SustaiNLP 2021

2nd Workshop on Simple and Efficient Natural Language Processing (SustaiNLP)

Proceedings of SustaiNLP

November 10, 2021
Online workshop

©2021 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 978-1-955917-01-8

Introduction

It is our great pleasure to welcome you to the second edition of SustaiNLP: Workshop on Simple and Efficient Natural Language Processing.

The Natural Language Processing community has, in recent years, demonstrated a notable focus on improving higher scores on standard benchmarks and taking the lead on community-wide leaderboards (e.g., GLUE, SentEval). While this aspiration has led to improvements in benchmark performance of (predominantly neural) models, it has also came at a cost, i.e., increased model complexity and the evergrowing amount of computational resources required for training and using the current state-of-the-art models. Moreover, the recent research efforts have, for the most part, failed to identify sources of empirical gains in models, often failing to empirically justify the model complexity beyond benchmark performance.

Because of these easily observable trends, we organized the SustaiNLP workshop with the goal of promoting more sustainable NLP research and practices, with two main objectives: (1) encouraging development of more efficient NLP models; and (2) providing simpler architectures and empirical justification of model complexity. For both aspects, we encouraged submissions from all topical areas of NLP.

Besides the original research papers (short and long), we encouraged cross-submissions of work that has been published at other events as well as extended abstracts of work in progress that fit the scope and aims of the workshop (only the original research papers, however, are included in these workshop proceedings).

We received 51 submissions, proposing a multitude of viable resource-efficient NLP methods and spanning a wide range of NLP applications. We have selected 21 submissions for presentation at the workshop, yielding an acceptance rate of $\sim$41%.

Many thanks to our program committee for their thorough and thoughtful reviews. We would also like to thank to our panelists and invited speakers whose discussions and talks we strongly believe will make the workshop exciting and memorable.

We are looking forward to the second edition of the SustaiNLP workshop!

SustaiNLP Organizers
September 2021

Organizers:
Nafise Sadat Moosavi, TU Darmstadt
Iryna Gurevych, TU Darmstadt
Angela Fan, INRIA Nancy and Facebook AI Research
Thomas Wolf, Huggingface Inc.
Yufang Hou, IBM Research
Ana Marasović, Allen Institute for AI and University of Washington
Zornitsa Kozareva, Facebook AI Research
Sujith Ravi, SliceX AI

Program Committee:
Dennis Aumiller, Emily M. Bender, Rajarshi Bhowmik, Samuel Cahyawijaya, Han Cai, Qingqing Cao, Rotem Dror, gerard DUPONT, Marzieh Fadaee, Benjamin Heinzerling, Ari Holtzman, Zhiqi Huang, Peter Izsak, Ehsan Kamalloo, Abhinav Ramesh Kashyap, Young Jin Kim, Gyuwan Kim, Tuan Lai, Anne Lauscher, Phong Le, Ji-Ung Lee, Zhaojiang Lin, Zeyu Liu, Alexandra Luccioni, Sandya Mannarswamy, Mohsen Mesgar, Seyed Abolghasem Mirroshandel, Marius Mosbach, Tapas Nayak, Myle Ott, Nikolaos Pappas, Daraksha Parveen, Mohammad Taher Pilehvar, Barbara Plank, Roi Reichart, Victor Sanh, Roy Schwartz, Sheng Shen, Edwin Simpson, Sameer Singh, Urmish Thakker, Rakesh Verma, Ivan Vulić, Moshe Wasserblat, Genta Indra Winata, Mengzhou Xia, Patrick Xia, Canwen Xu, Peng Xu, Yan Xu, Makoto Yamada, Yichun Yin, Naoki Yoshinaga, Tiezheng Yu, Zhuosheng Zhang

Invited Speakers:
Jacob Andreas, MIT
Colin Raffel, University of North Carolina, Chapel Hill
Nazneen Rajani, Salesforce Research
Dan Roth, University of Pennsylvania
Roy Schwartz, Hebrew University of Jerusalem
Yulia Tsvetkov, University of Washington

Panelists:
Emily M. Bender, University of Washington
Goran Glavaš, Universität Mannheim
Perez Ogayo, Carnegie Mellon University
Colin Raffel, University of North Carolina, Chapel Hill
Roy Schwartz, Hebrew University of Jerusalem
Moshe Wasserblat, Intel

Table of Contents

Low Resource Quadratic Forms for Knowledge Graph Embeddings
Zachary Zhou, Jeffery Kline, Devin Conathan and Glenn Fung 1

Evaluating the carbon footprint of NLP methods: a survey and analysis of existing tools
Nesrine bannour, Sahar Ghannay, Aurélie Névéol and Anne-Laure Ligozat 11

Limitations of Knowledge Distillation for Zero-shot Transfer Learning
Saleh Soltan, Haidar Khan and Wael Hamza 22

Countering the Influence of Essay Length in Neural Essay Scoring
Sungho Jeon and Michael Strube 32

Memory-efficient Transformers via Top-k Attention
Ankit Gupta, Guy Dar, Shaya Goodman, David Ciprut and Jonathan Berant 39

BioCopy: A Plug-And-Play Span Copy Mechanism in Seq2Seq Models
Yi Liu, Guoan Zhang, Puning Yu, Jianlin Su and Shengfeng Pan 53

Combining Lexical and Dense Retrieval for Computationally Efficient Multi-hop Question Answering
Georgios Sidiropoulos, Nikos Voskarides, Svitlana Vakulenko and Evangelos Kanoulas 58

Learning to Rank in the Age of Muppets: Effectiveness–Efficiency Tradeoffs in Multi-Stage Ranking
Yue Zhang, ChengCheng Hu, Yuqi Liu, Hui Fang and Jimmy Lin 64

Improving Synonym Recommendation Using Sentence Context
Maria Glenski, William I. Sealy, Kate Miller and Dustin Arendt 74

Semantic Categorization of Social Knowledge for Commonsense Question Answering
Gengyu Wang, Xiaochen Hou, Diyi Yang, Kathleen McKeown and Jing Huang 79

Speeding Up Transformer Training By Using Dataset Subsampling - An Exploratory Analysis
Lovre Torbarina, Velimir Mihelčić, Bruno Šarlija, Lukasz Roguski and Željko Kraljević 86

Hyperparameter Power Impact in Transformer Language Model Training
Lucas Høyberg Puvis de Chavannes, Mads Guldborg Kjeldgaard Kongsbak, Timmie Rantzau and Leon Derczynski 96

Distiller: A Systematic Study of Model Distillation Methods in Natural Language Processing
Haoyu He, Xingjian Shi, Jonas Mueller, Sheng Zha, Mu Li and George Karypis 119

Shrinking Bigfoot: Reducing wav2vec 2.0 footprint
Zilun Peng, Akshay Budhkar, Ilana Tuil, Jason Levy, Parinaz Sobhani, Raphael Cohen and Jumana Nassour 134

On the Role of Corpus Ordering in Language Modeling
Ameeta Agrawal, Suresh Singh, Lauren Schneider and Michael Samuels 142

Efficient Domain Adaptation of Language Models via Adaptive Tokenization
Vin Sachidananda, Jason Kessler and Yi-An Lai 155

Unsupervised Contextualized Document Representation
Ankur Gupta and Vivek Gupta 166

Conference Program

November 10, 2021

Low Resource Quadratic Forms for Knowledge Graph Embeddings
Zachary Zhou, Jeffery Kline, Devin Conathan and Glenn Fung

Evaluating the carbon footprint of NLP methods: a survey and analysis of existing tools
Nesrine bannour, Sahar Ghannay, Aurélie Névéol and Anne-Laure Ligozat

Limitations of Knowledge Distillation for Zero-shot Transfer Learning
Saleh Soltan, Haidar Khan and Wael Hamza

Countering the Influence of Essay Length in Neural Essay Scoring
Sungho Jeon and Michael Strube

Memory-efficient Transformers via Top-k Attention
Ankit Gupta, Guy Dar, Shaya Goodman, David Ciprut and Jonathan Berant

BioCopy: A Plug-And-Play Span Copy Mechanism in Seq2Seq Models
Yi Liu, Guoan Zhang, Puning Yu, Jianlin Su and Shengfeng Pan

Combining Lexical and Dense Retrieval for Computationally Efficient Multi-hop Question Answering
Georgios Sidiropoulos, Nikos Voskarides, Svitlana Vakulenko and Evangelos Kanoulas

Learning to Rank in the Age of Muppets: Effectiveness–Efficiency Tradeoffs in Multi-Stage Ranking
Yue Zhang, ChengCheng Hu, Yuqi Liu, Hui Fang and Jimmy Lin

Improving Synonym Recommendation Using Sentence Context
Maria Glenski, William I. Sealy, Kate Miller and Dustin Arendt

Semantic Categorization of Social Knowledge for Commonsense Question Answering
Gengyu Wang, Xiaochen Hou, Diyi Yang, Kathleen McKeown and Jing Huang

Speeding Up Transformer Training By Using Dataset Subsampling - An Exploratory Analysis
Lovre Torbarina, Velimir Mihelčić, Bruno Šarlija, Lukasz Roguski and Željko Kraljević

November 10, 2021 (continued)

Length-Adaptive Transformer: Train Once with Length Drop, Use Anytime with Search
Gyuwan Kim and Kyunghyun Cho

Hyperparameter Power Impact in Transformer Language Model Training
Lucas Høyberg Puvis de Chavannes, Mads Guldborg Kjeldgaard Kongsbak, Timmie Rantzau and Leon Derczynski

Distiller: A Systematic Study of Model Distillation Methods in Natural Language Processing
Haoyu He, Xingjian Shi, Jonas Mueller, Sheng Zha, Mu Li and George Karypis

Simple and Efficient ways to Improve REALM
Vidhisha Balachandran, Ashish Vaswani, Yulia Tsvetkov and Niki Parmar

Shrinking Bigfoot: Reducing wav2vec 2.0 footprint
Zilun Peng, Akshay Budhkar, Ilana Tuil, Jason Levy, Parinaz Sobhani, Raphael Cohen and Jumana Nassour

On the Role of Corpus Ordering in Language Modeling
Ameeta Agrawal, Suresh Singh, Lauren Schneider and Michael Samuels

Efficient Domain Adaptation of Language Models via Adaptive Tokenization
Vin Sachidananda, Jason Kessler and Yi-An Lai

Unsupervised Contextualized Document Representation
Ankur Gupta and Vivek Gupta

Logistic Regression Trained on Learner Data Outperformed Neural Language Models in Unsupervised Automatic Readability Assessment
Yo Ehara