Second Workshop on Insights from Negative Results in NLP 2021

Punta Cana, Dominican Republic and Online
10 November 2021

ISBN: 978-1-7138-9182-6

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© (2021) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

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

Corrected CBOW Performs as well as Skip-gram
Ozan İrsoy, Adrian Benton and Karl Stratos 1

Does Commonsense help in detecting Sarcasm?
Somnath Basu Roy Chowdhury and Snigdha Chaturvedi 9

BERT Cannot Align Characters
Antonis Maronikolakis, Philipp Dufter and Hinrich Schütze 16

Two Heads are Better than One? Verification of Ensemble Effect in Neural Machine Translation
chanjun park, Sungjin Park, Seolhwa Lee, Taesun Whang and Heuiseok Lim 23

Finetuning Pretrained Transformers into Variational Autoencoders
Seongmin Park and Jihwa Lee 29

Are BERTs Sensitive to Native Interference in L2 Production?
Zixin Tang, Prasenjit Mitra and David Reitter 36

Zero-Shot Cross-Lingual Transfer is a Hard Baseline to Beat in German Fine-Grained Entity Typing
Sabine Weber and Mark Steedman 42

Comparing Euclidean and Hyperbolic Embeddings on the WordNet Nouns Hypernymy Graph
Sameer Bansal and Adrian Benton 49

When does Further Pre-training MLM Help? An Empirical Study on Task-Oriented Dialog Pre-training
Qi Zhu, Yuxian Gu, Lingxiao Luo, Bing Li, Cheng LI, Wei Peng, Minlie Huang and Xiaoyan Zhu
54

Recurrent Attention for the Transformer
Jan Rosendahl, Christian Herold, Frithjof Petrick and Hermann Ney 62

On the Difficulty of Segmenting Words with Attention
Ramon Sanabria, Hao Tang and Sharon Goldwater 67

The Highs and Lows of Simple Lexical Domain Adaptation Approaches for Neural Machine Translation
Nikolay Bogoychev and Pinzhen Chen 74

Backtranslation in Neural Morphological Inflection
Ling Liu and Mans Hulden 81

Learning Data Augmentation Schedules for Natural Language Processing
Daphné Chopard, Matthias S. Treder and Irena Spasić 89

An Investigation into the Contribution of Locally Aggregated Descriptors to Figurative Language Identification
Sina Mahdipour Saravani, Ritwik Banerjee and Indrakshi Ray 103

Blindness to Modality Helps Entailment Graph Mining
Liane Guillou, Sander Bijl de Vroe, Mark Johnson and Mark Steedman 110

Investigating the Effect of Natural Language Explanations on Out-of-Distribution Generalization in Few-shot NLI
Yangqiaoyu Zhou and Chenhao Tan 117

Generalization in NLI: Ways (Not) To Go Beyond Simple Heuristics
Prajjwal Bhargava, Aleksandr Drozd and Anna Rogers 125

Challenging the Semi-Supervised VAE Framework for Text Classification
Ghazi Felhi, Joseph Le Roux and Djamé Seddah 136

Active Learning for Argument Strength Estimation
Nataliia Kees, Michael Fromm, Evgeniy Faerman and Thomas Seidl 144