# MULTIPLE SCLEROSIS: FROM PREDICTIONS TO SENSITIVITY ANALYSIS AND ROBUST MODELS

SUZANA IACOB AND EL GHALI AHMED ZERHOUNI

1. PROBLEM STATEMENT

- **Multiple Sclerosis (MS)** - disease of the nervous system affecting 1m US patients
- Not curable - treatments used to manage symptoms
- **Data**: 10k observations of MS patient utility data (e.g. risk of cardiac arrest)
- **3 treatment options**: High Dose* (1), Low Dose* (2), or No Treatment (3) (* of methylprednisolone)
- **Problem**: No personalized treatment, High Dose treatment prescribed for all patients
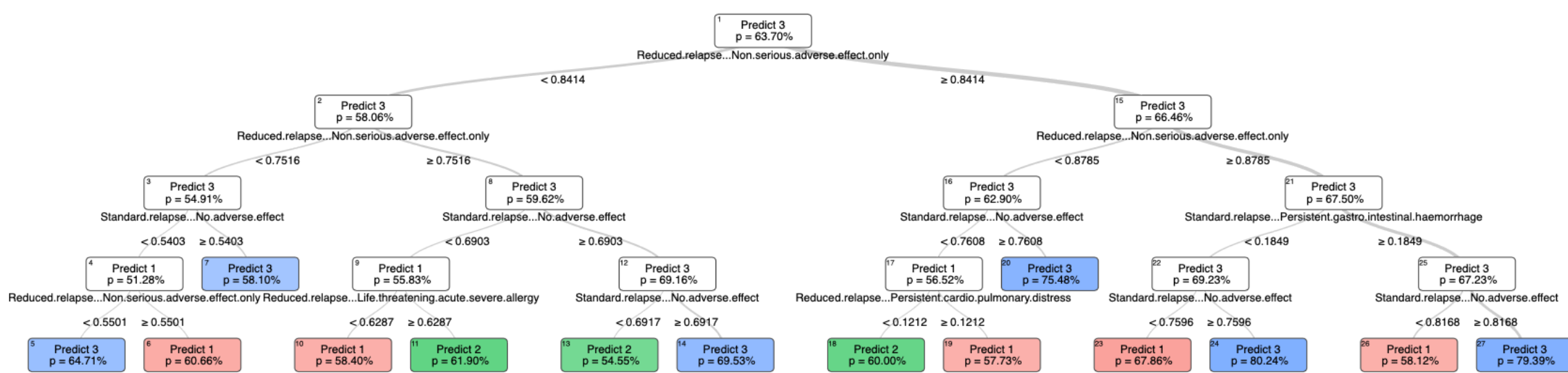
## 2. MODELS FOR BEST TREATMENT

Prescription Models - Unsupervized KNN
- Prescribe most common treatment of the cluster

$$\max_{z_p \in \{1,2,3\}} \sum_{p=1}^{T} \sum_{i=1}^{N} w_{pi} 1_{z_p = y_i}$$
$$\text{s.t.} \quad w_{pi} = 1 \quad \text{if } p \text{ and } i \text{ are in the same cluster}$$

Prediction Models - Trees
- Accuracy scores: CART (70%), OCT(69%), OCT-H (70%)
- Feature importance: No adverse effect, Diabetes
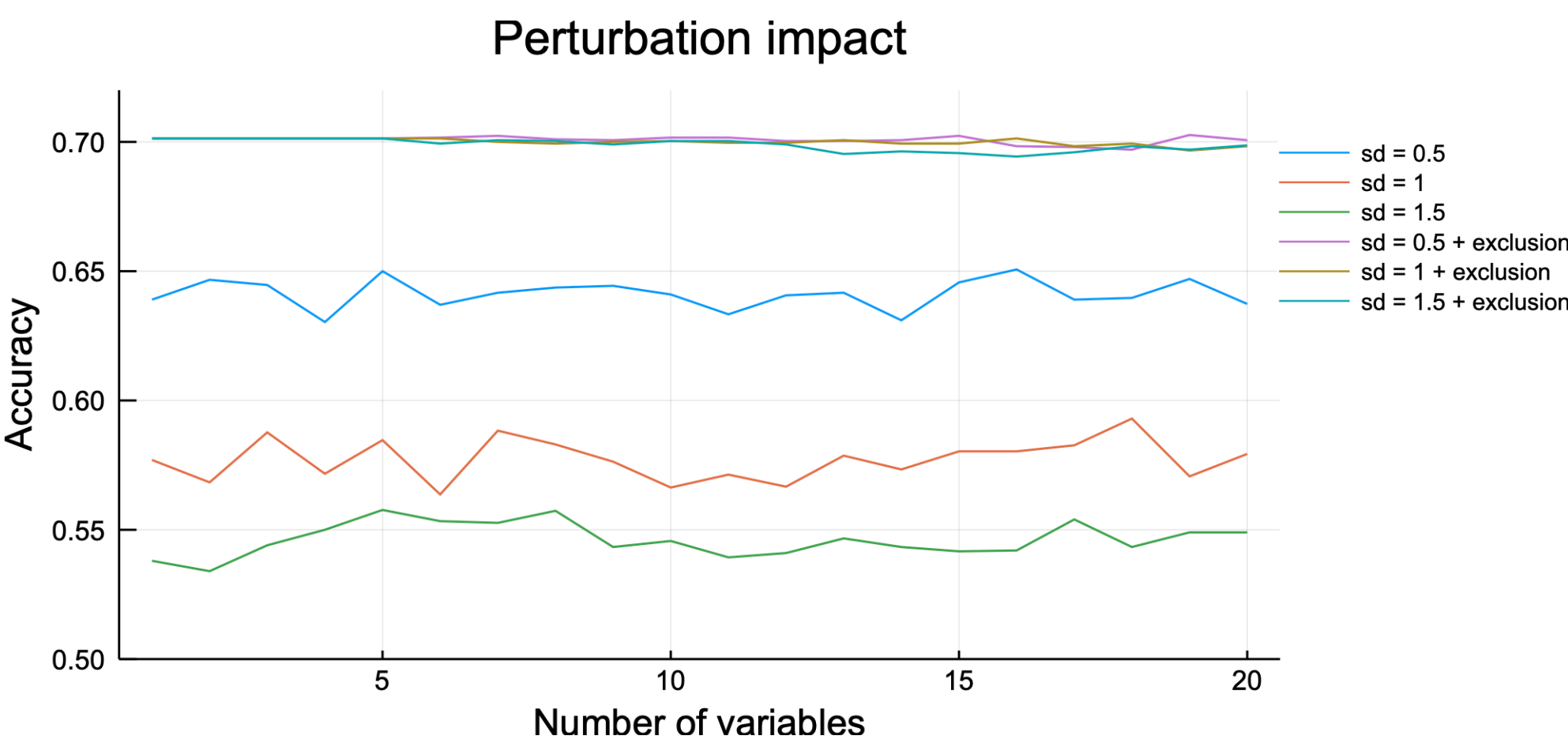- Models rarely predict Low Dose treatment because of unbalanced dataset

Accuracy evolution with perturbation

Legend: OCT - depth 5, OCT - depth 7, OCT - depth 10
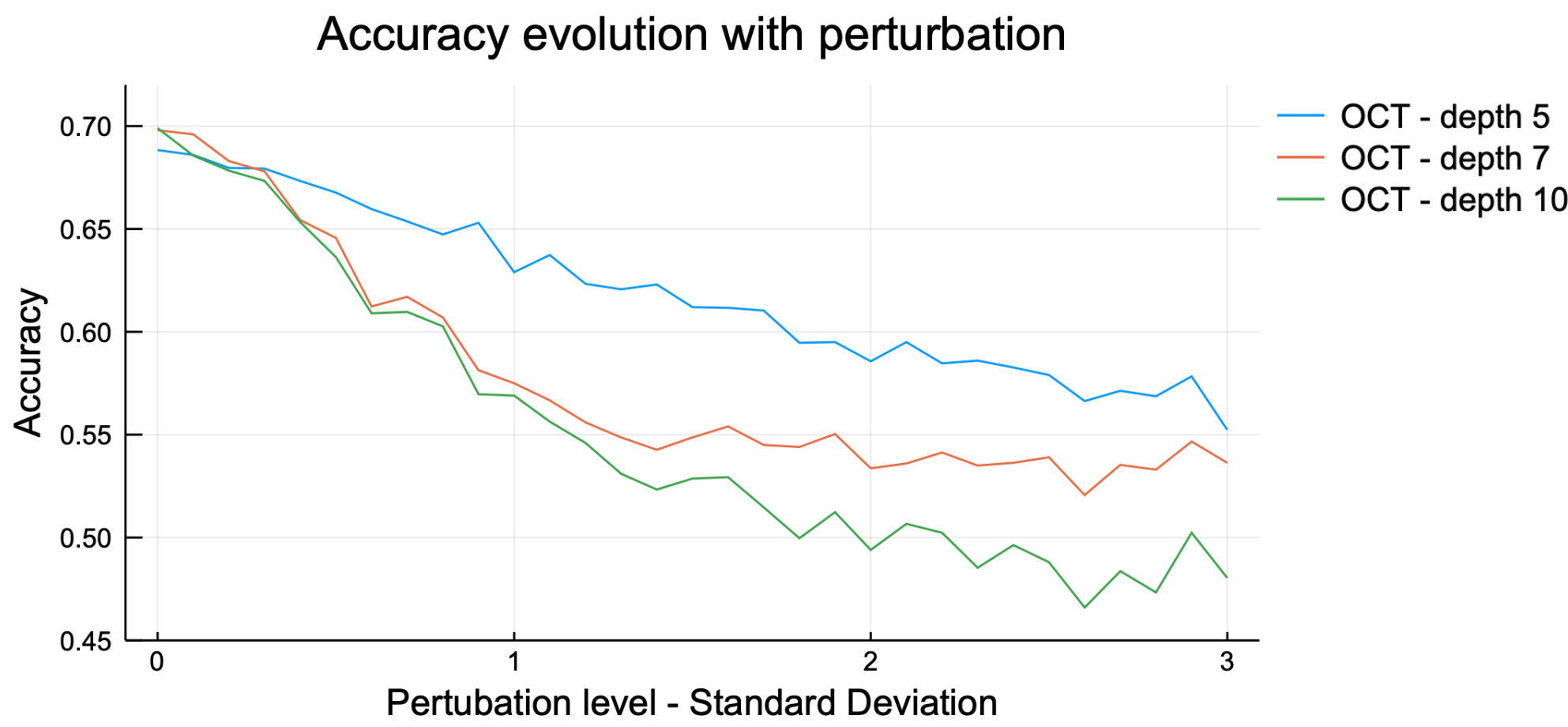
3. SENSITIVITY ANALYSIS

Progressive perturbation of patient utility data within ±3σ of true values

Process	Results
1. Perturb all data	1. Accuracy decreases by 15%
2. Perturb important features	2. Global decrease by 6-15%
3. Perturb other features	3. No effect

Perturbation impact

Legend: sd = 0.5, sd = 1, sd = 1.5, sd = 0.5 + exclusion, sd = 1 + exclusion, sd = 1.5 + exclusion

4. MODELS ROBUSTIFICATION WITH PERTURBED DATA GENERATION

- OCT, OCT-H and CART models **highly sensitive to pertubations in data**
- The more complex the model the more sentive
- **Robustification**: **small**, **medium** and **large** perturbing important features in the train set, and **merge the sets**
- **Results**: Superior stability under perturbations, **+10% accuracy** in highly perturbed scenarios

Robustification of Train Set

Comparison or robust and classic models

Legend: OCT, OCT_rob
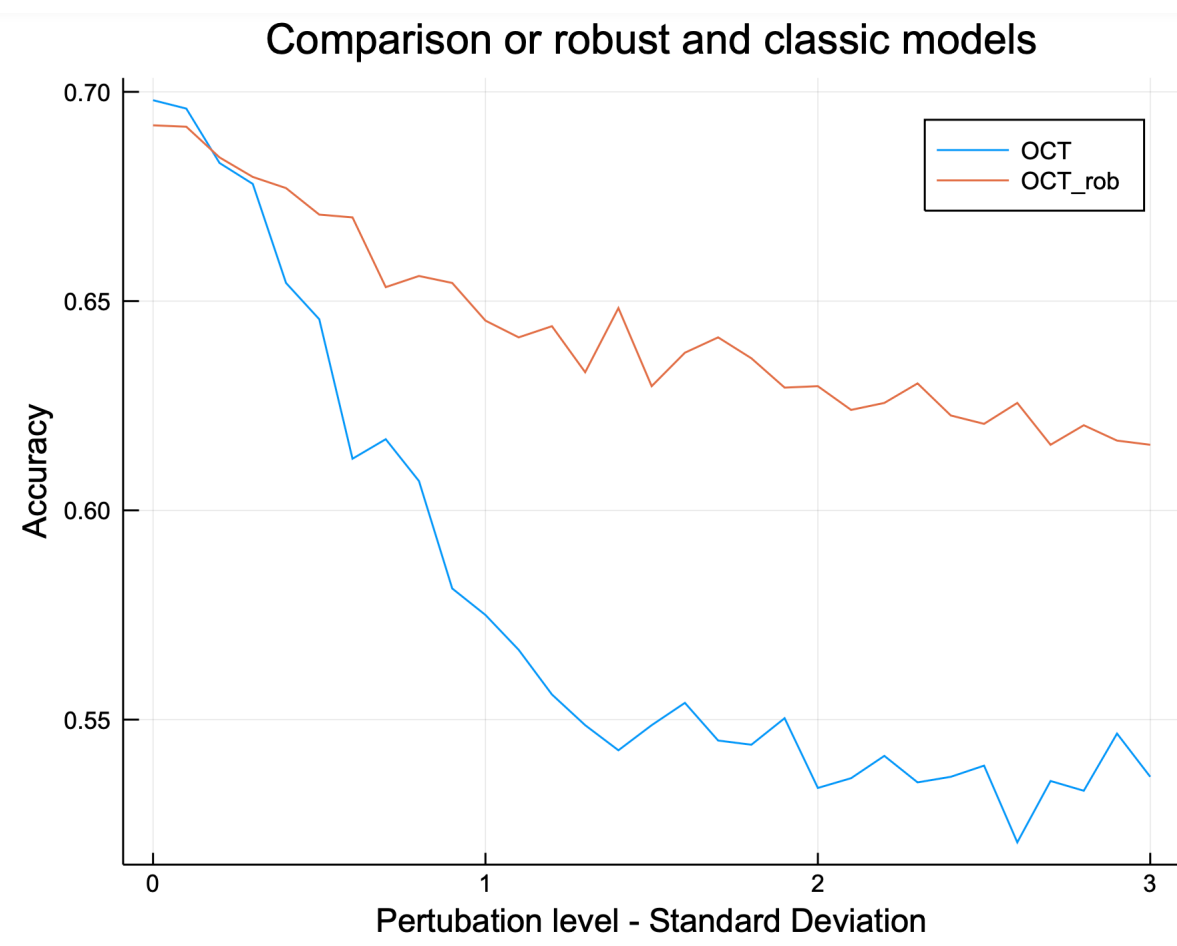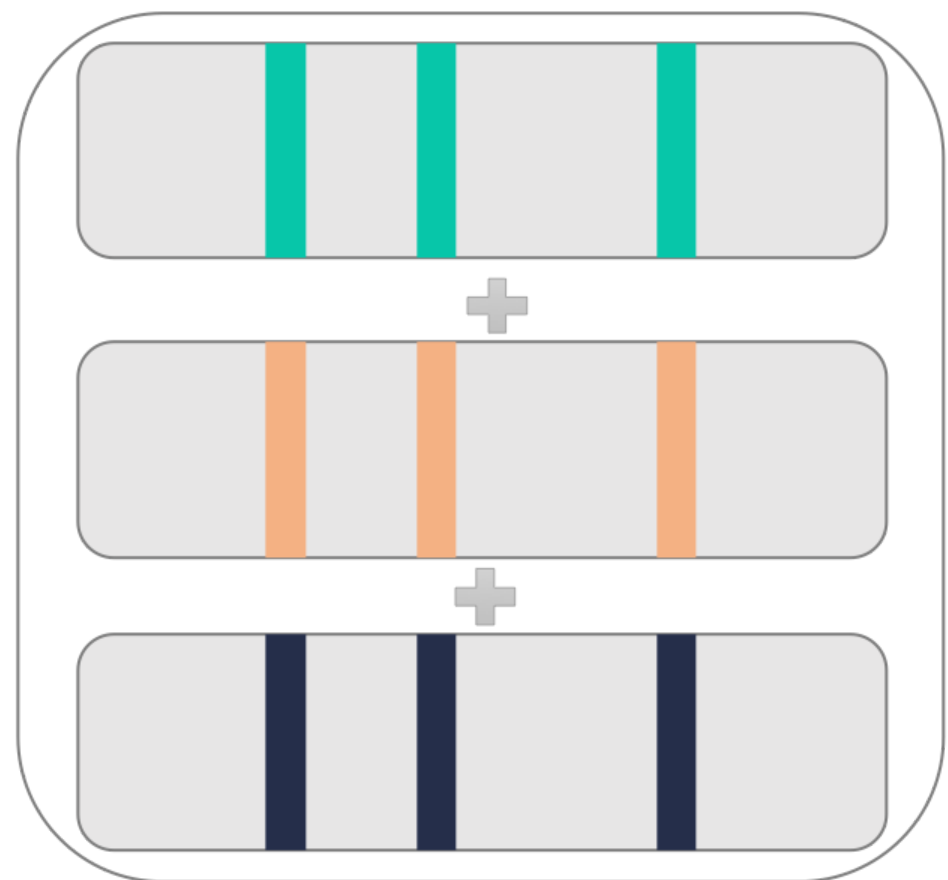
5. THE PRICE OF INTERPRETABILITY

Trade-off between model's interpretability and robustification
- Due to data generation process, we no longer have an interpretable path in the tree models
- Models' structure remains similar in terms of hyperparameters
- Accuracy is similar to original accuracy without perturbation

---

10K Patient Observations

70% Best Model Accuracy

±3σ Data Perturbation

-15% Accuracy Decrease

+10% Accuracy Improvement for Robust Models