

# **First Workshop on Insights from Negative Results in NLP (Insights 2020)**

Online  
19 November 2020

ISBN: 978-1-7138-1996-7

**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. (2021)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Domain adaptation challenges of BERT in tokenization and sub-word representations of Out-of-Vocabulary words</i> Anmol Nayak, Hariprasad Timmapathini, Karthikeyan Ponnalagu and Vijendran Gopalan Venkoparao	1
<i>Q. Can Knowledge Graphs be used to Answer Boolean Questions? A. It's complicated!</i> Daria Dzendzik, Carl Vogel and Jennifer Foster	6
<i>How Far Can We Go with Data Selection? A Case Study on Semantic Sequence Tagging Tasks</i> Samuel Louvan and Bernardo Magnini	15
<i>Evaluating the Effectiveness of Efficient Neural Architecture Search for Sentence-Pair Tasks</i> Ansel MacLaughlin, Jwala Dhamala, Anoop Kumar, Sriram Venkatapathy, Ragav Venkatesan and Rahul Gupta	22
<i>Which Matters Most? Comparing the Impact of Concept and Document Relationships in Topic Models</i> Silvia Terragni, Debora Nozza, Elisabetta Fersini and Messina Enza	32
<i>On Task-Level Dialogue Composition of Generative Transformer Model</i> Prasanna Parthasarathi, Sharan Narang and Arvind Neelakantan	41
<i>How Effectively Can Machines Defend Against Machine-Generated Fake News? An Empirical Study</i> Meghana Moorthy Bhat and Srinivasan Parthasarathy	47
<i>Label Propagation-Based Semi-Supervised Learning for Hate Speech Classification</i> Ashwin Geet D'Sa, Irina Illina, Dominique Fohr, Dietrich Klakow and Dana Ruiters	53
<i>Layout-Aware Text Representations Harm Clustering Documents by Type</i> Catherine Finegan-Dollak and Ashish Verma	59
<i>An Analysis of Capsule Networks for Part of Speech Tagging in High- and Low-resource Scenarios</i> Andrew Zupon, Faiz Rafique and Mihai Surdeanu	65
<i>Can Knowledge Graph Embeddings Tell Us What Fact-checked Claims Are About?</i> Valentina Beretta, Sébastien Harispe, Katarina Boland, Luke Lo Seen, Konstantin Todorov and Andon Tchechmedjiev	70
<i>Do Transformers Dream of Inference, or Can Pretrained Generative Models Learn Implicit Inferential Rules?</i> Zhengzhong Liang and Mihai Surdeanu	75
<i>Counterfactually-Augmented SNLI Training Data Does Not Yield Better Generalization Than Unaugmented Data</i> William Huang, Haokun Liu and Samuel R. Bowman	81
<i>NMF Ensembles? Not for Text Summarization!</i> Alka Khurana and Vasudha Bhatnagar	87
<i>If You Build Your Own NER Scorer, Non-replicable Results Will Come</i> Constantine Lignos and Marjan Kamyab	93

<i>HINT3: Raising the bar for Intent Detection in the Wild</i>	
Gaurav Arora, Chirag Jain, Manas Chaturvedi and Krupal Modi . . . . .	99
<i>The Extraordinary Failure of Complement Coercion Crowdsourcing</i>	
Yanai Elazar, Victoria Basmov, Shauli Ravfogel, Yoav Goldberg and Reut Tsarfaty . . . . .	105
<i>Embedding Structured Dictionary Entries</i>	
Steven Wilson, Walid Magdy, Barbara McGillivray and Gareth Tyson . . . . .	116