NAACL HLT 2019

The 2nd Clinical Natural Language Processing Workshop

Proceedings of the Workshop

June 7, 2019
Minneapolis, Minnesota, USA

Sponsored by:

©2019 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 978-1-948087-94-0

Preface

This volume contains papers from the 2nd Workshop on Clinical Natural Language Processing (ClinicalNLP), held at NAACL 2019.

Clinical text offers unique challenges that differentiate it not only from open-domain data, but from other types of text in the biomedical domain as well. Notably, clinical text contains a significant number of abbreviations, medical terms, and other clinical jargon. Clinical narratives are characterized by non-standard document structures that are often critical to overall understanding. Narrative provider notes are designed to communicate with other experts while at the same time serving as a legal record. Finally, clinical notes contain sensitive patient-specific information that raise privacy and security concerns that present special challenges for natural language systems. This workshop focuses on the work that develops methods to address the above challenges, with the goal to advance state-of-the-art in clinical NLP.

This year, we received the total of 28 submissions, out of which 10 were accepted as oral presentations and 10 as posters.

Organizers:

Anna Rumshisky, UMass Lowell
Kirk Roberts, University of Texas Health Science Center at Houston
Steven Bethard, University of Arizona
Tristan Naumann, Microsoft Research

Program Committee:

Sabine Bergler, Concordia University
Parminder Bhatia, Amazon
Vivek Datla, Philips Research North America
Dina Demner-Fushman, National Library of Medicine
Dmitriy Dligach, Loyola University
Jungwei Fan, Mayo Clinic
Sadid Hasan, Philips North America
Lynette Hirschman, The MITRE Corporation
Yoshinobu Kano, Shizuoka University
Kathy Lee, Philips Research North America
Stephane Meystre, University of Utah School of Medicine
Timothy Miller, Boston Children's Hospital
Martha Palmer, UC Boulder
Hoifung Poon, Microsoft Research
Ashequl Qadir, Philips Research North America
Chaitanya Shivade, IBM
Weiyi Sun, Nuance
Sumithra Velupillai, KTH Royal Institute of Technology
Karin Verspoor, The University of Melbourne
Byron Wallace, Northeastern University
Ben Wellner, The MITRE Corporation
Jenna Wiens, University of Michigan
Stephen Wu, University of Texas Health Science Center at Houston

Invited Speaker:

Heng Ji, University of Illinois at Urbana-Champaign

Invited Panelists:

Hongfang Liu, Mayo Clinic
Piet de Groen, University of Minnesota
Elmer Bernstam, University of Texas Health Science Center at Houston

Table of Contents

Effective Feature Representation for Clinical Text Concept Extraction
Yifeng Tao, Bruno Godefroy, Guillaume Genthial and Christopher Potts 1

An Analysis of Attention over Clinical Notes for Predictive Tasks
Sarthak Jain, Ramin Mohammadi and Byron C. Wallace 15

Extracting Adverse Drug Event Information with Minimal Engineering
Timothy Miller, Alon Geva and Dmitriy Dligach 22

Hierarchical Nested Named Entity Recognition
Zita Marinho, Afonso Mendes, Sebastião Miranda and David Nogueira 28

Towards Automatic Generation of Shareable Synthetic Clinical Notes Using Neural Language Models
Oren Melamud and Chaitanya Shivade 35

A Novel System for Extractive Clinical Note Summarization using EHR Data
Jennifer Liang and Ching-Huei Tsou 46

Study of lexical aspect in the French medical language. Development of a lexical resource
Agathe Pierson and Cédrick Fairon 55

A BERT-based Universal Model for Both Within- and Cross-sentence Clinical Temporal Relation Extraction
Chen Lin, Timothy Miller, Dmitriy Dligach, Steven Bethard and Guergana Savova 65

Publicly Available Clinical BERT Embeddings
Emily Alsentzer, John Murphy, William Boag, Wei-Hung Weng, Di Jindi, Tristan Naumann and Matthew McDermott 72

A General-Purpose Annotation Model for Knowledge Discovery: Case Study in Spanish Clinical Text
Alejandro Piad-Morffis, Yoan Guitérrez, Suilan Estevez-Velarde and Rafael Muñoz 79

Predicting ICU transfers using text messages between nurses and doctors
Faiza Khan Khattak, Chloe Pou-Prom, Robert Wu and Frank Rudzicz 89

Medical Entity Linking using Triplet Network
Ishani Mondal, Sukannya Purkayastha, Sudeshna Sarkar, Pawan Goyal, Jitesh Pillai, Amitava Bhattacharyya and Mahanandeeshwar Gattu 95

Annotating and Characterizing Clinical Sentences with Explicit Why-QA Cues
Jungwei Fan 101

Extracting Factual Min/Max Age Information from Clinical Trial Studies
Yufang Hou, Debasis Ganguly, Lea Deleris and Francesca Bonin 107

Distinguishing Clinical Sentiment: The Importance of Domain Adaptation in Psychiatric Patient Health Records
Eben Holderness, Philip Cawkwell, Kirsten Bolton, James Pustejovsky and Mei-Hua Hall 117

Medical Word Embeddings for Spanish: Development and Evaluation
Felipe Soares, Marta Villegas, Aitor Gonzalez-Agirre, Martin Krallinger and Jordi Armengol-Estapé 124

Attention Neural Model for Temporal Relation Extraction
Sijia Liu, Liwei Wang, Vipin Chaudhary and Hongfang Liu 134

Automatically Generating Psychiatric Case Notes From Digital Transcripts of Doctor-Patient Conversations
Nazmul Kazi and Indika Kahanda 140

Clinical Data Classification using Conditional Random Fields and Neural Parsing for Morphologically Rich Languages
Razieh Ehsani, Tyko Niemi, Gaurav Khullar and Tiina Leivo 149

Conference Program

Friday, June 7, 2019

09:00–09:15 *Opening remarks*

09:15–10:30 *Invited speaker: Heng Ji*

10:30–11:20 *Coffee break*

11:20–12:30 ***Oral session 1***

11:20–11:40 *Effective Feature Representation for Clinical Text Concept Extraction*
Yifeng Tao, Bruno Godefroy, Guillaume Genthial and Christopher Potts

11:40–11:55 *An Analysis of Attention over Clinical Notes for Predictive Tasks*
Sarthak Jain, Ramin Mohammadi and Byron C. Wallace

11:55–12:10 *Extracting Adverse Drug Event Information with Minimal Engineering*
Timothy Miller, Alon Geva and Dmitriy Dligach

12:10–12:25 *Hierarchical Nested Named Entity Recognition*
Zita Marinho, Afonso Mendes, Sebastião Miranda and David Nogueira

12:30–14:00 *Lunch*

14:00–15:30 *Oral session 2*

14:00–14:20 *Towards Automatic Generation of Shareable Synthetic Clinical Notes Using Neural Language Models*
Oren Melamud and Chaitanya Shivade

14:20–14:40 *A Novel System for Extractive Clinical Note Summarization using EHR Data*
Jennifer Liang and Ching-Huei Tsou

Friday, June 7, 2019 (continued)

14:40–14:55 *Study of lexical aspect in the French medical language. Development of a lexical resource*
Agathe Pierson and Cédrick Fairon

14:55–15:10 *A BERT-based Universal Model for Both Within- and Cross-sentence Clinical Temporal Relation Extraction*
Chen Lin, Timothy Miller, Dmitriy Dligach, Steven Bethard and Guergana Savova

15:10–15:25 *Publicly Available Clinical BERT Embeddings*
Emily Alsentzer, John Murphy, William Boag, Wei-Hung Weng, Di Jindi, Tristan Naumann and Matthew McDermott

15:30–16:00 *Coffee break*

16:00–16:45 ***Poster session***

A General-Purpose Annotation Model for Knowledge Discovery: Case Study in Spanish Clinical Text
Alejandro Piad-Morffis, Yoan Guitérrez, Suilan Estevez-Velarde and Rafael Muñoz

Predicting ICU transfers using text messages between nurses and doctors
Faiza Khan Khattak, Chloe Pou-Prom, Robert Wu and Frank Rudzicz

Medical Entity Linking using Triplet Network
Ishani Mondal, Sukannya Purkayastha, Sudeshna Sarkar, Pawan Goyal, Jitesh Pillai, Amitava Bhattacharyya and Mahanandeeshwar Gattu

Annotating and Characterizing Clinical Sentences with Explicit Why-QA Cues
Jungwei Fan

Extracting Factual Min/Max Age Information from Clinical Trial Studies
Yufang Hou, Debasis Ganguly, Lea Deleris and Francesca Bonin

Distinguishing Clinical Sentiment: The Importance of Domain Adaptation in Psychiatric Patient Health Records
Eben Holderness, Philip Cawkwell, Kirsten Bolton, James Pustejovsky and Mei-Hua Hall

Medical Word Embeddings for Spanish: Development and Evaluation
Felipe Soares, Marta Villegas, Aitor Gonzalez-Agirre, Martin Krallinger and Jordi Armengol-Estapé

Friday, June 7, 2019 (continued)

Attention Neural Model for Temporal Relation Extraction
Sijia Liu, Liwei Wang, Vipin Chaudhary and Hongfang Liu

Automatically Generating Psychiatric Case Notes From Digital Transcripts of Doctor-Patient Conversations
Nazmul Kazi and Indika Kahanda

Clinical Data Classification using Conditional Random Fields and Neural Parsing for Morphologically Rich Languages
Razieh Ehsani, Tyko Niemi, Gaurav Khullar and Tiina Leivo

16:45–17:30 *Panel discussion: NLP vs. structured data in the clinical domain (with panelists: Hongfang Liu, Piet de Groen, Elmer Bernstam)*

17:30–17:45 *Closing remarks*