ACL 2018

Natural Language Processing for Social Media

Proceedings of the Sixth Workshop
AFNLP SIG SocialNLP

ACL 2018 Workshop
July 20, 2018
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 978-1-948087-48-3

SocialNLP 2018@ACL Chairs' Welcome

It is our great pleasure to welcome you to the Sixth Workshop on Natural Language Processing for Social Media-SocialNLP 2018, associated with ACL 2018. SocialNLP is an inter-disciplinary area of natural language processing (NLP) and social computing. We hold SocialNLP twice a year: one in the NLP venue, the other in the associated venue such as those for web technology or artificial intelligence. There are three plausible directions of SocialNLP: (1) addressing issues in social computing using NLP techniques; (2) solving NLP problems using information from social media; and (3) handling new problems related to both social computing and natural language processing. Through this workshop, we anticipate to provide a platform for research outcome presentation and head-to-head discussion in the area of SocialNLP, with the hope to combine the insight and experience of prominent researchers from both NLP and social computing domains to contribute to the area of SocialNLP jointly. The submissions to this year's workshop were again of high quality and we had a competitive selection process. We received submissions from Asia, Europe, and the United States, and due to a rigorous review process, we only accepted 7 long oral papers among 19. Thus the acceptance rate was 37 percent. In addition, we are having our first research challenge: EmotionX in the SocialNLP workshop. A total of 18 groups registered for the dataset and at last 5 groups submitted their results successfully. Therefore these year we have 7 research papers, 1 challenge overview paper as well as 5 challenge papers, that is, a total of 13 papers published in the proceedings of the ACL chapter for the SocialNLP 2018 workshop.

This year, we are excited to have Dr. Saif Mohammad from National Research Council Canada, and Dr. Yi-Chia Wang from Uber as our keynote speakers. We also encourage attendees to attend the keynote talk presentations to have more discussions with outstanding researchers. Their valuable and insightful talk can and will guide us to a better understanding of the future. Putting together SocialNLP 2018 was a team effort. We first thank the authors for providing the quality content of the program. We are grateful to the program committee members, who worked very hard in reviewing papers and providing feedback for authors. Finally, we especially thank the Workshop Committee Chairs Prof. Brendan O'Connor and Prof. Eva Maria Vecchi.

We hope you join our community and enjoy the workshop!

Organizers

Lun-Wei Ku, Academia Sincia, Taiwan
Cheng-Te Li, National Cheng Kung University, Taiwan

Organizers

Lun-Wei Ku, Academia Sincia, Taiwan
Cheng-Te Li, National Cheng Kung University, Taiwan

Program Committee:

Sabine Bergler, Concordia University
Yung-Chun Chang, Taipei Medical University
Hsin-Hsi Chen, National Taiwan University
Freddy Chua, Singapore Management University
Danilo Croce, University of Roma, Tor Vergata
Lei Cui, Microsoft Research
Ronan Cummins, University of Cambridge
Min-Yuh Day, Tamkang University, Taiwan
Ann Devitt, Trinity College Dublin
Koji Eguchi, Kobe University
Michael Elhadad, Ben-Gurion University
Wei Gao, Qatar Computing Research Institute
Spandana Gella, University of Edinburgh
Weiwei Guo, LinkedIn
William Hamilton, Stanford University
Graeme Hirst, University of Toronto
Wen-Lian Hsu, Academia Sinica
Hen-Hsen Huang, National Taiwan University
Diana Inkpen, University of Ottawa
David Jurgens, Stanford University
Roman Klinger, University of Stuttgart
Tsung-Ting Kuo, University of California, San Diego
Cheng-Te Li, National Cheng Kung University
Chuan-Jie Lin, National Taiwan Ocean University
Shou-De Lin, National Taiwan University
Zhiyuan Liu, Tsinghua University
Bin Lu, Google Inc.
Zhunchen Luo, China Defense Science and Technology Information Center
Bruno Martins, University of Lisbon
Yelena Mejova, Qatar Computing Research Institute
Rada Mihalcea, University of Michigan
Manuel Montes-y-Gómez, INAOE, Mexico
Dong Nguyen, University of Twente
Haris Papageorgiou, ATHENA Research and Innovation Center
Georgios Petasis, NCSR "Demokritos"
Stephen Pulman, Oxford University
Saurav Sahay, Intel Labs
Mário J. Silva, Universidade de Lisboa
Yanchuan Sim, Institute for Infocomm Research
Jan Snajder, University of Zagreb
Xavier Tannier, Université Paris-Sud, LIMSI, CNRS

Mike Thelwall, University of Wolverhampton
Ming-Feng Tsai, National Chengchi University
Svitlana Volkova, PNNL
Xiaojun Wan, Peking University
Hsin-Min Wang, Academia Sinica
Jenq-Haur Wang, National Taipei University of Technology
William Yang Wang, UC Santa Barbara
Ingmar Weber, Qatar Computing Research Institute, HBKU
Shih-Hung Wu, Chaoyang University of Technology
Ruifeng Xu, Harbin Institute of Technology
Yi Yang, Georgia Tech
Liang-Chih Yu, Yuan Ze University
Zhe Zhang, IBM Watson

Keynote Speaker:

Saif Mohammad, National Research Council Canada
Yi-Chia Wang, Uber

Challenge Chair:

Chao-Chun Hsu, Academia Sinica
Lun-Wei Ku, Academia Sinica

Keynote Speech I

The Search for Emotions, Creativity, and Fairness in Language

Speaker: Saif M. Mohammad, Senior Research Scientist, National Research Council Canada

Abstract: Emotions are central to human experience, creativity, and behavior. They are crucial for organizing meaning and reasoning about the world we live in. They are ubiquitous and everyday, yet complex and nuanced. In this talk, I will describe our work on the search for emotions in language – by humans (through data annotation projects) and by machines (in automatic emotion detection systems).

I will outline ways in which emotions can be represented, challenges in obtaining reliable annotations, and approaches that lead to high-quality annotations. The lexicons thus created have entries for tens of thousands of terms. They provide fine-grained scores for basic emotions as well as for valence, arousal, and dominance (argued by some to be the core dimensions of meaning). They have wide-ranging applications in natural language processing, psychology, social sciences, digital humanities, and computational creativity. I will highlight some of the applications we have explored in literary analysis and automatic text-based music generation. I will also discuss new sentiment analysis tasks such as inferring fine-grained emotion intensity and stance from tweets, as well as detecting emotions evoked by art. I will conclude with work on quantifying biases in the way language is used and the impact of such biases on automatic emotion detection systems. From social media to home assistants, from privacy concerns to neuro-cognitive persuasion, never has natural language processing been more influential, more fraught with controversy, and more entrenched in everyday life. Thus as a community, we are uniquely positioned to make substantial impact by building applications that are not only compelling and creative but also facilitators of social equity and fairness.

Bio: Dr. Saif M. Mohammad is Senior Research Scientist at the National Research Council Canada (NRC). He received his Ph.D. in Computer Science from the University of Toronto. Before joining NRC, Saif was a Research Associate at the Institute of Advanced Computer Studies at the University of Maryland, College Park. His research interests are in Emotion and Sentiment Analysis, Computational Creativity, Psycholinguistics, Fairness in Machine Learning, Crowdsourced Human Annotations, Social Media Language, and Information Visualization. Saif is a co-organizer of WASSA (a sentiment analysis workshop) and co-chair of SemEval (the largest platform for semantic evaluations). He has also served as the area chair for Sentiment Analysis in ACL conferences. His work on emotions has garnered media attention, with articles in Time, Washington Post, Slashdot, LiveScience, The Physics arXiv Blog, PC World, Popular Science, etc. Webpage: http://saifmohammad.com

Keynote Speech II

Understanding Online Social Behaviors through Automatic Language Analysis

Speaker: Yi-Chia Wang, Data Scientist, Uber

Abstract: In online environments, people accomplish their social goals through the use of language - for example, presenting themselves appropriately on social networking sites, attracting followers in social media, or eliciting support in health support groups. In order to understand how people accomplish these goals and further design interventions to help people achieve them, we need sophisticated and scalable approaches to language analysis. My research investigates communication dynamics in online social environments. The goal is to understand how people use language to communicate with others online and its social outcomes, how language presentations are different in various types of online environments and provide guidance for practitioners to improve their services.

In this talk, I will present studies examining user behaviors in online environments. My research method consists of two phases: (1) developing machine learning models to automatically measure language concepts and (2) applying the models to analyze text at scale and quantitatively relate language concepts to user behaviors. The presentation will cover how I applied a similar research method to answer very different research questions and provide implications for practitioners.

Bio:

Yi-Chia Wang received her Ph.D. from the Language Technologies Institute in School of Computer Science at Carnegie Mellon University. Her research interests and skills are to combine language processing technologies, machine learning methodologies, and social science theories to statistically analyze large-scale data and understand user behaviors in online environments. Her thesis developed a machine learning model to study self-disclosure on Facebook. She also had experience on question answering and information extraction. She is currently a Data Scientist at Uber, focusing on customer support and conversational AI domains.

Table of Contents

Sociolinguistic Corpus of WhatsApp Chats in Spanish among College Students
Alejandro Dorantes, Gerardo Sierra, Tlauhlia Yamín Donohue Pérez, Gemma Bel-Enguix and Mónica Jasso Rosales 1

A Crowd-Annotated Spanish Corpus for Humor Analysis
Santiago Castro, Luis Chiruzzo, Aiala Rosá, Diego Garat and Guillermo Moncecchi 7

A Twitter Corpus for Hindi-English Code Mixed POS Tagging
Kushagra Singh, Indira Sen and Ponnurangam Kumaraguru 12

Detecting Offensive Tweets in Hindi-English Code-Switched Language
Puneet Mathur, Rajiv Shah, Ramit Sawhney and Debanjan Mahata 18

SocialNLP 2018 EmotionX Challenge Overview: Recognizing Emotions in Dialogues
Chao-Chun Hsu and Lun-Wei Ku 27

EmotionX-DLC: Self-Attentive BiLSTM for Detecting Sequential Emotions in Dialogues
Linkai Luo, Haiqin Yang and Francis Y. L. Chin 32

EmotionX-AR: CNN-DCNN autoencoder based Emotion Classifier
Sopan Khosla 37

EmotionX-SmartDubai_NLP: Detecting User Emotions In Social Media Text
Hessa AlBalooshi, Shahram Rahmanian and Rahul Venkatesh Kumar 45

EmotionX-Area66: Predicting Emotions in Dialogues using Hierarchical Attention Network with Sequence Labeling
Rohit Saxena, savita bhat and Niranjan Pedanekar 50

EmotionX-JTML: Detecting emotions with Attention
Johnny Torres 56

Towards Automation of Sense-type Identification of Verbs in OntoSenseNet
Sreekavitha Parupalli, Vijjini Anvesh Rao and Radhika Mamidi 61

Improving Classification of Twitter Behavior During Hurricane Events
Kevin Stowe, Jennings Anderson, Martha Palmer, Leysia Palen and Ken Anderson 67

Political discourse classification in social networks using context sensitive convolutional neural networks
Aritz Bilbao-Jayo and Aitor Almeida 76

Conference Program

July 20, 2018

09:20–10:30 *Keynote Speech (I):The Search for Emotions, Creativity, and Fairness in Language*
Dr. Saif Mohammad (NSF)

10:30–11:00 Coffee Break

11:00–12:20 Technical Session 1

Sociolinguistic Corpus of WhatsApp Chats in Spanish among College Students
Alejandro Dorantes, Gerardo Sierra, Tlauhlia Yamín Donohue Pérez, Gemma Bel-Enguix and Mónica Jasso Rosales

A Crowd-Annotated Spanish Corpus for Humor Analysis
Santiago Castro, Luis Chiruzzo, Aiala Rosá, Diego Garat and Guillermo Moncecchi

A Twitter Corpus for Hindi-English Code Mixed POS Tagging
Kushagra Singh, Indira Sen and Ponnurangam Kumaraguru

Detecting Offensive Tweets in Hindi-English Code-Switched Language
Puneet Mathur, Rajiv Shah, Ramit Sawhney and Debanjan Mahata

12:20–13:20 Lunch

13:20–14:30 *Keynote Speech (II): Understanding Online Social Behaviors through Automatic Language Analysis*
Dr. Yi-Chia Wang (Uber)

July 20, 2018 (continued)

14:30–15:30 EmotionX Challenge Session

SocialNLP 2018 EmotionX Challenge Overview: Recognizing Emotions in Dialogues
Chao-Chun Hsu and Lun-Wei Ku

EmotionX-DLC: Self-Attentive BiLSTM for Detecting Sequential Emotions in Dialogues
Linkai Luo, Haiqin Yang and Francis Y. L. Chin

EmotionX-AR: CNN-DCNN autoencoder based Emotion Classifier
Sopan Khosla

EmotionX-SmartDubai_NLP: Detecting User Emotions In Social Media Text
Hessa AlBalooshi, Shahram Rahmanian and Rahul Venkatesh Kumar

EmotionX-Area66: Predicting Emotions in Dialogues using Hierarchical Attention Network with Sequence Labeling
Rohit Saxena, savita bhat and Niranjan Pedanekar

EmotionX-JTML: Detecting emotions with Attention
Johnny Torres

15:30–16:00 Coffee Break

July 20, 2018 (continued)

16:00–17:00 Technical Session 2

Towards Automation of Sense-type Identification of Verbs in OntoSenseNet
Sreekavitha Parupalli, Vijjini Anvesh Rao and Radhika Mamidi

Improving Classification of Twitter Behavior During Hurricane Events
Kevin Stowe, Jennings Anderson, Martha Palmer, Leysia Palen and Ken Anderson

Political discourse classification in social networks using context sensitive convolutional neural networks
Aritz Bilbao-Jayo and Aitor Almeida

17:00–17:10 Closing