

Reto | Mercadotecnia telefónica con aprendizaje supervisado

Entregable final del reto

Documento editable por el participante

Nombre del participante:

Indicaciones:

A continuación, se describen los pasos a realizar en la presente actividad, en lo que llamaremos el Plan del Reto:

1. Deberás utilizar el archivo llamado **bank_marketing.csv** con los datos de problema.
2. Utilizar el archivo **bank-names.txt** para obtener información de cada una de las variables.
3. Crear un proyecto tipo Jupyter Notebook en Google-Colab llamado **Solucion_Reto_SC_63_<nombre_y_apellido_del_estudiante>.ipynb**.
4. Incluye las librerías que consideres adecuadas y carga los datos del archivo en una variable llamada "data".
5. Obtener la información de dicha base de datos que incluya el número de registros, el total de variables, el tipo de cada variable, la cantidad de datos perdidos de cada variable en caso de que existan.
6. Transforma las variables categóricas de manera que puedan ser tratadas numéricamente. Justifica si utilizas **LabelEncoder** o **OneHotEncoder**.
7. Transforma las variables numéricas en los casos que se tenga algún tipo de sesgo.
8. Considera la variable "y" como la variable de salida y el resto de las variables como las variables de entrada.
9. Particiona los datos en los conjuntos de entrenamiento, validación y prueba en **60%, 20% y 20%**, respectivamente.
10. Aplica el modelo **Regresión Logística** en el conjunto de entrenamiento. Valida el modelo con las predicciones del conjunto de validación y su matriz de confusión. Ajusta los parámetros del modelo hasta obtener tu mejor resultado.
11. Aplica el modelo **Red Neuronal** en el conjunto de entrenamiento. Valida el modelo con las predicciones del conjunto de validación y su matriz de confusión. Ajusta los parámetros del modelo hasta obtener tu mejor modelo, entre ellos el número de neuronas y capas ocultas.
12. Selecciona **el mejor modelo encontrado** en los incisos anteriores y utiliza el conjunto de prueba para obtener el desempeño final del modelo y su matriz de confusión.

13. Incluye tus conclusiones del problema, en particular, ¿qué puedes decir acerca del uso de técnicas de inteligencia artificial en problemas de mercadotecnia?
14. Descarga tu script (archivo con extensión .ipynb) y guárdalo siguiendo la nomenclatura que se te indica en Formato de entrega de actividad.
15. Sube el archivo a GitHub.
16. Copia y pega en un archivo de edición de texto la liga de tu archivo.

6.a Justifica el uso de **LabelEncoder** o **OneHotEncoder**.

Se utilizó un **OrdinalEncoder** para convertir variables categóricas a numéricas para que los algoritmos de ML pudieran establecer relaciones entre los valores numéricos que representan clases dentro de las variables categóricas y la clase de interés que representa el éxito o fracaso de los programas de telemarketing.

13.a Incluye tus conclusiones de este problema en particular, ¿qué puedes decir acerca del uso de técnicas de inteligencia artificial en problemas de mercadotecnia?

Para el caso específico de la clasificación de casos donde se tiene una variable que puede ser utilizada como referencia para establecer relaciones entre la misma y variables independientes, es posible establecer la relación mencionada. Sin embargo, no es suficiente el tener las etiquetas para que los algoritmos de ML puedan encontrar patrones subyacentes dentro de los datos que permitan establecer las relaciones mencionadas, es indispensable que los datos permitan establecerlas. Por más que tengamos las etiquetas que permitan establecer relaciones, si las variables independientes no tienen relación, no se podrá encontrar ninguna, por más que se ajusten los parámetros de los modelos.

