**Gradient Boosting Machine**

**1º Corrida**

Se evalúa el algoritmo con combinaciones de parámetros que van desde 350 a 650 árboles, 2 y a 4 niveles de profundidad y Shrinkage con valores de 0,05 y 0,08. A cada una de las combinaciones “Hipótesis” se les llama control

Para cada combinación se calcula adicionalmente el error en test para cada uno de los umbrales de probabilidad definidos entre 0,20 y 0,42 aumentando de 0,02.

Una vez corridas todas las combinaciones de controles con los parámetros mencionados, seleccionamos mediante una función los 12 mejores resultados de error en test con su correspondiente umbral de probabilidad y le calculamos el error de Cross Validation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Control** | **Tree** | **Depth** | **Shrinkage** | **Rank. Test\_Err** | **Err\_CV** |
| 11 | 650 | 3 | 0,05 | 1 | -1,84385 |
| 8 | 550 | 3 | 0,05 | 2 | -1,834988 |
| 3 | 350 | 4 | 0,05 | 3 | -1,838405 |
| 15 | 350 | 4 | 0,08 | 4 | -1,866272 |
| 9 | 550 | 4 | 0,05 | 5 | -1,863816 |
| 22 | 650 | 2 | 0,08 | 6 | -1,831945 |
| 12 | 650 | 4 | 0,05 | 7 | -1,893658 |
| 16 | 450 | 2 | 0,08 | 8 | -1,849669 |
| 24 | 650 | 4 | 0,08 | 9 | -1,88186 |
| 5 | 450 | 3 | 0,05 | 10 | -1,848441 |
| 21 | 550 | 4 | 0,08 | 11 | -1,867927 |
| 23 | 650 | 3 | 0,08 | 12 | -1,853246 |
| 1 | 350 | 2 | 0,05 |  |  |
| 2 | 350 | 3 | 0,05 |  |  |
| 4 | 450 | 2 | 0,05 |  |  |
| 6 | 450 | 4 | 0,05 |  |  |
| 7 | 550 | 2 | 0,05 |  |  |
| 10 | 650 | 2 | 0,05 |  |  |
| 13 | 350 | 2 | 0,08 |  |  |
| 14 | 350 | 3 | 0,08 |  |  |
| 17 | 450 | 3 | 0,08 |  |  |
| 18 | 450 | 4 | 0,08 |  |  |
| 19 | 550 | 2 | 0,08 |  |  |
| 20 | 550 | 3 | 0,08 |  |  |

Graficamos los errores de test para cada Control y para cada valor de probabilidad dentro del umbral analizado



A partir de los resultados se pueden realizar las siguientes observaciones:

* Los errores tienen similar comportamiento respecto al umbral de probabilidad obteniendo mejores resultados en el rango de 0,28 y 0,32
* Para hipótesis que comprenden un gran número de árboles (550 y 650) y un shrinkage de 0,05 se observa que las que tienen mejores desempeños son las de profundidad 3 concluyendo que en 4 hay overfiting y en 2 underfiting.
* Sin embargo para igual hipótesis respecto a cantidad de árboles pero con un shrinkage más elevado (0,08) se ven malos desempeños en la profundidad 3 respecto de las otras dos, concluyendo que este es un parámetro sensible para el algoritmo.
* Para un shrinkage más pequeño (0,05) y una profundidad de 3 se observa que la mejor hipótesis es la de 650 árboles y queda ordenada en el ranking a medida que disminuye la cantidad de árboles, reflejando la propiedad de este algoritmo, que para un learning rate pequeño necesita mayor cantidad de árboles para tener mejor desempeño.