**PARCIAL 1: Machine Learning**

**Fecha de entrega: 6 de abril**

**Utilice la base de datos de retiros de la clase 8.**

**Problema de clasificación**

1. Realizar un análisis exploratorio con la información, identifique potenciales variables para pronosticar si un empleado se retira o no.
2. Pruebe diferentes modelos de clasificación sólo haciendo la muestra de entrenamiento de prueba y la muestra de test (regresión logística , Random Forest, Gradient Boosting, Adaptative Boosting). Cuantifique las siguientes métricas para evaluar la capacidad predictiva del modelo AUC, Accuracy, Sensibiilidad, Especiicidad. Investigue otras métricas (Kolmogorov - Smirnov). Pruebe con umbrales diferentes a 0.5 para evaluar que la sensibilidad y especificidad sean altas.
3. Utilice estrategias de validación cruzada para seleccionar un modelo ganador.

**Problema de predicción**

1. Elabore un análisis exploratorio de los datos que le permita identificar potenciales variables para el tiempo de duración en términos de las demás variables.
2. Ajuste diferentes modelos de pronostico, regresión líneal múltiple, regresión ridge, regresión lasso, árboles de decisión, random forest, gradient boosting ajustando su modelo a una muestra de entranamiento y una de test. Verifique el RMSE, MSE, investigue al menos una métrica adicional y estimela.
3. Utilice estrategias de validación cruzada para seleccionar un modelo ganador.