Identifying Depressive Disorder in the Twitter Population

Goonmeet Bajaj  
*Wright State University - Main Campus*, bajaj.3@wright.edu

Amir Hossein Yazdavar

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

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

Follow this and additional works at: [https://corescholar.libraries.wright.edu/knoesis](https://corescholar.libraries.wright.edu/knoesis)

Part of the Bioinformatics Commons, Communication Technology and New Media Commons, Databases and Information Systems Commons, OS and Networks Commons, and the Science and Technology Studies Commons

Repository Citation

https://corescholar.libraries.wright.edu/knoesis/1140

This Article is brought to you for free and open access by the The Ohio Center of Excellence in Knowledge-Enabled Computing (Kno.e.sis) at CORE Scholar. It has been accepted for inclusion in Kno.e.sis Publications by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.
Identifying Depressive Disorder in the Twitter Population

Goonmeet Bajaj, Amir Hossein Yazdavar, T.K. Prasad, Amit Sheth
Kno.e.sis Center, Wright State University, Dayton, OH

Overview
Depression is a highly prevalent public health challenge and a major cause of disability across the globe.

- Annually 6.7% of Americans (that is, more than 16 million)
- Traditional approaches to curb depression involve survey-based methods via phone or online questionnaires
- Large temporal gaps and cognitive bias

Social media provides a method for learning users’ feelings, emotions, behaviors, and decisions in real-time.

1) How well do tweets express depressive behavior and can they be detected automatically?
2) How well does geographical information serve as the basis for effective community-level management of depression and location of mental health services?

Dataset

<table>
<thead>
<tr>
<th># of Self Reported Users</th>
<th># of users in the Gold Standard dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>33,719</td>
<td>16,194</td>
</tr>
</tbody>
</table>

Profile classification

- # of tweets
- # of followers
- # of friends
- Tweet content
- Levenshtein distance between screen name and lexicon of depressive symptoms
- Text from image
- Emotion from images
- Tweeting time
- Ego-network

Feature Engineering

Prediction model:
- Automatically detect depressed users leveraging a multimodal feature set:
  - # of tweets
  - # of followers
  - # of friends
  - Tweet content
  - Levenshtein distance between screen name and lexicon of depressive symptoms
  - Text from image
  - Emotion from images
  - Tweeting time
  - Ego-network

System Architecture

Geographical Analysis

Map of hospital inpatient and outpatient mental health centers from the Substance Abuse and Mental Health Services Administration (SAMHSA)

Map of user profiles collected by our platform

Map Key:
- Red pointer = users with geo-enabled tweets
- Teal pointer = users with places or locations in tweets or profile
- Green pointer = location determined by Pigeon, A Python Geotagging Tool

Project Funded by:
NIH/NIMH, R21MH099828

References: