

PEA Advanced Data Science Supervised Methods

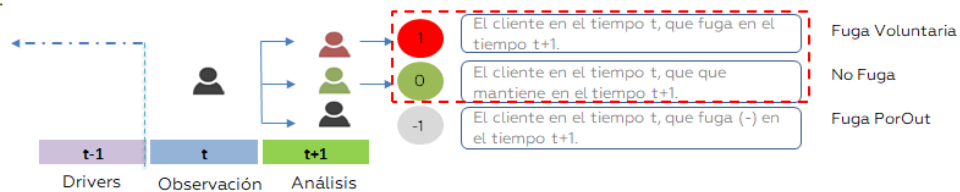
Se entrega el siguiente set de datos de una empresa del sector de telecomunicaciones, TelcoCustomerChurn.csv, el cual es un dataset que cuenta con las principales características de una empresa del sector de telecomunicaciones respecto a sus campañas, facturación y vínculo con el cliente.

Esquema metodológico de la fuga de clientes

Cosechas o periodos de estudio del Churn

Ene-19	Feb-19	Mar-19	Abr-19	May-19	Jun-19	Jul-19
t-3	t-2	t-1	t	t+1	t+2	
	t-3	t-2	t-1	t	t+1	t+2
		t-3	t-2	t-1	t	t+1

Donde :



La empresa en el último mes y debido a la coyuntura del COVID – 19 ha incrementado en + 0.5 pp su ratio de churn de clientes, pasando de 2% en promedio de los últimos 6 meses a 2.5% en los últimos 6 siguientes. Como sabemos el costo de adquisición de un nuevo cliente es aproximadamente 6 a 7 veces mayor que retener uno que ya tenemos.

¿Cómo puede la institución de telecomunicaciones tener una mayor efectividad para identificar a los clientes propensos a la fuga?

Para responder a esto, tenemos que analizar la información de toda la facturación y vínculo que tiene la institución o empresa con los clientes para identificar los patrones que nos ayudarán a encontrar conclusiones para desarrollar estrategias futuras. La empresa necesita saber dimensionar y bajar el ratio de fuga futuro de los clientes.

- Las variables que se disponibilizan son:

Variable	Descripción de la Variable.
customerID	Identificador único del cliente.
gender	Género de la persona o empresa.
SeniorCitizen	Estado civil del cliente o persona.
Partner	El cliente es una persona mayor o no?
Dependents	Número de hijos de la persona o cliente?
tenure	Número de meses que el cliente ha estado con la compañía.
PhoneService	El cliente tiene servicio de teléfono o no?
MultipleLines	El cliente tiene varios servicios de teléfono o no?
InternetService	Tipo de servicio de internet del cliente.
OnlineSecurity	El cliente tiene seguridad en la línea o no?
OnlineBackup	El cliente tiene respaldo en la línea o no?
DeviceProtection	El cliente tiene protección en el dispositivo o no?
TechSupport	El cliente tiene soporte técnico o no?
StreamingTV	El cliente tiene servicio de streaming de TV o no?
StreamingMovies	El cliente tiene servicio de streaming de películas o no?
Contract	Plazo del contrato de cliente.
PaperlessBilling	El cliente tiene facturación electrónica.
PaymentMethod	Método de pago del cliente.
MonthlyCharges	Importe cobrado al cliente mensualmente.
TotalCharges	Importe cobrado total al cliente.
Churn	El cliente se fue de la empresa o no?

Entregables del TA:

- 1.- Definir el problema de la naturaleza que se tiene a continuación, además de los objetivos de negocio bien definidos.
- 2.- ¿Qué tipo de variables se utilizan en el problema de negocio?
- 3.- ¿Desarrollar un pequeño análisis de exploratorio?
- 4—Aplicar las técnicas de balanceo previamente particionando el dataset en train 80% y 20% en test.
- 5– Probar usando el mismo modelo de clasificación que fue Random Forest y validar cual de los 3 tiene el mejor accuracy y la mejor precision en la clase objetivo que

1

La entrega hacerlo en un jupyter notebook y subirlo a la plataforma poniendo trabajo1_nombre, cualquier duda o consulta avisarme. Por otro lado el trabajo tien plazo máximo 1 semana de entrega (tomar precauciones).