8-2016

Analyzing Clinical Depressive Symptoms in Twitter

Amir Hossein Yazdavar
Wright State University - Main Campus, yazdavar.2@wright.edu

Hussein S. Al-Olimat
Wright State University - Main Campus

Tanvi Banerjee
Wright State University - Main Campus, tanvi.banerjee@wright.edu

Krishnaprasad Thirunarayan
Wright State University - Main Campus, t.k.prasad@wright.edu

Amit P. Sheth
Wright State University - Main Campus, amit@sc.edu

Analyzing Clinical Depressive Symptoms in Twitter

[1]Amir Hossein Yazdavar, [1]Hussein S. Al-Olimat, [1]Tanvi Banerjee, [1]Krishnaprasad Thirunarayan, [2]Jyotishman Pathak, [1]Amit Sheth
[1]Ohio Center of Excellence in Knowledge-enabled Computing (Kno.e.sis), Wright State University, Dayton OH, USA
[2] Division of Health Informatics, Weill Cornell Medical College, New York, NY, USA
[1] {amir,tanvi,hussein,tkprasad,amit}@knoesis.org, [2] jyp2001@med.cornell.edu

Motivation

- **350 million** people are suffering from clinical depression **worldwide**.
- **27 million Americans** are diagnosed with clinical depression that is responsible for more than **30,000 suicides** each year.
- Over **90%** of people who commit suicide have been diagnosed with clinical depression or another diagnosable mental illness.
- According to the World Mental Health Survey conducted in **17** countries, about **5%** of people reported having an episode of depression.
- Depression remains undiagnosed, untreated or under-treated phenomenon due to various reasons such as the *denial of illness* or the *social stigma* associated with it.
- *Early recognition* of depression symptoms and their treatment through timely intervention can **prevent** the *onset of major depression*.
- A common global effort to manage depression involves detecting depression through *survey-based* methods via phone or online questionnaires . However, these studies suffer from *under-representation*, *sampling biases* and *incomplete* information. Additionally, *large temporal gaps* between data collection and dissemination of findings can delay administration of timely and appropriate remedial measures.

Research Aims

(Aim-1) *Study* and identify clinical depressive disorders using explicit and implicit expression of depression on social streams.
(Aim-2) *Build* a reliable platform to automatically detect depressive behavior in social media that emulates and extends the functionality of PHQ-9 to monitor user depressive behaviors.
(Aim-3) *Evaluate* our approach on self-reported profiles on social media.

Related Works

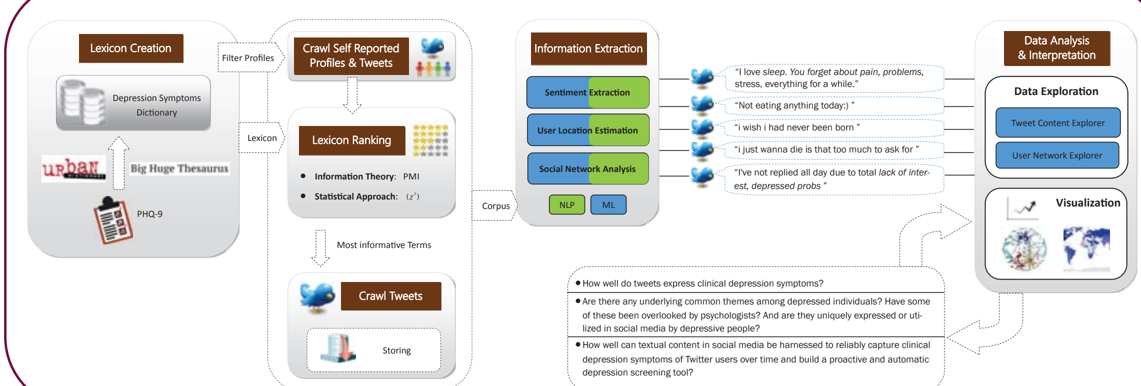
Recently, there has been a wealth of research on studying depression in several disciplines such as psychology, psychiatry, medicine, and sociolinguistics. Linguistic analysis of speech can be used to distinguish depressed subjects from normal people. Park et al. [8] shows that depressed users use Twitter for emotional interaction and social awareness while the non-depressed ones use it for information consumption and sharing. Several efforts have attempted to automatically detect depression in social media content using machine learning.

Nguyen et al. [9] leverage affective aspect, linguistic style and topics as features for detecting depressed communities. De Choudhury et al. [10] characterize depression based on factors such as language, emotion, style, ego-network, and user engagement. They utilize these distinguishing characteristics to build a classifier to predict the likelihood of depression in a post or in an individual.

Approach

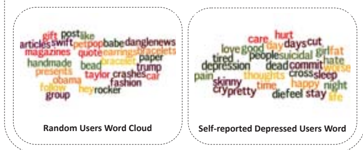
- Twitter provides a rich source for studying *people's mood* in order to detect *depressive behaviors*.
- We developed a novel technique to unobtrusively analyzes individual posts in social media to detect signs of depression that can be utilized to build a proactive and automatic screening tool for *early recognition* of clinical depression.
- Leveraging clinical definition of depression, we build a depression lexicon that contains common depression symptoms determined by experts such as from the established clinical assessment questionnaires PHQ-9.
- We expanded the terms expressing the nine PHQ-9 depression symptoms categories using Urban Dictionary and Big Huge Thesaurus.
- The lexicon contains depression-related symptoms that are likely to appear in the tweets of individuals either having depressive-like symptoms or suffering from depression.
- A subset of highly informative seed terms are selected from this depression lexicon for crawling depression-related tweets. For each lexical term, we calculate its association with all of the variations of the term "depress" using Pointwise Mutual Information (PMI) and Chi-squared (χ^2) test to quantify their correlation and thereby rank order them.
- We leverage Twitris, our social media analysis platform, to study language, sentiment, emotions, topics and people content-network of depressed individuals.

Architecture



Analysis

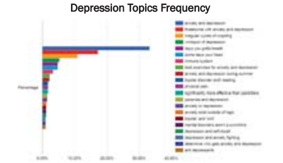
Twitter Users Word Usage



Random Users Word Cloud

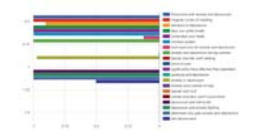
Self-reported Depressed Users Word

Depression Topics Frequency

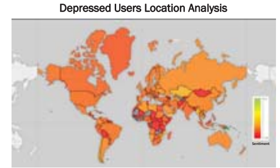


Emotion of Depressive Indicator Tweets

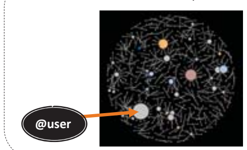
Sentiment associated with Topics



Depressed Users Location Analysis



Follower Network of Depressed Users



Conclusion

- We demonstrated the potential of social media data for extracting clinical depression symptoms in individuals that can be leveraged to improve the current questionnaire driven diagnostic tools in its ability to glean clinical depression symptoms in a natural setting and in a continuous and unobtrusive manner.
- Our results reveals that there are significant differences in the topic preferences and word usage patterns of self-declared depressed group from random users in our dataset.
- This analysis framework could benefit future research on building a warning system which can predict the onset of major depression.
- Current clinical approaches is mainly focused on individuals; our findings can be utilized to automatically assess public health at community level for determining health risks behavior, suicidal tendency and help-seeking behavior.

Acknowledgement

We thank our colleagues at Kno.e.sis, Surendra B. Marupudi and Ankita Saxena, and our collaborator at Weill Cornell Medical College for providing information for creating lexicon of symptoms. This study was supported in part by funding from NIH R01MH105384.

NIH National Institutes of Health

References

[1] Parvathy Mini, P. (2014). *Assessment of Factors Influencing the Choice and Use of Temporary Contraceptive Methods for Birth Spacing among Married Women in Rural Trivandrum (Doctoral dissertation, SCTIMST).*
[2] Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The Phq-9. *Journal of general internal medicine, 16(9), 606-613.*
[3] First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1997). *User's guide for the Structured clinical interview for DSM-IV axis I disorders SCID-I: clin ver. American Psychiatric Pub.*
[4] Dictionary, U. (2011). Urban dictionary. Web.
[5] Watson, J. (2014). Big huge thesaurus.
[6] Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. *Proceedings of GSCL, 31-40*
[7] Amit Sheth, Ashutosh Jadhav, Pavan Kapanipathi, Chen Lu, Hemant Purohit, Gary Alan Smith, Wenbo Wang, 'Twitris- a System for Collective Social Intelligence', Encyclopedia of Social Network Analysis and Mining (ESNAM), 2014.
[8] Park, Minsu, David W. McDonald, and Meeyoung Cha. "Perception Differences between the Depressed and Non-Depressed Users in Twitter." *ICWSM. 2013*
[9]Nguyen, Thin, et al. "Affective and content analysis of online depression communities." *IEEE Transactions on Affective Computing 5.3 (2014): 217-226.*
[10] M. De Choudhury, S. Counts, and E. Horvitz. Social media as a measurement tool of depression in populations.