



Tecnológico de Monterrey

**Momento de Retroalimentación: Módulo 2 Uso de framework o biblioteca de
aprendizaje máquina para la implementación de una solución**

María Fernanda Ramírez Barragán

A01750879

ITC

Fecha: 12/09/22

Instituto Tecnológico y de Estudios Superiores de Monterrey

TC3006C.101

Árbol de decisión sklearn

Al utilizar un árbol de decisión en un dataset con una Y categórica, el resultado (sin alterar de primeras hiperparametros) ha sido de un accuracy de 0.78, es decir, que tiene un error de 0.22 ya que el valor esperado (o 'perfecto') es de 1.0, que significa que el 100% de las predicciones han sido correctas.

En este ejemplo el alpha es de 0.0 (valor default), podemos ver que el árbol tiene 164 y la distancia máxima entre la root y cualquier hoja es de 24, sin embargo.

```
Exactitud del modelo: 0.784375
/-----/
Exactitud del modelo (train): 1.0
/-----/
Distancia máxima entre la raíz y cualquier hoja = 24
/-----/
Número de hojas del arbol = 164
```

Sin embargo, al aumentar el alpha a 0.002, el número de hojas y las distancia máxima entre la raíz y cualquier hoja han disminuido, mientras que la exactitud de nuestro modelo ha aumentado un 5.3%, dando como resultado un accuracy de 83.75%

```
Exactitud del modelo: 0.8375
/-----/
Exactitud del modelo (train): 0.90852228303362
/-----/
Distancia máxima entre la raíz y cualquier hoja = 9
/-----/
Número de hojas del arbol = 29
```

Así se ven los primeros resultados de ambas pruebas:

Alpha: 0.0

/-----/		
Predicciones:	Valor esperado	Valor obtenido
803	Regular	Regular
124	Regular	Regular
350	Regular	Regular
682	Regular	Regular
1326	Regular	Regular
976	Regular	Regular
1493	Regular	Regular
706	Regular	Regular
613	Regular	Regular
1587	Regular	Regular
513	Buena	Buena
1505	Mala	Regular
1525	Regular	Regular
1568	Regular	Regular
220	Regular	Regular
1010	Buena	Mala
30	Regular	Regular
942	Buena	Regular
481	Buena	Buena
135	Regular	Regular
65	Regular	Regular
99	Regular	Regular
1547	Regular	Regular
1311	Regular	Regular
29	Regular	Regular
745	Regular	Regular
538	Buena	Regular
168	Regular	Regular

Alpha: 0.002

/-----/		
Predicciones:	Valor esperado	Valor obtenido
803	Regular	Regular
124	Regular	Regular
350	Regular	Regular
682	Regular	Regular
1326	Regular	Regular
976	Regular	Regular
1493	Regular	Regular
706	Regular	Regular
613	Regular	Regular
1587	Regular	Regular
513	Buena	Buena
1505	Mala	Regular
1525	Regular	Regular
1568	Regular	Regular
220	Regular	Regular
1010	Buena	Buena
30	Regular	Regular
942	Buena	Regular
481	Buena	Buena
135	Regular	Regular
65	Regular	Regular
99	Regular	Regular
1547	Regular	Regular
1311	Regular	Regular
29	Regular	Regular
745	Regular	Regular
538	Buena	Regular
168	Regular	Regular

Regresión lineal sklearn

Para predecir el precio de una casa tomando en cuenta ciertas variables, utilice un modelo de regresión lineal, al entrenarlo y probarlo obtuve un accuracy de 0.68, aunque no es un valor lo suficientemente alto para decir que es un “buen” modelo, si tiene un accuracy mayor al 50%, esto nos dice que no es un resultado al azar y que si es capaz de dar una predicción.

```
Parametros utiles para el analisis del modelo
/-----/
Exactitud del modelo: 0.6810580555095946
/-----/
Exactitud del modelo (train): 0.558119727460709
/-----/
m (coeficientes)= [ 1.52963113 -3.06269441 -5.78692561  3.2188727  2.85510785 -0.4410086 ]
/-----/
b = 38.39154078549889
/-----/
Error medio cuadrado: 53.51
/-----/
Coeficiente de determinación (score): 0.68
/-----/
```

Finalmente, así es como se ven los valores esperados en comparación a las predicciones del modelo.

```
/-----/
Predicciones:      Valor esperado  Valor arrojado
358                45.1            47.886254
350                42.3            41.164046
373                52.2            44.273014
399                37.3            40.197615
369                22.8            27.513265
..                ...            ...
31                 25.0            40.368699
283                23.5            33.091396
157                35.6            42.327834
203                46.1            43.353985
208                26.2            28.439798
```