Texto

Descripción generada automáticamente

Siguiendo los hiperparametros en default y sin realizar tratamiento de los datos se obtuvo un accuracy del 0.71875, Lo cual nos dice que nuestro modelo es aceptable pero que aún tiene aspectos por mejorar

En esta ocasión variaremos el parámetro de la distribución de los datos para train y test, ahora será de 70% para train y 30% para test. (Antes 80%-20%)



Texto

Descripción generada automáticamente

Al realizar el cambio anteriormente mencionado, el accuracy del modelo aumentó considerablemente ya que ahora estamos en 79.166, esto es algo que nos podría estar hablando de overfiting al momento de darle una mayor cantidad de datos para train, lo cual puede estar provocando que nuestro modelo memorice en lugar de que aprenda.

En entregas posteriores analizaremos más estos fenómenos dentro de este modelo.

Cabe mencionar que esta exactitud se obtiene por el método score, el cual nos dice el accuracy del modelo para predecir la variable target.