Explainable Boosting Machines

Also known as Linear Regression on Steroids

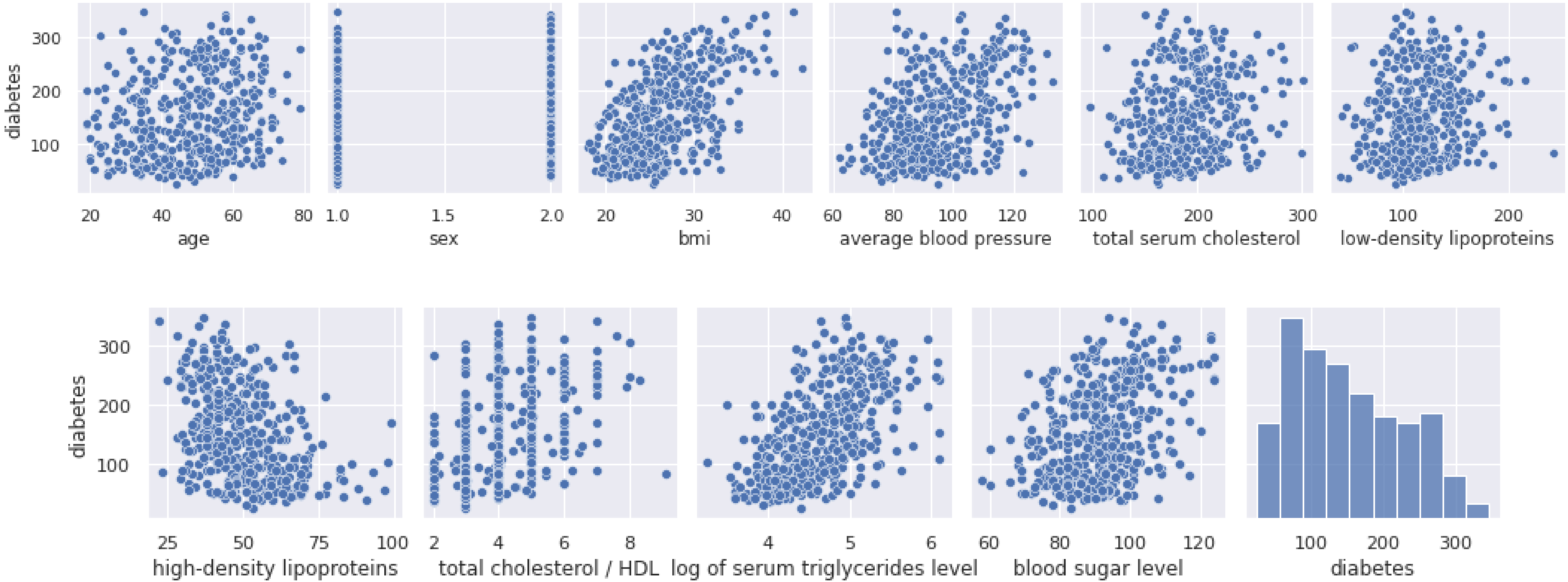
Raquel
raquel.perez@bsc.es

Linear Regression

Predicting Diabetes Progression

	age	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
mean	48.518	26.376	94.647	189.140	115.439	49.788	4.070	4.641	91.260	152.133
std	13.109	4.418	13.831	34.608	30.413	12.934	1.290	0.522	11.496	77.093
min	19.000	18.000	62.000	97.000	41.600	22.000	2.000	3.258	58.000	25.000
25%	38.250	23.200	84.000	164.250	96.050	40.250	3.000	4.277	83.250	87.000
50%	50.000	25.700	93.000	186.000	113.000	48.000	4.000	4.620	91.000	140.500
75%	59.000	29.275	105.000	209.750	134.500	57.750	5.000	4.997	98.000	211.500
max	79.000	42.200	133.000	301.000	242.400	99.000	9.090	6.107	124.000	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

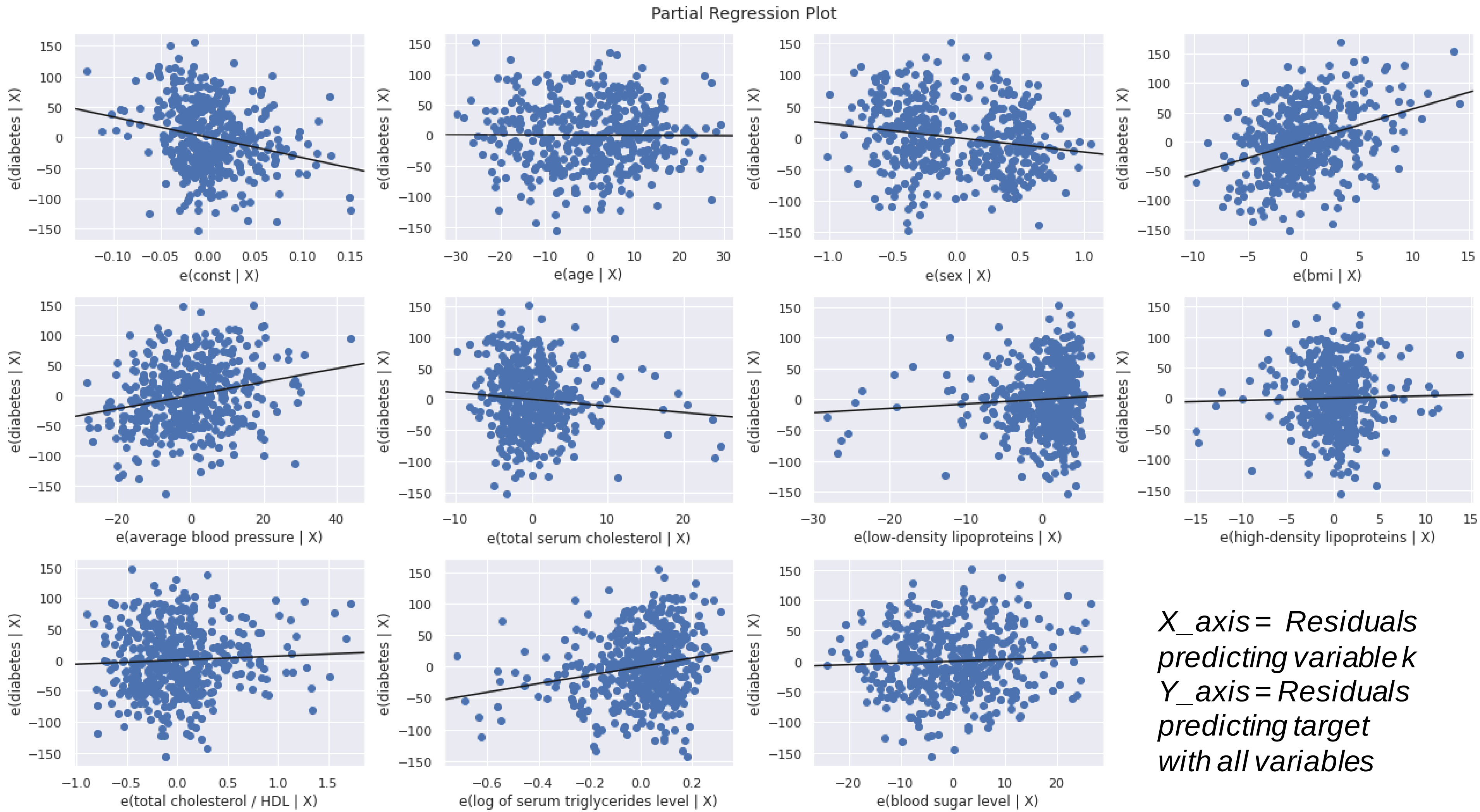
HIGH PERFORMANCE ARTIFICIAL INTELLIGENCE

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 predicting variable k
Y_axis = Residuals predicting target 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.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación
HIGH PERFORMANCE ARTIFICIAL INTELLIGENCE

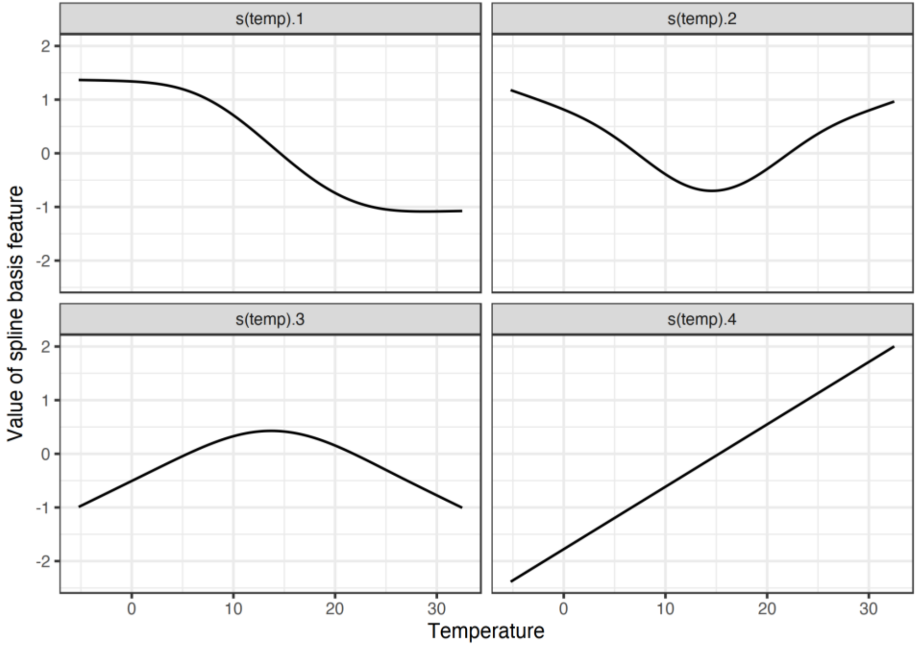
Generalized Additive Models

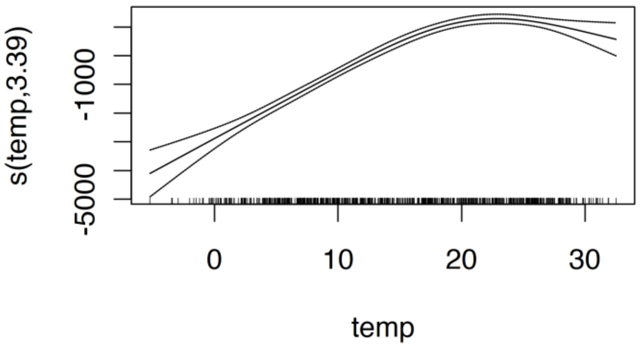
$$y = \epsilon + f_0(x_0) + f_1(x_1) + \cdots + f_n(x_n)$$

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.~~

Example of a Decision Tree

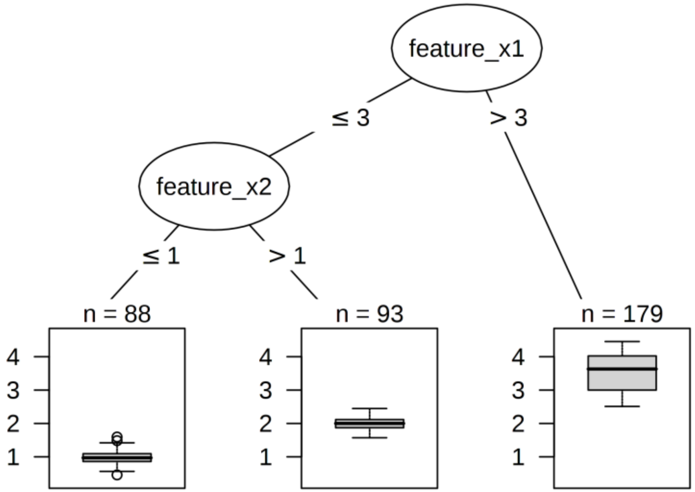
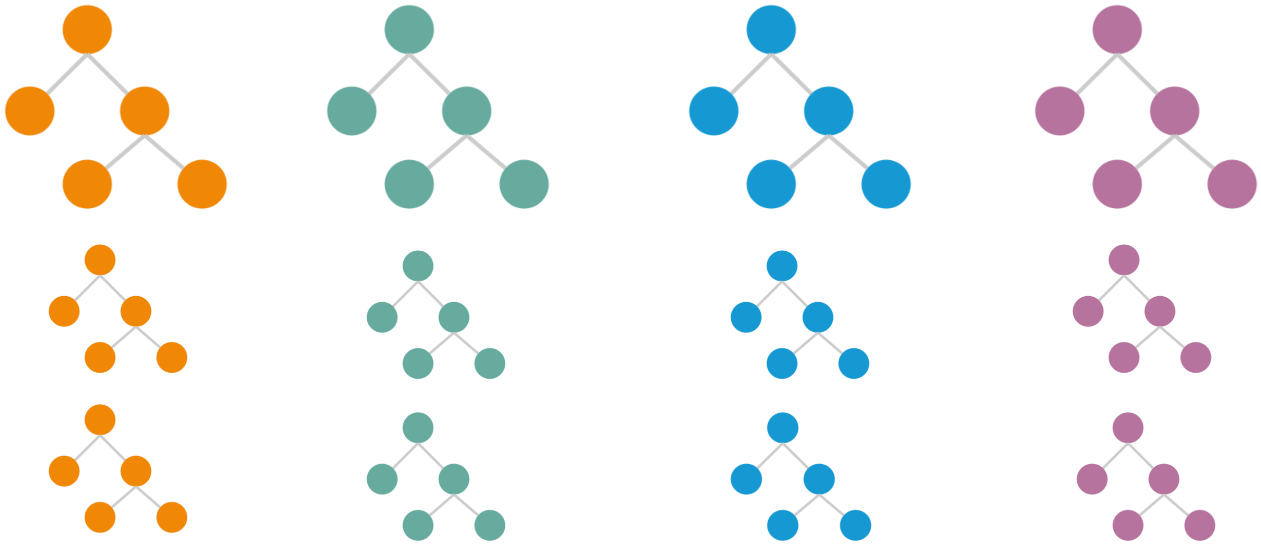


Figure from [1]

Functions are estimated using small trees

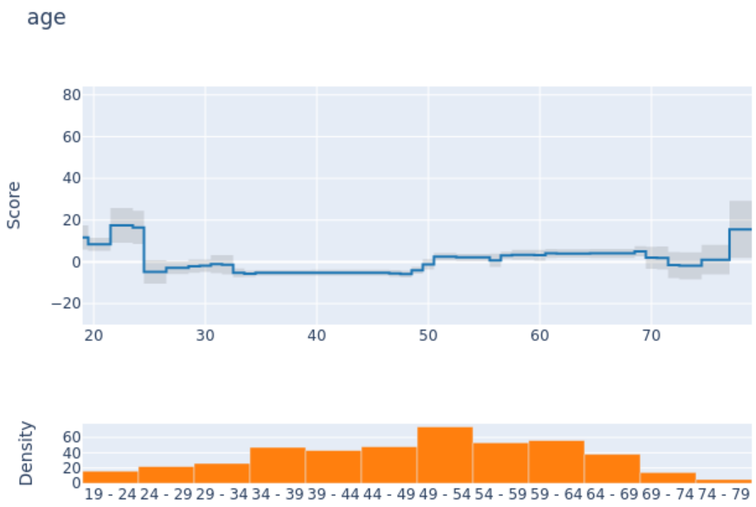
$$g(y) = f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4)$$



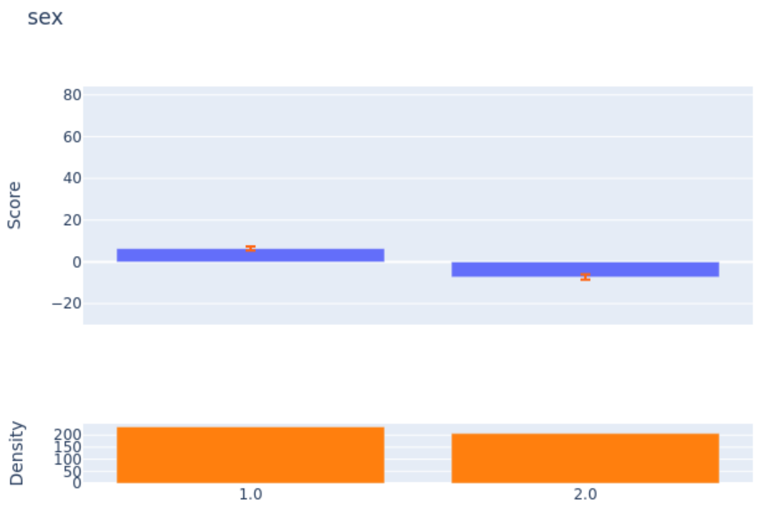
EBM Example

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

EBM Results

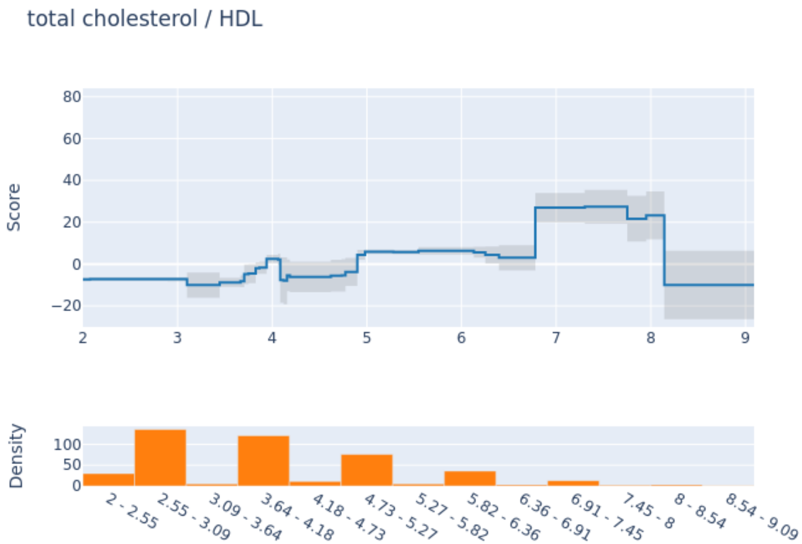


age
80
60
40
20
0
−20
Score
20
30
40
50
60
70
Density
60
40
20
0
19 - 24
24 - 29
29 - 34
34 - 39
39 - 44
44 - 49
49 - 54
54 - 59
59 - 64
64 - 69
69 - 74
74 - 79

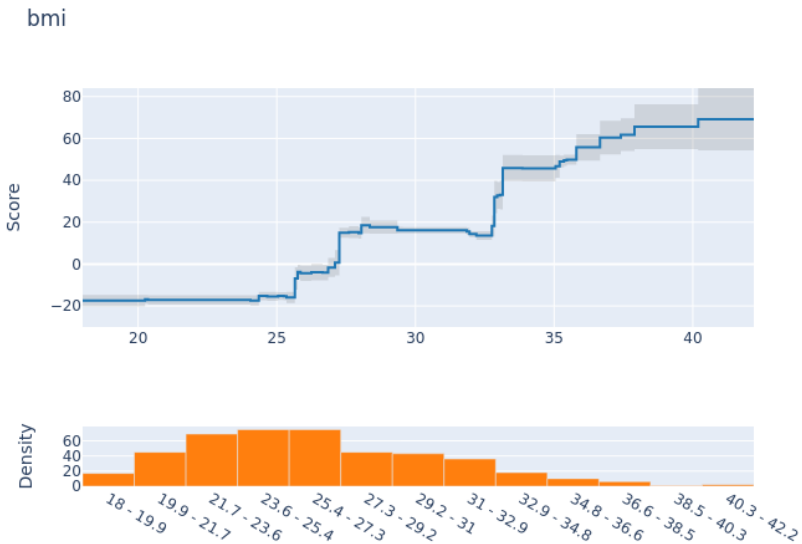


sex
80
60
40
20
0
−20
Score
Density
200
150
100
50
0
1.0
2.0

EBM Results

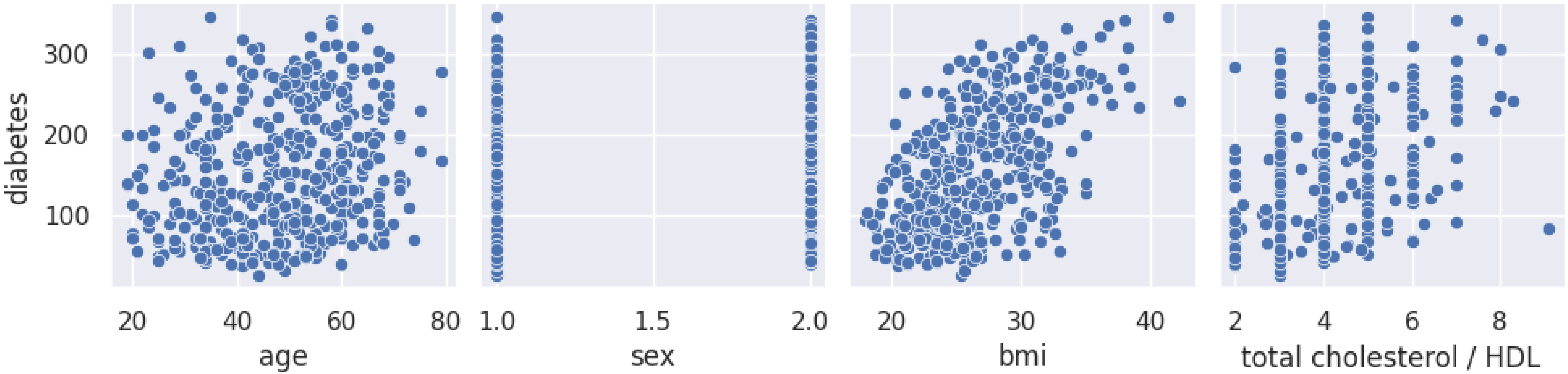


total cholesterol / HDL
Score
80
60
40
20
0
−20
2
3
4
5
6
7
8
9
Density
100
50
0
2 - 2.55
2.55 - 3.09
3.09 - 3.64
3.64 - 4.18
4.18 - 4.73
4.73 - 5.27
5.27 - 5.82
5.82 - 6.36
6.36 - 6.91
6.91 - 7.45
7.45 - 8
8 - 8.54
8.54 - 9.09



bmi
Score
80
60
40
20
0
−20
20
25
30
35
40
Density
60
40
20
0
18 - 19.9
19.9 - 21.7
21.7 - 23.6
23.6 - 25.4
25.4 - 27.3
27.3 - 29.2
29.2 - 31
31 - 32.9
32.9 - 34.8
34.8 - 36.6
36.6 - 38.5
38.5 - 40.3
40.3 - 42.2

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).

Raquel Pérez-Arnal

raquel.perez@bsc.es

Barcelona Supercomputing Center

Centro Nacional de Supercomputación

HIGH PERFORMANCE ARTIFICIAL INTELLIGENCE