**Preguntas de repaso 2 DATA 13**

1. Si tenemos una variable respuesta que es categórica, ¿Cuál o cuáles de las siguientes métricas de rendimiento son apropiadas para dicha variable?
2. El AUC, (el área bajo la curva ROC)
3. El coeficiente de Determinación (R2) 🡪 Numérica
4. La exactitud
5. El error cuadrático medio (RMSE) 🡪 Numérica
6. Para un algoritmo de clasificación, se aplica una validación cruzada 4-fold a la muestra de entrenamiento, obteniendo las siguientes matrices de confusión:

Texto, Carta

Descripción generada automáticamente

1. La información de las matrices confusión no es correcta, ya que las columnas no son consistentes en sus sumas, es decir, la suma de la columna 1 en la matriz 1 es diferente a la suma de la columna 1 en la matriz 2.
2. En la matriz de confusión 1 se calcula el Accuracy, en la matriz 2 se calcula la precisión, en la matriz 3 se calcula el recall y en la matriz 4 se calcula la especificidad. Luego, se promedian estos cuatro valores para tener una única métrica.
3. Fijar una métrica de la matriz de confusión, como Accuracy, calcularla para cada una de las 4 matrices y luego promediar dicha métrica.
4. La información de las matrices de confusión no es suficiente, ya que se necesitan los valores individuales para calcular el RMSE (raíz del error cuadrático medio).
5. Para un algoritmo de regresión, se aplica una validación cruzada 4-fold a la muestra de entrenamiento, obteniendo las siguientes métricas:

RMSE MAE R^2

Fold 1 2,95 1,99 0,8629

Fold 2 2,73 1,75 0,8826

Fold 3 3,19 2,25 0,8409

Fold 4 2,29 1,34 0,9174

A partir de esta información se puede concluir:

1. El fold 3 tiene el mejor rendimiento (los valores más altos en RMSE y MAE, y el más bajo en R^2) por lo que es apropiado solo usar los datos de dicha carpeta.
2. El RMSE promedio es 1.83
3. El fold 3 tiene el mejor rendimiento (los valores más bajos en RMSE y MAE, y el más alto en R^2) por lo que es apropiado solo usar los datos de dicha carpeta.
4. El coeficiente de determinación promedio es 0.8759
5. Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

   Descripción generada automáticamente¿Qué sucede en el proceso de aprendizaje automático cuando una clase es desbalanceada?