Optimización de parámetro a través de gridsearch en un

El sábado a las 15Hs, se lanza el desafío caza talentos

Los que quieran van a poder participar en el desafío que está al final del documento (puede ser en R, Python y Julia), quien entregue un programa que resuelva esto tiene uno o dos puntos. Quienes estén interesados lo hacen.

Problema clave,

Definir la función ganancia:

Se intenta predecir quien se va a ir dentro de dos meses. (Tener en cuenta los muertos)

El estímulo tiene un costo de 2000 (Según el valor de lo que va a dar el cliente es lo que uno tiene que estar dispuesto a gastar en mantenerlos)

Discriminar, para resolver.

En el mundo real, cada problema debe recibir un estímulo distinto.

Ganancia = 0,5 \* $160.000 \* aciertos - $2.000 \* envíos

Algo muy importante es saber definir bien la clase.

Ganancia $78.000 \* baja

Envío -$2.000 \* envío

En las campañas de marketing es muy común que solo compre el 1% y son buenas

Para entrenar (no para hacer el cálculo de ganancia) conviene unificar BAJA+1 y BAJA+2

Un solo train y test es una locura por la variabilidad.

Uno tiene una probabilidad de que un modelo sea mejor que otro