

Universidad Juárez del Estado de Durango
Facultad de Ciencias Exactas
Maestría en Estadística Aplicada
Modelos Lineales
Semestre A 2023

Instrucciones

De acuerdo con el Syllabus, se otorgará una base de datos a los estudiantes que consista en un conjunto de entrenamiento y un conjunto de prueba. De manera individual, cada alumno ajustará un modelo al conjunto de entrenamiento. Con dicho modelo deberá predecir la variable de respuesta para el conjunto de prueba.

Base de datos

En la plataforma se encuentran el archivo **DatosParaProyecto.RData** con dos conjuntos de datos: **Train** y **Test**. Como se vio en clase, el primer conjunto de datos deberá ser usado para ajustar el modelo para predecir la variable *Loan.Status*, que no se encuentra presente en *Test*.

Entregables

Podrá realizar cualquier análisis que desee (correlaciones, gráficos, modelos, etc.). No será necesario entregar un informe redactado, solo diapositivas con los resultados de esos análisis (Salidas de consola, gráficos, etc.) en buena calidad (imágenes con buena definición, tablas con formato, títulos, etc.).

Las predicciones para *Test* deberán ir en un dataframe en R que contenga, al menos, dos columnas: *ID*, identificador de la predicción (que coincida con el ID del objeto *Test*) y *Predicción* que contenga uno de dos valores: "*Fully Paid*" o "*Charged Off*".

Deberá subir a la plataforma, además, los siguientes archivos:

Código: **ML – ABCD – Proyecto Final.R**

Predicciones: **ML – ABCD – Proyecto Final.RData**

Presentación: **ML – ABCD – Proyecto Final.pdf**

Presentación

La fecha para la presentación de resultados se definirá el día 21 de abril de 2023, dependiendo de los avances del grupo.

Calificación

La calificación será el porcentaje de predicciones correctas de la variable de respuesta en el objeto *Test*. Se darán puntos extra por la inclusión de análisis (gráfico o numérico) no revisado en clase.

Dudas

Para dudas o comentarios, puede hacerlos llegar al correo diana.barraza@ujed.mx o por WhatsApp.