

Sex	
Male	→ 0
Female	→ 1

→ { 'M': 0, 'F': 1 }

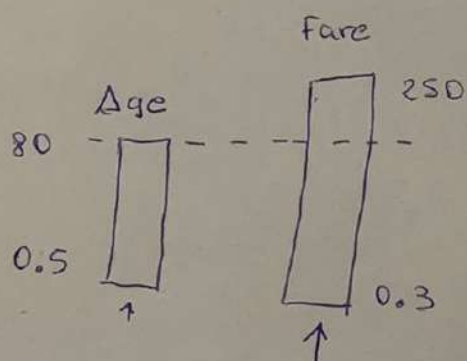
Sex	Sex - m	Sex - f
male	1	0
female	0	1
female	0	1
male	1	0

pd.get_dummies()

drop_first = True

Pclass	Pcl 1	Pcl 2	Pcl 3
1	1	0	0
2	0	1	0
3	0	0	1
2	0	1	0
1	1	0	0

Ahora tenemos 1 col mas por lo que habra que hacer lo mismo en el test.csv



Escalado de datos

Standard Scaler

MinMax Scaler

80
79
17
≡
5
≡
6

→ NaN
m.v

Cuando son numero se calcula la media

.5
6
Q
S
S
•

Cuando son strings se suele poner el que mas se repite

→ S