

2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (AAACL-IJCNLP 2022)

Short Papers

Online
20-23 November 2022

ISBN: 978-1-7138-6310-6

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© (2022) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2023)

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>Transfer Learning for Humor Detection by Twin Masked Yellow Muppets</i> Aseem Arora, Gaël Dias, Adam Jatowt and Asif Ekbal	1
<i>A Unified Model for Reverse Dictionary and Definition Modelling</i> Pinzhen Chen and Zheng Zhao	8
<i>Benchmarking the Covariate Shift Robustness of Open-world Intent Classification Approaches</i> Sopan Khosla and Rashmi Gangadharaiyah	14
<i>Number Theory Meets Linguistics: Modelling Noun Pluralisation Across 1497 Languages Using 2-adic Metrics</i> Gregory Baker and Diego Molla	24
<i>CLIP4IDC: CLIP for Image Difference Captioning</i> Zixin Guo, Tzu-Jui Wang and Jorma Laaksonen	33
<i>Towards Modeling Role-Aware Centrality for Dialogue Summarization</i> Xinnian Liang, Chao Bian, Shuangzhi Wu and Zhoujun Li	43
<i>Robust Hate Speech Detection via Mitigating Spurious Correlations</i> Kshitiz Tiwari, Shuhan Yuan and Lu Zhang	51
<i>FAD-X: Fusing Adapters for Cross-lingual Transfer to Low-Resource Languages</i> Jaeseong Lee, Seung-won Hwang and Taesup Kim	57
<i>Combining Argumentation Structure and Language Model for Generating Natural Argumentative Dialogue</i> Koh Mitsuda, Ryuichiro Higashinaka and Kuniko Saito	65
<i>Every word counts: A multilingual analysis of individual human alignment with model attention</i> Stephanie Brandl and Nora Hollenstein	72
<i>Analyzing Biases to Spurious Correlations in Text Classification Tasks</i> Adian Liusie, Vatsal Raina, Vyas Raina and Mark Gales	78
<i>BERTSeg: BERT Based Unsupervised Subword Segmentation for Neural Machine Translation</i> Haiyue Song, Raj Dabre, Zhuoyuan Mao, Chenhui Chu and Sadao Kurohashi	85
<i>NERDz: A Preliminary Dataset of Named Entities for Algerian</i> Samia Touileb	95
<i>An Effective Post-training Embedding Binarization Approach for Fast Online Top-K Passage Matching</i> Yankai Chen, Yifei Zhang, Huifeng Guo, Ruiming Tang and Irwin King	102
<i>Addressing Segmentation Ambiguity in Neural Linguistic Steganography</i> Jumon Nozaki and Yugo Murawaki	109
<i>Parsing linearizations appreciate PoS tags - but some are fussy about errors</i> Alberto Muñoz-Ortiz, Mark Anderson, David Vilares and Carlos Gómez-Rodríguez	117
<i>EmoNoBa: A Dataset for Analyzing Fine-Grained Emotions on Noisy Bangla Texts</i> Khondoker Ittehadul Islam, Tanvir Yuvraz, Md Saiful Islam and Enamul Hassan	128

<i>Exploring Universal Sentence Encoders for Zero-shot Text Classification</i>	
Souvika Sarkar, Dongji Feng and Shubhra Kanti Karmaker Santu	135
<i>The Effects of Language Token Prefixing for Multilingual Machine Translation</i>	
Rachel Wicks and Kevin Duh	148
<i>How Relevant is Selective Memory Population in Lifelong Language Learning?</i>	
Vladimir Araujo, Helena Balabin, Julio Hurtado, Alvaro Soto and Marie-Francine Moens	154
<i>An Improved Baseline for Sentence-level Relation Extraction</i>	
Wenxuan Zhou and Muhao Chen	161
<i>Multi-Type Conversational Question-Answer Generation with Closed-ended and Unanswerable Questions</i>	
Seonjeong Hwang, Yunsu Kim and Gary Geunbae Lee	169
<i>Improving Chinese Story Generation via Awareness of Syntactic Dependencies and Semantics</i>	
Henglin Huang, Chen Tang, Tyler Loakman, Frank Guerin and Chenghua Lin	178
<i>NGEP: A Graph-based Event Planning Framework for Story Generation</i>	
Chen Tang, Zhihao Zhang, Tyler Loakman, Chenghua Lin and Frank Guerin	186
<i>A Simple Yet Effective Hybrid Pre-trained Language Model for Unsupervised Sentence Acceptability Prediction</i>	
Yang Zhao and Issei Yoshida	194
<i>Post-Training with Interrogative Sentences for Enhancing BART-based Korean Question Generator</i>	
Gyu-Min Park, Seong-Eun Hong and Seong-Bae Park	202
<i>Do ever larger octopi still amplify reporting biases? Evidence from judgments of typical colour</i>	
Fangyu Liu, Julian Eisenschlos, Jeremy Cole and Nigel Collier	210
<i>Adversarially Improving NMT Robustness to ASR Errors with Confusion Sets</i>	
Shuaibo Wang, Yufeng Chen, Songming Zhang, Deyi Xiong and Jinan Xu	221
<i>Improving Graph-Based Text Representations with Character and Word Level N-grams</i>	
Wenzhe Li and Nikolaos Aletras	228
<i>Risk-graded Safety for Handling Medical Queries in Conversational AI</i>	
Gavin Abercrombie and Verena Rieser	234
<i>Performance-Efficiency Trade-Offs in Adapting Language Models to Text Classification Tasks</i>	
Laura Aina, Nikos Voskarides and Roi Blanco	244
<i>Seeking Diverse Reasoning Logic: Controlled Equation Expression Generation for Solving Math Word Problems</i>	
Yibin Shen, Qianying Liu, Zhuoyuan Mao, Zhen Wan, Fei Cheng and Sadao Kurohashi	254
<i>BanglaParaphrase: A High-Quality Bangla Paraphrase Dataset</i>	
Ajwad Akil, Najrin Sultana, Abhik Bhattacharjee and Rifat Shahriyar	261
<i>NepBERTa: Nepali Language Model Trained in a Large Corpus</i>	
Sulav Timilsina, Milan Gautam and Binod Bhattarai	273
<i>Local Structure Matters Most in Most Languages</i>	
Louis Clouatre, Prasanna Parthasarathi, Amal Zouaq and Sarath Chandar	285

<i>Transformer-based Localization from Embodied Dialog with Large-scale Pre-training</i> Meera Hahn and James M. Rehg	295
<i>CSS: Combining Self-training and Self-supervised Learning for Few-shot Dialogue State Tracking</i> Haoning Zhang, Junwei Bao, Haipeng Sun, Huaishao Luo, Wenye Li and Shuguang Cui	302
<i>Demographic-Aware Language Model Fine-tuning as a Bias Mitigation Technique</i> Aparna Garimella, Rada Mihalcea and Akhash Amarnath	311
<i>Towards Simple and Efficient Task-Adaptive Pre-training for Text Classification</i> Arnav Ladkat, Aamir Miyajiwala, Samiksha Jagadale, Rekha A. Kulkarni and Raviraj Joshi ..	320
<i>Extractive Entity-Centric Summarization as Sentence Selection using Bi-Encoders</i> Ella Hofmann-Coyle, Mayank Kulkarni, Lingjue Xie, Mounica Maddela and Daniel Preotiu-Pietro	326
<i>Towards Unsupervised Morphological Analysis of Polysynthetic Languages</i> Sujay Khandagale, Yoann Léveillé, Samuel Miller, Derek Pham, Ramy Eskander, Cass Lowry, Richard Compton, Judith Klavans, Maria Polinsky and Smaranda Muresan	334
<i>Self-Repetition in Abstractive Neural Summarizers</i> Nikita Salkar, Thomas Trikalinos, Byron Wallace and Ani Nenkova	341
<i>Domain Specific Sub-network for Multi-Domain Neural Machine Translation</i> Amr Hendy, Mohamed Abdelghaffar, Mohamed Afify and Ahmed Y. Tawfik	351
<i>Modeling Document-level Temporal Structures for Building Temporal Dependency Graphs</i> Prafulla Kumar Choubey and Ruihong Huang	357
<i>Evaluating Pre-Trained Sentence-BERT with Class Embeddings in Active Learning for Multi-Label Text Classification</i> Lukas Wertz, Jasmina Bogojeska, Katsiaryna Mirylenka and Jonas Kuhn	366
<i>MiQA: A Benchmark for Inference on Metaphorical Questions</i> Iulia Comşa, Julian Eisenschlos and Srinu Narayanan	373
<i>GCDT: A Chinese RST Treebank for Multigenre and Multilingual Discourse Parsing</i> Siyao Peng, Yang Janet Liu and Amir Zeldes	382
<i>Assessing Combinational Generalization of Language Models in Biased Scenarios</i> Yanbo Fang, Zuohui Fu, Xin Dong, Yongfeng Zhang and Gerard de Melo	392
<i>Controllable Text Simplification with Deep Reinforcement Learning</i> Daiki Yanamoto, Tomoki Ikawa, Tomoyuki Kajiwara, Takashi Ninomiya, Satoru Uchida and Yuki Arase	398
<i>Vector Space Interpolation for Query Expansion</i> Deepanway Ghosal, Somak Aditya, Sandipan Dandapat and Monojit Choudhury	405
<i>SchAman: Spell-Checking Resources and Benchmark for Endangered Languages from Amazonia</i> Arturo Oncevay, Gerardo Cardoso, Carlo Alva, César Lara Ávila, Jovita Vásquez Balarezo, Saúl Escobar Rodríguez, Delio Siticonatzi Camaiteri, Esaú Zumaeta Rojas, Didier López Francis, Juan López Bautista, Nimia Acho Rios, Remigio Zapata Cesareo, Héctor Erasmo Gómez Montoya and Roberto Zariquiey	411

<i>CoFE: A New Dataset of Intra-Multilingual Multi-target Stance Classification from an Online European Participatory Democracy Platform</i>	
Valentin Barriere, Guillaume Guillaume Jacquet and Leo Hemamou	418
<i>Exploring the Effects of Negation and Grammatical Tense on Bias Probes</i>	
Samia Touileb	423
<i>Promoting Pre-trained LM with Linguistic Features on Automatic Readability Assessment</i>	
Shudi Hou, Simin Rao, Yu Xia and Sujian Li	430
<i>An Empirical Study of Pipeline vs. Joint approaches to Entity and Relation Extraction</i>	
Zhaohui Yan, Zixia Jia and Kewei Tu	437
<i>CLASP: Few-Shot Cross-Lingual Data Augmentation for Semantic Parsing</i>	
Andy Rosenbaum, Saleh Soltan, Wael Hamza, Marco Damonte, Isabel Groves and Amir Saffari	444
<i>Plug and Play Knowledge Distillation for kNN-LM with External Logits</i>	
Xuyang Jin, Tao Ge and Furu Wei	463
<i>How Well Do Multi-hop Reading Comprehension Models Understand Date Information?</i>	
Xanh Ho, Saku Sugawara and Akiko Aizawa	470
<i>Dodging the Data Bottleneck: Automatic Subtitling with Automatically Segmented ST Corpora</i>	
Sara Papi, Alina Karakanta, Matteo Negri and Marco Turchi	480
<i>How to tackle an emerging topic? Combining strong and weak labels for Covid news NER</i>	
Aleksander Ficek, Fangyu Liu and Nigel Collier	488