

15th Workshop on Innovative Use of NLP for Building Educational Applications 2020

Online
10 July 2020

ISBN: 978-1-7138-1382-8

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2020) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2020)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006

Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

| | |
|--|-----|
| <i>Linguistic Features for Readability Assessment</i> Tovly Deutsch, Masoud Jasbi and Stuart Shieber | 1 |
| <i>Using PRMSE to evaluate automated scoring systems in the presence of label noise</i> Anastassia Loukina, Nitin Madnani, Aoife Cahill, Lili Yao, Matthew S. Johnson, Brian Riordan and Daniel F. McCaffrey | 18 |
| <i>Multiple Instance Learning for Content Feedback Localization without Annotation</i> Scott Hellman, William Murray, Adam Wiemerslage, Mark Rosenstein, Peter Foltz, Lee Becker and Marcia Derr | 30 |
| <i>Complementary Systems for Off-Topic Spoken Response Detection</i> Vatsal Raina, Mark Gales and Kate Knill | 41 |
| <i>CIMA: A Large Open Access Dialogue Dataset for Tutoring</i> Katherine Stasaski, Kimberly Kao and Marti A. Hearst | 52 |
| <i>Becoming Linguistically Mature: Modeling English and German Children’s Writing Development Across School Grades</i> Elma Kerz, Yu Qiao, Daniel Wiechmann and Marcus Ströbel | 65 |
| <i>Annotation and Classification of Evidence and Reasoning Revisions in Argumentative Writing</i> Tazin Afrin, Elaine Lin Wang, Diane Litman, Lindsay Clare Matsumura and Richard Correnti .. | 75 |
| <i>Can Neural Networks Automatically Score Essay Traits?</i> Sandeep Mathias and Pushpak Bhattacharyya | 85 |
| <i>Tracking the Evolution of Written Language Competence in L2 Spanish Learners</i> Alessio Miaschi, Sam Davidson, Dominique Brunato, Felice Dell’Orletta, Kenji Sagae, Claudia Helena Sanchez-Gutierrez and Giulia Venturi | 92 |
| <i>Distractor Analysis and Selection for Multiple-Choice Cloze Questions for Second-Language Learners</i> Lingyu Gao, Kevin Gimpel and Arnar Jensson | 102 |
| <i>Assisting Undergraduate Students in Writing Spanish Methodology Sections</i> Samuel González-López, Steven Bethard and Aurelio Lopez-Lopez | 115 |
| <i>Applications of Natural Language Processing in Bilingual Language Teaching: An Indonesian-English Case Study</i> Zara Maxwell-Smith, Simón González Ochoa, Ben Foley and Hanna Suominen | 124 |
| <i>An empirical investigation of neural methods for content scoring of science explanations</i> Brian Riordan, Sarah Bichler, Allison Bradford, Jennifer King Chen, Korah Wiley, Libby Gerard and Marcia C. Linn | 135 |
| <i>An Exploratory Study of Argumentative Writing by Young Students: A transformer-based Approach</i> Debanjan Ghosh, Beata Beigman Klebanov and Yi Song | 145 |
| <i>Should You Fine-Tune BERT for Automated Essay Scoring?</i> Elijah Mayfield and Alan W Black | 151 |

| | |
|---|-----|
| <i>GECToR – Grammatical Error Correction: Tag, Not Rewrite</i> | |
| Kostiantyn Omelianchuk, Vitaliy Atrasevych, Artem Chernodub and Oleksandr Skurzshanskyi | 163 |
| <i>Interpreting Neural CWI Classifiers’ Weights as Vocabulary Size</i> | |
| Yo Ehara | 171 |
| <i>Automated Scoring of Clinical Expressive Language Evaluation Tasks</i> | |
| Yiyi Wang, Emily Prud’hommeaux, Meysam Asgari and Jill Dolata | 177 |
| <i>Context-based Automated Scoring of Complex Mathematical Responses</i> | |
| Aoife Cahill, James H Fife, Brian Riordan, Avijit Vajpayee and Dmytro Galochkin | 186 |
| <i>Predicting the Difficulty and Response Time of Multiple Choice Questions Using Transfer Learning</i> | |
| Kang Xue, Victoria Yaneva, Christopher Runyon and Peter Baldwin | 193 |
| <i>A Comparative Study of Synthetic Data Generation Methods for Grammatical Error Correction</i> | |
| Max White and Alla Rozovskaya | 198 |