

Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2022)

Student Research Workshop

Online
10 - 15 July 2022

ISBN: 978-1-7138-5621-4

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. (2022)

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>Systematicity Emerges in Transformers when Abstract Grammatical Roles Guide Attention</i> Ayush K Chakravarthy, Jacob Labe Russin and Randall O’Reilly	1
<i>Grounding in social media: An approach to building a chit-chat dialogue model</i> Ritvik Choudhary and Daisuke Kawahara	9
<i>ExtraPhrase: Efficient Data Augmentation for Abstractive Summarization</i> Mengsay Loem, Sho Takase, Masahiro Kaneko and Naoaki Okazaki	16
<i>Regularized Training of Nearest Neighbor Language Models</i> Jean-Francois Ton, Walter Talbott, Shuangfei Zhai and Joshua M. Susskind	25
<i>Again, Dozens of Refugees Drowned: A Computational Study of Political Framing Evoked by Presuppositions</i> Qi Yu	31
<i>Methods for Estimating and Improving Robustness of Language Models</i> Michal Stefanik	44
<i>Retrieval-augmented Generation across Heterogeneous Knowledge</i> Wenhao Yu	52
<i>Neural Retriever and Go Beyond: A Thesis Proposal</i> Man Luo	59
<i>Improving Classification of Infrequent Cognitive Distortions: Domain-Specific Model vs. Data Augmentation</i> Xiruo Ding, Kevin Lybarger, Justin Tauscher and Trevor Cohen	68
<i>Generate, Evaluate, and Select: A Dialogue System with a Response Evaluator for Diversity-Aware Response Generation</i> Ryoma Sakaeda and Daisuke Kawahara	76
<i>Impact of Training Instance Selection on Domain-Specific Entity Extraction using BERT</i> Eileen Salhofer, Xing Lan Liu and Roman Kern	83
<i>Analysing the Correlation between Lexical Ambiguity and Translation Quality in a Multimodal Setting using WordNet</i> Ali Hatami, Paul Buitelaar and Mihael Arcan	89
<i>Building a Personalized Dialogue System with Prompt-Tuning</i> Tomohito Kasahara, Daisuke Kawahara, Nguyen Tung, Shengzhe Li, Kenta Shinzato and Toshi-nori Sato	96
<i>MM-GATBT: Enriching Multimodal Representation Using Graph Attention Network</i> Seung Byum Seo, Hyoungwook Nam and Payam Delgosha	106
<i>Simulating Feature Structures with Simple Types</i> Valentin D. Richard	113
<i>Dr. Livingstone, I presume? Polishing of foreign character identification in literary texts</i> Aleksandra Konovalova, Antonio Toral and Kristiina Taivalkoski-Shilov	123

<i>Zuo Zhuan Ancient Chinese Dataset for Word Sense Disambiguation</i>	
Xiaomeng Pan, Hongfei Wang, Teruaki Oka and Mamoru Komachi	129
<i>ViT5: Pretrained Text-to-Text Transformer for Vietnamese Language Generation</i>	
Long Phan, Hieu Tran, Hieu Nguyen and Trieu H. Trinh	136
<i>Compositional Generalization in Grounded Language Learning via Induced Model Sparsity</i>	
Sam Spilsbury and Alexander Ilin	143
<i>How do people talk about images? A study on open-domain conversations with images.</i>	
Yi-Pei Chen, Nobuyuki Shimizu, Takashi Miyazaki and Hideki Nakayama	156
<i>Text Style Transfer for Bias Mitigation using Masked Language Modeling</i>	
Ewoenam Kwaku Tokpo and Toon Calders	163
<i>Differentially Private Instance Encoding against Privacy Attacks</i>	
Shangyu Xie and Yuan Hong	172
<i>A Simple Approach to Jointly Rank Passages and Select Relevant Sentences in the OBQA Context</i>	
Man Luo, Shuguang Chen and Chitta Baral	181
<i>Multimodal Modeling of Task-Mediated Confusion</i>	
Camille Mince, Skye Rhomberg, Cecilia Alm, Reynold Bailey and Alex Ororbia	188
<i>Probe-Less Probing of BERT’s Layer-Wise Linguistic Knowledge with Masked Word Prediction</i>	
Tatsuya Aoyama and Nathan Schneider	195
<i>Multimodal large language models for inclusive collaboration learning tasks</i>	
Armanda Lewis	202
<i>Neural Networks in a Product of Hyperbolic Spaces</i>	
Jun Takeuchi, Noriki Nishida and Hideki Nakayama	211
<i>Explicit Use of Topicality in Dialogue Response Generation</i>	
Takumi Yoshikoshi, Hayato Atarashi, Takashi Kodama and Sadao Kurohashi	222
<i>Automating Human Evaluation of Dialogue Systems</i>	
Sujan Reddy A	229
<i>Strong Heuristics for Named Entity Linking</i>	
Marko Čuljak, Andreas Spitz, Robert West and Akhil Arora	235
<i>Static and Dynamic Speaker Modeling based on Graph Neural Network for Emotion Recognition in Conversation</i>	
Prakhar Saxena, Yin Jou Huang and Sadao Kurohashi	247
<i>Few-shot fine-tuning SOTA summarization models for medical dialogues</i>	
David Fraile Navarro, Mark Dras and Shlomo Berkovsky	254
<i>Unifying Parsing and Tree-Structured Models for Generating Sentence Semantic Representations</i>	
Antoine Simoulin and Benoit Crabbé	267
<i>Multiformer: A Head-Configurable Transformer-Based Model for Direct Speech Translation</i>	
Gerard Sant, Gerard I. Gállego, Belen Alastruey and Marta Ruiz Costa-jussà	277
<i>Defending Compositionality in Emergent Languages</i>	
Michal Auersperger and Pavel Pecina	285
<i>Exploring the Effect of Dialect Mismatched Language Models in Telugu Automatic Speech Recognition</i>	
Aditya Yadavalli, Ganesh Sai Mirishkar and Anil Vuppala	292