



Universidad Nacional  
de La Matanza



## Predicción BAJA+2 paquete Premium

Bettachini, Víctor A.

V Encuentro *Mejora de las Estrategias Pedagógicas*



# Contenido

1 Objetivo

2 Journal



## Objetivo

Predecir bajas de clientes del Paquete Premium con dos meses de anticipación

## Herramientas, datos, métricas

- Algoritmo: LightGBM(lightgbm\_3.3.5) en lenguaje R(R6\_2.5.1)
- Datos: 152 atributos, 154k clientes, fotos 201901 a 202109
- Evaluación: BAJA+2 de la foto de 202109 solo con foto 202107



# Análisis exploratorio del conjunto de datos

Cosas dataset crudo Anotar limpieza con NULL Coalescencia de atributos  
Creación de nuevos atributos Data engineering histórico lags deltas medias  
móviles



# Análisis exploratorio del conjunto de datos

Del análisis exploratorio:



# Feature engineering: nuevas variables



# Feature engineering: históricos



## Parámetros

- sin undersampling
- semilla\_azar: primera 777787
- lgb\_semilla: segunda 274837
- testing 20216
- valdation 20215
- meses entrenamiento: 201901 a 202104 excepto “confinamiento”: 202003 a 202011 inclusive

## Resultados iteración 26

- learning\_rate: pegado al límite inferior 0.02 → explorar menores
- num\_leaves: pegado a superior 2048 → explorar mayores
- min\_data\_in\_leaf: 3499





## Parámetros

- meses entrenamiento: ídem 823\_1123

No genera archivos de envíos:



