

TP4

November 16, 2021

```
<IPython.core.display.HTML object>
ToggleButton(value=False, description='Show code')
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
```

1 Parte I

1.1 1

1.2 2

1.3 3

1.4 4

```
0      0.500000
1      0.500000
2      0.500000
3      0.500000
4      0.000000
...
3314   0.000000
3315   0.000000
3316   0.333333
3317   0.333333
3318   0.333333
Name: Prop_menores, Length: 3319, dtype: float64
```

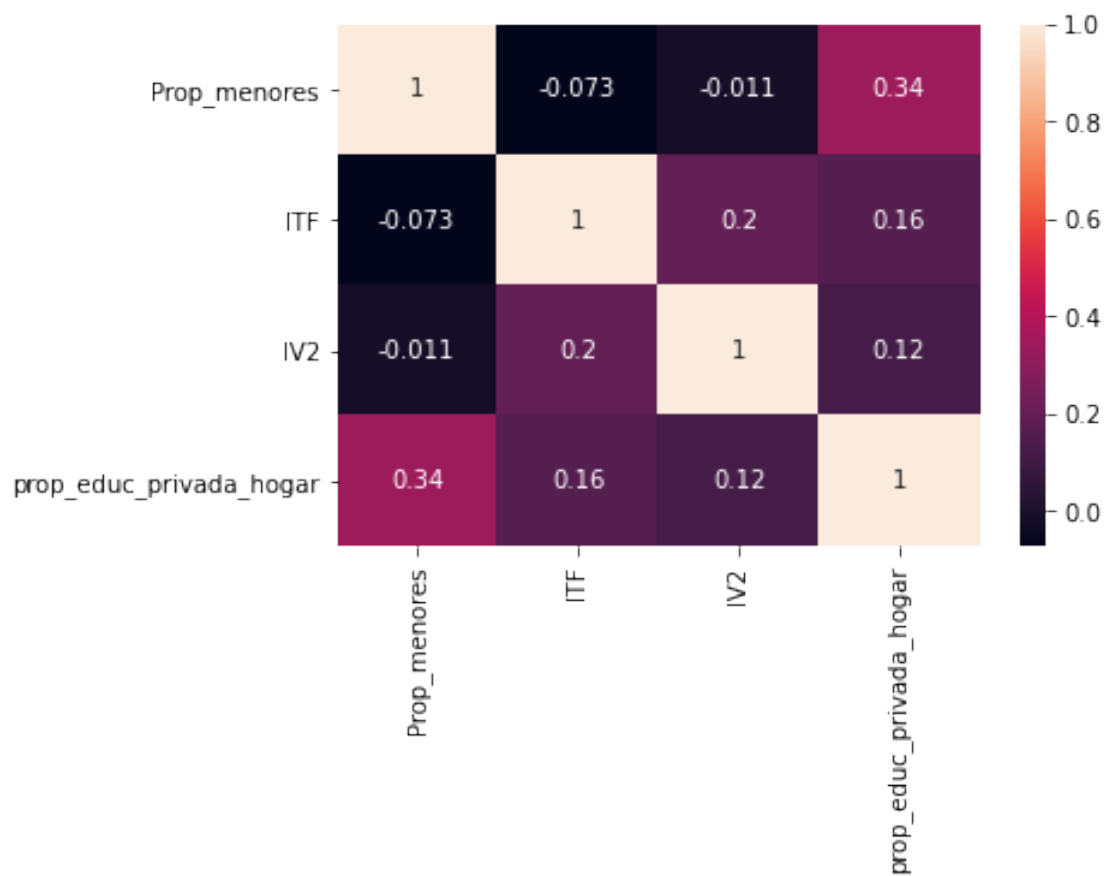
```
count    3319.000000
mean      0.107864
std       0.310255
min       0.000000
25%       0.000000
50%       0.000000
75%       0.000000
max       1.000000
Name: educ_privada, dtype: float64
```

```

count      3319.000000
mean        0.095917
std         0.181742
min         0.000000
25%         0.000000
50%         0.000000
75%         0.166667
max         1.000000
Name: prop_educ_privada_hogar, dtype: float64

```

1.5 5



1.6 6

1.7 7

1.8 8

```
PONDIH      0.365065
dtype: float64
```

Por lo tanto, la tasa de hogares bajo la línea de pobreza para el GBA es del 36,5%. Por su parte, para el periodo que estamos analizando el Indec reporta que el porcentaje de hogares por debajo de la línea de pobreza (LP) alcanzó el 31,2%; en los cuales reside el 40,6% de las personas.

2 Parte II

2.1 1

2.2 2

2.3 3

2.4 4

3 Parte III

3.1 1

3.2 2

3.3 3

	modelo	ecm	parámetro	auc	accuracy \
0	Lineal	0.197071	100.0	0.860638	0.802929
0	KNN	0.280959	15.0	0.761761	0.719041
0	Logit	0.174434	100.0	0.878208	0.825566
0	CART	0.173103	100.0	0.807253	0.826897
0	SVM	0.231691	10.0	NaN	0.768309
0	Bagging	0.202397	10000.0	0.852271	0.797603
0	RandomForest	0.127830	100000.0	0.932036	0.872170
0	Boosting	0.079893	10.0	0.969579	0.920107
0	Lasso	0.191744	10.0	0.858634	0.808256
0	Ridge	0.203728	1.0	0.869533	0.796272

	fpr \
0	[0.0, 0.0, 0.0, 0.00202020202020202, 0.0020202...
0	[0.0, 0.0, 0.0020491803278688526, 0.0040983606...
0	[0.0, 0.0, 0.0, 0.0020408163265306124, 0.00204...
0	[0.0, 0.13253012048192772, 1.0]
0	NaN
0	[0.0, 0.03143418467583497, 0.06679764243614932...
0	[0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, ...
0	[0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0019841269841...
0	[0.0, 0.0, 0.0, 0.001949317738791423, 0.001949...
0	[0.0, 0.0, 0.0, 0.001996007984031936, 0.001996...

```

                                tpr
0  [0.0, 0.00390625, 0.12890625, 0.12890625, 0.22...
0  [0.0, 0.0076045627376425855, 0.015209125475285...
0  [0.0, 0.0038314176245210726, 0.137931034482758...
0  [0.0, 0.7470355731225297, 1.0]
0  NaN
0  [0.0, 0.3925619834710744, 0.47520661157024796,...
0  [0.0, 0.007722007722007722, 0.0115830115830115...
0  [0.0, 0.06882591093117409, 0.10121457489878542...
0  [0.0, 0.004201680672268907, 0.0672268907563025...
0  [0.0, 0.004, 0.072, 0.072, 0.088, 0.088, 0.12,...

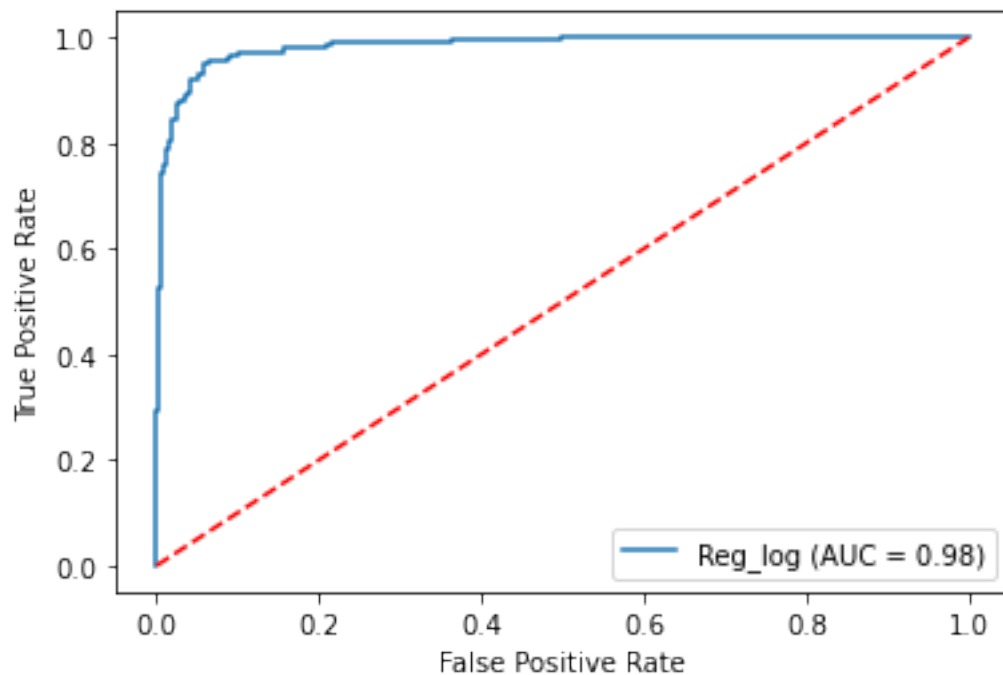
```

El método que predice mejor es Boosting, dado que presenta el mínimo error cuadrático medio.

3.4 4

Logramos mejorar las predicciones con respecto al TP3 dado que pasamos de un ecm 0.11 a uno de 0.079.

3.5 5



0.43137254901960786

El 43,13% de los hogares son pobres dentro de la base no respondieron.