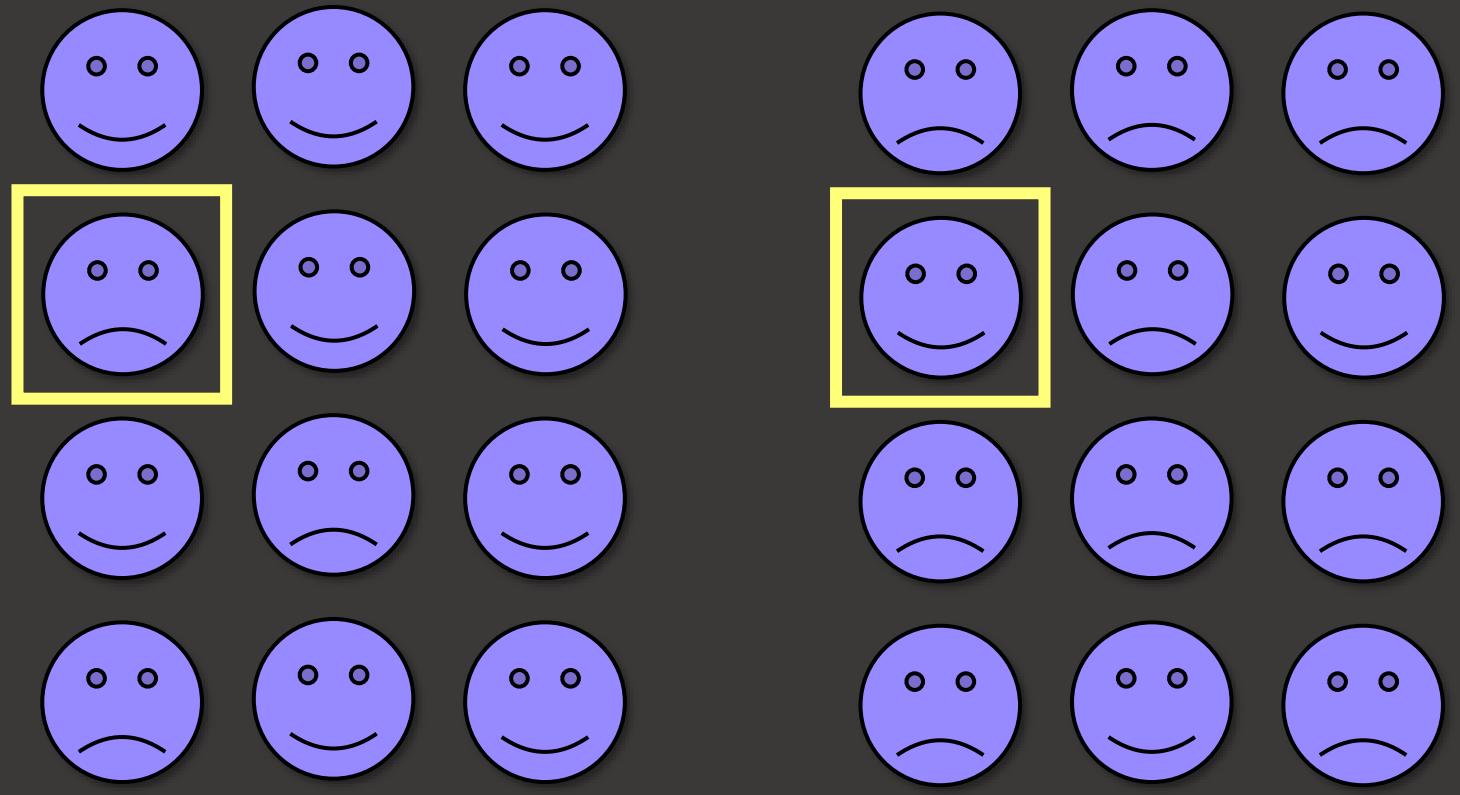
Using HTC to develop precision mental health algorithms

Gaylen Fronk, Hannah Moshontz, & John Curtin
University of Wisconsin-Madison
HTCondor Week
May 27, 2021

@gaylenfronk

Background: Traditional Treatment Assignment

75% Effective



25% Effective

Background: Precision Mental Health

The application of precision medicine to mental health conditions

Goal: to use individual differences to select the treatment with the highest predicted treatment efficacy for a given patient

1. **Increase likelihood of treatment success within an individual**

2. **Improve treatment effectiveness rates across the population**

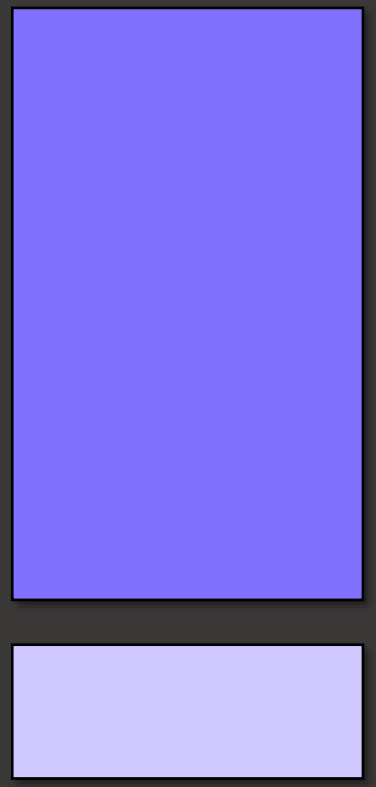
Background: Machine Learning

- Offers a path forward for precision mental health
- Maps onto complexity of people and clinical phenomena
- Prioritizes out-of-sample prediction

Model Fitting

Models include:

- **Treatment condition**
- **~400 individual difference predictors**
- **Outcome (treatment success at 6 months)**



Data used for model fitting

Data used for model evaluation

Model Fitting

- **Model configurations include:**
  - **Statistical algorithms (e.g., elastic net logistic regression, random forest, k-nearest neighbors, neural networks)**
  - **Algorithm-specific characteristics (e.g., hyperparameters, number of hidden layers)**
  - **Feature sets & feature engineering decisions**

Model Fitting

- Example:
  - Statistical algorithm: random forest
  - Algorithm-specific characteristics: 3 hyperparameters with varying levels, total of 100 combinations
  - Feature sets & feature engineering decisions: models with self-report items or self-report scales

- Total: 200 models to fit

Jobs Setup

- To maximize parallelism of CHTC, we break down model fitting into the smallest jobs possible
  - Single combination of statistical algorithm, algorithm characteristics, and feature characteristics run as one job
- Jobs are run across CHTC and Open Science Grid machines
- We aggregate jobs locally to select the model configuration that performed best across cross-validated iterations

Computing Time

- Each model takes anywhere from 1 – 5 minutes to run
- Even running in parallel locally, it would take days to fit single scenarios
  - This is ~650 computing hours just for the previous example scenario!
- Computing time multiplied by “intended” scenarios as well as testing and iteration

Using CHTC & HTCondor

- **Most important feature is the CHTC support team!**
- **Also helpful are clear documentation and walkthroughs**
- **Have also benefitted from within-lab collaboration across multiple projects using CHTC**

Impact

- Precision mental health research and algorithm building will only be possible with this kind of computing power
- Allows me to maximize time, resources, and person power
- Allows me to expand my research and funding opportunities
  - Incorporating genetics
  - UW Seed Grant
  - Fellowship application


Addiction
Research Center
UNIVERSITY OF WISCONSIN – MADISON

Thank you!