ACL 2024

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

Proceedings of the Workshop

August 15, 2024

©2024 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-157-5

Introduction

The 9th Workshop on Representation Learning for NLP, organized by SIGREP, aims to continue the success of the first eight workshops in the series. The RepL4NLP workshop was introduced as a synthesis of several years of independent *CL workshops focusing on vector space models of meaning, compositionality, and the application of deep neural networks and spectral methods to NLP. With the widespread adoption of neural network architectures (and especially Transformer models), representation learning has become a central concept in NLP research. RepL4NLP provides a forum for discussing recent advances on these topics, ranging from the development of new representations for various tasks and applications to the analysis of existing representations, e.g., with respect to generalization and robustness.

In most NLP applications, the goal is to understand, interpret, and generate human language text, extract linguistic information from raw text or to transform linguistic observations into an alternative form, e.g., from speech to text or from one language to another. Traditional statistical models relied heavily on discrete categories that are parameterized individually (e.g., each word or n-gram in a language model gets its own parameter); Representation learning—whether using spectral methods, probabilistic models, or deep neural networks—offers an alternative: by learning continuous, multidimensional representations of discrete objects (e.g., words, tags, labels, phrases, sentences), models can find representations that enable them to solve tasks more effectively. As learning representations has now been adopted as the de-facto approach to building models in NLP, their widespread use raises important questions towards further understanding them.

Organizing Committee

Workshop Organizers

Chen Zhao, New York University Shanghai
Marius Mosbach, Mila - McGill University
Pepa Atanasova, University of Copenhagen
Seraphina Goldfarb-Tarrent, Cohere
Peter Hase, University of North Carolina at Chapel Hill
Arian Hosseini, Mila - University of Montreal
Maha Elbayad, Meta AI
Sandro Pezzelle, University of Amsterdam
Maximilian Mozes, Cohere

Program Committee

Reviewers

Taichi Aida, Tokyo Metropolitan University
Gábor Berend, University of Szeged
Ondřej Bojar, Charles University Prague
Bridger Bridger, Holistic Intelligence for Global Good
Necva Bölücü, CSIRO
Muthu Kumar Chandrasekaran, Accenture
Anni Chen
Danlu Chen
Lin Chen, Facebook
Yue Chen, Microsoft and Indiana University
Catherine Chen, University of California Berkeley
Minhao Cheng, Pennsylvania State University
Heeyoul Choi, Handong Global University
Manuel Rafael Ciosici, USC/ISI
Wenliang Dai, NVIDIA
Christopher Davis, University of Cambridge
Vladimir Eidelman, FiscalNote, Inc.
Maha Elbayad, FAIR
Sergey Feldman, Allen Institute for Artificial Intelligence and Data Cowboys
Matthias Gallé, Cohere
Félix Gaschi, University of Lorraine
Shahriar Golchin, University of Arizona
Ashim Gupta
Gustavo Hernandez Abrego, Google
Julia Hockenmaier, University of Illinois, Urbana Champaign, University of Illinois, Urbana Champaign and University of Illinois, Urbana Champaign
Valentin Hofmann, Allen Institute for Artificial Intelligence
Dae Yon Hwang, Amazon AGI
Katharina Hämmerl, CIS, LMU Munich
Etsuko Ishii, The Hong Kong University of Science and Technology
Hamish Ivison, University of Washington
Fran Jelenić
Josip Jukić, Faculty of Electrical Engineering and Computing, University of Zagreb
Hirotaka Kameko, Kyoto University
Jaeyoung Kim, Seoul National University
Hyunjong Kim
Shankar Kumar
Matthieu Labeau, Télécom ParisTech
Seungil Chad Lee
Fei-Tzin Lee, Columbia University
Yitong Li, Huawei Technologies Co., Ltd.
Tao Li, Google DeepMind
Kai Liu, Clemson University
Zhu Liu
Chen Cecilia Liu
Holy Lovenia, AI Singapore

Yiwei Lyu
Kimberly Mai, University College London, University of London
Danil Malaev, MTS
Sheikh Mannan
Sneha Mehta
Antonio Valerio Miceli Barone, Università di Palermo, Università di Palermo, Università di Pisa and University of Edinburgh
Ashutosh Modi, IIT Kanpur
Vishvak Murahari, Princeton University
Alberto Muñoz-Ortiz
Inderjeet Jayakumar Nair, University of Michigan - Ann Arbor
Nikita Nangia
Vivi Nastase, University of Geneva
Abhijnan Nath
Chau Nguyen, University of Western Australia
Stephen Obadinma, Queen's University
Tsuyoshi Okita, Kyushu Institute of Technology
Ankur Padia, Philips Research North America
Jungsoo Park, NAVER
ChaeHun Park
Vipul Raheja, Columbia University, Grammarly and International Institute of Information Technology Hyderabad
Surangika Ranathunga, Massey University
Narutatsu Ri
Frank Rudzicz, Dalhousie University
Kaushik Ram Sadagopan, Meta
Hassan Sajjad, Dalhousie University
Avneesh Saluja, Netflix
Lingfeng Shen, ByteDance Inc.
Keisuke Shirai, Kyoto University
Shashwat Singh, International Institute of Information Technology Hyderabad
Karolina Stanczak, Mila - Quebec Artificial Intelligence Institute and McGill University, McGill University
Sho Takase, LINE Corporation
Shuai Tang, Amazon Web Services
Olga Tsymboi, Sber AI Lab and Moscow Institute of Physics and Technology
Sowmya Vajjala, National Research Council Canada
Eva Maria Vecchi
Qingyun Wang, University of Illinois, Urbana Champaign
Fanyu Wang, Monash University
Hai Wang, Samsung
Rodrigo Wilkens, UCL
Chenghao Xiao
Zhenping Xie, Jiangnan University
Zhikun Xu, Arizona State University
Tiezheng YU, The Hong Kong University of Science and Technology
Takateru Yamakoshi, The University of Tokyo
Sohee Yang, University College London, University of London, Department of Computer Science, University College London, University of London, DeepMind and Google
Majid Yazdani, Leeroo
Hee Suk Yoon, Korea Advanced Institute of Science & Technology

Eunseop Yoon, KAIST
Hong Yu, Apple
Kai Zhang
Wenbo Zhao, Amazon
Wenxuan Zhou, Zoom
Dong Zhou, Guangdong University of Foreign Studies
Vilém Zouhar, Department of Computer Science, ETHZ - ETH Zurich

Invited Speakers

Mor Geva, Tel Aviv University
Dieuwke Hupkes, Meta
Hassan Sajjad, Dalhousie University
Edoardo Ponti, University of Edinburgh

Table of Contents

Learning Contextualized Box Embeddings with Prototypical Networks
Kohei Oda, Kiyoaki Shirai and Natthawut Kertkeidkachorn 1

DomainInv: Domain Invariant Fine Tuning and Adversarial Label Correction For Unsupervised QA Domain Adaptation
Anant Khandelwal 13

Relevance-aware Diverse Query Generation for Out-of-domain Text Ranking
Jia-Huei Ju, Chao-Han Huck Yang, Szu-Wei Fu, Ming-Feng Tsai and Chuan-Ju Wang 26

Learning from Others: Similarity-based Regularization for Mitigating Dataset Bias.
Reda Igbaria and Yonatan Belinkov 37

Prior Knowledge-Guided Adversarial Training
Lis Pereira, Fei Cheng, Wan Jou She, Masayuki Asahara and Ichiro Kobayashi 51

IT-Tuning : Parameter Efficient Information Token Tuning for Language Model
Jungu Kim and Hyeoncheol Kim 58

Bridging the Gap: Transfer Learning from English PLMs to Malaysian English
Mohan Raj, Lay-Ki Soon, Huey Fang Ong and Bhawani Selvaretnam 69

Unified Interpretation of Smoothing Methods for Negative Sampling Loss Functions in Knowledge Graph Embedding
Xincan Feng, Hidetaka Kamigaito, Katsuhiko Hayashi and Taro Watanabe 78

How Useful is Continued Pre-Training for Generative Unsupervised Domain Adaptation?
Rheeya Uppaal, Yixuan Li and Junjie Hu 99

Word Boundary Information Isn't Useful for Encoder Language Models
Edward Gow-Smith, Dylan Phelps, Harish Tayyar Madabushi, Carolina Scarton and Aline Villavicencio 118

Beyond Link Prediction: On Pre-Training Knowledge Graph Embeddings
Daniel Ruffinelli and Rainer Gemulla 136

Learn it or Leave it: Module Composition and Pruning for Continual Learning
Mingyang Wang, Heike Adel, Lukas Lange, Jannik Strötgen and Hinrich Schuetze 163

Text-Guided Alternative Image Clustering
Andreas Stephan, Lukas Miklautz, Collin Leiber, Pedro Henrique Luz De Araujo, Dominik Répás, Claudia Plant and Benjamin Roth 177

QAVSA: Question Answering using Vector Symbolic Algebras
Ryan Laube and Chris Eliasmith 191

Tracking linguistic information in transformer-based sentence embeddings through targeted sparsification
Vivi Nastase and Paola Merlo 203

Learning New Tasks from a Few Examples with Soft-Label Prototypes
Avyav Kumar Singh, Ekaterina Shutova and Helen Yannakoudakis 215

Learned Transformer Position Embeddings Have a Low-Dimensional Structure
Ulme Wennberg and Gustav Eje Henter 237

Multi-label Learning with Random Circular Vectors
Ken Nishida, Kojiro Machi, Kazuma Onishi, Katsuhiko Hayashi and Hidetaka Kamigaito . . . 245

Mitigating Semantic Leakage in Cross-lingual Embeddings via Orthogonality Constraint
Dayeon Ki, Cheonbok Park and Hyunjoong Kim 256

On the Pathological Path-star Task for Language Models (Extended Abstract)
Arvid Frydenlund 274

Whitening Not Recommended for Classification Tasks in LLMs
Ali Forooghi, Shaghayegh Sadeghi and Jianguo Lu 285

LLM Circuit Analyses Are Consistent Across Training and Scale
Curt Tigges, Michael Hanna, Qinan Yu and Stella Biderman 290