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

Diheng Zhang[1], Yuling Shi[2], Hongning Wang[1], Bethany A. Teachman[1]
[1]University of Virginia, [2]Wuhan 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.
- MindTrails is an online study platform designed to test interpretation training (a form of cognitive bias modification) programs for anxiety and related rigid, negative thinking patterns.

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).

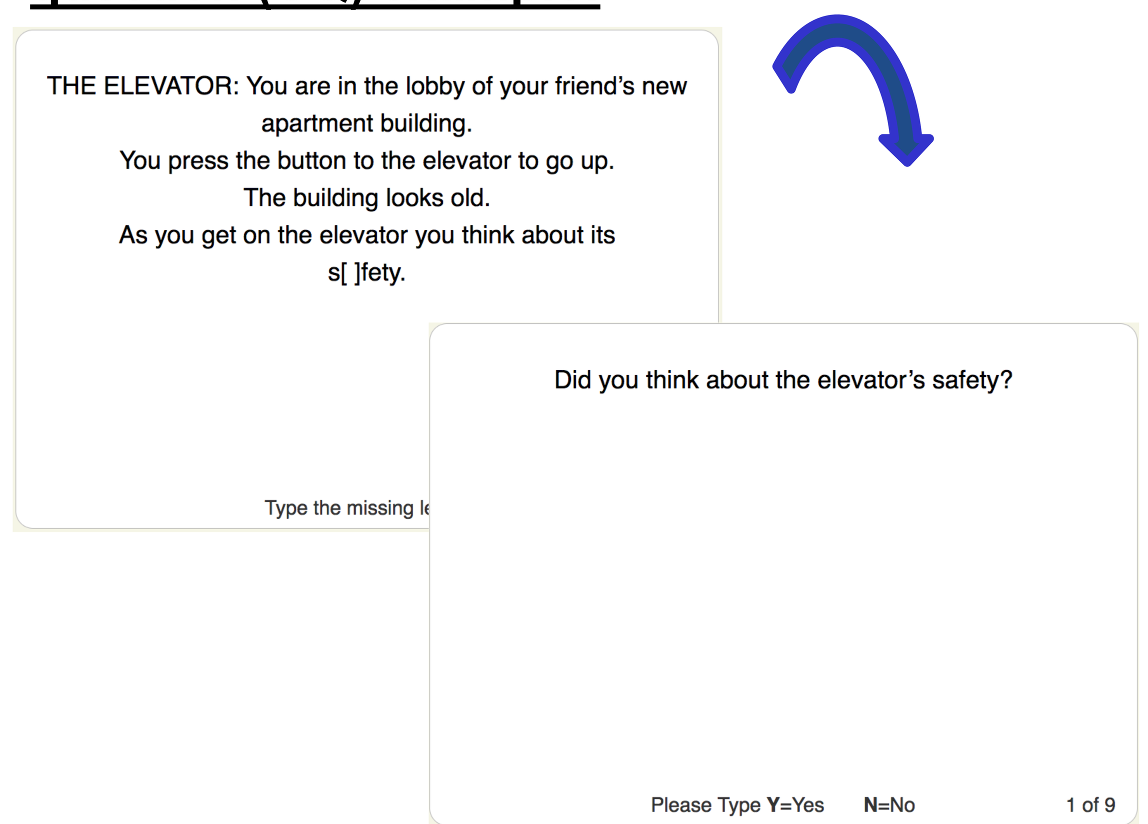
Tested algorithm:
- Support Vector Machine (SVM)

Methods (Cont.)

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

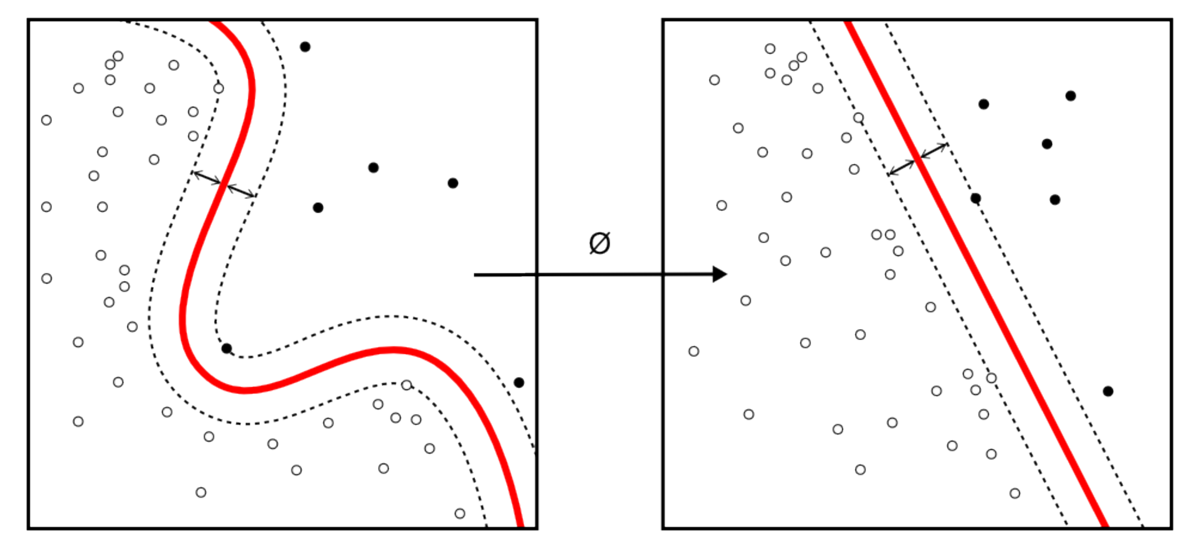
Tasks

Missing-letter-filling and comprehension question (CQ) example:



SVM

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



Results

Performance of SVM

Predict dropout of	Sample size	Number of features	F1 score		Accuracy	
			Mean	SD	Mean	SD
Session 1	409	28	0.835	0.021	0.821	0.021

Weight of features (selected)

Top features predicting dropout (0/-)	w (weight) of each feature		Top features predicting stay (1/+)	w (weight) of each feature	
	Mean	SD		Mean	SD
Average time spent (ATS) on scales	-0.338	0.183	OASIS	0.266	0.02
Errors (count) on tasks	-0.333	0.208	SD of time spent on CQ	0.151	0.149
ATS on missing-letter-filling trials	-0.236	0.162	QOL	0.123	0.036
ATS on the second guess for CQ	-0.159	0.029	Time spent on DASS21_DS	0.112	0.266
SD of time spent on correcting Missing-letter-filling trials	-0.109	0.046	SD of time spent on the second guess for CQ	0.105	0.078
Time spent on daily drinking habit questionnaire	-0.075	0.009	Time spent on Anxiety Triggers questionnaire	0.074	0.023

Results (Cont.)

- The algorithm was able to predict participants' dropout at an accuracy rate of 82.10% on average (variance = 2.1%) based on their responses from the baseline assessment.
- Online behavioral markers were stronger predictors of dropping out after the baseline assessment (before they had started interpretation training) than participants' scores on most questionnaires (beside OASIS).
- Specifically, higher error rates, higher standard deviations, and longer mean time to correctly respond on the missing-letter-filling task all predicted higher likelihood of dropout before or during the following session.
- Demographic variables were relatively weak predictors (feature weights < 0.04)

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 implant 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.

The MindTrails PROJECT

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