Contenido

[Conceptos básicos 3](#_Toc192684048)

[Técnicas de modelización 3](#_Toc192684049)

[Preprocesamiento de datos 4](#_Toc192684050)

[Modelización y aprendizaje 4](#_Toc192684051)

[Inferencia y aplicación de modelos 4](#_Toc192684052)

[Casos de aplicación 4](#_Toc192684053)

**Ejercicio sobre Minería de Datos y Análisis Predictivo**

**Instrucciones:**

Lee atentamente cada pregunta y responde con base en los conceptos del Capítulo 7 "Análisis Predictivo: Minería de Datos".

**Preguntas:**

# Conceptos básicos

a) ¿Qué es la minería de datos y cuál es su principal objetivo en el análisis de Big Data?

b) Explica la diferencia entre análisis descriptivo y análisis predictivo dentro del contexto de minería de datos.

# Técnicas de modelización

a) ¿Cuál es la diferencia entre un modelo supervisado y un modelo no supervisado?

b) Clasifica los siguientes ejemplos dentro de modelos supervisados o no supervisados:

* + Segmentación de clientes según patrones de compra.
  + Predicción de la probabilidad de que un cliente cancele su suscripción.
  + Detección de fraude en transacciones bancarias.
  + Identificación de grupos de productos que suelen comprarse juntos.

# Preprocesamiento de datos

a) ¿Por qué es importante la etapa de preprocesamiento en un proyecto de minería de datos?

b) Menciona al menos tres técnicas utilizadas en el preprocesamiento de datos y su utilidad.

# Modelización y aprendizaje

a) ¿Qué significa el término "sobreajuste" (overfitting) en un modelo de aprendizaje supervisado?

b) ¿Qué estrategias pueden utilizarse para evitar el sobreajuste?

# Inferencia y aplicación de modelos

a) ¿Qué diferencia hay entre la inferencia por lotes y la inferencia en tiempo real en la puesta en producción de un modelo?

b) ¿Por qué es importante monitorear un modelo una vez que ha sido desplegado en producción?

# Casos de aplicación

a) Explica cómo la minería de datos podría aplicarse en un hospital para mejorar la atención a los pacientes.

b) Imagina que una empresa de e-commerce quiere mejorar sus recomendaciones de productos. ¿Qué tipo de modelo de minería de datos recomendarías y por qué?