ACL 2022

The 60th Annual Meeting of the Association for Computational Linguistics

Proceedings of the Student Research Workshop

May 22-27, 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 978-1-955917-23-0

Introduction

Welcome to the ACL 2022 Student Research Workshop!

The ACL 2022 Student Research Workshop (SRW) is a forum for student researchers in computational linguistics and natural language processing. The workshop provides a great opportunity for student participants to take part in a mentorship program, present their work and receive valuable feedback from the international research community.

Following the tradition of the previous student research workshops, we have two tracks: research papers and thesis proposals. The research paper track is a venue for students to describe completed work or work-in-progress along with preliminary results. The thesis proposal track is offered for Ph.D. students who have decided on a thesis topic and are interested in getting feedback on their proposal and ideas about future directions for their work.

Mentoring is at the heart of the SRW. In keeping with previous years, we had a pre-submission mentoring program before the submission deadline. Excluding 4 withdrawals and duplicates, a total of 29 papers participated in the pre-submission mentoring program. This program offered students the opportunity to receive feedback from a mentor to improve the writing style and presentation of their submissions.

This year, the student research workshop has again received wide attention. Excluding 7 withdrawals and duplicates, we received 100 submissions including 91 research papers (56 long papers and 35 short papers) and 9 thesis proposals. 3 submissions were desk rejected. 43 submissions (4 thesis proposals, 28 long papers and 11 short papers)were accepted. 1 long paper was withdrawn after acceptance. Excluding non-archival papers, 39 papers appear in these proceedings. All the accepted papers will be presented in person and/or virtually in the poster sessions of the main conference. Some will also have oral presentations.

We are deeply grateful to ACL for providing funds that covered registrations for part of the authors as volunteer students. We thank our program committee members for their careful reviews of each paper and all of our mentors for donating their time to provide feedback to our student authors. Thank you to our faculty advisors, Cecile Paris, Siva Reddy and German Rigau, for their advice and to the ACL 2022 organizing committee for their support. Finally, thank you to our student participants!

Program Committee

Organizers

Samuel Louvan, Fondazione Bruno Kessler
Andrea Madotto, Meta, Reality Lab
Brielen Madureira, University of Potsdam

Faculty Advisors

Cecile Paris, CSIRO, Australia
Siva Reddy, McGill University, Canada
German Rigau, Basque Country University, Spain

Program Committee

Talha olakolu, AI Enablement Department, Huawei Turkey Research and Development Center
Sree Harsha Ramesh, UMass Amherst
Barry Haddow, University of Edinburgh
Mariana Neves, German Federal Institute for Risk Assessment
Omri Abend, The Hebrew University of Jerusalem
Chris Develder, Ghent University
Alina Karakanta, Fondazione Bruno Kessler (FBK), University of Trento
Antoine Venant, Universite Toulouse 3/ IRIT
Glorianna Jagfeld, Spectrum Centre for Mental Health Research, University of Lancaster
Chia-Hui Chang, National Central University
Olga Zamaraeva, University of A Corua
Arkaitz Zubiaga, Queen Mary University of London
Masaaki Nagata, NTT Corporation
Ahsaas Bajaj, University of Massachusetts Amherst
Tatiana Anikina, DFKI / Saarland Informatics Campus
Nina Hosseini-Kivanani, University of Luxembourg
Sarthak Mittal, Mila
Silviu Oprea, University of Edinburgh
Ian Stewart, University of Michigan
Zhengbao Jiang, Carnegie Mellon University
Steven Wilson, Oakland University
Denis Emelin, The University of Edinburgh
Jana Gtze, University of Potsdam
Melissa Roemmele, SDL
Katira Soleymanzadeh, Ege University
Naoki Otani, Carnegie Mellon University
Tomoyuki Kajiwara, Ehime University
Dan Goldwasser, Purdue University
Denis Newman-Griffis, University of Pittsburgh
Arlene Casey, University of Edinburgh
Christopher Homan, Rochester Institute of Technology
Vikas Raunak, Microsoft
Anne Beyer, University of Potsdam

Mamoru Komachi, Tokyo Metropolitan University
Laurie Burchell, University of Edinburgh
Surangika Ranathunga, university of moratuwa
Chuan-Jie Lin, National Taiwan Ocean University
Kemal Kurniawan, University of Melbourne
Kornraphop Kawintiranon, Georgetown University
Diana Galvan-Sosa, Tohoku University
Nihal V. Nayak, Brown University
Xiang Dai, University of Copenhagen
Ronald Cardenas, University of Edinburgh
Gosse Minnema, University of Groningen
Jingzhou Liu, Carnegie Mellon University
Jad Kabbara, McGill University - MILA
Xinyi Zheng, University of Michigan, Ann Arbor
Carolina Scarton, University of Sheffield
Kirk Roberts, University of Texas Health Science Center at Houston
Junxian He, Carnegie Mellon University
Yasumasa Onoe, The University of Texas at Austin
Parisa Kordjamshidi, Michigan State University
Vasu Sharma, Carnegie Mellon University
Jannis Vamvas, Department of Computational Linguistics, University of Zurich
Kiet Nguyen, University of Information Technology, VNU-HCM
Alok Debnath, Trinity College, Dublin
Vivek Gupta, School of Computing, University of Utah
Vincent Ng, University of Texas at Dallas
Lucy Lin, University of Washington
Farig Sadeque, Educational Testing Service
Bharat Ram Ambati, Apple Inc.
Jeff Jacobs, Columbia University
Durgesh Nandini, University of Bamberg
Sudipta Kar, Amazon Alexa AI
Shruti Rijhwani, Carnegie Mellon University
Miguel A. Alonso, Universidade da Corua
Merel Scholman, Saarland University
Valentina Pyatkin, Bar-Ilan University
Bonnie Webber, University of Edinburgh
Marco Antonio Sobrevilla Cabezudo, University of So Paulo
Amita Misra, IBM
Deeksha Varshney, Indian Institute of Technology, Patna, India
Abhilasha Ravichander, Carnegie Mellon University
Adithya Pratapa, Carnegie Mellon University
Micha Elsner, The Ohio State University
Haoran Zhang, University of Pittsburgh
Guy Rotman, Faculty of Industrial Engineering and Management, Technion, IIT
Jakob Prange, Georgetown University
Maria Antoniak, Cornell University
Yusu Qian, New York University
Labiba Jahan, Florida International University
Cesare Spinoso-Di Piano, McGill University
Valentin Malykh, Huawei Noah's Ark Lab / Kazan Federal University
Oscar Sainz, University of the Basque Country (UPV/EHU)

ABULIKEMU ABUDUWEILI, Carnegie Mellon University
Bonaventure F. P. Dossou, Jacobs University Bremen
Jifan Chen, UT Austin
Vincent Nguyen, Australian National University & CSIRO Data61
Zihan Liu, Hong Kong University of Science and Technology
Koji Mineshima, Keio University
Rajaswa Patil, TCS Research
Ian Porada, Mila, McGill University
Valeria de Paiva, Topos Institute
Elizabeth Salesky, Johns Hopkins University
Iker Garca-Ferrero, HiTZ Center - Ixa, University of the Basque Country UPV/EHU
Devang Kulshreshtha, MILA Lab, McGill University
Ivan Vuli, University of Cambridge
Zhong Zhou, Carnegie Mellon University
eljko Agi, Unity Technologies
Najoung Kim, New York University
Meaghan Fowlie, Utrecht University
Gabriel Doyle, San Diego State University
Dimosthenis Kontogiorgos, PhD Student
Aitor Ormazabal, University of the Basque Country
Taiwo Kolajo, Federal University Lokoja
Kevin Small, Amazon
Tejas Srinivasan, University of Southern California
Badr Abdullah, Saarland University
Hai Pham, Carnegie Mellon University
Mandy Korpusik, Loyola Marymount University
Dayne Freitag, SRI International
Fangyu Liu, University of Cambridge
Ruken Cakici, METU
Marija Stanojevic, Center for Data Analytics and Biomedical Informatics, Temple University
Di Lu, Dataminr
Sina Sheikholeslami, KTH Royal Institute of Technology
Jorge Balazs, Amazon
Sowmya Vajjala, National Research Council
Da Yin, University of California, Los Angeles (UCLA)
Zeerak Talat, University of Sheffield
Eduardo Blanco, Arizona State University
David Demeter, Northwestern University
David Adelani, Saarland University
Ivaylo Radev, IICT, Bulgarian Academy of Sciences
Zi-Yi Dou, UCLA
Tom Hosking, University of Edinburgh
Abeer Aldayel, King Saud University
Piush Aggarwal, FernUniversitt in Hagen, Computational Linguistics
Endang Pamungkas, Universitas Muhammadiyah Surakarta
David Trye, University of Waikato
Bruno Martins, IST and INESC-ID
Alexandra Lavrentovich, Amazon Alexa
Gunhee Kim, Seoul National University
Aditi Chaudhary, Carnegie Mellon University
Amir Zeldes, Georgetown University

Jasy Suet Yan Liew, School of Computer Sciences, Universiti Sains Malaysia
Marius Mosbach, Saarland University
Carlos Escolano, Universitat Politcnica de Catalunya
Dat Quoc Nguyen, VinAI Research
Evelin Amorim, UFMG
Zara Kancheva, IICT-BAS
Maike Paetzel-Prsmann, University of Potsdam
Shane Steinert-Threlkeld, University of Washington
Marcos Garcia, Universidade de Santiago de Compostela
Manex Agirrezabal, University of Copenhagen
Tina Fang, University of Waterloo
Richard Sproat, Google, Japan
Rachel Bawden, Inria
Philipp Sadler, University of Potsdam
Yumo Xu, University of Edinburgh
Tatjana Scheffler, Ruhr-Universitt Bochum
Jonathan K. Kummerfeld, University of Michigan
Neat Dereli, Adjust GmbH
Meishan Zhang, Harbin Institute of Technology (Shenzhen), China
Valerio Basile, University of Turin
Shabnam Tafreshi, UMD:ARLIS
Anjali Bhavan, University of Washington
Samuel Pecar, PwC
Omid Moradiannasab, Saarland University
Tatiana Bladier, Heinrich Heine University Dsseldorf
Jonas Groschwitz, Saarland University
Christoph Teichmann, Bloomberg LP

Mentor

Mihai Surdeanu, University of Arizona
Yansong Feng, Peking University
Tatjana Scheffler, Ruhr-Universität Bochum
Hanna Suominen, Australian National University
Vincent Ng, utdallas
Rajaswa Patil Patil, TCS Research
Raffaella Bernardi, University of Trento
Jindřich Libovický, LMU Munich
Marco Basaldella, Amazon Alexa AI
Arkaitz Zubiaga Zubiaga, Queen Mary University of London
Clara Vania Vania, Amazon Alexa AI
Yi-Ling Chung Chung, University of Trento/Fondazione Bruno Kessler
Duygu Ataman, New York University
Fajri Koto, University of Melbourne
Miryam de Lhoneux, Uppsala University
Genta Indra Winata Indra Winata, Bloomberg
Alan Ramponi, Fondazione Bruno Kessler
Jinwook Choi, Seoul National University
Derry Tanti Wijaya, Boston University
Valerio Basile, U of Turin
Niranjan Balasubramanian, Stonybrook University

Marco Gaido, University of Trento/Fondazione Bruno Kessler
Greg Durrett, University of Texas, Austin
Sara Tonelli, Fondazione Bruno Kessler
Bonnie Webber, University of Edinburgh
Shujian Huang, Nanjing University
Malihe Alikhani, University of Pittsburgh
Peng Xu, Hong Kong University of Technology
Zhaojiang Lin, Facebook
Xiang Dai, University of Sydney

Table of Contents

Evaluating zero-shot transfers and multilingual models for dependency parsing and POS tagging within the low-resource language family Tupían
Frederic Blum 1

RFBFN: A Relation-First Blank Filling Network for Joint Relational Triple Extraction
Zhe Li, Luoyi Fu, Xinbing Wang, Haisong Zhang and Chenghu Zhou 10

Building a Dialogue Corpus Annotated with Expressed and Experienced Emotions
Tatsuya Ide and Daisuke Kawahara 21

Darkness can not drive out darkness: Investigating Bias in Hate SpeechDetection Models
Fatma Elsafoury 31

Ethical Considerations for Low-resourced Machine Translation
Levon Haroutunian 44

Integrating Question Rewrites in Conversational Question Answering: A Reinforcement Learning Approach
Etsuko Ishii, Bryan Wilie, Yan Xu, Samuel Cahyawijaya and Pascale Fung 55

What Do You Mean by Relation Extraction? A Survey on Datasets and Study on Scientific Relation Classification
Elisa Bassignana and Barbara Plank 67

Logical Inference for Counting on Semi-structured Tables
Tomoya Kurosawa and Hitomi Yanaka 84

GNNer: Reducing Overlapping in Span-based NER Using Graph Neural Networks
Urchade Zaratiana, Nadi Tomeh, Pierre Holat and Thierry Charnois 97

Compositional Semantics and Inference System for Temporal Order based on Japanese CCG
Tomoki Sugimoto and Hitomi Yanaka 104

Combine to Describe: Evaluating Compositional Generalization in Image Captioning
George Pantazopoulos, Alessandro Suglia and Arash Eshghi 115

Towards Unification of Discourse Annotation Frameworks
Yingxue Fu 132

AMR Alignment for Morphologically-rich and Pro-drop Languages
K. Elif Oral and Gülşen Eryiğit 143

Sketching a Linguistically-Driven Reasoning Dialog Model for Social Talk
Alex Lu 153

Scoping natural language processing in Indonesian and Malay for education applications
Zara Maxwelll-Smith, Michelle Kohler and Hanna Suominen 171

English-Malay Cross-Lingual Embedding Alignment using Bilingual Lexicon Augmentation
Ying Hao Lim and Jasy Suet Yan Liew 229

Towards Detecting Political Bias in Hindi News Articles
Samyak Agrawal, Kshitij Gupta, Devansh Gautam and Radhika Mamidi 239

Restricted or Not: A General Training Framework for Neural Machine Translation
Zuchao Li, Masao Utiyama, Eiichiro Sumita and Hai Zhao 245

What do Models Learn From Training on More Than Text? Measuring Visual Commonsense Knowledge
Lovisa Hagström and Richard Johansson 252

TeluguNER: Leveraging Multi-Domain Named Entity Recognition with Deep Transformers
Suma Reddy Duggenpudi, Subba Reddy Oota, Mounika Marreddy and Radhika Mamidi 262

Using Neural Machine Translation Methods for Sign Language Translation
Galina Angelova, Eleftherios Avramidis and Sebastian Möller 273

Flexible Visual Grounding
Yongmin Kim, Chenhui Chu and Sadao Kurohashi 285

A large-scale computational study of content preservation measures for text style transfer and paraphrase generation
Nikolay Babakov, David Dale, Varvara Logacheva and Alexander Panchenko 300

Explicit Object Relation Alignment for Vision and Language Navigation
Yue Zhang and Parisa Kordjamshidi 322

Mining Logical Event Schemas From Pre-Trained Language Models
Lane Lawley and Lenhart Schubert 332

Exploring Cross-lingual Text Detoxification with Large Multilingual Language Models.
Daniil Moskovskiy, Daryna Dementieva and Alexander Panchenko 346

MEKER: Memory Efficient Knowledge Embedding Representation for Link Prediction and Question Answering
Viktoriia Chekalina, Anton Razzhigaev, Albert Sayapin, Evgeny Frolov and Alexander Panchenko 355

Discourse on ASR Measurement: Introducing the ARPOCA Assessment Tool
Megan Merz and Olga Scrivner 366

Pretrained Knowledge Base Embeddings for improved Sentential Relation Extraction
andrea papaluca, Daniel Krefl, Hanna Suominen and Artem Lenskiy 373

Improving Cross-domain, Cross-lingual and Multi-modal Deception Detection
Subhadarshi Panda and Sarah Ita Levitan 383

Automatic Generation of Distractors for Fill-in-the-Blank Exercises with Round-Trip Neural Machine Translation
Subhadarshi Panda, Frank Palma Gomez, Michael Flor and Alla Rozovskaya 391

On the Locality of Attention in Direct Speech Translation
Belen Alastruey, Javier Ferrando, Gerard I. Gállego and Marta R. Costa-jussà 402

Extraction of Diagnostic Reasoning Relations for Clinical Knowledge Graphs
Vimig Socrates 413

Scene-Text Aware Image and Text Retrieval with Dual-Encoder
Shumpei Miyawaki, Taku Hasegawa, Kyosuke Nishida, Takuma Kato and Jun Suzuki 422

Towards Fine-grained Classification of Climate Change related Social Media Text
Roopal Vaid, Kartikey Pant and Manish Shrivastava 434

Deep Neural Representations for Multiword Expressions Detection
Kamil Kanclerz and Maciej Piasecki 444

A Checkpoint on Multilingual Misogyny Identification
Arianna Muti and Alberto Barrón-Cedeño 454

Using dependency parsing for few-shot learning in distributional semantics
Stefania Preda and Guy Emerson 461

A Dataset and BERT-based Models for Targeted Sentiment Analysis on Turkish Texts
Mustafa Melih Mutlu and Arzucan Özgür 467