

# **5th Workshop on Representation Learning for NLP (RepL4NLP-2020)**

Online  
9 July 2020

ISBN: 978-1-7138-1389-7

**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. (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](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>Zero-Resource Cross-Domain Named Entity Recognition</i> Zihan Liu, Genta Indra Winata and Pascale Fung .....	1
<i>Encodings of Source Syntax: Similarities in NMT Representations Across Target Languages</i> Tyler A. Chang and Anna Rafferty .....	7
<i>Learning Probabilistic Sentence Representations from Paraphrases</i> Mingda Chen and Kevin Gimpel .....	17
<i>Word Embeddings as Tuples of Feature Probabilities</i> Siddharth Bhat, Alok Debnath, Souvik Banerjee and Manish Shrivastava .....	24
<i>Compositionality and Capacity in Emergent Languages</i> Abhinav Gupta, Cinjon Resnick, Jakob Foerster, Andrew Dai and Kyunghyun Cho .....	34
<i>Learning Geometric Word Meta-Embeddings</i> Pratik Jawanpuria, Satya Dev N T V, Anoop Kunchukuttan and Bamdev Mishra .....	39
<i>Improving Bilingual Lexicon Induction with Unsupervised Post-Processing of Monolingual Word Vector Spaces</i> Ivan Vulić, Anna Korhonen and Goran Glavaš .....	45
<i>Adversarial Training for Commonsense Inference</i> Lis Pereira, Xiaodong Liu, Fei Cheng, Masayuki Asahara and Ichiro Kobayashi .....	55
<i>Evaluating Natural Alpha Embeddings on Intrinsic and Extrinsic Tasks</i> Riccardo Volpi and Luigi Malagò .....	61
<i>Exploring the Limits of Simple Learners in Knowledge Distillation for Document Classification with DocBERT</i> Ashutosh Adhikari, Achyudh Ram, Raphael Tang, William L. Hamilton and Jimmy Lin .....	72
<i>Joint Training with Semantic Role Labeling for Better Generalization in Natural Language Inference</i> Cemil Cengiz and Deniz Yuret .....	78
<i>A Metric Learning Approach to Misogyny Categorization</i> Juan Manuel Coria, Sahar Ghannay, Sophie Rosset and Hervé Bredin .....	89
<i>On the Choice of Auxiliary Languages for Improved Sequence Tagging</i> Lukas Lange, Heike Adel and Jannik Strötgen .....	95
<i>Adversarial Alignment of Multilingual Models for Extracting Temporal Expressions from Text</i> Lukas Lange, Anastasiia Iurshina, Heike Adel and Jannik Strötgen .....	103
<i>Contextual and Non-Contextual Word Embeddings: an in-depth Linguistic Investigation</i> Alessio Miaschi and Felice Dell’Orletta .....	110
<i>Are All Languages Created Equal in Multilingual BERT?</i> Shijie Wu and Mark Dredze .....	120
<i>Staying True to Your Word: (How) Can Attention Become Explanation?</i> Martin Tutek and Jan Snajder .....	131

<i>Compressing BERT: Studying the Effects of Weight Pruning on Transfer Learning</i> Mitchell Gordon, Kevin Duh and Nicholas Andrews .....	143
<i>On Dimensional Linguistic Properties of the Word Embedding Space</i> Vikas Raunak, Vaibhav Kumar, Vivek Gupta and Florian Metze .....	156
<i>A Cross-Task Analysis of Text Span Representations</i> Shubham Toshniwal, Haoyue Shi, Bowen Shi, Lingyu Gao, Karen Livescu and Kevin Gimpel .	166
<i>Enhancing Transformer with Sememe Knowledge</i> Yuhui Zhang, Chenghao Yang, Zhengping Zhou and Zhiyuan Liu .....	177
<i>Evaluating Compositionality of Sentence Representation Models</i> Hanoz Bhathena, Angelica Willis and Nathan Dass .....	185
<i>Supertagging with CCG primitives</i> Aditya Bhargava and Gerald Penn .....	194
<i>What's in a Name? Are BERT Named Entity Representations just as Good for any other Name?</i> Sriram Balasubramanian, Naman Jain, Gaurav Jindal, Abhijeet Awasthi and Sunita Sarawagi .	205