Enabling Socially Aware Passive Measurements of Health and Well-Being Outcomes

Munmun De Choudhury
Assistant Professor
School of Interactive Computing + GVU Center, Georgia Tech
CCC Smart and Pervasive Health Workshop | December 5, 2016
Social Dynamics and Well-Being Lab (SocWeB)
big data + machine learning + social and behavioral science
= methods and artifacts for personal and societal well-being
Social Media Derived Behavioral and Affective Markers Predict Postpartum Changes

<table>
<thead>
<tr>
<th>volume</th>
<th>replies</th>
<th>NA</th>
<th>activation</th>
<th>1st pp.</th>
</tr>
</thead>
</table>

(-ve change) → New mothers → (+ve change)

Normalized changes in each figure represents estimated time of childbirth.
<table>
<thead>
<tr>
<th>depression</th>
<th>+30%</th>
<th>differences</th>
<th>-57%</th>
</tr>
</thead>
<tbody>
<tr>
<td>useless</td>
<td>+51%</td>
<td>always a</td>
<td>-56%</td>
</tr>
<tr>
<td>suicidal</td>
<td>+34%</td>
<td>be working</td>
<td>-56%</td>
</tr>
<tr>
<td>medicine</td>
<td>+52%</td>
<td>keep your</td>
<td>-56%</td>
</tr>
<tr>
<td>locked</td>
<td>+51%</td>
<td>preferred</td>
<td>-56%</td>
</tr>
<tr>
<td>no friends</td>
<td>+51%</td>
<td>awesome I</td>
<td>-56%</td>
</tr>
<tr>
<td>loneliness</td>
<td>+50%</td>
<td>and enjoy</td>
<td>-55%</td>
</tr>
<tr>
<td>alone I</td>
<td>+34%</td>
<td>it work</td>
<td>-55%</td>
</tr>
</tbody>
</table>
Social media depression index

actual (CDC data -- BRFSS)  predicted (SMDI)
The CampusLife Project

- Pilot study at Georgia Tech involving students
- Multimodal behavior, mood, activity tracked (actively and passively) over two weeks

- Voluntarily shared Facebook data predicts stress, anxiety, and depression with over 75% accuracy
Open Questions - Methods

- What kind of new computational approaches can be developed that glean meaningful signals relating to health outcomes from noisy social media data?

- How can social media derived variables relating to health outcomes be integrated with conventional and state-of-the-art measures?

- How we measure social and community level (causal) effects of health with social media data?
Open Questions - Validity

• In community-centric studies employing large repositories of social media data, how can ground truth data on health outcomes be obtained at scale?

• How can social media based measurements of community health and well-being be evaluated in “in the wild” settings?
Open Questions - Translation

• How can we ensure that the social media data source and its measurements are usable, actionable and interpretable to different stakeholders (community-centric sensing, intervention design etc.)?

• What kind of risks exist in developing new research programs (e.g., intervention design or deployment, clinical trials) that could use social media data?