ACTIVIDAD ALGORITMOS DE CLASIFICACIÓN

La empresa DataVision te ha contratado como analista de datos para desarrollar un modelo de clasificación de clientes. Se te ha proporcionado un conjunto de datos con información sobre clientes, incluyendo variables como edad, ingresos, historial de compras y suscripciones.

Tu tarea es aplicar distintos algoritmos de clasificación para predecir si un cliente potencial contratará un nuevo servicio.

INSTRUCCIONES:

1. Carga de datos (1 punto)

- Descarga el archivo clientes.csv.
- Carga el conjunto de datos utilizando Python y realiza una exploración inicial.

2. Aplicación de modelos de clasificación (6 puntos)

Implementa los siguientes modelos de clasificación y evalúa su desempeño:

- Regresión logística: Implementa el modelo y analiza cómo la función sigmoidea afecta la clasificación.
- K-Nearest Neighbors (K-NN): Prueba distintos valores de k y analiza su impacto en la exactitud.
- Árbol de decisión: Ajusta los hiperparámetros del modelo y analiza las medidas de impureza de los nodos.
- **Bosques aleatorios:** Aplica bagging para mejorar la predicción y analiza la importancia de las variables.
- Support Vector Machine (SVM): Experimenta con distintos tipos de kernel y compara los resultados.

Para cada modelo:

- Divide los datos en entrenamiento y prueba.
- Normaliza los datos si es necesario.

• Evalúa la precisión con métricas como accuracy.

3. Análisis de resultados (3 puntos)

- Comparación de modelos: Explica cuál modelo tuvo mejor desempeño y por qué.
- Impacto de los hiperparámetros: Describe cómo afectaron los hiperparámetros en cada modelo.
- Aplicabilidad: Indica en qué casos sería recomendable utilizar cada algoritmo.

INSTRUCCIONES ADICIONALES:

- Puntos totales = 10 puntos.
- Descarga el material complementario que se encuentra disponible en la plataforma.
- Comprime el archivo en formato .zip o .rar.
- Incluye un documento de texto con las reflexiones analíticas.
- Sube el archivo a la plataforma.