

Prueba Técnica: Creación de un Modelo de Segmentación de Clientes

Objetivo:

El objetivo de esta prueba técnica es evaluar tus habilidades en el diseño, desarrollo y evaluación de un modelo de segmentación de clientes utilizando técnicas de aprendizaje automático. Deberás crear un modelo que sea capaz de segmentar a los clientes en diferentes grupos según sus características y comportamientos. El desarrollo de esta prueba puede ser en Python o R.

Datos:

Utilizaremos un conjunto de datos ficticio para esta prueba. El conjunto de datos contiene información sobre varios clientes y sus características.

El conjunto de datos incluye las siguientes columnas:

1. `ID`: Identificador único del cliente.
2. `Edad`: Edad del cliente.
3. `Género`: Género del cliente (Masculino/Femenino).
4. `Ingresos Anuales`: Ingresos anuales del cliente.
5. `Gasto Mensual`: Cantidad promedio gastada por el cliente cada mes.
6. `Puntuación de Crédito`: Puntuación de crédito del cliente basada en su historial crediticio.
7. `Región`: Región geográfica del cliente.
8. `Comportamiento en Línea`: Puntuación que indica la actividad del cliente en línea (0-10).
9. `Compras en el Último Año`: Número de compras realizadas por el cliente en el último año.

Tareas:

1. Análisis Exploratorio de Datos: Realiza un análisis exploratorio de los datos para entender la distribución de las características, identificar valores faltantes y posibles relaciones entre variables.
2. Preprocesamiento de Datos: Limpia los datos, trata los valores faltantes y codifica las variables categóricas si es necesario.
3. Selección de Características: Decide qué características utilizarás para el modelo

de segmentación y justifica tu elección.

4. Modelo de Segmentación: Implementa un algoritmo de clustering para segmentar a los clientes en diferentes grupos.

5. Evaluación del Modelo: Evalúa la calidad de los clústeres utilizando métricas relevantes como la inercia, el coeficiente de silueta u otras métricas que consideres adecuadas.

6. Visualización de Resultados: Visualiza los clústeres en un gráfico de dispersión (scatter plot) utilizando una técnica de reducción de dimensionalidad como PCA (Análisis de Componentes Principales).

7. Informe y Presentación: Realiza una presentación concisa (diapositivas) en la que puedas comunicar tus hallazgos y resultados de manera efectiva. Además anexa la URL del repositorio que contiene el código fuente.