Driving Action with Machine Learning

Nasir Bhanpuri, PhD
Online Type 2 Diabetes Reversal Clinic

Coach continually monitors data

PATIENT

HEALTH COACH
Provide daily input and support

BIOMARKER TRACKING

RESOURCES
Recipes, videos, guides

PATIENT COMMUNITY

PHYSICIAN
Regular telemed appointments
Early in Journey

- Preparation & Core Concepts
- Dietary Changes Start
- Metabolic Changes & Personalized Guidance

~ Day 5

Day 14
High retention, but can improve

94% Retained at 90 Days (6% Dropout)

82% Retained at 1 Year

Source: IU Health Arnett - Virta Clinical Trial Data (Hallberg et al., 2017)
N = 158
*Those not “retained” either requested to terminate Virta services (usually because of unrelated health/family issues or undisclosed personal choice) or were removed from the study due to noncompliance and concerns related to safety.
Machine Learning to Drive Clinical Action and Decision-Making

- **Goal:** Increase long-term retention rate of patients
- **Questions:** Who is at risk of dropping out? *Why* are they dropping out?
- **Users:** Clinicians
- **Model Requirements:**
  1. Easy to communicate to clinicians
  2. Accuracy
- **Model Benefits:** Prioritization and insight into underlying factors
“All models are wrong, but some are useful”

George Box, Statistician
Feature Selection & Algorithm Selection

Based on Clinical Data and Research
- 108 Features (Characteristics)
- First 14 days of data
- Feature selection for Top 15
  - Random Forest Out-of-bag error
- Logistic regression

Text Messages
- Length
- Count/Freq
- Topic
- Urgency

App Data
- Weight
- Glucose
- Symptoms
- Medications

Dedicated health coach

App and biomarker tracking tools
## Model Results: Feature Directionality

<table>
<thead>
<tr>
<th>Increase Dropout Risk</th>
<th>Decrease Dropout Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to dietary change</td>
<td>Age</td>
</tr>
<tr>
<td>Texts about discomfort</td>
<td>Texts about challenges</td>
</tr>
<tr>
<td>Fatigue</td>
<td>Urgent texts</td>
</tr>
<tr>
<td>Opiate/Pain meds</td>
<td>Weight loss</td>
</tr>
</tbody>
</table>
Model Results: Dropout Risk

Distribution of Dropout Risk

- Average: 18%
- 50th percentile: 16%
- 90th percentile: 37%
## Model Validation

### Risk Level vs. Dropout Risk

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>Dropout Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0 - 10%</td>
</tr>
<tr>
<td>Medium</td>
<td>10 - 25%</td>
</tr>
<tr>
<td>High</td>
<td>25 - 45%</td>
</tr>
<tr>
<td>Very High</td>
<td>45 - 70%</td>
</tr>
<tr>
<td>Extremely High</td>
<td>70 - 100%</td>
</tr>
</tbody>
</table>

Anecdotal confirmation from coaches.
# Actionable Insights

## Who?

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Dropout Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.9 %</td>
</tr>
<tr>
<td>2</td>
<td>30.4 %</td>
</tr>
<tr>
<td>3</td>
<td>27.4 %</td>
</tr>
</tbody>
</table>

## Who? Why? (What to do?)

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Dropout Risk</th>
<th>Text Discomfort</th>
<th>Fatigue Count</th>
<th>Weight Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>High ~35%</td>
<td>9</td>
<td>5</td>
<td>-4.3</td>
</tr>
<tr>
<td>3</td>
<td>High ~35%</td>
<td>14.5</td>
<td>0</td>
<td>-4.0</td>
</tr>
<tr>
<td>1</td>
<td>High ~35%</td>
<td>9</td>
<td>2</td>
<td>-1.0</td>
</tr>
</tbody>
</table>
Impact

- Coach impressions:
  - Prioritize additional outreach
  - Focus efforts
- Dropout rate down 66%
- Important caveats
  - Different population
  - Evolving product
Key Takeaways

- Interpretable ML is critical to creating actionable insights
- Collaborate with users (clinicians) early and often during model development and delivery
- High frequency data + ML + high touch model = tight feedback loop
- Why we are really here