

Modelo de attrition para la campaña de retención proactiva de clientes

Objetivo

- Bajas del *Paquete Premium*, 154 k clientes
- Predicción dos meses previos (BAJA+2)

Estimación del modelo finalizado

- # Envíos de ofertas de retención = 9.5k
- Ganancia = 125.4 M AR\$

Algoritmo

- LightGBM (lightgbm_3.3.5)
- Entrenamiento y ensayo contra foto 202107
→ reutilización código z824
- Optimización Bayesiana de hiperparámetros
→ z823

Insumo

Datos crudos

- 152 atributos de 154k clientes
- fotos 201901 a 202109

Baseline \Rightarrow 456 atributos

- lag 1, 3 y 6 meses
- Extensión JupySQL operando sobre base SQL DuckDB (duckdb_0.2.6)

sql_eng_baseline.ipynb

FEH 1 = baseline + \Rightarrow 912 atributos

- lags 2, 4 y 5 meses \rightarrow semestre
- min, máx, media móvil semestre
- drop reiterados
- $\text{atm_oth} > \text{atm_oth}$
- tarjeta = visa + mc (m/c)
- proxy actividad = $\text{sum}[\text{abs}(m^*)]$

sql_eng_all6cat.ipynb

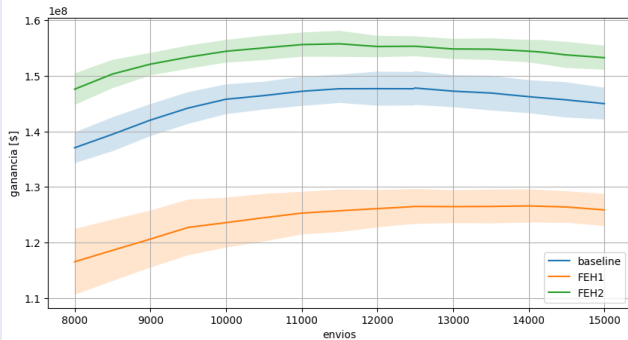
FEH 2 = FEH 1 + \Rightarrow 2128 atributos

- normalización semestral: $\text{atr} / \text{media semestral}$
- Δ normalizado $(\text{atr} - \text{lag\#}) / \text{media semestral}$

sql_eng_all6.ipynb

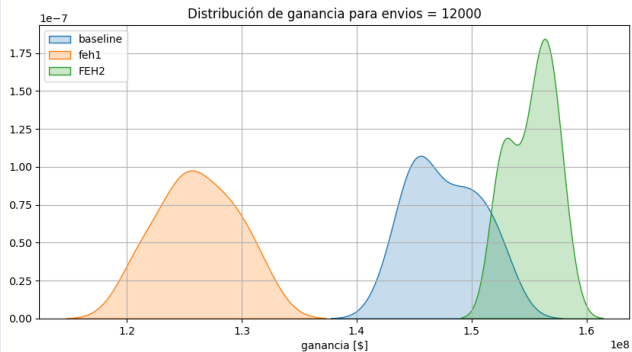
Rendimientos FEH 1 y FEH 2

Ganancia vs # Envíos



Distribución de ganancias

Envíos con máxima ganancia en FEH2



Catástrofes

Análisis de catástrofes

- Atributos montos (m^*) [flotantes]
- o cantidad (c^*) [enteros]
- >10 k registros $\implies 0.0$ / 0 en foto
- \implies NULL todos los registros en foto

sql_eng_all6cat.ipynb

sale

dritte824_1129.r

1129_null_sem.r