

# Potabilidad del agua

By Massobrio, Gutierrez y Tourn

Machine Learning II

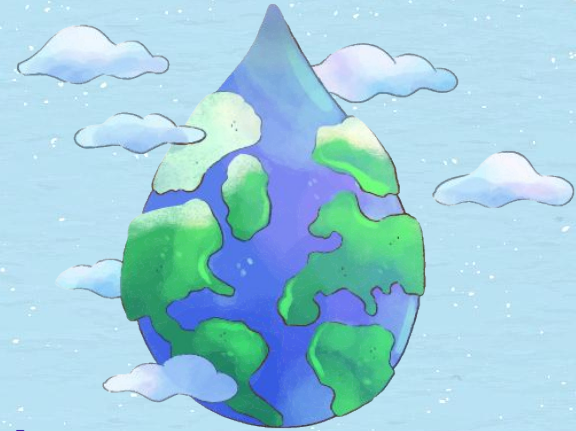


- 01 **Review**
- 02 **Esquema de implementación**
- 03 **Arquitectura**
- 04 **Conclusiones – Tareas futuras**

# 01 Review – Datos

## Predicción de potabilidad del agua

Desarrollo de modelo ML a partir de muestras experimentales de agua



### **PH**

Nivel de acidez del agua

### **Dureza**

Da cuenta de la concentración de calcio y magnesio

### **Sólidos**

Minerales y sales disueltos en agua

### **Cloraminas**

Presencia de sustancias desinfectantes

### **Turbidez**

Da cuenta de la cantidad de materia en suspensión

### **Trihalometanos**

Presencia de sustancias químicas con cloro

### **Carbón orgánico**

Materia en descomposición

### **Conductividad**

Da cuenta de la concentración de iones

### **Sulfatos**

Precipitados de minerales

# 01 Review – Primeros resultados

## Métrica: Precision

Se prioriza evitar falsos positivos. Incluso, es más importante que el hecho de detectar agua potable (recall alto).

## Mejor modelo inicial: Random Forest

Resultó el que mejor performance mostró para todas las métricas evaluadas. Aunque se puso el foco en *precision*.

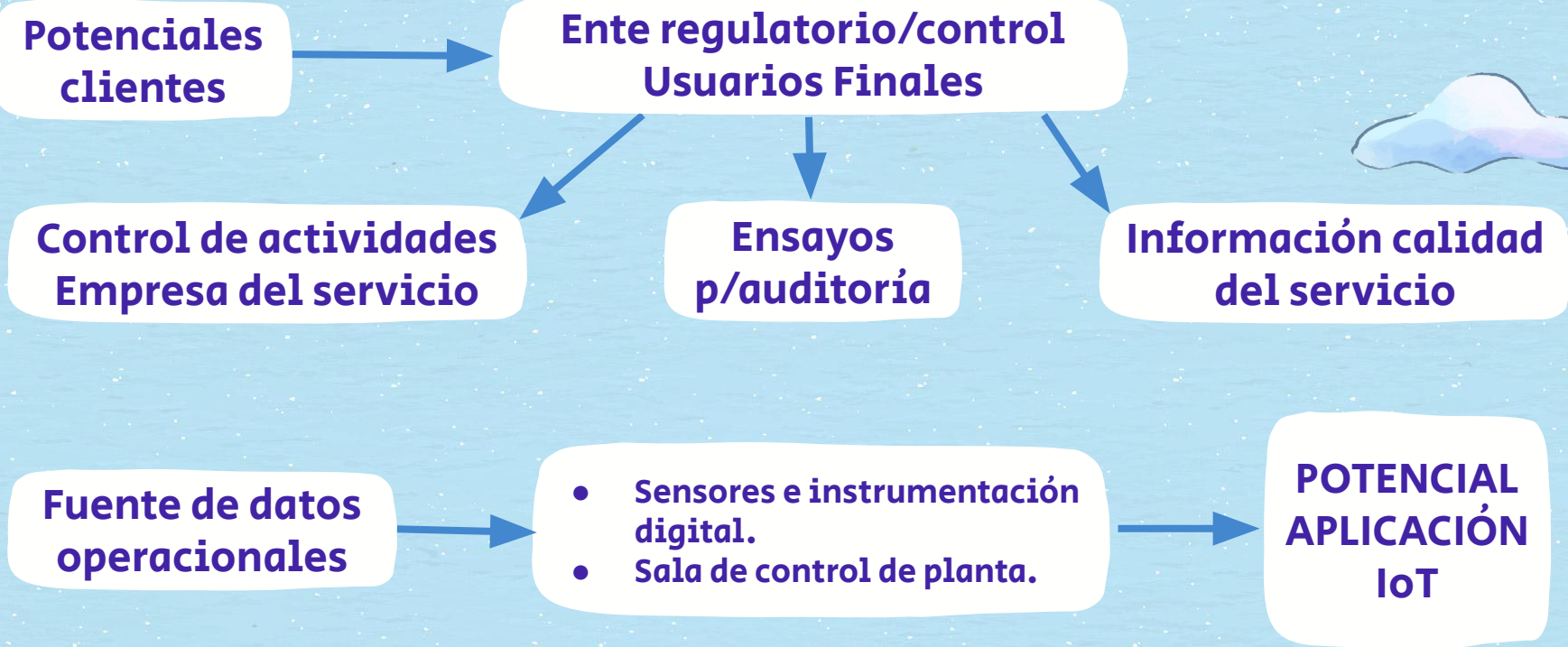
## Performance general – Despliegue:

No muy buena (mejor métrica 66%). Clase desbalanceada. Calidad de información. Modelos adicionales a entrenar durante despliegue:

- Decision Tree
- SVC
- Logistic Regression



# 02 Esquema de Implementación



# 02 Esquema de implementación

Un mismo despliegue – Dos modelos de trabajo asincrónicos

## OPERACIONES

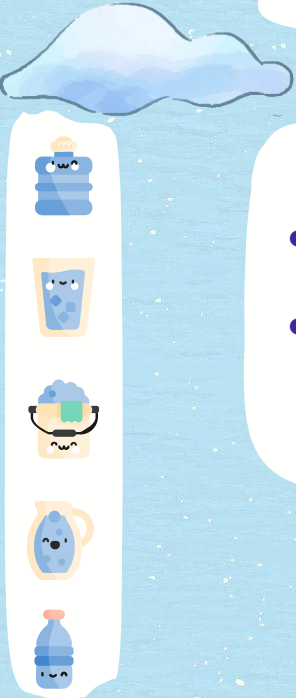
- Es quien requiere realizar inferencias
- Como dato de entrada emplea la información directa de planta

## CONTROL

- Ensayos de laboratorio desacoplado de operaciones.
- Los resultados son insumo del dataset de entrenamiento

## MLFLOW

- Gestión de inferencias.
- Gestión de promoción de modelos por reentrenamiento



## Arquitectura de solución Ente gestión de agua

Google Cloud Platform

Planta agua 1 - Inferencia



Sensor agua

### OPERACIONES

Planta agua 2 - Inferencia



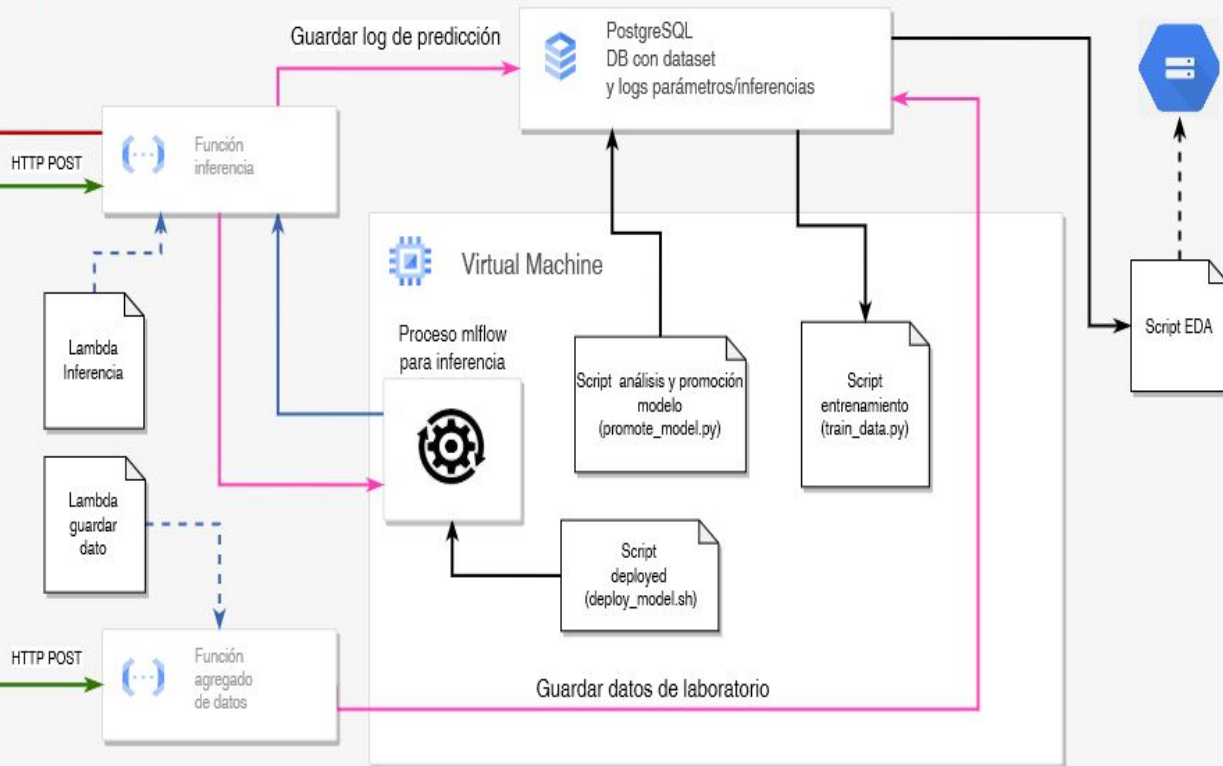
Sensor agua

Planta agua 1 - Laboratorio



### CONTROL

Planta agua 2 - Laboratorio



## 04 Conclusiones – Tareas futuras

- Se logró el deploy del modelo.
- Se logró gestionar los cambios de estado de procesos de entrenamientos mediante MLFLOW

### Tareas futuras

- Finalizar integración con Bucket para data analysis.
- Implementación de interfaz para usuarios finales de operaciones y control.



# Gracias!

¿Tiene alguna pregunta?

[cgmasso3@gmail.com](mailto:cgmasso3@gmail.com)

[lionelgutierrez@gmail.com](mailto:lionelgutierrez@gmail.com)

[b.tourn.86@gmail.com](mailto:b.tourn.86@gmail.com)

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

