

Ejercicio de Clase.

Se entregan los datos de un experimento de optimización del índice de blancura de cascarilla de café. En este experimento se varió la concentración de NaOH, la concentración de peróxido de hidrógeno, y el tamaño del material (D_{medio}), y se midió como variable respuesta al índice de blancura (WI). Se quiere encontrar un modelo que permita predecir el valor de WI ante cambios en la concentración de NaOH, el peróxido y el D_{medio} .

1. Intente una regresión lineal para la construcción del modelo, y evalúelo con uso de la validación cruzada por k-folds, midiendo el RMSE y el R^2 . Hágalo tanto con, como sin incluir las interacciones.
2. Realice una regularización por Ridge y por Lasso. Según la regularización Lasso, ¿hay *features* que podrían eliminarse?
3. Ahora intente una regresión polinomial. Determine el orden más apropiado de la regresión usando una curva de validación
4. Ahora realice una regularización por Lasso. ¿hay *features* que podrían eliminarse?