

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
FACULTAD DE INGENIERÍA Y CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA [METODOLOGÍA VIRTUAL]
TALLER 3 **Análisis de Regresión**
Modelo de Regresión Logística

PROFESOR: SÉBASTIEN LOZANO FORERO
slozanof@libertadores.edu.co

INSTRUCCIONES

- Determine si es de su interés someter sus eventuales resultados a la plataforma Kaggle, en el desafío [Titanic](#).
- Si le resulta interesante la idea de someter sus resultados a Kaggle, utilice los datos dados [aquí](#). En este caso, el entregable constará únicamente del comprobante de la sumisión junto con los resultados dados por la plataforma.
- Caso contrario, utilice los datos disponibles [aquí](#). Aquí, el entregable constará de un pdf o Jupyter notebook con la presentación de su solución a los siguientes ítems
- Plazo Máximo de entrega: Domingo 5 de Diciembre de 2021 (23h59).

Considere la base de datos **Titanic** seleccionada. A partir de la misma resuelva:

1. Realice una corta presentación del conjunto de datos asegurándose de entender cada variable.
2. Presente los principales resultados de un Análisis Descriptivo y Exploratorio de los datos. Haga énfasis en la relación que pueda existir entre las variables.
3. Realice una partición de la base de datos: 70% como base de entrenamiento (**Train**) y 30% como base de evaluación (**Test**).
4. Proponga el mejor modelo de regresión logística para la variable **survival**, basado en el conjunto de datos **Train**.
5. Reporte los valores de varias medidas de bondad de ajuste presentado.
6. Genere la predicción para los individuos de la base de datos **Test**.
7. Presente la matriz de confusión, Presicion y Recall del modelo obtenido. Es un buen modelo?
8. Interprete su modelo en términos de las variables que incluye.