# ICTNET at Trec 2019 Incident Streams Track

Kuang Guangsheng[1,2], *ZhangKun*[1,2], *ZhangJiabao*[1,2], *ZhengXin*[1,2]
[1].*UniversityofChineseAcademyofSciences, Beijing, China*
[2].*CASKeyLabofNetworkDataScienceandTechnology, InstituteofComputingTechnology*
{KuangGuangsheng,ZhangKun18z,ZhangJiabao18s,ZhengXin18s}@ict.ac.cn

## Abstract

Social medial become our public ways to share our information in our lives. Crisis management via social medial is becoming indispensable for its tremendous information. While deep learning shows surprising outcome in many tasks. So in this paper, we cope this learning task with neural network in the view of classification problem.

## 1 Introduction

Social media information process plays an important role in our lives. We can acquire a new complete way to get the things done like crisis management. However, it is very difficult to get the right or the important things we need because of the data volume and noise.

In terms of performance, classification applying this task can perform better because its rich supervised information. So classification methods with the supervised label information can satisfied our needs ,and we will adopt it.

## 2 Our Method

In this section, we will introduce our understanding on this task and state our method to tackle this problem. we explore to learn information types from the tweets. We can process it with different ways, the social network structure can be used to acquire the whole understanding on crisis, the time series in tweets have abundant information. The text released and other static information in the tweets is provided for us.

The task here will be regard as classic classification in our way. our classification object is to label different information type for different tweets. The data streams will be balanced by the attribute of data priority. Features from users and tweets will processed by neural network with supervised information. The lstm net process the tweets text and the multiple layers net process other tweet information. For the priority, we a recognized it as a classification.

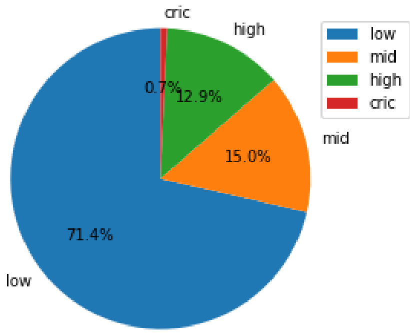


Figure 1: Proportion of different type in priority

**Data Enhancement.** We use the dataset provided by the host, while the data is unbalanced which is showed in Fig 1. The proportion of different type in priority is very unbalanced.To balance the attribute of the priority, We copy the data which is less than others according the attribute of data priority. The final size of data is similar in the view of attribute of priority. For the other classification types, we use the F1 score as the cost function to balance the data. These data enhancement methods are introduced in [1, 2].

## 3 Proposed Model

The tweets text will be cleaned for next procedure, then tokenized and put into our embedding lays with the weights of glove.840B.300d.txt to represent their seman-

Table 1: Performance Comparison

| run | Alerting | | Information Feed | | | Prioritization | |
|---|---|---|---|---|---|---|---|
| | Accumulated Alert Worth | | Info.Type Postive F1 | | Info. Type Accuracy | Priorty RMSE | |
| | High Priority | All | Actionable | All | All | Actionable | All |
| *ict_dl* | -0.3979 | -0.3103 | 0.0347 | 0.0787 | 0.7285 | 0.1254 | 0.1451 |
| median score | -0.9197 | -0.4609 | 0.0386 | 0.1055 | 0.8583 | 0.1767 | 0.1028 |

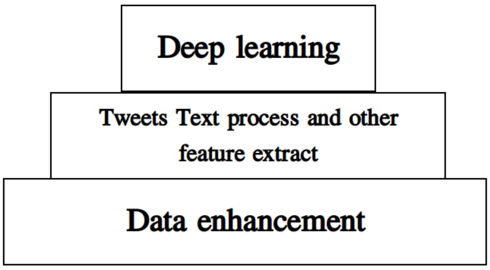


Figure 2: Proportion of different type in priority

tic content. Inspired by the last best system, the contents and their sentiment will considered into our features.

After features engineering,the text streams will be processed by the bidirectional lstm ,nand the forward network will process the network features like friends numbers.and they will be jointed and pass through several forward network layers to extract higher features,like most of text process work [3, 4].We will predict the priority and information type independently with the same model except the output numbers and cost function. Two model are regarded this task classification problem.

Our deep learning model shows in the Fig 2.

# 4 Experiment

## 4.1 Model Training and Evaluation

We train our model on the dataset provided by the TREC-IS organizer and implement the model in keras and run it on linux. For the model, we use categorical loss to train the priority measured by the F1 and mean squared error. the other categories is optimized by F1 loss and measured by the F1 score. by the limit of our computer resource,we iterate them 2 epochs. In our train process, f1 score is 0.2200 in the aspect of the priority. For the other information type, the overall f1 score is 0.1129, these shows our method help us make great progress to tackle this problem.

However, in this track,the result was once evaluated by F1 score among all the categories. A new metric called AAW(Accumulated alert worth) is introduced to capture and evaluate important information. So our data enhancement will make the important information stand out and answer for this metric.we haven't optimise the information feed according the information categories.And we just use F1 cost function to ease our noise during the data enhancement process.

## 4.2 Results

The result from the host is shown in Table 1. There are seven indicators to evaluate our method. Meanwhile, we can get better understanding on our method according to the track median scores the host transmits to us.

AAW score ranges between -1 and 1,higher is better.our alerting score is -0.3979 in the high priority and -0.3103 over all. contrast to the median score in term of this view,The run has better performing. median alert score will make the lower alerts missing important things. The result of our method originate from our data enhancement that it bring much attention into the rare types making the data balanced that bring a good score in alerting.

As showed in the result, it produce much poor performance than the median on the information feed metrics. information type f1 score is slightly lowered than the median score, but the overall accuracy is much lower than the median score which will mistake more categories than median.much lower accuracy result from the error produced by the data process. during the process, we only balance the priority other than the information feed. it has been eased by the f1 loss function during the training.

Prioritization score shows our performance is about the median level. Since we balance,category them and measured by RMSE, the result is about median rank because we train them by categorical loss.

In short, our data enhancement make alerting score better and bring about lower f1 score.

# 5 Conclusion

In this paper,we have implement a category algorithm for this incident stream track.

In a word, we should have a long way to go on this topic according our performance. We can improve our data with other data sampling methods or other ways like reinforcement to expand our data. In the feature representation, we can extract the social structure and use some short texture processing methods. In the last, we should try more ways to model the data and train the model.

# References


[1] N. V. Chawla. Data mining for imbalanced datasets: An overview. In *Data mining and knowledge discovery handbook*, pages 875–886. Springer, 2009.

[2] H. He and E. A. Garcia. Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9):1263–1284, 2009.

[3] M. B. Lazreg, M. Goodwin, and O.-C. Granmo. Deep learning for social media analysis in crises situations (position paper). In *The 29th Annual Workshop of the Swedish Artificial Intelligence Society (SAIS), 2–3 June 2016, Malmö, Sweden*, number 129. Linköping University Electronic Press, 2016.

[4] P. Zhou, Z. Qi, S. Zheng, J. Xu, H. Bao, and B. Xu. Text classification improved by integrating bidirectional lstm with two-dimensional max pooling. *arXiv preprint arXiv:1611.06639*, 2016.