

“PROGRAMA ESPECIALIZADO CIENCIA DE DATOS CON PYTHON”

PROYECTO INDIVIDUAL O GRUPAL

RETO: CIENCIA DE DATOS - Modelo de Siniestro



RETO:

Se tiene un conjunto de datos simulado basado en seis meses de la información recogida por una aseguradora mundial llamada “SALUD Y SEGUROS”, cuyo objetivo fue desarrollar un modelo para determinar la probabilidad de siniestro en los próximos 12 meses.

La variable de interés es la siniestralidad , que representa el estado binario de No Siniestro 0 y Siniestro 1.

DESCRIPCIÓN DEL PROBLEMA

Los modelos analíticos para el manejo de los seguros de accidentes se están usando por muchas instituciones y están dando resultados exitosos en todo el mundo. Los modelos analíticos se pueden definir como un conjunto de métodos y técnicas cuantitativas usados para predecir la probabilidad de que un cliente falle (Sea siniestroso) y en consecuencia no se recupere el rédito otorgado por la empresa.

DESCRIPCIÓN DE LAS VARIABLES



Cliente_ID. : Identificador de cliente.

Antigüedad_Maxima: Antigüedad del cliente en meses.

Nivel_Ingresos: Nivel de ingresos de 0 (bajo) a 5 (alto).

Saldo_Pendiente: Saldo pendiente del cliente.

Puntaje_Morosidad : Tasa de morosidad en el seguro. 0: No Dq., 1:1-29... 6:150-179.

Saldo_Pendiente_Seg. : Saldo pendiente de la póliza del seguro

Siniestros: Número de siniestros en la historia del cliente.

Estado_Siniestro. 0: No Siniestro; 1: Siniestro.

ENTREGABLE:

- Desarrollar el modelo basado en la metodología CRISP-DM.
 - a. Realizar el análisis exploratorio de datos (EDA) con la variable dependiente (Target) e interpretar los hallazgos.
 - b. Identificar las variables más importantes
 - c. 1. Realizar análisis exploratorio de datos
 - d. 2. Realizar la segmentación de clientes (creando la variable cualitativa de segmentación).
 - e. 3. Desarrollar 1 modelo incluyendo la variable de segmentación como variable cualitativa o desarrollar 1 modelo dentro de cada segmento.
 - f. Interpretación, conclusiones y recomendaciones
- Subir a Google Classroom la solución en Python Notebook (documentado). ●
- Realizar una presentación en PPT y enviar el link en youtube de la sustentación de la solución. El equipo debe de sustentar paso a paso los resultados.
- Grupo: máximo 02 integrantes.

Data: [Siniestros.csv](#)

[Subir al Proyecto - Google Classroom.](#)

Fecha máxima de entrega: 26 de diciembre de 2024

Instructora: Msc (e) Ing. Est. E Inf. Yaneth Ccahuantico Mendoza