

Actividad 3 – Aprendizaje Supervisado

Brayan Steven Bonilla Castellanos
Juan Carlos Monsalve Gómez

Universidad Ibero.
Ingeniería de Software
Inteligencia artificial

Enlace repositorio GitHub

https://github.com/Brayan-Bonilla1224/IA_Unidad3_act3

Enlace video-tutorial

<https://youtu.be/fwdl91zUJCw>

Descripción del ejercicio

Escribir en Python un ejercicio que cumpla con el modelo de aprendizaje supervisado.

Ejecución del ejercicio

Importamos las librerías necesarias para la ejecución del programa

```
import matplotlib.pyplot as plt
import numpy as np
from sklearn.datasets import load_diabetes
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
```

Obtenemos un dataset de la librería sklearn, que corresponde al número de personas que padecen diabetes con relación a la edad, en donde X funciona como variable independiente (Rango de edad) y Y como variable dependiente (Número de personas con diabetes), trabajamos con un 20% de datos para pruebas y el 80% restante para entrenamiento

```
diabetes = load_diabetes()
X = diabetes.data
y = diabetes.target

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Definimos el modelo de regresión en el cual le pasamos los valores previamente establecidos como parámetros de entrada, de tal forma que nos permita obtener los valores predictivos de la información suministrada

```
model = LinearRegression()

model.fit(X_train, y_train)

y_pred = model.predict(X_test)
```

Validamos el coeficiente de determinación con el fin de verificar si el modelo se ajusta perfectamente a los datos

```
r2 = r2_score(y_test, y_pred)

print("Coeficiente de determinación (R²):", r2)
```

Definimos el tipo de grafo que vamos a utilizar, con el fin de que nos permita realizar un análisis de las predicciones obtenidas en el ejercicio

```
x_range = np.arange(len(y_test))

plt.plot(x_range, y_test, label='Valores reales')
plt.plot(x_range, y_pred, label='Predicciones')
plt.xlabel('Índice de muestra')
plt.ylabel('Progresión de la enfermedad')
plt.title('Comparación de valores reales y predicciones')
plt.legend()
plt.show()
```

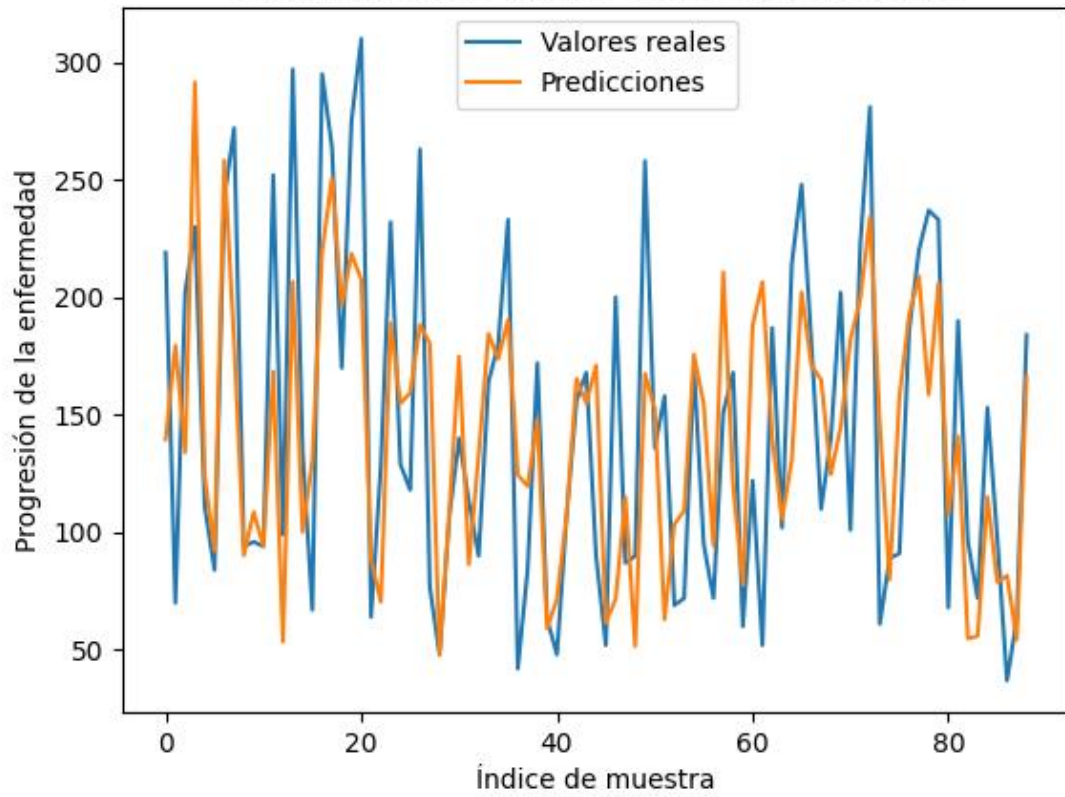
Ejecución del ejercicio

Obtenemos una relación óptima entre el modelo y los datos

```
Coeficiente de determinación (R²): 0.4526027629719197
```

Por último, obtenemos el grafo plasmado con los valores reales y las predicciones correspondiente al modelo de aprendizaje supervisado basado en la información suministrada

Comparación de valores reales y predicciones



Conclusiones

- Todo análisis de datos parte de unos objetivos, siendo entonces lo más importante el diseño de la investigación o la estrategia de análisis, no la herramienta, es fundamental tener en cuenta que la herramienta es solo un medio.
- Las herramientas de inteligencia artificial tienen un gran potencial de uso en el análisis de datos, mejorando de esta manera significativamente las posibilidades de análisis más rápidos, más potentes, más precisos, con otras alternativas que redundarán en mejoras en diferentes ámbitos.
- Los métodos de aprendizaje supervisado, de manera especial, permiten a los sistemas desarrollados un aprendizaje a partir de los datos suministrados, mejorando continuamente a través de los datos que se propician a este tipo de sistemas, teniendo resultados cada vez más precisos.

Bibliografía

¿Qué es el aprendizaje supervisado? | IBM. (s. f.). <https://www.ibm.com/mx-es/topics/supervised-learning>