ACL 2020

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

Proceedings of the Workshop

July 9, 2020

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-952148-15-6

Introduction

The 5th Workshop on Representation Learning for NLP (RepL4NLP-2020) will be hosted at ACL 2020. The workshop is being organised by Spandana Gella, Johannes Welbl, Marek Rei, Fabio Petroni, Patrick Lewis, Emma Strubell, Minjoon Seo and Hannaneh Hajishirzi; and advised by Isabelle Augenstein, Kyunghyun Cho, Edward Grefenstette, Karl Moritz Hermann, and Chris Dyer. The workshop is organised by the ACL Special Interest Group on Representation Learning (SIGREP).

The 5th Workshop on Representation Learning for NLP aims to continue the success of the 1st Workshop on Representation Learning for NLP (about 50 submissions and over 250 attendees; second most attended collocated event at ACL'16 after WMT), 2nd Workshop on Representation Learning for NLP, 3rd Workshop on Representation Learning for NLP, and 4th Workshop on Representation Learning for NLP. The 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. It provides a forum for discussing recent advances on these topics, as well as future research directions in linguistically motivated vector-based models in NLP.

Organizers:

Spandana Gella, Amazon AI
Johannes Welbl, University College London
Marek Rei, Imperial College London
Fabio Petroni, Facebook AI Research
Patrick Lewis, University College London & FAIR
Emma Strubell, Carnegie Mellon University & FAIR
Minjoon Seo, University of Washington & Naver
Hannaneh Hajishirzi, University of Washington

Senior Advisors:

Kyunghyun Cho, NYU and Facebook AI Research
Edward Grefenstette, Facebook AI Research & University College London
Karl Moritz Hermann, DeepMind
Laura Rimell, DeepMind
Chris Dyer, DeepMind
Isabelle Augenstein, University of Copenhagen

Keynote Speakers:

Kristina Toutanova, Google Research
Ellie Pavlick, Brown University & Google
Mike Lewis, Facebook AI Research
Evelina Fedorenko, Massachusetts Institute of Technology

Program Committee:

Muhammad Abdul-Mageed
Guy Aglionby
Roee Aharoni
Arjun Akula
Julio Amador Díaz López
Mikel Artetxe
Yoav Artzi
Miguel Ballesteros
Gianni Barlacchi
Max Bartolo
Joost Bastings
Federico Bianchi
Rishi Bommasani
Samuel R. Bowman
Andrew Caines
Claire Cardie
Haw-shiuan Chang

Lin Chen
Danlu Chen
Yue Chen
Yu Cheng
Manuel R. Ciosici
William Cohen
Christopher Davis
Eliezer de Souza da Silva
Luciano Del Corro
Zhi-Hong Deng
Leon Derczynski
Shehzaad Dhuliawala
Giuseppe Antonio Di Luna
Kalpit Dixit
Aleksandr Drozd
Kevin Duh
Necati Bora Edizel
Guy Emerson
Eraldo Fernandes
Orhan Firat
Rainer Gemulla
Kevin Gimpel
Hongyu Gong
Ana Valeria González
Batool Haider
He He
Ji He
Jiaji Huang
Sung Ju Hwang
Robin Jia
Mark Johnson
Arzoo Katiyar
Santosh Kesiraju
Douwe Kiela
Ekaterina Kochmar
Julia Kreutzer
Shankar Kumar
John P. Lalor
Carolin Lawrence
Kenton Lee
Xiang Li
Shaohua Li
Tao Li
Bill Yuchen Lin
Chu-Cheng Lin
Peng Liu
Feifan Liu
Fei Liu
Suresh Manandhar
Luca Massarelli

Sneha Mehta
Todor Mihaylov
Tsvetomila Mihaylova
Swaroop Mishra
Ashutosh Modi
Lili Mou
Maximilian Mozes
Khalil Mrini
Phoebe Mulcaire
Nikita Nangia
Shashi Narayan
Thien Huu Nguyen
Tsuyoshi Okita
Ankur Padia
Ashwin Paranjape
Tom Pelsmaeker
Aleksandra Piktus
Vassilis Plachouras
Edoardo Maria Ponti
Ratish Puduppully
Leonardo Querzoni
Chris Quirk
Vipul Raheja
Muhammad Rahman
Natraj Raman
Surangika Ranathunga
Siva Reddy
Sravana Reddy
Roi Reichart
Devendra Sachan
Marzieh Saeidi
Avneesh Saluja
Hinrich Schütze
Tianze Shi
Vered Shwartz
Kyungwoo Song
Daniil Sorokin
Lucia Specia
Mark Steedman
Karl Stratos
Ming Sun
Jörg Tiedemann
Ivan Titov
Nadi Tomeh
Shubham Toshniwal
Kristina Toutanova
Lifu Tu
Lyle Ungar
Menno van Zaanen
Andrea Vanzo

Shikhar Vashishth
Eva Maria Vecchi
Elena Voita
Yogarshi Vyas
Hai Wang
Bonnie Webber
Dirk Weissenborn
Rodrigo Wilkens
Yuxiang Wu
Yadollah Yaghoobzadeh
Haiqin Yang
Majid Yazdani
Wen-tau Yih
Hong Yu
Wenxuan Zhou
Dong Zhou
Xiangyang Zhou
Imed Zitouni
Diarmuid Ó Séaghdha
Robert Östling

Table of Contents

Workshop Program

Thursday, July 9, 2020

9:30–9:45 Welcome and Opening Remarks

9:45–14:45 Keynote Session

9:45–10:30 *Invited talk 1*
Kristina Toutanova

10:30–11:00 Coffee Break

11:00–11:45 *Invited talk 2*
Ellie Pavlick

11:45–12:30 *Invited talk 3*
Mike Lewis

12:30–14:00 Lunch

14:00–14:45 *Invited talk 4*
Evelina Fedorenko

14:45–15:00 Outstanding Papers Spotlight Presentations

Thursday, July 9, 2020 (continued)

15:00–16:30 Poster Session

Zero-Resource Cross-Domain Named Entity Recognition
Zihan Liu, Genta Indra Winata and Pascale Fung

Encodings of Source Syntax: Similarities in NMT Representations Across Target Languages
Tyler A. Chang and Anna Rafferty

Learning Probabilistic Sentence Representations from Paraphrases
Mingda Chen and Kevin Gimpel

On the Ability of Self-Attention Networks to Recognize Counter Languages
Satwik Bhattamishra, Kabir Ahuja and Navin Goyal

Word Embeddings as Tuples of Feature Probabilities
Siddharth Bhat, Alok Debnath, Souvik Banerjee and Manish Shrivastava

Compositionality and Capacity in Emergent Languages
Abhinav Gupta, Cinjon Resnick, Jakob Foerster, Andrew Dai and Kyunghyun Cho

Learning Geometric Word Meta-Embeddings
Pratik Jawanpuria, Satya Dev N T V, Anoop Kunchukuttan and Bamdev Mishra

Variational Inference for Learning Representations of Natural Language Edits
Edison Marrese-Taylor, Machel Reid and Yutaka Matsuo

Improving Bilingual Lexicon Induction with Unsupervised Post-Processing of Monolingual Word Vector Spaces
Ivan Vulić, Anna Korhonen and Goran Glavaš

Adversarial Training for Commonsense Inference
Lis Pereira, Xiaodong Liu, Fei Cheng, Masayuki Asahara and Ichiro Kobayashi

Evaluating Natural Alpha Embeddings on Intrinsic and Extrinsic Tasks
Riccardo Volpi and Luigi Malagò

Thursday, July 9, 2020 (continued)

Exploring the Limits of Simple Learners in Knowledge Distillation for Document Classification with DocBERT
Ashutosh Adhikari, Achyudh Ram, Raphael Tang, William L. Hamilton and Jimmy Lin

Joint Training with Semantic Role Labeling for Better Generalization in Natural Language Inference
Cemil Cengiz and Deniz Yuret

A Metric Learning Approach to Misogyny Categorization
Juan Manuel Coria, Sahar Ghannay, Sophie Rosset and Hervé Bredin

On the Choice of Auxiliary Languages for Improved Sequence Tagging
Lukas Lange, Heike Adel and Jannik Strötgen

Adversarial Alignment of Multilingual Models for Extracting Temporal Expressions from Text
Lukas Lange, Anastasiia Iurshina, Heike Adel and Jannik Strötgen

Contextual and Non-Contextual Word Embeddings: an in-depth Linguistic Investigation
Alessio Miaschi and Felice Dell'Orletta

Are All Languages Created Equal in Multilingual BERT?
Shijie Wu and Mark Dredze

Staying True to Your Word: (How) Can Attention Become Explanation?
Martin Tutek and Jan Snajder

Compressing BERT: Studying the Effects of Weight Pruning on Transfer Learning
Mitchell Gordon, Kevin Duh and Nicholas Andrews

On Dimensional Linguistic Properties of the Word Embedding Space
Vikas Raunak, Vaibhav Kumar, Vivek Gupta and Florian Metze

A Simple Approach to Learning Unsupervised Multilingual Embeddings
Pratik Jawanpuria, Mayank Meghwanshi and Bamdev

A Cross-Task Analysis of Text Span Representations
Shubham Toshniwal, Haoyue Shi, Bowen Shi, Lingyu Gao, Karen Livescu and Kevin Gimpel

Thursday, July 9, 2020 (continued)

Enhancing Transformer with Sememe Knowledge
Yuhui Zhang, Chenghao Yang, Zhengping Zhou and Zhiyuan Liu

Evaluating Compositionality of Sentence Representation Models
Hanoz Bhathena, Angelica Willis and Nathan Dass

AI4Bharat-IndicNLP Dataset: Monolingual Corpora and Word Embeddings for Indic Languages: Monolingual Corpora and Word Embeddings for Indic Languages
Anoop Kunchukuttan, Divyanshu Kakwani, Satish Golla, Gokul N.C., Avik Bhattacharyya, Mitesh M. Khapra and Pratyush Kumar

Supertagging with CCG primitives
Aditya Bhargava and Gerald Penn

What's in a Name? Are BERT Named Entity Representations just as Good for any other Name?
Sriram Balasubramanian, Naman Jain, Gaurav Jindal, Abhijeet Awasthi and Sunita Sarawagi

16:30–17:30 Panel Discussion

17:30–17:40 Closing Remarks and Best Paper Announcement