

Fourth Conference on Machine Translation (WMT 2019)

Volume 3: Shared Task Papers, Day 2

Florence, Italy
1-2 August 2019

ISBN: 978-1-5108-9291-0

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 by Curran Associates, Inc. (2019)

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>Findings of the WMT 2019 Shared Tasks on Quality Estimation</i> Erick Fonseca, Lisa Yankovskaya, André F. T. Martins, Mark Fishel and Christian Federmann . . .	1
<i>Findings of the WMT 2019 Shared Task on Automatic Post-Editing</i> Rajen Chatterjee, Christian Federmann, Matteo Negri and Marco Turchi	11
<i>Findings of the WMT 2019 Biomedical Translation Shared Task: Evaluation for MEDLINE Abstracts and Biomedical Terminologies</i> Rachel Bawden, Kevin Bretonnel Cohen, Cristian Grozea, Antonio Jimeno Yepes, Madeleine Kitner, Martin Krallinger, Nancy Mah, Aurelie Neveol, Mariana Neves, Felipe Soares, Amy Siu, Karin Verspoor and Maika Vicente Navarro	29
<i>Findings of the WMT 2019 Shared Task on Parallel Corpus Filtering for Low-Resource Conditions</i> Philipp Koehn, Francisco Guzmán, Vishrav Chaudhary and Juan Pino	54
<i>RTM Stacking Results for Machine Translation Performance Prediction</i> Ergun Biçici	73
<i>Unbabel’s Participation in the WMT19 Translation Quality Estimation Shared Task</i> Fabio Kepler, Jonay Trénous, Marcos Treviso, Miguel Vera, António Góis, M. Amin Farajian, António V. Lopes and André F. T. Martins	78
<i>QE BERT: Bilingual BERT Using Multi-task Learning for Neural Quality Estimation</i> Hyun Kim, Joon-Ho Lim, Hyun-Ki Kim and Seung-Hoon Na	85
<i>MIPT System for World-Level Quality Estimation</i> Mikhail Mosyagin and Varvara Logacheva	90
<i>NJU Submissions for the WMT19 Quality Estimation Shared Task</i> Hou Qi	95
<i>Quality Estimation and Translation Metrics via Pre-trained Word and Sentence Embeddings</i> Elizaveta Yankovskaya, Andre Tättar and Mark Fishel	101
<i>SOURCE: SOURce-Conditional Elmo-style Model for Machine Translation Quality Estimation</i> Junpei Zhou, Zhisong Zhang and Zecong Hu	106
<i>Transformer-based Automatic Post-Editing Model with Joint Encoder and Multi-source Attention of Decoder</i> WonKee Lee, Jaehun Shin and Jong-Hyeok Lee	112
<i>Unbabel’s Submission to the WMT2019 APE Shared Task: BERT-Based Encoder-Decoder for Automatic Post-Editing</i> António V. Lopes, M. Amin Farajian, Gonçalo M. Correia, Jonay Trénous and André F. T. Martins	118
<i>USAAR-DFKI – The Transference Architecture for English–German Automatic Post-Editing</i> Santanu Pal, Hongfei Xu, Nico Herbig, Antonio Krüger and Josef van Genabith	124
<i>APE through Neural and Statistical MT with Augmented Data. ADAPT/DCU Submission to the WMT 2019 APE Shared Task</i> Dimitar Shterionov, Joachim Wagner and Félix do Carmo	132

<i>Effort-Aware Neural Automatic Post-Editing</i> Amirhossein Tebbifakhr, Matteo Negri and Marco Turchi	139
<i>UDS Submission for the WMT 19 Automatic Post-Editing Task</i> Hongfei Xu, Qihui Liu and Josef van Genabith	145
<i>Terminology-Aware Segmentation and Domain Feature for the WMT19 Biomedical Translation Task</i> Casimiro Pio Carrino, Bardia Rafeian, Marta R. Costa-jussà and José A. R. Fonollosa	151
<i>Exploring Transfer Learning and Domain Data Selection for the Biomedical Translation</i> Noor-e- Hira, Sadaf Abdul Rauf, Kiran Kiani, Ammara Zafar and Raheel Nawaz	156
<i>Huawei's NMT Systems for the WMT 2019 Biomedical Translation Task</i> Wei Peng, Jianfeng Liu, Liangyou Li and Qun Liu	164
<i>UCAM Biomedical Translation at WMT19: Transfer Learning Multi-domain Ensembles</i> Danielle Saunders, Felix Stahlberg and Bill Byrne	169
<i>BSC Participation in the WMT Translation of Biomedical Abstracts</i> Felipe Soares and Martin Krallinger	175
<i>The MLLP-UPV Spanish-Portuguese and Portuguese-Spanish Machine Translation Systems for WMT19 Similar Language Translation Task</i> Pau Baquero-Arnal, Javier Iranzo-Sánchez, Jorge Civera and Alfons Juan	179
<i>The TALP-UPC System for the WMT Similar Language Task: Statistical vs Neural Machine Translation</i> Magdalena Biesialska, Lluís Guardia and Marta R. Costa-jussà	185
<i>Machine Translation from an Intercomprehension Perspective</i> Yu Chen and Tania Avgustinova	192
<i>Utilizing Monolingual Data in NMT for Similar Languages: Submission to Similar Language Translation Task</i> Jyotsana Khatri and Pushpak Bhattacharyya	197
<i>Neural Machine Translation: Hindi-Nepali</i> Sahinur Rahman Laskar, Partha Pakray and Sivaji Bandyopadhyay	202
<i>NICT's Machine Translation Systems for the WMT19 Similar Language Translation Task</i> Benjamin Marie, Raj Dabre and Atsushi Fujita	208
<i>Panlingua-KMI MT System for Similar Language Translation Task at WMT 2019</i> Atul Kr. Ojha, Ritesh Kumar, Akanksha Bansal and Priya Rani	213
<i>UDS-DFKI Submission to the WMT2019 Czech-Polish Similar Language Translation Shared Task</i> Santanu Pal, Marcos Zampieri and Josef van Genabith	219
<i>Neural Machine Translation of Low-Resource and Similar Languages with Backtranslation</i> Michael Przystupa and Muhammad Abdul-Mageed	224
<i>The University of Helsinki Submissions to the WMT19 Similar Language Translation Task</i> Yves Scherrer, Raúl Vázquez and Sami Virpioja	236
<i>Dual Monolingual Cross-Entropy Delta Filtering of Noisy Parallel Data</i> Amittai Axelrod, Anish Kumar and Steve Sloto	245

<i>NRC Parallel Corpus Filtering System for WMT 2019</i>	
Gabriel Bernier-Colborne and Chi-kiu Lo	252
<i>Low-Resource Corpus Filtering Using Multilingual Sentence Embeddings</i>	
Vishrav Chaudhary, Yuqing Tang, Francisco Guzmán, Holger Schwenk and Philipp Koehn	261
<i>Quality and Coverage: The AFRL Submission to the WMT19 Parallel Corpus Filtering for Low-Resource Conditions Task</i>	
Grant Erdmann and Jeremy Gwinnup	267
<i>Webinterpret Submission to the WMT2019 Shared Task on Parallel Corpus Filtering</i>	
Jesús González-Rubio	271
<i>Noisy Parallel Corpus Filtering through Projected Word Embeddings</i>	
Murathan Kurfalı and Robert Östling	277
<i>Filtering of Noisy Parallel Corpora Based on Hypothesis Generation</i>	
Zuzanna Pancheta, Germán Sanchis-Trilles and Francisco Casacuberta	282
<i>Parallel Corpus Filtering Based on Fuzzy String Matching</i>	
Sukanta Sen, Asif Ekbal and Pushpak Bhattacharyya	289
<i>The University of Helsinki Submission to the WMT19 Parallel Corpus Filtering Task</i>	
Raúl Vázquez, Umut Sulubacak and Jörg Tiedemann	294