

Ejercicio de Data Scientist

Objetivo

El objetivo del ejercicio es poder predecir quiénes de los usuarios que instalaron en el día de hoy van a dejar de jugar el día siguiente.

El método de evaluación de performance del modelo queda a criterio del candidato.

Datos

El dataset en cuestión cuenta de los siguientes campos:

- user_id: identificador unívoco anonimizado del usuario.
- install_time: indica la fecha y horario en que se registró la instalación (UTC).
- platform: plataforma del usuario, puede ser: Android o iOS.
- country region: sería el equivalente a provincia.
- city: ciudad.
- gender: género del usuario, puede ser: hombre o mujer.
- min_age_range: cota inferior del rango de edad.
- max age range: cota superior del rango de edad.
- event 1: cantidad de eventos de un tipo anonimizado.
- event 2: cantidad de eventos de un tipo anonimizado.
- event_3: cantidad de eventos de un tipo anonimizado.
- event_4: cantidad de eventos de un tipo anonimizado.
- event_5: cantidad de eventos de un tipo anonimizado.
- target_churn_indicator: es la variable a predecir. Si es 1 significa que no jugó el día siguiente, en cambio, si es 0 significa que jugó.

Entregable

Se necesita que se explique el modelo entrenado y validado, mediante una presentación. No hay ninguna restricción en cuanto a las herramientas para desarrollar el modelo y la presentación.

Por último, sería recomendable que en la presentación se incluyan algunos gráficos o conclusiones sobre el análisis exploratorio de los datos.

La presentación junto con el código para crear el modelo y validarlo debe ser enviada por mail como respuesta al envío de este ejercicio.