

Homework (opcional)

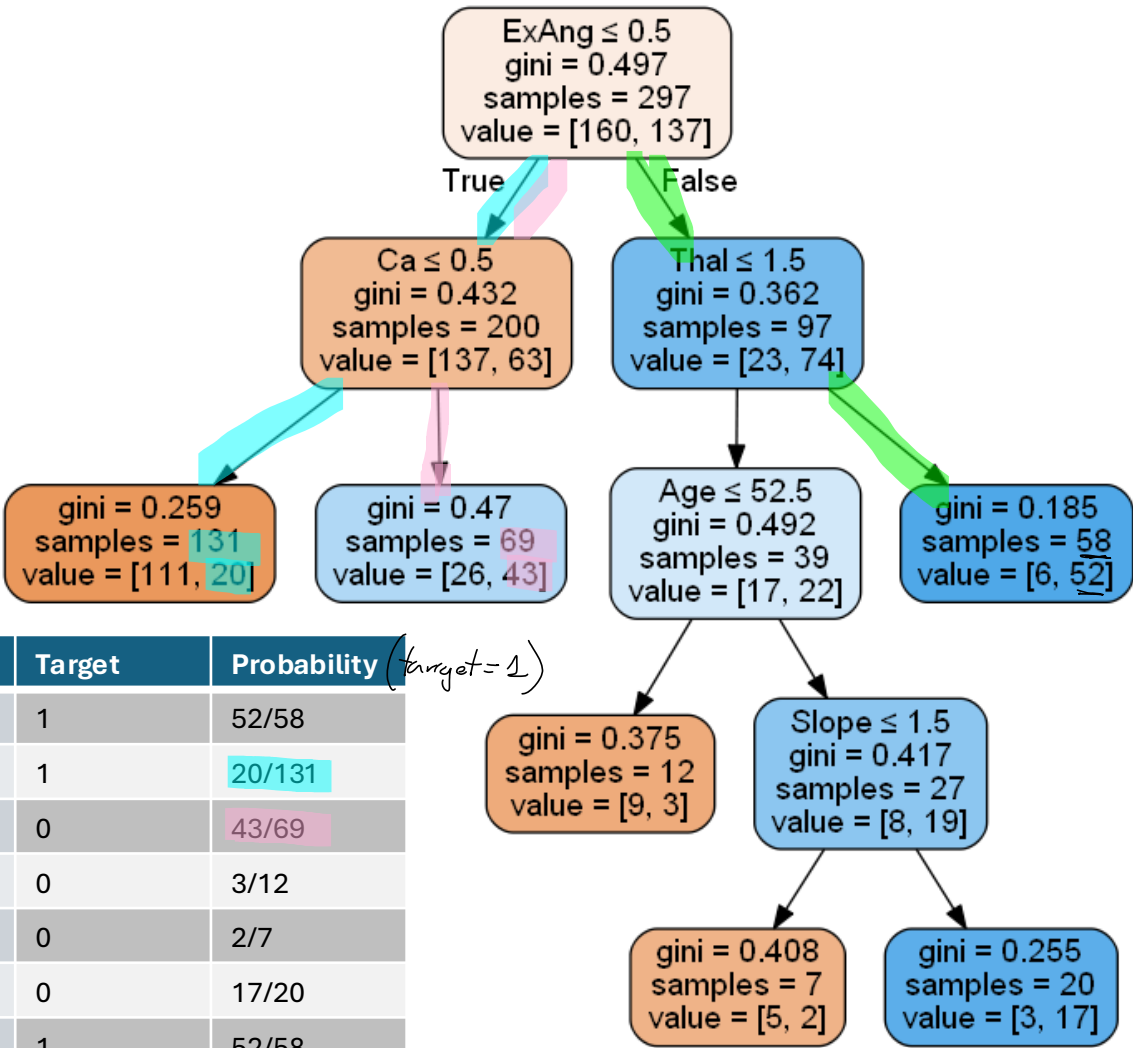
A partir del siguiente árbol, indica:

- Profundidad del árbol: **4**
- Min samples split que utilizamos en el entrenamiento: **7**
- Con los siguientes datos, dibuja la ROC Curve resultante del modelo.

Pistas:

- **Calcula** la probabilidad que predice el árbol para cada fila
- Ordena las filas de mayor a menor probabilidad
- Calcula el TPR y FPR para diferentes cortes del threshold (al menos 5 cortes distintos)
- Dibuja estos valores en el gráfico de TPR/FPR

Id fila	ExAng	Ca	Thal	Age	Slope	Target	Probability
1	1	0	8	65	1	1	52/58
2	0	0	5	25	8	1	20/131
3	0	1	2	55	5	0	43/69
4	1	1	1	45	4	0	3/12
5	1	0	0	60	1	0	2/7
6	1	0	0	65	3	0	17/20
7	1	1	3	28	7	1	52/58
8	0	1	0	39	2	1	43/69
9	1	0	0	87	4	1	17/20
10	1	1	1	15	8	0	3/12



recl

Id fila	ExAng	Ca	Thal	Age	Slope	Target	Probability	Th = 0,86	Th = 0,8	Th = 0,6	Th = 0,28	Th = 0,11
1	1	0	8	65	1	1 +	52/58 (89%)	1 TP	1	1	1 TP	1
2	0	0	5	25	8	1 +	20/131 (15%)	0 FN	0	0	0 FN	1
3	0	1	2	55	5	0 -	43/69 (62%)	0 TN	0	1	1 FP	1
4	1	1	1	45	4	0 -	3/12 (25%)	0 TN	0	0	0 TN	1
5	1	0	0	60	1	0 -	2/7 (29%)	0 TN	0	0	1 FP	1
6	1	0	0	65	3	0 -	17/20 (85%)	0 TN	1	1	1 FP	1
7	1	1	3	28	7	1 +	52/58 (89%)	1 TP	1	1	1 TP	1
8	0	1	0	39	2	1 +	43/69 (62%)	0 FN	0	1	1 TP	1
9	1	0	0	87	4	1 +	17/20 (85%)	0 FN	1	1	1 TP	1
10	1	1	1	15	8	0 -	3/12 (25%)	0 TN	0	0	0 TN	
						TP+FN = 5	TPR	2/5 = 0,4	3/5 = 0,6	4/5 = 0,8	4/5 = 0,8	5/5 = 1
						TN+FP = 5	FPR	0/5 = 0	1/5 = 0,2	2/5 = 0,4	3/5 = 0,6	5/5 = 1

$$TPR = \frac{TP}{TP + FN} \quad \left. \begin{array}{l} \text{Gijó} \\ \text{reclas} \end{array} \right\}$$

$$FPR = \frac{FP}{TN + FP} \quad \left. \begin{array}{l} \text{Gijó} \\ \text{reclas} \end{array} \right\}$$

$prob \geq 0,86 \rightarrow pred = 1$
 $prob < 0,86 \rightarrow pred = 0$

