Explainable Boosting Machines

Also known as Linear Regression in Steroids

Raquel
raquel.perez@bsc.es

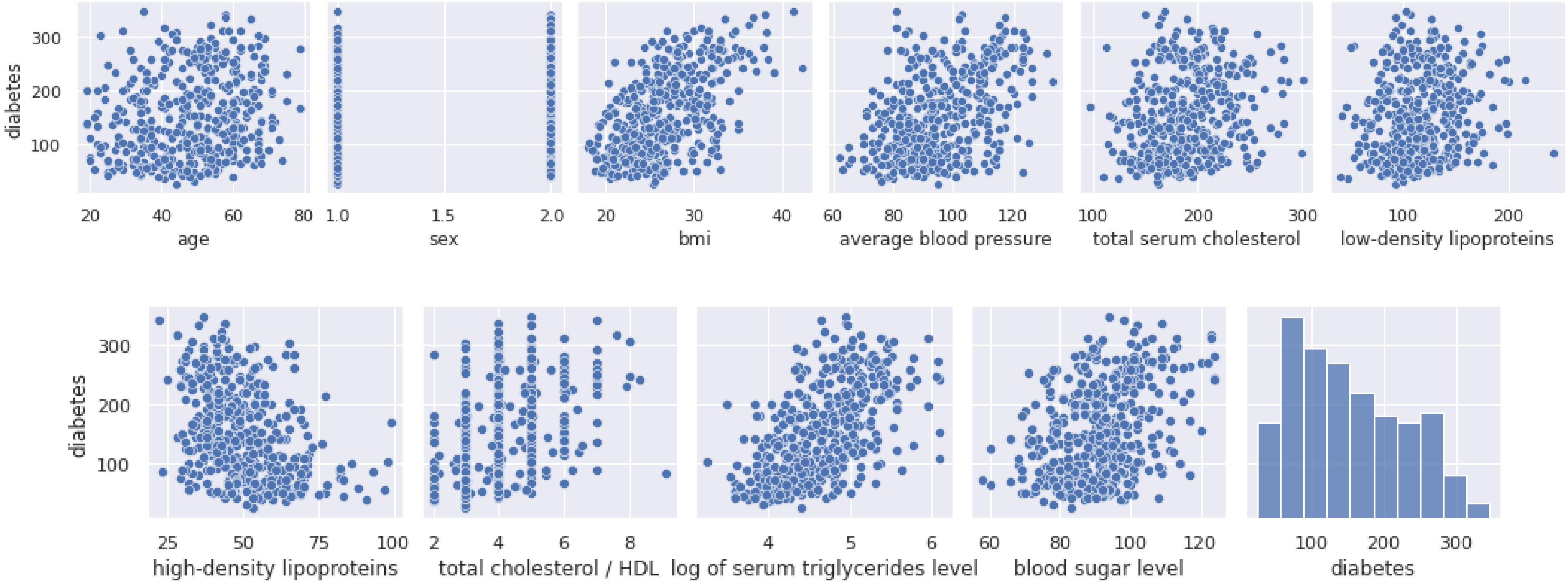
other collaborators

Linear Regression

Predicting Diabetes Progression

	age	sex	bmi	average blood pressure	total serum cholesterol	low-density lipoproteins	high-density lipoproteins	total cholesterol / HDL	log of serum triglycerides level	blood sugar level	diabetes
count	442.000	442.000	442.000	442.000	442.000	442.000	442.000	442.000	442.000	442.000	442.000
mean	-0.000	0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	0.000	0.000	152.133
std	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	77.093
min	-0.107	-0.045	-0.090	-0.112	-0.127	-0.116	-0.102	-0.076	-0.126	-0.138	25.000
25%	-0.037	-0.045	-0.034	-0.037	-0.034	-0.030	-0.035	-0.039	-0.033	-0.033	87.000
50%	0.005	-0.045	-0.007	-0.006	-0.004	-0.004	-0.007	-0.003	-0.002	-0.001	140.500
75%	0.038	0.051	0.031	0.036	0.028	0.030	0.029	0.034	0.032	0.028	211.500
max	0.111	0.051	0.171	0.132	0.154	0.199	0.181	0.185	0.134	0.136	346.000

Diabetes Progression



Linear Regression

$$y = \epsilon + w_0 x_0 + w_1 x_1 + \cdots + w_n x_n$$

Assumptions

- Target variable is Gaussian.
- Variables do not interact between them.
- There is a linear relationship between variables and target.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

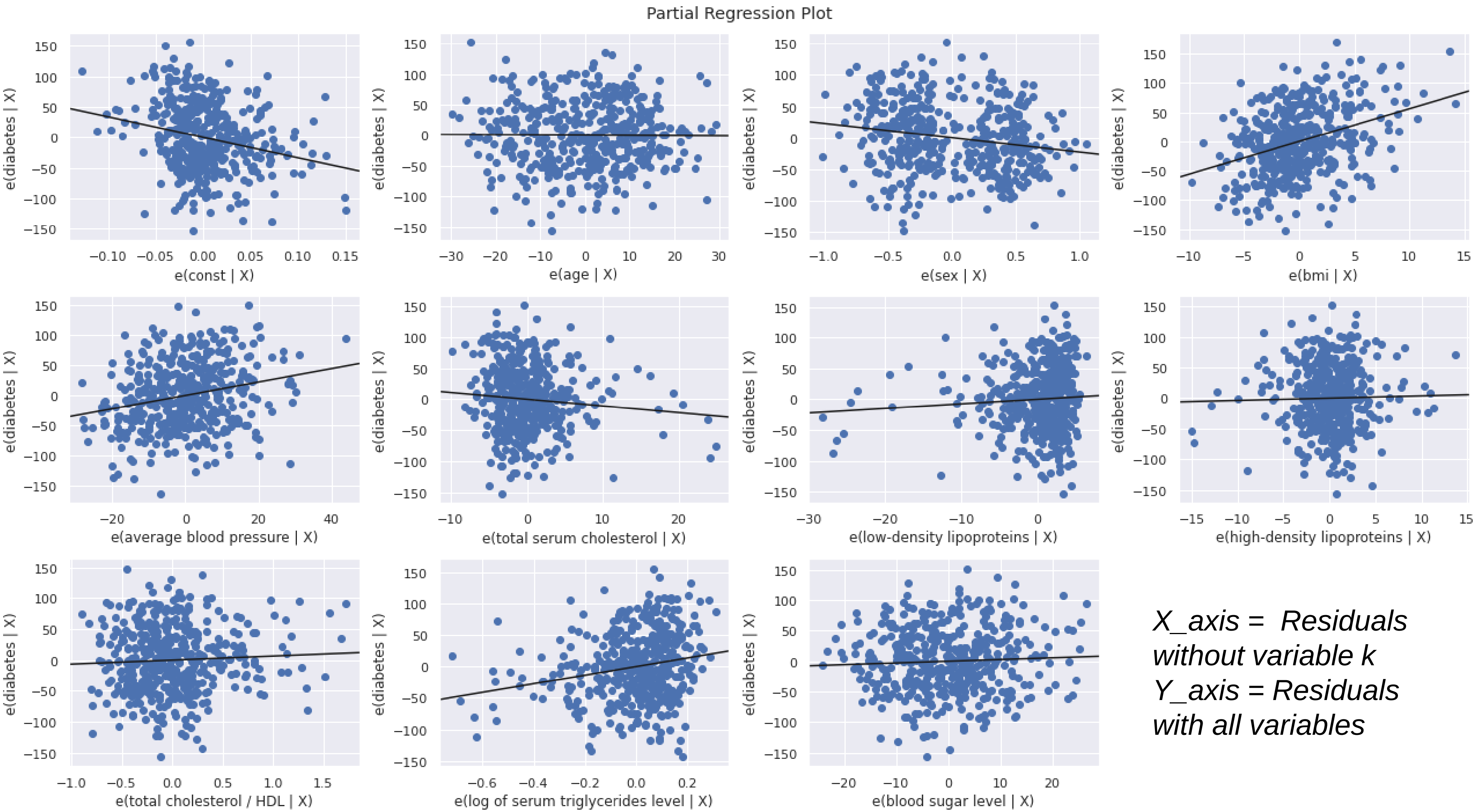
HIGH PERFORMANCE ARTIFICIAL INTELLIGENCE

Linear Regression Interpretability

- Increasing a variable by one unit increases the target by its weight.
- You can compute the feature importance (t statistic) as the weight divided by its standard error. The bigger the std error the bigger the uncertainty about the feature weight.

Linear Regression

	coef	std err	t	P>|t|	[0.025	0.975]
const	-334.5671	67.455	-4.960	0.000	-467.148	-201.986
age	-0.0364	0.217	-0.168	0.867	-0.463	0.390
sex	-22.8596	5.836	-3.917	0.000	-34.330	-11.389
bmi	5.6030	0.717	7.813	0.000	4.194	7.012
average blood pressure	1.1168	0.225	4.958	0.000	0.674	1.560
total serum cholesterol	-1.0900	0.573	-1.901	0.058	-2.217	0.037
low-density lipoproteins	0.7465	0.531	1.406	0.160	-0.297	1.790
high-density lipoproteins	0.3720	0.782	0.475	0.635	-1.166	1.910
total cholesterol / HDL	6.5338	5.959	1.097	0.273	-5.178	18.245
log of serum triglycerides level	68.4831	15.670	4.370	0.000	37.685	99.282
blood sugar level	0.2801	0.273	1.025	0.306	-0.257	0.817



X_axis = Residuals without variable k
Y_axis = Residuals with all variables

Linear Regression Extensions

- Generalized Linear Model
- Adding interactions Manually
- Generalized Additive Model
- Explainable Boosting Machines

Generalized Linear Models

$$g(y) = \epsilon + w_0 x_0 + w_1 x_1 + \cdots + w_n x_n$$

Assumptions

- ~~Target variable is Gaussian.~~
- Variables do not interact between them.
- There is a linear relationship between variables and target.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

HIGH PERFORMANCE ARTIFICIAL INTELLIGENCE

Adding Interactions Manually

$$y = \epsilon + \sum w_i x_i + \sum w_{i,j} x_i x_j$$

Assumptions

- Target variable is Gaussian.
- ~~Variables do not interact between them~~.
- There is a linear relationship between variables and target.

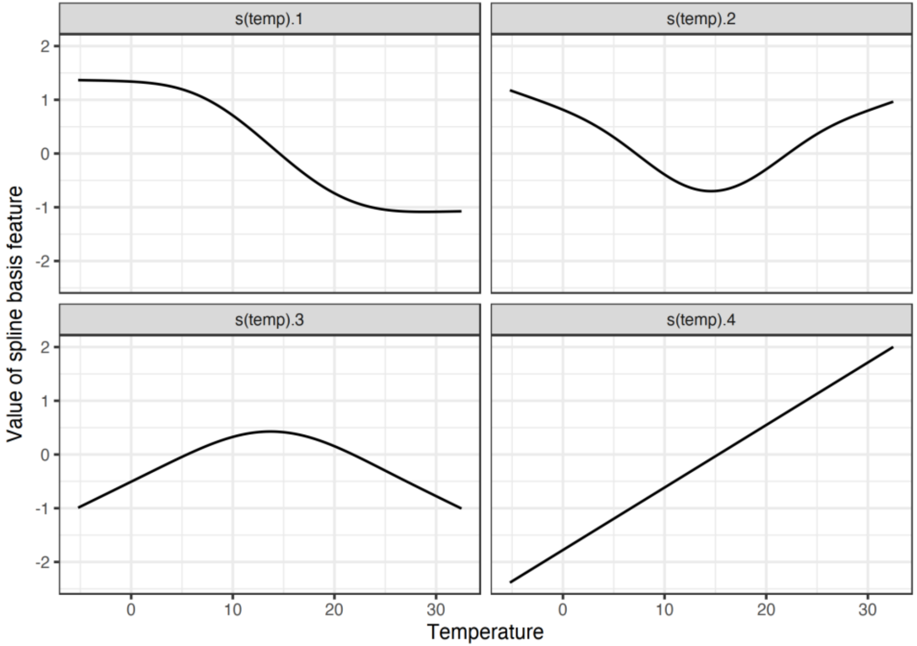
Generalized Additive Models

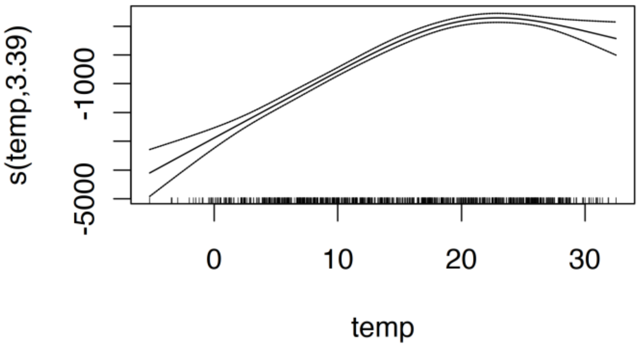
$$y = \epsilon + f_0(x_0) + f_1(x_1) + \cdots + f_n(x_n)$$

Assumptions

- Target variable is Gaussian.
- Variables do not interact between them.
- ~~There is a linear relationship between variables and target.~~

Example of splines





Figures from [1]

Explainable Boosting Machines [2]

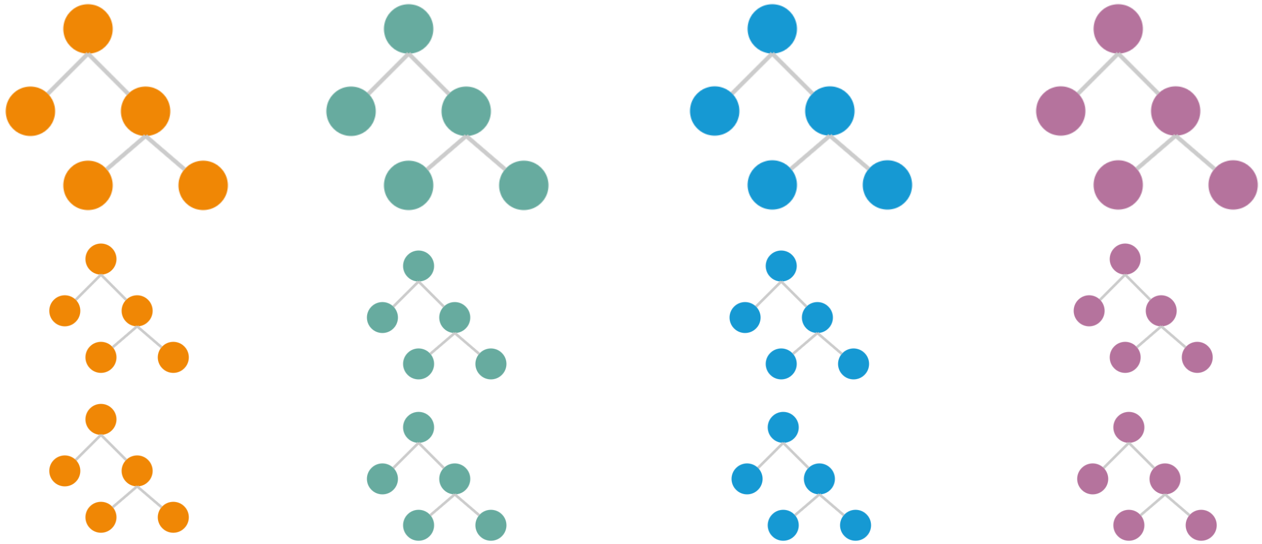
$$g(y) = \epsilon + \sum f_i(x_i) + \sum f_{i,j}(x_i, x_j)$$

Assumptions

- ~~Target variable is Gaussian.~~
- ~~Variables do not interact between them.~~
- ~~There is a linear relationship between variables and target.~~

Functions are estimated using small trees

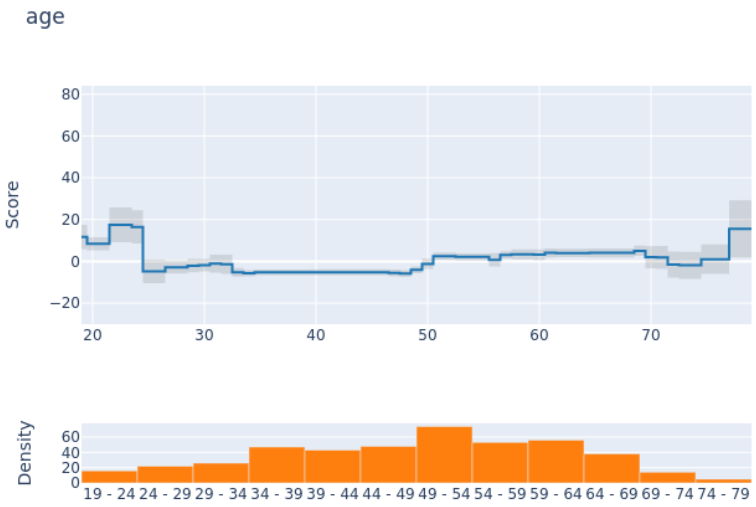
$$g(y) = f_1(x_1) + f_2(x_2) + f_3\,(x_3) + f_4(x_4)$$



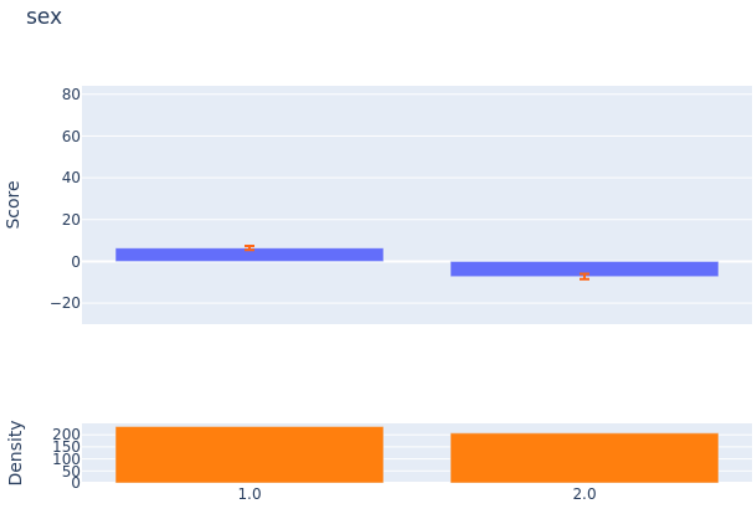
EBM Example

https://github.com/HPAI-BSC/interpretable_models_seminar/blob/main/src/EBM.ipynb

EBM Results

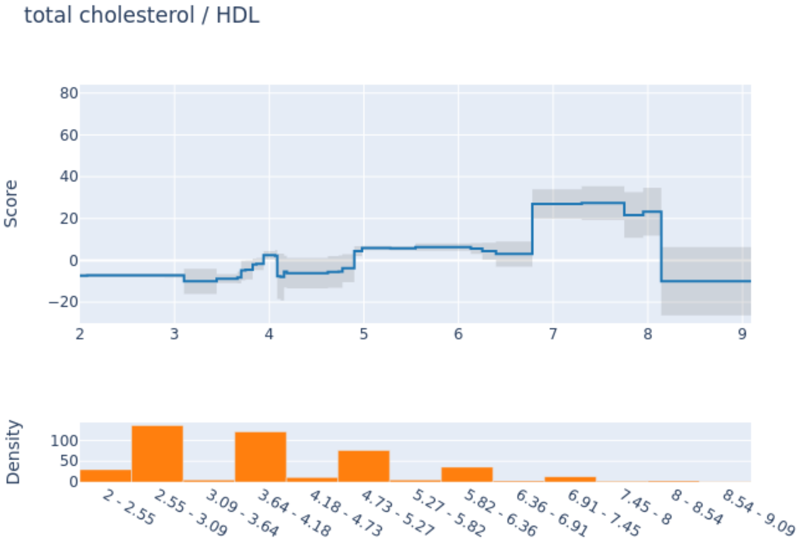


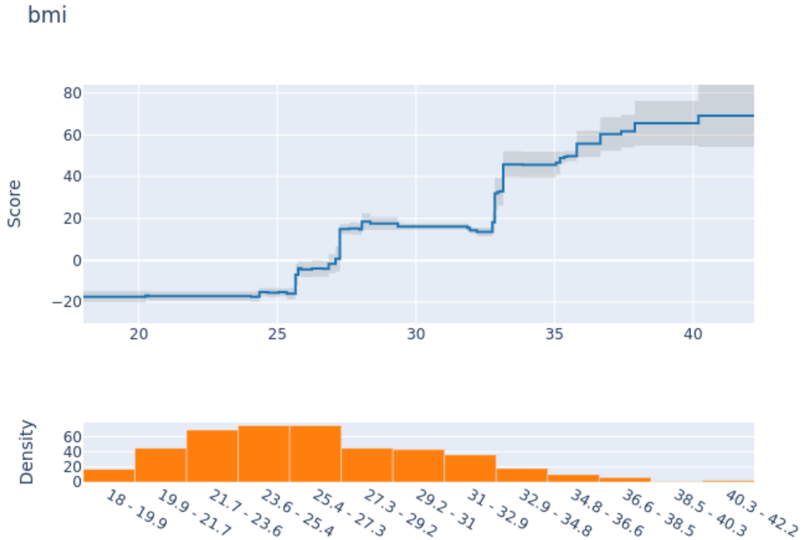
age
Score
80
60
40
20
0
−20
20
30
40
50
60
70
Density
60
40
20
0
19 - 24 24 - 29 29 - 34 34 - 39 39 - 44 44 - 49 49 - 54 54 - 59 59 - 64 64 - 69 69 - 74 74 - 79



sex
Score
80
60
40
20
0
−20
Density
200
150
100
50
0
1.0
2.0

EBM Results

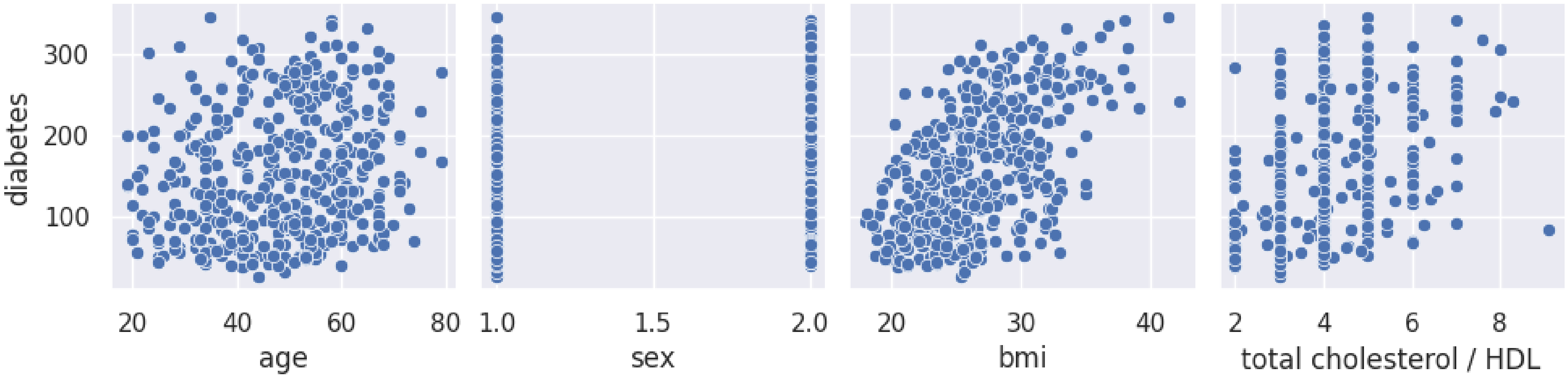




Barcelona Supercomputing Center
Centro Nacional de Supercomputación

HIGH PERFORMANCE ARTIFICIAL INTELLIGENCE

EBM Results



diabetes
300
200
100
20
40
60
80
age
1.0
1.5
2.0
sex
20
30
40
bmi
2
4
6
8
total cholesterol / HDL

References

[1] Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.
[2] Lou, Y., Caruana, R., Gehrke, J., & Hooker, G. (2013, August). Accurate intelligible models with pairwise interactions. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 623-631).

Barcelona Supercomputing Center
Centro Nacional de Supercomputación
HIGH PERFORMANCE ARTIFICIAL INTELLIGENCE

Raquel Pérez-Arnal
raquel.perez@bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

HIGH PERFORMANCE ARTIFICIAL INTELLIGENCE