

# Propuesta de Tesis

MASTER IN MANAGEMENT + ANALYTICS

PREDICCIÓN DE UPSELLING CON TÉCNICAS DE MACHINE  
LEARNING PARA EMPRESA DE TELECOMUNICACIONES

**Darmograj, Andrés**

27/12/2019

## Propuesta

El objetivo es desarrollar un modelo predictivo que, analizando el comportamiento pasado de clientes móviles del segmento B2B que realizaron una operación de cambio de plan ascendente, asigne una probabilidad para cada cliente de realizar dicha operación durante el próximo mes.

## Modelo Predictivo

El modelo intentará predecir con alguna técnica de Machine Learning la probabilidad de que cada uno de nuestros clientes móviles B2B haga una operación de cambio de plan ascendente.

Se deberá generar el dataset con la respectiva clase a predecir, aplicar técnicas de data cleansing, crear o transformar variables, seleccionar variables relevantes que expliquen el fenómeno a predecir, evaluar cómo separar el dataset en train-test, analizar performance de distintos modelos para luego ejecutar la primera prueba piloto. Debido al volumen de datos, es necesario que el mismo pueda correrse en un cluster dedicado, apoyados por la librería PySpark.

El modelo deberá entrenarse para predecir a 2 meses vista, ya que la predicción del mes  $n$  se ejecuta en  $n-1$  con información hasta el cierre de  $n-2$ .

## Información disponible

La empresa vela por el respeto absoluto a la confidencialidad y privacidad de los datos personales de cada uno de nuestros clientes. Por ello, a efectos de este trabajo, los datos serán anonimizados y algunas variables críticas para el negocio estandarizadas.

Se utilizarán las siguientes fuentes de información:

- ABT\_Suscripcion\_Movil: base que incluye datos mensuales pre-procesados de todo el parque móvil. Contiene actualmente 377 variables, y se disponibiliza todos los días 12 de cada mes. El histórico del segmento B2B comenzó a cargarse en este tablón a partir de Mayo-2019, por lo cual hasta el día de la fecha ya contamos con más de 6 meses de historia.
- ICN\_Campanias\_Envios\_His: alberga el histórico de todas las bases enviadas a campañas, con su respectivo período de vigencia. Está alojada en otro sistema (datawarehouse móvil), por lo cual será necesario importar la información vía *scoop* a Hadoop.

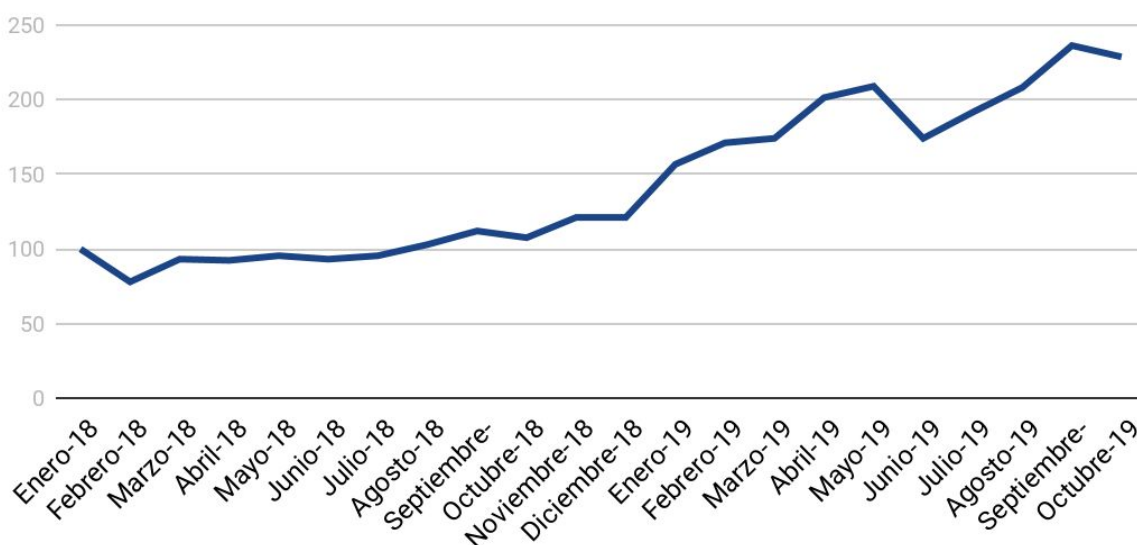
- Campanias\_B2B: alojará la información de campañas importada del datawarehouse móvil a Hadoop. Incluirá todas las campañas, para facilitar la elaboración de futuros modelos.
- PDB\_Operaciones\_Detalle: incluye detalle de todas las operaciones comerciales realizadas por nuestros clientes. Cada registro corresponde a una operación. Contiene en total 90 variables, entre las que se puede consultar información del cliente, información de la operación realizada y canal por el cual se gestionó.

## Aplicación de negocio e impacto

A partir de la sanción del [Decreto 1340/2017](#), y con el fin de promover la competencia y la “convergencia tecnológica”, se habilitó a las telefónicas a ofrecer servicios audiovisuales por cable. Esto inició la era del cuádruple play (4P) en Argentina, con las empresas ofreciendo descuentos por la contratación de ofertas multi-productos. Estas ofertas 4P se promueven con el fin de blindar al cliente desde la parte fija, dificultando a la competencia ingresar al hogar.

Desde enero 2019, fecha de plena puesta en vigencia de la reglamentación, el mercado se recalienta, con ofertas de captación agresivas y multi-producto para determinadas regiones. A continuación se puede apreciar cómo evolucionó la tasa de churn debido a este fenómeno:

### Evolución de Tasa de Churn - Individuos Móviles



En este contexto, con una competencia cada vez más agresiva y con un mercado móvil saturado, se tornó relevante la adquisición de ingresos alternativos vía desarrollo de clientes. Para lo cual se diseñaron inicialmente distintos modelos predictivos enfocados en el segmento B2C, ya que este segmento acumula la mayor cantidad de ingresos operativos.

Con los mismos se logró, por ejemplo, incrementar los ingresos unitarios por contacto de campañas en un 10%, sin aumentar la cantidad de representantes destinados a la tarea.

En una segunda etapa, se propuso comenzar a trabajar sobre B2B, con el fin de optimizar la base mensual enviada a campañas para ser gestionada. La meta es elevar la efectividad de las campañas, sin aumentar los recursos destinados a las mismas. Una vez implementados los nuevos modelos, se espera tener un impacto similar al visto en B2C.

Se empezará por el modelo de Upselling, ya que es la campaña que genera mayor cantidad de ingresos.