Future of Mental Health Sensing on College Campuses

Andrew T. Campbell, Dartmouth College
HotMobile, February 12, 2018
what happens if life throws you a googly?
what is mental illness?
fast forward 35 years
Dartmouth's depression curve
19% of Dartmouth students were diagnosed with depression in 2016.

24% reported depression has having an impact on academic performance.

28% have seen a mental health counselor.
why do students burn out, drop classes, do poorly, even drop out of college when others excel?

what is the impact of stress, mood, workload, sociability, sleep and mental health on academic performance?

is there a set of behavioral trends or signature to the semester?

can mobile sensing and predictive modeling assess student mental wellbeing?
most faculty are unaware that their students are struggling beyond grades
we subjectively know there is a cycle to the term or semester
but there is no objective data
48 students over 10 week Spring 2013 term

• 10 female, 38 male CS students
• 30 undergraduates, 18 graduates
• 8 seniors, 14 juniors, 6 sophomores, 2 freshmen, 3 Ph.D students, 1 second-year Masters student, and 13 first-year Masters students
• 23 Caucasians, 23 Asians and 2 African-Americans.
• Gave out Androids to students in my Android programming class
• No feedback to students, only collection.
• No funding so begged and borrowed and just did it — you have believe.
sensing system
automatic continuous sensing
- accelerometer
- microphone
- light Sensor
- GPS/Bluetooth

self-reports
- SurveyMonkey
- mobile EMA

behavioral classifiers
- activity
- conversation
- sleep
- location/co-location

StudentLife cloud

statistical analysis
- mental health
- academic performance
- behavioral term trends
- management scripts

Android phone

cloud
classifiers
activity

- sitting
- standing
- walking
- running
around conversation

face-to-face conversation: duration and frequency
<table>
<thead>
<tr>
<th>Activity Feature</th>
<th>Feature</th>
<th>Stationary duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound Feature</td>
<td>Feature</td>
<td>Silence duration</td>
</tr>
<tr>
<td>Light Feature</td>
<td>Feature</td>
<td>Darkness duration</td>
</tr>
<tr>
<td>Phone Usage Features</td>
<td>Features</td>
<td>Phone-lock, charging, phone-off duration</td>
</tr>
</tbody>
</table>

**Linear regression model**

\[
Sl = \sum_{i=1}^{6} \alpha_i \cdot F_i, \quad \alpha_i \geq 0
\]

\[
\min_{\alpha_i} \sum_{j=1}^{4} (Sl^j - \sum_{i=1}^{6} \alpha_i \cdot F_i^j)^2
\]

Sleep duration
we also computed

- activity duration
- outdoor mobility
- indoor mobility
- location and co-location
- phone usage (apps, lock/unlock)
ecological momentary assessment (EMA)
32,000 EMAs
>9000 facelog images
### Patient Health Questionnaire (PHQ-9)

**NAME:**

**DATE:**

Over the last 2 weeks, how often have you been bothered by any of the following problems?  
(use "x" to indicate your answer)

<table>
<thead>
<tr>
<th>Problem</th>
<th>Not at all</th>
<th>Several days</th>
<th>More than half the days</th>
<th>Nearly every day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Little interest or pleasure in doing things</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2. Feeling down, depressed, or hopeless</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3. Trouble falling or staying asleep, or sleeping too much</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4. Feeling tired or having little energy</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5. Poor appetite or overeating</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7. Trouble concentrating on things, such as reading the newspaper or watching television</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8. Moving or speaking so slowly that other people could have noticed, Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9. Thoughts that you would be better off dead, or of hurting yourself</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

### Depression Severity

<table>
<thead>
<tr>
<th>Severity</th>
<th>Minimal</th>
<th>Minor</th>
<th>Moderate</th>
<th>Moderately Severe</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PHQ-9 Score</strong></td>
<td>1-4</td>
<td>5-9</td>
<td>10-14</td>
<td>15-19</td>
<td>20-27</td>
</tr>
</tbody>
</table>

### Number of Students (Pre-Survey)

| Depression Severity | 17 | 15 | 6 | 1 | 1 |

### Number of Students (Post-Survey)

| Depression Severity | 19 | 12 | 3 | 2 | 2 |

(Higher scores indicate worse depression. Refer to the interpretation matrix above.)

---

**Note:** For interpretation of TOTAL score, refer to the accompanying interpretation matrix.
pre-post questionnaires

- depression scale
- perceived stress scale
- loneliness scale
- flourishing scale
- Big 5
study design
recruiting participants
recruiting participants → orientation

consent form

mental health pre-surveys
recruiting participants → orientation → data collection

passive sensing

- accelerometer
- microphone
- light Sensor
- GPS/Bluetooth

activity → conversation → sleep → location/co-location

ecological momentary assessment (EMA)

[Image of app interface]
recruiting participants → orientation → data collection → exit interview/survey

exit interview
mental health post-surveys
educational data
**Dataset size**
53 GB of data, 32,000 EMAs, 48 pre-post surveys, interviews

**Passive sensor data from phone**
- activity, sleep, face-to-face conversation frequency/duration, indoor and outdoor mobility, location, distance travelled, co-location, light, app usage, calendar, call logs

**Experience sampling**
- pam (affect), behavioral, class, campus events, social events, sleep quality, exercise, comments, mood

**Pre-post surveys from Survey Monkey**
- stress, personality, mental and physical health, loneliness, mood, sleep

**Transcripts:** educational stats

**Other:** Facebook (not released), face log images (not released), dining details, seating data

**Entry-exit interviews** (not released)
- social net in class, classes information (deadlines for all classes), group review, study specific questions.
StudentLife is the first study that uses passive and automatic sensing data from the phones of a class of 48 Dartmouth students over the academic year to assess their mental health (e.g., depression, loneliness, stress), academic performance (grades across all their courses and their cumulative GPA), and behavioral trends (e.g., how stress, sleep, visits to the gym, etc. change in response to college life).
results
<table>
<thead>
<tr>
<th>Time</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
</tr>
</thead>
<tbody>
<tr>
<td>08</td>
<td>9</td>
<td>9</td>
<td>9L</td>
<td>9L</td>
<td>9S</td>
</tr>
<tr>
<td>09</td>
<td>9S</td>
<td>9S</td>
<td>9S</td>
<td>9x</td>
<td>9S</td>
</tr>
<tr>
<td>10</td>
<td>10A</td>
<td>9A</td>
<td>10</td>
<td>9Lx</td>
<td>9S</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>10x</td>
<td>12</td>
<td>10x</td>
<td>10A</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>2x</td>
<td>2</td>
<td>2</td>
<td>11x</td>
</tr>
<tr>
<td>01</td>
<td>2</td>
<td>2A</td>
<td>2</td>
<td>2</td>
<td>12x</td>
</tr>
<tr>
<td>02</td>
<td>3</td>
<td>3A</td>
<td>3</td>
<td>3</td>
<td>2A</td>
</tr>
<tr>
<td>03</td>
<td>2A</td>
<td>3B</td>
<td>2A</td>
<td>3A</td>
<td>3A/Bx</td>
</tr>
<tr>
<td>04</td>
<td>2</td>
<td>3B</td>
<td>2A</td>
<td>3A</td>
<td>3A/Bx</td>
</tr>
<tr>
<td>05</td>
<td>3A</td>
<td>3B</td>
<td>3A</td>
<td>3A/Bx</td>
<td></td>
</tr>
<tr>
<td>06</td>
<td>3A/Bx</td>
<td></td>
<td>3A/Bx</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Time of the day (8 am - 6 pm)*
day: 9 am - 6 pm
night: 12 am - 9 am

- 9 pm~10 pm
- 10 pm~11 pm
- 12 am
- 1 am~2 am
- 2 am~3 am
- 3 am~4 am
- 4 am~5 am
- 5 am~6 am
- 6 am~7 am

number of instances

0 50 100 150 200 250

time of day
evening: 6 pm-12 am
mental health
depression

- sleep duration *
- conversation frequency (day) **
- conversation frequency (evening) *
- number of co-locations *

*p ≤ 0.05, **p ≤ 0.01
loneliness

activity duration *
activity duration (evening) **
outdoor mobility (day) *
indoor mobility (day) *

R value

*p ≤ 0.05, **p ≤ 0.01
stress

conversation duration during day *
conversation freq during day **
conversation freq during evening *
sleep duration *

*p ≤ 0.05, **p ≤ 0.01
behavioral trends
sleep

mid-term

deadlines

day

number of deadlines
face-to-face conversation

mid-term
activity duration

mid-term

deadlines

activity duration
stress and affect

deadlines positive affect (PA) stress

mid-term
class attendance

The graph shows the number of deadlines and attendance over the course of the semester. The shaded area indicates the mid-term period. The blue line represents deadlines, and the orange line represents attendance. Attendance generally decreases as the number of deadlines increases.
there is no correlation between class attendance and grade!
idea

- look at individual differences between high and low performers
- define higher level behaviors for studying and partying
- track behavior changes using time series analysis
capturing behavioral change

behavioral slope and breakpoints
studying
Your phone knows when you are studying.
labelled study areas
attending classes and studying

![Graph showing attendance and study duration over weeks. The x-axis represents weeks from 1 to 9, the y-axis represents attendance from 0 to 1, and the z-axis represents study duration (hours) from 0 to 24. The graph highlights the midterm period from week 4 to week 6, during which attendance decreases and study duration increases.]
partying
partying
party places
sound
activity co-location
partying
drinking
when do students party?

when do students study?
partying trends across the term

<table>
<thead>
<tr>
<th>week 1</th>
<th>week 3 (mid term)</th>
<th>week 5</th>
<th>week 7 (green key)</th>
<th>week 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>party duration (hours)</td>
<td>180</td>
<td>135</td>
<td>90</td>
<td>45</td>
</tr>
</tbody>
</table>
GPA

- activity post-slope **
- conversation duration *
- indoor mobility term-slope **
- class attendance pre-slope *
- party duration term-slope*
- study duration mean *
- study focus pre-slope *

*p ≤ 0.05, **p ≤ 0.01

R value

-0.7 to 0.525
Predict GPA?
We use the Lasso regularized linear regression model to determine lasso’s hyperparameter and to obtain the prediction results.

leave-one-out cross validation to determine lasso's hyperparameter and to obtain the prediction results
prediction performance

MAE = 0.179

goodness of fit:
• $R^2 = 0.559$
• $r = 0.81$, $p < 0.01$
a professor’s intervention
why no signal?
it has to be there
somewhere, no?
College age young adults 18–25 are more likely to have major depressive episodes than any other age groups.

Vast majority of college students who screen positively for depression never seek mental health services.

Suicide is the second leading cause of death among people ages 15–24.
depression sensing
83 undergraduate students across two 10 week term in the winter and spring terms 2016
• 44 female, 40 male students
• 26 Asians, 5 African-Americans, 24 Caucasians, 1 multiracial, 26 not specified
• average age 20.33, std=2.31
• Users’ Android, iPhone, microsoft band 2
Diagnostic Criteria for Major Depressive Disorder and Depressive Episodes

DSM-5 Criteria for Major Depressive Disorder (MDD)

- Depressed mood or a loss of interest or pleasure in daily activities for more than two weeks.
- Mood represents a change from the person's baseline.
- Impaired function: social, occupational, or personal.
- Specific symptoms, at least 5 present and present nearly every day:
  1. Depressed mood or irritability nearly every day, as indicated by either subjective report (e.g., Feeling sad or empty) or observation made by others (e.g., appears tearful).
  2. Decreased interest or pleasure in most activities, most of each day.
  3. Significant weight change (10% or more) or change in appetite.
  4. Change in sleep: Insomnia, hypersomnia.
  5. Change in activity: Psychomotor agitation or retardation.
  6. Fatigue or loss of energy.
  7. Guilt or worthlessness; feelings of worthlessness or excessive or inappropriate guilt.
  8. Concentration difficulties, inability to think or concentrate, or more indecisiveness.
  9. Suicide ideation, plan, or suicide attempt.

DSM-V proposed (not in the text): Anxiety symptoms that may indicate depression: Irrational worry, preoccupation with worst case, trouble relaxing, feeling tense, fear that something awful might happen.

Screen for conditions that may mimic or coexist with Major Depressive Disorder:
- Substance abuse causing depressed mood (e.g., drugs, alcohol, medications).
- Medical illness causing depressed mood.
- Other psychiatric disorders: mania, hypomania, bipolar, schizoaffective, schizophrenia, etc.

Episode unless sx persist for > 2 months or show marked functional impairment, morbid preoccupation with depressive ideation, psychotic symptoms, or psychomotor retardation.

Mood Episodes (may be part of Major Depressive Disorder OR an isolated episode)

- Loss of interest and enjoyment in usual activities.
- Reduced energy and increased activity.
- Reduced self-esteem and confidence.
- Ideas of guilt and unworthiness.
- Pessimistic thoughts.
- Disturbed sleep.
- Diminished appetite.
- Ideas of self-harm.

Severity of Depressive Episodes

Mild:
1 sx from column A, and 0 or 1 sx from column B. Or 2-6 sx but mild in severity and functional impairment.

Moderate:
2 or 3 sx from column A plus 0-2 sx from column B. Or 3-6 sx but moderate functional impairment.

Severe:
All 3 from column A plus > 3 sx from column B. Or fewer sx but any of these: Severe functional impairment, psychotic sx, recent suicide attempt, or has specific suicide plan or clear intent.

<table>
<thead>
<tr>
<th>Functional Domain</th>
<th>Mildly Impaired</th>
<th>Severely Impaired</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Relationships</td>
<td>Current</td>
<td>Withdrawn, won't talk, brusque, angry, aggressive</td>
</tr>
<tr>
<td>School &amp; Academics/Work</td>
<td>Grades deteriorate, frequent absences, cutting classes or work, decreased effort, moderate academic or work stress</td>
<td>Failing performance, missing school or work, doesn't care about work, oppositional, argumentative, high academic or work stress</td>
</tr>
</tbody>
</table>
we hypothesize that mobile sensor data from phones and wearables represent proxies for the DSM-5 depression symptoms in college students; that is, we design a set of behavioral features to capture the characteristics of the depression symptoms that “take into account lifestyles of students” (e.g., going to class, working in study areas, socializing on campus)
Symptoms to phone behavioral features

Change in sleep
- sleep onset, sleep wakeup, sleep duration

Decreased interest in activities
- number of places visited, physical activity, time spent in study places and dorms, and social engagement

Loss of concentration
- lock/unlock during the day, at study places and at dorms

Depressed mood/fatigue
- heart rate/heart rate variability
self-reported pre-post PHQ8 and weekly PHQ4
PATIENT HEALTH QUESTIONNAIRE (PHQ-9)

NAME: __________________________ DATE: __________________________

Over the last 2 weeks, how often have you been bothered by any of the following problems? (Use "X" to indicate your answers)

1. Little interest or pleasure in doing things
2. Feeling down, depressed, or hopeless
3. Trouble falling or staying asleep, or sleeping too much
4. Feeling tired or having little energy
5. Poor appetite or overeating
6. Feeling bad about yourself—either that you are a failure or that people are against you
7. Trouble concentrating on things, such as reading or watching television
8. Moving or speaking so slowly that other people have noticed. Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual
9. Thoughts that you would be better off dead, or of hurting yourself

(a) The distribution of pre PHQ-8 scores
(b) The distribution of post PHQ-8 scores
(c) The distribution of PHQ-4 depression subscale scores
(d) The distribution of mean PHQ-4 depression subscale scores

(Healthcare professional: For interpretation, please refer to www.medscape.com/section/nutrition)
correlation matrix of symptom features and PHQ-8 item scores

- sleep duration -
- sleep end -
- sleep start -
- unlock duration -
- unlock duration at dorm -
- unlock duration at study places -
- stationary time -
- conversation duration -
- number of places visited -
- time at on-campus health facilities -
- time at dorm -
- time at study places -
- heart rate -

p≤0.05
PHQ-8 score: sleep changes; and diminished ability to concentrate

- sleep start time variance*
- sleep end time variance*
- unlock duration*
- unlock duration @ study places**
- unlock duration @ dorms*

*\( p \leq 0.05 \), **\( p \leq 0.01 \)
PHQ-8 score: diminished interest in activities; and depressed mood, fatigue

stationary time*
conversation duration term slope**
number of places visited*
time at on-campus health facilities
heart rate/heart rate variability

*p ≤ 0.05, **p ≤ 0.01
We use ANOVA to determine whether or not the mean of the two groups (non-depressed and depressed) significantly differ.
sleep

Box plots showing sleep duration (hours) for depression groups before and after intervention. Sleep duration is categorized into two groups: less than 10 hours and 10 hours or more. The sleep duration decreases from pre to post in both groups.
predicting weekly PHQ-4

We use lasso regularized logistic regression to determine if a student is depressed or not; selected features are:

- sleep end
- unlock duration at study places
- stationary time, number of conversations
- number of places visited, time at dorm, time at study places
- median heart rate
predicting weekly PHQ-4 ROC curve of logistic regression model obtained from 10-fold cross validation for binary classification of depression

true positive rate  
false positive rate

AUC=0.805

recall=81.5%
precision=69.1%
PHQ-4 depression

number of conversations

sleep duration

sleep time

wake time

visited places

student case study
24/7 passive sensing is here!
we have a mountain to climb
• validity of sensors and diagnostic predictive models
• flaky signals, real-world noise
• interventions
• over burdened clinicians
• privacy, privacy, privacy
new forms of intervention
personalized mental health sensing and intervention models are essential
People may experience:

**Mood:** mood swings, sadness, elevated mood, anger, anxiety, apathy, apprehension, euphoria, general discontent, guilt, hopelessness, loss of interest, or loss of interest or pleasure in activities

**Behavioral:** irritability, risk taking behaviors, disorganized behavior, aggression, agitation, crying, excess desire for sex, hyperactivity, impulsivity, restlessness, or self-harm

**Cognitive:** unwanted thoughts, delusion, lack of concentration, racing thoughts, slowness in activity, or false belief of superiority

**Psychological:** depression, manic episode, agitated depression, or paranoia

**Sleep:** difficulty falling asleep or excess sleepiness

**Weight:** weight gain or weight loss

**Also common:** fatigue or rapid and frenzied speaking
Predicting relapse in schizophrenia
sensing time series from two relapse patients
200 x 4 study

follow 200 first years undergraduate students over their four years of college

• behavioral sensing all year
• brain imagine 2 per year
• focus: anxiety and depression
• User’s phones
• largest and longest campus life study
Depression is associated with network differences in the medial prefrontal cortex. Initial analysis of the resting state functional connectivity (RSFC) data shows a relationship between RSFC and depression scale (PHQ).
Big thanks

Dror Ben-Zeev and Rachel Brian (UW), Tanzeem Choudhury and Hane Aung (Cornell), Ethan Berke (was DHMC), Randy Colvin and Stefanie Tignor (Northeastern), Sam Gosling (UT Austin), Gabriella Harari (Stanford), Catherine Norris (Swarthmore), Rui Wang, Kizito Masaba, Shayan Mirjafari, Weichen Wang, and Xia Zhou, Emily Scherer (Dartmouth)
CampusLife consortium
“StudentLife”, UbiComp, 2014
“SmartGPA”, UbiComp, 2015
“CrossCheck”, UbiComp 2016
“Predicting Symptoms”, UbiComp 2017
“Depression Tracking”, UbiComp 2018
FUTURE

mobile mental health sensing
• No major leaps forward yet
• Lots of small scale studies (N=100, T=20 weeks) with differing features/models and results
• No generalizable results, no repeated studies, and no randomized control trials (RCTs)
• Things are inconclusive but progress is being made and there is commercial activity — always a good sign
• Significant challenges remain
• Widely accepted feature set/models will emerge
• Breakthrough in personalized depression models and interventions proven out
• Advances in deep learning and wearables for mobile mental health — the phone will be a step toward that
• Multiple large scale RCT studies will confirm clinical validity of depression sensing across the general population
• Clinician acceptance of mobile PHQ9 prediction. Biz models emerge for payers/providers
Depression sensing driven intervention (e.g., next gen Cognitive behavioral therapy) is common.

Personalized depression and intervention models embedded not only mobile — maybe everything you touch (from social media, future wearables, your car).

Mental health assistant — mental health sensing based AI.

Reactive model start the move:
  - “Hey Siri how am I doing today?”

Proactive model will dominate:
  - “Andrew, go for a long run today — your open between 1-2 pm”
  - “I’ve booked an appointment for you to see Mark Tuesday”
  - “You didn’t take your meds yesterday — call Mark”
silent tsunami
44 million Americans live with mental health disorders

55% of the US counties have no MH provider – 45% of psychiatrists do not accept insurance – 57% do not accept medicaid
There are 10X more people with serious mental health (SMI) in prisons than in hospitals

25% of homeless people have SMI.
44,000 Americans die by suicide each year. There are 13.8 deaths by suicide per 100,000 persons each year.

It’s the 2nd leading cause of death among people aged 10–24.
what happens if life throws you a googly?
Mobile will detect it.
And deflect it.
StudentLife is the first study that uses passive and automatic sensing data from the phones of a class of 48 Dartmouth students over a single academic term to assess their mental health (e.g., depression, loneliness, stress), academic performance (grades across all their courses and cumulative GPA) and behavioral trends (e.g., how much stress, sleep, visits to the gym, etc. change in response to college challenges).