

Trabajo Práctico 4:

Clustering

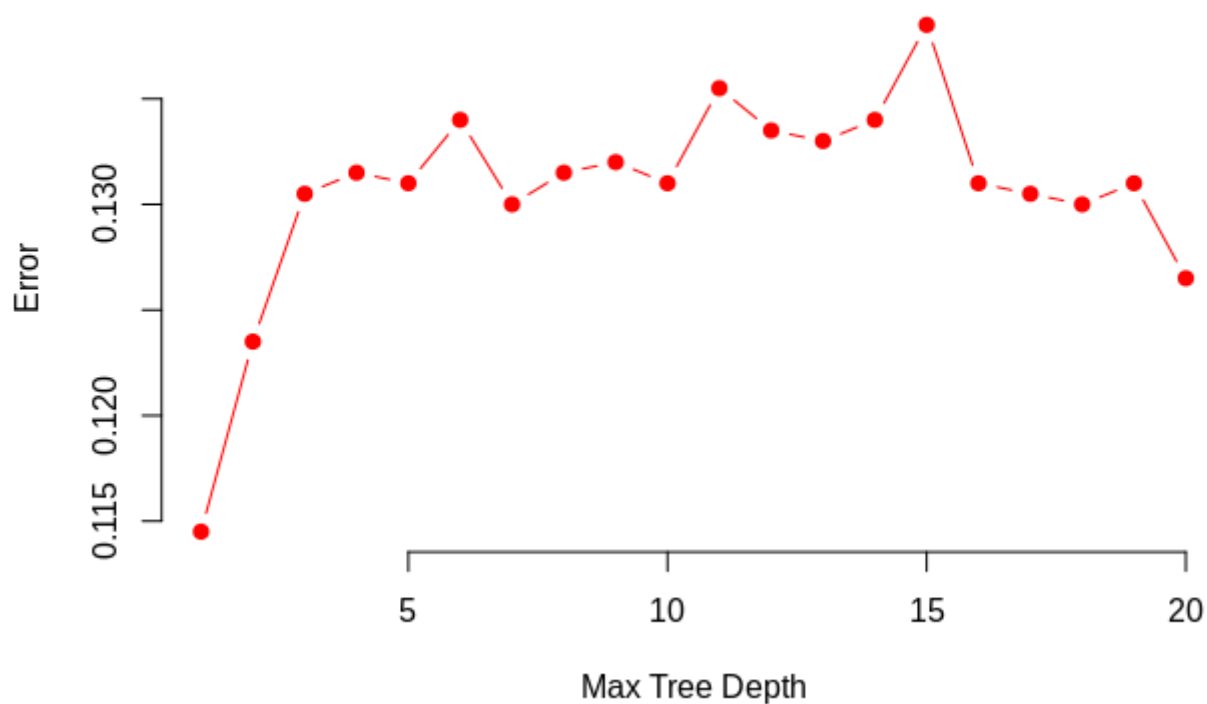
Minería de Datos



Alumno: Navall, Nicolás Uriel. N-1159/2.

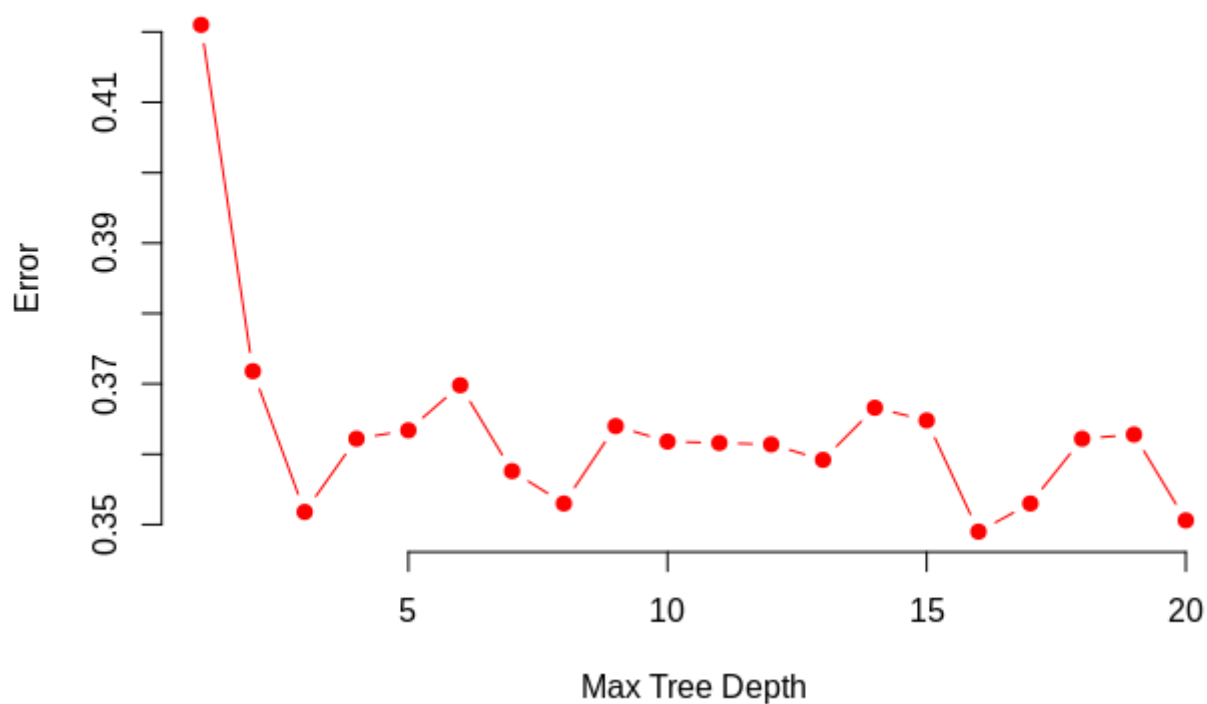
1)

Diagonal boosting classification



El error que surge al aumentar la máxima profundidad a los árboles de boosting puede atribuirse a que el dataset no es tan complejo por lo que probablemente no requiera árboles de gran profundidad. Igualmente el error en general es considerablemente bajo más allá del overfitting, lo cual puede deberse a que estamos realizando una "votación" con boosting.

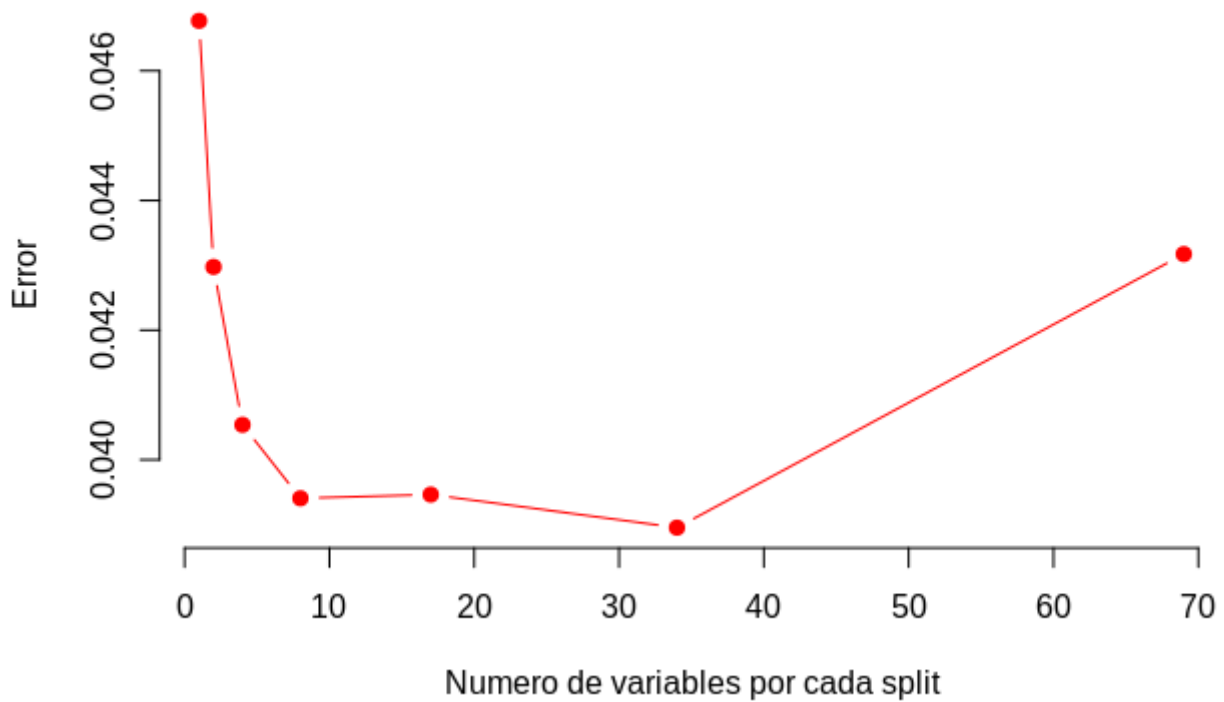
Espirales boosting classification



Como el método de árboles no funciona muy bien con el dataset de espirales andidados (más aun teniendo en cuenta que este dataset en particular tiene ruido), tiene sentido que los resultados no sean muy buenos al utilizar múltiples árboles.

2)

RRL randomForest classification



Tiene sentido que aumente el error al saltar de 35 a 69 en el valor de mtry, ya que como no hay podado ni criterio de parada esto puede interpretarse como un overfitting.

3)

svmradi	0.153
svmpoly	0.162
boosting	0.2
randomForest	0.418

Tiene sentido que randomForest nos de mas que los otros ya que se suele utilizar como cota superior de error.