PARKINSON'S TEST RECOMMENDATION ENGINE

Elder Research developed predictive models for the Michael J. Fox Foundation to score and rank medical tests and biomarkers to find those that offer the greatest value in predicting Parkinson's disease.

INDUSTRY

» Biotechnology & Healthcare

BUSINESS NEED

» Determine which medical test, or combination of tests, best predicts Parkinson's disease

SOLUTION

» Identified patient clusters based on their available test results

» Built and validated models that identified biomarkers important to predicting Parkinson's disease

BENEFIT

» Enables doctors to make more effective treatment decisions

» Optimizes patient testing requirements to improve disease diagnosis

» Our models identified several biomarkers that are important in predicting Parkinson's disease

The Challenge

Recent research supported by the Michael J. Fox Foundation (MJFF) (and other benefactors) collected multifaceted data sets from patients with Parkinson's Disease (PD) that can include:

- Clinical assessments
- Bio-samples (analytes from sources like plasma, cerebral spinal fluid, and saliva)
- Genomics
- Imaging data

Research programs, treatment clinics, and physician's offices vary in the types of data and medical test results they collect on PD patients. The BioFIND observational clinic study standardizes the longitudinal collection of biospecimens. It contains results from nearly 2,500 biomarker and other medical tests, but none of its approximately 200 participants have results for all of the tests. The number of completed test results for a given participant range from 119 to over 2,450, but more than a third of subjects have results for fewer than 300 tests. Despite the extreme variability in patient data, the client was very eager to determine: "Which tests offer the greatest value in disease prediction?" And, "Could the importance of a key test be affected when data from further medical tests becomes available?"

The Solution

Data for the project was provided through the MJFF and the BioFIND observational clinic study, which was designed to discover and verify biomarkers of Parkinson's disease. A proprietary clustering method was used to identify twelve clusters of patients based on the test results available for each patient.

Prior to model selection, we used two feature selection algorithms to produce two lists of medical tests for each cluster. A random forest algorithm (called Boruta) was used to produce an all relevant tests list. Potentially useful for scientific research, this list included all tests the algorithm found useful for disease prediction. We then used a recursive elimination algorithm to generate a minimal test list, representing the smallest set of tests needed for accurate disease prediction. Each

ELDER RESEARCH
— DATA SCIENCE · AI · MACHINE LEARNING —

Office Locations

Charlottesville, VA
Washington, DC
Baltimore, MD
Raleigh, NC
London, UK

cluster was further divided into ten subgroups, and algorithms were trained on each set of 90% and validated on the remaining, held-out 10%. This 10-fold cross-validation allowed us to:

1. Accurately score how well our algorithms performed

2. Determine how often a test was chosen within each cluster, as well as between clusters, enabling us to rank the importance of each medical test for disease prediction.

Our models yielded better accuracy and specificity metrics than those obtained from target shuffling nearly 100% of the time for 10 of the 12 clusters. This comes at very little cost to sensitivity in clusters 1 through 10, which had performance above 95% in six of the ten clusters and above 84% among the remaining four clusters. We observed reduced performance in clusters 11 and 12, where very few patient observations or available tests provided limited data for model training.
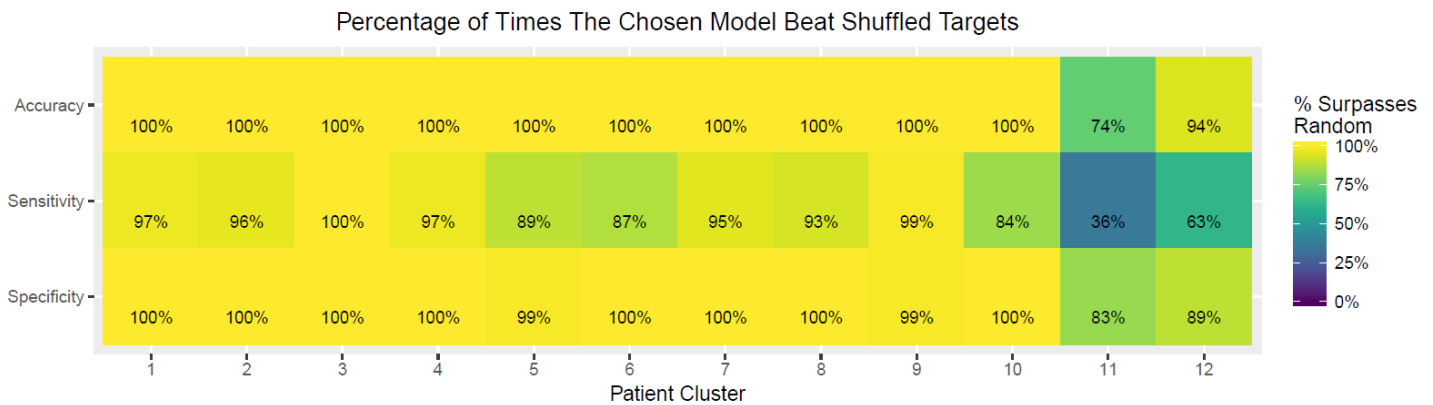

Figure 1. Model results from target shuffling. The percentages indicate how often the each model outperformed the random (shuffled) data sets.

Several models were trained on each cluster and the most accurate model was selected. Target shuffling was then used to further validate the statistical accuracy of our model results.

Model ensembles outperformed single models in 6 of the 12 clusters. Extreme Gradient Boosting (XGBoost) achieved superior performance in 4 of 12 clusters, while the C5.0 decision tree and feed-forward neural net each won for a single cluster. By target shuffling over 300 times, we confidently estimated the performance of our models (Figure 1).

Our models identified several biomarkers that are important in predicting Parkinson's disease. For example, chemical_id_100004634 (cerebrospinal fluid origin) was determined to be critical for accurate prediction more than 80% of the time in clusters where the test result was available. Plasma_unknown_1 was determined to be critical 100% of the time in all but one cluster in which it was available. In the remaining cluster it was found to be critical only 50% of the time. Initial analysis suggests that a combination of other tests in this cluster may prove more valuable in PD prediction, suggesting that indeed the importance of a key test can be affected when data from other medical tests are available.

Results

Deployed in a clinic, the Elder Research solution will help clinicians diagnose Parkinson's based on available tests, and recommend the fewest additional (or next best) tests to improve disease prediction. The analytical processes summarized here are applicable to many other classification targets, beyond the prediction of Parkinson's disease.

They could be applied to predict the occurrence of other diseases, the speed of disease progression, the effectiveness of treatments, and indicate dominant or worsening symptoms, enabling doctors to improve treatment decisions.