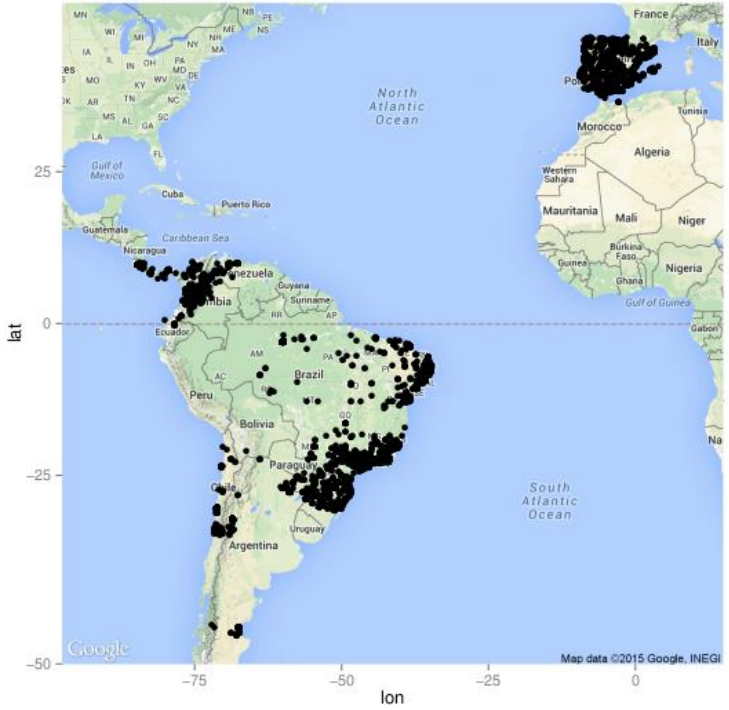
Explainable ML (XAI)

INFORMAÇÃO,

TECNOLOGIA

& INOVAÇÃO

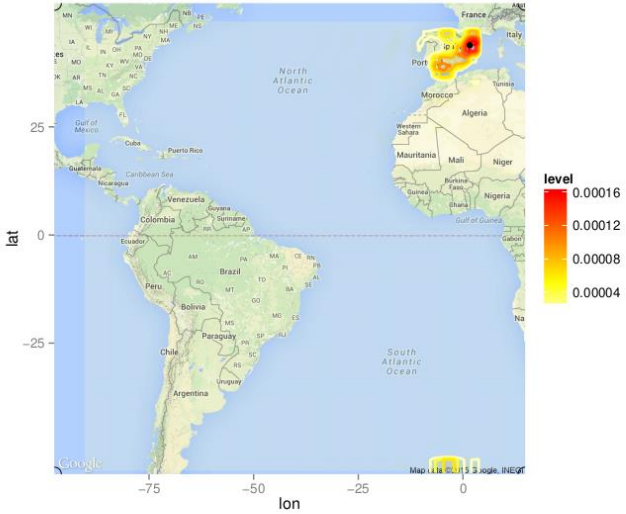
Para que explicações?



Onde ele foi postado?

Vanessa Regueira @VanessaRegueira · Jul 5

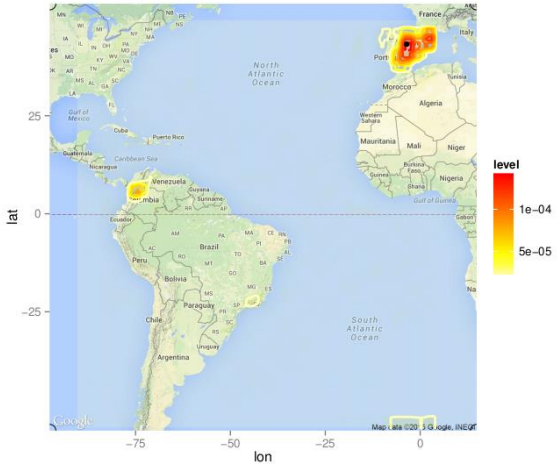
Contra la ola de calor.... Un chapuzón en la playa tempranito y ahora... Reclusión en casa… instagram.com/p/4wA3agvmOC/



- calor
- verano

Gaspar Sanchez @thaiklbali · Jul 4

Combatiendo el calor #verano #lacuevadekrusty #elmolar @ La Cueva de Krusty instagram.com/p/4tvDJrOwo_/



Explicações são úteis para avaliar se o modelo vai de fato funcionar.

Ex:

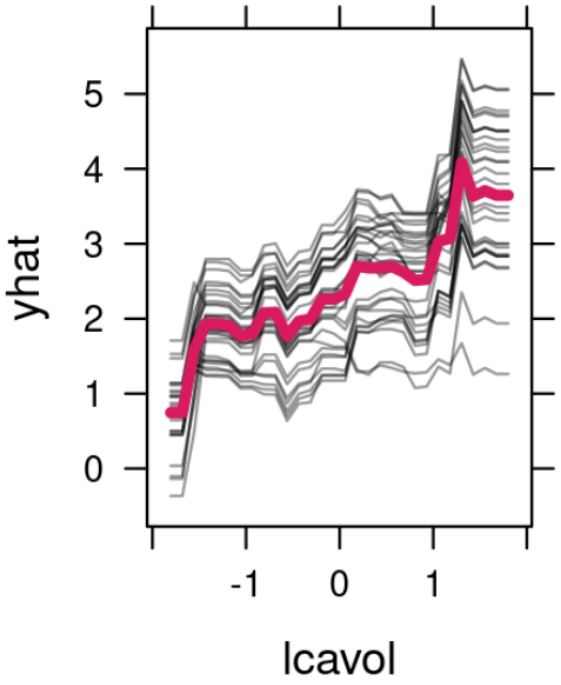
- Problemas de novos dados não serem i.i.d.'s aos dados disponíveis
- Covariáveis que não deviam ter sido utilizadas

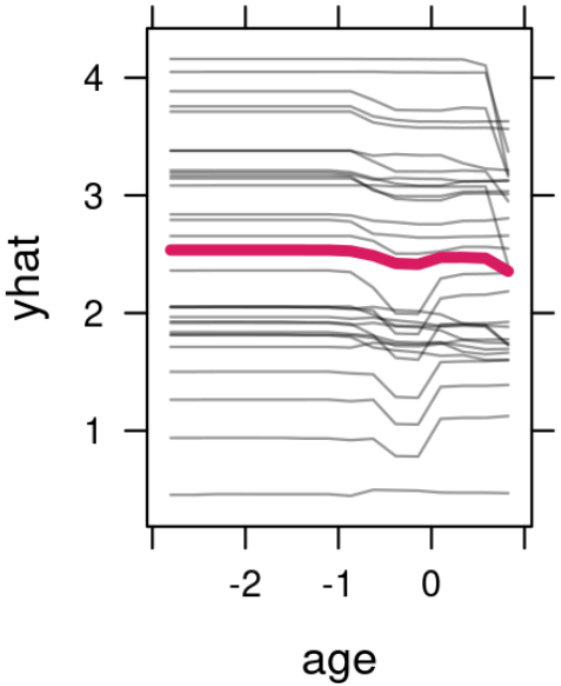
PDP - Partial dependence plot

$$h_i(x) = \frac{1}{n}\sum_{j=1}^{n} g(x_{j,1}, \ldots, x_{j,i-1}, x, x_{j,i+1}, \ldots, x_{j,d})$$

$$\mathbb{E}[g(X_1, \ldots, X_{i-1}, x, X_{i+1}, \ldots, X_d)]$$

PDP - Partial dependence plot



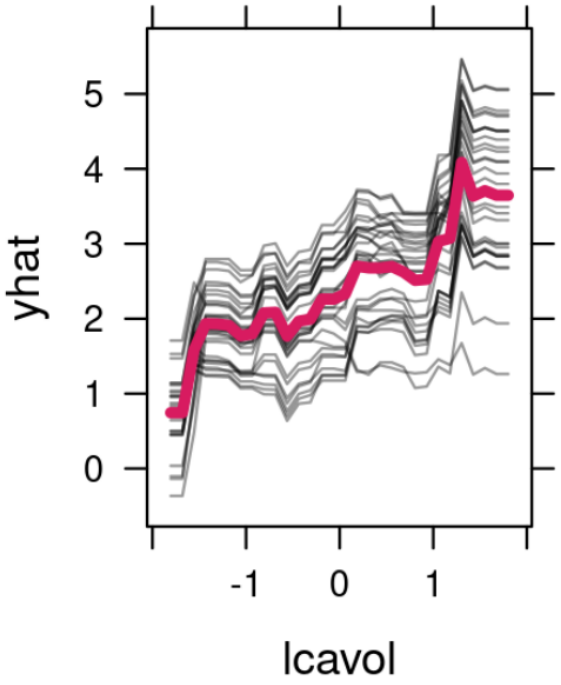


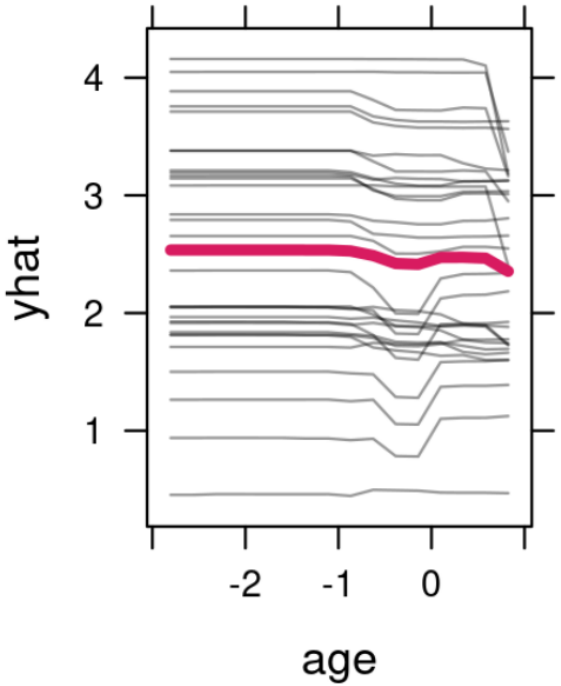
Amazon

ICE - Individual conditional expectation

$$h_{i,j}(x) := g(x_{j,1}, \ldots, x_{j,i-1}, x, x_{j,i+1}, \ldots, x_{j,d})$$

ICE - Individual conditional expectation





Amazon

Lime

- Por que o algoritmo me forneceu a predição $g(\mathbf{x}^*)$ para a nova amostra $\mathbf{x}^*$?
- aproxima a solução $g(\mathbf{x}^*)$ *localmente* via uma regressão linear ajustada por lasso.

Amazon

Titulo