Interpretable Random Forests for Industrial Applications

Clément Bénard
Sorbonne Université - Safran Tech

Supervisor(s): Gérard Biau (Sorbonne Université), Sébastien Da Veiga (Safran Tech) and Erwan Scornet (École Polytechnique)

Ph.D. expected duration: Dec. 2018 - Dec. 2021

Adress: Safran Tech, Modelling & Simulation, Rue des Jeunes Bois, Chateaufort, 78114 Magny-Les-Hameaux, France

Email: clement.benard@safrangroup.com

Abstract:

In the manufacturing industry, the core of production processes involves complex physical and chemical phenomena. Their control and efficiency is of critical importance. In practice, data is collected along the manufacturing line, characterizing both the production conditions and its quality. State-of-art learning algorithms can successfully catch patterns of such complex processes, characterized by non-linear effects and low-order interactions between parameters. However, any decision impacting a production process has long term and heavy consequences, and thus, cannot simply rely on blind stochastic modelling. Models have to be interpretable, i.e. provide an understanding of the internal mechanism that builds a relation between inputs and ouputs, to provide insights to guide the physical analysis. There is no agreement in statistics and machine learning communities about a rigorous definition of interpretability [7]. It is yet possible to define minimum requirements for interpretable algorithms: simplicity, stability [9] and predictivity.

Decision trees [2] can model highly non-linear patterns while having a simple structure and are therefore widely used when interpretability is required. Decision trees are also highly unstable to small data pertubation, which is a very strong limitation to their practical use. The random forest method [1] stabilizes decision trees by aggregating many of them, it strongly improves accuracy but the model is a black box. Another class of supervised learning method can model non-linear patterns while having a simple structure: rule models. A rule is a conjunction of constraints on inputs variables that form a hyper-rectangle in the input space, where the estimated output is constant. As an example, an elementary rule is typically

$$\textbf{If } \begin{cases} X^{(3)} < 2.54 \\ \textbf{\&} \ X^{(5)} \geq 0.061 \end{cases} \textbf{ then } \hat{Y} = 21.1 \textbf{ else } \hat{Y} = 0.41 \textbf{ .}$$

A collection of rules is combined to form a model. Many algorithms were developped, among them: SLIPPER [3], Rulefit [6], and Node Harvest [8]... They share the same drawback as trees: instability. Besides, rule learning is almost exclusively dedicated to classification problems.

In this work, we design a new regression rule algorithm which inherits the accuracy of random forests, the simplicity of decision trees while having a stable structure for problems with low-order interaction effects. Our algorithm is based on random forests [1], and its general principle is as follows: since each node of a decision tree can be turned into an elementary rule, the core idea is to extract rules from a tree ensemble based on their frequency of appearance. Thus, we use the principle of random forests, but the structure of the forest itself is aggregated instead of predictions. The most frequent rules, which represent robust and strong patterns in the data, are linearly combined to form predictions through a ridge regression.

Dataset	RuleFit	Node Harvest	Our algorithm
Ozone	0.22	0.30	**0.62**
Mpg	0.25	0.43	**0.83**
Prostate	0.32	0.23	**0.48**
Housing	0.19	0.40	**0.80**
Diabetes	0.18	0.39	**0.66**
Machine	0.23	0.29	**0.88**
Galaxy	0.40	0.39	**0.77**
Abalone	0.31	0.38	**0.82**
Bones	0.59	0.52	**0.89**

Table 1: Mean stability over a 10-fold cross-validation for various public datasets.

Experiments on various public datasets [4, 5] show good performance of the procedure in terms of both predictive accuracy and stability. The model error is quantified with the unexplained variance. To evaluate stability, a 10-fold cross-validation is run, and, for each pair of folds the relative size of the intersection between the two lists of rules is computed. The accuracy is similar to competitors while stability is considerably improved as shown in Table 1.

References

[1] L. Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

[2] L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone. *Classification and Regression Trees*. Chapman & Hall/CRC, Boca Raton, 1984.

[3] W.W. Cohen and Y. Singer. A simple, fast, and effective rule learner. In *Proceedings of the 16th National Conference on Artificial Intelligence and 11th Conference on Innovative Applications of Artificial Intelligence*, pages 335–342, Palo Alto, 1999. AAAI Press.

[4] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017.

[5] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *The Annals of statistics*, 32:407–499, 2004.

[6] J.H. Friedman and B.E. Popescu. Predictive learning via rule ensembles. *The Annals of Applied Statistics*, 2:916–954, 2008.

[7] Z.C. Lipton. The mythos of model interpretability. *arXiv:1606.03490*, 2016.

[8] N. Meinshausen. Node harvest. *The Annals of Applied Statistics*, 4:2049–2072, 2010.

[9] B. Yu. Stability. *Bernoulli*, 19:1484–1500, 2013.

Short biography – Clément Bénard is a research engineer at Safran Tech and a second year PhD student in statistics, in collaboration with LPSM, Sorbonne Université.