

Estudio de la fractura de pozos de hidrocarburos mediante ciencia de datos

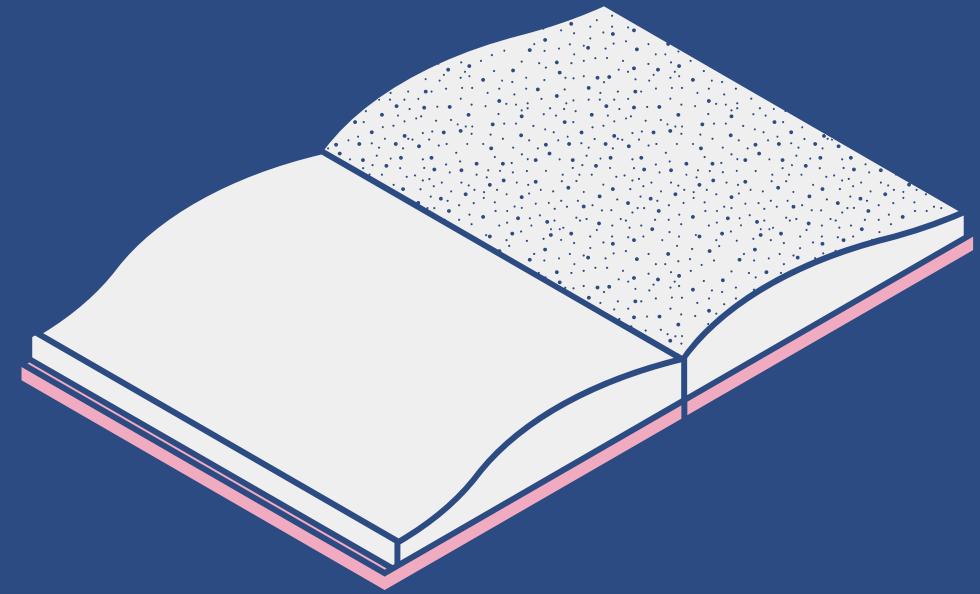


Grupo 1

Laura S. Gómez Velázquez, María Candela Herrera, Sofía Vaucelles

Problemática

Durante la perforación de los pozos se utiliza grandes cantidades significativas de agua para perforar la roca para luego realizar la extracción de petróleo.



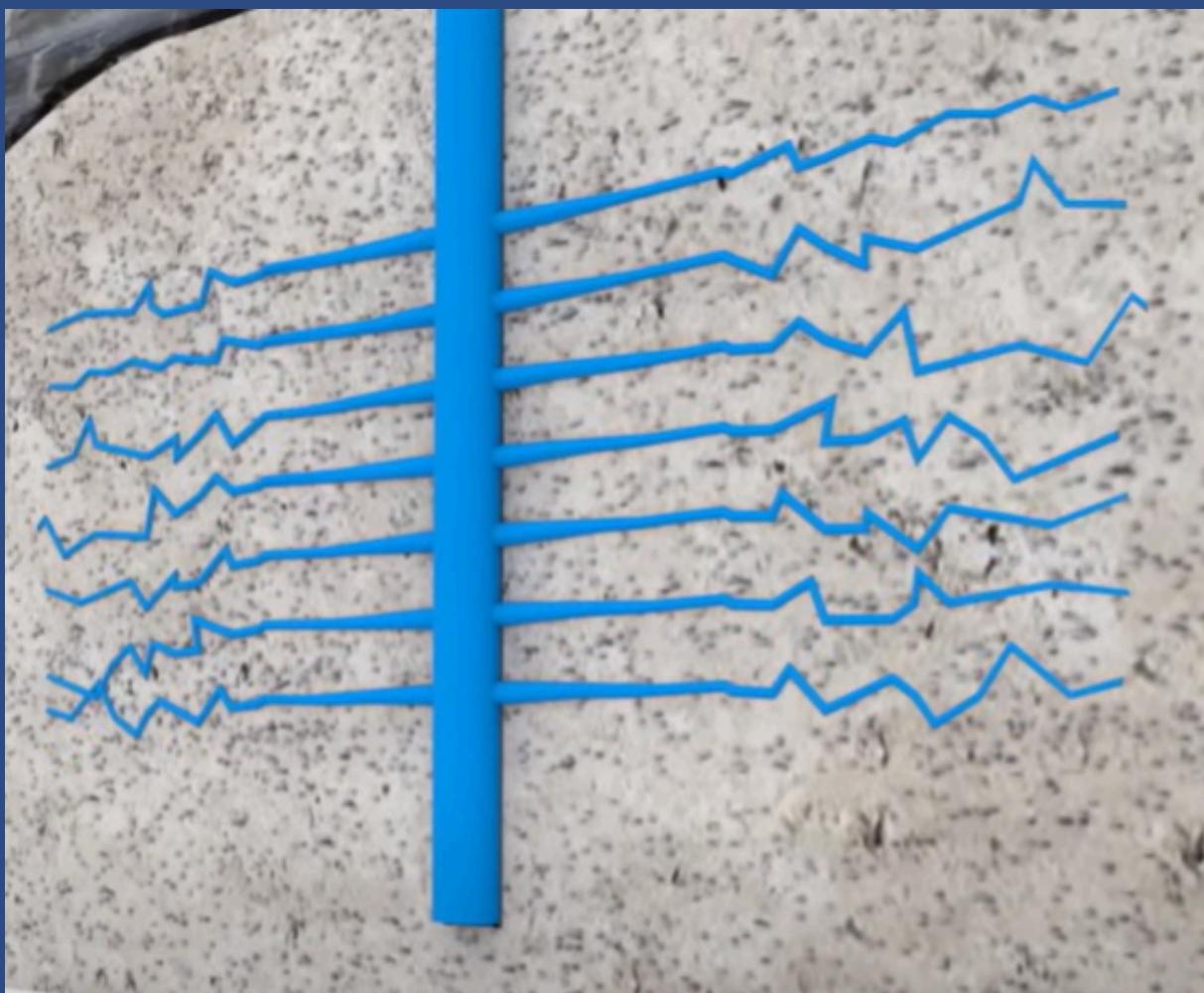
Relevancia

Al utilizar solo la cantidad mínima necesaria de agua se puede:

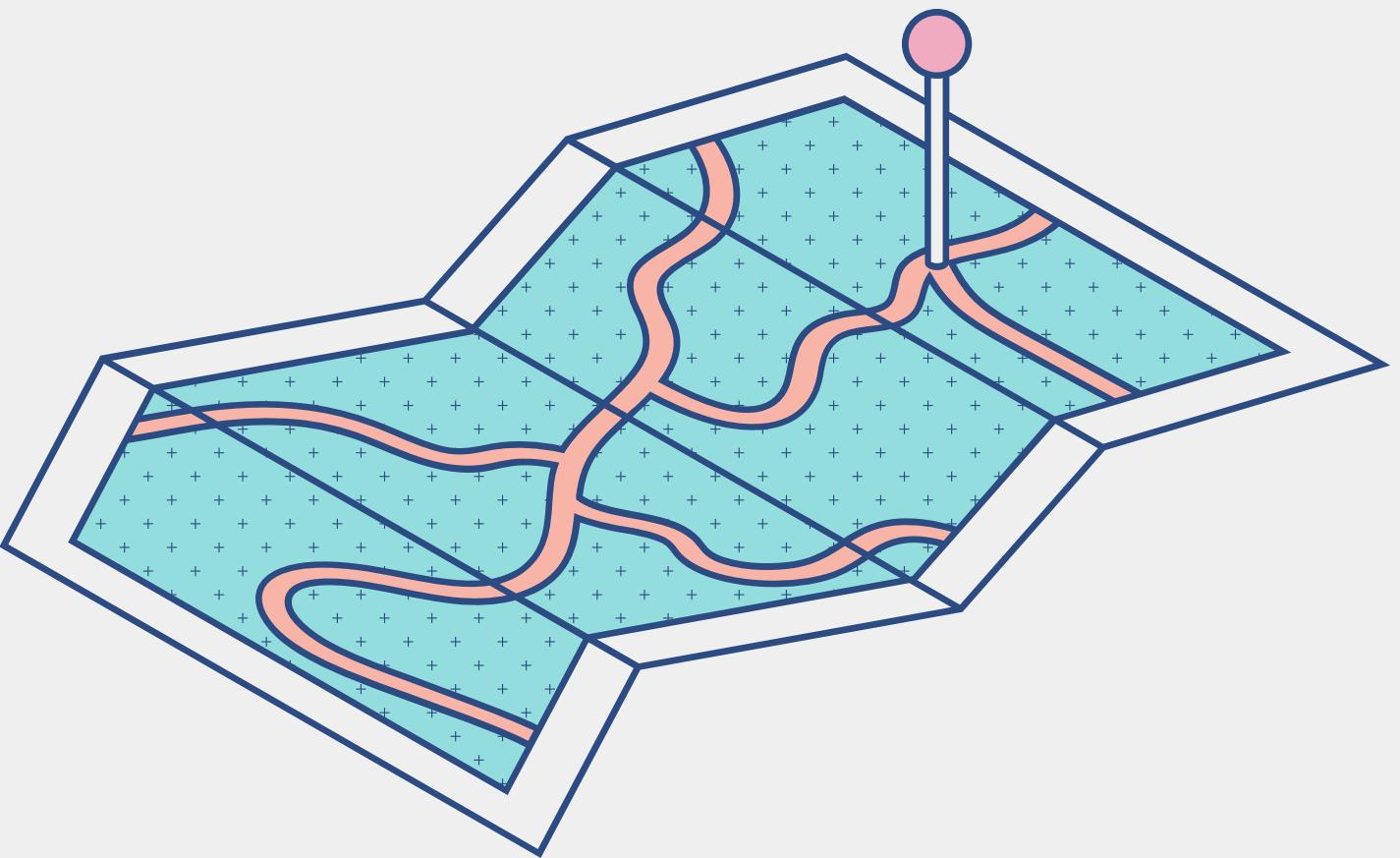
- Mitigar el impacto ambiental
- Reducir el costo (transporte)

Dataset

<http://datos.energia.gob.ar/dataset/>



Objetivos



Análisis Exploratorio de Datos: Identificar patrones y relaciones clave entre las variables del dataset de fractura de pozos de hidrocarburos.

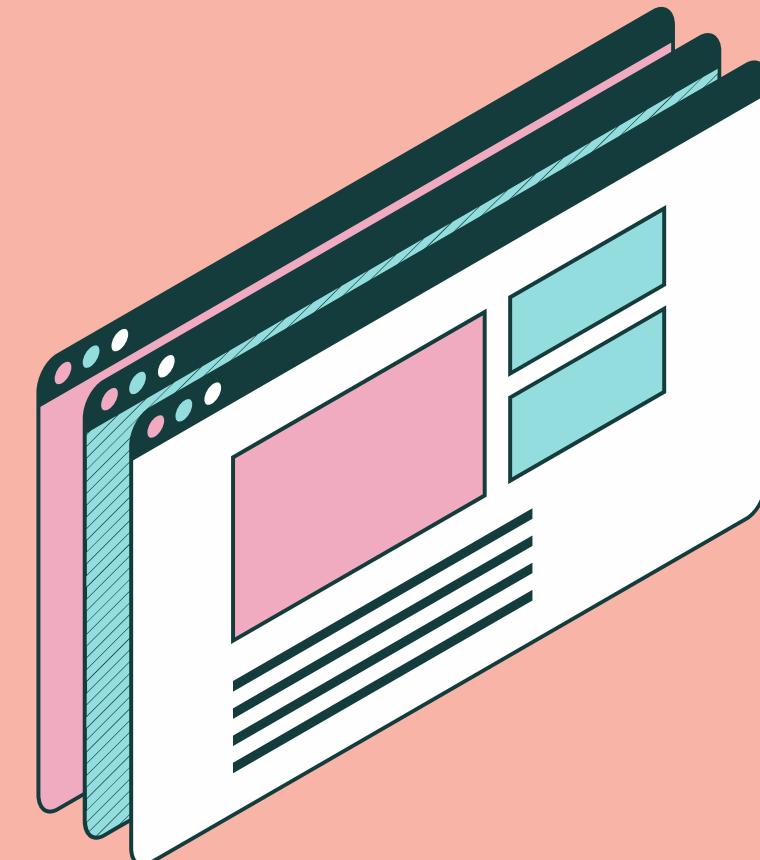
Modelos de Aprendizaje Supervisado:
Entrenar modelos de machine learning para predecir la cantidad necesaria de agua a inyectar, para mejorar la eficiencia del proceso, reducir el impacto ambiental y minimizar los costos operativos.

Modelos de Aprendizaje No Supervisados: Aplicar técnicas de clustering para segmentar los pozos en grupos homogéneos, identificando patrones operacionales y de rendimiento.

Descripción del dataset

Categoría	Columnas
Geología y Yacimiento	<ul style="list-style-type: none">- Cuenca- Yacimiento- Formación Productiva- Tipo de Reservorio- Subtipo de Reservorio
Operación del Pozo	<ul style="list-style-type: none">- Longitud de Rama Horizontal- Cantidad de Fracturas- Tipo de Terminación
Inyección y Bombeo	<ul style="list-style-type: none">- Arena Bombeada Nacional- Arena Bombeada Importada- Agua Inyectada- CO2 Inyectado

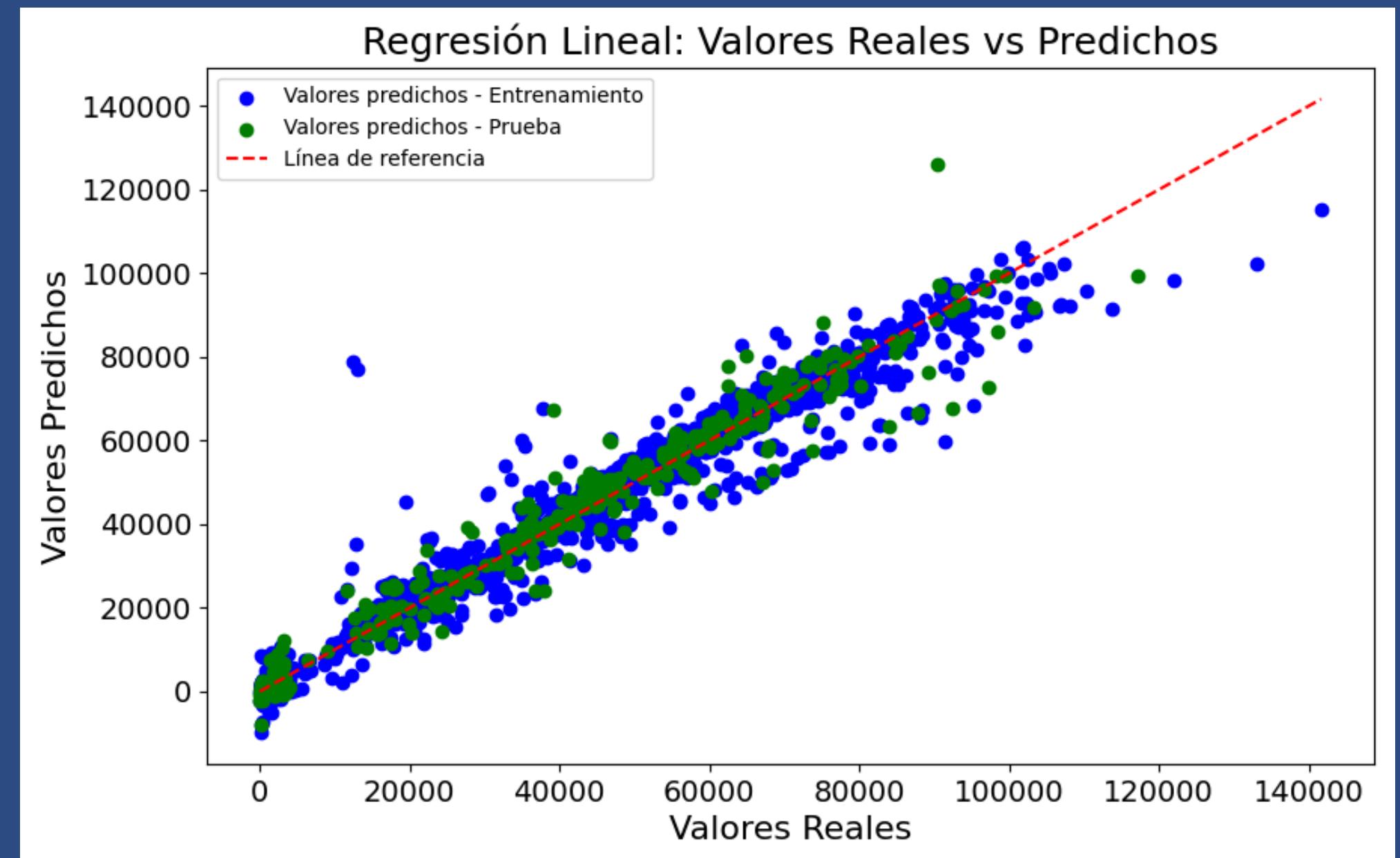
INFORMACIÓN SOBRE LAS VARIABLES CLAVE DEL DATASET



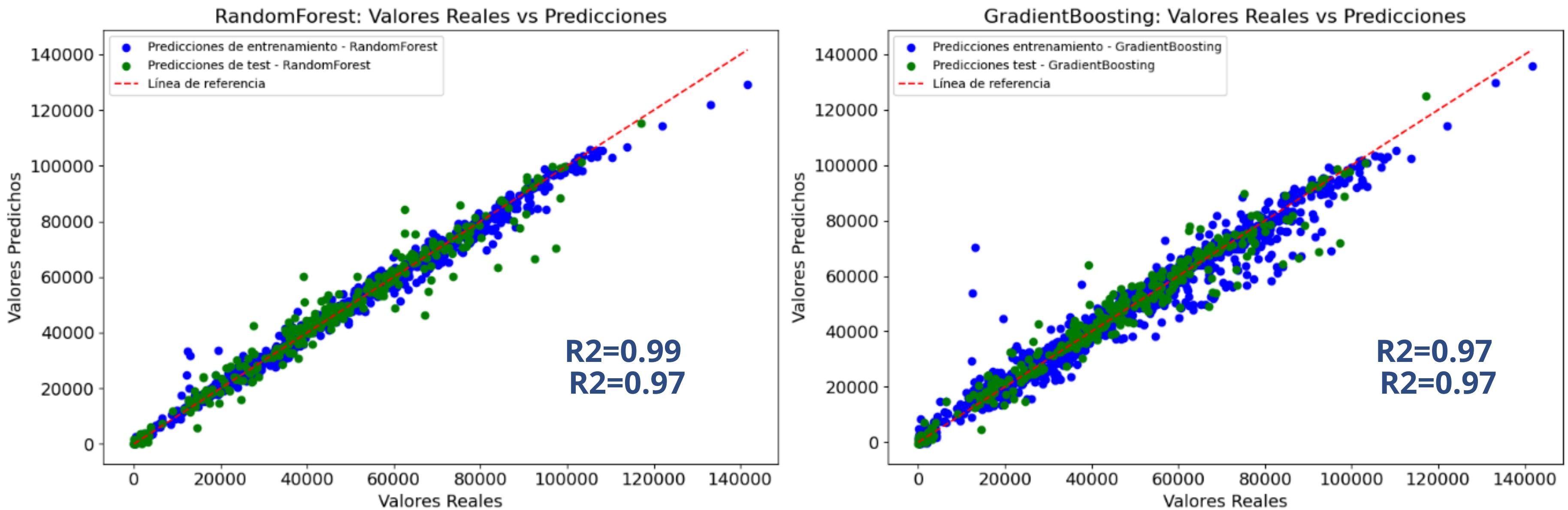
Categoría	Columnas
Rendimiento Operacional	<ul style="list-style-type: none">- Presión Máxima- Potencia de Equipos de Fractura
Temporalidad	<ul style="list-style-type: none">- Fecha de Inicio de Fractura- Fecha de Fin de Fractura

Modelo de Aprendizaje Supervisado

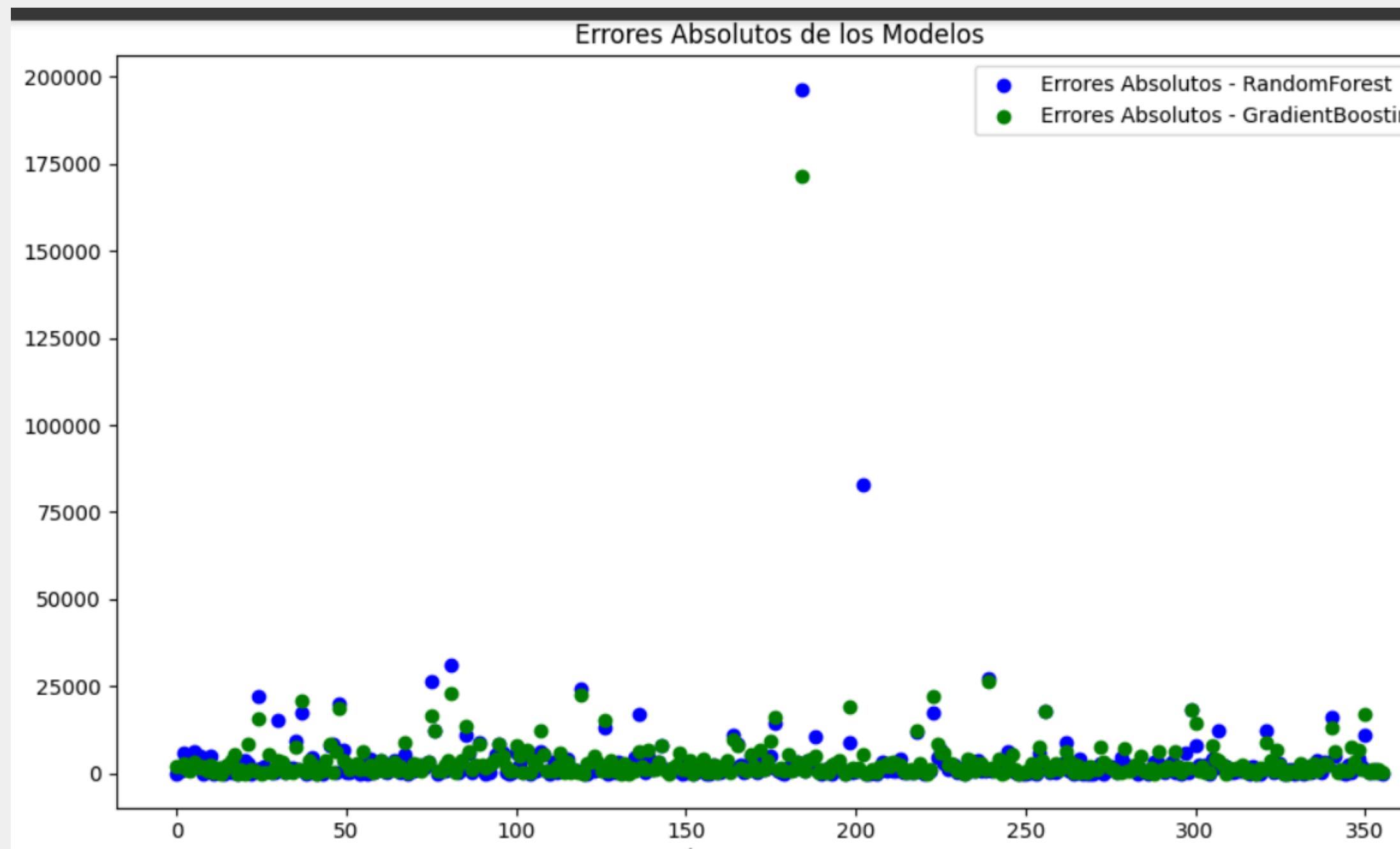
- Aplicamos tres modelos:
 - 1) Regresión lineal
 - 2) Random Forest
 - 3) Gradient Boosting Regresor



Interpretación de los resultados



Interpretación de los resultados



Podemos observar que los errores están distribuidos uniformemente y que además son bajos.

Modelos de Aprendizaje No Supervisado

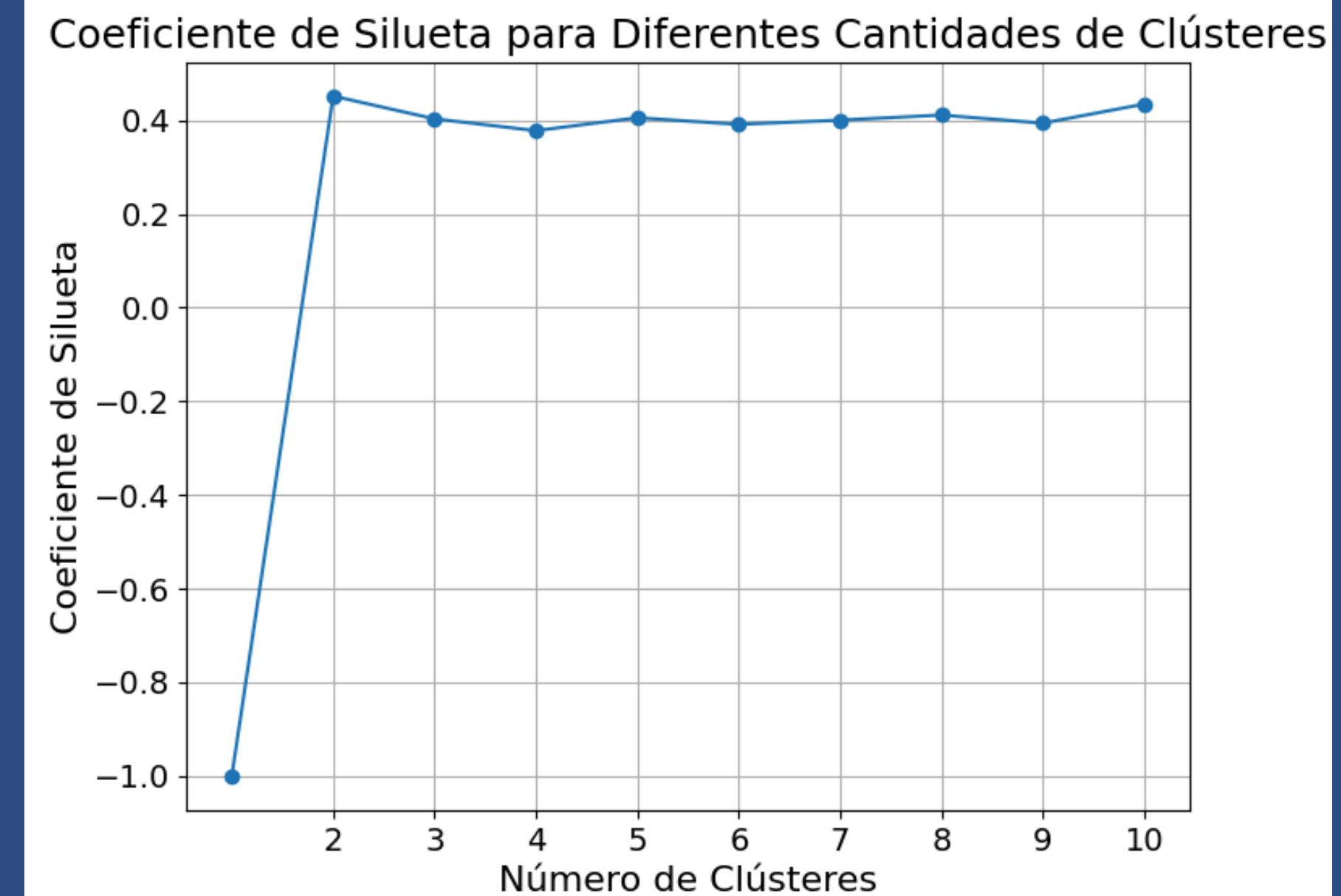
Implementación del Modelo

Kmeans

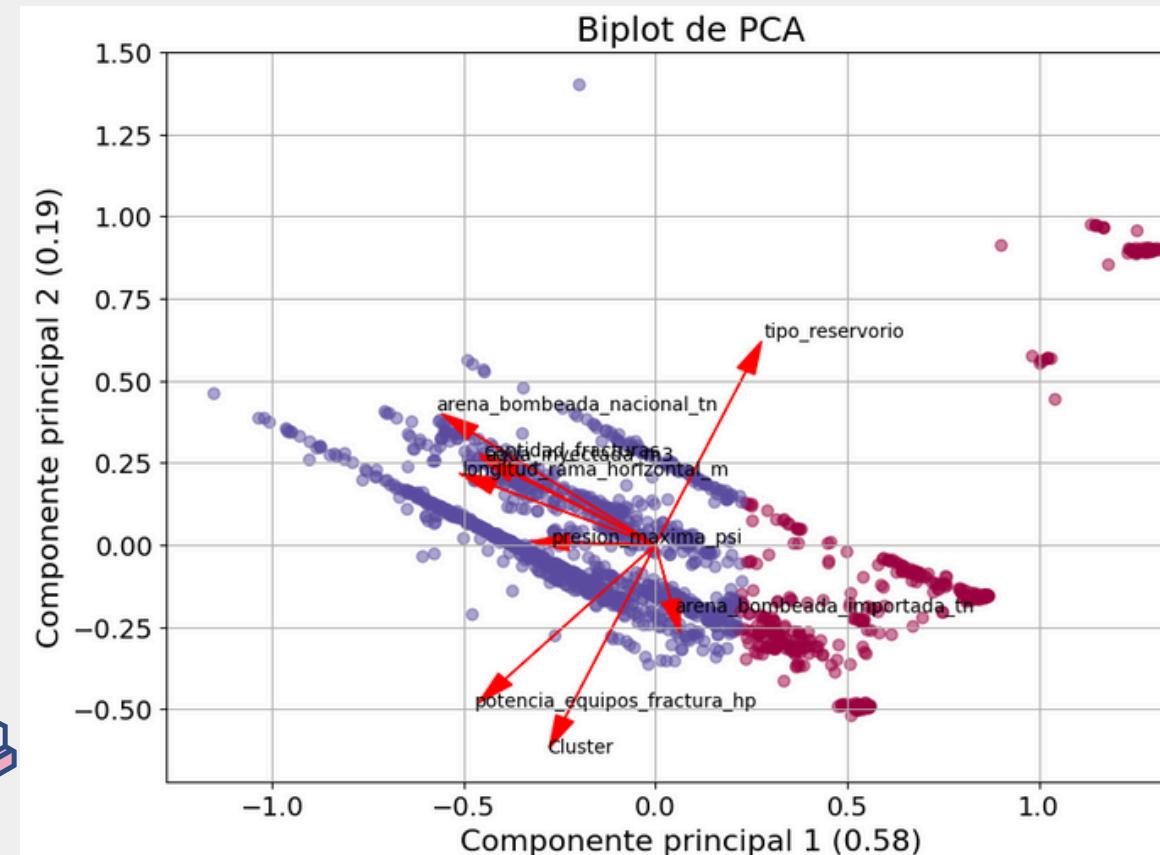
Número óptimo de clústeres: 2

Coeficiente de Silueta máximo: 0.45

Implementación de Análisis de Componentes principales PCA



Interpretación de los resultados

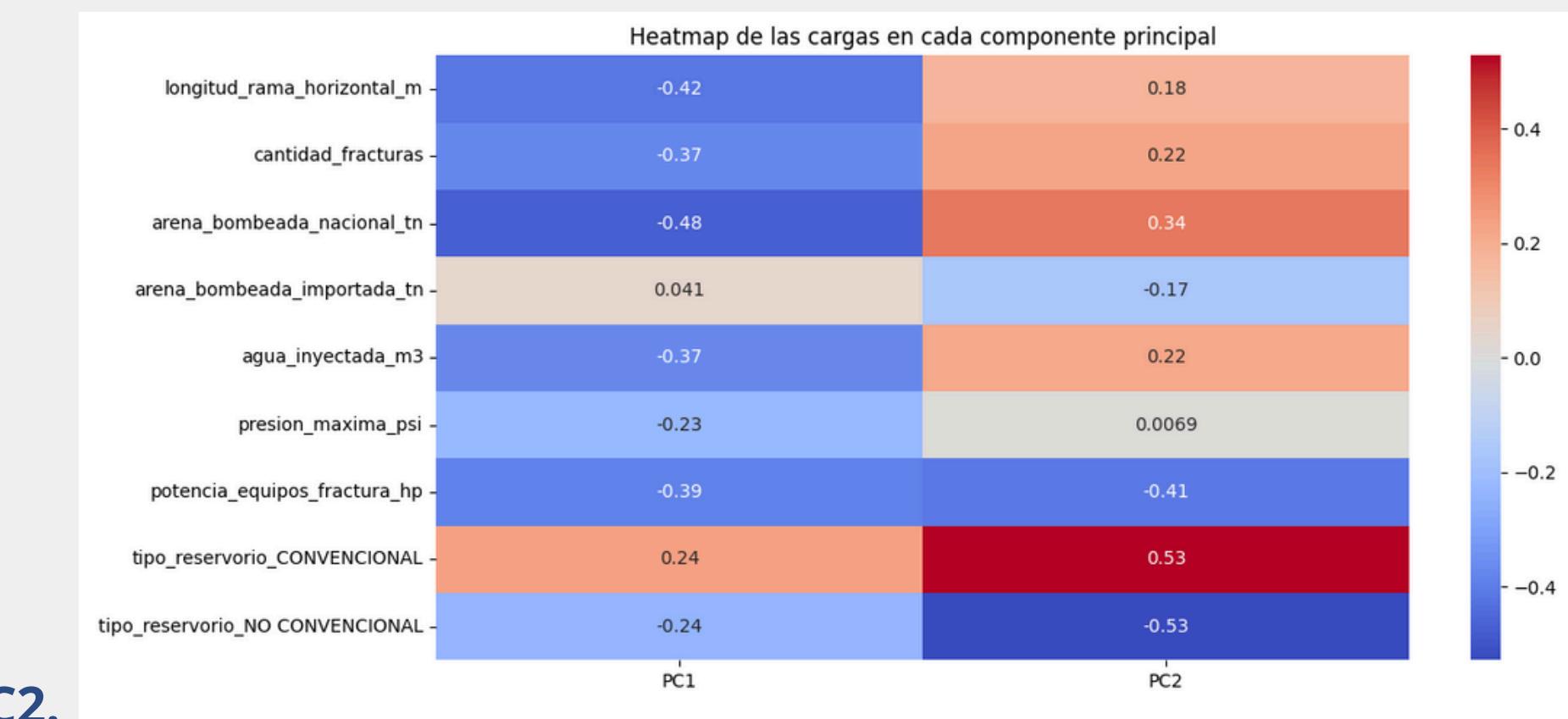


Clúster 1: Agrupa pozos que tienden a tener una mayor intensidad en las operaciones de fracturación hidráulica, caracterizados por valores negativos en **PC1**.

Clúster 2: Agrupa pozos asociados con reservorios convencionales y operaciones menos intensivas, con valores positivos en **PC2**.

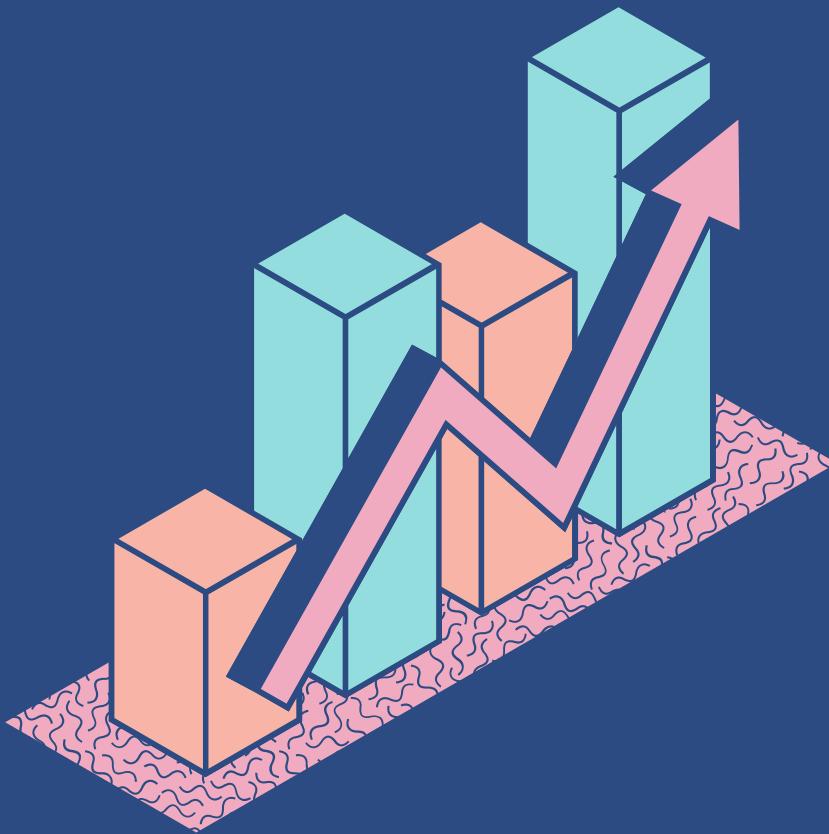
Componente 1: Está más relacionada con las características físicas y operativas

Componente 2: Realiza la distinción entre tipos de reservorios, con un enfoque en cómo estos tipos se asocian con la cantidad de arena bombeada y la potencia de los equipos de fractura



Impacto en la toma de decisiones

- Optimización del uso de recursos
- Planificación operativa
- Mejora en la toma de decisiones ambientales
- Costos operativos y eficiencia económica
- Toma de decisiones estratégicas de inversión



Beneficios Ambientales

La OMS considera que la cantidad adecuada de agua para consumo humano (beber, cocinar, higiene personal y limpieza del hogar) es de 50 litros (0.05 m^3) por habitante/día.



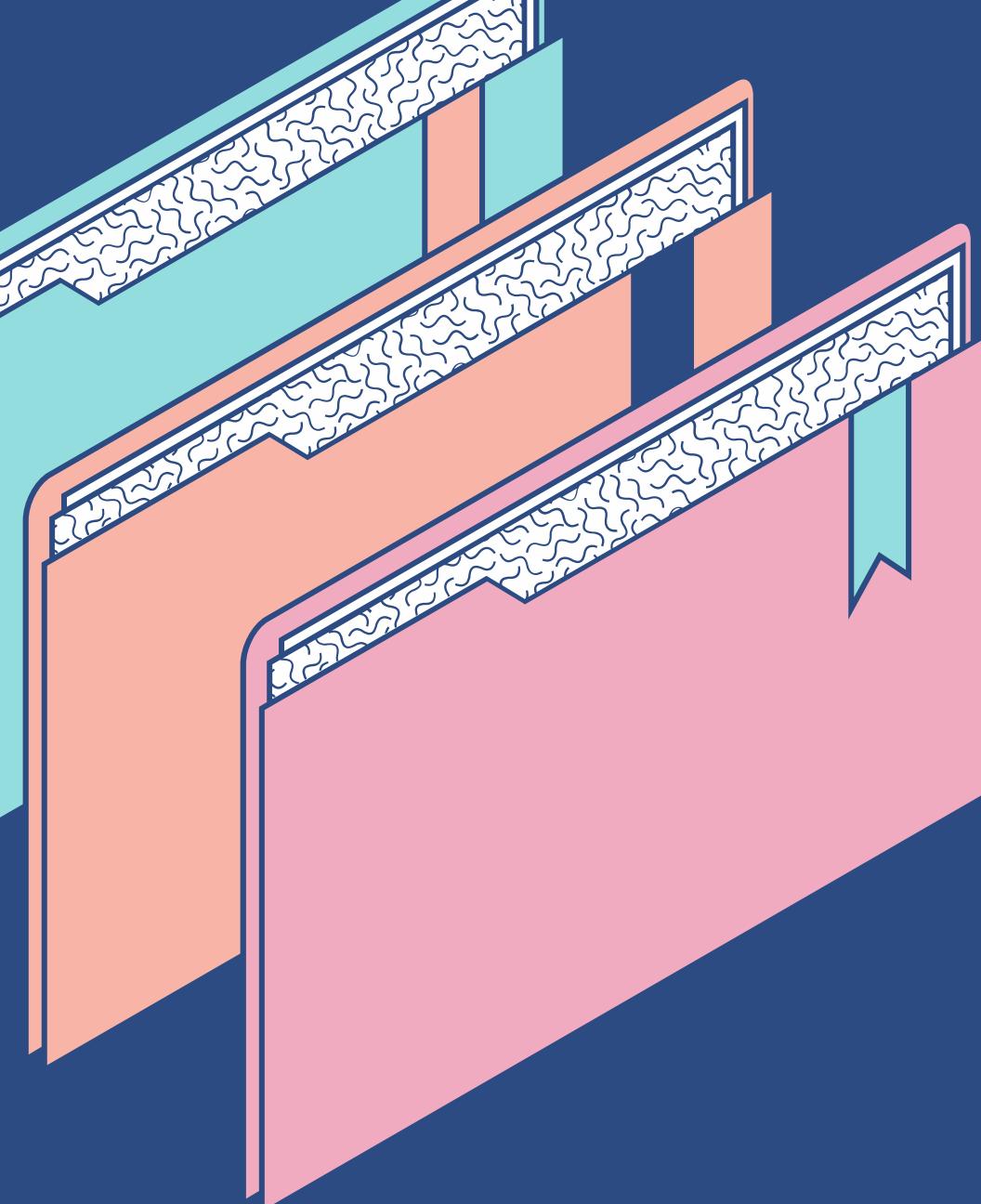
Ciudad mediana (100,000 habitantes): En promedio, consumiría 5000 m^3 de agua por día. En tan solo 10 días consumiría la misma cantidad de agua promedio que la fractura de un pozo.

Preservación de recursos hídricos

Menor generación de aguas residuales

Mitigación del impacto en zonas con estrés hídrico

Conclusiones



Modelo Supervisado

El hecho de que tanto en los datos de entrenamiento como en los de prueba se obtenga un R^2 de 0.97 demuestra que los modelos no están sobreajustados (overfitting) ni infraajustados (underfitting). Los resultados consistentes indican que los modelos tienen una buena capacidad de generalización y son capaces de realizar predicciones precisas con nuevos datos.

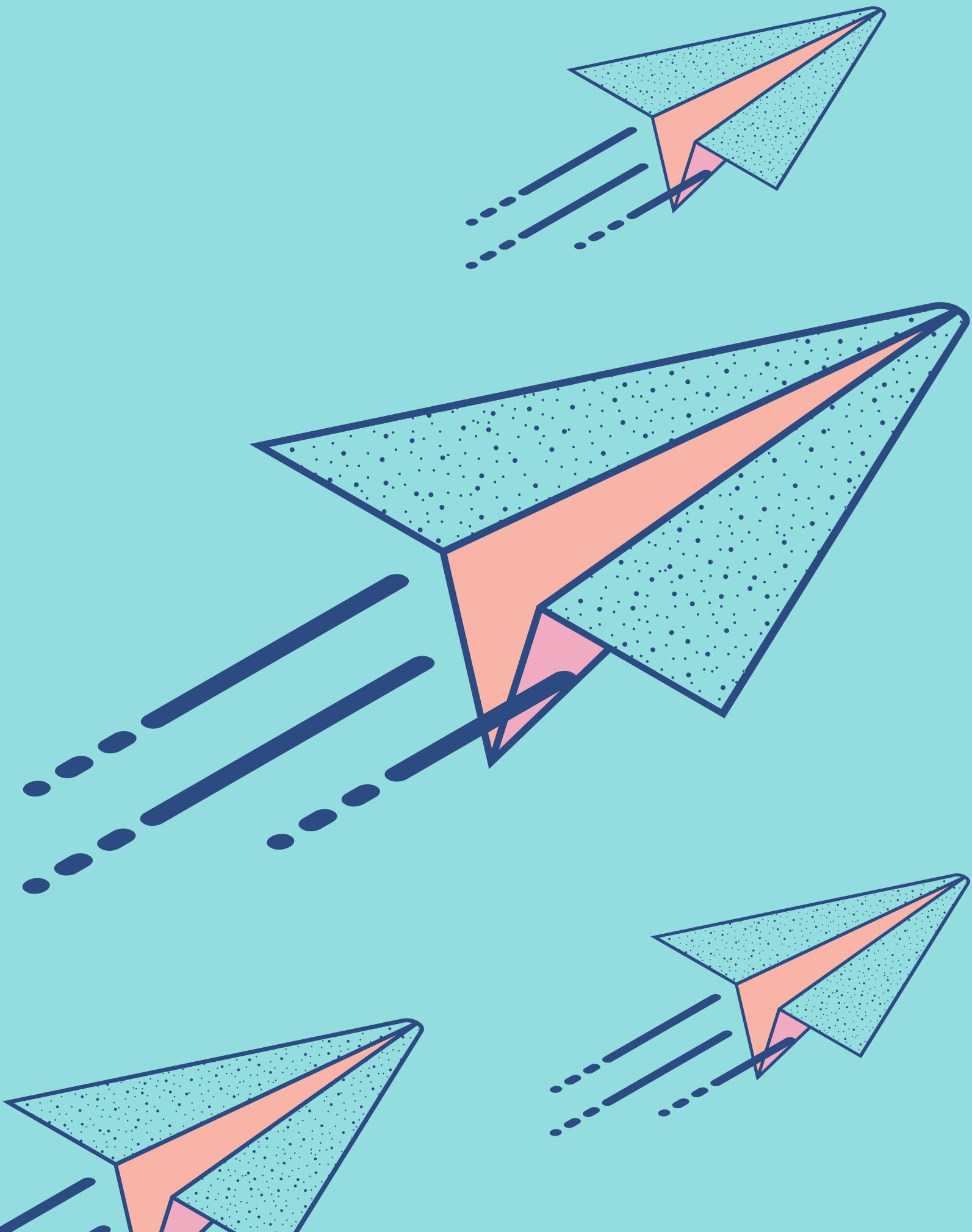
Modelo No Supervisado

Al identificar patrones de operación y rendimiento, se puede ajustar el uso de agua de manera más precisa, evitando un uso excesivo en pozos que no lo requieren y optimizando el recurso en aquellos donde es clave para el éxito operativo.

Cluster 1. Agrupa pozos que tienden a tener una mayor intensidad en las operaciones de fracturación hidráulica, caracterizados por valores PC1 (areaan bombeada, longitud de la rama y potencia de equipo).

Perspectivas a futuro para la mejora de modelos

Una posible mejora en el entrenamiento de los modelos de aprendizaje sería la incorporación de más características operativas y geológicas. Por ejemplo, agregar información sobre el tipo de roca en las fracturas podría aportar mayor precisión a las predicciones, al capturar mejor las variaciones en el comportamiento de los pozos según las propiedades del terreno.





¡Gracias por su atención!

Estamos disponibles para responder cualquier duda o comentario que tengan sobre el análisis presentado



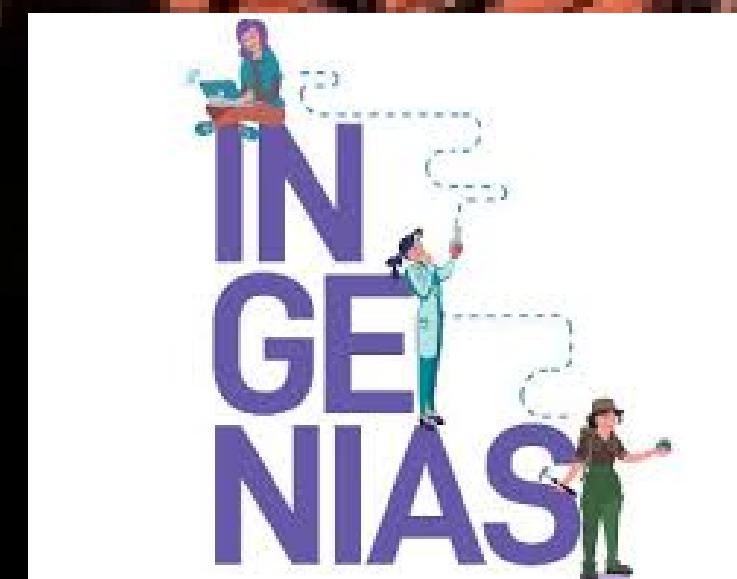
Contacto:

laugomezv13@gmail.com
sofivaucelles@gmail.com
herreracandela2196@gmail.com



Agradecimientos

FUNDACIÓN
YPF



jump.