

Tarea 2

1. MSE (Mean Squared Error) - Error cuadrático medio

Fórmula:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Donde:

- y_i es el valor real.
- \hat{y}_i es el valor predicho.
- n es el número total de observaciones.

Definición:

El MSE calcula la media de los cuadrados de las diferencias entre los valores reales y los valores predichos. Esta medida penaliza fuertemente los errores grandes debido al término cuadrático. Es muy útil cuando se quiere evitar errores grandes.

Aplicación:

Se utiliza para medir la calidad de un modelo de regresión, siendo una medida estándar en la evaluación de la precisión de los modelos. El MSE es comúnmente usado en problemas de predicción donde se busca minimizar el error cuadrático.

2. MAE (Mean Absolute Error) - Error absoluto medio

Fórmula:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Donde:

- y_i es el valor real.
- \hat{y}_i es el valor predicho.
- n es el número total de observaciones.

Definición:

El MAE mide la media de las diferencias absolutas entre los valores reales y los valores predichos. A diferencia del MSE, no penaliza los errores grandes de manera tan fuerte, lo que lo hace más robusto a outliers.

Aplicación:

Se usa cuando se quiere tener una interpretación más directa del error promedio, ya que se mide en las mismas unidades que los datos. Es útil en contextos donde se requiere un error medio "sencillo", sin que los errores grandes tengan un efecto desproporcionado.

3. RMSE (Root Mean Squared Error) - Raíz cuadrada del error medio cuadrado

Fórmula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Donde:

- y_i es el valor real.
- \hat{y}_i es el valor predicho.
- n es el número total de observaciones.

Definición:

El RMSE es simplemente la raíz cuadrada del MSE, lo que lo convierte en una medida de error que tiene las mismas unidades que los datos originales. Al igual que el MSE, penaliza más fuertemente los errores grandes, pero en una escala más comparable con los datos.

Aplicación:

Es ampliamente utilizado para comparar modelos y evaluar la calidad de las predicciones en tareas de regresión. Es particularmente útil cuando los errores deben estar en la misma escala que los datos originales.

4. R-squared (R^2) - Coeficiente de determinación

Fórmula:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Donde:

- y_i es el valor real.
- \hat{y}_i es el valor predicho.
- \bar{y} es el valor medio de los valores reales.
- n es el número total de observaciones.

Definición:

El R^2 mide la proporción de la varianza de la variable dependiente que es explicada por el modelo. Un valor de R^2 cercano a 1 indica que el modelo explica bien la variabilidad de los datos, mientras que un valor cercano a 0 sugiere que el modelo no tiene capacidad predictiva.

Aplicación:

Se utiliza comúnmente para evaluar qué tan bien un modelo de regresión se ajusta a los datos. Es una medida clave en problemas de regresión lineal y se usa para entender la calidad general del modelo.