

1st Workshop on Evaluating Vector-Space Representations for NLP 2016

Held at ACL 2016

Berlin, Germany
7 – 12 August 2016

ISBN: 978-1-5108-2770-7

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© (2016) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2016)

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>Intrinsic Evaluation of Word Vectors Fails to Predict Extrinsic Performance</i> Billy Chiu, Anna Korhonen and Sampo Pyysalo	1
<i>A critique of word similarity as a method for evaluating distributional semantic models</i> Miroslav Batchkarov, Thomas Kober, Jeremy Reffin, Julie Weeds and David Weir	7
<i>Issues in evaluating semantic spaces using word analogies</i> Tal Linzen	13
<i>Evaluating Word Embeddings Using a Representative Suite of Practical Tasks</i> Neha Nayak, Gabor Angeli and Christopher D. Manning	19
<i>Story Cloze Evaluator: Vector Space Representation Evaluation by Predicting What Happens Next</i> Nasrin Mostafazadeh, Lucy Vanderwende, Wen-tau Yih, Pushmeet Kohli and James Allen	24
<i>Problems With Evaluation of Word Embeddings Using Word Similarity Tasks</i> Manaal Faruqui, Yulia Tsvetkov, Pushpendre Rastogi and Chris Dyer	30
<i>Intrinsic Evaluations of Word Embeddings: What Can We Do Better?</i> Anna Gladkova and Aleksandr Drozd	36
<i>Find the word that does not belong: A Framework for an Intrinsic Evaluation of Word Vector Representations</i> José Camacho-Collados and Roberto Navigli	43
<i>Capturing Discriminative Attributes in a Distributional Space: Task Proposal</i> Alicia Krebs and Denis Paperno	51
<i>An Improved Crowdsourcing Based Evaluation Technique for Word Embedding Methods</i> Farhana Ferdousi Liza and Marek Grzes	55
<i>Evaluation of acoustic word embeddings</i> Sahar Ghannay, Yannick Estève, Nathalie Camelin and Paul Deleglise	62
<i>Evaluating Embeddings using Syntax-based Classification Tasks as a Proxy for Parser Performance</i> Arne Köhn	67
<i>Evaluating vector space models using human semantic priming results</i> Allyson Ettinger and Tal Linzen	72
<i>Evaluating embeddings on dictionary-based similarity</i> Judit Ács and Andras Kornai	78
<i>Evaluating multi-sense embeddings for semantic resolution monolingually and in word translation</i> Gábor Borbély, Márton Makrai, Dávid Márk Nemeskey and Andras Kornai	83
<i>Subsumption Preservation as a Comparative Measure for Evaluating Sense-Directed Embeddings</i> Ali Seyed	90
<i>Evaluating Informal-Domain Word Representations With UrbanDictionary</i> Naomi Saphra	94

<i>Thematic fit evaluation: an aspect of selectional preferences</i> Asad Sayeed, Clayton Greenberg and Vera Demberg	99
<i>Improving Reliability of Word Similarity Evaluation by Redesigning Annotation Task and Performance Measure</i> Oded Avraham and Yoav Goldberg	106
<i>Correlation-based Intrinsic Evaluation of Word Vector Representations</i> Yulia Tsvetkov, Manaal Faruqui and Chris Dyer	111
<i>Evaluating word embeddings with fMRI and eye-tracking</i> Anders Søgaard	116
<i>Defining Words with Words: Beyond the Distributional Hypothesis</i> Iuliana-Elena Parasca, Andreas Lukas Rauter, Jack Roper, Aleksandar Rusinov, Guillaume Bouchard, Sebastian Riedel and Pontus Stenetorp	122
<i>A Proposal for Linguistic Similarity Datasets Based on Commonality Lists</i> Dmitrijs Milajevs and Sascha Griffiths	127
<i>Probing for semantic evidence of composition by means of simple classification tasks</i> Allyson Ettinger, Ahmed Elgohary and Philip Resnik	134
<i>SLEDDDED: A Proposed Dataset of Event Descriptions for Evaluating Phrase Representations</i> Laura Rimell and Eva Maria Vecchi	140
<i>Sentence Embedding Evaluation Using Pyramid Annotation</i> Tal Baumel, Raphael Cohen and Michael Elhadad	145