYNlNUTTlSS21pQmM/edit

Method Glove Pre-trained vector with 300 dimension is used to initialize weight matrix of the embedding layer.. We trained our Convolution neural network on this training corpus with 10-fold cross validation. Model gives validation accuracy around 94%. Finally, we run the model on the entire corpus and sorted the tweet based upon the predicted probability of the objective- class. The results of all the methods are shown in Table 1.

Table 1. Evaluation results of sub-task 1- identifying fact-checkable tweets by IRMiDis FIRE 2018 [2]

Run*	Run Type	P@100	Recall@100	MAP@100	MAP	NDCG@100	NDCG
1	Auto.	0.1500	0.0670	0.0014	0.0139	0.0930	0.1330
2	Auto.	0.0100	0.0670	0.0000	0.0104	0.0033	0.1271
3	Semi-Auto.	0.4000	0.2002	0.0129	0.1471	0.4021	0.7492

* **The Run IDs are named DAIICT-Hildesheim-mod**i **for the** i^{th} **run**.

References

1. Arora, S., Liang, Y., Ma, T.: A simple but tough-to-beat baseline for sentence embeddings (2017)
2. Basu, M., Ghosh, S., Ghosh, K.: Overview of the FIRE 2018 track: Information Retrieval from Microblogs during Disasters (IRMiDis). In: Proceedings of FIRE 2018 - Forum for Information Retrieval Evaluation (December 2018)
3. Socher, R., Perelygin, A., Wu, J.Y., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C.: Recursive deep models for semantic compositionality over a sentiment treebank