



Reto | Mercadotecnia telefónica con aprendizaje supervisado

Entregable final del reto

Documento editable por el participante

Nombre del participante: Bernardo Lozano Wise

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**.

Para la transformación de variables categóricas, se ha seleccionado el método **One-Hot Encoding** en lugar de Label Encoding. La razón fundamental es que la mayoría de las variables categóricas en este dataset son nominales, es decir, no tienen un orden intrínseco (por ejemplo, job o marital). Usar Label Encoder asignaría valores enteros (ej. "admin"=0, "student"=1, "technician"=2), lo que introduciría una relación ordinal artificial que no existe en la realidad. El modelo podría interpretar erróneamente que "technician" es "mayor que" "student", lo cual es incorrecto. One-Hot Encoding evita este problema al crear columnas binarias separadas para cada categoría, tratando a cada una como una entidad distinta sin imponer un orden jerárquico.

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?

Tras comparar los modelos de Regresión Logística y Red Neuronal, se concluyó que la Red Neuronal fue el mejor modelo para este problema, debido a su mayor capacidad para identificar correctamente a los clientes que Sí aceptarían la inversión (mayor Recall). Este modelo logró un Recall final de 72.72% para la clase "Acepta", lo cual es vital para el negocio, ya que el error más costoso es perder una oportunidad de venta. El uso de técnicas de Inteligencia Artificial en mercadotecnia transforma el telemarketing tradicional. La IA nos permite dejar de hacer llamadas a ciegas y, en su lugar, enfocar los recursos humanos y el tiempo del equipo de ventas solo en la pequeña porción de clientes que el



modelo predice como altamente probables de decir "sí". Esto aumenta drásticamente el retorno de la inversión de la campaña, haciendo que los esfuerzos de marketing sean más eficientes, más baratos y, lo más importante, más inteligentes.