

*Universidad de Oriente.*

*Facultad de Ingeniería en Telecomunicaciones, Informática y Biomédica.*

*Departamento de Informática.*

# ***TAREA INDEPENDIENTE DE INTELIGENCIA ARTIFICIAL II***

**Autora:** Nineth Maza Arroyo

**Año:** 2do      **Semestre:** II

**grupo:** INF-21

*CARRERA INGENIERÍA INFORMÁTICA*

*2023*

## ORIENTACIONES:

1-Principales características de las estrategias de regresión lineal y los árboles de decisión.

La estrategia de regresión lineal es un método popular para modelar la relación entre una variable dependiente y una o más variables independientes. Las principales características son:

- a. Linealidad: La regresión lineal asume una relación lineal entre las variables independientes (predictoras) y la variable dependiente (objetivo). Esto significa que los cambios en las variables independientes están linealmente relacionados con los cambios en la variable dependiente.
- b. Modelo paramétrico: La regresión lineal es un modelo paramétrico, lo que significa que se asume que la relación entre las variables puede describirse en términos de un conjunto fijo de parámetros. Estos parámetros se estiman a partir de los datos de entrenamiento.
- c. Coeficientes de regresión: El modelo de regresión lineal estima los coeficientes (pendientes) que cuantifican la magnitud y dirección de la relación entre las variables independientes y la variable dependiente. Estos coeficientes representan el cambio esperado en la variable dependiente por unidad de cambio en la variable independiente, manteniendo constante el resto de las variables independientes.

d. Suposiciones: La regresión lineal se basa en varias suposiciones, incluyendo la linealidad, la independencia de los errores, la homocedasticidad (varianza constante de errores), la normalidad de los errores, y la ausencia de multicolinealidad entre las variables independientes.

e. Método de mínimos cuadrados: La regresión lineal busca minimizar la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos por el modelo. Este enfoque se conoce como el método de mínimos cuadrados.

f. Interpretación de coeficientes: Los coeficientes estimados en la regresión lineal pueden interpretarse para entender la relación cuantitativa entre las variables independientes y la variable dependiente. Esto permite evaluar el impacto de cada variable independiente sobre la variable dependiente.

g. Flexibilidad: Aunque se basa en una relación lineal, la regresión lineal puede adaptarse para incluir interacciones entre variables, transformaciones no lineales de las variables, y manejo de multicolinealidad para modelar relaciones más complejas.

Los árboles de decisión son una herramienta que nos ayuda a visualizar los posibles resultados de decisiones relacionadas entre sí. Comúnmente comienzan en un solo nodo y se van partiendo conforme las decisiones derivadas van sucediendo. Finalmente nos queda la forma de un árbol, de ahí es que proviene el nombre. Son usados por individuos o compañías para determinar cuál es la mejor ruta o camino a tomar para obtener el máximo beneficio. También podemos usarlos para clasificar cosas, por ejemplo: Determinar si un animal es un pez.

Son contruidos por analistas, dentro de una empresa y hasta por nosotros, pero si esta tarea la realiza una computadora estamos hablando de Machine Learning. La forma más común en la que una computadora puede generar un árbol de decisiones a partir de un dataset es comenzando desde arriba: comienza en un nodo raíz con todo el dataset, después el algoritmo selecciona una variable y luego un valor específico de esta variable que nos da la mejor división de los datos. Por cada división el algoritmo genera 2 nuevos nodos que contienen una versión reducida de nuestro dataset. Finalmente se aplica el mismo procedimiento que al nodo raíz, es un proceso recursivo que se repite hasta llegar a un nodo para el cual ya no sea posible realizar una división de nuestras observaciones. Cuando queremos hacer una predicción tomamos la nueva instancia y seguimos las decisiones que el árbol creó para llegar al resultado.

2-Realizar un programa en Python donde se realice un experimento con una de estas estrategias.

Este código crea datos de ejemplo, ajusta un modelo de regresión lineal a los datos utilizando la biblioteca scikit-learn para construirlo y visualiza tanto los datos como la línea de regresión resultante con matplotlib.

```
# Importar las bibliotecas necesarias
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# Crear datos de ejemplo
```

```
np.random.seed(0)
x = 2 * np.random.rand(100, 1)
y = 3 + 4 * x + np.random.randn(100, 1)
```

```
# Visualizar los datos de ejemplo
plt.scatter(x, y, color='b', label='Datos de ejemplo')
plt.xlabel('X')
plt.ylabel('Y')
plt.title('Datos de ejemplo para regresión lineal')
plt.show()
```

```
# Crear y ajustar el modelo de regresión lineal
model = LinearRegression()
model.fit(x, y)
```

```
# Visualizar la línea de regresión
plt.scatter(x, y, color='b', label='Datos de ejemplo')
plt.plot(x, model.predict(x), color='r', label='Regresión lineal')
plt.xlabel('X')
plt.ylabel('Y')
plt.title('Regresión lineal')
plt.legend()
plt.show()
```