NAACL HLT 2015

Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality

Proceedings of the Second Workshop

June 5, 2015
Denver, Colorado, USA

Sponsor:

©2015 The Association for Computational Linguistics

Order print-on-demand copies from:

Curran Associates
57 Morehouse Lane
Red Hook, New York 12571
USA
Tel: +1-845-758-0400
Fax: +1-845-758-2633
curran@proceedings.com

Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality
ISBN 978-1-941643-43-3

Introduction

In the United States, mental health problems are among the most costly challenges we face. The numbers are staggering: An estimated $57.5B was spent on mental health care in 2006. Some 25 million American adults will have an episode of major depression this year, and suicide is the third leading cause of death for people between 10 and 24 years old. The importance of clinical psychology as a problem space cannot be overstated.

For clinical psychologists, language plays a central role in diagnosis, and many clinical instruments fundamentally rely on manual coding of patient language. Applying language technology in this domain can have an enormous impact: Many individuals under-report psychiatric symptoms, such as active duty soldiers; or lack the self-awareness to report accurately, such as individuals involved in substance abuse who do not recognize their own addiction. Many people cannot even obtain access to a clinician who is qualified to perform a psychological evaluation, such as those without adequate insurance or who live in rural areas. Bringing language technology to bear on these problems could lead to inexpensive screening measures that may be administered by a wider array of healthcare professionals, suited to the realities of healthcare practice.

Researchers have begun targeting such issues, applying computational linguistic methods to clinical psychology with compelling results. Prior to this workshop series, research had looked at identifying emotion in suicide notes, analyzing the language of those with autistic spectrum disorders, aiding the diagnosis of dementia, and screening for depression.

ACL 2014 hosted the first Computational Linguistics and Clinical Psychology Workshop, which brought together the researchers in this nascent field. This workshop was a great success, with accepted papers proposing methods for predicting veteran suicide risk, aiding the diagnosis of dementia, and predicting depression and post-traumatic stress order in social media.

NAACL 2015 hosts the second Computational Linguistics and Clinical Psychology Workshop. Members of the community have come together to organize a hackathon, with a data release and shared task for detecting mental illness as part of this workshop. We hope to build the momentum towards releasing tools and data that can be used by clinical psychologists, and as such, we diverge from the conventional "mini-conference" workshop format, including practicing clinical psychologists on our program committee and as discussants in the workshop. The ability to communicate relevant computational methods and results clearly, connecting the work to clinical practice, is as important as the quality of the work itself, and more important than research novelty.

We received 15 submissions for the main workshop and 3 for the shared task. Of the main workshop submissions, 12 (80%) were accepted: 6 for oral and 6 for poster presentation. Oral presentations will be followed by discussions led by several experts on working with patients and clinical data: Shandra M. Brown Levey, Loring J. Ingraham, John P. Pestian, and Kytja K. S. Voeller. We also have an invited talk from Munmun De Choudhury, an expert in computational social science who has done pioneering work on understanding mental health in social media.

We wish to thank everyone who showed interest and submitted a paper, all of the authors for their contributions, the members of the Program Committee for their thoughtful reviews, our clinical discussants for their helpful insights, and all the attendees of the workshop. We also wish to extend thanks to the Association for Computational Linguistics for making this workshop possible, and to Microsoft Research for its generous sponsorship.

– Meg, Glen, and Kristy

Organizers:

Margaret Mitchell, Microsoft Research (MSR)
Glen Coppersmith, Qntfy
Kristy Hollingshead, Florida Institute for Human and Machine Cognition (IHMC)

Clinical Discussants:

Shandra M. Brown Levey, University of Colorado Denver
Loring J. Ingraham, George Washington University
John P. Pestian, Cincinnati Children's Hospital Medical Center
Kytja K. S. Voeller, Western Institute for Neurodevelopmental Studies and Interventions

Program Committee:

Steven Bedrick, Oregon Health & Science University
Wei Chen, Nationwide Children's Hospital
Glen Coppersmith, Qntfy
Mark Dredze, Johns Hopkins University
Michael Gamon, Microsoft Research
Kimberly Glasgow, Johns Hopkins Applied Physics Laboratory
Dan Goldwasser, University of Maryland
Graeme Hirst, University of Toronto
Christopher Homan, Rochester Institute of Technology
Loring J. Ingraham, George Washington University
William Jarrold, Nuance Communications
Yangfeng Ji, Georgia Institute of Technology
Tong Liu, Rochester Institute of Technology
Antolin Llorente, Mt. Washington Pediatric Hospital
Aimee Mooney, Oregon Health & Science University
Eric Morley, Oregon Health & Science University
Sylvester Olubolu Orimaye, Monash University Malaysia
Cecilia Ovesdotter Alm, Rochester Institute of Technology
Craig Pfeifer, The MITRE Corporation
Matthew Purver, Queen Mary University of London
Philip Resnik, University of Maryland
Rebecca Resnik, Mindwell Psychology Bethesda
Brian Roark, Google
Masoud Rouhizadeh, Oregon Health & Science University
Ronald Schouten, Harvard Medical School
H. Andrew Schwartz, University of Pennsylvania
Richard Sproat, Google
Hiroki Tanaka, NAIST
Paul Thompson, Dartmouth College
Jan van Santen, Oregon Health & Science University

Invited Speaker:

Munmun De Choudhury, Georgia Tech

Table of Contents

From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses
Glen Coppersmith, Mark Dredze, Craig Harman and Kristy Hollingshead 1

Quantifying the Language of Schizophrenia in Social Media
Margaret Mitchell, Kristy Hollingshead and Glen Coppersmith 11

The role of personality, age, and gender in tweeting about mental illness
Daniel Preoţiuc-Pietro, Johannes Eichstaedt, Gregory Park, Maarten Sap, Laura Smith, Victoria Tobolsky, H. Andrew Schwartz and Lyle Ungar 21

CLPsych 2015 Shared Task: Depression and PTSD on Twitter
Glen Coppersmith, Mark Dredze, Craig Harman, Kristy Hollingshead and Margaret Mitchell . 31

Mental Illness Detection at the World Well-Being Project for the CLPsych 2015 Shared Task
Daniel Preoţiuc-Pietro, Maarten Sap, H. Andrew Schwartz and Lyle Ungar 40

Screening Twitter Users for Depression and PTSD with Lexical Decision Lists
Ted Pedersen 46

The University of Maryland CLPsych 2015 Shared Task System
Philip Resnik, William Armstrong, Leonardo Claudino and Thang Nguyen 54

Computational cognitive modeling of inflectional verb morphology in Spanish-speakers for the characterization and diagnosis of Alzheimer's Disease
M. Dolores del Castillo, J. Ignacio Serrano and Jesús Oliva 61

Recursive Neural Networks for Coding Therapist and Patient Behavior in Motivational Interviewing
Michael Tanana, Kevin Hallgren, Zac Imel, David Atkins, Padhraic Smyth and Vivek Srikumar71

Putting Feelings into Words: Cross-Linguistic Markers of the Referential Process
Sean Murphy, Bernard Maskit and Wilma Bucci 80

Towards Developing an Annotation Scheme for Depressive Disorder Symptoms: A Preliminary Study using Twitter Data
Danielle Mowery, Craig Bryan and Mike Conway 89

Beyond LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter
Philip Resnik, William Armstrong, Leonardo Claudino, Thang Nguyen, Viet-An Nguyen and Jordan Boyd-Graber 99

Automated morphological analysis of clinical language samples
Kyle Gorman, Steven Bedrick, Geza Kiss, Eric Morley, Rosemary Ingham, Metrah Mohammed, Katina Papadakis and Jan van Santen 108

Similarity Measures for Quantifying Restrictive and Repetitive Behavior in Conversations of Autistic Children
Masoud Rouhizadeh, Richard Sproat and Jan van Santen 117

Practical issues in developing semantic frameworks for the analysis of verbal fluency data: A Norwegian data case study
Mark Rosenstein, Peter Foltz, Anja Vaskinn and Brita Elvevåg 124

A Computer Program for Tracking the Evolution of a Psychotherapy Treatment
Bernard Maskit, Wilma Bucci and Sean Murphy 134

Workshop Program

2015/06/05

09:00–09:15 *Opening Remarks*
Margaret Mitchell, Glen Coppersmith, Kristy Hollingshead

09:15–11:00 Oral Presentations, Session 1

From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses
Glen Coppersmith, Mark Dredze, Craig Harman and Kristy Hollingshead

Quantifying the Language of Schizophrenia in Social Media
Margaret Mitchell, Kristy Hollingshead and Glen Coppersmith

The role of personality, age, and gender in tweeting about mental illness
Daniel Preoţiuc-Pietro, Johannes Eichstaedt, Gregory Park, Maarten Sap, Laura Smith, Victoria Tobolsky, H. Andrew Schwartz and Lyle Ungar

11:00–11:15 Break

11:15–11:45 Shared Task

CLPsych 2015 Shared Task: Depression and PTSD on Twitter
Glen Coppersmith, Mark Dredze, Craig Harman, Kristy Hollingshead and Margaret Mitchell

Mental Illness Detection at the World Well-Being Project for the CLPsych 2015 Shared Task
Daniel Preoţiuc-Pietro, Maarten Sap, H. Andrew Schwartz and Lyle Ungar

Screening Twitter Users for Depression and PTSD with Lexical Decision Lists
Ted Pedersen

The University of Maryland CLPsych 2015 Shared Task System
Philip Resnik, William Armstrong, Leonardo Claudino and Thang Nguyen

11:35–11:45 *Discussion*
Philip Resnik

2015/06/05 (continued)

11:45–12:45 Poster Presentations

Computational cognitive modeling of inflectional verb morphology in Spanish-speakers for the characterization and diagnosis of Alzheimer's Disease
M. Dolores del Castillo, J. Ignacio Serrano and Jesús Oliva

Recursive Neural Networks for Coding Therapist and Patient Behavior in Motivational Interviewing
Michael Tanana, Kevin Hallgren, Zac Imel, David Atkins, Padhraic Smyth and Vivek Srikumar

Putting Feelings into Words: Cross-Linguistic Markers of the Referential Process
Sean Murphy, Bernard Maskit and Wilma Bucci

Towards Developing an Annotation Scheme for Depressive Disorder Symptoms: A Preliminary Study using Twitter Data
Danielle Mowery, Craig Bryan and Mike Conway

Beyond LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter
Philip Resnik, William Armstrong, Leonardo Claudino, Thang Nguyen, Viet-An Nguyen and Jordan Boyd-Graber

Automated morphological analysis of clinical language samples
Kyle Gorman, Steven Bedrick, Geza Kiss, Eric Morley, Rosemary Ingham, Metrah Mohammed, Katina Papadakis and Jan van Santen

12:45–14:00 Lunch

14:00–14:45 Invited Talk

14:00–14:45 *Invited Talk*
Munmun De Choudhury

2015/06/05 (continued)

14:45–15:00 Break

15:00–16:45 Oral Presentations, Session 2

Similarity Measures for Quantifying Restrictive and Repetitive Behavior in Conversations of Autistic Children
Masoud Rouhizadeh, Richard Sproat and Jan van Santen

Practical issues in developing semantic frameworks for the analysis of verbal fluency data: A Norwegian data case study
Mark Rosenstein, Peter Foltz, Anja Vaskinn and Brita Elvevåg

A Computer Program for Tracking the Evolution of a Psychotherapy Treatment
Bernard Maskit, Wilma Bucci and Sean Murphy

16:45–17:00 Closing Remarks