

Análisis de resultados de ensayos

CONFIDENCIAL



Objetivo

La empresa tiene que tomar la decisión de sacar un modelo a producción y para ello se pide realizar un análisis sobre los resultados de los últimos ensayos. [Datos - Raw.csv](#)

El objetivo es que usted analice los datos brindados, y con ellos elija el modelo productivo.

Contexto

Los datos brindados proporcionan los resultados de las detecciones de cuatro modelos evaluados a campo de diferentes clientes, presentados de la siguiente manera:

Objeto	Clase	SubClase	Especie	Detección			
				Modelo A	Modelo B	Modelo C	Modelo D
1	Weed	Grass	Specie 5	Si	Si	Si	Si
2	Weed	Broadleaf	Specie 5	No	Si	No	Si
3	Weed	Grass	Specie 3	Si	Si	No	Si

Se indican:

- **Clase:** clase objetivo.
- **Subclase:** familia de maleza asociada.
- **Especie:** especie de maleza asociada.
- **Detección Modelo:** detección del modelo X ("Si": fue detectada)

Si bien se agregan datos específicos asociados a cada objeto, el detector NO es multiclase, es decir, el modelo solamente arroja como resultado "Maleza" o "No maleza".

Se espera que el modelo productivo final tenga la mejor performance posible, tanto en porcentaje de detección como en ahorro de producto, brindando la mejor experiencia de uso a CADA cliente, entendiendo que el mismo espera detectar todas las malezas de su lote y no desperdiciar nada de producto.

Datos adicionales

- Tener en cuenta que el mismo modelo final se va a enfrentar a diferentes situaciones, y la distribución de cantidad y especies de malezas no es la misma para cada cliente. Por ejemplo podemos encontrar un Cliente A, que contenga 6 especies de malezas en su lote y a su vez podemos tener un Cliente B que solo tenga malezas de una única especie en su lote.



- Dentro de nuestro conjunto de datos sepáramos a todas las malezas en dos grandes grupos:
 - GRASS, y más específicamente estas en especies 1 a 3
 - BROADLEAF, y más específicamente estas en especies 4 a 6
- Los falsos positivos pueden llevar a que se desperdicie productos químicos que son muy costosos. Para medir esto, se realizaron sobre los ensayos una medición de True Negative Rate arrojando los siguientes resultados:
 - MODELO A: 95% TNR
 - MODELO B: 91% TNR
 - MODELO C: 92% TNR
 - MODELO D: 75% TNR

Nota

Deberá justificar su elección en base a su análisis y métricas utilizadas.