NAACL HLT 2019

The 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

Proceedings of the Student Research Workshop

June 3 - 5, 2019
Minneapolis, Minnesota

This workshop is supported by the National Science Foundation under Grant No. 1907573. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

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 978-1-950737-15-4

Introduction

Welcome to the NAACL-HLT 2019 Student Research Workshop! This year's submissions were organized in two tracks: research papers and thesis proposals.

- Research papers may describe completed work, or work in progress with preliminary results. For these papers, the first author must be a current graduate or undergraduate student.
- Thesis proposals are geared towards students who have decided on a thesis topic and wish to get feedback on their proposal and broader ideas for their continuing work.

This year, we received a total of 54 submissions: 50 of these were research papers, and 4 were thesis proposals. We accepted 22 research papers and 2 thesis proposals, resulting in an overall acceptance rate of 44%. The main authors of the accepted papers represent a variety of countries: Australia, Canada, India (6 papers), Ireland, Japan (2 papers), Hong Kong, South Korea, Taiwan, USA (10 papers). These papers span topics across a wide range of subdisciplines and topics within NLP and CL.

Accepted research papers will be presented as either talks or posters within the NAACL main conference. We have 9 accepted papers being presented as talks and 14 as posters.

Following previous editions of the Student Research Workshop, we have offered students the opportunity to get mentoring feedback before submitting their work for review. Each student that requested pre-submission mentorship was assigned to an experienced researcher who read the paper and provided some comments on how to improve the quality of writing and presentation of the student's work. A total of 22 students participated in the mentorship program. During the workshop itself, we will also provide an on-site mentorship program. Each mentor will meet with their assigned students to provide feedback on their poster or oral presentation, and to discuss their research careers.

We would like to express our gratitude for the financial support from the National Science Foundation (NSF), the Computing Research Association Computing Community Consortium (CRA-CCC), the National Research Council (NRC) Canada, and Google. Thanks to their support, this year's SRW is able to assist students with their registration, travel, and lodging expenses.

We would like to thank the mentors for dedicating their time to help students improve their papers prior to submission, and we thank the members of the program committee for the constructive feedback they have provided for each submitted paper.

This workshop would not have been possible without the help from our faculty advisors, and we thank them for their guidance along this year of workshop preparation. We also thank the organizers of NAACL-HLT 2019 for their continuous support. Finally, we would like to thank all students who have submitted their work to this edition of the Student Research Workshop. We hope our collective effort will be rewarded in the form of an excellent workshop!

Student Chairs:

Sudipta Kar, University of Houston (USA)
Farah Nadeem, University of Washington (USA)
Laura Burdick, University of Michigan (USA)

Faculty Advisers:

Greg Durrett, University of Texas at Austin (USA)
Na-Rae Han, University of Pittsburgh (USA)

Mentors:

Mohit Bansal, University of North Carolina at Chapel Hill
Micha Elsner, The Ohio State University
Katrin Erk, University of Texas at Austin
Jonathan Kummerfeld, University of Michigan
Junyi Jessy Li, University of Texas at Austin
Yang Liu, Liulishuo Inc.
Marie de Marneffe, The Ohio State University
Saif Mohammed, National Research Council Canada
Mari Ostendorf, University of Washington
Verónica Pérez-Rosas, University of Michigan
Swapna Somasundaran, Educational Testing Service

Program Committee:

Gustavo Aguilar, University of Houston
Bharat Ram Ambati, Apple
Antonios Anastasopoulos, Carnegie Mellon University
Maria Antoniak, Cornell University
Rachel Bawden, University of Edinburgh
Jifan Chen, University of Texas at Austin
Volkan Cirik, Carnegie Mellon University
Ahmed Elgohary, University of Maryland, College Park
Hady Elsahar, Université de Lyon
Saadia Gabriel, University of Washington
David Gaddy, University of California, Berkeley
Kevin Gimpel, Toyota Technological Institute at Chicago
Alvin Grissom II, Ursinus College
Na-Rae Han, University of Pittsburgh
Ji He, Citadel
Eric Holgate, University of Texas at Austin
Dirk Hovy, Bocconi University
Aaron Jaech, Facebook Research
Labiba Jahan, Florida International University
Youxuan Jiang, University of Michigan

Daniel Khashabi, University of Pennsylvania
Philipp Koehn, Johns Hopkins University
Jasy Suet Yan Liew, Universiti Sains Malaysia
Angela Lin, University of Texas at Austin
Kevin Lybarger, University of Washington
Suraj Maharjan, Capital One
Amita Misra, IBM Watson
Kenton Murray, University of Notre Dame
Denis Newman-Griffis, The Ohio State University
Yasumasa Onoe, University of Texas at Austin
Aishwarya Padmakumar, University of Texas at Austin
Sheena Panthaplackel, University of Texas at Austin
Sudha Rao, University Of Maryland, College Park
Mohammad Sadegh Rasooli, Facebook AI
Farig Sadeque, The University of Arizona
Niloofar Safi Samghabadi, University of Houston
Maarten Sap, University of Washington
Kevin Small, Amazon
Sandeep Soni, Georgia Institute of Technology
Akshay Srivatsan, Carnegie Mellon University
Alane Suhr, Cornell University
Jeniya Tabassum, The Ohio State University
Trang Tran, University of Washington
Sowmya Vajjala, Iowa State University
John Wieting, Carnegie Mellon University
Adina Williams, Facebook AI Research
Jiacheng Xu, University of Texas at Austin
Fan Yang, Facebook
Vicky Zayats, University of Washington
Ramon Ziai, University of Tübingen

Table of Contents

Conference Program

Monday, June 3, 2019

Session 1F: Question Answering, Sentiment, Machine Translation, Resources and Evaluation (Posters)

Is It Dish Washer Safe? Automatically Answering "Yes/No" Questions Using Customer Reviews
Daria Dzendzik, Carl Vogel and Jennifer Foster

Session 2B: Ethics, Bias and Fairness

16:15–16:30 *Identifying and Reducing Gender Bias in Word-Level Language Models*
Shikha Bordia and Samuel R. Bowman

Session 2C: Style and Sentiment

16:15–16:30 *Emotion Impacts Speech Recognition Performance*
Rushab Munot and Ani Nenkova

Session 2D: Summarization and Information Retrieval

16:15–16:30 *The Strength of the Weakest Supervision: Topic Classification Using Class Labels*
Jiatong Li, Kai Zheng, Hua Xu, Qiaozhu Mei and Yue Wang

Monday, June 3, 2019 (continued)

Session 2E: Tagging, Chunking, Syntax and Parsing

16:15–16:30 *Handling Noisy Labels for Robustly Learning from Self-Training Data for Low-Resource Sequence Labeling*
Debjit Paul, Mittul Singh, Michael A. Hedderich and Dietrich Klakow

Session 2F: Information Extraction, Generation and Semantics (Posters)

Opinion Mining with Deep Contextualized Embeddings
Wen-Bin Han and Noriko Kando

A Bag-of-concepts Model Improves Relation Extraction in a Narrow Knowledge Domain with Limited Data
Jiyu Chen, Karin Verspoor and Zenan Zhai

Generating Text through Adversarial Training Using Skip-Thought Vectors
Afroz Ahamad

A Partially Rule-Based Approach to AMR Generation
Emma Manning

Computational Investigations of Pragmatic Effects in Natural Language
Jad Kabbara

Monday, June 3, 2019 (continued)

Session 3F: Applications, Social Media, Biomedical NLP and Clinical Text Processing (Posters)

SEDTWik: Segmentation-based Event Detection from Tweets Using Wikipedia
Keval Morabia, Neti Lalita Bhanu Murthy, Aruna Malapati and Surender Samant

Tuesday, June 4, 2019

Session 5B: Machine Translation

12:15–12:30 *Multimodal Machine Translation with Embedding Prediction*
Tosho Hirasawa, Hayahide Yamagishi, Yukio Matsumura and Mamoru Komachi

Session 6F: Phonology, Speech and Text Mining (Posters)

Deep Learning and Sociophonetics: Automatic Coding of Rhoticity Using Neural Networks
Sarah Gupta and Anthony DiPadova

Data Augmentation by Data Noising for Open-vocabulary Slots in Spoken Language Understanding
Hwa-Yeon Kim, Yoon-Hyung Roh and Young-Kil Kim

Expectation and Locality Effects in the Prediction of Disfluent Fillers and Repairs in English Speech
Samvit Dammalapati, Rajakrishnan Rajkumar and Sumeet Agarwal

Gating Mechanisms for Combining Character and Word-level Word Representations: an Empirical Study
Jorge Balazs and Yutaka Matsuo

A Pregroup Representation of Word Order Alternation Using Hindi Syntax
Alok Debnath and Manish Shrivastava

Wednesday, June 5, 2019

Session 9B: Applications

16:15–16:30 *Speak up, Fight Back! Detection of Social Media Disclosures of Sexual Harassment*
Arijit Ghosh Chowdhury, Ramit Sawhney, Puneet Mathur, Debanjan Mahata and Rajiv Ratn Shah

Session 9D: Cognitive Modeling and Psycholinguistics

15:30–15:45 *SNAP-BATNET: Cascading Author Profiling and Social Network Graphs for Suicide Ideation Detection on Social Media*
Rohan Mishra, Pradyumn Prakhar Sinha, Ramit Sawhney, Debanjan Mahata, Puneet Mathur and Rajiv Ratn Shah