Respuestas:

Una vez realizada tu solución, responde a las siguientes preguntas:

1. De las deficiencias en los datos, ¿cuáles y como las identificaste?

Durante la fase inicial se encontró lo siguiente:

* Existencia de datos fuera de especificación en las variables: Educasión, genero, y estado civil. Se detectó, validando que fuera consistente con la descripción de los datos. Dado que son pocos, 396, se tomó la decisión de eliminarlos.
* Las variables de pagos (PAY\_0, …, PAY\_6). Se observó que el rango parece estar rezagado un periodo, ya que en la descripción menciona que va de -1 a 9, y los datos muestra un desfase que va de -2 hasta 8, por lo cual se consideró hacer lo siguiente: los valores menores a 0 se etiquetaran como -1, y os mayores o iguales a 0, se les sumará 1.

1. De realizar creación de variables, explica cuales hiciste y por qué.

No se realizó creación de variables.

1. De los modelos realizados, ¿cómo seleccionaste al mejor?

Verificando las métricas

1. ¿Qué desafíos encontraste y cómo los superaste?

El desbalance de clases, es un problema que a veces no es tan sencillo como parece, sin embargo, para el problema en concreto, funcionó considerablemente bien aplicar un oversampling.