ECNLP 2022

The 5th Workshop on e-Commerce and NLP

Proceedings of the Workshop

May 26, 2022

©2022 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 978-1-955917-35-3

Introduction

It is our great pleasure to welcome you to the Fifth Workshop on e-Commerce and NLP (ECNLP).

This workshop focuses on intersection of Natural Language Processing (NLP) and e-Commerce. NLP and information retrieval (IR) have been powering e-Commerce applications since the early days of the fields. Today, NLP and IR already play a significant role in e-Commerce tasks, including product search, recommender systems, product question answering, machine translation, sentiment analysis, product description and review summarization, and customer review processing. With the exploding popularity of chatbots and shopping assistants -– both text- and voice-based -– NLP, IR, question answering, and dialogue systems research is poised to transform e-Commerce once again.

The ECNLP workshop series was designed to provide a venue for the dissemination of late-breaking research results and ideas related to e-commerce and online shopping, as well as a forum where new and unfinished ideas could be discussed. After four successful editions since 2019, we are happy to host ECNLP 5 at ACL 2022 and once again bring together researchers from both academia and industry.

We have received a larger number of submissions than we could accept for presentation. ECNLP 5 received 52 submissions of long and short research papers. In total, ECNLP 5 featured 29 accepted papers (56% acceptance rate). The selection process was competitive and we believe it resulted in a balanced and varied program that is appealing to audiences from the various sub-areas of e-Commerce.

We would like to thank everyone who submitted a paper to the workshop. We would also like to express our gratitude to the members of the Program Committee for their timely reviews, and for supporting the tight schedule by providing reviews at short notice.

We hope that you enjoy the workshop!

The ECNLP Organizers May 2022

Organizing Committee

Organizing Committee

Shervin Malmasi, Amazon, USA
Surya Kallumadi, Lowe's Companies, Inc, USA
Nicola Ueffing, eBay, Germany
Eugene Agichtein, Emory University, USA
Oleg Rokhlenko, Amazon, USA
Ido Guy, Meta AI, Israel

Program Committee

Program Committee

Ahsaas Bajaj
Adrian Benton, Google
Giuseppe Castellucci, Amazon
Dumitru Clementin Cercel, University Politehnica of Bucharest
Lei Chen, Rakuten Institute of Technology, The University of Tokyo
Young-joo Chung
Marcus D. Collins, Amazon
Leonard Dahlmann
Li Dong, Amazon
Besnik Fetahu, Amazon
Ciro Greco
Jongseok Han
Ohnmar Htun
Fei Hu, Microsoft
Manoj Balaji J
Ajinkya Kale
Sudipta Kar, Amazon
Shahram Khadivi, eBay Research
Tracy Holloway King, Adobe Systems
Vinayshekhar Bannihatti Kumar, Amazon
Saar Kuzi, Amazon
Gal Lavee
Jeong Min Lee, Facebook AI
Amita Misra, Amazon
Sudipto Mukherjee, Microsoft
Varun Nagaraj Rao, University of Southern California
Indraneil Paul, Amazon
Arushi Prakash
Alexandre Salle, VTEX
Salim Sazzed
Venkat Srinivasan
Daniel Stein
Sandesh Swamy, Amazon
Jacopo Tagliabue, Coveo
Liling Tan, Amazon
Chen Zhang
Jie Zhao

Table of Contents

DEFTri: A Few-Shot Label Fused Contextual Representation Learning For Product Defect Triage in e-Commerce
Ipsita Mohanty 1

Interactive Latent Knowledge Selection for E-Commerce Product Copywriting Generation
Zeming Wang, Yanyan Zou, Yuejian Fang, Hongshen Chen, Mian Ma, Zhuoye Ding and Bo Long 8

Leveraging Seq2seq Language Generation for Multi-level Product Issue Identification
Yang Liu, Varnith Chordia, Hua Li, Siavash Fazeli Dehkordy, Yifei Sun, Vincent Gao and Na Zhang 20

Data Quality Estimation Framework for Faster Tax Code Classification
Ravi Kondadadi, Allen Mathew Williams and Nicolas Nicolov 29

CML: A Contrastive Meta Learning Method to Estimate Human Label Confidence Scores and Reduce Data Collection Cost
Bo Dong, Yiyi Wang, Hanbo Sun, Yunji Wang, Alireza Hashemi and Zheng Du 35

Improving Relevance Quality in Product Search using High-Precision Query-Product Semantic Similarity
Alireza Bagheri Garakani, Fan Yang, Wen-Yu Hua, Yetian Chen, Michinari Momma, Jingyuan Deng, Yan Gao and Yi Sun 44

Comparative Snippet Generation
Saurabh Jain, Yisong Miao and Min-Yen Kan 49

Textual Content Moderation in C2C Marketplace
Yusuke Shido, Hsien-Chi Toby Liu and Keisuke Umezawa 58

Spelling Correction using Phonetics in E-commerce Search
Fan Yang, Alireza Bagheri Garakani, Yifei Teng, Yan Gao, Jia Liu, Jingyuan Deng and Yi Sun 63

Logical Reasoning for Task Oriented Dialogue Systems
Sajjad Beygi, Maryam Fazel-Zarandi, Alessandra Cervone, Prakash Krishnan and Siddhartha Jonnalagadda 68

CoVA: Context-aware Visual Attention for Webpage Information Extraction
Anurendra Kumar, Keval Morabia, William Wang, Kevin Chang and Alex Schwing 80

Product Titles-to-Attributes As a Text-to-Text Task
Gilad Fuchs and Yoni Acriche 91

Product Answer Generation from Heterogeneous Sources: A New Benchmark and Best Practices
Xiaoyu Shen, Gianni Barlacchi, Marco Del Tredici, Weiwei Cheng, Bill Byrne and Adrià de Gispert 99

semiPQA: A Study on Product Question Answering over Semi-structured Data
Xiaoyu Shen, Gianni Barlacchi, Marco Del Tredici, Weiwei Cheng and Adrià de Gispert 111

Improving Specificity in Review Response Generation with Data-Driven Data Filtering
Tannon Kew and Martin Volk 121

Extreme Multi-Label Classification with Label Masking for Product Attribute Value Extraction
Wei-Te Chen, Yandi Xia and Keiji Shinzato 134

Enhanced Representation with Contrastive Loss for Long-Tail Query Classification in e-commerce
Lvxing Zhu, Hao Chen, Chao Wei and Weiru Zhang 141

Domain-specific knowledge distillation yields smaller and better models for conversational commerce
Kristen Howell, Jian Wang, Akshay Hazare, Joseph Bradley, Chris Brew, Xi Chen, Matthew T. Dunn, Beth Ann Hockey, Andrew Maurer and Dominic Widdows 151

OpenBrand: Open Brand Value Extraction from Product Descriptions
Kassem Sabeh, Mouna Kacimi and Johann Gamper 161

Robust Product Classification with Instance-Dependent Noise
Huy V. Nguyen and Devashish Khatwani 171

Structured Extraction of Terms and Conditions from German and English Online Shops
Tobias Michael Schamel, Daniel Braun and Florian Matthes 181

"Does it come in black?" CLIP-like models are zero-shot recommenders
Patrick John Chia, Jacopo Tagliabue, Federico Bianchi, Ciro Greco and Diogo Goncalves . . . 191

Clause Topic Classification in German and English Standard Form Contracts
Daniel Braun and Florian Matthes 199

Investigating the Generative Approach for Question Answering in E-Commerce
Kalyani Roy, Vineeth Kumar Balapanuru, Tapas Nayak and Pawan Goyal 210

Utilizing Cross-Modal Contrastive Learning to Improve Item Categorization BERT Model
Lei Chen and Hou Wei Chou 217

Towards Generalizeable Semantic Product Search by Text Similarity Pre-training on Search Click Logs
Zheng Liu, Wei Zhang, Yan Chen, Weiyi Sun, Tianchuan Du and Benjamin Schroeder 224

Can Pretrained Language Models Generate Persuasive, Faithful, and Informative Ad Text for Product Descriptions?
Fajri Koto, Jey Han Lau and Timothy Baldwin 234

A Simple Baseline for Domain Adaptation in End to End ASR Systems Using Synthetic Data
Raviraj Bhuminand Joshi and Anupam Singh 244

Lot or Not: Identifying Multi-Quantity Offerings in E-Commerce
Gal Lavee and Ido Guy 250