Second Workshop on Deep Learning Approaches for Low-Resource Natural Language Processing 2019

Hong Kong, China
3 November 2019

ISBN: 978-1-7138-0077-4

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© (2019) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2020)

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

A Closer Look At Feature Space Data Augmentation For Few-Shot Intent Classification
Varun Kumar, Hadrien Glaude, Cyprien de Lichy and Wlliam Campbell 1

A Comparative Analysis of Unsupervised Language Adaptation Methods
Gil Rocha and Henrique Lopes Cardoso 11

A logical-based corpus for cross-lingual evaluation
Felipe Salvatore, Marcelo Finger and Roberto Hirata Jr 22

Bad Form: Comparing Context-Based and Form-Based Few-Shot Learning in Distributional Semantic Models
Jeroen Van Hautte, Guy Emerson and Marek Rei 31

Bag-of-Words Transfer: Non-Contextual Techniques for Multi-Task Learning
Seth Ebner, Felicity Wang and Benjamin Van Durme 40

BERT is Not an Interlingua and the Bias of Tokenization
Jasdeep Singh, Bryan McCann, Richard Socher and Caiming Xiong 47

Cross-lingual Joint Entity and Word Embedding to Improve Entity Linking and Parallel Sentence Mining
Xiaoman Pan, Thamme Gowda, Heng Ji, Jonathan May and Scott Miller 56

Deep Bidirectional Transformers for Relation Extraction without Supervision
Yannis Papanikolaou, Ian Roberts and Andrea Pierleoni 67

Domain Adaptation with BERT-based Domain Classification and Data Selection
Xiaofei Ma, Peng Xu, Zhiguo Wang, Ramesh Nallapati and Bing Xiang 76

Empirical Evaluation of Active Learning Techniques for Neural MT
Xiangkai Zeng, Sarthak Garg, Rajen Chatterjee, Udhyakumar Nallasamy and Matthias Paulik . . 84

Fast Domain Adaptation of Semantic Parsers via Paraphrase Attention
Avik Ray, Yilin Shen and Hongxia Jin 94

Few-Shot and Zero-Shot Learning for Historical Text Normalization
Marcel Bollmann, Natalia Korchagina and Anders Søgaard 104

From Monolingual to Multilingual FAQ Assistant using Multilingual Co-training
Mayur Patidar, Surabhi Kumari, Manasi Patwardhan, Shirish Karande, Puneet Agarwal, Lovekesh Vig and Gautam Shroff 115

Generation-Distillation for Efficient Natural Language Understanding in Low-Data Settings
Luke Melas-Kyriazi, George Han and Celine Liang 124

Unlearn Dataset Bias in Natural Language Inference by Fitting the Residual
He He, Sheng Zha and Haohan Wang 132

Metric Learning for Dynamic Text Classification
Jeremy Wohlwend, Ethan R. Elenberg, Sam Altschul, Shawn Henry and Tao Lei 143

Evaluating Lottery Tickets Under Distributional Shifts
Shrey Desai, Hongyuan Zhan and Ahmed Aly 153

Cross-lingual Parsing with Polyglot Training and Multi-treebank Learning: A Faroese Case Study
James Barry, Joachim Wagner and Jennifer Foster 163

Inject Rubrics into Short Answer Grading System
Tianqi Wang, Naoya Inoue, Hiroki Ouchi, Tomoya Mizumoto and Kentaro Inui 175

Instance-based Inductive Deep Transfer Learning by Cross-Dataset Querying with Locality Sensitive Hashing
Somnath Basu Roy Chowdhury, Annervaz M and Ambedkar Dukkipati 183

Multimodal, Multilingual Grapheme-to-Phoneme Conversion for Low-Resource Languages
James Route, Steven Hillis, Isak Czeresnia Etinger, Han Zhang and Alan W Black 192

Natural Language Generation for Effective Knowledge Distillation
Raphael Tang, Yao Lu and Jimmy Lin 202

Neural Unsupervised Parsing Beyond English
Katharina Kann, Anhad Mohananey, Samuel R. Bowman and Kyunghyun Cho 209

Reevaluating Argument Component Extraction in Low Resource Settings
Anirudh Joshi, Timothy Baldwin, Richard Sinnott and Cecile Paris 219

Reinforcement-based denoising of distantly supervised NER with partial annotation
Farhad Nooralahzadeh, Jan Tore Lønning and Lilja Øvrelid 225

Samvaadhana : A Telugu Dialogue System in Hospital Domain
Suma Reddy Duggenpudi, Kusampudi Siva Subrahamanyam Varma and Radhika Mamidi 234

Towards Zero-resource Cross-lingual Entity Linking
Shuyan Zhou, Shruti Rijhwani and Graham Neubig 243

Transductive Auxiliary Task Self-Training for Neural Multi-Task Models
Johannes Bjerva, Katharina Kann and Isabelle Augenstein 253

Weakly Supervised Attentional Model for Low Resource Ad-hoc Cross-lingual Information Retrieval
Lingjun Zhao, Rabih Zbib, Zhuolin Jiang, Damianos Karakos and Zhongqiang Huang 259

X-WikiRE: A Large, Multilingual Resource for Relation Extraction as Machine Comprehension
Mostafa Abdou, Cezar Sas, Rahul Aralikatte, Isabelle Augenstein and Anders Søgaard 265

Zero-Shot Cross-lingual Name Retrieval for Low-Resource Languages
Kevin Blissett and Heng Ji 275

Zero-shot Dependency Parsing with Pre-trained Multilingual Sentence Representations
Ke Tran and Arianna Bisazza 281