ACL 2018

The 56th Annual Meeting of the Association for Computational Linguistics

Proceedings of the Student Research Workshop

July 15 - July 20
Melbourne, Australia

©2018 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 978-1-948087-36-0 (Student Research Workshop)

Introduction

Welcome to the ACL 2018 Student Research Workshop!

The ACL 2018 Student Research Workshop (SRW) is a forum for student researchers in computational linguistics and natural language processing. The workshop provides a unique opportunity for student participants to present their work and receive valuable feedback from the international research community as well as from selected mentors.

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

We received 66 submissions in total: 12 research proposals and 54 research papers. We accepted 26 papers, for an acceptance rate of 39%. After withdrawals, 22 papers are appearing in these proceedings, including 6 research proposals and 16 research papers. All of the accepted papers will be presented as posters in lunchtime sessions as a part of the main conference, split across two days (July 17th and 18th).

Mentoring is at the heart of the SRW. In keeping with previous years, students had the opportunity to participate in a pre-submission mentoring program prior to the submission deadline. This program offered students a chance to receive comments from an experienced researcher, in order to improve the quality of the writing and presentation before making their submission. Eleven authors participated in the pre-submission mentoring. In addition, authors of accepted SRW papers are matched with mentors who will meet with the students in person during the workshop. Each mentor prepares in-depth comments and questions prior to the student's presentation, and provides discussion and feedback during the workshop.

We are deeply grateful to our sponsors whose support will enable a number of students to attend the conference, including the Don and Betty Walker Scholarship Fund, Roam Analytics, Google, and the National Science Foundation under award No. 1827830. We would also like to 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 Marie-Catherine de Marneffe, Wanxiang Che, and Malvina Nissim for their essential advice and guidance, and to the members of the ACL 2018 organizing committee, in particular Claire Cardie, Yusuke Miyao, and Priscilla Rasmussen for their helpful support. Finally, thank you to our student participants!

Organizers:

Vered Shwartz, Bar-Ilan University
Jeniya Tabassum, Ohio State University
Rob Voigt, Stanford University

Faculty Advisors:

Wanxiang Che, Harbin Institute of Technology
Marie-Catherine de Marneffe, Ohio State University
Malvina Nissim, University of Groningen

Program Committee:

Roee Aharoni, Bar Ilan University
Fernando Alva-Manchego, University of Sheffield
Bharat Ram Ambati, Apple Inc.
Jacob Andreas, University of California, Berkeley
Isabelle Augenstein, Department of Computer Science, University of Copenhagen
Petr Babkin, Rensselaer Polytechnic Institute
Fan Bai, The Ohio State University
Kalika Bali, Microsoft Research Labs
Miguel Ballesteros, IBM Research AI
David Bamman, University of California, Berkeley
Yonatan Belinkov, MIT CSAIL
Yonatan Bisk, University of Washington
Facundo Carrillo, Buenos Aires University
Arun Chaganty, Stanford University
Kuan-Yu Chen, NTUST
Tanmay Chinchore, M S Ramaiah Institute of Technology
Grzegorz Chrupała, Tilburg University
Daria Dzendzik, ADAPT Centre Dublin City University
Kurt Junshean Espinosa, University of the Manchester
Tina Fang, University of Waterloo
Silvia García Méndez, AtlantTIC Center - University of Vigo
Matt Gardner, Allen Institute for Artificial Intelligence
Marjan Ghazvininejad, University of Southern California
Hila Gonen, Bar-Ilan University
Alvin Grissom II, Ursinus College
Vivek Gupta, Microsoft Research
He He, Stanford University
Mohit Iyyer, Allen Institute for Artificial Intelligence
Ganesh Jawahar, IIIT Hyderabad
Vasu Jindal, University of Texas at Dallas
Sudipta Kar, University of Houston
Alina Karakanta, Saarland Univesity
Eliyahu Kiperwasser, Bar-Ilan University
Thomas Kober, University of Edinburgh
Alexandar Konovalov, The Ohio State University

Alice Lai, University of Illinois at Urbana-Champaign
Wuwei Lan, The Ohio State University
Eric Laporte, Université Paris-Est Marne-la-Vallée
Haixia Liu, The University of Nottingham
Haijing Liu, Institute of Software, Chinese Academy of Sciences & University of Chinese Academy of Sciences
Edison Marrese-Taylor, The University of Tokyo
Prashant Mathur, eBay
Brian McMahan, Rutgers University
Paul Michel, Carnegie Mellon University
Saif Mohammad, National Research Council Canada
Taesun Moon, IBM Research
Maria Moritz, University of Goettingen
Nihal V. Nayak, Stride.AI
Denis Newman-Griffis, The Ohio State University
Dat Quoc Nguyen, The University of Melbourne
Vlad Niculae, Cornell University
Suraj Pandey, The Open University
Mohammad Taher Pilehvar, University of Cambridge
Daniel Preoţiuc-Pietro, University of Pennsylvania
Emily Prud'hommeaux, Boston College
Avinesh PVS, UKP Lab, Technische Universität Darmstadt
Sudha Rao, University Of Maryland, College Park
Will Roberts, Humboldt-Universität zu Berlin
Stephen Roller, Facebook
Julia Romberg, Heinrich Heine University Düsseldorf
Farig Sadeque, University of Arizona
Enrico Santus, MIT
Shoetsu Sato, Graduate School of Information Science and Technology, The University of Tokyo
Sabine Schulte im Walde, University of Stuttgart
Sebastian Schuster, Stanford University
Vasu Sharma, Carnegie Mellon University
Sunayana Sitaram, Microsoft Research India
Kevin Small, Amazon
Katira Soleymanzadeh, Ege University
R.Sudhesh Solomon, Indian Institute of Information Technology
Richard Sproat, Google
Gabriel Stanovsky, University of Washington, Allen Institute for Artificial Intelligence
Shang-Yu Su, National Taiwan University
Sandesh Swamy, The Ohio State University
Ivan Vulić, University of Cambridge
Olivia Winn, Columbia University
Zhou Yu, University of California, Davis
Omnia Zayed, Insight Centre for Data Analytics, National University of Ireland Galway
Meishan Zhang, Heilongjiang University, China
Yuan Zhang, Google

Table of Contents

Towards Opinion Summarization of Customer Reviews
Samuel Pecar 1

Sampling Informative Training Data for RNN Language Models
Jared Fernandez and Doug Downey 9

Learning-based Composite Metrics for Improved Caption Evaluation
Naeha Sharif, Lyndon White, Mohammed Bennamoun and Syed Afaq Ali Shah 14

Recursive Neural Network Based Preordering for English-to-Japanese Machine Translation
Yuki Kawara, Chenhui Chu and Yuki Arase 21

Pushing the Limits of Radiology with Joint Modeling of Visual and Textual Information
Sonit Singh 28

Recognizing Complex Entity Mentions: A Review and Future Directions
Xiang Dai 37

Automatic Detection of Cross-Disciplinary Knowledge Associations
Menasha Thilakaratne, Katrina Falkner and Thushari Atapattu 45

Language Identification and Named Entity Recognition in Hinglish Code Mixed Tweets
Kushagra Singh, Indira Sen and Ponnurangam Kumaraguru 52

German and French Neural Supertagging Experiments for LTAG Parsing
Tatiana Bladier, Andreas van Cranenburgh, Younes Samih and Laura Kallmeyer 59

SuperNMT: Neural Machine Translation with Semantic Supersenses and Syntactic Supertags
Eva Vanmassenhove and Andy Way 67

Unsupervised Semantic Abstractive Summarization
Shibhansh Dohare, Vivek Gupta and Harish Karnick 74

Biomedical Document Retrieval for Clinical Decision Support System
Jainisha Sankhavara 84

A Computational Approach to Feature Extraction for Identification of Suicidal Ideation in Tweets
Ramit Sawhney, Prachi Manchanda, Raj Singh and Swati Aggarwal 91

BCSAT : A Benchmark Corpus for Sentiment Analysis in Telugu Using Word-level Annotations
Sreekavitha Parupalli, Vijjini Anvesh Rao and Radhika Mamidi 99

Reinforced Extractive Summarization with Question-Focused Rewards
Kristjan Arumae and Fei Liu 105

Graph-based Filtering of Out-of-Vocabulary Words for Encoder-Decoder Models
Satoru Katsumata, Yukio Matsumura, Hayahide Yamagishi and Mamoru Komachi 112

Exploring Chunk Based Templates for Generating a subset of English Text
Nikhilesh Bhatnagar, Manish Shrivastava and Radhika Mamidi 120

Trick Me If You Can: Adversarial Writing of Trivia Challenge Questions
Eric Wallace and Jordan Boyd-Graber 127

Alignment Analysis of Sequential Segmentation of Lexicons to Improve Automatic Cognate Detection
Pranav A 134

Mixed Feelings: Natural Text Generation with Variable, Coexistent Affective Categories
Lee Kezar 141

Automatic Spelling Correction for Resource-Scarce Languages using Deep Learning
Pravallika Etoori, Manoj Chinnakotla and Radhika Mamidi 146

Automatic Question Generation using Relative Pronouns and Adverbs
Payal Khullar, Konigari Rachna, Mukul Hase and Manish Shrivastava 153

Conference Program

Tuesday July 17, 2018

Towards Opinion Summarization of Customer Reviews
Samuel Pecar

Sampling Informative Training Data for RNN Language Models
Jared Fernandez and Doug Downey

Learning-based Composite Metrics for Improved Caption Evaluation
Naeha Sharif, Lyndon White, Mohammed Bennamoun and Syed Afaq Ali Shah

Recursive Neural Network Based Preordering for English-to-Japanese Machine Translation
Yuki Kawara, Chenhui Chu and Yuki Arase

Pushing the Limits of Radiology with Joint Modeling of Visual and Textual Information
Sonit Singh

Recognizing Complex Entity Mentions: A Review and Future Directions
Xiang Dai

Automatic Detection of Cross-Disciplinary Knowledge Associations
Menasha Thilakaratne, Katrina Falkner and Thushari Atapattu

Language Identification and Named Entity Recognition in Hinglish Code Mixed Tweets
Kushagra Singh, Indira Sen and Ponnurangam Kumaraguru

German and French Neural Supertagging Experiments for LTAG Parsing
Tatiana Bladier, Andreas van Cranenburgh, Younes Samih and Laura Kallmeyer

SuperNMT: Neural Machine Translation with Semantic Supersenses and Syntactic Supertags
Eva Vanmassenhove and Andy Way

Unsupervised Semantic Abstractive Summarization
Shibhansh Dohare, Vivek Gupta and Harish Karnick

Wednesday July 18, 2018

Biomedical Document Retrieval for Clinical Decision Support System
Jainisha Sankhavara

A Computational Approach to Feature Extraction for Identification of Suicidal Ideation in Tweets
Ramit Sawhney, Prachi Manchanda, Raj Singh and Swati Aggarwal

BCSAT : A Benchmark Corpus for Sentiment Analysis in Telugu Using Word-level Annotations
Sreekavitha Parupalli, Vijjini Anvesh Rao and Radhika Mamidi

Reinforced Extractive Summarization with Question-Focused Rewards
Kristjan Arumae and Fei Liu

Graph-based Filtering of Out-of-Vocabulary Words for Encoder-Decoder Models
Satoru Katsumata, Yukio Matsumura, Hayahide Yamagishi and Mamoru Komachi

Exploring Chunk Based Templates for Generating a subset of English Text
Nikhilesh Bhatnagar, Manish Shrivastava and Radhika Mamidi

Trick Me If You Can: Adversarial Writing of Trivia Challenge Questions
Eric Wallace and Jordan Boyd-Graber

Alignment Analysis of Sequential Segmentation of Lexicons to Improve Automatic Cognate Detection
Pranav A

Mixed Feelings: Natural Text Generation with Variable, Coexistent Affective Categories
Lee Kezar

Automatic Spelling Correction for Resource-Scarce Languages using Deep Learning
Pravallika Etoori, Manoj Chinnakotla and Radhika Mamidi

Automatic Question Generation using Relative Pronouns and Adverbs
Payal Khullar, Konigari Rachna, Mukul Hase and Manish Shrivastava