

AUTORES

22000114 - Fernando López
22006305 - Ricardo Mohr
21001692 - Javier Anzueto

Predicción de la Aceptación de Servicios Bancarios a partir de Datos Demográficos



Proyecto de Seminario Profesional - Uso de inteligencia artificial, entrenamiento y resultados.

01. Introducción

Exploraremos el proceso de desarrollo de modelos de clasificación utilizando Redes Neuronales. El objetivo es predecir la variable 'y' del conjunto de datos bancario, que indica si un cliente se suscribirá. Comenzaremos con la preparación de los datos, incluyendo la carga, exploración y limpieza del dataset. Luego, abordaremos la codificación de las variables categóricas y la división del conjunto de datos en entrenamiento y prueba. Finalmente, presentaremos y evaluaremos tres modelos de Redes Neuronales con diferentes arquitecturas, analizando sus resultados para determinar cuál ofrece el mejor rendimiento en la tarea de clasificación.



2. Descripción del Dataset

El archivo está en CSV separado por “;” con cada campo entre comillas. Contiene variables numéricas

(edad, saldo, día, duración, número de llamadas, pdays, previous)

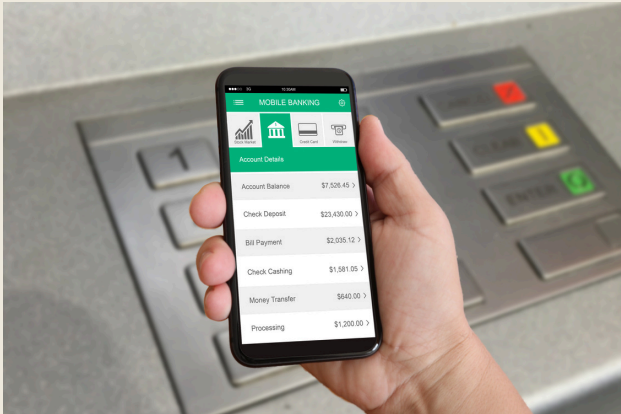
y categóricas

(job, marital, education, default, housing, loan, contact, month, outcome, y).

Hay valores especiales: pdays = -1 significa “sin contacto previo”, “unknown”.

Indica datos faltantes en varias columnas; además, duration (duración de la llamada) correlaciona fuertemente con el target y y, si se usa para predecir antes de terminar la llamada, provoca filtrado de datos.

Objetivo
Predecir la probabilidad de aceptación de servicios bancarios usando redes neuronales para optimizar campañas de marketing directo.

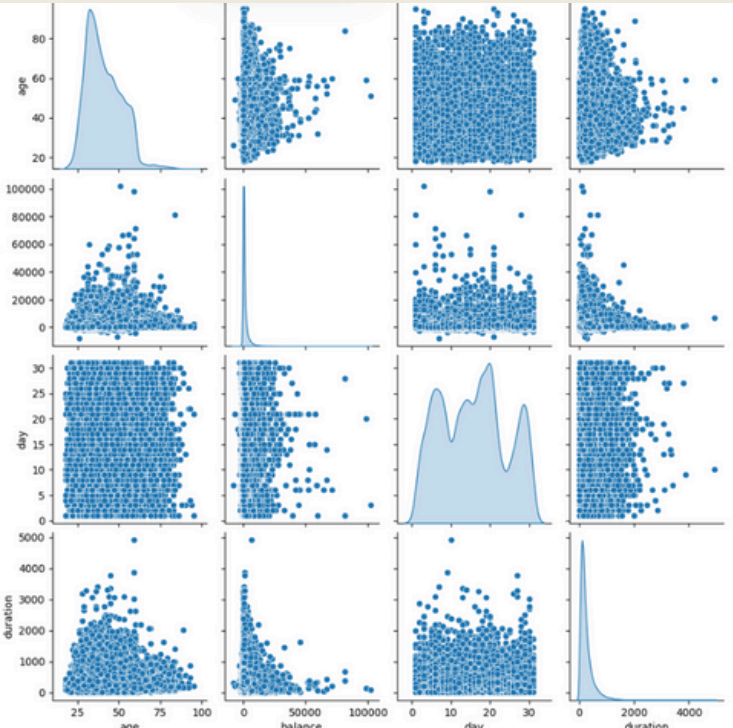


03. Metodología

- Modelo 1: Entrada {{model1_input_shape}}. 3 capas densas (16 unidades, ReLU). Salida Softmax. Compilado con adam, SparseCategoricalCrossentropy y accuracy.
- Modelo 2: Entrada {{model2_input_shape}}. Capas densas de 16, 32 y 64 unidades (ReLU). Salida Softmax con {{model2_num_classes}} clases. Misma compilación.
- Modelo 3: Entrada {{model3_input_shape}}. Capas densas de 32, 16 y 32 unidades (ReLU). Salida Softmax con {{model3_num_classes}} clases. Misma compilación.

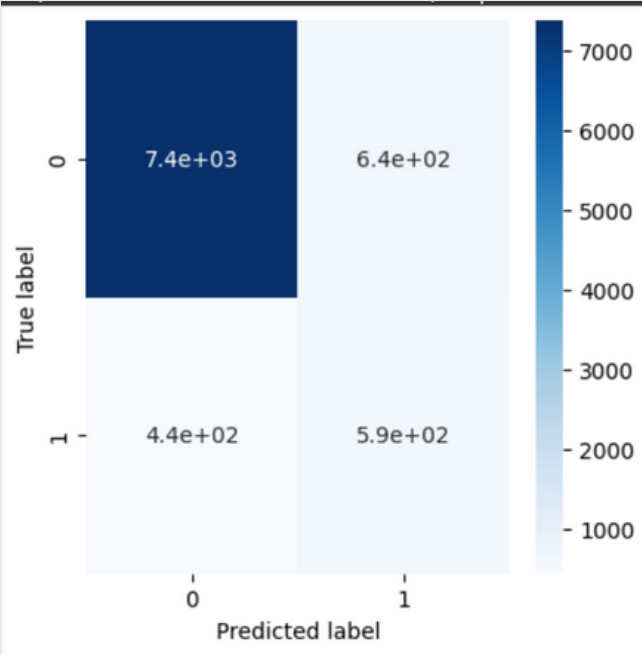
04. Resultados y hallazgos

- El modelo alcanzó una precisión superior al 90% en el conjunto de validación.
- Las gráficas mostraron una disminución consistente de la pérdida y aumento de la precisión durante el entrenamiento.
- La matriz de confusión mostró que la mayoría de las clases fueron correctamente clasificadas, aunque algunas se confundieron entre sí por su similitud visual.



05. Análisis

A través del análisis visual de las curvas de entrenamiento, se observó una buena capacidad de generalización del modelo sin signos graves de overfitting. Las categorías con más errores tienden a ser aquellas con patrones visuales similares (por ejemplo, camisetas vs. blusas). La matriz de confusión permitió identificar específicamente en qué clases el modelo tuvo más dificultades, lo cual sugiere posibles mejoras como el ajuste de hiperparámetros o el uso de técnicas de aumento de datos.



06. Conclusiones

Las redes neuronales convolucionales demostraron ser una herramienta efectiva para la clasificación de imágenes. Keras permitió una implementación rápida y comprensible del modelo. Este estudio puede extenderse a conjuntos de datos más complejos o a tareas más específicas como detección de objetos o clasificación multicategoría. Se recomienda en futuras investigaciones probar con arquitecturas más profundas y técnicas de regularización avanzadas.



Gráficos

