Meet-up 2019 | Doctorants & Industrie

Task-driven forecasting with random forests

Frédéric Logé[1,2]

Erwan Le Pennec[1], Eric Moulines[1], Habiboulaye Amadou-Boubacar[2]

1 : Centre de Mathématiques Appliquées, Polytechnique, Palaiseau, France

2 : Computational & Data Science Lab, Paris Innovation Campus, Air Liquide R&D, France

1. CONTEXT

In supervised learning, theory and applications focus on the predictability of a target Y given some feature vector X, measured by classic metrics such as RMSE, MAE or MAPE. Yet, in practice, such metrics do not reflect the impact of prediction error, since the way those predictions are used is not taken into account. Research has been done in this direction with neural networks [1]. We focused on a rather simple use-case for which we simulated data. We proceeded to build a forecast of Y, optimized for the task, based on the knowledge of how the data was simulated, which makes for an oracle benchmark and we also build adapted decision trees (and by natural extension random forests) for this specific task [2,3]. Nets can perform significantly better than bagging methods on some tasks, yet they loose interpretability and most importantly, are usually regarded as a difficult-to-implement to solution, contrary to random forest.

2. PROPOSAL

Let

$(Y_i, X_i) \in \mathcal{Y} \times \mathcal{X}$	some target and features
Θ	control based on $\hat{Y}_i$
$\mathcal{F}$	admissible predictors
$\mathbf{L} : \Theta \times \mathcal{Y} \to \mathbb{R}$	our loss function
$\mathbf{R} : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$	classic risk function

Agnostic

$$\arg\min_{f \in \mathcal{F}} \sum_{i=1}^{n} \mathbf{R}(f(X_i), Y_i)$$

Task-Driven

$$\arg\min_{f \in \mathcal{F}} \sum_{i=1}^{n} \mathbf{L}\left(\arg\min_{\theta \in \Theta} \mathbf{L}\left(\theta, f(X_i)\right), Y_i\right)$$

Let $\theta^*(y) := \arg\min_{\theta \in \Theta} \mathbf{L}\left(\theta, y\right)$. We consider two ways to adapt decision trees to loss $\mathbf{L}$.

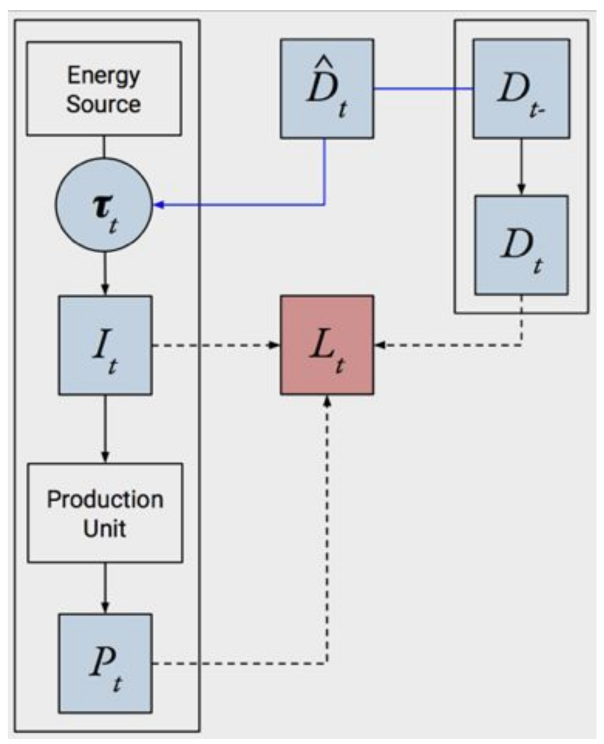
Prediction

$$y^*\left(\left(Y_i\right)_{i \in \text{leaf}}\right) := \arg\min_{y \in \mathcal{Y}} \sum_{i \in \text{leaf}} \mathbf{L}(\theta^*(y); Y_i)$$

Splitting

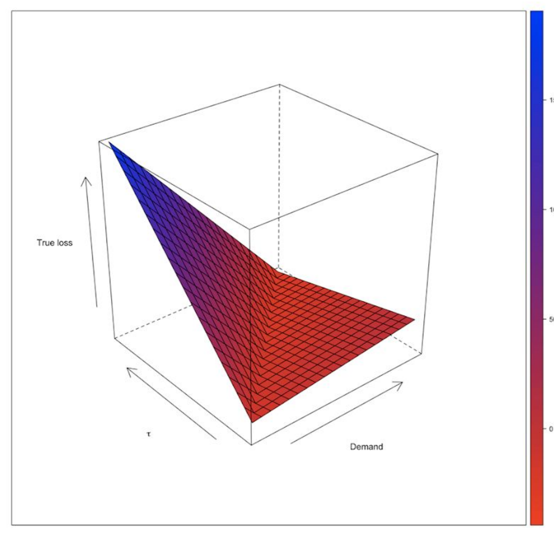
$$\arg\min_{\nu} \sum_{\text{leaves}(\nu)} \sum_{i \in \text{leaf}} \mathbf{L}\left(\theta^*\left(y^*\left(\left(Y_i\right)_{i \in \text{leaf}}\right)\right); Y_i\right)$$

3. CASE STUDY: Production unit problem, with asymmetric costs



$$\mathbf{L}(D_t, P_t, I_t) := C_d \cdot I_t + C_e \cdot \max\{D_t - P_t; 0\} - C_f \cdot \min\{P_t; D_t\}$$

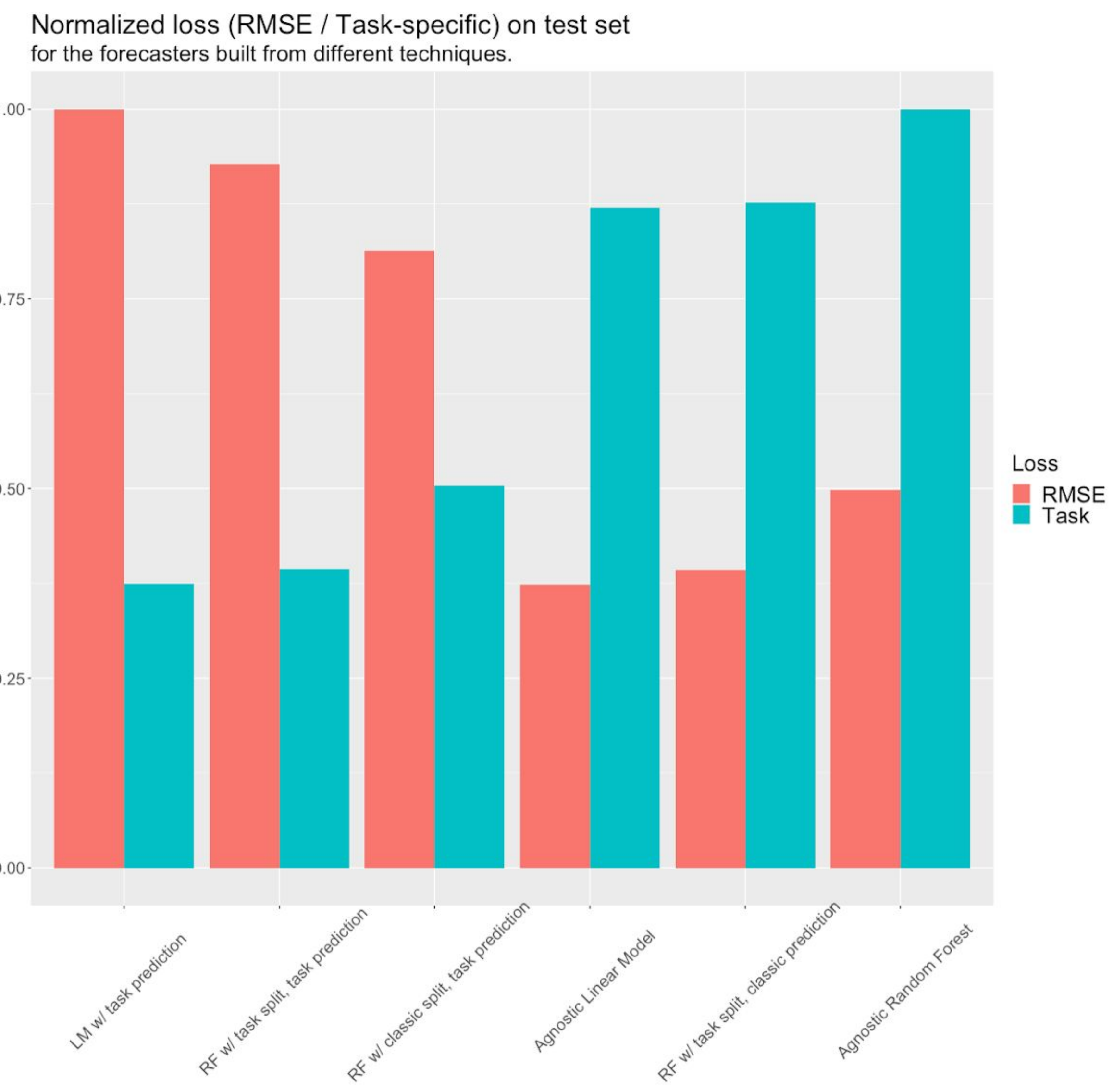
Input cost — Unmet demand cost — Revenue



Demand was simulated as an AR(1) process.

4. RESULTS

The prediction criterion must be changed in order to respond to the task (at least 50% improvement), changing the splitting criterion is not sufficient in itself (83% loss) but it does help when combined with prediction (task loss drops to 37% and matches that of task-based LM).



5. FUTURE WORK

Code robustness and development. A comparison with the work on neural nets [1] would be interesting.

REFERENCES

[1] Donti, Priya and Amos, Brandon and Kolter, J. Zico. Task-based End-to-end Model Learning in Stochastic Optimization in Advances in Neural Information Processing Systems 30.

[2] Github, project taskDrivenRandomForest https://github.com/FredericLoge/taskDrivenRandomForest

[3] Github, project taskDrivenForecast https://github.com/FredericLoge/taskDrivenForecast/

EDMH

Air Liquide R&D; CMAP, Polytechnique

Frédéric Logé

frederic.logemunerel@gmail.com