

ACL 2018

Representation Learning for NLP

Proceedings of the Third Workshop

July 20, 2018
Melbourne, Australia

Sponsors:

facebook



DeepMind



Microsoft Research

ASAPP

NAVER

©2018 The Association for Computational Linguistics

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 978-1-948087-43-8

Introduction

The ACL 2018 Workshop on Representation Learning for NLP (Repl4NLP) takes place on Friday, July 20, 2018 in Melbourne, Australia, immediately following the 56th Annual Meeting of the Association for Computational Linguistics (ACL). The workshop is generously sponsored by Facebook, Salesforce, ASAPP, DeepMind, Microsoft Research, and Naver.

Repl4NLP is organised by Isabelle Augenstein, Kris Cao, He He, Felix Hill, Spandana Gella, Jamie Kiros, Hongyuan Mei and Dipendra Misra, and advised by Kyunghyun Cho, Edward Grefenstette, Karl Moritz Hermann and Laura Rimell.

The 3rd Workshop on Representation Learning for NLP aims to continue the success of the 1st Workshop on Representation Learning for NLP, which received about 50 submissions and over 250 attendees and was the second most attended collocated event at ACL 2016 in Berlin, Germany after WMT; and the 2nd Workshop on Representation Learning for NLP at ACL 2017 in Vancouver, Canada.

The workshop has a focus 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:

Isabelle Augenstein, University of Copenhagen
Kris Cao, University of Cambridge
He He, Stanford University
Felix Hill, DeepMind
Spandana Gella, University of Edinburgh
Jamie Kiros, University of Toronto
Hongyuan Mei, Johns Hopkins University
Dipendra Misra, Cornell University

Senior Advisors:

Kyunghyun Cho, NYU and FAIR
Edward Grefenstette, DeepMind
Karl Moritz Hermann, DeepMind
Laura Rimell, DeepMind

Program Committee:

Eneko Agirre, University of the Basque Country
Yoav Artzi, Cornell University
Mohit Bansal, UNC Chapel Hill
Meriem Beloucif, HKUST
Jonathan Berant, Tel-Aviv
Johannes Bjerva, University of Copenhagen
Jan Buys, Oxford University
Xilun Chen, Cornell University
Eunsol Choi, University of Washington
Heeyoul Choi, Handong Global University
Junyoung Chung, University of Montreal
Manuel Ciosici, Aarhus University
Stephen Clark, DeepMind
Marco Damonte, University of Edinburgh
Desmond Elliot, University of Edinburgh
Katrin Erk, University of Texas
Orhan Firat, Middle East Technical University
Lucie Flekova, Amazon Research
Kevin Gimpel, TTI-Chicago
Caglar Gulcehre, University of Montreal
Gholamreza Haffari, Monash University
Mohit Iyyer, AI2
Katharina Kann, LMU
Arzoo Katiyar, Cornell University
Miryam de Lhoneux, Uppsala University
Tegan Maharaj, Polytechnique Montreal
Ana Marasovic, Heidelberg, University
Yishu Miao, Oxford University
Todor Mihaylov, Heidelberg University

Pasquale Minervini, UCL
Nikita Nangia, NYU
Shashi Narayan, University of Edinburgh
Thien Huu Nguyen, NYU
Robert Östling, Stockholm University
Alexander Panchenko, University of Hamburg
Matthew Peters, AI2
Barbara Plank, University of Groningen
Marek Rei, University of Cambridge
Roi Reichart, Technion
Alan Ritter, Ohio State University
Diarmuid Ó Séaghdha, Apple
Holger Schwenk, Facebook Research
Tianze Shi, Cornell University
Vered Shwartz, Bar-Ilan University
Ashudeep Singh, Cornell University
Richard Socher, Salesforce
Mark Steedman, University of Edinburgh
Karl Stratos, Columbia University
Sam Thomson, CMU
Ivan Titov, University of Edinburgh
Shubham Toshniwal, TTIC
Andreas Vlachos, Sheffield
Pontus Stenetorp, UCL
Anders Søgaard, University of Copenhagen
Jörg Tiedemann, University of Helsinki
Chris Quirk, Microsoft Research
Lyle Ungar, University of Pennsylvania
Eva Maria Vecchi, University of Cambridge
Dirk Weissenborn, German Research Center for AI
Tsung-Hsien Wen, University of Cambridge
Yi Yang, Bloomberg LP
Helen Yannakoudakis, University of Cambridge

Invited Speaker:

Yejin Choi, University of Washington
Trevor Cohn, University of Melbourne
Margaret Mitchell, Google Research
Yoav Goldberg, Bar Ilan University

Table of Contents

<i>Corpus Specificity in LSA and Word2vec: The Role of Out-of-Domain Documents</i> Edgar Altszyler, Mariano Sigman and Diego Fernandez Slezak	1
<i>Hierarchical Convolutional Attention Networks for Text Classification</i> Shang Gao, Arvind Ramanathan and Georgia Tourassi	11
<i>Extrofitting: Enriching Word Representation and its Vector Space with Semantic Lexicons</i> Hwiyeol Jo and Stanley Jungkyu Choi	24
<i>Chat Discrimination for Intelligent Conversational Agents with a Hybrid CNN-LMTGRU Network</i> Dennis Singh Moirangthem and Minhoo Lee	30
<i>Text Completion using Context-Integrated Dependency Parsing</i> Amr Rekaby Salama, Özge Alacam and Wolfgang Menzel	41
<i>Quantum-Inspired Complex Word Embedding</i> Qiuchi Li, Sagar Uprety, Benyou Wang and Dawei Song	50
<i>Natural Language Inference with Definition Embedding Considering Context On the Fly</i> Kosuke Nishida, Kyosuke Nishida, Hisako Asano and Junji Tomita	58
<i>Comparison of Representations of Named Entities for Document Classification</i> Lidia Pivovarova and Roman Yangarber	64
<i>Speeding up Context-based Sentence Representation Learning with Non-autoregressive Convolutional Decoding</i> Shuai Tang, Hailin Jin, Chen Fang, Zhaowen Wang and Virginia de Sa	69
<i>Connecting Supervised and Unsupervised Sentence Embeddings</i> Gil Levi	79
<i>A Hybrid Learning Scheme for Chinese Word Embedding</i> Wenfan Chen and Weiguo Sheng	84
<i>Unsupervised Random Walk Sentence Embeddings: A Strong but Simple Baseline</i> Kawin Ethayarajh	91
<i>Evaluating Word Embeddings in Multi-label Classification Using Fine-Grained Name Typing</i> Yadollah Yaghoobzadeh, Katharina Kann and Hinrich Schütze	101
<i>A Dense Vector Representation for Open-Domain Relation Tuples</i> Ade Romadhony, Alfian Farizki Wicaksono, Ayu Purwarianti and Dwi Hendratmo Widyanoro	107
<i>Exploiting Common Characters in Chinese and Japanese to Learn Cross-Lingual Word Embeddings via Matrix Factorization</i> Jilei Wang, Shiyang Luo, Weiyan Shi, Tao Dai and Shu-Tao Xia	113
<i>WordNet Embeddings</i> Chakaveh Saedi, António Branco, João António Rodrigues and João Silva	122
<i>Knowledge Graph Embedding with Numeric Attributes of Entities</i> Yanrong Wu and Zhichun Wang	132

<i>Injecting Lexical Contrast into Word Vectors by Guiding Vector Space Specialisation</i> Ivan Vulić	137
<i>Characters or Morphemes: How to Represent Words?</i> Ahmet Üstün, Murathan Kurfalı and Burcu Can	144
<i>Learning Hierarchical Structures On-The-Fly with a Recurrent-Recursive Model for Sequences</i> Athul Paul Jacob, Zhouhan Lin, Alessandro Sordoni and Yoshua Bengio	154
<i>Limitations of Cross-Lingual Learning from Image Search</i> Mareike Hartmann and Anders Søgaard	159
<i>Learning Semantic Textual Similarity from Conversations</i> Yinfei Yang, Steve Yuan, Daniel Cer, Sheng-Yi Kong, Noah Constant, Petr Pilar, Heming Ge, Yun-hsuan Sung, Brian Strope and Ray Kurzweil	164
<i>Multilingual Seq2seq Training with Similarity Loss for Cross-Lingual Document Classification</i> Katherine Yu, Haoran Li and Barlas Oguz	175
<i>LSTMs Exploit Linguistic Attributes of Data</i> Nelson F. Liu, Omer Levy, Roy Schwartz, Chenhao Tan and Noah A. Smith	180
<i>Learning Distributional Token Representations from Visual Features</i> Samuel Broscheit	187
<i>Jointly Embedding Entities and Text with Distant Supervision</i> Denis Newman-Griffis, Albert M. Lai and Eric Fosler-Lussier	195
<i>A Sequence-to-Sequence Model for Semantic Role Labeling</i> Angel Daza and Anette Frank	207
<i>Predicting Concreteness and Imageability of Words Within and Across Languages via Word Embeddings</i> Nikola Ljubešić, Darja Fišer and Anita Peti-Stantić	217

Workshop Program

Friday, July 20, 2018

09:30–09:45 *Welcome and Opening Remarks*

09:45–14:45 **Keynote Session**

09:45–10:30 *Invited Talk 1*
Yejin Choi

10:30–11:00 *Coffee Break*

11:00–11:45 *Invited Talk 2*
Trevor Cohn

11:45–12:30 *Invited Talk 3*
Margaret Mitchell

12:30–14:00 *Lunch*

14:00–14:45 *Invited Talk 4*
Yoav Goldberg

14:45–15:00 **Outstanding Papers Spotlight Presentations**

Friday, July 20, 2018 (continued)

15:00–16:30 **Poster Session (including Coffee Break from 15:30-16:00) + Drinks Reception**

Corpus Specificity in LSA and Word2vec: The Role of Out-of-Domain Documents

Edgar Altszyler, Mariano Sigman and Diego Fernandez Slezak

Hierarchical Convolutional Attention Networks for Text Classification

Shang Gao, Arvind Ramanathan and Georgia Tourassi

Extrofitting: Enriching Word Representation and its Vector Space with Semantic Lexicons

Hwiyeol Jo and Stanley Jungkyu Choi

Chat Discrimination for Intelligent Conversational Agents with a Hybrid CNN-LMTGRU Network

Dennis Singh Moirangthem and Minhoo Lee

Text Completion using Context-Integrated Dependency Parsing

Amr Rekaby Salama, Özge Alacam and Wolfgang Menzel

Quantum-Inspired Complex Word Embedding

Qiuchi Li, Sagar Uprety, Benyou Wang and Dawei Song

Natural Language Inference with Definition Embedding Considering Context On the Fly

Kosuke Nishida, Kyosuke Nishida, Hisako Asano and Junji Tomita

Comparison of Representations of Named Entities for Document Classification

Lidia Pivovarova and Roman Yangarber

Speeding up Context-based Sentence Representation Learning with Non-autoregressive Convolutional Decoding

Shuai Tang, Hailin Jin, Chen Fang, Zhaowen Wang and Virginia de Sa

Connecting Supervised and Unsupervised Sentence Embeddings

Gil Levi

A Hybrid Learning Scheme for Chinese Word Embedding

Wenfan Chen and Weiguo Sheng

Friday, July 20, 2018 (continued)

Unsupervised Random Walk Sentence Embeddings: A Strong but Simple Baseline

Kawin Ethayarajh

Evaluating Word Embeddings in Multi-label Classification Using Fine-Grained Name Typing

Yadollah Yaghoobzadeh, Katharina Kann and Hinrich Schütze

A Dense Vector Representation for Open-Domain Relation Tuples

Ade Romadhony, Alfian Farizki Wicaksono, Ayu Purwarianti and Dwi Hendratmo Widyantoro

Exploiting Common Characters in Chinese and Japanese to Learn Cross-Lingual Word Embeddings via Matrix Factorization

Jilei Wang, Shiyang Luo, Weiyan Shi, Tao Dai and Shu-Tao Xia

WordNet Embeddings

Chakaveh Saedi, António Branco, João António Rodrigues and João Silva

Knowledge Graph Embedding with Numeric Attributes of Entities

Yanrong Wu and Zhichun Wang

Injecting Lexical Contrast into Word Vectors by Guiding Vector Space Specialisation

Ivan Vulić

Characters or Morphemes: How to Represent Words?

Ahmet Üstün, Murathan Kurfalı and Burcu Can

Learning Hierarchical Structures On-The-Fly with a Recurrent-Recursive Model for Sequences

Athul Paul Jacob, Zhouhan Lin, Alessandro Sordani and Yoshua Bengio

Limitations of Cross-Lingual Learning from Image Search

Mareike Hartmann and Anders Søgaard

Learning Semantic Textual Similarity from Conversations

Yinfei Yang, Steve Yuan, Daniel Cer, Sheng-Yi Kong, Noah Constant, Petr Pilar, Heming Ge, Yun-hsuan Sung, Brian Strope and Ray Kurzweil

Multilingual Seq2seq Training with Similarity Loss for Cross-Lingual Document Classification

Katherine Yu, Haoran Li and Barlas Oguz

Friday, July 20, 2018 (continued)

LSTMs Exploit Linguistic Attributes of Data

Nelson F. Liu, Omer Levy, Roy Schwartz, Chenhao Tan and Noah A. Smith

Learning Distributional Token Representations from Visual Features

Samuel Broscheit

Jointly Embedding Entities and Text with Distant Supervision

Denis Newman-Griffis, Albert M. Lai and Eric Fosler-Lussier

A Sequence-to-Sequence Model for Semantic Role Labeling

Angel Daza and Anette Frank

*Predicting Concreteness and Imageability of Words Within and Across Languages
via Word Embeddings*

Nikola Ljubešić, Darja Fišer and Anita Peti-Stantić

16:30–17:30 Panel Discussion

17:30–17:40 Closing Remarks + Best Paper Awards Announcement