Today

• Depression
• Schizophrenia
• Dementia – Alzheimer’s Disease
• Ethics
Language as a marker
Quantifying Mental Health Signals in Twitter (Coppersmith et al., 2014)

- Automatically identify self-expressions of mental illness diagnoses
  - Leverage these to construct a labeled dataset
Quantifying Mental Health Signals in Twitter (Coppersmith et al., 2014)

• Genuine statements of diagnosis:
  – @USER The VA diagnosed me with PTSD, so I can’t go in that direction anymore
  – I wanted to share some things that have been helping me heal lately. I was diagnosed with severe complex PTSD and... LINK

• Disingenuous statement of diagnosis:
  – LOL omg my bro the “psychologist” just diagnosed me with seasonal ADHD AHAHAHA********IM DYING.
Quantifying Mental Health Signals in Twitter (Coppersmith et al., 2014)

Table 2: Number of users matching the diagnosis regular expression, *users* labeled with genuine diagnoses and *tweets* retrieved from diagnosed users for each mental health condition.

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Users</th>
<th>Tweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipolar</td>
<td>6k</td>
<td>394</td>
<td>992k</td>
</tr>
<tr>
<td>Depression</td>
<td>5k</td>
<td>441</td>
<td>1.0m</td>
</tr>
<tr>
<td>PTSD</td>
<td>477</td>
<td>244</td>
<td>573k</td>
</tr>
<tr>
<td>SAD</td>
<td>389</td>
<td>159</td>
<td>421k</td>
</tr>
<tr>
<td>Control</td>
<td>10k</td>
<td>5728</td>
<td>13.7m</td>
</tr>
</tbody>
</table>
Features

• LIWC
• Language Models (LMs)
  – Unigrams (ULM)
  – Characters (CLM)
• Pattern of life
  – Social engagement
  – Insomnia
  – Exercise
  – Sentiment
Depression

• Mood disorder that causes a persistent feeling of sadness and loss of interest

• Statistics
  – 16 million adults had at least 1 major depressive episode in 2012 (NIMH)
  – 350 million people worldwide suffer from depression (WHO)
  – Depression is the cause of over 2/3 of suicides in the US each year (White House Conference on Mental Health)
  – Women experience depression at 2x the rate of men (Journal of AMA)
TellTale Signs of Depression

Everyone feels sad, lonely or depressed at times. But when these feelings last for a long time and become overwhelming, it maybe time to seek medical help.

- Loss of interest in hobbies and activities
- Difficulty remembering, concentrating or making decisions
- Loss of appetite or over-eating
- Insomnia or excessive sleeping
- Continuous sad, anxious, or negative thoughts
- Thoughts of suicide, or suicide attempts
- Feeling worthless, guilty or helpless
- Feeling irritable or having a short temper
- Feeling hopeless or pessimistic
- Feeling fatigued

Sources: webmd.com | helpguide.org | mayoclinic.com
Diagnosis

• Diagnostic assessment by GP, psychologist, or psychiatrist
  – Examine biological, psychological, social factors
• Rating scales
  – Hamilton rating scale for depression
  – Beck depression inventory
  – Suicide behaviors questionnaire
Treatments

• Psychotherapy
• Mediation – antidepressants
• Electroconvulsive therapy
• Lifestyle
  – Exercise
  – Smoking cessation
  – Diet
Predicting Depression via Social Media
(De Choudhury et al., 2013)

• Use crowdsourcing to identify Twitter users with clinical depression
• Measure behavioral attributes from tweets 1 yr prior to diagnosis
• Estimate risk of depression *before* diagnosis
Data

• Depression screening test: CES-D questionnaire
  – http://cesd-r.com
  – Amazon Mechanical Turk
  – Workers could opt in to share Twitter username
  – Quality control

• Self-reported information
  – Clinical diagnosis of depression
  – Estimated time of onset
  – Using antidepressants
Data

• 1,583 responses; 40% shared Twitter feeds
• 476 users with depression diagnosis (after removing noisy responses)
  – 243 male, 233 female
• 171 users who scores positive for depression on CES-D
• 2,157,992 tweets retrieved
Data

Having a job again makes me happy. Less time to be depressed and eat all day while watching sad movies.

“Are you okay?” Yes…. I understand that I am upset and hopeless and nothing can help me… I’m okay… but I am not alright

“empty” feelings I WAS JUST TALKING ABOUT HOW I HAVE EMOTION OH MY GOODNESS I FEEL AWFUL

I want someone to hold me and be there for me when I’m sad.

Reloading twitter till I pass out. *lonely* *anxious* *butthurt* *frustrated* *dead*
Measuring Depressive Behavior

- Engagement
- Insomnia index
- Egocentric network measures
- Emotion
- Linguistic style
- Depression language
Tweet Activity and Depression

\[ f(x) = 0.00070393x^2 + -0.015183x + 0.10547 \]
\[ f(x) = -6.9926e-05x^2 + 0.0045778x + 0.03255 \]
Depression prediction

<table>
<thead>
<tr>
<th>Feature</th>
<th>Precision</th>
<th>Recall</th>
<th>Acc. (+ve)</th>
<th>Acc. (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>engagement</td>
<td>0.542</td>
<td>0.439</td>
<td>53.212%</td>
<td>55.328%</td>
</tr>
<tr>
<td>ego-network</td>
<td>0.627</td>
<td>0.495</td>
<td>58.375%</td>
<td>61.246%</td>
</tr>
<tr>
<td>emotion</td>
<td>0.642</td>
<td>0.523</td>
<td>61.249%</td>
<td>64.325%</td>
</tr>
<tr>
<td>linguist. style</td>
<td>0.683</td>
<td>0.576</td>
<td>65.124%</td>
<td>68.415%</td>
</tr>
<tr>
<td>dep. language</td>
<td>0.655</td>
<td>0.592</td>
<td>66.256%</td>
<td>69.244%</td>
</tr>
<tr>
<td>demographics</td>
<td>0.452</td>
<td>0.406</td>
<td>47.914%</td>
<td>51.323%</td>
</tr>
<tr>
<td>all features</td>
<td>0.705</td>
<td>0.614</td>
<td>68.247%</td>
<td>71.209%</td>
</tr>
<tr>
<td>dim. reduced</td>
<td>0.742</td>
<td>0.629</td>
<td><strong>70.351%</strong></td>
<td><strong>72.384%</strong></td>
</tr>
</tbody>
</table>
Schizophrenia

• Chronic mental disorder

• Symptoms:
  – Positive: hallucination, delusion, thought disorders, movement disorders
  – Negative: “flat affect”, reduced feelings of pleasure, reduced speaking
  – Cognitive: poor executive functioning, trouble focusing or paying attention, problems with working memory
Schizophrenia and language

• “Negative thought disorder”
  – Alogia – poverty of speech

• “Positive thought disorder”
  – Derailment
  – Tangentiality
Schizophrenia and language

• Derailment

“I always liked geography. My last teacher in that subject was Professor August A. He was a man with black eyes. I also like black eyes. There are also blue and grey eyes and other sorts, too...” (Bleuler, 1950)

• Tangentiality

“Well, in myself I have been okay what with the prices in the shops being what they are and my flat is just round the corner. I keep a watch for the arbiters most of the time since it is just round the corner. There is not all that much to do otherwise.”
Schizophrenia and language

• Word-level abnormalities

“I got so angry I picked up a dish and threw it at the geshinker”
“So I sort of bawked the whole thing up”
“They’re destroying too many cattle and oil just to make soap. If we need soap when you can jump into a pool of water, and then when you go to buy your gasoline, my folks always thought they should, get pop but the best thing to get, is motor oil, and, money. May may as well go there and, trade in some, pop caps and, uh, tires, and tractors to grup, car garages, so they can pull cars away from wrecks, is what I believe in. So I didn’t go there to get no more pop when my folks said it. I just went there to get a ice-cream cone, and some pop, in cans, or we can go over there to get a cigarette”
Quantifying the Language of Schizophrenia in Social Media (Mitchell et al., 2015)

• Data: 174 Twitter users with self-stated diagnosis of schizophrenia
  – Age and gender matched controls, balanced dataset

• Features:
  – LIWC
  – Open-vocabulary approaches
    • LDA, Brown clustering, character n-grams, perplexity
LIWC analysis

– Schizophrenia users had more words from these categories: Cognitive mechanisms, death, function words, negative emotion

– And fewer words from these categories: home, leisure, positive emotion
## Classification results

<table>
<thead>
<tr>
<th>Features</th>
<th>SVM</th>
<th>MAXENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perplexity (ppl)</td>
<td>52.0</td>
<td>51.4</td>
</tr>
<tr>
<td>Brown-Cluster Dist (BDist)</td>
<td>53.3</td>
<td>72.3</td>
</tr>
<tr>
<td>LIWC</td>
<td>68.8</td>
<td>70.8</td>
</tr>
<tr>
<td>CLM</td>
<td>77.1</td>
<td>77.2</td>
</tr>
<tr>
<td>LIWC+CLM</td>
<td>78.2</td>
<td>77.2</td>
</tr>
<tr>
<td>LDA Topic Dist (TDist)</td>
<td>80.4</td>
<td>80.4</td>
</tr>
<tr>
<td>CLM+TDist+BDist+ppl</td>
<td>81.2</td>
<td>79.7</td>
</tr>
<tr>
<td>CLM+TDist</td>
<td>81.5</td>
<td>81.8</td>
</tr>
<tr>
<td>LIWC+TDist</td>
<td><strong>82.3</strong></td>
<td><strong>81.9</strong></td>
</tr>
</tbody>
</table>
Linguistic Analysis of Schizophrenia in Reddit Posts

Reddit:
- No limits on post length
- Subreddits
- Python API Wrapper (PRAW)
Data

- Subreddits:
  - r/schizophrenia
  - r/schizophrenic
  - r/AskReddit: “Any Redditors With Schizophrenia?”
- Manual inspection for formal diagnosis
  - e.g. “my diagnosis of schizophrenia”
Data

- SZ: 159 users; 66,454 comments
- Control: 159 users; 113,570 comments

At least 10 posts per user
LIWC Analysis

• Comparison of LIWC findings across 5 studies in multiple domains (Reddit, Twitter, discussion boards)

• SZ indicators across studies:
  - increased: negative emotion, first person singular pronouns, tentative, cognitive process, health, anxiety, third person plural pronouns
  - decreased: leisure

• SZ indicators in Reddit:
  - increased: word count
LIWC Analysis

Logistic Regression: 81.56% accuracy

<table>
<thead>
<tr>
<th>Control (CTL)</th>
<th></th>
<th>Schizophrenia (SZ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>Feature</td>
<td>Weight</td>
</tr>
<tr>
<td>-1.2748</td>
<td>Sadness</td>
<td>1.6105</td>
</tr>
<tr>
<td>-1.1109</td>
<td>Quotation mark</td>
<td>1.0717</td>
</tr>
<tr>
<td>-0.8715</td>
<td>3rd person singular</td>
<td>1.0614</td>
</tr>
<tr>
<td>-0.7956</td>
<td>Feel</td>
<td>0.9825</td>
</tr>
<tr>
<td>-0.7949</td>
<td>Articles</td>
<td>0.9426</td>
</tr>
<tr>
<td>-0.7302</td>
<td>Nonfluencies</td>
<td>0.9304</td>
</tr>
<tr>
<td>-0.6705</td>
<td>Adjectives</td>
<td>0.8021</td>
</tr>
<tr>
<td>-0.6329</td>
<td>See</td>
<td>0.7642</td>
</tr>
<tr>
<td>-0.6214</td>
<td>Motion</td>
<td>0.6975</td>
</tr>
<tr>
<td>-0.6182</td>
<td>Present focus</td>
<td>0.6478</td>
</tr>
</tbody>
</table>
Dementia

• Broad category of brain diseases that cause long term decrease in ability to think and remember
• Most common type - Alzheimer’s disease (AD)
  – 60%-70% of cases
• Affects 27.5 million people
  – 3% of people 65-74
  – 19% of people 75-84
  – ~50% of people >85
• $604 billion in costs per year
AD Symptoms

- Short-term memory loss
- Problems with language
- Disorientation
- Mood swings
- Decreased motivation
Cause

• 70% of risk is believed to be genetic
• Risk factors:
  – Head injuries
  – Depression
  – Hypertension
Pathology

• Amyloid plaques
• Tau tangles
Amyloid plaques
Diagnosis

• Can only be definitively diagnosed in a postmortem brain tissue examination
• Clinical assessment
Clinical Assessment

• Medical history
• Physical exam
• Neurological exam
• Mental status tests
Mental status tests

• Mini-mental state exam (MMSE)

• Mini-cog
Normal Score 10

Mild Cognitive Impairment (Numbers error and placement of hands) Score 8

Moderate Cognitive Impairment Score 4

Severe Cognitive Impairment Score 2

Sunderland, 1989
Stages of AD

Mild Cognitive Impairment
- Duration: 7 years
- Disease begins in Medial Temporal Lobe
- Symptoms: Short-term memory loss

Mild Alzheimer’s
- Duration: 2 years
- Disease spreads to Lateral Temporal & Parietal Lobes
- Symptoms include:
  - Reading problems
  - Poor object recognition
  - Poor direction sense

Moderate Alzheimer’s
- Duration: 2 years
- Disease spreads to Frontal Lobe
- Symptoms include:
  - Poor judgment
  - Implusivity
  - Short attention

Severe Alzheimer’s
- Duration: 3 years
- Disease spreads to Occipital Lobe
- Symptoms include: Visual problems
Stages of AD

• Pre-dementia
  – Mild Cognitive Impairment (MCI)
  – Short-term memory loss

• Early
  – Increased impairment of learning and memory

• Moderate
  – Increased impairment of learning and memory
Treatment

• Medication for symptoms
• Clinical trials
Language Indicators of AD

• Can we predict AD from writing analysis?
Case study: Agatha Christie
Nun Study

• Longitudinal study, 1986- (David Snowden)
• 678 Roman Catholic sisters
• Homogenous group
• Autobiographical essays
Nun Study - excerpts

• "It was about a half hour before midnight between February 28 and 29 of the leap year 1912 when I began to live, and to die, as the third child of my mother, whose maiden name is Hilda Hoffman, and my father, Otto Schmidt..."

• "I was born in Eau Claire, Wisconsin on May 24, 1913, and was baptized in St. James Church..."
Nun Study - findings

• Sisters who scored poorly on
  – Idea density
  – Grammatical complexity
  were much more likely to develop dementia

• E.g. sisters in lower third of idea density were 60 times more likely to develop AD than sister in upper third

• Using essays, predict with 92% accuracy whether the brain would contain plaques post-mortem
Vector-space topic models for detecting Alzheimer’s disease (Yancheva & Rudzicz, 2016)

• Measure semantic content deficiency
  – Idea density – ratio of semantic units to total number of words
  – Efficiency – rate of semantic units over duration of speech sample
Measuring semantic content

• Picture description task
• Human-supplied information content units (hsICUs)
Method

• Data: DementiaBank
  – Samples:
    http://talkbank.org/browser/index.php?url=DementiaBank/English/Pitt/Control/cookie/013-0.cha
## DementiaBank Subset

<table>
<thead>
<tr>
<th>Class</th>
<th>Subjects</th>
<th>Samples</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>168</td>
<td>255</td>
<td>24,753</td>
</tr>
<tr>
<td>CT</td>
<td>98</td>
<td>241</td>
<td>26,654</td>
</tr>
<tr>
<td>Total</td>
<td>266</td>
<td>496</td>
<td>51,407</td>
</tr>
</tbody>
</table>

*Table 1: Distribution of dataset transcriptions.*
Automatic ICU generation

- Train word vector model on large general purpose corpus (GloVe)
- Extract vector representations of words in DementiaBank corpus
- Cluster vectors separately for AD and control group -> represent “topics”
<table>
<thead>
<tr>
<th>ID</th>
<th>Cluster words</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0</td>
<td>window, floor, curtains, plate, kitchen</td>
</tr>
<tr>
<td>C1</td>
<td>dishes, dish</td>
</tr>
<tr>
<td>C2</td>
<td>running, standing, action, hand, counter</td>
</tr>
<tr>
<td>C3</td>
<td>water, sink, drying, overflowing, washing</td>
</tr>
<tr>
<td>C4</td>
<td>stool, legged</td>
</tr>
<tr>
<td>C5</td>
<td>mother, boy, girl, sister, children</td>
</tr>
<tr>
<td>C6</td>
<td>cookie, cookies, sakes, cream</td>
</tr>
<tr>
<td>C7</td>
<td>jar, cups, lid, dried, bowl</td>
</tr>
<tr>
<td>C8</td>
<td>see, going, getting, looks, know</td>
</tr>
<tr>
<td>C9</td>
<td>reaching, falling, fall, summer, growing</td>
</tr>
<tr>
<td>D0</td>
<td>cookie, cookies, cake, baking, apples</td>
</tr>
<tr>
<td>D1</td>
<td>dishes, dish, eating, bowls, dinner</td>
</tr>
<tr>
<td>D2</td>
<td>boy, girl, mother, sister, lady</td>
</tr>
<tr>
<td>D3</td>
<td>going, see, getting, get, know</td>
</tr>
<tr>
<td>D4</td>
<td>stool, floor, window, chair, curtains</td>
</tr>
<tr>
<td>D5</td>
<td>jar, cups, jars, dried, honey</td>
</tr>
<tr>
<td>D6</td>
<td>sink, drying, washing, spilling, overflowing</td>
</tr>
<tr>
<td>D7</td>
<td>mama, huh, alright, johnny, ai</td>
</tr>
<tr>
<td>D8</td>
<td>running, fall, falling, reaching, hand</td>
</tr>
<tr>
<td>D9</td>
<td>water, dry, food</td>
</tr>
</tbody>
</table>
Cluster analysis

• Do clusters match hsICUs?
• Do the topics differ between groups?
Cluster analysis

• Do clusters match hsICUs?
  – Yes (except “dishcloth”)

• Do the topics differ between groups?
  – Yes!
  – Control group: overflowing, sink, indifferent, mother, apron, window, curtain, cupboard, counter
  – AD group: brother, sister, son, daughter
Quantifying Irrelevance

• Align pairs of clusters between 2 models
• All control clusters are recalled by AD model
• D7 not recalled by control model – “extraneous terms”
Are topics discussed in same contexts?

• Augment word vectors with local context windows

• Results:
  – All control cluster words were used in the same contexts by both groups
  – Frequency of control words is higher in control group than AD group
Classification

• Features:
  – Distance metrics for AD and control clusters
  – Idea density
  – Idea efficiency

• Random Forest classifier, 10-fold cross-validation

• Compare models (control, AD, combined), feature sets, and context

• Baseline: hsICU features
## Classification results

<table>
<thead>
<tr>
<th>Model</th>
<th>Features</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>hsICUs</td>
<td>0.73</td>
<td>0.74</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>Baseline</td>
<td>LS&amp;A</td>
<td>0.76</td>
<td>0.77</td>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>Baseline</td>
<td>hsICUs + LS&amp;A</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td><strong>0.80</strong></td>
</tr>
<tr>
<td>control</td>
<td>distance-based</td>
<td>0.68</td>
<td>0.69</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>dementia</td>
<td>distance-based</td>
<td>0.66</td>
<td>0.67</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>combined</td>
<td>distance-based</td>
<td>0.68</td>
<td>0.69</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>control</td>
<td>distance-based + idea density + idea efficiency</td>
<td>0.74</td>
<td>0.76</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>dementia</td>
<td>distance-based + idea density + idea efficiency</td>
<td>0.74</td>
<td>0.75</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>combined</td>
<td>distance-based + idea density + idea efficiency</td>
<td>0.74</td>
<td>0.75</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>control</td>
<td>distance-based + idea density + idea efficiency + LS&amp;A</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>dementia</td>
<td>distance-based + idea density + idea efficiency + LS&amp;A</td>
<td>0.77</td>
<td>0.78</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>combined</td>
<td>distance-based + idea density + idea efficiency + LS&amp;A</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td><strong>0.80</strong></td>
</tr>
</tbody>
</table>
Ethics
Ethics in NLP workshop

• http://www.ethicsinnlp.org/accepted-papers
Conclusions

• NLP and speech processing are very useful tools for prediction of mental illness!
• Social media approaches for data collection
• Ethical considerations