**MACHINE LEARNING SCIENTIST CAREER - DATACAMP**

**Enunciado Challenge N°1**

**Detección de transacciones fraudulentas en operaciones con tarjetas de crédito**

**Challenge N°1:**

**Problema**

Detección de transacciones fraudulentas en operaciones con tarjetas de crédito.

**Asignación**

Diseñar e implementar un modelo predictivo basado en las técnicas de machine learning vistas en los módulos 1 a 4 (preferiblemente) del curso de ML Scientist de DataCamp para identificar transacciones fraudulentas en operaciones con tarjetas de crédito.

**Requerimientos**

* **Lenguaje:** Python.
* **Tiempo:** 2 semanas (acordar inicio).
* **Métrica:** Balanced accuracy o alguna basada en Confusion Matrix o ROC Curve (a definir).

**Perturbación de los datos:**

Cada integrante debe sugerir una modificación a realizar sobre el dataset original, de las cuales se seleccionarán 3 según concenso y definición del grupo.

Ejemplos:

* Reemplazar valores por NaNs en algún feature.
* Agregar una columna nueva.

**Evaluación y Puntuación**

Se evaluarán los siguientes aspectos:

* Prolijidad del código (25%).
* Diseño del modelo (25%).
* Performance alcanzada por el modelo (50%).

La evaluación se realizará de manera grupal en donde, para cada entrega, todos los miembros del grupo asignarán una puntuación entre 0 y 10 a cada aspecto de evaluación. Claramente, los integrantes no sabrán a quien están evaluando, salvo el dueño de la entrega, quien debe realizar una auto-evaluación de su entrega a partir de los comentarios de sus pares.

**Entrega**

Finalizado el “deadline” de entrega, los códigos deben ser commiteados al repositorio de GitHub principal, en donde no pueden realizarse modificaciones posteriores. Incumplir lo anterior es motivo de descalificación y se asignará una penalización a acordar entre todos.

**Premiación**

El ganador será elegido como el modelo con el mejor puntaje total a partir de los 3 aspectos evaluados. El premio DEBE ser enmarcado dentro de una salida a comer/birras/etc.