The Role of Personality, Age and Gender in Tweeting about Mental Illnesses

Daniel Preoţiuc-Pietro, Johannes Eichstaedt, Gregory Park, Maarten Sap
Laura Smith, Victoria Tobolsky, H. Andrew Schwartz and Lyle Ungar
Problem

- Mental illnesses are underdiagnosed
Problem

- Mental illnesses are underdiagnosed

This Study:

- Explore the predictive power of demographic and personality based features.
- Find insights provided by each feature.
Data

- Twitter self-reports
  - ‘I have been diagnosed with depression’
  - depression: 483
  - PTSD: 370
  - controls: 1104

- each user has avg. 3400 messages

(Coppersmith et. al, CLPsych 2014)
Study Setup

Twitter -> language -> mental illness

classification
Study Setup

Twitter -> language -> age, gender, personality?

classification -> mental illness
Age, Gender

- Model from FB and Twitter data
  (Sap et. al, EMNLP 2014)
Age, Gender

- Model from FB and Twitter data
  (Sap et. al, EMNLP 2014)
Age, Gender

![Graph showing age and gender distribution with density plots for depressed, control, and ptsd groups.](image-url)
Personality

- Big 5 Personality Traits
  - openness
  - conscientiousness
  - extraversion
  - agreeableness
  - neuroticism

- Model from Facebook data
  (Park et. al 2014)
Personality

- **Big 5 Personality Traits**
  - openness
  - conscientiousness
  - extraversion
  - agreeableness
  - neuroticism

- **Model from Facebook data**
  (Park et. al 2014)
Personality

- mentally ill users:
  1. high on neuroticism
  2. more introverted
  3. less agreeable
Personality

● mentally ill users:
1. high on neuroticism
2. more introverted
3. less agreeable

● controlling for age and gender
Personality

![Bar chart showing personality traits and their comparisons.](chart.png)

- **Gender**
  - Depressed vs. Control: 0.56
  - PTSD vs. Control: 0.52
  - Depressed vs. PTSD: 0.56

- **Age**
  - Depressed vs. Control: 0.74
  - PTSD vs. Control: 0.78
  - Depressed vs. PTSD: 0.78

- **Big5**
  - Depressed vs. Control: 0.80
  - PTSD vs. Control: 0.81
  - Depressed vs. PTSD: 0.78

[Logo: Penn World Well-Being Project]
Age, Gender, Personality

![Bar chart showing correlation coefficients for Depressed vs Control, PTSD vs Control, and Depressed vs PTSD for Gender, Age, Big5, and Age, Gender, Big5.](chart.png)
Affect and Intensity • Model trained on 3000 annotated FB posts and applied to all user posts (to be published)

- circumplex model similar to valence & arousal (ANEW)
Affect and Intensity • Model trained on 3000 annotated FB posts and applied to all user posts (to be published) • circumplex model similar to valence & arousal (ANEW)
Affect and Intensity

- mentally ill users are less aroused and less positive
LIWC

- standard psychologically inspired dictionaries
- 64 categories such as:
  - parts-of-speech
  - topical categories
  - emotions
- standard baseline for open vocabulary approaches
LIWC
LIWC

7 features

64 features

PPenn | World Well-Being Project
Topics

- posteriors computed using Latent Dirichlet Allocation (LDA)
- underlying set of Facebook statues (same data as personality model)
- 2000 topics in total
Topics
Topics

7 features
64 features
2000 features
Topics: Depression

Topics controlled for age and gender
Topics: PTSD

Topics controlled for age and gender
Topics: PTSD, Depression, & Neuroticism
+ Dep, +++ PTSD
++ Dep, ++ PTSD
+++ Dep, 0 PTSD
Topics

![Bar chart showing topics and their comparisons]

- Gender
- Age
- Big5
- Age, Gender, Big5
- Affect, Intensity
- LIWC
- Topics

*Depressed vs. Control* | *PTSD vs. Control* | *Depressed vs. PTSD*
1-3 grams
1-3 grams

7

64

2k

~25k

Penn | World Well-Being Project
1-3 grams: Depressed vs. Controls
1-3 grams: PTSD vs. Controls
1-3 grams: Depressed vs. PTSD

Almost nothing left when controlling for age and gender
Other features…

- use metadata features
  # friends, #statuses

- use different word clusters
  Brown clustering, NPMI Spectral clustering, Word2Vec/GloVe embeddings

- linear ensemble of logistic regression classifiers

Mental Illness detection at the World Well-Being Project for the CLPsych 2015 Shared Task
D. Preotiuc-Pietro, M. Sap, H.A. Schwartz, L. Ungar
ROC Curve
Depressed vs. Controls
ROC Curve

PTSD vs. Controls
ROC Curve
Depressed vs. PTSD
Take Home

- Control the analysis for age & gender
Take Home

- Control the analysis for age & gender

- Personality plays an important role in mental illnesses
  (depression auc: 7 features -> .78; 25k features-> .86)
Take Home

● Control the analysis for age & gender

● Personality plays an important role in mental illnesses (depression auc: 7 features -> .78; 25k features-> .86)

● Language use of depressed/PTSD reveals symptoms, emotions, and cognitive processes.

Penn | World Well-Being Project
Thank you!