Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2022)

Industry Papers

Online
10 – 15 July 2022

ISBN: 978-1-7138-5619-1

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© (2022) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2022)

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

Scalable and Robust Self-Learning for Skill Routing in Large-Scale Conversational AI Systems
Mohammad Kachuee, Jinseok Nam, Sarthak Ahuja, Jin-Myung Won and Sungjin Lee 1

CREATER: CTR-driven Advertising Text Generation with Controlled Pre-Training and Contrastive Fine-Tuning
Penghui Wei, Xuanhua Yang, ShaoGuo Liu, Liang Wang and Bo Zheng 9

Augmenting Poetry Composition with Verse by Verse
David Uthus, Maria Voitovich and R.J. Mical 18

AB/BA analysis: A framework for estimating keyword spotting recall improvement while maintaining audio privacy
Raphael Petegrosso, VasistaKrishna Baderdinnni, Thibaud Senechal and Benjamin Bullough . 27

Temporal Generalization for Spoken Language Understanding
Judith Gaspers, Anoop Kumar, Greg Ver Steeg and Aram Galstyan 37

An End-to-End Dialogue Summarization System for Sales Calls
Abedelkadir Asi, Song Wang, Roy Eisenstadt, Dean Geckt, Yarin Kuper, Yi Mao and Royi Ronen 45

Controlled Data Generation via Insertion Operations for NLU
Manoj Kumar, Yuval Merhav, Haidar Khan, Rahul Gupta, Anna Rumshisky and Wael Hamza 54

Easy and Efficient Transformer: Scalable Inference Solution For Large NLP Model
Gongzheng Li, Yadong Xi, Jingzhen Ding, Duan Wang, Ziyang Luo, Rongsheng Zhang, Bai Liu, Changjie Fan, Xiaoxi Mao and Zeng Zhao 62

Aspect-based Analysis of Advertising Appeals for Search Engine Advertising
Soichiro Murakami, Peinan Zhang, Sho Hoshino, Hidetaka Kamigaito, Hiroya Takamura and Manabu Okumura 69

Self-supervised Product Title Rewrite for Product Listing Ads
Xue Zhao, Dayiheng Liu, Junwei Ding, Liang Yao, Mahone Yan, Huibo Wang and Wenqing Yao 79

Efficient Semi-supervised Consistency Training for Natural Language Understanding
George Leung and Joshua Tan 86

Distantly Supervised Aspect Clustering And Naming For E-Commerce Reviews
Prateek Sircar, Aniket Chakrabarti, Deepak Gupta and Anirban Majumdar 94

Local-to-global learning for iterative training of production SLU models on new features
Yulia Grishina and Daniil Sorokin 103

CULG: Commercial Universal Language Generation
Haonan Li, Yameng Huang, Yeyun Gong, Jian Jiao, Ruofei Zhang, Timothy Baldwin and Nan Duan 112

Constraining word alignments with posterior regularization for label transfer
Thomas Gueudre and Kevin Martin Jose 121

Explaining the Effectiveness of Multi-Task Learning for Efficient Knowledge Extraction from Spine MRI Reports
Arijit Sehanobish, McCullen Sandora, Nabila Abraham, Jayashri Pawar, Danielle Torres, Anasuya Das, Murray Becker, Richard Herzog, Benjamin Odry and Ron Vianu 130

FPI: Failure Point Isolation in Large-scale Conversational Assistants
Rinat Khaziev, Usman Shahid, Tobias Röding, Rakesh Chada, Emir Kapanci and Pradeep Natarajan 141

Asynchronous Convergence in Multi-Task Learning via Knowledge Distillation from Converged Tasks
Weiyi Lu, Sunny Rajagopalan, Priyanka Nigam, Jaspreet Singh, Xiaodi Sun, Yi Xu, Belinda Zeng and Trishul Chilimbi 149

Augmenting Training Data for Massive Semantic Matching Models in Low-Traffic E-commerce Stores
Ashutosh Joshi, Shankar Vishwanath, Choon Hui Teo, Vaclav Petricek, Vishy Vishwanathan, Rahul Bhagat and Jonathan May 160

Retrieval Based Response Letter Generation For a Customer Care Setting
Biplob Biswas, Renhao Cui and Rajiv Ramnath 168

Medical Coding with Biomedical Transformer Ensembles and Zero/Few-shot Learning
Angelo Ziletti, Alan Akbik, Christoph Berns, Thomas Herold, Marion Legler and Martina Viell 176

Knowledge extraction from aeronautical messages (NOTAMs) with self-supervised language models for aircraft pilots
Alexandre Arnold, Fares Ernez, Catherine Kobus and Marion-Cécile Martin 188

Intent Discovery for Enterprise Virtual Assistants: Applications of Utterance Embedding and Clustering to Intent Mining
Minhua Chen, Badrinath Jayakumar, Michael Johnston, S. Eman Mahmoodi and Daniel Pressel 197

ReFinED: An Efficient Zero-shot-capable Approach to End-to-End Entity Linking
Tom Ayoola, Shubhi Tyagi, Joseph Fisher, Christos Christodoulopoulos and Andrea Pierleoni209

Lightweight Transformers for Conversational AI
Daniel Pressel, Wenshuo Liu, Michael Johnston and Minhua Chen 221

NER-MQMRC: Formulating Named Entity Recognition as Multi Question Machine Reading Comprehension
Anubhav Shrimal, Avi Jain, Kartik Mehta and Promod Yenigalla 230

What Do Users Care About? Detecting Actionable Insights from User Feedback
Kasturi Bhattacharjee, Rashmi Gangadharaiah, Kathleen McKeown and Dan Roth 239

CTM - A Model for Large-Scale Multi-View Tweet Topic Classification
Vivek Kulkarni, Kenny Leung and Aria Haghighi 247

Developing a Production System for Purpose of Call Detection in Business Phone Conversations
Elena Khasanova, Pooja Hiranandani, Shayna Gardiner, Cheng Chen, Simon Corston-Oliver and Xue-Yong Fu 259

Adversarial Text Normalization
Joanna Bitton, Maya Pavlova and Ivan Evtimov 268

Constraint-based Multi-hop Question Answering with Knowledge Graph
Sayantan Mitra, Roshni Ramnani and Shubhashis Sengupta 280

Fast Bilingual Grapheme-To-Phoneme Conversion
Hwa-Yeon Kim, Jong-Hwan Kim and Jae-Min Kim 289

Knowledge Extraction From Texts Based on Wikidata
Anastasia Shimorina, Johannes Heinecke and Frédéric Herledan 297

AIT-QA: Question Answering Dataset over Complex Tables in the Airline Industry
Yannis Katsis, Saneem Ahmed Chemmengath, Vishwajeet Kumar, Samarth Bharadwaj, Mustafa Canim, Michael Glass, Alfio Gliozzo, Feifei Pan, Jaydeep Sen, Karthik Sankaranarayanan and Soumen Chakrabarti 305

Parameter-efficient Continual Learning Framework in Industrial Real-time Text Classification System
Tao Zhu, Zhe Zhao, Weijie Liu, Jiachi Liu, Yiren Chen, Weiquan Mao, Haoyan Liu, Kunbo Ding, Yudong Li and Xuefeng Yang 315

Self-Aware Feedback-Based Self-Learning in Large-Scale Conversational AI
Pragaash Ponnusamy, Clint Solomon Mathialagan, Gustavo Aguilar, Chengyuan Ma and Chenlei Guo 324

Fast and Light-Weight Answer Text Retrieval in Dialogue Systems
Hui Wan, Siva Sankalp Patel, J William Murdock, Saloni Potdar and Sachindra Joshi 334

BLINK with Elasticsearch for Efficient Entity Linking in Business Conversations
Md Tahmid Rahman Laskar, Cheng Chen, Aliaksandr Martsinovich, Jonathan Johnston, Xue-Yong Fu, Shashi Bhushan TN and Simon Corston-Oliver 344

Q2R: A Query-to-Resolution System for Natural-Language Queries
Shiau Hong Lim and Laura Wynter 353

Identifying Corporate Credit Risk Sentiments from Financial News
Noujoud Ahbali, Xinyuan Liu, Albert Aristotle Nanda, Jamie Stark, Ashit Talukder and Rupinder Paul Khandpur 362