

WUST CRF-Based System at NTCIR-13 ECA Task

Maofu Liu, Linxu Xia, Zhenlian Zhang, Yang Fu
School of Computer Science and Technology
Wuhan University of Science and Technology
liumaofu@wust.edu.cn



Contents

- Introduction
- Related work
- System Architecture
- Experiments
- Conclusions

Introduction

- In NTCIR-13, Emotion Cause Analysis (ECA) is an evaluation task aiming to detect the emotion cause description from the text, which is a common problem shared widely among researchers of Natural Language Processing (NLP) and information access.
- We regard the task that the emotion cause detection at clause as a clause-level binary text classification problem. The clauses will be classified according to containing emotion cause or not.
- In our system, we apply the Conditional Random Field (CRF) classification model to identifying the emotion cause description with a series of features, such as basic word features, POS features, distance features, and contextual features.

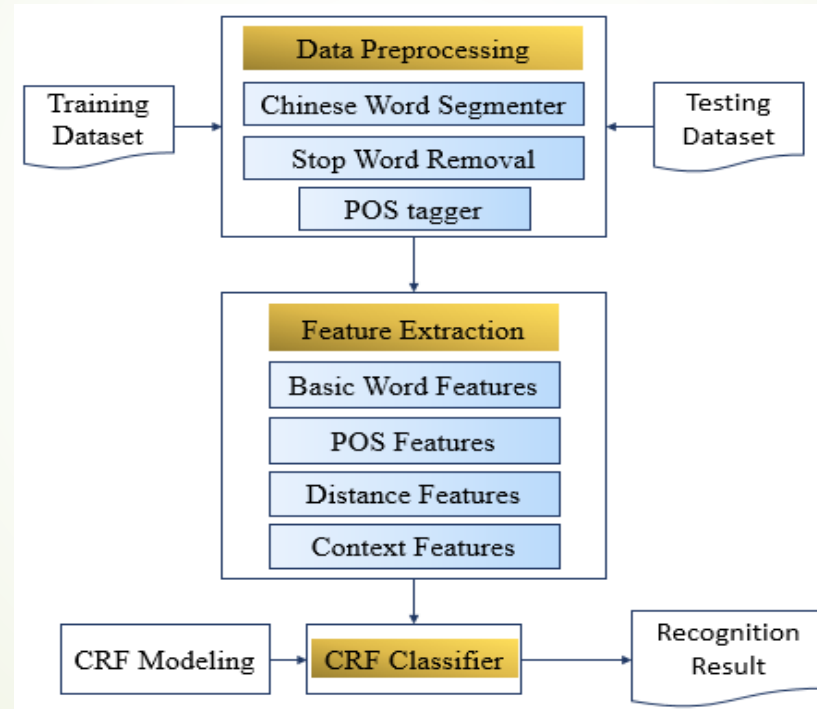


Related work

- The text emotion analysis has become a research hotspot in the field of NLP. At present, the study of emotion analysis mainly focuses on the emotion recognition and emotion cause analysis.
- The emotion recognition focuses on two tasks
 - emotion classification
 - emotion information extraction
- For emotion cause analysis, the majority of these works are based on rules and statistical model.
- In this paper, we choose the CRF model to solve the problem of the ECA task.

System Architecture

- Our system consist of three main modules, i.e. data preprocessing, feature extraction, and CRF classifier.



Feature extraction

- ▶ In this subsection, we divide feature sets into basic features and contextual features.

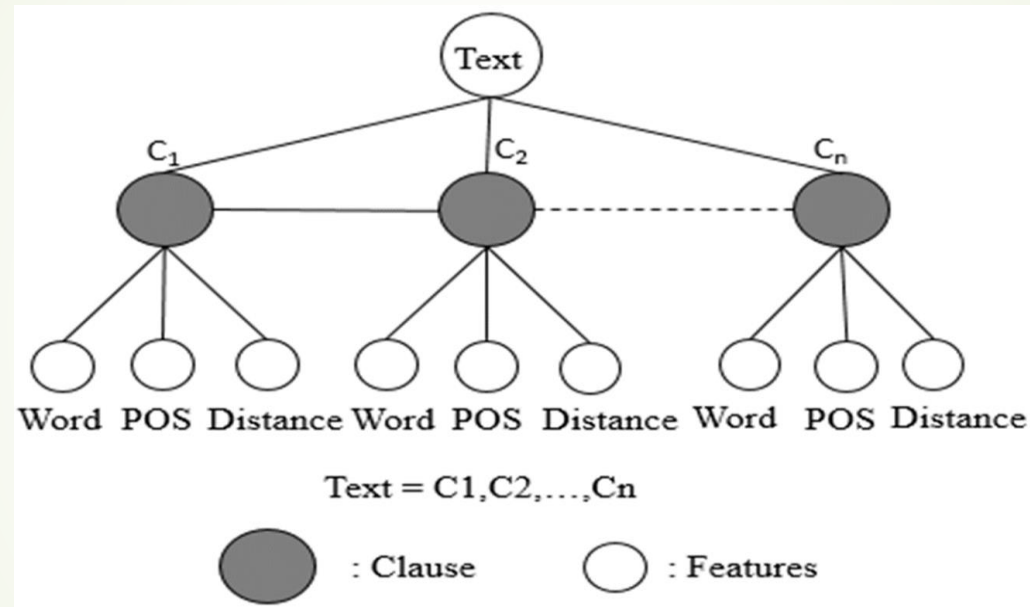
- Basic features

Features	Description
nouns	The nouns in the present sentence, if not, should be filled with "NULL".
verbs	The verbs in the present sentence, if not, should be filled with "NULL".
numbers of nouns	The number of nouns containing in the present sentence.
numbers of verbs	The number of verbs containing in the present sentence.
distance	The distance between present sentence and emotional keyword, and the values represented by Left2, Left1, Keyword, Right1 and Right2 are -2, -1, 0, 1 and 2 respectively.

- The contextual features mainly contain the syntactic ones of the previous sentence and the next sentence, and they can be represented by feature templates.

CRF Classifier

- The schematic diagram of CRF classification model in our system .



- The specific formula is shown in formula (1).

$$\mathbf{P}_{\lambda}(\mathbf{Y}|\mathbf{W}) = \frac{1}{Z(\mathbf{W})} \exp\left(\sum_{t \in T} \sum_k \lambda_k f_k(\mathbf{y}_{t-1}, \mathbf{W}, t)\right) \quad (1)$$

Experiments

- In this paper, we use the data sets provided by NTCIR-13. This dataset contains about 2,200 training instances and 10,000 instances for testing.
- There are three values to evaluate the experimental results, namely precision (P), recall (R) and f -score (F).

$$\begin{aligned} \textit{Precision}_{\textit{clause}} &= \frac{\# \textit{correct cause relevant clauses}}{\# \textit{detected cause relevant clause}} \\ \textit{Recall}_{\textit{clause}} &= \frac{\# \textit{correct cause relevant clauses}}{\# \textit{annotated cause relevant clause}} \\ \textit{F - measure}_{\textit{clause}} &= \frac{2 \times \textit{Precision}_{\textit{clause}} \times \textit{Recall}_{\textit{clause}}}{\textit{Precision}_{\textit{clause}} \times \textit{Recall}_{\textit{clause}}} \end{aligned}$$

Experimental results

- ▶ We submitted one system result to NTCIR-13 of emotion cause detection at clause level. The official evaluation results of performance are listed in the Table 1.

Table 1. Performances of the CRF model

Team	P	R	F
WUST	0.6930	0.6399	0.6654

Conclusions

- ▶ In our work, we construct the classification model based on conditional random field to recognize emotion cause in Chinese text pair using the hybrid features, including words, POS, distance and context.
- ▶ In the future, we will introduce more features into the Emotion Cause Analysis, such as coherent and semantic features. The other methods, like neural network based on attention model, will be also applied to detecting the emotion cause in our future work.



END

Thanks for listening!

2017.11.28