



# A Concrete Strength Regression

Juan Carlos Soriano (1493037) - Jorge Giménez (1493035) - Roger Boadella ( 1494330)

Dataset link: <https://www.kaggle.com/maajdl/yeh-concret-data>



# Apartado C: Exploración del dataset

# Descripción estadística

	cement	slag	flyash	water	superplasticizer	coarseaggregate	fineaggregate	age	csMPa
count	1030.000	1030.000	1030.000	1030.000	1030.000	1030.000	1030.000	1030.000	1030.000
mean	281.168	73.896	54.188	181.567	6.205	972.919	773.580	45.662	35.818
std	104.506	86.279	63.997	21.354	5.974	77.754	80.176	63.170	16.706
min	102.000	0.000	0.000	121.800	0.000	801.000	594.000	1.000	2.330
25%	192.375	0.000	0.000	164.900	0.000	932.000	730.950	7.000	23.710
50%	272.900	22.000	0.000	185.000	6.400	968.000	779.500	28.000	34.445
75%	350.000	142.950	118.300	192.000	10.200	1029.400	824.000	56.000	46.135
max	540.000	359.400	200.100	247.000	32.200	1145.000	992.600	365.000	82.600

# Primeras 5 entrades

	cement	slag	flyash	water	superplasticizer	coarseaggregate	fineaggregate	age	csMPa
0	540.000	0.000	0.000	162.000	2.500	1040.000	676.000	28	79.990
1	540.000	0.000	0.000	162.000	2.500	1055.000	676.000	28	61.890
2	332.500	142.500	0.000	228.000	0.000	932.000	594.000	270	40.270
3	332.500	142.500	0.000	228.000	0.000	932.000	594.000	365	41.050
4	198.600	132.400	0.000	192.000	0.000	978.400	825.500	360	44.300

# Últimas 5 entradas

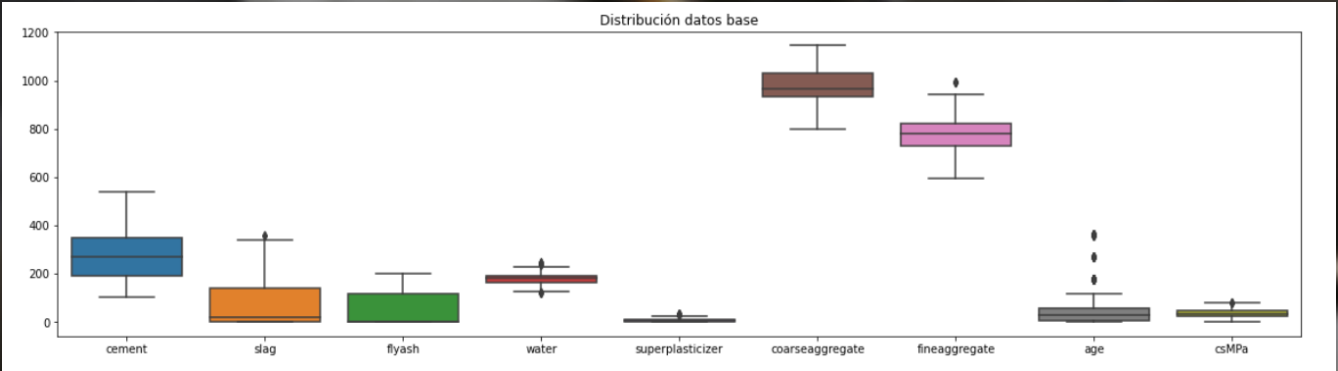
	cement	slag	flyash	water	superplasticizer	coarseaggregate	fineaggregate	age	csMPa
1025	276.400	116.000	90.300	179.600	8.900	870.100	768.300	28	44.280
1026	322.200	0.000	115.600	196.000	10.400	817.900	813.400	28	31.180
1027	148.500	139.400	108.600	192.700	6.100	892.400	780.000	28	23.700
1028	159.100	186.700	0.000	175.600	11.300	989.600	788.900	28	32.770
1029	260.900	100.500	78.300	200.600	8.600	864.500	761.500	28	32.400



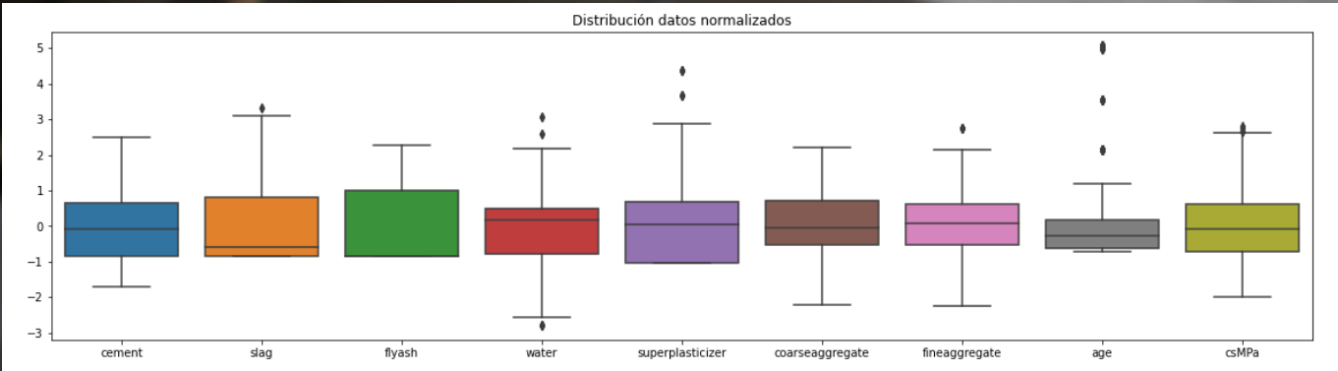
# Datos estandarizados

	cement	slag	flyash	water	superplasticizer	coarseaggregate	fineaggregate	age	csMPa
count	1030.000	1030.000	1030.000	1030.000	1030.000	1030.000	1030.000	1030.000	1030.000
mean	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	-0.000	0.000	-0.000
std	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
min	-1.715	-0.857	-0.847	-2.800	-1.039	-2.212	-2.241	-0.707	-2.006
25%	-0.850	-0.857	-0.847	-0.781	-1.039	-0.527	-0.532	-0.612	-0.725
50%	-0.079	-0.602	-0.847	0.161	0.033	-0.063	0.074	-0.280	-0.082
75%	0.659	0.801	1.002	0.489	0.669	0.727	0.629	0.164	0.618
max	2.478	3.311	2.281	3.066	4.354	2.214	2.733	5.058	2.802

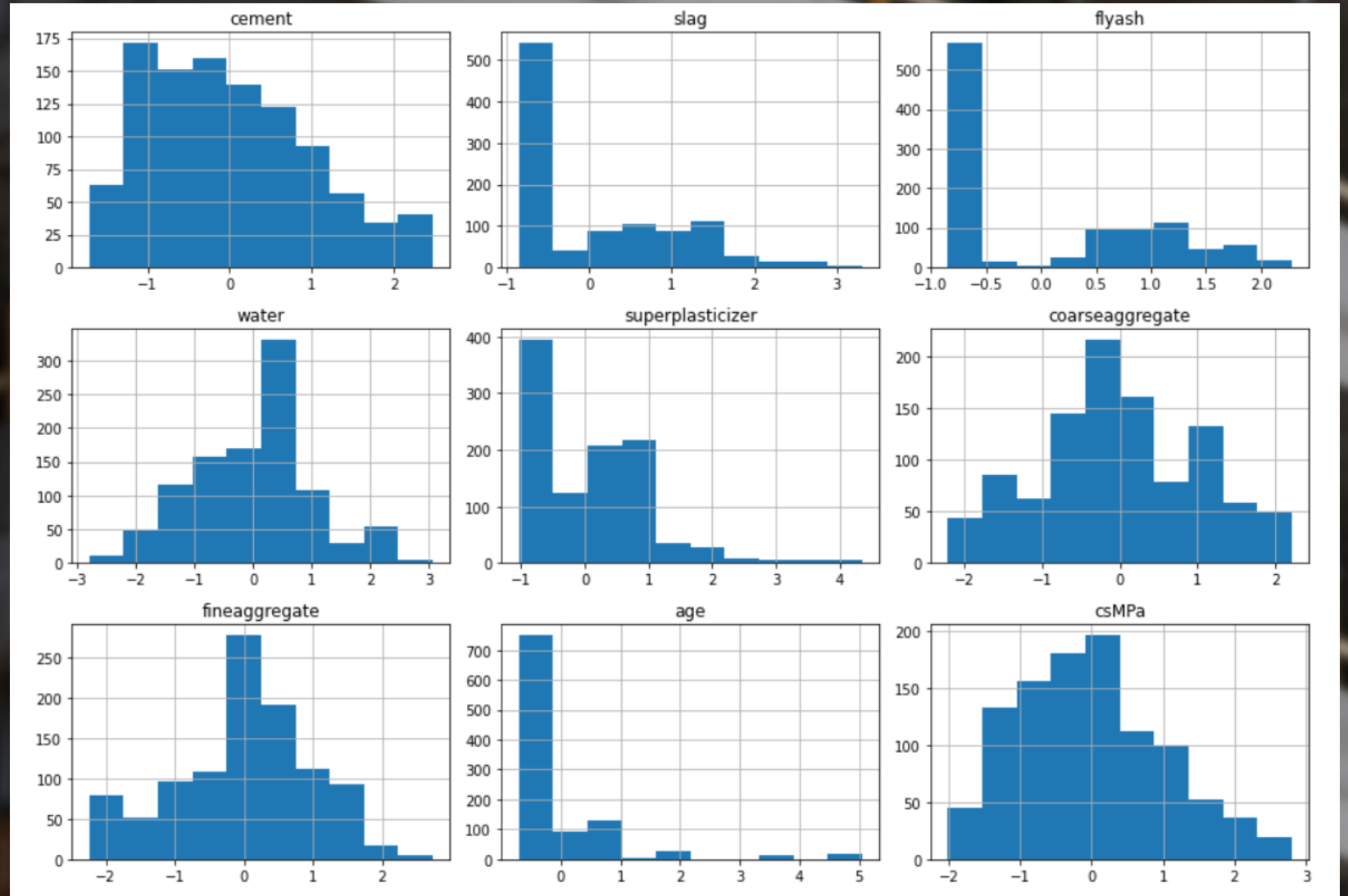
# Distribución anterior



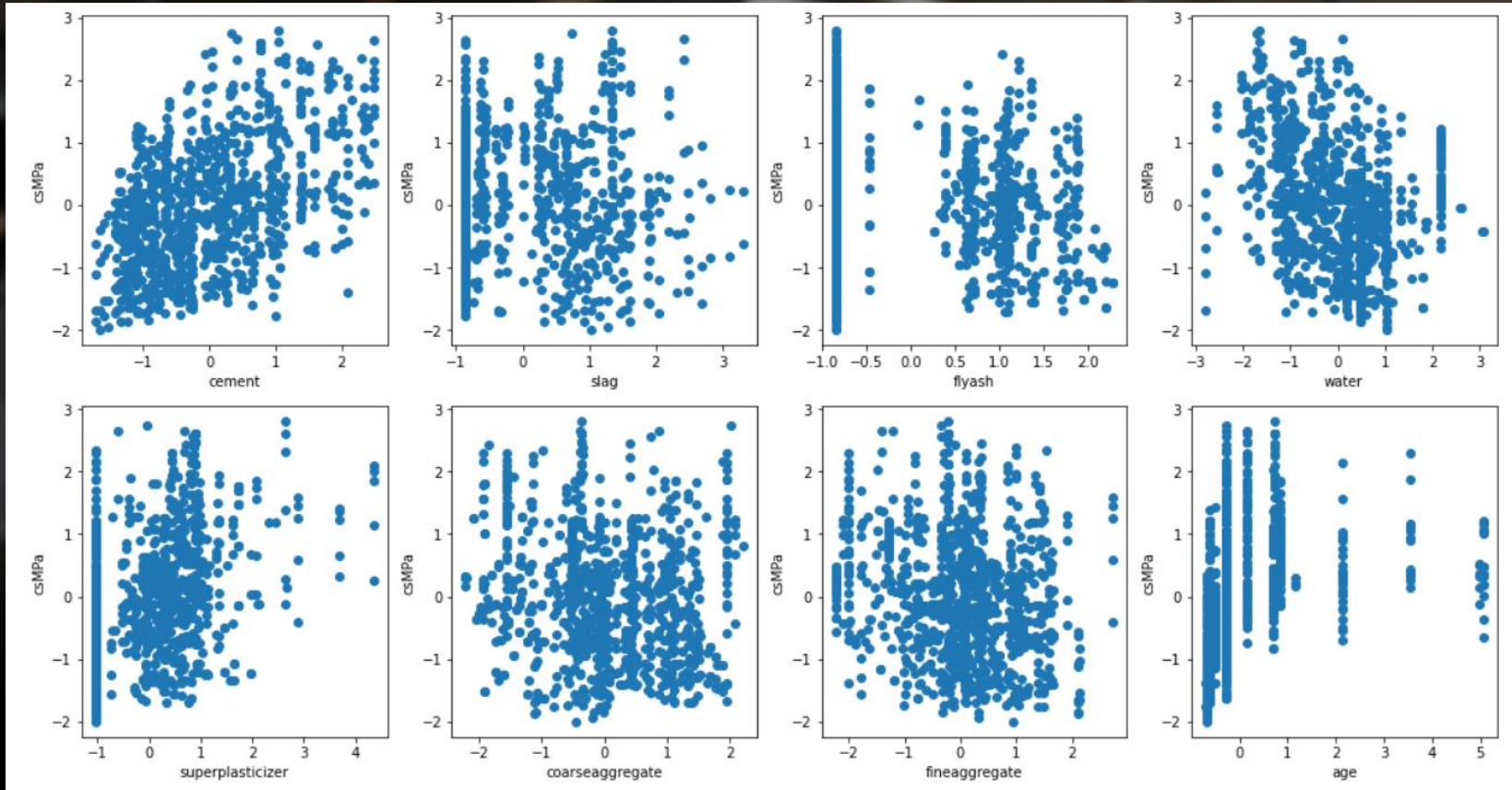
# Distribución estandarizada



# Distribución de los valores de las variables

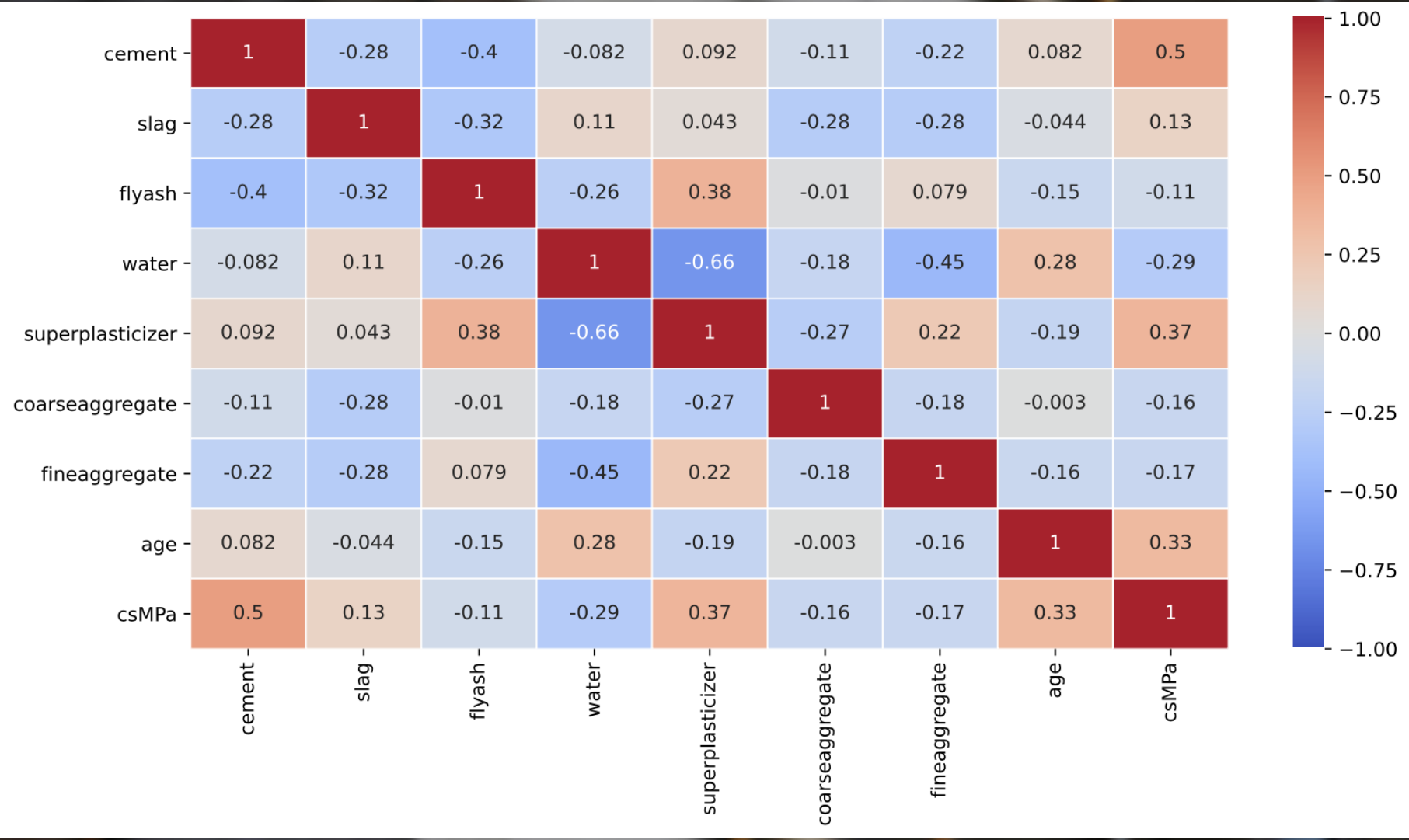



# Visualització de la variable output para cada valor de las demás de variables





# Correlación entre las variables

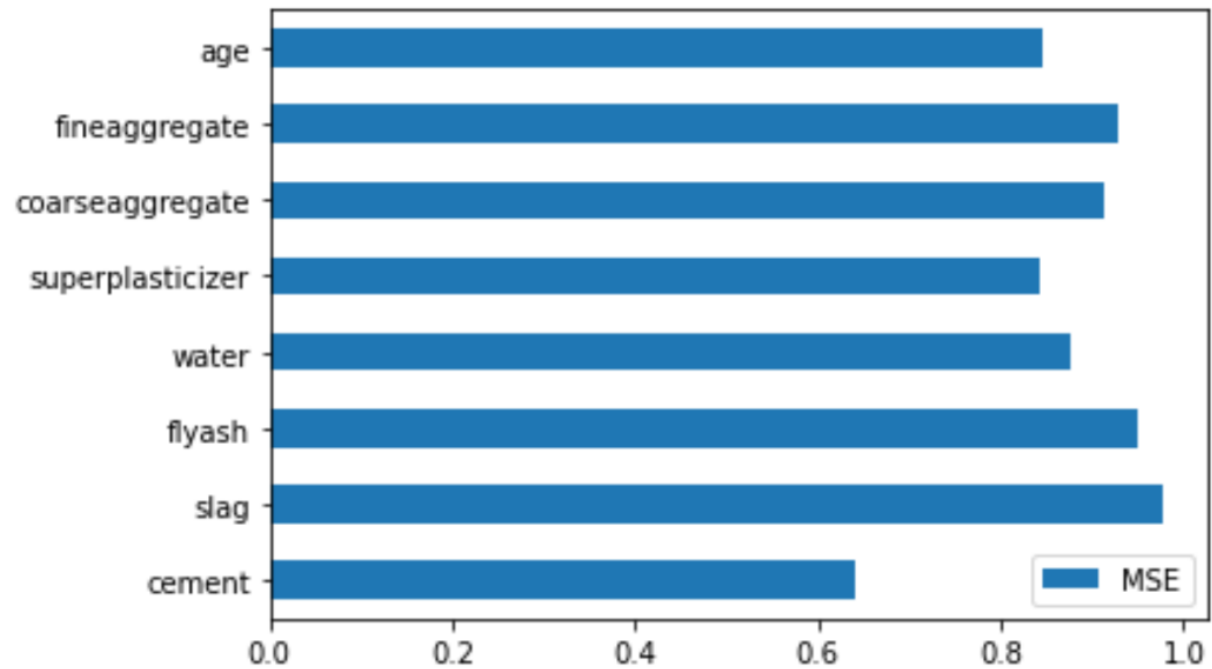


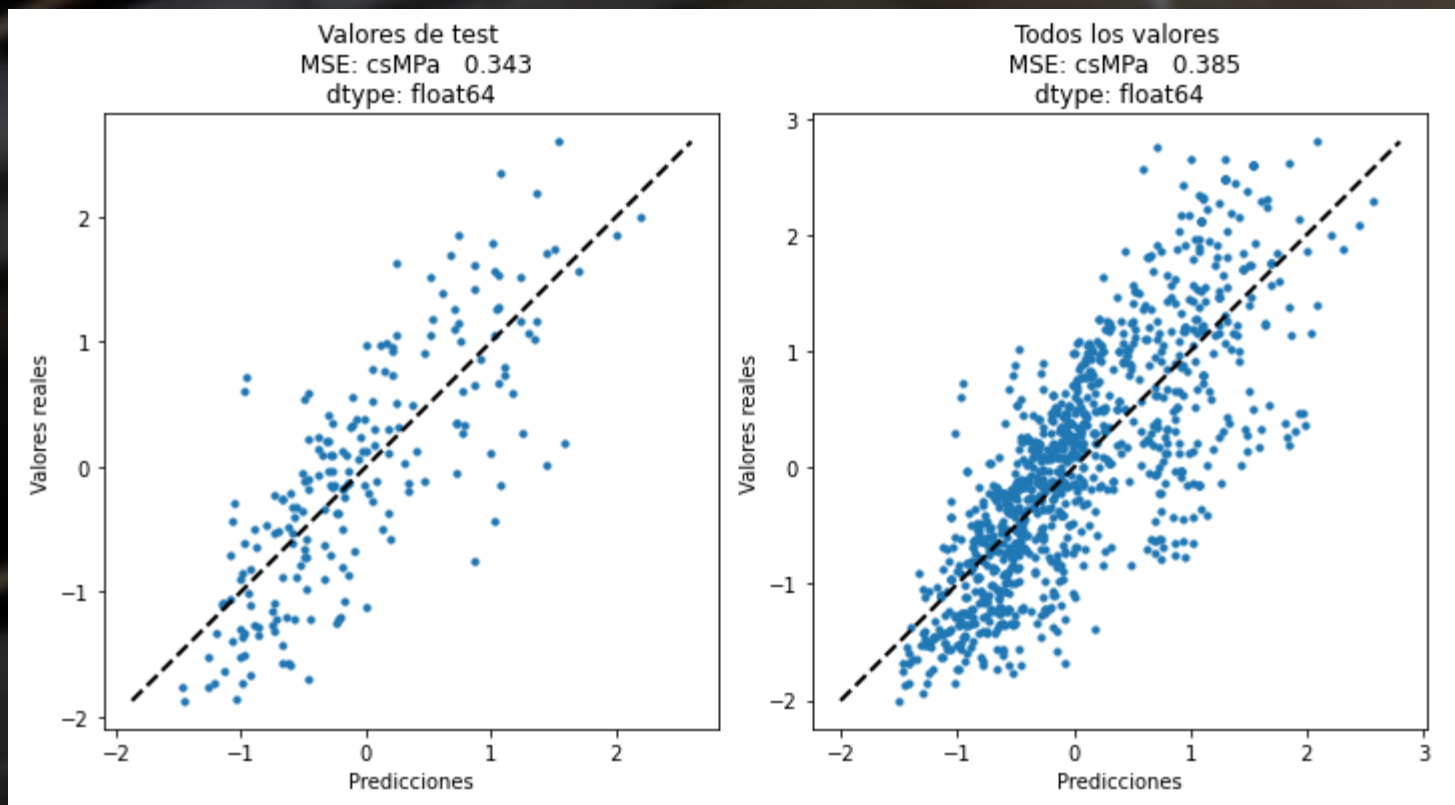


# Apartado B: Modelos de regresión lineal



# Regresión lineal con un solo atributo

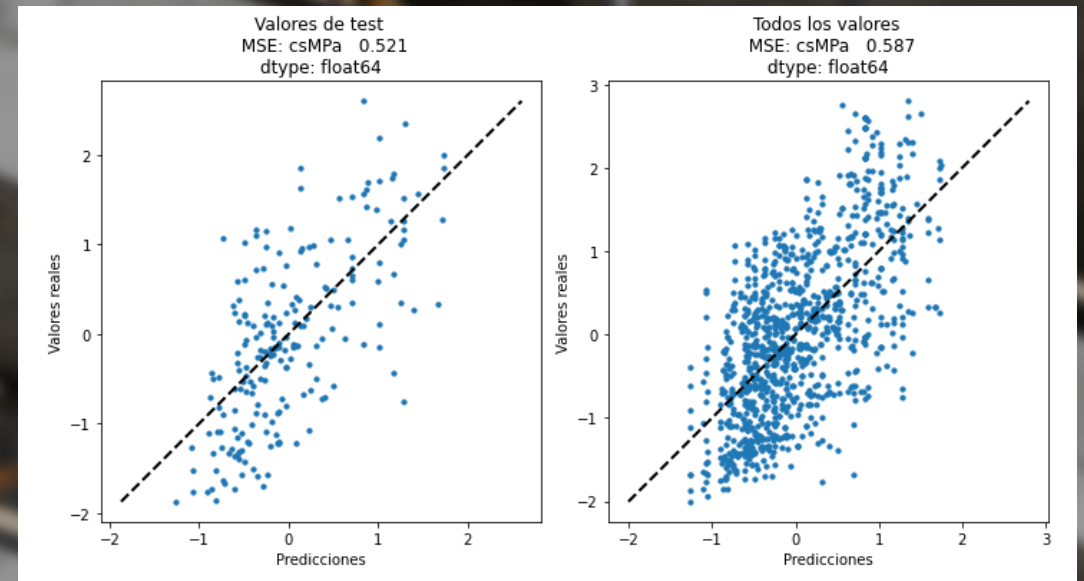
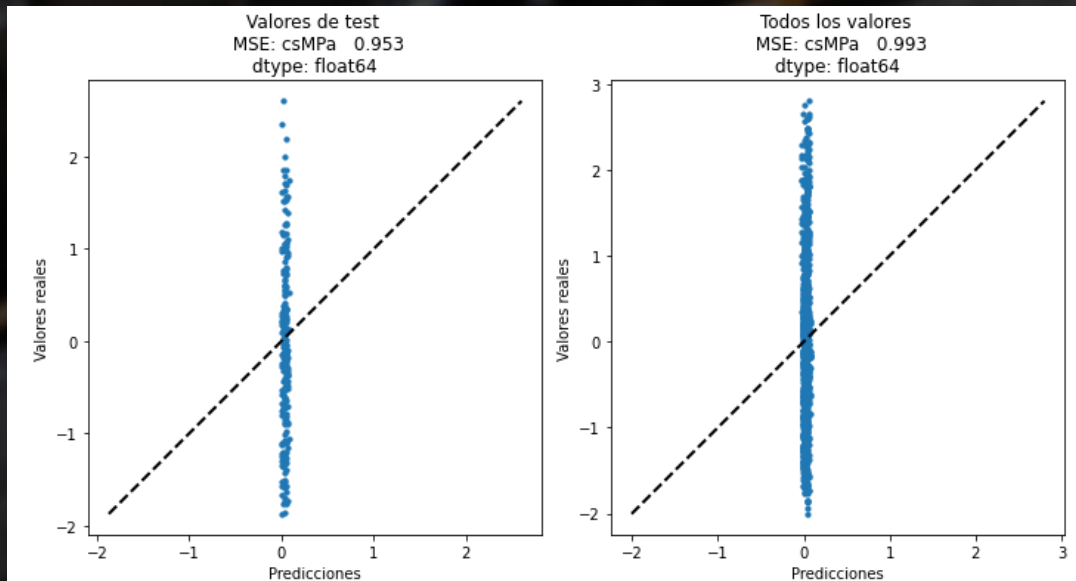




Regresión lineal  
con todos los  
atributos  
Modelo base

# Modelo con PCA

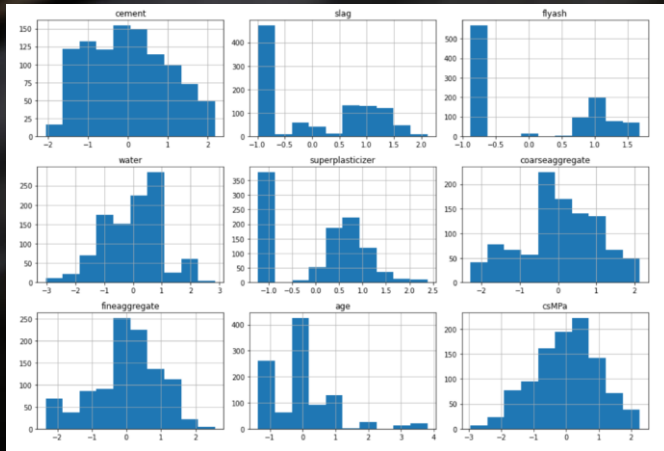
# Modelo con los atributos con distribución normal



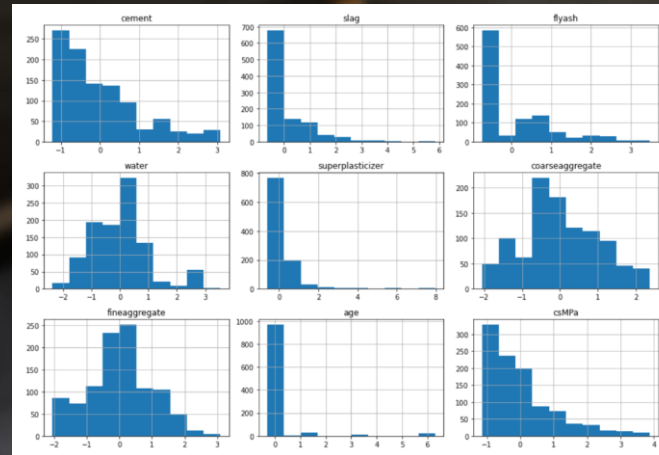


# Modelos con transformaciones de datos

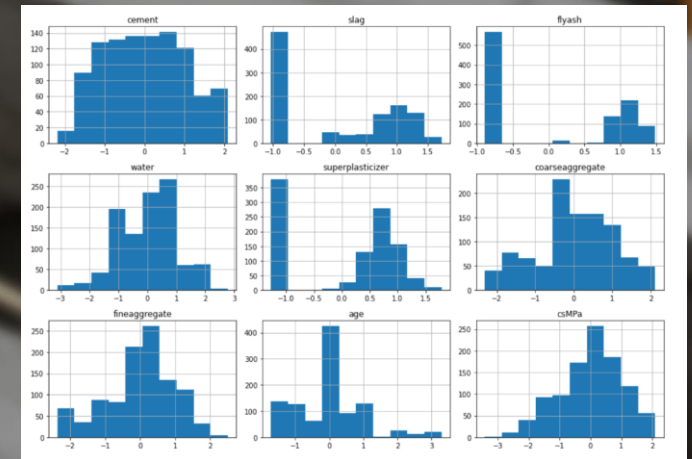
SQRT



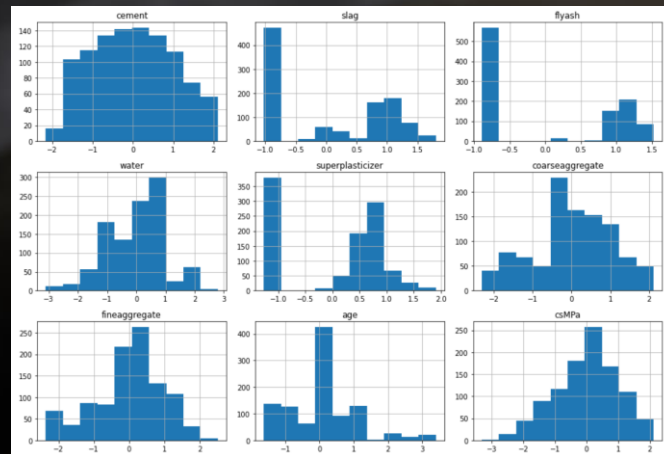
SQUARE



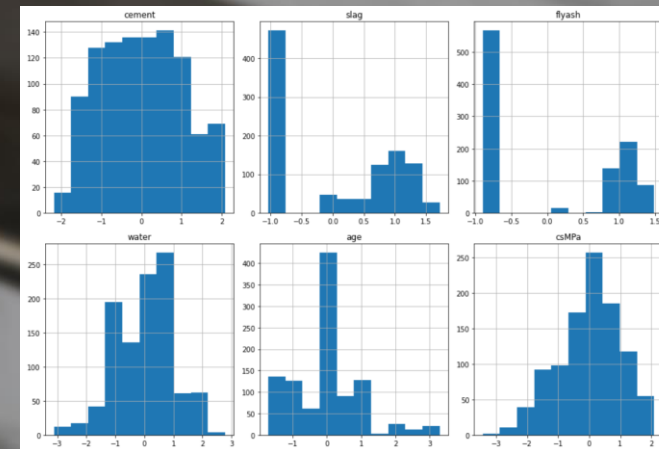
CBRT



LOG

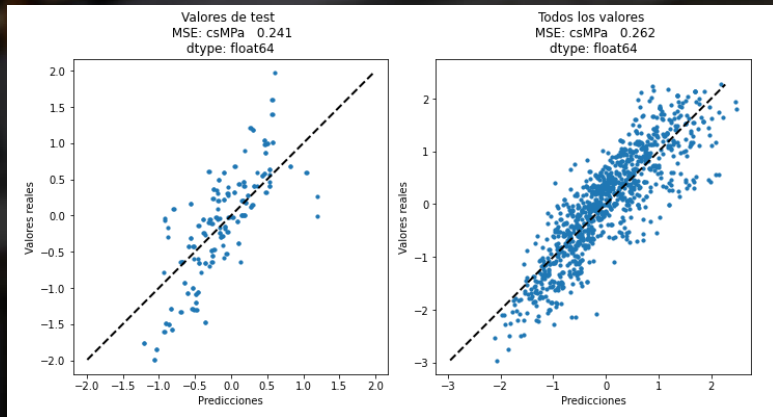


CBRT con selección

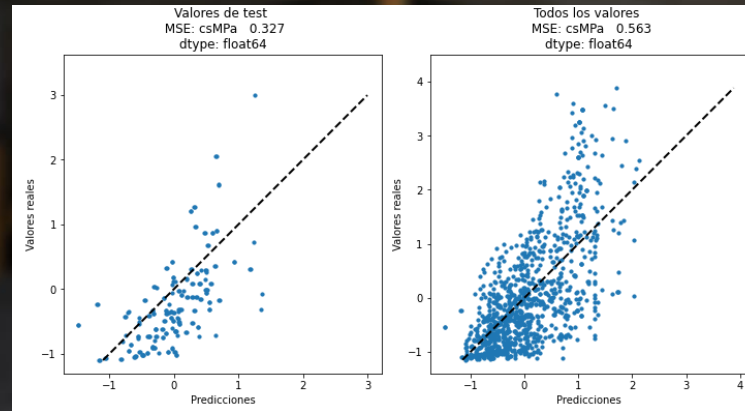


# Modelos con transformaciones de datos

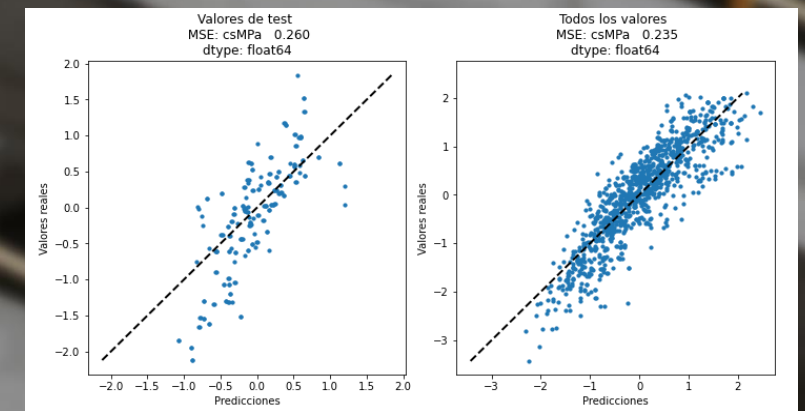
## SQRT



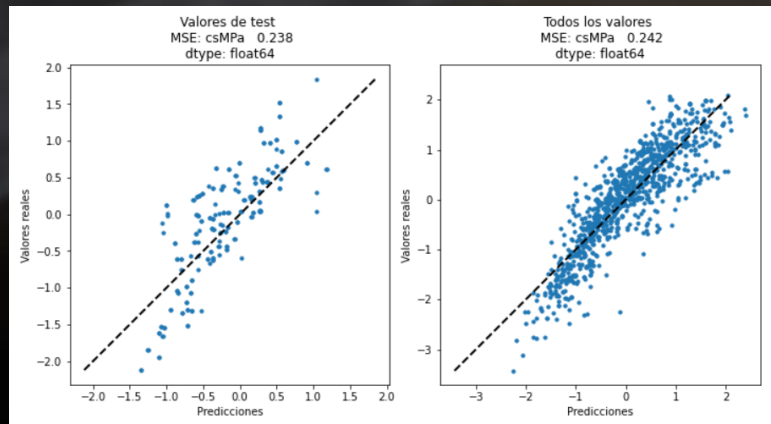
## SQUARE



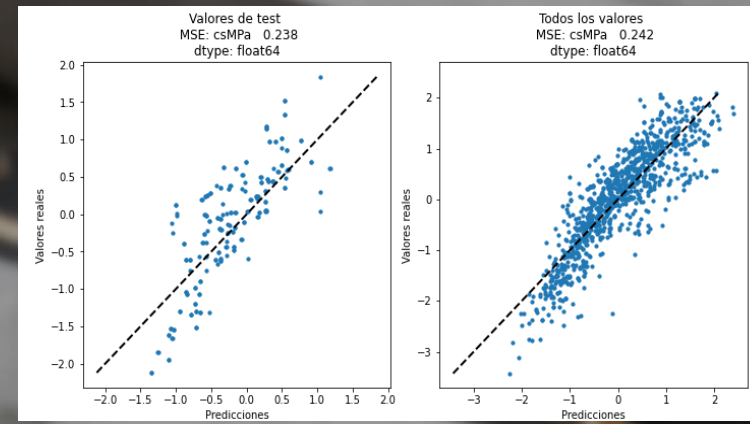
## CBRT



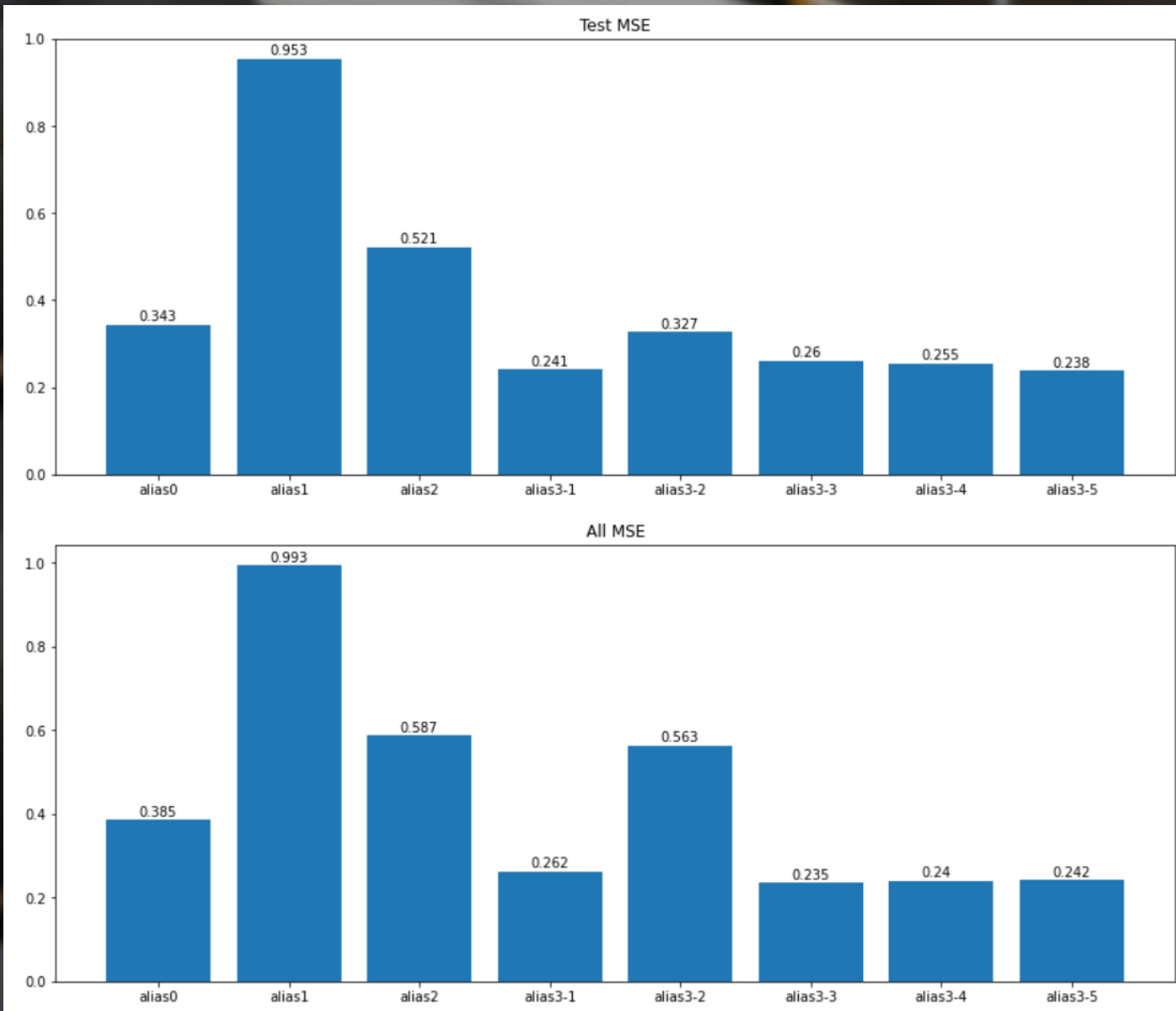
## LOG




## CBRT con selección



MSEs de  
cada modelo  
en test y ALL

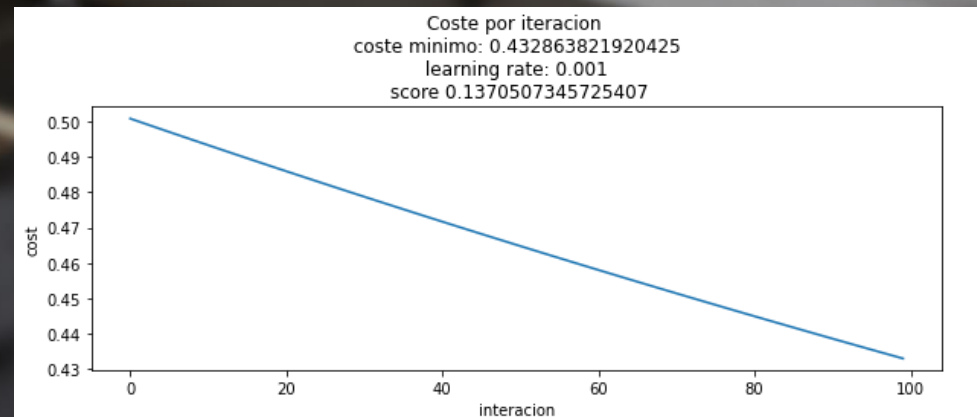
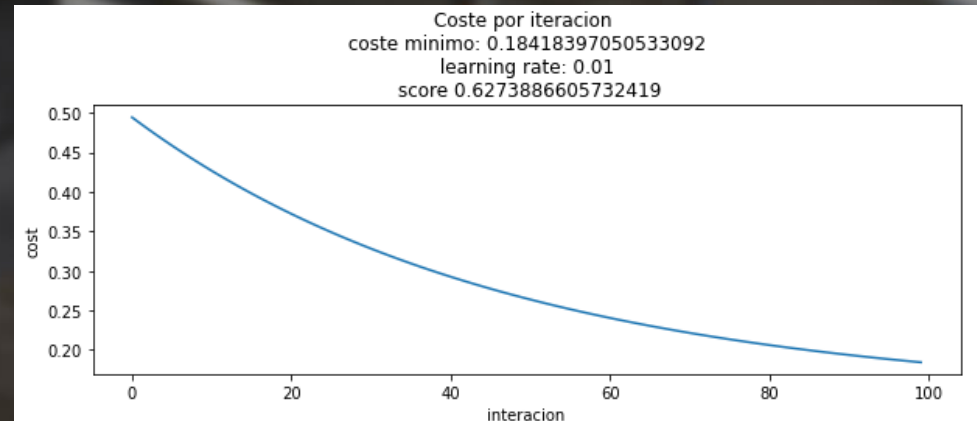
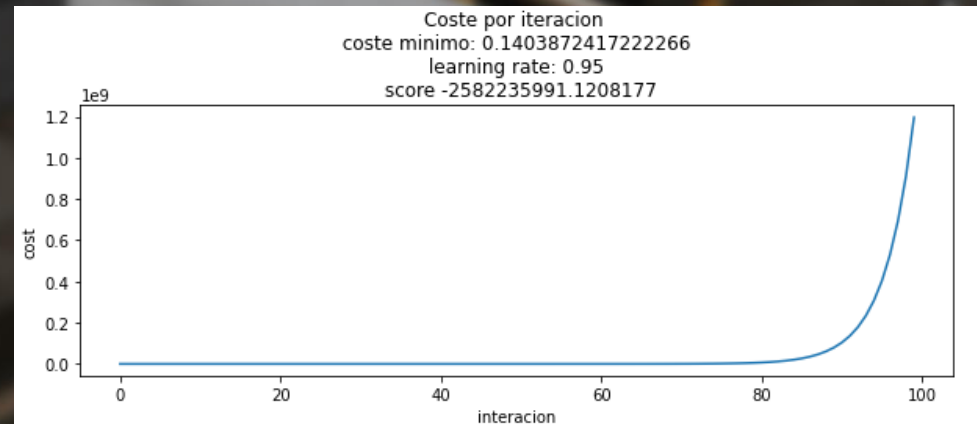


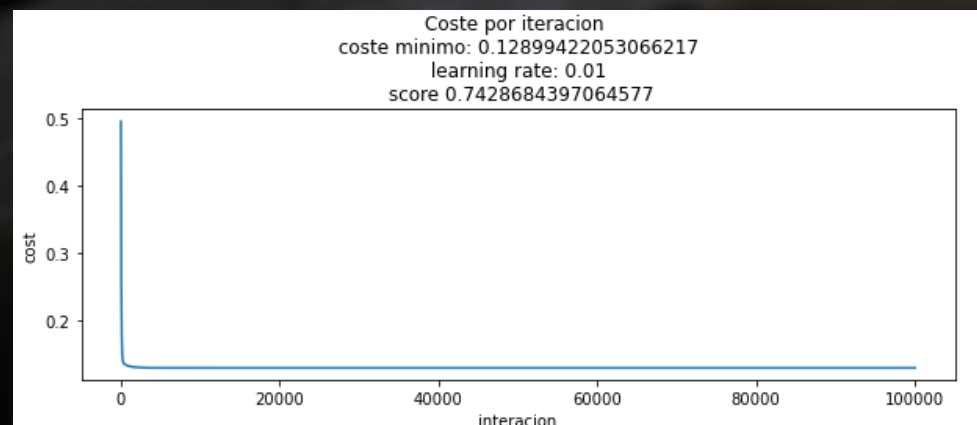
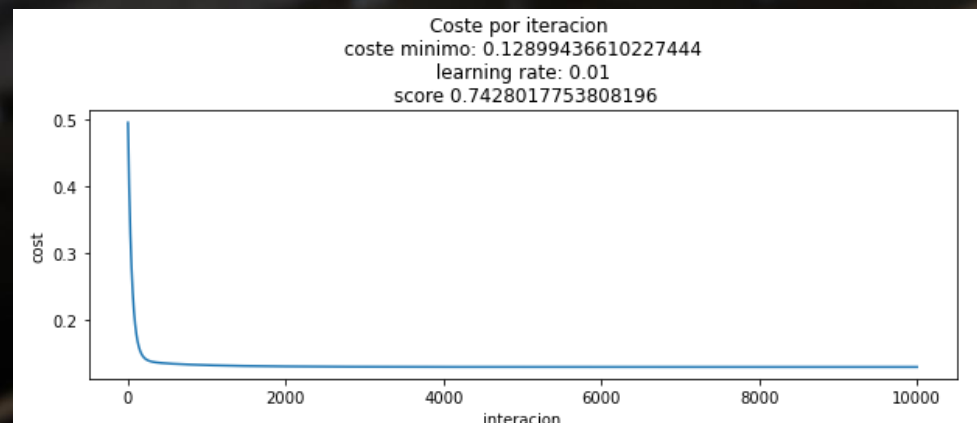
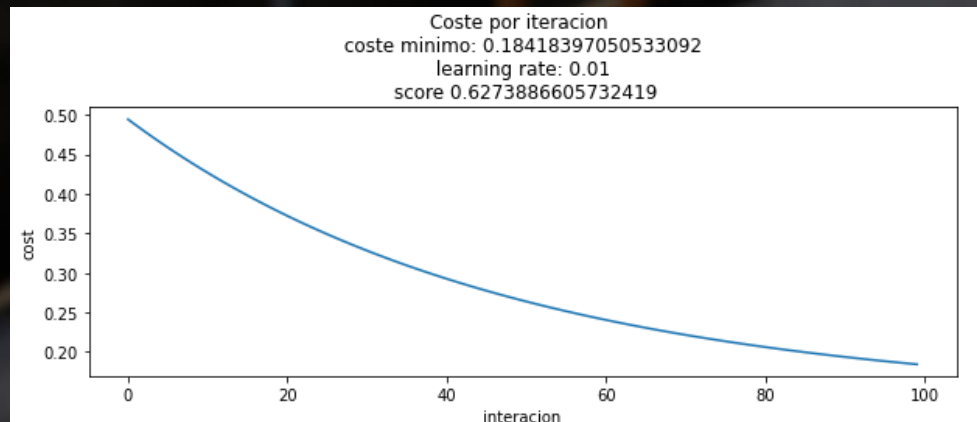




# Apartado C: Implementación del descenso del gradiente

# Descenso de gradiente Diferentes learning rates





Descenso de  
gradiente  
Diferentes  
iteraciones



# Pesos de cada variable

