Prediction of Oncotype DX using DRFs

Romain Pic [1] Zeina Al Masry [2] Clément Dombry [1] Christine Devalland [3]

[1]Laboratoire de Mathématiques de Besançon (UBFC/CNRS) [2]FEMTO-ST (UBFC/CNRS/ENSMM) [3]Hôpital Nord Franche-Comté

Introduction

Oncotype DX (ODX) is a multi-gene expression signature for **breast cancer** that provides prognostic and predictive breast cancer recurrence information for estrogen receptor (ER)-positive and HER2 negative patients. This test can predict if a chemotherapy treatment would be beneficial. However, this test is **expensive** and several studies have shown its link with certain clinico-pathological data. The result of the test is a score between 0 and 100 and guides clinicians regarding **prescription of chemotherapy** (e.g. for score higher than 25).
Aim: Predict the ODX score based on histopathological variables.
Proposed methodology: Probabilistic forecast using Distributional Regression Forest (DRF; [1])

Presentation of the cohort

The cohort is a retrospective study between 2012 and 2019 with 334 cases that underwent ODX assay from three hospitals: Besançon, Belfort and Dijon. All patients have ER-positive and HER2-negative early breast cancer. We used **9 variables** for the prediction of the ODX: age, tumor size, SBR grade, Nottingham grade, ER status, PR status, Ki67 index proliferation, P53 protein and lymph node status. In order to remove missing values of ODX, we had to remove one patient and we finally keep **333 patients**.

Flowchart of Distributional Regression Forests

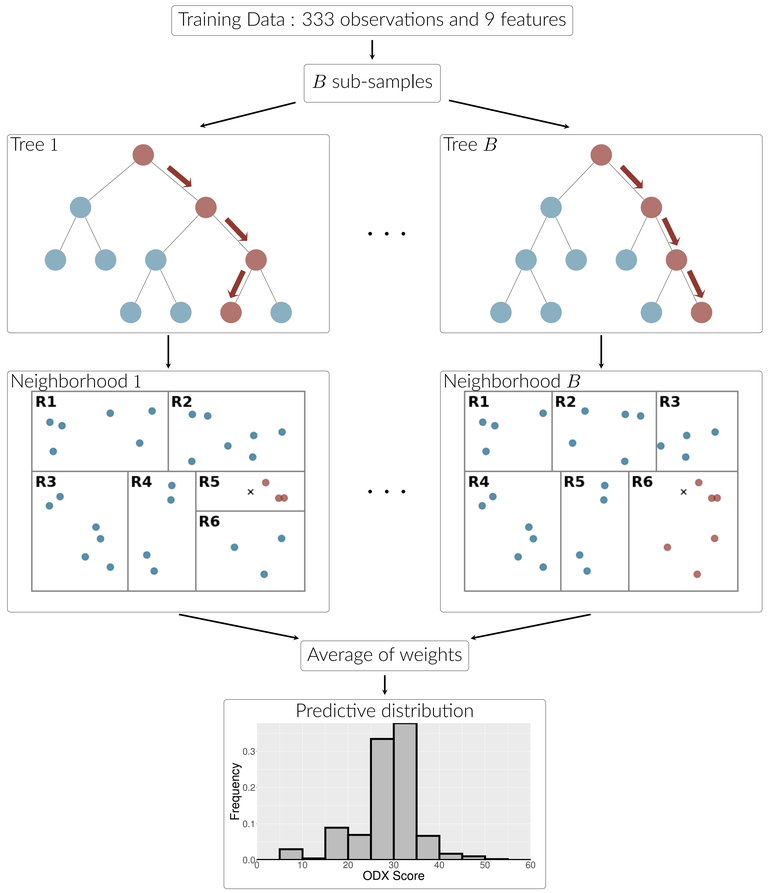

Figure 1. Flowchart of distributional random forest. Starting from the training data, a large number of subsamples are chosen and binary trees are constructed on each subsamples; the neighborhoods and weights at the point to predict (cross) are computed in each tree and then averaged so as to give the forest weights; the predictive distribution corresponds to the weighted sample of the original training data with these forest weights.

Outputs of DRF

DRF is a powerful tool and is able to cover numerous methods such as classification, mean regression and quantile regression. Moreover, DRFs assess the uncertainty of a prediction.

- **Histogram** of the predictive distribution;
- **Probability** of classes of interest;
- **Mean** or **median** prediction;
- **Uncertainty** associated to the prediction (standard deviation or confidence interval);
- **Similar/influential patients** for meaningful and informative comparison. (*very important for practitioners*)

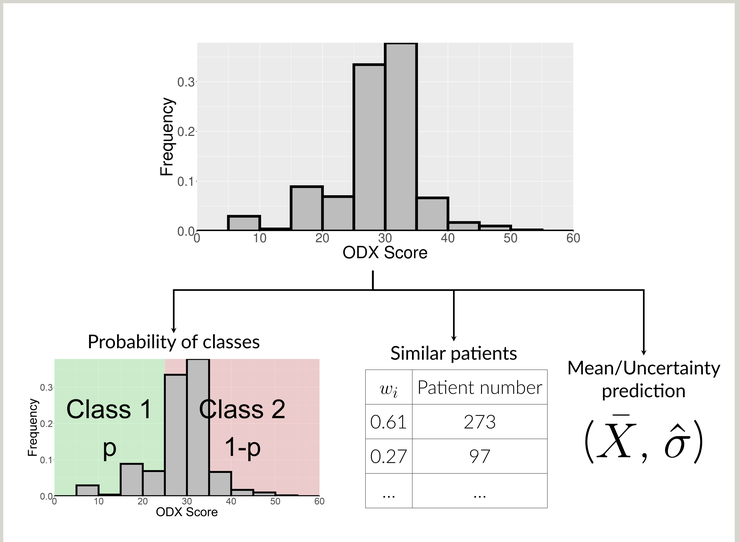

Figure 2. Applications of DRF.

Evaluation of the predictive distribution

The Continuous Ranked Probability Score (CRPS; [5]) is a scoring rule used to compare a predictive distribution F and the outcome y. The CRPS is used in Figure 3 to evaluate the performances of DRF for each patient in an out-of-bag prediction. The CRPS evaluates the dispersion and the bias; thus, a **higher CRPS** correspond to **higher bias** and/or **higher dispersion** of the predictive distribution.

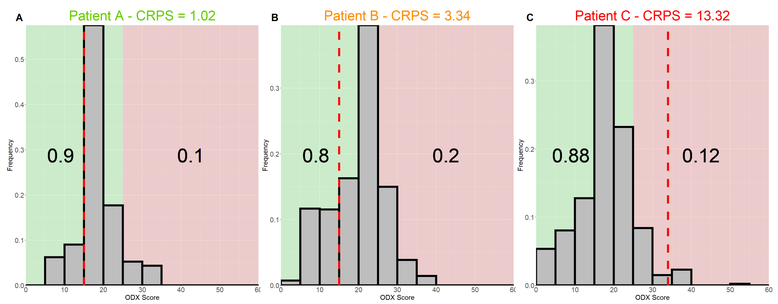

Figure 3. We highlight three cases : a "good" prediction, a "medium" prediction and a "bad" prediction associated with low, medium and high values of CRPS respectively. The red dashed line shows the true ODX, the regions in green and red corresponds to $ODX \leq 25$ and $ODX > 25$ respectively and the numbers at the center of these regions are the predicted probabilities of each class.

Comparison with state-of-the-art techniques

		Predicted	
		$\text{ODX} \leq 25$	$\text{ODX} > 25$
True	$\text{ODX} \leq 25$	231	20
	$\text{ODX} > 25$	49	33

Accuracy	79.3%
Sensitivity	92.0%
Specificity	40.2%
Positive Predictive Value	62.3%
Negative Predictive Value	82.5%
F1-score	0.870
Area Under Curve	0.759

Table 1. Evaluation of DRF with classification diagnostics. Confusion matrix (left) and standard metrics (right).

Study	Data set (n_{train}, n_{test}, p)	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Kim et al. (2019)	$(208, 76, 20)$	88.0	17.0	96.0	0.917
Orucevic et al. (2019)	$(65,754,\ 18,585,\ 5)$	87.2	18.3	99.2	0.81
Pawloski et al. (2021)	$(2587, 1184, 6)$	96.3	48.3	96.3	NA
DRF	$(333, \text{OOB}, 9)$	79.3	40.2	92.0	0.759

Table 2. Comparison (on different datasets) of the DRF method with 3 other methods for the 2-class classification problem ($\text{ODX} \leq 25$ and $\text{ODX} > 25$). The methods are taken from Kim et al. (2019), Orucevic et al. (2019) and Pawloski et al. (2021) [4, 6, 7]. The high class is taken as *positive*.

Study	Acc.	Sensi.	Speci.	PPV	NPV	AUC
Hou et al. (2017)	68.7	85.7	41.4	72.6	61.5	NA
Baltres et al. (2017)	46.3	55.0	78.0	58.3	62.4	0.63
DRF	57.7	58.5	83.6	87.0	51.9	0.748

Table 3. Comparison (on different datasets) of the DRF method with 2 other methods (Hou et al. (2017) and Baltres et al. (2017) [3, 2]) for the 3-class classification problem: $\text{ODX} \leq 18$ (low), $18 < \text{ODX} \leq 30$ (medium) and $\text{ODX} > 30$ (high). The sensitivity, specificity, PPV, NPV and AUC are given only for the lower class.

Conclusion

- Novel approach: **distribution prediction**;
- Various outputs providing information for practitioners (e.g. **similar profiles**);
- **Performance similar to state-of-the-art** techniques in terms of classification;
- **Explainable** and transparent method;
- Operational clinical `RShiny` app for practitioners.
- *Perspectives*: robustness to missing values and noise.
- *Limitation*: predictions are only as good as the training dataset they are based on.

[1] S. Athey et al. "Generalized random forests". In: *The Annals of Statistics* 47.2 (2019).
[2] A. Baltres et al. "Prediction of Oncotype DX recurrence score using deep multi-layer perceptrons in estrogen receptor-positive, HER2-negative breast cancer". In: *Breast Cancer* 27.5 (2020).
[3] Y. Hou et al. "Comparison of Oncotype DX With Modified Magee Equation Recurrence Scores in Low-Grade Invasive Carcinoma of Breast". In: *American Journal of Clinical Pathology* 148.2 (2017).
[4] I. Kim et al. "A predictive model for high/low risk group according to oncotype DX recurrence score using machine learning". In: *European Journal of Surgical Oncology* 45.2 (2019).
[5] J. E. Matheson and R. L. Winkler. "Scoring Rules for Continuous Probability Distributions". In: *Management Science Series A-theory* 22.10 (1976).
[6] A. Orucevic et al. "Nomogram update based on TAILORx clinical trial results - Oncotype DX breast cancer recurrence score can be predicted using clinicopathologic data". In: *The Breast* 46 (2019).
[7] K. R. Pawloski et al. "Supervised machine learning model to predict oncotype DX risk category in patients over age 50". In: *Breast Cancer Research and Treatment* 191.2 (2021).