Overview of NLPCC 2018 Shared Task 1: Emotion Detection in Code-Switching Text

Zhongqing Wang(✉), Shoushan Li, Fan Wu, Qingying Sun, and Guodong Zhou

Natural Language Processing Lab, Soochow University, Suzhou, China
{wangzq, lishoushan, gdzhou}@suda.edu.cn,
20174227039@stu.suda.edu.cn, qingying.sun@foxmail.com

Abstract. This paper presents the overview of the shared task, emotion detection in code-switching text, in NLPCC 2018. The submitted systems are expected to automatically determine the emotions in the Chinese-English code-switching text. Different from monolingual text, code-switching text contains more than one language, and the emotion can be expressed by either monolingual or bilingual form. Hence, the challenges are: how to integrate both monolingual and bilingual forms to detect emotion, and how to bridge the gap to between two languages. Our shared task has 19 team participants. The highest F-score was 0.515. In this paper, we introduce the task, the corpus, the participating teams, and the evaluation results.

Keywords: Emotion detection · Code-switching text
Annotation and evaluation

1 Introduction

With the rapid development of Web 2.0, emotion analysis in social media has become of great value to market predictions and analysis [1–3]. Previous researches on emotion analysis have mainly focused on emotion expressions in monolingual texts [5–7]. However, in informal settings such as micro-blogs, emotions are often expressed by a mixture of different natural languages. Such a mixture of language is called *code-switching*. Specifically, code-switching text is defined as text that contains more than one language (‘*code*’). It is a common phenomenon in multilingual communities [4].

In this shared task, the submitted systems are expected to automatically determine the emotions in the Chinese-English code-switching text. It is more difficult to detect emotions in code-switching texts than in monolingual ones since emotions in code-switching posts can be expressed through one or two languages [9–12]. Hence, traditional automatic emotion detection methods which simply consider monolingual texts would not be readily applicable.

2 Data

In this task, we provide training, development, and testing data. Each post in dataset contains two language text (Chinese and English), we call such post as code-switched text. Following are some examples of code-switched posts with different emotions.

M. Zhang et al. (Eds.): NLPCC 2018, LNAI 11109, pp. 429–433, 2018.
https://doi.org/10.1007/978-3-319-99501-4_39

[E1] 婚礼上新娘大秀鼓技, high 翻全场! (**happiness**)
(The bride played the drums at the wedding, everyone was very high!)

[E2] 美好假期已经开始, have a nice time. (**happiness**)
(The happy vacation has begun, have a nice time.)

[E3] 上了一天的课, 嗓子 hold 不住了啊。(**sadness**)
(I have been teaching the whole day, my throat can't take it anymore.)

For each post, five emotions were annotated, namely *happiness*, *sadness*, *fear*, *anger* and *surprise*. The number of samples of different emotions in training, development, and testing data is in Table 1, respectively.

Table 1. Samples of different emotions

	Train	Dev	Test
Happiness	1824	220	490
Sadness	1086	120	296
Anger	570	84	111
Fear	648	85	37
Surprise	651	92	68

3 Participants

There are 46 teams registered the share task, and 19 teams submitted their final results. The teams that have submitted the results are shown in Table 2.

Table 2. Introduction of participating teams

Team name	Organization name
DeepIntell	Research team of DeepIntell co., Ltd.
DUTIR_938	Dalian University of Technology
Shining	University of South China
lxzlx624	University of South China
xiamx-rali	Université de Montréal – Laboratoire RALI
zzuhhjx	University of Zhengzhou
Team_1	Beijing Guangnian Infinity Technology Co., Ltd.
CASIA-ED	Research Center for Brain-inspired Intelligence
cmos_nlp	Move online text algorithm group
Yang_NEU	Northeastern University
rax	Peking University
NLP@WUST	Wuhan University of Science and Technology
scau_geek	South China Agricultural University
Lab1010	Southwest University
The Dream Team of NLP	Nanjing University Science and Technology
GDUFSLEC	Guangdong University of Foreign Studies
CSUNLP	Guangdong University of Foreign Studies
CQUT_301_1	Chongqing University of Technology
BISTU_IIPI	Beijing Information Science and Technology University

4 Evaluation

4.1 Evaluation Metric and Baseline

We compute precision (P), recall (R), and F1-Score for each emotion separately, and compute the macro averaged P, R and F1 with all emotions. Our official scoring metric is macro-averaged F1-Score.

We implement a simple baseline on training data, and test on the development data. The performance is shown in Table 3. We use SVM[1] as the classification method, and unigram as features.

Table 3. Performance of baseline

	P	R	F1
Happiness	0.610	0.691	0.648
Sadness	0.307	0.650	0.417
Anger	0.299	0.667	0.413
Fear	0.256	0.788	0.386
Surprise	0.142	0.391	0.208

4.2 Submission Results and Discussions

There are 19 teams submitted their valid results, the results are shown in Table 4. As Table 4 given, DeepIntell, DUTIR_938 and Shining have better results than others. In particular, **DeepIntell** converts multi-label classification into binary classification task and employ ensemble learning for code-switching text with sampling and emotion lexicon. **DUTIR_938** also use ensemble separately trains several models include CNN, RCNN and Attention based LSTM model. Then combine their classification results to improve the performance. **Shining** considers the relationship between different emotions in a single post, at the same time, the attention mechanism is employed to find the importance of different features and predict all emotions expressed by each post.

[1] http://svmlight.joachims.org/.

Table 4. Evaluation results

Team	Happiness	Sadness	Anger	Fear	Surprise	Marco-F1
DeepIntell	**0.734**	0.616	**0.543**	0.264	**0.418**	**0.515**
DUTIR_938	0.715	0.521	0.541	0.166	0.396	0.468
Shining	0.710	**0.652**	0.540	**0.292**	0.139	0.467
lxzlx624	0.734	0.637	0.570	0.204	0.164	0.462
xiamx-rali	0.624	0.494	0.457	0.200	0.366	0.428
zzuhhjx	0.692	0.428	0.406	0.186	0.376	0.418
Team_1	0.594	0.510	0.440	0.143	0.310	0.399
CASIA-ED	0.596	0.417	0.510	0.130	0.337	0.398
cmos_nlp	0.632	0.414	0.352	0.161	0.265	0.365
Yang_NEU	0.568	0.432	0.351	0.207	0.255	0.363
rax	0.576	0.421	0.392	0.103	0.306	0.360
NLP@WUST	0.630	0.374	0.287	0.146	0.354	0.358
scau_geek	0.615	0.480	0.398	0.000	0.230	0.345
Baseline	0.587	0.500	0.390	0.108	0.128	0.342
Lab1010	0.385	0.458	0.271	0.061	0.063	0.248
The Dream Team of NLP	0.480	0.350	0.160	0.062	0.137	0.238
GDUFSLEC	0.491	0.131	0.190	0.045	0.299	0.231
CSUNLP	0.441	0.354	0.033	0.093	0.149	0.214
CQUT_301_1	0.246	0.161	0.093	0.053	0.049	0.121
BISTU_IIPI	0.456	0.007	0.018	0.000	0.000	0.096

5 Conclusion

This paper briefly introduces the overview of emotion detection in code-switching text shared task at NLPCC 2018. There are 19 participants having submitted final results. And some participants get exciting results in this corpus. Meanwhile, we release a large code-switching text emotion corpus for more large-scale research in emotion and bilingual analysis.

Acknowledgment. We would like to thank the participants for their valuable feedback and results. We should thank Dr. Sophia Yat Mei Lee and Helena Yan Ping Lau for their excellent works on corpus annotation and analysis. The work is supported by the National Natural Science Foundation of China (61331011, 61751206), and the Early Career Scheme (ECS) sponsored by the Research Grants Council of Hong Kong (No. PolyU 5593/13H).

References

1. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? Sentiment classification using machine learning techniques. In: Empirical Methods in Natural Language Processing, pp. 79–86 (2002)
2. Yat, S., Lee, M., Li, S., Huang, C.: Annotating events in an emotion corpus. In: Proceedings of the Ninth International Conference on Language Resources and Evaluation, pp. 3511–3516 (2014)
3. Solorio, T., Liu, Y.: Learning to predict code-switching points. In: Conference on Empirical Methods in Natural Language Processing, pp. 973–981(2008)
4. Adel, H., Vu, N., Schultz, T.: Combination of recurrent neural networks and factored language models for code-switching language modeling. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, pp. 206–211 (2013)
5. Mohammad, S.M., Kiritchenko, S.: Using hashtags to capture fine emotion categories from tweets. Comput. Intell. **31**(2), 301–326 (2015)
6. Crossley, S.A., Kyle, K., McNamara, D.S.: Sentiment analysis and social cognition engine (SEANCE): an automatic tool for sentiment, social cognition, and social-order analysis. Behav. Res. Methods **49**, 1–19 (2016)
7. Gupta U., Chatterjee A., Srikanth R., Agrawal P.: A sentiment-and-semantics-based approach for emotion detection in textual conversations (2017)
8. Sun, R., Wilson, N., Lynch, M.: Emotion: a unified mechanistic interpretation from a cognitive architecture. Cogn. Comput. **8**(1), 1–14 (2016)
9. Jamatia, A., Das, A.: Part-of-speech tagging for code-mixed English-Hindi Twitter and Facebook chat messages. In: Recent Advances in Natural Language Processing, pp. 139–248 (2015)
10. Lee, S., Wang, Z.: Emotion in code-switching texts: corpus construction and analysis. In: Eighth Sighan Workshop on Chinese Language Processing, pp. 91–99 (2015)
11. Wang, Z., Zhang, Y., Lee, S., Li, S., Zhou, G.: A bilingual attention network for code-switched emotion prediction. In: Proceeding of COLING (2016)
12. Wang, Z., Lee, S., Li, S., Zhou, G.: Emotion analysis in code-switching text with joint factor graph model. In: Transactions on Audio Speech and Language Processing, pp. 469–480 (2017)