USING SOCIAL MEDIA TO UNDERSTAND MENTAL HEALTH AT THE POPULATION LEVEL

MIKE CONWAY

mike.conway@utah.edu

DECART DATA SCIENCE SUMMER SCHOOL 2018
UNIVERSITY OF UTAH
ABOUT ME…

• Currently an assistant professor in the Department of Biomedical Informatics at the University of Utah

RESEARCH INTERESTS…

• NLP for disease surveillance, focusing on mental health
• EHR-oriented NLP
• Ethical Issues in Biomedical Informatics

APPLYING NLP TO CRITICAL PROBLEMS IN CLINICAL & PUBLIC HEALTH INFORMATICS
TALK OUTLINE

• Introduction:
  • Depression as a public health problem
  • Public social media as a resource for public mental health

• Twitter, NLP & Depression:
  • Developing an annotated Twitter corpus of depression symptoms
  • Developing NLP algorithms to automatically identify depression symptoms in Twitter

• Ethical Issues:
  • Focus group work
PART ONE: PUBLIC MENTAL HEALTH & SOCIAL MEDIA
WHAT IS POPULATION HEALTH?

“Understanding health outcomes, patterns of health determinants, policies and interventions as they relate to a group of individuals”

Kindig (2003)

PUBLIC HEALTH SURVEILLANCE

“The ongoing, systematic collection, analysis, and interpretation of health-related data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know. The final link in the surveillance chain is the application of these data to prevention and control”

CDC Definition
DSM-5 CRITERIA - MAJOR DEPRESSIVE DISORDER

Five or more of the following symptoms during a two week period:

1. Depressed mood
2. Anhedonia
3. Weight change
4. Changes in sleep
5. Psychomotor agitation or retardation
6. Fatigue or loss of energy
7. Feelings of worthlessness/inappropriate guilt
8. Diminished concentration or indecisiveness
9. Recurrent thoughts of death or suicidal ideation
MAJOR DEPRESSIVE DISORDER

• MDD has a lifetime prevalence of 16% (12 month prevalence of 6.6%)
• Fifth biggest contributor to the US burden of disease (after lung cancer, heart disease, back pain and COPD)
• Global economic cost of mental illness: US$2.5 trillion (estimated to rise to US$5 trillion by 2030)
• Suicide and intentional self harm:
  • 42,733 suicide deaths per year in the US
  • Almost 500,000 people with self-inflicted injuries treated in the US ER per year
WHY DO PUBLIC HEALTH SURVEILLANCE?

• Estimate the magnitude of a health problem in the population
• Understand the natural history of a disease or injury
• Detect outbreaks/epidemics
• Generate hypotheses regarding aetiology
• Evaluate control strategies
• Detect changes in health practice
• Assess health care quality
• Assess the safety of drugs, devices, diagnostics, or procedures
EXISTING MENTAL HEALTH SURVEILLANCE SYSTEMS

• Telephone-based surveys
  • Fewer people have landlines
  • Conducted infrequently (e.g. BRFSS is conducted once per year)
  • Expensive - requires a team of phone operators
• Interview/Examination based systems
  • Limited number of participants
  • Expensive - personal interviews/tests

PUBLIC SOCIAL MEDIA IN CONJUNCTION WITH NLP MAY PROVIDE A COST EFFECTIVE AND FLEXIBLE APPROACH TO AUGMENTING CURRENT SURVEY BASED RISK FACTOR MONITORING METHODS
PART TWO: EXTRACTING MENTAL HEALTH SIGNALS FROM TWITTER

Dr Albert Park, University of Utah
Dr Danielle Mowery, University of Utah
Dr Nicholas Perry, University of Utah
Dr Craig Bryan, University of Utah
Mr Gregg Stoddard, University of Utah
Mr Tyler Cheney, University of Utah
Ms Hilary Smith, University of Utah
Dr Glenn Coppersmith, Johns Hopkins University/Qntfy
PROJECT GOALS

Overarching goal: Develop NLP classifiers to automatically identify depression symptoms and psychosocial stressors from Twitter data.

- Develop an annotation scheme for depression symptoms in Twitter
- Build a corpus based on this annotation scheme
- Develop classifiers to automatically tag tweets that are indicative of depression
- Apply classifiers to national data
COMPREHENSIVE ANNOTATION SCHEME
*Coloured symptoms and stressors are most prevalent

TOP LEVEL ANNOTATION SCHEME
LIWC TERMS

- Associated with symptoms and stressors
- e.g., “die” → Recurrent thoughts of death/suicidal ideation
- Maintained natural distribution by random sampling

Most frequent: N= 120 unique terms

- die
- cry
- kill
- sad
- hurt
- tired

<table>
<thead>
<tr>
<th>Depression Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms</td>
</tr>
<tr>
<td>Depressed mood</td>
</tr>
<tr>
<td>Anhedonia</td>
</tr>
<tr>
<td>Weight change or change in appetite</td>
</tr>
<tr>
<td>Disturbed sleep</td>
</tr>
<tr>
<td>Psychomotor agitation or retardation</td>
</tr>
<tr>
<td>Fatigue or loss of energy</td>
</tr>
<tr>
<td>Feelings of worthlessness or excessive inappropriate guilt</td>
</tr>
<tr>
<td>Diminished ability to think or concentrate, indecisiveness</td>
</tr>
<tr>
<td>Recurrent thoughts of death, suicidal ideation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problems with expected life course with respect to self</td>
</tr>
<tr>
<td>Problems with primary support group</td>
</tr>
<tr>
<td>Problems related to the social environment</td>
</tr>
<tr>
<td>Educational problems</td>
</tr>
<tr>
<td>Occupational problems</td>
</tr>
<tr>
<td>Housing problems</td>
</tr>
<tr>
<td>Economic problems</td>
</tr>
<tr>
<td>Problems with access to healthcare</td>
</tr>
<tr>
<td>Problems related to the legal system/crime</td>
</tr>
<tr>
<td>Other psychosocial and environmental problems</td>
</tr>
<tr>
<td>Weather</td>
</tr>
<tr>
<td>Media</td>
</tr>
</tbody>
</table>

LIWC “sad” keyword

- abandon*, ache*, aching, agoni*,
- alone, broke*, cried, cries, crushed,
- cry, damag*, defeat*, depress*,
- depriv*, despair*, devastat*,
- disadvantage*, disappoint*,
- discourag*, dishearten*, disillusion*,
- dissatisf*, doom*, dull*.
# RELIABILITY STUDY

**N=300 tweets**

<table>
<thead>
<tr>
<th>Depression Categories</th>
<th>A1/A2</th>
<th>A2/A3</th>
<th>A1/A3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall</strong></td>
<td>81%</td>
<td>78%</td>
<td>76%</td>
</tr>
<tr>
<td><strong>No evidence of depression</strong></td>
<td>89%</td>
<td>86%</td>
<td>87%</td>
</tr>
<tr>
<td><strong>Symptoms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed mood</td>
<td>38%</td>
<td>60%</td>
<td>48%</td>
</tr>
<tr>
<td>Anhedonia</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Weight change or change in appetite</td>
<td>–</td>
<td>0%</td>
<td>–</td>
</tr>
<tr>
<td>Disturbed sleep</td>
<td>100%</td>
<td>50%</td>
<td>100%</td>
</tr>
<tr>
<td>Psychomotor agitation or retardation</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Fatigue or loss of energy</td>
<td>74%</td>
<td>78%</td>
<td>94%</td>
</tr>
<tr>
<td>Feelings of worthlessness or excessive inappropriate guilt</td>
<td>0%</td>
<td>29%</td>
<td>68%</td>
</tr>
<tr>
<td>Diminished ability to think or concentrate, indecisiveness</td>
<td>100%</td>
<td>–</td>
<td>0%</td>
</tr>
<tr>
<td>Recurrent thoughts of death, suicidal ideation</td>
<td>100%</td>
<td>100%</td>
<td>75%</td>
</tr>
<tr>
<td><strong>Stressors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems with expected life course with respect to self</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Problems with primary support group</td>
<td>0%</td>
<td>40%</td>
<td>36%</td>
</tr>
<tr>
<td>Problems related to the social environment</td>
<td>23%</td>
<td>42%</td>
<td>58%</td>
</tr>
<tr>
<td>Educational problems</td>
<td>–</td>
<td>50%</td>
<td>–</td>
</tr>
<tr>
<td>Occupational problems</td>
<td>–</td>
<td>0%</td>
<td>–</td>
</tr>
<tr>
<td>Housing problems</td>
<td>–</td>
<td>0%</td>
<td>–</td>
</tr>
<tr>
<td>Economic problems</td>
<td>–</td>
<td>67%</td>
<td>50%</td>
</tr>
<tr>
<td>Problems with access to healthcare</td>
<td>–</td>
<td>0%</td>
<td>–</td>
</tr>
<tr>
<td>Problems related to the legal system/crime</td>
<td>–</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Other psychosocial and environmental problems</td>
<td>–</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Weather</td>
<td>–</td>
<td>100%</td>
<td>–</td>
</tr>
<tr>
<td>Media</td>
<td>50.0%</td>
<td>0%</td>
<td>67%</td>
</tr>
</tbody>
</table>
CORRELATIONS BETWEEN SYMPTOMS & STRESSORS

- Media: 21
- Weather: 20
- Other psychosocial and environmental problems: 19
- Problems related to the legal system or crime: 18
- Problems with access to healthcare: 17
- Economic problems: 16
- Housing problems: 15
- Occupational problems: 14
- Educational problems: 13
- Problems related to the social environment: 12
- Problems with primary support group: 11
- Problems with expected life course with respect to self: 10
- Recurrent thoughts of death suicidal ideation: 9
- Diminished ability to think or concentrate indecisiveness: 8
- Feelings of worthlessness or excessive inappropriate guilt: 7
- Fatigue or loss of energy: 6
- Emotional withdrawal: 5
- Motor agitation or retardation: 5
- Disturbed sleep: 4
- Weight or appetite change: 3
- Anhedonia: 2
- Depressed mood: 1

Colors:
- > 0.50: greater than large effect
- 0.3-0.49: medium-large effect
- 0.1-0.29: small-medium effect
- < 0.09: less than small effect
CORRELATIONS

- Fatigue or loss of energy
  - disturbed sleep
  - education problems
- Depressed mood
  - feelings of worthlessness or guilt
- Education problems
  - diminished ability to think & concentrate
- Housing problems
  - economic problems

High correlations
[Pearson’s R >0.50]

“Exhausted and still can’t sleep”

“I feel so guilty and regret so much”

“I can’t concentrate on my homework”

“Never leave this broken house & still can’t pay my rent”
CLASSIFICATION STUDY

1. Discern whether or not a tweet contains no evidence of depression or evidence of depression.

2. Classify whether the tweet contained a subtype of depressed mood, disturbed sleep, fatigue or loss of energy, or none.
CLASSIFICATION STUDY

• Train classifiers using 9,473 annotations
• 5-fold cross validation utilising all features (17,000)
• Experiments:
  • Classifiers:
    • Linear Perceptron
    • Random Forests
    • Decision Trees
    • Naive Bayes
    • Support Vector Machines
  • Classifier analysis:
    • F1 scores for each classifier
    • Precision: LIWC vs most precise classifier
FEATURES USED

• Unigrams - e.g. “depressed”, “miserable”
• syntax - “cried” encoded as V for verb
• LIWC categories - e.g. negative affect, cognitive mechanism
• age - “this semester” as an indicator of 19-22 years of age
• emoticons - “:(“ encoded as SAD
• personality traits - e.g. “pissed off” implies neuroticism
• sentiment - “terrible” encoded as strongly negative and strongly subjective
WHICH CLASSIFIERS PERFORMED BEST?

Highest F scores achieved using Random Forests and Logistic Regression.
WHICH CLASSIFIERS PERFORMED BEST?

Support Vector Machines had the best precision for most classes.
Classifiers (often decision trees) are more often precise than LIWC terms.

Performance of best average precision classifier for depressive symptoms and for each subtype (a = no depth restriction, b = restriction depth of 5. c = L1 regularization)
ABLATION STUDY

- Train classifiers using 9,473 annotations
- 5-fold cross-validation ablating each feature group
- Classifiers:
  - Support Vector Machine
- Classifier analysis:
  - Baseline: all feature groups
  - F-score: change in F1 scores for each feature group ablated
WHICH FEATURE GROUPS WERE MOST INFORMATIVE?

Several feature groups are important for classifying disturbed sleep or fatigue or loss of energy.

Difference in F1 performance from baseline system produced by ablating each feature group.
CONCLUSIONS/FUTURE WORK

• Depressive symptom-related tweets can be classified with moderate to high accuracy
• In general, machine learning classifiers can produce higher precision than LIWC terms alone
• Simple lexical features are important for classifying tweets, although more complex features are also informative for some classes
• Reduced feature sets do not necessarily reduce performance
PART THREE: ETHICAL ISSUES IN UTILIZING SOCIAL MEDIA FOR MENTAL HEALTH SURVEILLANCE

Dr Dan O’Connor, Wellcome Trust
Dr Jude Mikal, University of Minnesota
Dr Samantha Hurst, University of California San Diego
BUT FIRST, SOME IMPORTANT DISTINCTIONS

• Observation vs Intervention
• Not all social media are the same:
  • password protected
  • terms and conditions differ
  • some social media forums (and topics) are inherently more sensitive than others (e.g. sexual abuse survivors)
ETHICAL CONSIDERATIONS - EYSENBACH (2001)

• Intrusiveness — Discuss to what degree the research is intrusive (e.g. passive analysis vs active involvement in community)
• Perceived privacy — Is it a closed group that requires registration?
• Vulnerability — e.g. fitness forum vs sexual abuse survivor forum
• Potential harm — Does the research have the potential to harm forum users?
• Informed consent — Is informed consent required, or can it be waived?
• Confidentiality — How can anonymity be protected?
• Intellectual property rights — Some participants may want publicity, not anonymity and require attribution
Public social media in combination with Natural Language Processing and Machine Learning has provided significant opportunities for public health, but also ethical challenges. IRBs typically grant exemptions for this work. Little guidance for researchers. Reviewed literature (PubMed, Psychinfo, Compendex) for papers that discussed social media mining and ethical concepts [2006-2014]. Goal of the paper was to map the terrain of ethical issues.
SYSTEMATIC REVIEW

PsycINFO
Retrieved: 51
Inclusion: 12
Not available: 7
Inclusion: 5

PubMed
Retrieved: 66
Inclusion: 49
Not available: 13
Inclusion: 36

Philosophers Index
Retrieved: 6
Inclusion: 0

Compendex
Retrieved: 219
Inclusion: 9
Not available: 1
Inclusion: 8

Screen title and abstract

Screen full text of 49 references

Final inclusion: 13
PROPOSED NORMATIVE RULES

- Avoid quoting directly from users when reporting research
- Informed consent should be gained from participants
- Metadata should not be disclosed
- Data collection should be logged and justified
- There should be parity between researcher and user
- Employment-related profiling for mental health conditions should not be routinely conducted
SOME COMMENTS...

- Few Twitter papers explicitly mention ethical issues
- Heterogenous views (e.g. requires consent, is [or is not] human subjects research, anything goes because the data is public)
- CS/Engineering - less familiar with regulatory oversight
- “It’s public data. There are no ethical issues in using it”
ETHICS FOCUS GROUPS

• We have examined the views of researchers, but what do social media users with depression think?
• Five 2-hour focus group interviews following a semi-structured protocol (26 participants)
• Recruited via notice boards, online discussion forums
• All participants were Twitter users (16 participants had a depression diagnosis)
SEMI-STRUCTURED PROTOCOL

• Introduction [name, age, occupation]
• General Twitter use (e.g. how often do you use Twitter? What are the things you are most likely to post on Twitter?)
• Privacy expectations (e.g. do you know who is able to monitor your Twitter use?)
• Privacy expectations and health (e.g. what do you think about using Twitter to monitor flu levels at the population level? Is mental health different?)
• Recommendations (how should Twitter data be used?)
• Other (e.g. are you the same person on Twitter as you are in everyday life?)
TWITTER & PRIVACY EXPECTATIONS

“I don’t pay to use Twitter. I sort of signed up with the expectation that it’s a free site and you just kind of throw things out publicly, [so] I don’t really have an expectation that anything I post is going to remain private”

Control Group, 29 year old male

“Exactly, like that’s what their product is. Their product is you. Because it’s free, you are the product”

Depression Group, 29 year old male

“I just acknowledged to myself a long time ago that whatever I put on the internet — whatever I put into my search engine, anything that I click on — is not private”

Depression Group, 21 year old female
USING TWITTER FOR PUBLIC HEALTH MONITORING

“I kind of think it’s cool when it’s stuff like the flu, because that that’s how they know to get the vaccines to a place”

Depression Group, 24 year old female

“It’s like fluoride in the water to me. They put fluoride in our water. We don’t really have a choice if we want to drink water, we’re going to get fluoride. But the benefits outweigh the risks”

Control Group, 26 year old female
USING TWITTER FOR POPULATION-LEVEL DEPRESSION MONITORING

“I’m OK as long as we can, you know, figure out ways to keep the data anonymous and completely, highly aggregated”
Depression Group, 24 year old female

“During my senior year, I would just tweet just because I wanted my friends to see it and to know that I didn’t feel good, or that i was upset or mad at someone. I think it would be very obvious, actually”
Control Group, 20 year old male

“It’s just the opposite for me. If I’m feeling down or anything, I just retreat back. There’d be a huge gap there”
Depression Group, 29 year old male
INTEGRATING TWITTER MENTAL HEALTH MONITORING WITH CLINICAL CARE

“I’m all for it. I know when I’ve gone to therapists or my doctor or whatever, I’m not the best at reporting how I’ve been doing when I’m actually at my appointment. That would be fantastic to have something else to support what I think, just because I’m not reliable about accurately assessing how I’m doing”

Depression Group, 24 year old female

“I think that sounds great! Especially, I think one common questions is ‘How long have you felt this way?’ I don’t know. But if you look at Twitter and just immediately generate a graph that shows mood swings over time. Absolutely!”

Control Group, 20 year old male
ACKNOWLEDGEMENTS

Dr Craig Bryan, University of Utah
Dr Annie Chen, University of Washington
Mr Tyler Cheney, California State University San Francisco
Dr Daniel O’Connor, Wellcome Trust
Dr Glen Coppersmith, Johns Hopkins University/Qntfy
Dr Samantha Hurst, University of California San Diego
Dr Jude Mikal, University of Minnesota
Dr Albert Park, University of Utah
Ms Hilary Smith, Brigham Young University
Mr Greg Stoddard, University of Utah
Dr Shu-Hong Zhu, University of California San Diego

Research reported in this presentation was supported by the NIH/National Library of Medicine under award numbers K99LM011393 & R00LM011393
THANKS!

mike.conway@utah.edu