

Multilingual Emotion Detection

Junyin Chen and Hanshu Ding and Zoe Fang and Yifan Jiang

{junyinc, hsdning99, zoekfang, yfjiang}@uw.edu

Department of Linguistics

University of Washington

Abstract

TBD

1 Introduction

TBD

2 Task Description

2.1 Primary Task

Our primary task is an emotion detection task on dialogues in the EmotionLines dataset. The dataset is in English and contains TV show scripts and text message dialogues. We need to classify the emotion for each utterance in dialogues.

2.1.1 Dataset

EmotionLines (Hsu et al., 2018)¹ dataset is an emotion dialogue dataset with emotion labels for each utterance. The dataset contains a total of 29, 245 utterances from 2, 000 dialogues, which are collected from Friends TV scripts and private Facebook messenger dialogues. Each utterance is labeled with one of Ekman’s six basic emotions plus the neutral emotion.

Hsu et al. (2018) have split the dataset into training, development, and testing set separately. We use the testing set and corresponding gold standard annotation for analysis.

2.2 Adaptation Task

Our adaptation task is to adapt our emotion detection model to Chinese dialogues in the Multi-party Dialogue Dataset (MPDD). Other dimensions for this task remain the same as the primary task.

¹<http://doraemon.iis.sinica.edu.tw/emotionlines/index.html>

2.2.1 Dataset

Multi-party Dialogue Dataset (MPDD) (Chen et al., 2020)² is a Chinese emotion dialogue dataset with emotion and relation labels for each utterance. The dataset contains a total of 25, 548 utterances from 4, 142 dialogues, which are collected from five TV series scripts from www.juban108.com. The dialogues have three types of labels: emotion, relation, and target listener. In particular, emotion labels are consistent with those in the Emotionlines dataset.

We will randomly split ten percent of the dataset as the test data for analysis.

2.3 Evaluation

Both tasks will be evaluated using standard evaluation metrics, including accuracy, precision, recall and F1-score.

3 System Overview

TBD

4 Approach

TBD

5 Results

TBD

6 Discussion

TBD

7 Conclusion

TBD

²<http://nlg.csie.ntu.edu.tw/nlpresource/MPDD>

8 References

References

- Yi-Ting Chen, Hen-Hsen Huang, and Hsin-Hsi Chen. 2020. [MPDD: A multi-party dialogue dataset for analysis of emotions and interpersonal relationships](#). In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 610–614, Marseille, France. European Language Resources Association.
- Chao-Chun Hsu, Sheng-Yeh Chen, Chuan-Chun Kuo, Ting-Hao Huang, and Lun-Wei Ku. 2018. [Emotion-Lines: An emotion corpus of multi-party conversations](#). In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).