Estimation of ML model uncertainty in particle physics event classifiers

Miguel Cárdenas Montes[1] miguel.cardenas@ciemat.es

José M. Hernández[1] jose.hernandez@ciemat.es **Julia Vázquez Escobar**[1] julia.vazquez@ciemat.es

[1]Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (**CIEMAT**)

EuCAIFCon 2024

May 30, 2024

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

cfp CIEMAT física de partículas

Julia Vázquez-Escobar, J.M. Hernández, Miguel Cárdenas-Montes, Computer Physics Communications, 268, (2021)

Methods

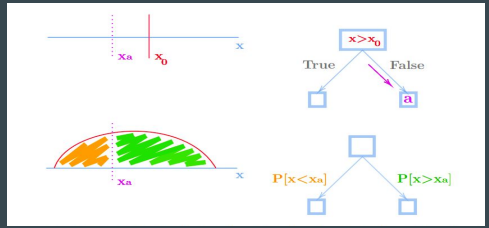
Bayesian approx.

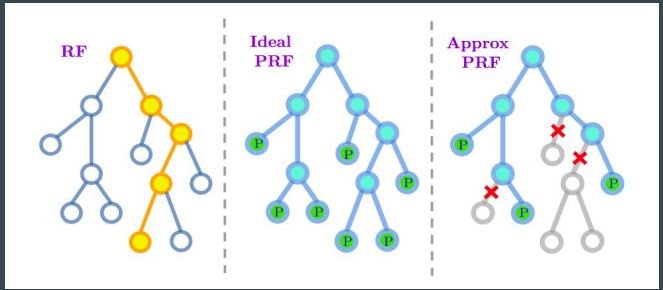
$$\widetilde{\mathbb{E}}\left[\mathbf{y}^*\right] := \frac{1}{T}\sum_{t=1}^{T}\mathbf{f}^{\hat{\omega}_t}\left(\mathbf{x}^*\right) \underset{T\to\infty}{\longrightarrow} \mathbb{E}_{q_\theta^*(\mathbf{y}^*|\mathbf{x}^*)}\left[\mathbf{y}^*\right]$$

$$\widetilde{\mathrm{Var}}\left[\mathbf{y}^*\right] := \tau^{-1}\mathbf{I} + \frac{1}{T}\sum_{t=1}^{T}\mathbf{f}^{\hat{\omega}t}\left(\mathbf{x}^*\right)^T\mathbf{f}^{\hat{\omega}_t}\left(\mathbf{x}^*\right) -$$

$$-\widetilde{\mathbb{E}}\left[\mathbf{y}^*\right]^T\widetilde{\mathbb{E}}\left[\mathbf{y}^*\right] \underset{T\to\infty}{\to} \mathrm{Var}_{q_\theta^*(\mathbf{y}^*|\mathbf{x}^*)}\left[\mathbf{y}^*\right]$$

Probabilistic RF





Local ensembles

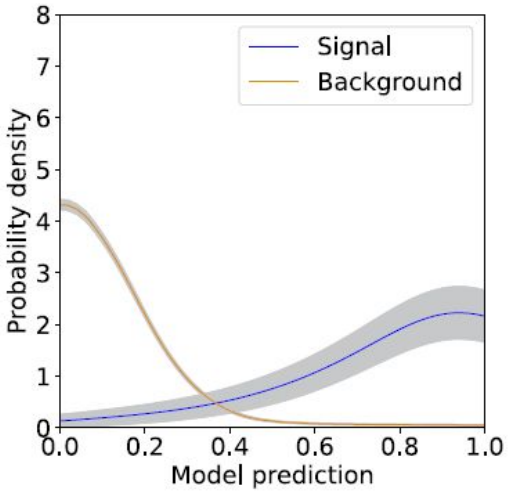
Proposition 1.1 *Let Δ_θ be the projection of a random perturbation with mean zero and covariance proportional to the identity $\epsilon \cdot I$ into the ensemble subspace spanned by $\{\xi_{(j)} : j > m\}$. Let P_Δ be the linearized change in prediction induced by the perturbation*

$$P_\Delta(x') := g_{\theta*}(x')^\top \Delta_\theta \approx \hat{y}(x', \theta * + \Delta_\theta) - \hat{y}(x', \theta*)$$

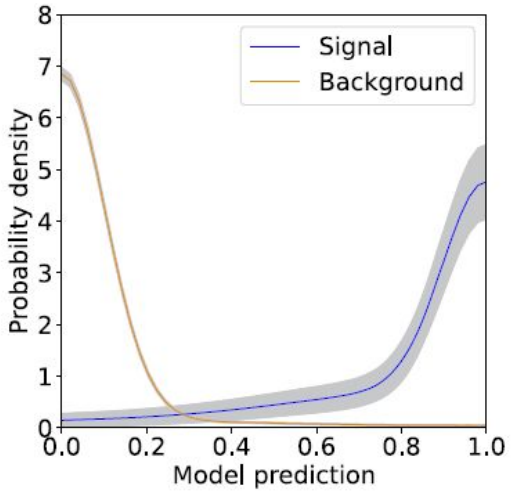
then $\mathcal{E}_m(x') = \varepsilon^{-1/2} \cdot SD(P_\Delta(x'))$.

Results

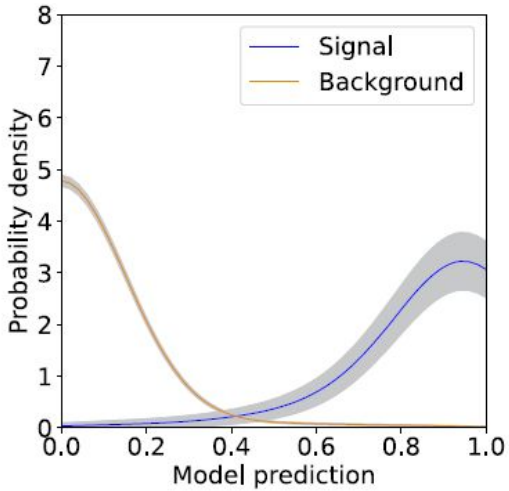
The probability density functions of the classification parameter for true signal and background events are shown.



(a) Local Ensembles



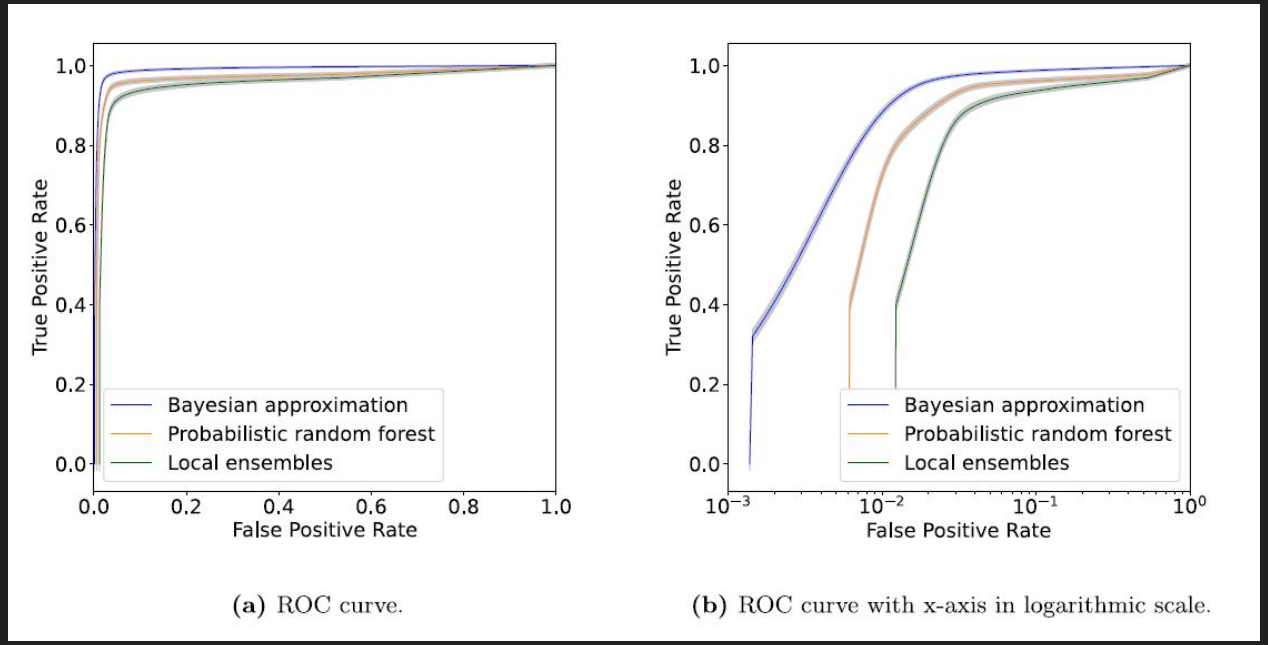
(b) Probabilistic Random Forest



(c) Bayesian approximation

Results

Model	AUC
Probabilistic Random Forest	0.969 ± 0.005
Local Ensembles	0.951 ± 0.006
Bayesian Approximation	0.990 ± 0.001



(a) ROC curve.

(b) ROC curve with x-axis in logarithmic scale.