

# exp\_propio

Piñera - Saintnom - Dediego

2023-08-19

## Ejercicio 3

En el ej\_1 realizamos una transformación para modificar ciertas categorías de la columna “NObeyesdad” (Insufficient Weight, Normal Weight, Overweight Level I, Overweight Level II) por 0 indicando que el individuo no tiene obesidad y (Obesity Type I, Obesity Type II y Obesity Type III) por 1 indicando que el individuo tiene obesidad. Ahora queremos explorar cómo la optimización del punto de corte utilizado para clasificar individuos como “obesos” o “no obesos” puede influir en el rendimiento del modelo. Nuestra hipótesis es que al ajustar este corte, si la cantidad de ceros y unos llegan a un numero parecido, podríamos mejorar la capacidad del modelo para predecir mejor. Sin embargo, también consideramos que cambios en el corte podrían aumentar los errores de clasificación. Nuestro objetivo es examinar cómo estos ajustes afectan las métricas de evaluación del modelo.

Luego de realizar las transformaciones probando distintos cortes de clasificación, vemos la frecuencia de ceros y unos en cada una.

Clasificamos desde data\_1, siendo el corte mas restrictivo (solo Obesity Type III considerado como obesidad), hasta data\_5, siendo el corte mas abarcativo (considerando a Overweight Level I, Overweight Level II, Obesity Type I, Obesity Type II y Obesity Type III como obesidad).

```
## Frecuencias para data_1:
```

```
##
##      0      1
## 1787   324
```

```
##
## Frecuencias para data_2:
```

```
##
##      0      1
## 1490   621
```

```
##
## Frecuencias para data_3:
```

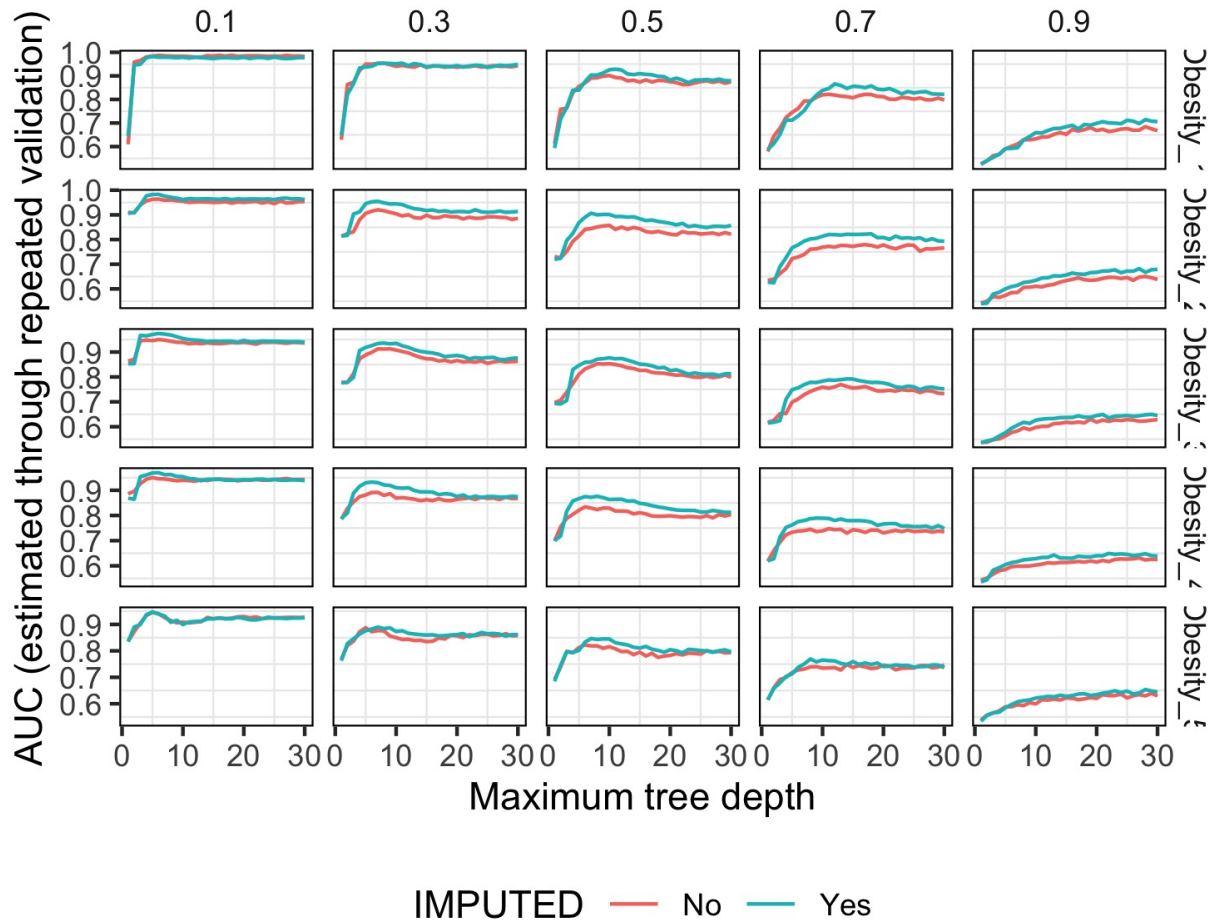
```
##
##      0      1
## 1139   972
```

```
##
## Frecuencias para data_4:
```

```
##
##      0      1
##   849 1262
```

```
##
## Frecuencias para data_5:
```

```
##
##      0      1
## 559 1552
```



Luego de realizar los experimentos, observamos un patrón interesante: el modelo obtuvo su mejor rendimiento cuando el punto de corte fue más restrictivo, clasificando a un grupo más pequeño como “obeso”. En este contexto, notamos que esta mejora podría relacionarse con la presencia de patrones extremos en los estilos de vida de los individuos clasificados como obesos. Esta agrupación de casos extremos podría ser más coherente con la forma en que el modelo está aprendiendo y generalizando, lo que a su vez mejora su capacidad predictiva para estos casos específicos.