Conceptual Representation for Crisis-Related Tweet Classification

Won-Gyu Choi and Kyung-Soon Lee[1]

Division of Computer Science and Engineering, CAIIT
Chonbuk National University, Republic of Korea
{smtj119, selfsolee}@chonbuk.ac.kr

Abstract. The importance of social media such as Twitter, as a conduit for actionable and tactical information during disasters is increasingly recognized. During crisis situations, rapid and effective response actions by emergency services are critical to assure the safety of the public. In this paper, we propose a conceptual representation for crisis-related tweet classification. In order to classify a stream of tweets related to the incident, the crisis-related terms in each tweet are represented as conceptual entities such as event entities, category indicator entities, information type entities, URL entities, and user entities. For tweet classification, we have compared support vector machines and deep learning model which combines class activation mapping with one-shot learning in convolutional neural networks. Experimental results on TREC 2018 Incident Streams test collection show significant improvement over the baseline system.

Keywords: Conceptual representation, Crisis-related tweets classification, Incident Streams, Convolutional neural networks, Class activation mapping, One-shot learning, Support Vector Machines.

1 Introduction

In recent years, media has become an important source of real-time information. People use social media to communicate, socialize, debate and engage in arguments. With the rapidly increasing amount of data on social networking service (SNS) like Twitter, researches on event detection and classification for the crisis-related tweet data are attracting more and more attention. During a crisis situation, people use SNS to post situational updates, look for useful information, and ask for help [1, 2]. In these cases, users can help the victims by classifying and understanding the posts as crisis-related tweets.

During natural disasters, SNS such as Twitter and Facebook are widely used by affected people to post crisis-related messages. These crisis-related posts disperse among multiple categories like information wanted about missing, injured and dead people,

---

1 Corresponding author

and infrastructure damage. The challenge here is to classify important situational updates from these crisis-related tweets, assign them appropriate informational categories [3, 4].

Related with this challenge, TREC 2018 introduced an Incident Streams (TREC-IS) track which is designed to categorize a stream of tweets related to each incident by information type in order to automatically process social media streams during emergency situations [5]. There have been several approaches to classify incident streams by using deep learning models and word representation methods. Chy et al. [6] used a method of combining convolutional neural networks(CNNs) and Long Short Term Memory (LSTM). Buntain [7] used a pre-trained word vector from Wikipedia and the FastText method. Miyazaki et al. [8] used Convolutional Neural Networks(CNNs) model and a simple Multilayer Perceptron (MLP) by inputting text BoW vector.

Deep learning has made major advances in areas such as image processing and natural language processing. Convolutional neural networks have recently achieved remarkably strong performance on the practically important task of sentence classification [9]. In recent image classification research, a technique for generating a class activation map (CAM) is introduced using the global average pooling in CNNs [10]. A class activation map for a specific class indicates the discriminative image regions which are relevant to that class. On the other hand, when learning with small number of training data, deep learning does not provide a satisfactory performance. One-shot learning aims to learn information about object categories from a few training images.

In this paper, in order to categorize a stream of tweets related to each incident by information types, we propose a conceptual representation for crisis-related tweet classification. The crisis-related terms in a tweet are represented as conceptual entities such as event entities, category indicator entities, information entities, URL entities, and user entities to represent the specific characteristics of information types in disaster and emergency situation. We have experimented using deep learning models which combines CAM and one-shot learning model. The CAM is applied to text classification by highly weighting significant term based on class activation mapping using global average pooling instead of fully connected layer. To deal with small number of training data to learn, we have used one-shot learning method using fully connected layer. To show the effectiveness of our method, experiments are conducted on a TREC-IS 2018 test collection.

The paper is organized as follows: Section 2 describes a conceptual representation for crisis-related tweets; Section 3 presents combining deep learning models. Section 4 describes experimental results. Finally, we conclude in Section 5.

2 Conceptual Representation for Crisis-Related Tweets

In order to classify a stream of tweets related to the incident, the terms in each tweet are represented as conceptual entities such as event entities, category indicator entities, information type entities, URL entities, and user entities. The conceptual entities are defined according to the information type categories.

2.1 Event Entities

In TREC-IS training data, hashtags and keywords used for collecting a stream of tweets related to each incident have been provided. For conceptual representation, the hashtags and keywords are represented as event entities: <EventLoc>, <EventCri>, and <EventClu> according to the information in such as a location, an incident and a clue. The followings are examples of event entity representation for each hashtag or keyword.

- <EventLoc>: #colorado wildfire, #abflood, Lax airport, #colorado, #costarica, …
- <EventCri>: #colorado wildfire, #abflood, #earthquake, #Wildfire, …
- <EventClu>: #StrongerPH, #prayfornsw, #SafeNow, …

Since each tweet includes event entities, an event entity itself may not give information. However, it can reduce the noise by the use of diverse terms.

2.2 Category Indicator Entities

In a training data, category indicator terms are provided for each tweet by human assessors. We have used the indicator terms for each category indicator entity which are provided by TREC-IS training data. The number of categories in TREC-IS and category indicator entities are 25 such as <InformationWanted>, <ServiceAvailable>, <Official>, <Donations>, <EmergingThreats> and etc.

Table 1. Some examples of category indicator entities.

Category Indicator Entity	Examples
<MultimediaShare>	#photo, camera image, videos, …
<InformationWanted>	does anyone know, report?, what is going on, …
<ServiceAvailable>	heavy air tanker, animal hospital, offering housing, …
<Official>	tsunami warning issued, briefing, emergency declaration, …
<Donations>	donations, items needed, sending, …
<Factoid>	kill, death, victims, …
<Weather>	#weather, degree, rain, …
<Movepeople>	evacuate, leave town, evacuation order, …
<EmergingThreats>	road blocked, sewage treatment disabled, landslide reported, …

The terms for 5 category indicator entities such as volunteering, donation, multimedia, and sentiment are expanded using terms in thesaurus.com and relatedwords.org.

2.3 Information Type Entities

In order to represent information type for each term, information type entities are defined as followings:
<Question>, <Communication>, <Organization>, <Evacuation>, <Location>, <Emergency>, <Person1st>, <Informal>, <Emotion/reaction>, <Narration>, <News>, <Disaster/threat>, <Change>, <Service>, <Provider>, <Fact>, <Unit>, <Government>, <Advice>, <Pastdate>, <Pastverb>

Table 2. Some examples of information type entities with expected category.

Information Type Entity	Examples	Expected Category
<Disaster-Threat>	Power is out, No power, People trapped, …	EmergingThreats
<Evacuation>	leave NOW, evacuation orders, closed the highway, …	MovePeople
<Government>	USGS, LarimerSheriff, Commander, …	Official
<Organization>	HumaneSociety, RedCrossDenever, NoCORedCross, …	
<Emergency>	be on ALERT, please don't wait, …	EmergingThreats
<Unit>	magnitude, dollars, acres, …	Factoid
<Provider>	Red Cross, Animal Hosptal, #Gov20, …	Volunteer
<Communication>	Call, text, hotline, dm, …	Official, Advice
<Service>	animal boarding pick up, disaster recovery center opened	ServiceAvailable
<Change>	Magnitude reduced, update, now cancelled, …	EmergingThreats

The seed terms we have defined are expanded using terms in thesaurus.com and relatedwords.org. We have assumed that each information type entity would give additional information for classification. For example, the entities such as <Government> and <Organization> would be helpful for the category 'Official', an entity <News> for the category 'ContinuingNews', an entity <Disaster/threat> for the category 'EmergingThreats', and an entity <Evacuation> for the category 'MovePeople'.

2.4 URL Entities

A URL itself has information as represented by the URL. We have defined URL entities such as <URLVideo>, <URLPhoto>, <URLNews>, <URLWeather>, <URLOrganization>, <URLDonation>, <URLDisaster> and <URLBlogMagazine>.

We have assumed that each URL entity would give information for a category. For example, the entity <URLDonation> would be helpful for the category ‘Donations’, an entity <URLWeather> for the category ‘Weather’, an entity <URLOrganization> for the category ‘Official’, an entity <URLNews> for the categories ‘ContinuingNews’ or ‘PastNews’, an entity <URLDisasterInfo> for the category ‘EmergingThreats’, and the entities <URLVideo> and <URLPhoto> for the category ‘MultimediaShare’.

In order to extract useful URLs, the TREC-IS training data and CrisisT26 dataset are used. For pre-processing, a short URL is converted to an original URL. The URLs with more than frequency 10 have been defined as 8 entities. The followings are term lists for each URL entity.

Table 3. Examples of URL entities with expected category.

URL Entity	Examples	Expected Category
<URLVideo>	youtube.com, vimeo.com	MultimediaShare
<URLPhoto>	instagram.com, twitpic.com, facebook.com/photo.php	
<URLNews>	yahoo.com/news, cbsnews.com, bbc.co.uk	ContinuingNews PastNews
<URLWeather>	wunderground.com, theweathernetwork.com	Weather
<URLOrganization>	larimer.org, usgs.gov, emergenza24.org	Official
<URLDonation>	worldvision.org.ph, prizeo.com, secure.redcross.ca	Donations
<URLDisaster>	sismos.cl, tenki.jp	EmergingThreats
<URLBlogMagazine>	livejournal.com, grist.org	KnownAlready

Table 4. User entity for expected category.

User Entity	Examples	Expected Category
<UserNews>	@9NEWS, @BreakingNews, @ANCALERTS	ContinuingNews PastNews
<UserWeather>	@dost_pagasa, @breakingstorm	Weather
<UserOrganization>	@NSWRFS, @oxfamgb, @NASA	Official
<UserDonation>	@RescuePH	Donations
<UserDisasterInfo>	@NewEarthquake, @INGVterremoti	EmergingThreats
<UserMultimedia>	@youtube	MultimediaShare

2.5 User Entities

Based on a user's characteristics, we have defined user entities such as <UserNews>, <UserWeather>, <UserOrganization>, <UserDonation>, <UserDisasterInfo>, and <UserMultimedia>.

In order to extract useful user information in tweets, the TREC-IS training data and CrisisT26 dataset are used. The users with more than frequency 10 have been defined as 6 user entities as the characteristics. We have assumed that a user entity <UserDonation> would be helpful for the category 'Donations', a user entity <UserWeather> for the category 'Weather', an entity <UserNews> for the categories 'Factoid', 'ContinuingNews' or 'PastNews', an entity <UserOrganization> for the category 'Official', and an entity <UserMultimedia> for the category 'MultimediaShare'.

3 Combining Deep Learning Models

In image localization, class activation mapping highlights the class-specific discriminative regions [10]. Class activation mapping can identify the importance of the image regions most relevant to the particular category. We expect that the class activation map for text classification detect the importance of terms.

On the other hand, there has been research using one-shot learning in Long-Short Term Memory. One-shot learning is effective when the number of training data is too small to learn [11, 12]. It improves classification accuracy.

We combine CAM with one-shot learning in CNNs as shown in Fig. 1. We expect crisis-related terms are continuously changed through the learning of term weights with high importance.

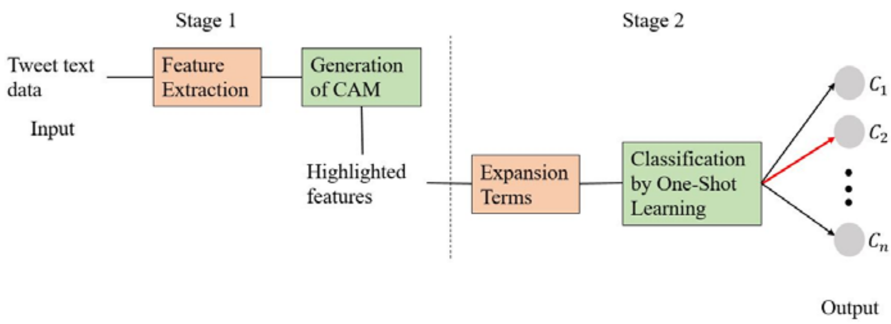

Fig. 1. Combining deep learning models with CAM and one-shot learning in CNNs.

3.1 Class Activation Mapping for Text Classification

In this paper, class activation mapping is applied for text classification. Fig. 2. shows CAM applied for image localization and text classification. The class activation mapping is generated by multiplying the weight from the convolution layer and the weight used for classification using the global average pooling in CNNs. This allows us to check the importance of terms in the input data. Through learning, global average pooling creates class activation mapping and represents terms. The proposed method readjusts the terms weight in order of importance and re-inputs them.

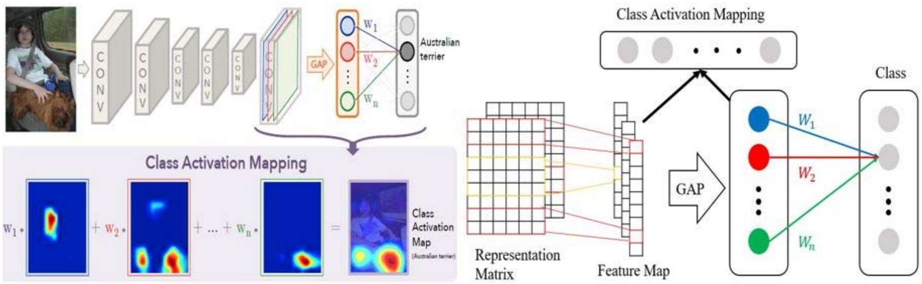

Fig. 2. Class activation mapping for image localization and text classification.

3.2 Features for One-shot Learning

CNNs consists of a convolutional layer, a maximum pooling layer, and a fully connected layer [9]. One-shot learning uses CNN's document vectors that come from fully connected layer and category vectors that represent each category. One-shot learning finds a category vector corresponding to the matched category, learns the index of similarity of the document vector and sets the value to 1. In the other case of finding, the category vector corresponding to the unmatched category, it sets the value of index of similarity to zero.

4 Experiments

4.1 Test Collection

In TREC Incident Streams (TREC-IS) 2018 test collection, a stream of tweets is collected by 19 events from the CrisisLex or CrisisNLP initiatives using hashtags and keyword monitoring [14, 15]. The tweets are related to the incidents of six different types, e.g. earthquakes, hurricanes, public or shootings [13]. The training set contains six event topics such as 2012 Colorado wildfires, 2012 Costa Rica Earthquake, 2013 Colorado floods, 2012 Typhoon Pablo, 2013 LA Airport Shooting, and 2013 West Texas Explosion. The test set contains 15 event topics such as 2012 Guatemala earthquake, 2012 Philippines floods, 2013 Australia bushfire, 2013 Boston bombings, 2011 Joplin tornado, 2015 Paris attacks and etc.

The TREC-IS task is classifying the tweets into 25 information types. The information type categories, training data and test data for each category are shown in Table 6. The number of tweets in a training set is 1335, whereas the number of tweets in a test set is 22160.

Since the human assessors were allowed to select as many information types when creating the ground truth, the performance is measured in two ways: any-type and multi-type. In any-type evaluation, a system receives a full score for a tweet if it assigned any of the categories that the human assessor selected for that tweet. In the multi-type evaluation, the performance per information type is calculated in a 1 vs. all manner [13]. As evaluation metrics, Precision (P), Recall (R), F_1 score (F_1), and Accuracy (A) are measured.

4.2 Comparison Experiments

In order to see the effectiveness of the proposed conceptual representation, we have compared our method with baseline term representation for two machine learning methods: support vector machines (SVM) [16] and combining deep learning models.

- baseSVM: SVM with term representation.
- baseComb: Combining deep learning models with term representation.
- conceptSVM: SVM with conceptual representation.
- conceptComb: Combining deep learning models with conceptual representation

We have compared our method with TREC-IS 2018 track participants with high performance.

- KDEIS4_DM: CNNs with bidirectional LSTM [6].
- umdhcilfasttext: FastText using pre-trained word vector datasets on Wikipedia [7].
- NHK: Simple Multilayer perceptron by inputting text BoW vector [8].

4.3 Experimental Results

Table 5. shows the experimental results of information type categorization for any-type. The proposed conceptual representation achieved significant improvement in SVM and deep learning models compared to the baseline representation. The proposed representation showed high performance in F_1 score compared to other participant methods. In this experiments, SVM showed better performance than deep learning models. We can guess it is due to the small number of a training data for each category even though we adopted one-shot learning method.

Table 6 shows results per each information type performance for multi-type.

5 Conclusion

In this paper, in order to classify a stream of tweets related to the incident, we proposed conceptual representation for terms, URLs and users in a tweet. The crisis-related terms were represented as conceptual entities such as event entities, category indicator entities, information type entities, URL entities, and user entities. We have assumed that each conceptual entity would give additional information for information type classification. Experimental results show the conceptual representation is effective for crisis-related tweet classification. In our experiments, SVM showed better performance than combining class activation mapping and one-shot learning in CNNs.

Table 5. Experimental results for information type categorization(any-type)

Methods	Precision	Recall	F_1	Accuracy
baseSVM	0.442	0.748	0.555	0.402
baseComb	0.416	0.616	0.497	0.340
conceptSVM	0.456	0.778	**0.575**	0.421
conceptComb	0.532	0.533	0.533	0.389
KDEIS4_DM	0.391	0.986	0.560	0.391
umdhcilfasttext	0.453	0.726	0.558	0.402
NHK	0.448	0.714	0.551	0.400

Table 6. Information type categorization (multi-type) per information type performance

				baseSVM			conceptSVM			baseComb			conceptComb		
	Categories	Train data	Test data	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1
Request	GoodsServices	0	126	0	0	0	0	0	0	0	0	0	0	0	0
	SearchAndRescue	0	286	0	0	0	0	0	0	0	0	0	0	0	0
	InformationWanted	10	172	0.07	0.02	0.03	0.17	0.01	0.02	0	0	0	0	0	0
Call-ToAc-tion	Volunteer	2	116	0	0	0	0	0	0	0	0	0	0	0	0
	Donations	15	804	0.46	0.47	0.46	0.48	0.47	0.47	0.30	0.05	0.09	0.62	0.46	0.53
	MovePeople	26	27	0.04	0.11	0.06	0.05	0.19	0.07	0.04	0.07	0.05	0.05	0.04	0.04
Report	FirstPartyObservation	28	3807	0.28	0.01	0.01	0.37	0.01	0.01	0	0	0	0.47	0	0
	ThirdPartyObservation	15	4160	0.44	0	0	0.30	0	0	0.38	0	0	0.25	0	0
	Weather	42	1325	0.38	0.10	0.15	0.53	0.17	0.25	0.40	0.07	0.12	0.62	0.15	0.24
	EmergingThreats	36	686	0.08	0.01	0.02	0.07	0.02	0.03	0.03	0	0	0.07	0.01	0.01
	SignificantEventChange	34	415	0.03	0	0.01	0.02	0.01	0.01	0.03	0	0	0.03	0	0.01
	MultimediaShare	127	3974	0.38	0.34	0.36	0.40	0.46	0.43	0.33	0.13	0.19	0.48	0.26	0.34
	ServiceAvailable	15	1076	0.36	0.03	0.05	0.37	0.09	0.15	0.20	0	0.01	0.57	0.03	0.06
	Factoid	140	2383	0.32	0.18	0.23	0.38	0.18	0.25	0.33	0.10	0.15	0.4	0.14	0.21
	Official	52	403	0.15	0.06	0.08	0.18	0.05	0.08	0.17	0.04	0.06	0.18	0.07	0.10
	CleanUp	2	62	0	0	0	0	0	0	0	0	0	0	0	0
	Hashtags	4	3363	0	0	0	1	0	0	0.08	0	0	0	0	0
Other	PastNews	12	1351	0.50	0	0	0.45	0	0.01	0.26	0.01	0.03	0	0	0
	ContinuingNews	251	4871	0.41	0.41	0.41	0.43	0.31	0.36	0.42	0.32	0.36	0.50	0.28	0.36
	Advice	39	1209	0.19	0.07	0.10	0.18	0.05	0.08	0.17	0.04	0.07	0.27	0.06	0.09
	Sentiment	132	6952	0.76	0.39	0.52	0.72	0.42	0.53	0.51	0.54	0.52	0.64	0.48	0.55
	Discussion	51	2060	0.22	0.05	0.08	0.23	0.04	0.07	0.15	0.04	0.06	0.20	0.02	0.04
	Irrelevant	163	2605	0.19	0.22	0.20	0.21	0.23	0.22	0.14	0.11	0.12	0.17	0.31	0.22
	Unknown	26	77	0	0	0	0	0	0	0.01	0.25	0.02	0.01	0.19	0.01
	KnownAlready	113	1101	0.13	0.09	0.10	0.13	0.10	0.11	0.16	0.06	0.08	0.26	0.07	0.11

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2017R1D1A1B03036275).

References

1. Vieweg S., Castillo C., Imran M.: Integrating Social Media Communications into the Rapid Assessment of Sudden Onset Disasters. In: Aiello L.M., McFarland D. (eds) Social Informatics. SocInfo 2014. Lecture Notes in Computer Science, vol 8851. Springer, Cham (2014).
2. Rudra K., Banerjee S., Ganguly N., Goyal P., Imran M., Mitra P.: Summarizing situational tweets in crisis scenario. In: Proceedings of the 27th ACM Conference on Hypertext and Social Media, pp. 137–147. ACM, (2016).
3. Nguyen D. T., Al-Mannai K., Joty S., Sajjad H., Imran M., Mitra P.: Robust Classification of Crisis-Related Data on Social Networks Using Convolutional Neural Networks. In: Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM), Montreal, QC, Canada, (2017).
4. Rudra K., Banerjee S., Ganguly N., Goyal P., Imran M., Mitra P.: Summarizing Situational Tweets in Crisis Scenario. In: Proceedings of the 27th ACM Conference on Hypertext and Social Media. Halifax, Nova Scotia, Canada (2016).
5. McCreadie R., Buntain C., Soboroff I.: TREC 2018 Incident Streams Track Guidelines, http://dcs.gla.ac.uk/~richardm/TREC_IS/, (2018)
6. Chy A. N., Siddiqua U. A., Aono M.: Neural Networks and Support Vector Machine based Approach for Classifying Tweets by Information Types at TREC 2018 Incident Streams Task. In: Proceedings of the 26th Text Retrieval Conference. Gaithersburg, Maryland (2018).
7. Buntain C.: Learning Information Types in Social Media for Crises. In: Proceedings of the 26th Text Retrieval Conference. Gaithersburg, Maryland (2018).
8. Miyazaki T., Makino K., Takei Y., Okamoto H., Goto J.: NHK STRL at TREC 2018 Incident Streams track. In: Proceedings of the 26th Text Retrieval Conference. Gaithersburg, Maryland (2018).
9. Kim Y.: Convolutional neural networks for sentence classification. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1746–1751, (2014).
10. Zhou B., Khosla A., Lapedriza A., Oliva A., Torralba A.: Learning Deep Features for Discriminative Localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2921–2929, (2016).
11. Vinyals O., Blundell C., Lillicrap T., Wierstra D.: Matching networks for one shot learning. In: Advances in Neural Information Processing Systems, pp. 3630–3638, (2016).
12. Akata Z., Perronnin F., Harchaoui Z., Schmid C.: Label-embedding for image classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1425–1438, (2015).

13. McCreadie R., Soboroff I.: TREC 2018 Incident Streams Track Overview Paper. In: Proceedings of the 26th Text Retrieval Conference. Gaithersburg, Maryland (2018).
14. Olteanu A., Castillo C., Diaz F., Vieweg S.: CrisisLex: A Lexicon for Collecting and Filtering Microblogged Communications in Crises. In: Proceedings of the AAAI Conference on Weblogs and Social Media (ICWSM'14). AAAI Press, Ann Arbor, MI, USA (2014).
15. Imran M., Mitra P., Castillo C.: Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages. In: Proceedings of the 10th Language Resources and Evaluation Conference (LREC), pp. 1638-1643. Portorož, Slovenia (2016).
16. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research (JMLR) 12, pp. 2825-2830, (2011).