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.sheth@wright.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](https://corescholar.libraries.wright.edu/knoesis/1140)
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
# of Self Reported Users: 33,719
# of users in the Gold Standard dataset*: 16,194
# of depressed profiles: 4823
# of non-depressed profiles: 7739

*Users judged depressed by human annotators

Self Reported Users
- Profiles collected using the Social-media Depression Detector (SDD)
- SDD is based on PHQ-9—a multipurpose instrument for screening, diagnosing, monitoring, and measuring the severity of depression.

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

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

Brainstorm
Please tell me your interest and your thoughts about this study!
- What are possible features that used to classifying profiles?

Email me @ goonmeet.bajaj@gmail.com
Project website: rebread.ly/depressionProject