Classical and Quantum Machine Learning Applied To Predicting Placebo Response For Clinical Trials In Bipolar Disorder: Recent Results

Joseph Geraci[1,2], Bernard Wong[1], Jalal Ziauddin[1], Siddhant Jain[1,3], Paul Leonczyk[1], Kim I. Bishop[4]

[1]NetraMark Corp, Toronto, Ontario, Canada; [2] Department of Molecular Medicine, Queen's University, Kingston, Ontario, Canada;
[3]University of Toronto, Ontario, Canada; [4]Global Pharma Consultancy, LLC, PA, USA

Question Being Explored

Can Machine Learning be used to predict placebo response in a consistent way for Central Nervous System (CNS) clinical trials?

Introduction

Placebo response for CNS clinical trials especially with mood disorders, namely Major Depressive Disorder (MDD) and Bipolar Disorder, is a cause for concern. Placebo responders can easily obfuscate potentially positive results and a method to correct this effect is a desirable technology that is currently missing. Here we utilized modern machine learning methods to create a model for *Placebo Non-Response* that can be used to demonstrate efficacy, even in the presence of a large number of placebo responders.

Methods

We obtained data from a failed trial meant to evaluate depression in Bipolar Disorder from a Pharmaceutical Company where all patient identification has been removed. Our relationship is collaborative on this project for now. The data was reshaped into a structure suitable for our machine learning platform. The data contained 64 placebo patients and approximately 150 variables consisting of clinical data in addition to clinical rating scale data including the Hamilton Rating Scale for Depression (HAM-D) and Montgomery–Åsberg Depression Rating Scale (MADRS). Despite this being a small sample size, our unique technology is capable of learning from data of this size. Validation for this has been provided by hundreds of tests but also with a validation we performed on a failed drug arm where the response rate was very poor. We will report on the result of this test here. Three methods were employed:

1. This placebo cohort data was passed through a proprietary expansion software program that created thousands of artificial "placebo arm patients" based on the distribution and values of the actual patients given placebo. This large set was used to train thousands of Random Forest + XGBoost + Deep Neural Network models, out of which a "best model" was tested on a data set it has never seen before.
2. A proprietary machine learning method called DeepCrush was used to learn from this data set because of its ability to learn from small data sets. This method created 12 models out of which a "best model" was used to test on a data set it has never seen before.
3. We ran our data through the D-Wave Quantum Computer and utilized the Quantum Boltzmann Machine approach.

The model in number 2 was queried and asked to explain what variables were driving the resulting clustering of patients. Due to the fact that the method used is not a black-box it was able to explain itself when queried. All the variables used to produce this model were items that evaluated attitudes towards the trial, including *excitement* and *belief in the protocol.* Other factors that were involved measured *forgetfulness and sleep quality*. We have created similar models for MDD patients utilizing a mixture of items from the HAM-D 17, MADRS, and Beck Depression Inventory (BDI) scales with other pharmaceutical groups.

We validated the model by testing it on a test set of 219 Bipolar I patients who were given a drug for depressive episodes. The drug was deemed ineffective and thus we asked the machine to predict non-responders.

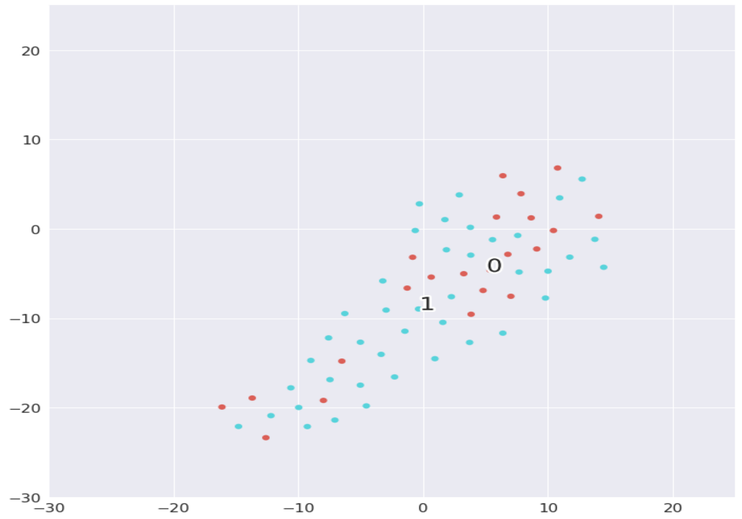

Figure 1. This graph is a two dimensional representation provided by t-SNE of a model using a machine learning algorithm based on boosted trees, which is very powerful and performed best out of all the classical models we used. As you can see this result does demonstrate that the machine was seeing some trend but this model ***would not be useful in a clinical trial*** due to a lack of separation.

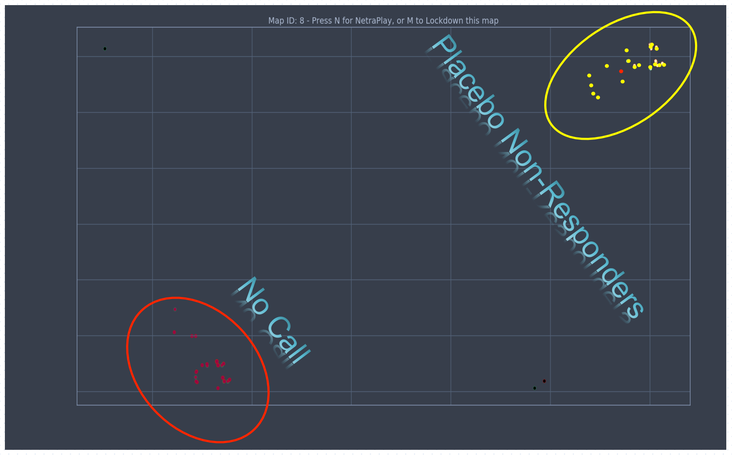

Figure 2. This is a representation of what NetraAI sees with the same data set. Two groups: the subpopulation captured by the yellow oval are all non-responders (except one) and the red cluster at the bottom are a mixed group. The machine is able to extract people who are explainable and reject others who are not.

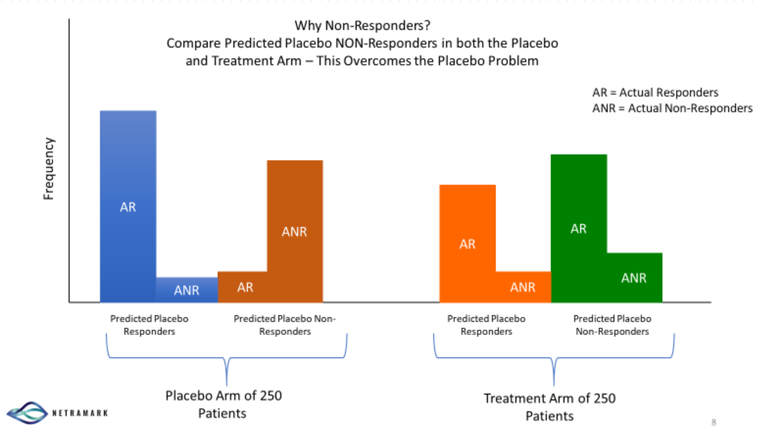

Figure 3. This image conveys the utility of using this technology to evaluate the efficacy of a drug by allowing the machine to determine **the tendency to not respond to placebo** in both drug and placebo arms. By comparing results one can make powerful inferences.

Results

We discovered a model (Figure 2) that is excellent at predicting placebo non-responders. In fact, if the machine claims that someone is a non-responder it will be correct 94% of the time. The fascinating thing about this result however are the variables chosen and they corroborate common sense regarding placebo response. In brief, there are a set of psychological attitudes which come together in a non-linear way to characterize a type of patient that will not respond to placebo, e.g., the desire to be part of the trial. No one variable alone is capable of having this efficacy.

As mentioned above, the model was tested on a failed cohort of 219 patients where the mechanism of action was deemed ineffective who happened to have the correct variables captured. ***The machine selected 55 people for whom it was very confidant should not respond and indeed 89% of those individuals were in fact non-responders.*** Less impressively, the machine was correct at predicting responders to this drug 71% of the time.

Introduction to Quantum Machine Learning

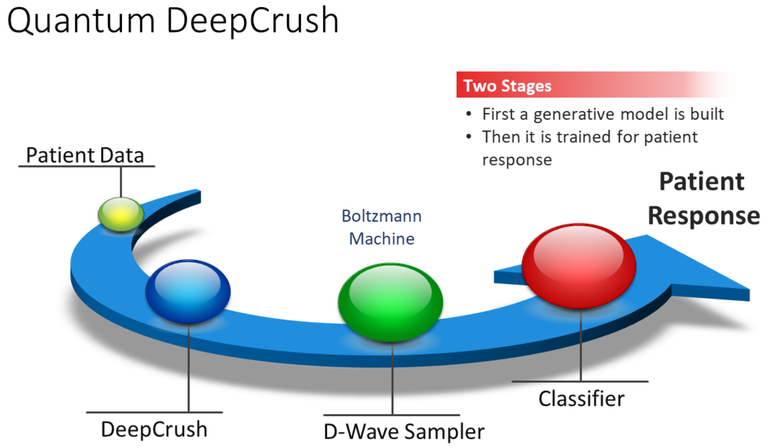

Figure 4. In conjunction with Queen's University, the University of Toronto, Creative Destruction Labs, and D-Wave, we have begun to explore the utility of using *near term* Quantum Computers to predict patient response. What we have discovered is that there is a way to utilize the power of these machine to augment modern machine learning algorithms by helping with the most difficult classical steps.

Conclusions

We have preliminary results that suggest it is possible to predict placebo response in mood disorder cohorts (bipolar disorder and depression), and especially placebo non-response, with a set of simple attitudinal based measures that we have identified as working together in a non-linear way. We have developed models that can predict placebo response as well by using items from clinical scales and gene expression but our current collaboration has allowed us to initiate a validation of the model shown in figure 2. This particular non-response model seems to be very well suited for CNS clinical trials. Cutting edge machine learning models that deviate from the typical tree and neural network based models are showing promise with the type of "small" and heterogeneous data found in clinical trials. Quantum ML will begin to have an impact in a few years.

Acknowledgements

We are grateful to the Creative Destruction Labs at the University of Toronto for this opportunity and the D-Wave team for their time, instruction, and access to their hardware. We are also grateful to the Ontario Brain Institute for their financial support. Special thanks to NetraPharma for the opportunity to work on this project.