Predictors of Attrition in a Public Online Interpretation Training Program for Anxiety

Diheng Zhang¹, Yuling Shi², Hongning Wang¹, Bethany A. Teachman¹
¹University of Virginia, ²Wuhan University

Introduction

- Approximately 40 million people suffer from anxiety disorders in any given year (Kessler et al., 2005).
- Despite the high prevalence, less than half of those suffering ever seek treatment (Wang et al., 2005).
- Stigma against seeking mental health services, fear of coming to treatment, and limited access to traditional treatment options all contribute to the low rates of help-seeking. Online interventions are a promising, low-cost approach to reach people who might not seek out or have access to traditional mental health services.
- However, many online interventions have very high attrition (Mathieu et al., 2012), limiting the impact of these potentially helpful approaches. It is thus critical to learn more about when and why participants drop out of online interventions.

Aim:
- In our current study, we used a machine learning approach to utilize MindTrails participants’ early responses and online behavior markers to predict participant dropout on an individual level.

Methods

- **Participants:**
  - A subset of the MindTrails sample from May 5th 2016 to Nov 10th 2017 (N=409)
  - Gender: Female: 73.6%
  - Age: Mean = 35.07, SD = 13.83
  - Race & ethnicity:
    - Hispanic: 7.3%;
    - White: 77.5%;
    - African: 3.4%;
    - Asian: 8.1%;
    - Other: 11.0%

- **Procedure:**
  - Train: Randomly assigned 80% of the sample into an algorithm training group (N= 327) to develop the prediction algorithm.
  - Validate: Assigned the rest to an algorithm testing group (N=82) to validate it.
  - Repeat: Repeated this process five times to stabilize our model (termed 5-Fold cross validation).

- **Tested algorithm:**
  - Support Vector Machine (SVM)

Included Features:

- Responses to questionnaires(scales):
  - The Overall Anxiety Severity and Impairment Scale (OASIS) (Norman et al., 2006)
  - The Quality of Life scale (QOL) (Burckhardt, 2003)
  - The depression sub-scale of Depression, Anxiety and Stress Scales (DASS-21) (Lovibond & Lovibond, 1995)
  - Daily Drinking Questionnaire (Collins et al., 1985)

- Online behavior markers:
  - Mean (M) and standard deviation (SD) of time spent on questionnaires and trials in missing-letter-filling task (see below for illustration)
  - Error rate on tasks trials
  - Time gap between questionnaires and tasks

- Demographic variables: Education level & Income

Tasks

- **Missing-letter-filling and comprehension question (CQ) example:**
  - The ELEVATOR: You are in the lobby of your new apartment building. You press the button for the elevator to go up.
  - As you get on the elevator you think about it.
  - Did you think about the elevator’s safety?

  - Please Type Yes or No.

- **Support Vector Machine:**
  - Finding the hyperplane separating the target groups that maximize the margin of support vectors

SVM

- **Input space**
  - **Feature space**

Results

- **Predictive performance of SVM**
  - **Feature space**
  - **Linear space**

- **Top features predicting dropout (0/-)**
  - w (weight) of each feature
  - Top features predicting stay (1/+)
  - w (weight) of each feature

- **Mean**
  - **SD**

- **Accuracy**
  - **Mean**
  - **SD**

- **Top features predicting dropout (0/-)**
  - Average time spent (ATS) on scales
  - Errors (count) on tasks
  - ATS on missing-letter-filling trials
  - ATS on the second guess for CQ
  - SD of time spent on correcting Missing-letter-filling trials
  - Time spent on daily drinking habit questionnaire

- **Feature space**
  - **Input space**

- **Mean**
  - **SD**

- **OASIS**
  - **QOL**

- **0.021**
  - **0.021**

- **0.026**
  - **0.02**

- **0.011**
  - **0.026**

- **0.105**
  - **0.078**

- **0.078**
  - **0.033**

- **0.074**
  - **0.078**

- **0.012**
  - **0.011**

- **0.012**
  - **0.011**

- **0.009**
  - **0.009**

- **0.007**
  - **0.007**

- **0.021**
  - **0.021**

Discussion

- This study suggests that, with reasonably-sized samples of early responses and online behavior markers, machine learning algorithms hold considerable promise to predict individual level dropout behavior based on algorithms trained at a group level.
- This could help researchers and clinicians identify those at high risk of dropout early on so they could try to tailor interventions accordingly to minimize dropout and help people get a full dose of treatment.
- This work can also provide helpful insights about how to improve online interventions for anxiety that encourage participant retention.

Limitations & Future Directions

- Only participants who finished the full initial assessment were included in this analysis. Attrition during the initial assessment was not accounted for in the analysis.
- The sample from the control condition was excluded because data from one of the baseline tasks was missing due to an administration error.
- Future research should replicate the model in other online studies to see if the same pattern remains for other online interventions or populations.
- Future online intervention studies could implement SVM-based algorithms into online web apps, to train the model with live data and send out notification or intervention when participants are at risk of dropping out.

This poster was supported by the NIH R01MH113752 and NIH R34MH106770 grants awarded to B. Teachman.