

Third Workshop on Insights from Negative Results in NLP (Insights 2022)

Dublin, Ireland
26 May 2022

ISBN: 978-1-7138-6741-8

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© (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>On Isotropy Calibration of Transformer Models</i> Yue Ding, Karolis Martinkus, Damian Pascual, Simon Clematide and Roger Wattenhofer	1
<i>Do Dependency Relations Help in the Task of Stance Detection?</i> Alessandra Teresa Cignarella, Cristina Bosco and Paolo Rosso	10
<i>Evaluating the Practical Utility of Confidence-score based Techniques for Unsupervised Open-world Classification</i> Sopan Khosla and Rashmi Gangadharaiah	18
<i>Extending the Scope of Out-of-Domain: Examining QA models in multiple subdomains</i> Chenyang Lyu, Jennifer Foster and Yvette Graham	24
<i>What Do You Get When You Cross Beam Search with Nucleus Sampling?</i> Uri Shaham and Omer Levy	38
<i>How Much Do Modifications to Transformer Language Models Affect Their Ability to Learn Linguistic Knowledge?</i> Simeng Sun, Brian Dillon and Mohit Iyyer	46
<i>Cross-lingual Inflection as a Data Augmentation Method for Parsing</i> Alberto Muñoz-Ortiz, Carlos Gómez-Rodríguez and David Vilares	54
<i>Is BERT Robust to Label Noise? A Study on Learning with Noisy Labels in Text Classification</i> Dawei Zhu, Michael A. Hedderich, Fangzhou Zhai, David Ifeoluwa Adelani and Dietrich Klakow 62	
<i>Ancestor-to-Creole Transfer is Not a Walk in the Park</i> Heather Lent, Emanuele Bugliarello and Anders Søgaard	68
<i>What GPT Knows About Who is Who</i> Xiaohan Yang, Eduardo Peynetti, Vasco Meerman and Chris Tanner	75
<i>Evaluating Biomedical Word Embeddings for Vocabulary Alignment at Scale in the UMLS Metathesaurus Using Siamese Networks</i> Goonmeet Bajaj, Vinh Nguyen, Thilini Wijesiriwardene, Hong Yung Yip, Vishesh Javangula, Amit P. Sheth, Srinivasan Parthasarathy and Olivier Bodenreider	82
<i>On the Impact of Data Augmentation on Downstream Performance in Natural Language Processing</i> Itsuki Okimura, Machel Reid, Makoto Kawano and Yutaka Matsuo	88
<i>Can Question Rewriting Help Conversational Question Answering?</i> Etsuko Ishii, Yan Xu, Samuel Cahyawijaya and Bryan Wilie	94
<i>Clustering Examples in Multi-Dataset Benchmarks with Item Response Theory</i> Pedro Rodriguez, Phu Mon Htut, John P. Lalor and João Sedoc	100
<i>On the Limits of Evaluating Embodied Agent Model Generalization Using Validation Sets</i> Hyoungun Kim, Aishwarya Padmakumar, Di Jin, Mohit Bansal and Dilek Hakkani-Tur	113
<i>Do Data-based Curricula Work?</i> Maxim K. Surkov, Vladislav D. Mosin and Ivan P. Yamshchikov	119

<i>The Document Vectors Using Cosine Similarity Revisited</i> Zhang Bingyu and Nikolay Arefyev	129
<i>Challenges in including extra-linguistic context in pre-trained language models</i> Ionut Teodor Sorodoc, Laura Aina and Gemma Boleda	134
<i>Label Errors in BANKING77</i> Cecilia Ying and Stephen Thomas	139
<i>Pathologies of Pre-trained Language Models in Few-shot Fine-tuning</i> Hanjie Chen, Guoqing Zheng, Ahmed Hassan Awadallah and Yangfeng Ji	144
<i>An Empirical study to understand the Compositional Prowess of Neural Dialog Models</i> Vinayshekhar Bannihatti Kumar, Vaibhav Kumar, Mukul Bhutani and Alexander Rudnicky ..	154
<i>Combining Extraction and Generation for Constructing Belief-Consequence Causal Links</i> Maria Alexeeva, Allegra A. Beal A. Beal and Mihai Surdeanu	159
<i>Replicability under Near-Perfect Conditions – A Case-Study from Automatic Summarization</i> Margot Mieskes	165
<i>BPE beyond Word Boundary: How NOT to use Multi Word Expressions in Neural Machine Translation</i> Dipesh Kumar and Avijit Thawani	172
<i>Pre-trained language models evaluating themselves - A comparative study</i> Philipp Koch, Matthias Aßenmacher and Christian Heumann	180