

Boosting



INFORMAÇÃO,
TECNOLOGIA
& INOVAÇÃO

Boosting

- Ensembles
- Weak Classifiers



Boosting

$$G(\mathbf{x}) = \text{sign} \left(\sum_{m=1}^M \alpha_m G_m(\mathbf{x}) \right)$$



AdaBoost.M1

1. Inicialize o peso das observações $w_i = 1/N$ com $i = 1, \dots, N$.
2. Para m variando de 1 a M :
 - 2.1 Ajuste um classificador $G_m(\mathbf{x})$ ao conjunto de treinamento utilizando os pesos w_i .
 - 2.2 Calcule $\text{err}_m = \frac{\sum_{i=1}^N w_i \mathbf{I}(y_i \neq G_m(\mathbf{x}))}{\sum_{i=1}^N w_i}$
 - 2.3 Calcule $\alpha_m = \log\left(\frac{1 - \text{err}_m}{\text{err}_m}\right)$.
 - 2.4 Atualize $w_i \leftarrow w_i \cdot \exp\{\alpha_m \cdot \mathbf{I}(y_i \neq G_m(\mathbf{x}))\}$, $i = 1, 2, \dots, N$
3. Retorne $G(\mathbf{x}) = \text{sign}\left[\sum_{m=1}^M \alpha_m G_m(\mathbf{x})\right]$.



dmlc
XGBoost

 LightGBM

 Yandex
CatBoost



	CatBoost		LightGBM		XGBoost		H2O	
	Tuned	Default	Tuned	Default	Tuned	Default	Tuned	Default
Adult	0.26974	0.27298 +1.21%	0.27602 +2.33%	0.28716 +0.46%	0.27542 +2.11%	0.28009 +3.84%	0.27510 +1.99%	0.27607 +2.35%
Amazon	0.13772	0.13811 +0.29%	0.16360 +18.80%	0.16716 +21.38%	0.16327 +18.56%	0.16536 +20.07%	0.16264 +18.10%	0.16950 +23.08%
Click prediction	0.39090	0.39112 +0.06%	0.39633 +1.39%	0.39749 +1.69%	0.39624 +1.37%	0.39764 +1.73%	0.39759 +1.72%	0.39785 +1.78%
KDD appetency	0.07151	0.07138 -0.19%	0.07179 +0.40%	0.07482 +4.63%	0.07176 +0.35%	0.07466 +4.41%	0.07246 +1.33%	0.07355 +2.86%
KDD churn	0.23129	0.23193 +0.28%	0.23205 +0.33%	0.23565 +1.89%	0.23312 +0.80%	0.23369 +1.04%	0.23275 +0.64%	0.23287 +0.69%
KDD Internet	0.20875	0.22021 +5.49%	0.22315 +6.90%	0.23627 +13.19%	0.22532 +7.94%	0.23468 +12.43%	0.22209 +6.40%	0.24023 +15.09%
KDD upselling	0.16613	0.16674 +0.37%	0.16682 +0.42%	0.17107 +2.98%	0.16632 +0.12%	0.16873 +1.57%	0.16824 +1.28%	0.16981 +2.22%
KDD 98	0.19467	0.19479 +0.07%	0.19576 +0.56%	0.19837 +1.91%	0.19568 +0.52%	0.19795 +1.69%	0.19539 +0.37%	0.19607 +0.72%
Kick prediction	0.28479	0.28491 +0.05%	0.29566 +3.82%	0.29877 +4.91%	0.29465 +3.47%	0.29816 +4.70%	0.29481 +3.52%	0.29635 +4.06%



Tuning Parameters

1. max_depth (3 – 10) e min_child_weight (5 – 10)
2. subsample (70% – 100%) e colsample_bytree (\sqrt{d} – 100%)
3. ETA (0.001 – 0.5)

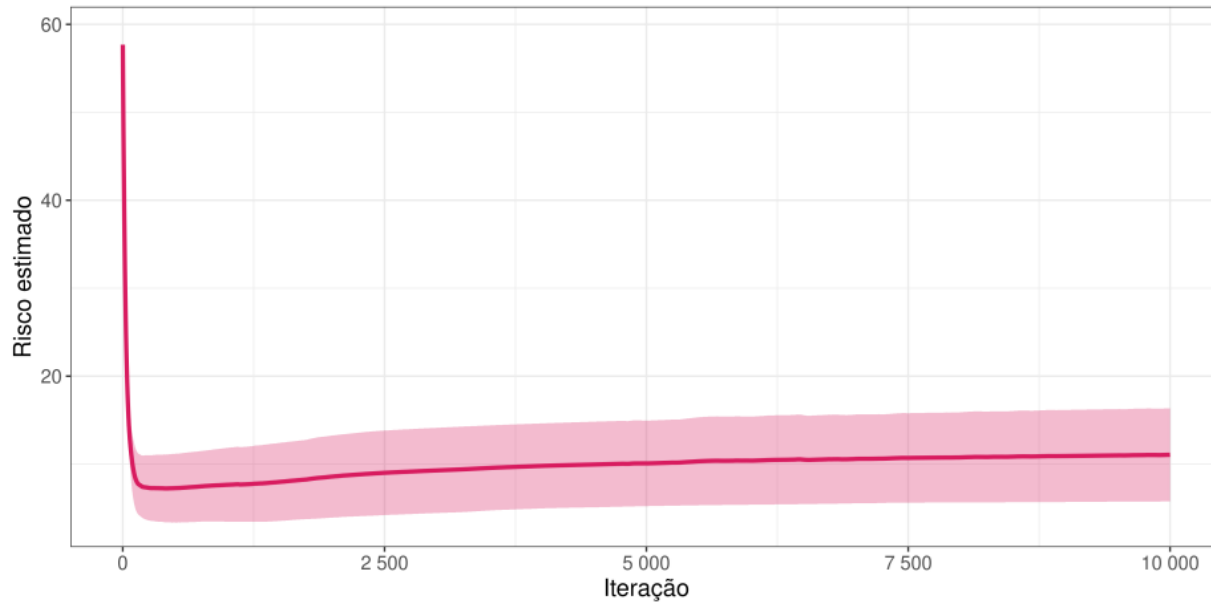
Grid Search; Randomized CV; Greedy Choice

Early Stopping



Tuning Parameters

Early Stopping



Amazon

