Hybrid AI Methods for Feature Selection in Personalized Treatment Effect Prediction

Pamela M. Chiroque-Solano[a], Manel Martínez-Ramón[b], Thomas Jaki[a,c]
pamela.chiroque-solano@ur.de

Universität Regensburg

Women in Data Science Worldwide | Regensburg

1. Overview

- Accurate **Predicted Individual Treatment Effects (PITE)** are crucial for personalizing treatment decisions in healthcare [Lamont et al., 2016].
- Clinical data is **high-dimensional and heterogeneous**, making it difficult to identify relevant features.
- **Bayesian Additive Regression Trees (BART)** are effective for modeling complex relationships but depend on careful feature selection to optimize performance [Chipman et al., 2010].
- **Meta-heuristic optimization algorithms** such as Genetic Algorithms [Holland, 1975] and Particle Swarm Optimization [J. and R.] can help identify the most relevant features.
- Using **hybrid AI approaches** combining meta-heuristics and BART enhances both **prediction accuracy** and **computational efficiency**.

2. Research Questions

1. How to select relevant, non-redundant features for PITE?
2. How can meta-heuristic identify the most relevant variables for PITE?
3. How effective are AI models (BART) at estimating PITE when trained on optimally selected feature subsets?

3. PITE Formalization

PITE consists of the difference between experimental(E) and control (C) prediction for each individual

$\text{PITE}_i = f_E(\boldsymbol{X}_i) - f_C(\boldsymbol{X}_i), \quad f(\cdot)$ is a predictor

- PITE is unobserved since we typically observe the outcome for a given patient only under either the E or the C condition.
- The best method to estimate the PITE is not necessarily the best fitted model.

5. Methodology Overview

Model: Bayesian Additive Regression Trees (BART)

$$Y_i = \sum_{j=1}^{m} g(X_i; T_j, M_j) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

Feature Selection with Meta-Heuristics:
Fitness $= \alpha \cdot \text{Accuracy} - \beta \cdot \frac{|S|}{|F|}$
Where $|S|$: number of selected features, $|F|$: total number of features, α, β: tunable parameters.
Algorithms: Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Grey Wolf Optimizer (GWO)

6. Results (Simulated Data)

We simulate a dataset with treatment effects, apply Bayesian Additive Regression Trees (BART) to estimate the Predicted Individual Treatment Effects (PITE), and integrate meta-heuristic algorithms (e.g., Genetic Algorithm [GA], Particle Swarm Optimization [PSO]) to perform feature selection. Results in Table .

Method	RMSE ↓[a]	AUUC ↑[b]	# Features ↓	Time (min) ↓
BART + GA	0.36	0.83	21	19.1
BART + PSO	0.34	0.84	18	17.4
BART (all)[c]	0.48	0.76	100	3.0

Performance of BART based PITE estimation using different feature selection methods. The RMSE is computed from model based two-step predictions. AUUC assesses discriminatory ability. Models trained on simulated data using the `bartMachine` package.

- Both BART + GA and BART + PSO significantly outperform the baseline BART model
- BART + PSO achieves the best overall performance
- The BART model without feature selection, BART (all), performs worst across all metrics

[a] RMSE (Root Mean Square Error)
[b] AUUC (Area Under Uplift Curve)
[c] BART (all) = BART trained on the entire feature set, used as a baseline to compare against BART models with optimized feature subsets.

7. Conclusion

- Flexible machine learning models are paired with optimization-driven feature selection techniques
- Hybrid AI approaches improve clinical decision-making and patient outcomes.
- In personalized medicine, this synergy provides accurate and efficient identification of patients who benefit most from a given treatment

8. Challenges and discussion

- Address treatment bias in clinical trials and observational datasets.
- Improve generalization across subgroups.
- Optimize computational efficiency of hybrid methods.
- Move toward interpretable models for clinicians.
- PITE estimation with different feature sets.
- Heuristic Methods may not be directly comparable.

9. References

References

Hugh A. Chipman, Edward I. George, and Robert E. McCulloch. Bart: Bayesian additive regression trees. *The Annals of Applied Statistics*, 2010. doi: 10.1214/09-aoas285.

John H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 1975. ISBN 9780262581110.

Kennedy J. and Eberhart R. Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4 of *ICNN-95*, page 1942–1948. IEEE. doi: 10.1109/icnn.1995.488968. URL `http://dx.doi.org/10.1109/ICNN.1995.488968`.

Andrea Lamont, Michael D Lyons, Thomas Jaki, Elizabeth Stuart, Daniel J Feaster, Kukatharmini Tharmaratnam, Daniel Oberski, Hemant Ishwaran, Dawn K Wilson, and M Lee Van Horn. Identification of predicted individual treatment effects in randomized clinical trials. *Statistical Methods in Medical Research*, pages 142–157, 2016. doi: 10.1177/0962280215623981.

Affiliations

a: Faculty of Informatics and Data Science, University of Regensburg, Germany
b: Department of Electrical and Computer Engineering, The University of New Mexico, USA
c: MRC Biostatistics Unit, Cambridge University, UK