

9th Workshop on Representation Learning for NLP (RepL4NLP-2024)

Bangkok, Thailand
15 August 2024

ISBN: 979-8-3313-0174-3

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© (2024) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

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>Learning Contextualized Box Embeddings with Prototypical Networks</i>	
Kohei Oda, Kiyooki Shirai and Natthawut Kertkeidkachorn	1
<i>DomainInv: Domain Invariant Fine Tuning and Adversarial Label Correction For Unsupervised QA Domain Adaptation</i>	
Anant Khandelwal	13
<i>Relevance-aware Diverse Query Generation for Out-of-domain Text Ranking</i>	
Jia-Huei Ju, Chao-Han Huck Yang, Szu-Wei Fu, Ming-Feng Tsai and Chuan-Ju Wang	26
<i>Learning from Others: Similarity-based Regularization for Mitigating Dataset Bias.</i>	
Reda Igbaria and Yonatan Belinkov	37
<i>Prior Knowledge-Guided Adversarial Training</i>	
Lis Pereira, Fei Cheng, Wan Jou She, Masayuki Asahara and Ichiro Kobayashi	51
<i>IT-Tuning : Parameter Efficient Information Token Tuning for Language Model</i>	
Jungu Kim and Hyeoncheol Kim	58
<i>Bridging the Gap: Transfer Learning from English PLMs to Malaysian English</i>	
Mohan Raj, Lay-Ki Soon, Huey Fang Ong and Bhawani Selvaretnam	69
<i>Unified Interpretation of Smoothing Methods for Negative Sampling Loss Functions in Knowledge Graph Embedding</i>	
Xincan Feng, Hidetaka Kamigaito, Katsuhiko Hayashi and Taro Watanabe	78
<i>How Useful is Continued Pre-Training for Generative Unsupervised Domain Adaptation?</i>	
Rheeya Uppaal, Yixuan Li and Junjie Hu	99
<i>Word Boundary Information Isn't Useful for Encoder Language Models</i>	
Edward Gow-Smith, Dylan Phelps, Harish Tayyar Madabushi, Carolina Scarton and Aline Villavicencio	118
<i>Beyond Link Prediction: On Pre-Training Knowledge Graph Embeddings</i>	
Daniel Ruffinelli and Rainer Gemulla	136
<i>Learn it or Leave it: Module Composition and Pruning for Continual Learning</i>	
Mingyang Wang, Heike Adel, Lukas Lange, Jannik Strötgen and Hinrich Schuetze	163
<i>Text-Guided Alternative Image Clustering</i>	
Andreas Stephan, Lukas Miklautz, Collin Leiber, Pedro Henrique Luz De Araujo, Dominik Répás, Claudia Plant and Benjamin Roth	177
<i>QAVSA: Question Answering using Vector Symbolic Algebras</i>	
Ryan Laube and Chris Eliasmith	191
<i>Tracking linguistic information in transformer-based sentence embeddings through targeted sparsification</i>	
Vivi Nastase and Paola Merlo	203
<i>Learning New Tasks from a Few Examples with Soft-Label Prototypes</i>	
Avyav Kumar Singh, Ekaterina Shutova and Helen Yannakoudakis	215

<i>Learned Transformer Position Embeddings Have a Low-Dimensional Structure</i>	
Ulme Wennberg and Gustav Eje Henter	237
<i>Multi-label Learning with Random Circular Vectors</i>	
Ken Nishida, Kojiro Machi, Kazuma Onishi, Katsuhiko Hayashi and Hidetaka Kamigaito . . .	245
<i>Mitigating Semantic Leakage in Cross-lingual Embeddings via Orthogonality Constraint</i>	
Dayeon Ki, Cheonbok Park and Hyunjoong Kim	256
<i>On the Pathological Path-star Task for Language Models (Extended Abstract)</i>	
Arvid Frydenlund	274
<i>Whitening Not Recommended for Classification Tasks in LLMs</i>	
Ali Forooghi, Shaghayegh Sadeghi and Jianguo Lu	285
<i>LLM Circuit Analyses Are Consistent Across Training and Scale</i>	
Curt Tigges, Michael Hanna, Qinan Yu and Stella Biderman	290