Introduction

In the United States, mental and neurological health problems are among the costliest challenges we face. Depression, Alzheimer’s disease, bipolar disorder, and attention deficit hyperactivity disorder (ADHD) are only a handful of the many illnesses that contribute to this cost. The global cost of mental health conditions alone was estimated at $2.5 trillion in 2010, with a projected increase to over $6 trillion in 2030. Neurological illnesses and mental disorders cost the U.S. more than $760 billion a year. The World Health Organization (WHO) estimates one out of four people worldwide will suffer from a mental illness at some point in their lives, while one in five Americans experience a mental health problem in any given year. Mental, neurological, and substance use disorders are the leading cause of disability worldwide, yet most public service announcements and government education programs remain focused on physical health issues such as cancer and obesity. Despite the substantial and rising burden of such disorders, there is a significant shortage of resources available to prevent, diagnose, and treat them; thus technology must be brought to bear.

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 the domain of mental and neurological health could lead to inexpensive screening measures that may be administered by a wider array of healthcare professionals. Researchers had begun targeting such issues prior to this workshop series, using language technology to identify emotion in suicide notes, analyze the language of those with autistic spectrum disorders, and aid the diagnosis of dementia.

The series of Computational Linguistics and Clinical Psychology (CLPsych) workshops began at ACL 2014, while NAACL 2015 hosted the second such workshop with a near-doubling in attendance. The 2015 workshop additionally hosted a Shared Task for detecting depression and post-traumatic stress disorder (PTSD) based on social media posts. The CLPsych workshops diverge from the conventional “mini-conference” workshop format by inviting clinical psychologists and researchers to join us at the workshop as discussants, to provide real-world points of view on the potential applications of NLP technologies presented during the workshop. We hope to build the momentum towards releasing tools and data that can be used by clinical psychologists.

NAACL 2016 hosts the third CLPsych workshop, with another shared task. Published papers in this proceedings propose methods for aiding the diagnosis of dementia, analyzing sentiment as related to psychotherapy, assessing suicide risk, and quantifying the language of mental health. The 2016 CLPsych Shared Task centered on the classification of posts from a mental health forum to assist forum moderators in triaging and escalating posts requiring immediate attention. We accepted 11 submissions for the main workshop and 16 for the shared task. Each oral presentation will be followed by discussions led by one of our discussants, subject matter experts working in the fields of behavioral and mental health and with clinical data, including: Dr. Loring Ingraham and Dr. Bart Andrews.

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 very generous sponsorship.

– Kristy and Lyle
Organizers:
  Kristy Hollingshead, IHMC
  Lyle Ungar, University of Pennsylvania

Clinical Discussants:
  Loring J. Ingraham, George Washington University
  Bart Andrews, Behavioral Health Response

Program Committee:
  Steven Bedrick, Oregon Health & Science University
  Archna Bhatia, IHMC
  Wilma Bucci, Adelphi University
  Wei Chen, Nationwide Children’s Hospital
  Leonardo Claudino, University of Maryland, College Park
  Mike Conway, University of Utah
  Glen Coppersmith, Qntfy
  Brita Elvevåg, Department of Clinical Medicine, University of Tromsø
  Peter Foltz, Pearson
  Dan Goldwasser, Purdue University
  Ben Hachey, University of Sydney
  Graeme Hirst, University of Toronto
  Christopher Homan, Rochester Institute of Technology
  Jena Hwang, IHMC
  Zac Imel, University of Utah
  Loring Ingraham, George Washington University
  William Jarrold, Nuance Communications
  Yangfeng Ji, School of Interactive Computing, Georgia Institute of Technology
  Dimitrios Kokkinakis, University of Gothenburg
  Tong Liu, Rochester Institute of Technology
  Shervin Malmasi, Harvard Medical School
  Bernard Maskit, Stony Brook University
  Margaret Mitchell, Microsoft Research
  Eric Morley, Oregon Health & Science University
  Danielle Mowery, University of Utah
  Sean Murphy, John Jay College of Criminal Justice; City University of New York
  Cecilia Oveddoter Alm, Rochester Institute of Technology
  Ted Pedersen, University of Minnesota
  Craig Pfeifer, MITRE
  Glen Pink, University of Sydney
  Daniel Preotiuc, University of Pennsylvania
  Emily Prud’hommeaux, Rochester Institute of Technology
## Table of Contents

*Detecting late-life depression in Alzheimer’s disease through analysis of speech and language*
  Kathleen C. Fraser, Frank Rudzicz and Graeme Hirst .................................................. 1

*Towards Early Dementia Detection: Fusing Linguistic and Non-Linguistic Clinical Data*
  Joseph Bullard, Cecilia Ovesdotter Alm, Xumin Liu, Qi Yu and Rubén Proaño ................. 12

*Self-Reflective Sentiment Analysis*
  Benjamin Shickel, Martin Heesacker, Sherry Benton, Ashkan Ebadi, Paul Nickerson and Parisa Rashidi ......................................................................................................................... 23

*Is Sentiment in Movies the Same as Sentiment in Psychotherapy? Comparisons Using a New Psychotherapy Sentiment Database*
  Michael Tanana, Aaron Dembe, Christina S. Soma, Zac Imel, David Atkins and Vivek Srikumar 33

*Building a Motivational Interviewing Dataset*
  Verónica Pérez-Rosas, Rada Mihalcea, Kenneth Resnicow, Satinder Singh and Lawrence An . 42

*Crazy Mad Nutters: The Language of Mental Health*
  Jena D. Hwang and Kristy Hollingshead .............................................................. 52

*The language of mental health problems in social media*
  George Gkotsis, Anika Oellrich, Tim Hubbard, Richard Dobson, Maria Liakata, Sumithra Velupillai and Rina Dutta ................................................................. 63

*Exploring Autism Spectrum Disorders Using HLT*
  Julia Parish-Morris, Mark Liberman, Neville Ryant, Christopher Cieri, Leila Bateman, Emily Ferguson and Robert Schultz ................................................................. 74

*Generating Clinically Relevant Texts: A Case Study on Life-Changing Events*
  Mayuresh Oak, Anil Behera, Titus Thomas, Cecilia Ovesdotter Alm, Emily Prud’hommeaux, Christopher Homan and Raymond Ptucha ................................................ 85

*Don’t Let Notes Be Misunderstood: A Negation Detection Method for Assessing Risk of Suicide in Mental Health Records*
  George Gkotsis, Sumithra Velupillai, Anika Oellrich, Harry Dean, Maria Liakata and Rina Dutta 95

*Exploratory Analysis of Social Media Prior to a Suicide Attempt*
  Glen Coppersmith, Kim Ngo, Ryan Leary and Anthony Wood ..................................... 106

*CLPsych 2016 Shared Task: Triaging content in online peer-support forums*
  David N. Milne, Glen Pink, Ben Hachey and Rafael A. Calvo ........................................ 118

*Data61-CSIRO systems at the CLPsych 2016 Shared Task*
  Sunghwan Mac Kim, Yufei Wang, Stephen Wan and Cecile Paris .................................. 128
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicting Post Severity in Mental Health Forums</td>
<td>Shervin Malmasi, Marcos Zampieri and Mark Dras</td>
<td>133</td>
</tr>
<tr>
<td>Classifying ReachOut posts with a radial basis function SVM</td>
<td>Chris Brew</td>
<td>138</td>
</tr>
<tr>
<td>Triaging Mental Health Forum Posts</td>
<td>Arman Cohan, Sydney Young and Nazli Goharian</td>
<td>143</td>
</tr>
<tr>
<td>Mental Distress Detection and Triage in Forum Posts: The LT3 CLPsych 2016 Shared Task System</td>
<td>Bart Desmet, Gilles Jacobs and Véronique Hoste</td>
<td>148</td>
</tr>
<tr>
<td>Text Analysis and Automatic Triage of Posts in a Mental Health Forum</td>
<td>Ehsaneddin Asgari, Soroush Nasiriany and Mohammad R.K. Mofrad</td>
<td>153</td>
</tr>
<tr>
<td>The UMD CLPsych 2016 Shared Task System: Text Representation for Predicting Triage of Forum Posts about Mental Health</td>
<td>Meir Friedenberg, Hadi Amiri, Hal Daumé III and Philip Resnik</td>
<td>158</td>
</tr>
<tr>
<td>Using Linear Classifiers for the Automatic Triage of Posts in the 2016 CLPsych Shared Task</td>
<td>Juri Opitz</td>
<td>162</td>
</tr>
<tr>
<td>The GW/UMD CLPsych 2016 Shared Task System</td>
<td>Ayah Zirikly, Varun Kumar and Philip Resnik</td>
<td>166</td>
</tr>
<tr>
<td>Semi-supervised CLPsych 2016 Shared Task System Submission</td>
<td>Nicolas Rey-Villamizar, Prasha Shrestha, Thamar Solorio, Farig Sadeque, Steven Bethard and Ted Pedersen</td>
<td>171</td>
</tr>
<tr>
<td>Combining Multiple Classifiers Using Global Ranking for ReachOut.com Post Triage</td>
<td>Chen-Kai Wang, Hong-Jie Dai, Chih-Wei Chen, Jitendra Jonnagaddala and Nai-Wen Chang</td>
<td>176</td>
</tr>
<tr>
<td>Classification of mental health forum posts</td>
<td>Glen Pink, Will Radford and Ben Hachey</td>
<td>180</td>
</tr>
<tr>
<td>Automatic Triage of Mental Health Online Forum Posts: CLPsych 2016 System Description</td>
<td>Hayda Almeida, Marc Queudot and Marie-Jean Meurs</td>
<td>183</td>
</tr>
<tr>
<td>Automatic Triage of Mental Health Forum Posts</td>
<td>Benjamin Shickel and Parisa Rashidi</td>
<td>188</td>
</tr>
<tr>
<td>Text-based experiments for Predicting mental health emergencies in online web forum posts</td>
<td>Hector-Hugo Franco-Penya and Liliana Mamani Sanchez</td>
<td>193</td>
</tr>
</tbody>
</table>
Conference Program

2016/06/16

09:00–09:20  Opening Remarks
Kristy Hollingshead and Lyle Ungar

09:20–10:30  Oral Presentations, Session 1

Detecting late-life depression in Alzheimer’s disease through analysis of speech and language
Kathleen C. Fraser, Frank Rudzicz and Graeme Hirst

Towards Early Dementia Detection: Fusing Linguistic and Non-Linguistic Clinical Data
Joseph Bullard, Cecilia Ovesdotter Alm, Xumin Liu, Qi Yu and Rubén Proaño

10:30–11:00  Break

11:00–11:45  Poster Presentations

Self-Reflective Sentiment Analysis
Benjamin Shickel, Martin Heesacker, Sherry Benton, Ashkan Ebadi, Paul Nickerson and Parisa Rashidi

Is Sentiment in Movies the Same as Sentiment in Psychotherapy? Comparisons Using a New Psychotherapy Sentiment Database
Michael Tanana, Aaron Dembe, Christina S. Soma, Zac Imel, David Atkins and Vivek Srikumar

Building a Motivational Interviewing Dataset
Verónica Pérez-Rosas, Rada Mihalcea, Kenneth Resnicow, Satinder Singh and Lawrence An

Crazy Mad Nutters: The Language of Mental Health
Jena D. Hwang and Kristy Hollingshead

The language of mental health problems in social media
George Gkotsis, Anika Oellrich, Tim Hubbard, Richard Dobson, Maria Liakata, Sumithra Velupillai and Rina Dutta

Exploring Autism Spectrum Disorders Using HLT
Julia Parish-Morris, Mark Liberman, Neville Ryant, Christopher Cieri, Leila Bate-man, Emily Ferguson and Robert Schultz
11:45–1:00 Lunch

1:00–2:45 Oral Presentations, Session 2

*Generating Clinically Relevant Texts: A Case Study on Life-Changing Events*
Mayuresh Oak, Anil Behera, Titus Thomas, Cecilia Ovesdotter Alm, Emily Prud’hommeaux, Christopher Homan and Raymond Ptucha

*Don’t Let Notes Be Misunderstood: A Negation Detection Method for Assessing Risk of Suicide in Mental Health Records*
George Gkotsis, Sumithra Velupillai, Anika Oellrich, Harry Dean, Maria Liakata and Rina Dutta

*Exploratory Analysis of Social Media Prior to a Suicide Attempt*
Glen Coppersmith, Kim Ngo, Ryan Leary and Anthony Wood

2:45–3:00 Break

3:00–3:15 Shared Task Introduction

*CLPsych 2016 Shared Task: Triaging content in online peer-support forums*
David N. Milne, Glen Pink, Ben Hachey and Rafael A. Calvo

3:15–3:40 Shared Task Poster Presentations, Session 1

*Data61-CSIRO systems at the CLPsych 2016 Shared Task*
Sunghwan Mac Kim, Yufei Wang, Stephen Wan and Cecile Paris

*Predicting Post Severity in Mental Health Forums*
Shervin Malmasi, Marcos Zampieri and Mark Dras

*Classifying ReachOut posts with a radial basis function SVM*
Chris Brew

*Triaging Mental Health Forum Posts*
Arman Cohan, Sydney Young and Nazli Goharian
Mental Distress Detection and Triage in Forum Posts: The LT3 CLPsych 2016 Shared Task System
Bart Desmet, Gilles Jacobs and Véronique Hoste

Text Analysis and Automatic Triage of Posts in a Mental Health Forum
Ehsaneddin Asgari, Soroush Nasiriany and Mohammad R.K. Mofrad

The UMD CLPsych 2016 Shared Task System: Text Representation for Predicting Triage of Forum Posts about Mental Health
Meir Friedenberg, Hadi Amiri, Hal Daumé III and Philip Resnik

3:40–4:00 Break

4:00–4:25 Shared Task Poster Presentations, Session 2

Using Linear Classifiers for the Automatic Triage of Posts in the 2016 CLPsych Shared Task
Juri Opitz

The GW/UMD CLPsych 2016 Shared Task System
Ayah Zirikly, Varun Kumar and Philip Resnik

Semi-supervised CLPsych 2016 Shared Task System Submission
Nicolas Rey-Villamizar, Prasha Shrestha, Thamar Solorio, Farig Sadeque, Steven Bethard and Ted Pedersen

Combining Multiple Classifiers Using Global Ranking for ReachOut.com Post Triage
Chen-Kai Wang, Hong-Jie Dai, Chih-Wei Chen, Jitendra Jonnagaddala and Nai-Wen Chang

Classification of mental health forum posts
Glen Pink, Will Radford and Ben Hachey

Automatic Triage of Mental Health Online Forum Posts: CLPsych 2016 System Description
Hayda Almeida, Marc Queudot and Marie-Jean Meurs

Automatic Triage of Mental Health Forum Posts
Benjamin Shickel and Parisa Rashidi

Text-based experiments for Predicting mental health emergencies in online web forum posts
Hector-Hugo Franco-Penya and Liliana Mamani Sanchez
4:25–4:45  Closing Remarks