

AAACL-IJCNLP 2020

**Natural Language Processing
Techniques for Educational Applications**

Proceedings of the Sixth Workshop

December, 2020
Suzhou, China

©2020 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-952148-99-6

Preface

Welcome to the 6th Workshop on Natural Language Processing Techniques for Educational Applications (NLPTEA 2020), with a Shared Task on Chinese Grammatical Error Diagnosis.

The development of Natural Language Processing (NLP) has advanced to a level that affects the research landscape of many academic domains and has practical applications in many industrial sectors. On the other hand, educational environment has also been changed deeply, due to the significant impact from the international society and emergency disasters like CoVID19. With the impact, this workshop focuses on the NLP techniques applied to the educational environment. Research issues in this direction have gained more and more attention.

This is the 6th workshop held in the Asian area, with the first one NLPTEA 2014 workshop being held in conjunction with the 22nd International Conference on Computer in Education (ICCE 2014) from Nov. 30 to Dec. 4, 2014 in Japan. The second edition NLPTEA 2015 workshop was held in conjunction with the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2015) from July 26- 31 in Beijing, China. The third version NLPTEA 2016 workshop was held in conjunction with the 26th International Conference on Computational Linguistics (COLING 2016) from December 11- 16 in Osaka, Japan. The fourth edition NLPTEA 2017 workshop was held in conjunction with the 8th International Joint Conference on Natural Language Processing (IJCNLP 2017) from November 27- December 1 in Taipei, China. The fifth NLPTEA 2018 workshop was colocated with 56th Annual Meeting of the Association for Computational Linguistics from July 15- 19 in Melbourne, Australia. This year, we continue to promote this research line by holding the workshop in conjunction with the first Asia Chapter of ACL conference and also holding the fifth shared task on Chinese Grammatical Error Diagnosis.

In this year, we received 28 valid submissions for research issues, each of which was reviewed by at least two experts, and had 17 teams participating in the shared task, submitting their task reports. In total, there are 11 oral papers and 7 posters accepted. We also organize a keynote speech session in this workshop. The invited speakers, Professor Yuming Li and Professor Jihua Song are expected to deliver great talks on language education across human and machine and Chinese intelligent education online.

We would like to thank the program committee members for their hard work in completing the review tasks. Their collective efforts achieved quality reviews of the submissions within a few weeks. Great thanks should also go to the speakers, authors, and participants for the tremendous supports in making the workshop a success.

It's a great pity that we cannot meet at Suzhou city as planned, due to the impact of CoVID19. Still we welcome you to the first online NLPTEA, and wish you enjoy the workshop.

NLPTEA 2020 Workshop Chairs
Erhong Yang, Beijing Language and Culture University
Endong Xun, Beijing Language and Culture University
Baolin Zhang, Beijing Language and Culture University
Gaoqi Rao, Beijing Language and Culture University

Workshop Organizers:

Erhong Yang, Beijing Language and Culture University
Endong Xun, Beijing Language and Culture University
Baolin Zhang, Beijing Language and Culture University
Gaoqi Rao, Beijing Language and Culture University

Program Committee:

Ting Liu, Harbin Institute of Technology
Miao Fan, Baidu Co., Ltd.
Ruiji Fu, iFlyTek Co., Ltd.
Tingting He, Center China Normal University
Po Hu, Center China Normal University
Renfen Hu, Beijing Normal University
Lung-hao Lee, National Center University
Bin Li, Nanjing Normal University
Changliang Li, Kingsoft Co., Ltd.
Zhenghua Li, Soochow University
Pengyuan Liu, Beijing Language and Culture University
Dawei Lu, Renmin University of China
Shan Wang, University of Macau
Wenxin Xiong, Beijing Foreign Studies University
Chi Xiu, Microsoft (China) Co., Ltd.
Dong Yu, Beijing Language and Culture University
Liang-Chih Yu, Yuan-Ze University
Tianlin Yang, Beijing Language and Culture University
Min Zhang, Soochow University
Yue Zhang, Westlake University

Additional Reviewers:

Chengwen Wang, Beijing Language and Culture University

Invited Speaker:

Yuming Li, Beijing Language and Culture University
Jihua Song, Beijing Normal University

Table of Contents

<i>Non-Autoregressive Grammatical Error Correction Toward a Writing Support System</i> Hiroki Homma and Mamoru Komachi	1
<i>Arabisc: Context-Sensitive Neural Spelling Checker</i> Yasmin Moslem, Rejwanul Haque and Andy Way	11
<i>LXPER Index 2.0: Improving Text Readability Assessment Model for L2 English Students in Korea</i> Bruce Lee and Jason Lee	20
<i>Overview of NLPTEA-2020 Shared Task for Chinese Grammatical Error Diagnosis</i> Gaoqi Rao, Erhong Yang and Baolin Zhang	25
<i>Combining ResNet and Transformer for Chinese Grammatical Error Diagnosis</i> Shaolei Wang, Baoxin Wang, Jiefu Gong, Zhongyuan Wang, Xiao Hu, Xingyi Duan, Zizhuo Shen, Gang Yue, Ruiji Fu, Dayong Wu, Wanxiang Che, Shijin Wang, Guoping Hu and Ting Liu	36
<i>Chinese Grammatical Error Diagnosis with Graph Convolution Network and Multi-task Learning</i> Yikang Luo, Zuyi Bao, Chen Li and Rui Wang	44
<i>Integrating BERT and Score-based Feature Gates for Chinese Grammatical Error Diagnosis</i> Yongchang Cao, Liang He, Robert Ridley and Xinyu Dai	49
<i>BERT Enhanced Neural Machine Translation and Sequence Tagging Model for Chinese Grammatical Error Diagnosis</i> Deng Liang, Chen Zheng, Lei Guo, Xin Cui, Xiuzhang Xiong, Hengqiao Rong and Jinpeng Dong	57
<i>A Hybrid System for NLPTEA-2020 CGED Shared Task</i> Meiyuan Fang, Kai Fu, Jiping Wang, Yang Liu, Jin Huang and Yitao Duan	67
<i>Chinese Grammatical Error Correction Based on Hybrid Models with Data Augmentation</i> Yi Wang, Ruibin Yuan, Yan'gen Luo, Yufang Qin, NianYong Zhu, Peng Cheng and Lihuan Wang	78
<i>TMU-NLP System Using BERT-based Pre-trained Model to the NLP-TEA CGED Shared Task 2020</i> Hongfei Wang and Mamoru Komachi	87
<i>CYUT Team Chinese Grammatical Error Diagnosis System Report in NLPTEA-2020 CGED Shared Task</i> Shih-Hung Wu and Junwei Wang	91
<i>Chinese Grammatical Error Diagnosis Based on RoBERTa-BiLSTM-CRF Model</i> Yingjie Han, Yingjie Yan, Yangchao Han, Rui Chao and Hongying Zan	97
<i>Chinese Grammatical Errors Diagnosis System Based on BERT at NLPTEA-2020 CGED Shared Task</i> Hongying Zan, Yangchao Han, Haotian Huang, Yingjie Yan, Yuke Wang and Yingjie Han	102
<i>Chinese Grammatical Error Detection Based on BERT Model</i> Yong Cheng and Mofan Duan	108
<i>Named-Entity Based Sentiment Analysis of Nepali News Media Texts</i> Birat Bade Shrestha and Bal Krishna Bal	114

<i>SEMA: Text Simplification Evaluation through Semantic Alignment</i> Xuan Zhang, Huizhou Zhao, KeXin Zhang and Yiyang Zhang.....	121
<i>A Corpus Linguistic Perspective on the Appropriateness of Pop Songs for Teaching Chinese as a Second Language</i> Xiangyu Chi and Gaoqi Rao.....	129

Conference Program

Friday, December 4, 2020

9:00–9:10 *Opening Remarks*

9:10–10:50 **Keynotes Session**

9:10–10:00 *Keynote: Vision and Impact of Intelligent Essay Scoring*
Yuming Li

10:00–10:50 *Keynote: Intelligent Techniques and International Chinese Education Online*
Jihua Song

10:50–11:00 *Coffe Break*

11:00–12:00 **Oral Seesion1: NLP Techniques for Education**

11:00–11:15 *Non-Autoregressive Grammatical Error Correction Toward a Writing Support System*
Hiroki Homma and Mamoru Komachi

11:15–11:30 *Arabisc: Context-Sensitive Neural Spelling Checker*
Yasmin Moslem, Rejwanul Haque and Andy Way

11:30–11:45 *LXPER Index 2.0: Improving Text Readability Assessment Model for L2 English Students in Korea*
Bruce Lee and Jason Lee

11:45–12:00 *Overview of NLPTEA-2020 Shared Task for Chinese Grammatical Error Diagnosis*
Gaoqi Rao, Erhong Yang and Baolin Zhang

12:00–14:00 *Lunch Break*

Friday, December 4, 2020 (continued)

14:00–16:20 Oral Session2: Chinese Grammatical Error Diagnosis

- 14:00–14:20 *Combining ResNet and Transformer for Chinese Grammatical Error Diagnosis*
Shaolei Wang, Baoxin Wang, Jiefu Gong, Zhongyuan Wang, Xiao Hu, Xingyi Duan, Zizhuo Shen, Gang Yue, Ruiji Fu, Dayong Wu, Wanxiang Che, Shijin Wang, Guoping Hu and Ting Liu
- 14:20–14:40 *Chinese Grammatical Error Diagnosis with Graph Convolution Network and Multi-task Learning*
Yikang Luo, Zuyi Bao, Chen Li and Rui Wang
- 14:40–15:00 *Integrating BERT and Score-based Feature Gates for Chinese Grammatical Error Diagnosis*
Yongchang Cao, Liang He, Robert Ridley and Xinyu Dai
- 15:00–15:20 *BERT Enhanced Neural Machine Translation and Sequence Tagging Model for Chinese Grammatical Error Diagnosis*
Deng Liang, Chen Zheng, Lei Guo, Xin Cui, Xiuzhang Xiong, Hengqiao Rong and Jinpeng Dong
- 15:20–15:40 *A Hybrid System for NLPTEA-2020 CGED Shared Task*
Meiyuan Fang, Kai Fu, Jiping Wang, Yang Liu, Jin Huang and Yitao Duan
- 15:40–16:00 *Chinese Grammatical Error Correction Based on Hybrid Models with Data Augmentation*
Yi Wang, Ruibin Yuan, Yan'gen Luo, Yufang Qin, NianYong Zhu, Peng Cheng and Lihuan Wang
- 16:00–16:20 *TMU-NLP System Using BERT-based Pre-trained Model to the NLP-TEA CGED Shared Task 2020*
Hongfei Wang and Mamoru Komachi

Friday, December 4, 2020 (continued)

16:20–17:45 Poster Session

- 16:20–16:32 *CYUT Team Chinese Grammatical Error Diagnosis System Report in NLPTEA-2020 CGED Shared Task*
Shih-Hung Wu and Junwei Wang
- 16:32–16:44 *Chinese Grammatical Error Diagnosis Based on RoBERTa-BiLSTM-CRF Model*
Yingjie Han, Yingjie Yan, Yangchao Han, Rui Chao and Hongying Zan
- 16:44–16:56 *Chinese Grammatical Errors Diagnosis System Based on BERT at NLPTEA-2020 CGED Shared Task*
Hongying Zan, Yangchao Han, Haotian Huang, Yingjie Yan, Yuke Wang and Yingjie Han
- 16:56–17:08 *Chinese Grammatical Error Detection Based on BERT Model*
Yong Cheng and Mofan Duan
- 17:08–17:20 *Named-Entity Based Sentiment Analysis of Nepali News Media Texts*
Birat Bade Shrestha and Bal Krishna Bal
- 17:20–17:32 *SEMA: Text Simplification Evaluation through Semantic Alignment*
Xuan Zhang, Huizhou Zhao, KeXin Zhang and Yiyang Zhang
- 17:32–17:44 *A Corpus Linguistic Perspective on the Appropriateness of Pop Songs for Teaching Chinese as a Second Language*
Xiangyu Chi and Gaoqi Rao

Friday, December 4, 2020 (continued)

17:45–17:55 Closing Remark