Experimental Design & Machine Learning Opportunities in Mobile Health

Susan Murphy
02.01.17

HeartSteps
JITAI

JITAI

MD2K
Center of Excellence for Mobile Sensor Data-to-Knowledge

The Methodology Center
advancing methods, improving health

INSTITUTE FOR SOCIAL RESEARCH
UNIVERSITY OF MICHIGAN

STATISTICS
The Dream!

“Continually Learning Mobile Health Intervention”

• Help you achieve, and maintain, your desired long term healthy behaviors
  – Provide sufficient short term reinforcement to enhance your ability to achieve your long term goal

• The ideal mobile health intervention
  – will engage you when you need it and will not intrude when you don’t need it.
  – will adjust to unanticipated life events
Setting

- Mobile health science studies involving clinical populations
- Machine learning/reinforcement learning tied closely to scientific inquiry in behavioral science
  - Experimental Design in ML/RL
  - Causal Inference in ML/RL
  - Interpretable ML/RL
Mobile Intervention Types

PUSH

PULL
Goal

Determine when and in which setting
- whether the mobile device/wearable should deliver a treatment push &
- which type of push to deliver.

Sequential decision making
- Inform behavior change science
- Development of treatment policy
Conceptual Data

• On each individual: $O_1, A_1, Y_2, \ldots, O_t, A_t, Y_{t+1}, \ldots$

• $t$: Decision point

• $O_t$: Observations at $t^{\text{th}}$ decision point (high dimensional)

• $A_t$: Treatment at $t^{\text{th}}$ decision point (pushes)

• $Y_{t+1}$: Proximal response (e.g., reward, utility)
Observations

• Commercial wearable wrist band (data each minute); Smartphone sensor data (6 times per day); Daily self-report

Pushes

• Activity planning for following day (each evening)
• In the moment tailored activity suggestions (5 pre-specified times per day)
Pushes

1) Types of treatments that can be provided at a decision point, $t$
2) Whether to provide a treatment
Observations

- Commercial wearable wrist band (data each minute); Smartphone sensor data (6 times per day); Daily self-report

Pushes

- 2x2 Factorial for morning greeting (each morning)
- Anti-sedentary message (at 5 min. intervals during day & only if sedentary over prior 30 minutes)
- Activity suggestions (5 pre-specified times per day)
- Evening greeting (each evening)
Observations

- Investigational wearable wrist and chest bands (data output < 1 second intervals); Self-report (6 times/day)

Pushes

- Reminder to utilize app directed stress-management exercises (every minute of 10 hour day & only if a stress classification is possible)
Availability

Treatments can only be delivered at a time $t$ if an individual is *available*.

Treatment effects at a decision point are conditional on availability.
Goal

Determine when and in which setting
• whether the mobile device/wearable should deliver a treatment push &
• which type of push to deliver.

Sequential decision making
• Inform behavior change science
• Development of treatment policy
Micro-Randomized Trial

On each participant and at each decision point, \( t \), randomize between treatment actions, \( a \)

Pre-specified algorithm for the randomization probability:

\[
P[A_t=a| H_t, I_t=1]
\]

- \( I_t=1 \) if available, \( I_t=0 \) if not
- \( H_t \) denotes data on participant through \( t \)
Some Challenges

- Experimental design → the formula for the randomization probabilities, $P[A_t = 1 | H_t, I_t = 1]$
  - All tuning parameters, entire trial protocol must be pre-specified prior to study

- The choice of proximal response, aka “reward.”

- Non-stationarity

- Need for multiple, interacting, treatment policies
Randomization

Heartsteps V1 activity suggestion
- Team decides to provide an average of 3 tailored activity suggestions per day (5 opportunities per day)
- Binary $a=0,1$

$$P[A_t=1| H_t, I_t=1]=.6$$
Randomization

Heartsteps V2 anti-sedentary message

– Randomize only if user has been sedentary for ≥ 30 min. Team decides approximately 1.5 messages per day.

– Using Heartsteps V1 data, built algorithm to predict, at each time point, the mean and variance of the number of remaining available, sedentary, decision times in day.

– Randomization probabilities use these predictions.
Randomization

Sense²Stop reminder

– Team decides approximately 1.5 reminders per day when currently stressed & 1.5 reminders per day when currently not-stressed.

– Using data from another smoking study (with no intervention), built algorithm based on a simple Markovian model to predict, at each time point, number of remaining available stressed and non-stressed episodes in day.

– Randomization probabilities use this prediction.
Randomization

Heartsteps V2 tailored activity suggestion

- Use a “Thompson Sampling Contextual Bandit” algorithm to randomize at each of 5 decision times per day.
- Thompson Sampling prior is based on Heartsteps V1 data.

Assess feasibility of algorithm
Randomization

Enhance Feasibility of Contextual Bandit Algorithm:

– Randomization probabilities from Thompson Sampling are clipped between .10 and .80

– How to select the proximal response (reward)—not just stepcount….
Proximal Response (Reward)

For analyses conducted after study ends:

Each type of push designed to operate on a different time scale

• Heartsteps activity suggestion
  – Stepcount over 30 min. following randomization

• Heartsteps V2 anti-sedentary message
  – Stepcount or heartrate over next 5 (?) min. following randomization
Proximal Response (Reward)

In different settings push is expected to operate on different time scales

- Sense²Stop reminder
  - % time stressed over subsequent hour if currently stressed
  - % time stressed over subsequent 4 hours if currently not stressed
Non-stationarity: Heartsteps V1

On each of $n=37$ participants:

- Tailored activity suggestion
  - Provide a suggestion with probability 0.6
  - Do nothing with probability 0.4

- 5 times per day * 42 days = 210 randomizations per participant
The data indicates that there is a causal effect of the activity suggestion vs no activity suggestion on step count in the succeeding 30 minutes.

- This effect deteriorates with time
- The walking activity suggestion initially increases step count over succeeding 30 minutes by ≈ 171 steps but by day 21 this increase is only ≈ 35 steps.
Non-stationarity; Heartsteps V1

The deteriorating effect of the walking activity suggestion on the subsequent 30 min. stepcount may be due to

- Habituation
- Burden

(e.g. unobserved/poorly observed variables)
Multiple Treatment Policies

Multiple causal pathways $\rightarrow$ need to learn multiple interacting treatment policies

- Treatment pushes and responses at weekly, daily, hourly, minute level time scales
- Engagement pushes
Last Comments

• Reinforcement learning involves learning causal inferences

• Randomization enables causal inferences based on minimal structural assumptions

• Theory for tracking in reinforcement learning
Collaborators!
Conceptual Models

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \beta_0 A_t \]
\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \alpha_2 d_t + \beta_0 A_t + \beta_1 A_t d_t \]

- \( t=1, \ldots, T=210 \)
- \( Y_{t+1} \) = log-transformed step count in the 30 minutes after the \( t^{th} \) decision point,
- \( A_t = 1 \) if an activity suggestion is delivered at the \( t^{th} \) decision point; \( A_t = 0 \), otherwise,
- \( Z_t \) = log-transformed step count in the 30 minutes prior to the \( t^{th} \) decision point,
- \( d_t \) = days in study; takes values in \((0,1,\ldots,41)\)
Pilot Study Analysis

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \beta_0 A_t, \text{ and} \]

\[ Y_{t+1} \sim \alpha_0 + \alpha_1 Z_t + \alpha_2 d_t + \beta_0 A_t + \beta_1 A_t d_t \]

<table>
<thead>
<tr>
<th>Causal Effect Term</th>
<th>Estimate</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 A_t ) (effect of an activity suggestion)</td>
<td>( \hat{\beta}_0 = .13 )</td>
<td>(-0.01, 0.27)</td>
<td>.06</td>
</tr>
<tr>
<td>( \beta_0 A_t + \beta_1 A_t d_t ) (time trend in effect of an activity suggestion)</td>
<td>( \hat{\beta}_0 = .51 )</td>
<td>(.20, .81)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta}_1 = -.02 )</td>
<td>(-.03, -.01)</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>
Goal: Develop an mobile activity coach for individuals who have coronary artery disease

Three iterative studies:

- 42 day micro-randomized pilot study with sedentary individuals,
- 90 day micro-randomized & personalized study,
- 365 day personalized study
Continually Learning Mobile Health Intervention

1) Trial Designs: Are there effects of the actions on the proximal response? *experimental design*

2) Data Analytics for use with trial data: Do effects vary by the user’s internal/external context? Are there delayed effects of the actions? *causal inference*

3) Learning Algorithms for use with trial data: Construct a “warm-start” treatment policy. *batch Reinforcement Learning*

4) Online Algorithms that personalize and continually update the mHealth Intervention. *online Reinforcement Learning*