




# PAYSIM

Detección de Fraude



# EQUIPO

## S15-35-T

Nefer Barreto Vergel - Data Analyst

Amalia Raimundo - Data Analystt

Mariana Soto - Data Analyst

Cristian Lepra - Data Science

Carolina Romero- Data Analyst

Yefferson Espinoza - Team Leader

# OBJETIVO

- Identificación de Patrones de Fraude: Detectar patrones y anomalías en los datos que puedan indicar actividades fraudulentas.
- Modelado Predictivo: Desarrollar y entrenar modelos de machine learning para predecir la probabilidad de fraude en nuevas transacciones.
- Mejora Continua: Iterar y mejorar continuamente los modelos y técnicas de detección de fraudes mediante la evaluación y ajuste de sus rendimientos.

# ETAPAS DEL PROYECTO

- Etl
- Eda
- Modelado
- Creación de api (Strealit)



# ETL

- El dataset cuenta con 6.362.620 filas y 11 columnas.
- El dataset se encontraba muy limpio no presentaba valores faltantes ni valores nulos.
- Solo se borro dos columnas que no eran relevantes.
- Del total de transacciones 0,13% son fraudulentas.



## Detección de Fraudes con Análisis de Datos

### Transacciones Registradas

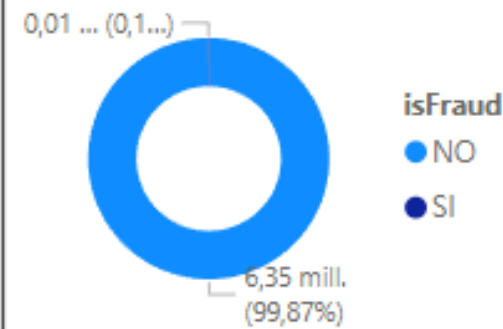
6,36262 mill.

### Es Fraude?

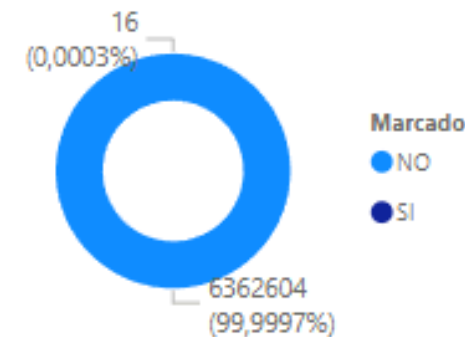
NO

SI

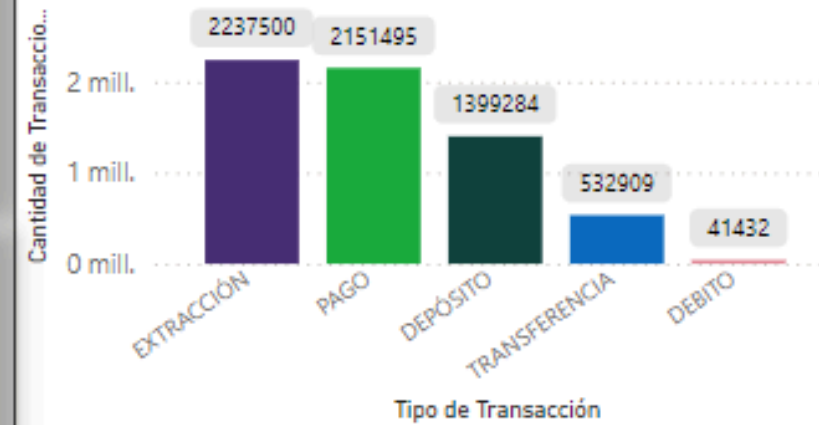
### Recuento de Fraudes



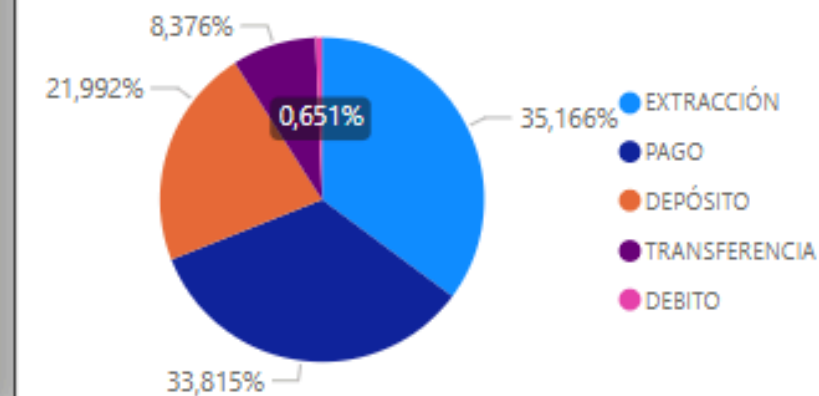
### Recuento de marcados como Fraude



### Tipos y Cantidades de Transacciones Registradas



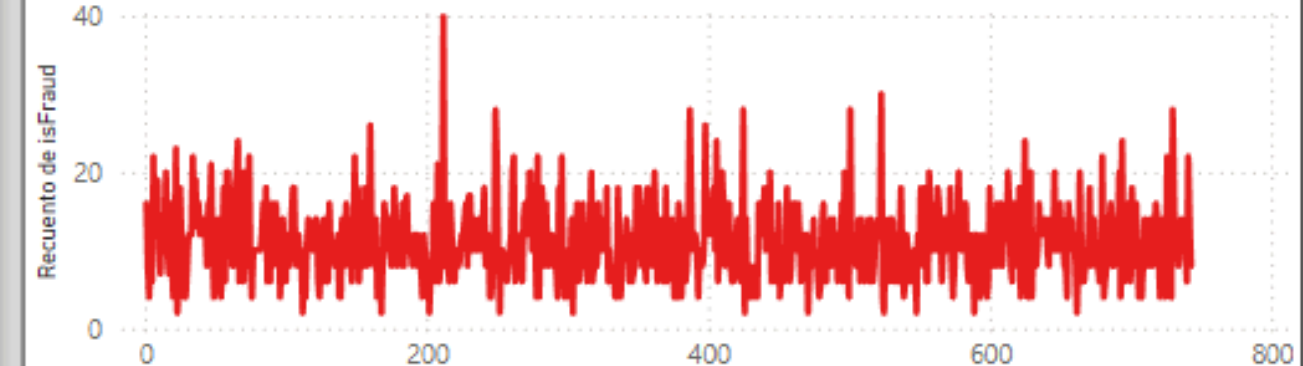
### Porcentaje de Transacciones Registradas



### Transacciones Válidas durante el tiempo simulado



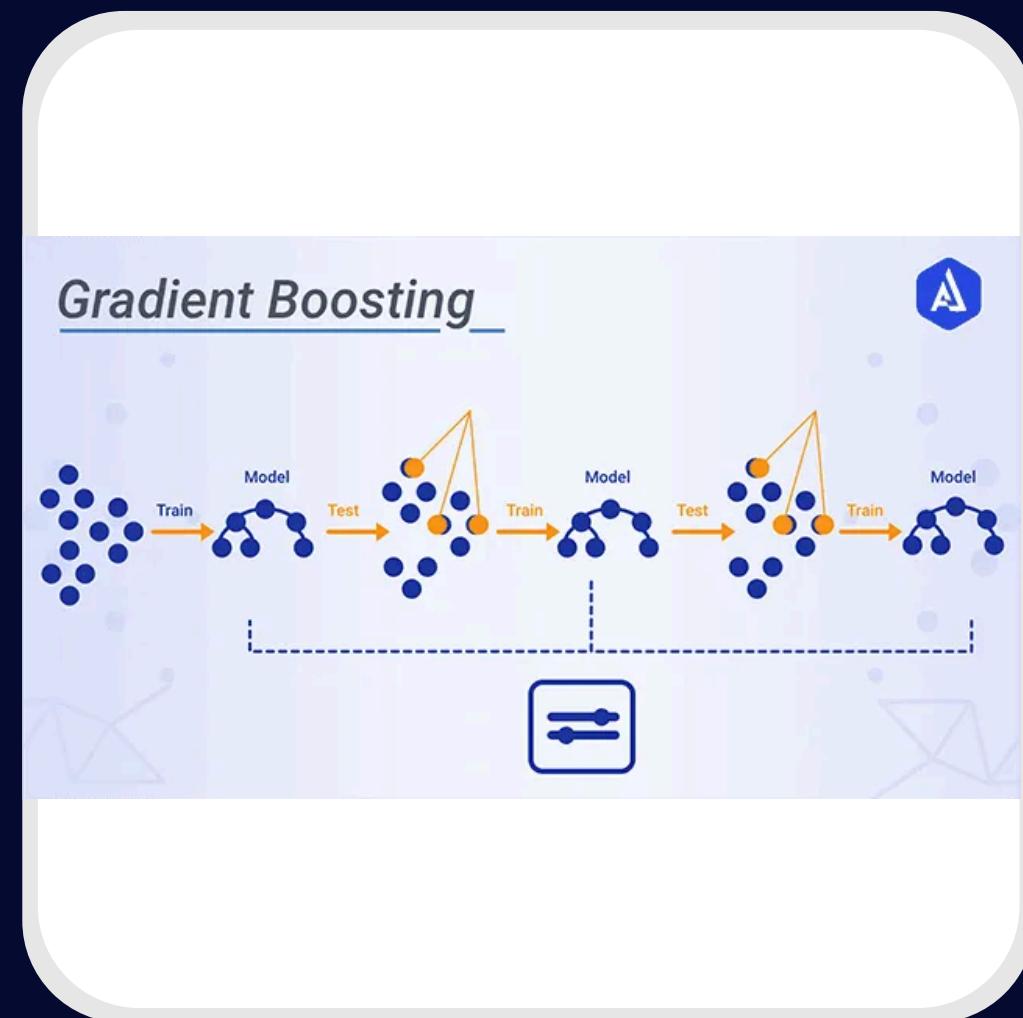
### Fraudes realizados durante el tiempo simulado



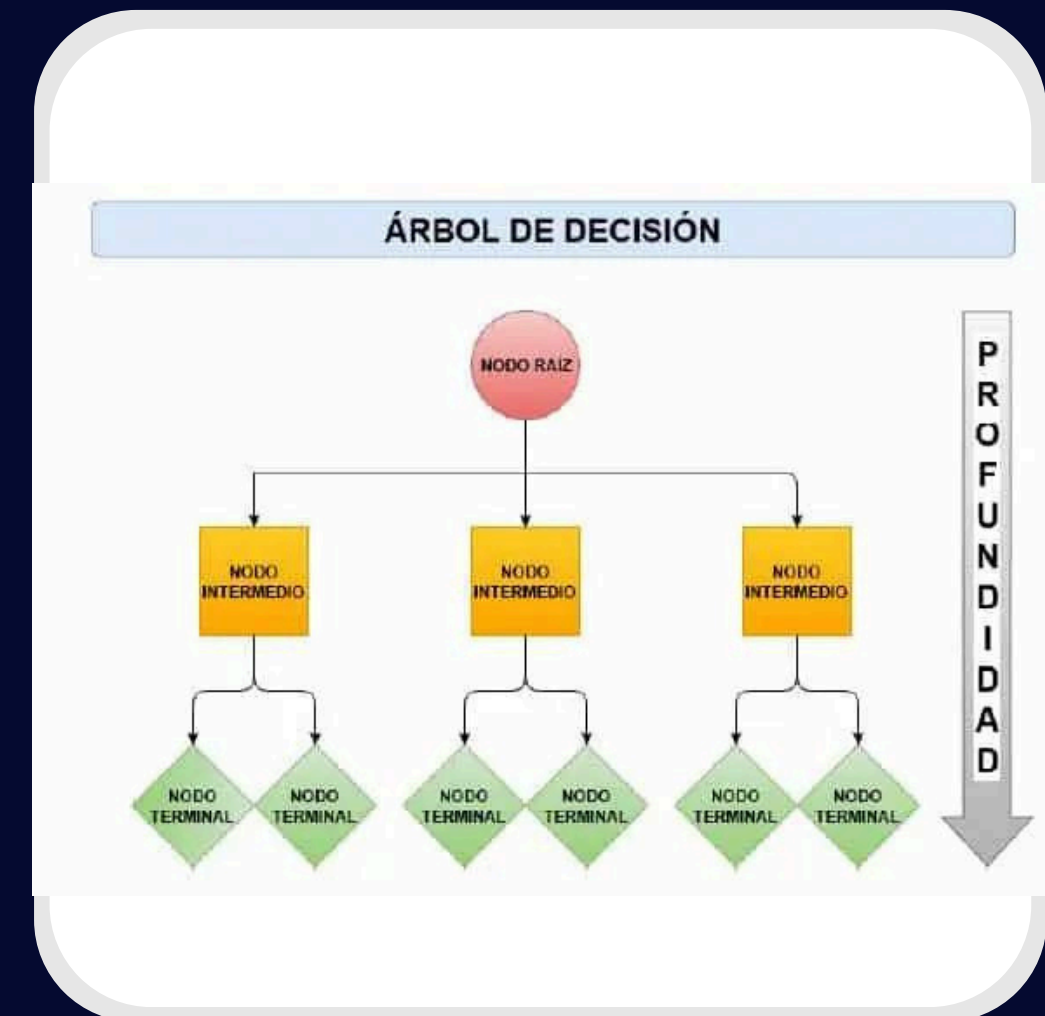
## Comparación de los modelos utilizados para predecir Fraude

	Model	Accuracy	Precision	Recall	F1 Score
0	Decision Tree	0.999716	0.999713	0.999716	0.999715
1	Gradient Boosting	0.999241	0.999220	0.999241	0.999089

# MODELADO



Potenciador de gradiente



Árbol de decisión

# Metricas

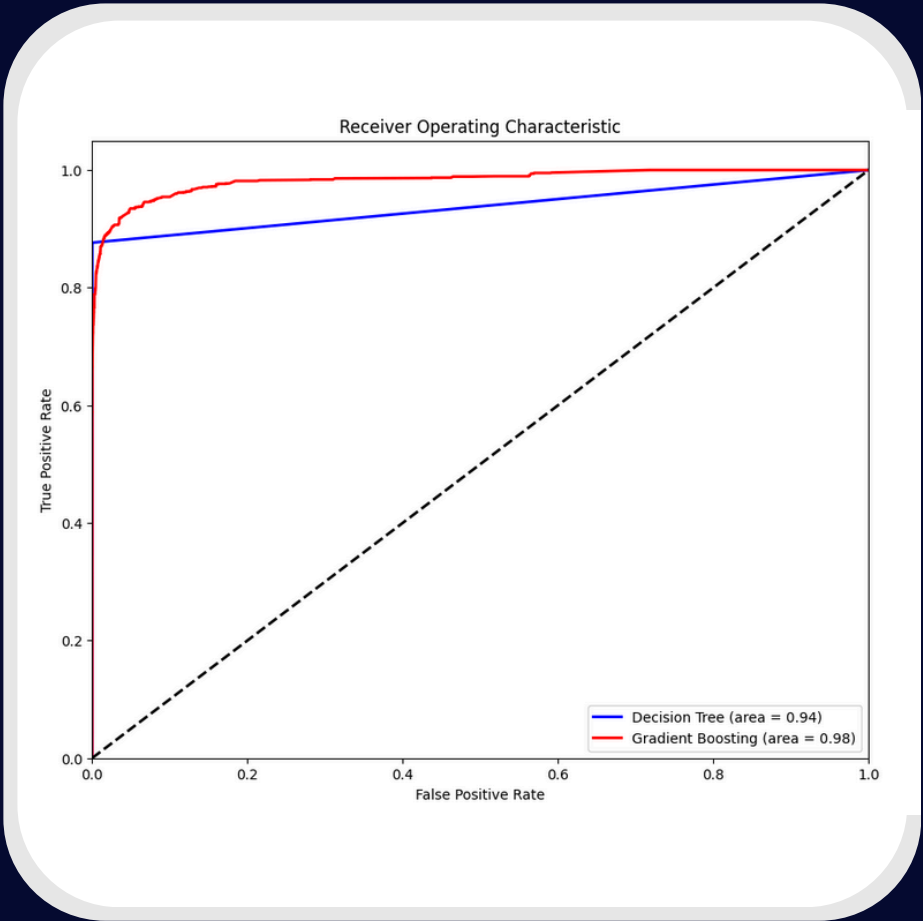
Curva Roc

Desision Tree Classifier

	Precisión	Recall	f1-score	Support
0	1.00	1.00	1.00	1270904
1	0.90	0.88	0.89	1620
accuracy			1.00	1272524
macro avg	0.95	0.94	0.94	1272524
weighted avg	1.00	1.00	1.00	1272524

Gradient Boosting Classifier

	Precisión	Recall	f1-score	Support
0	1.00	1.00	1.00	1270904
1	0.90	0.88	0.89	1620
accuracy			1.00	1272524
macro avg	1.00	0.95	0.95	1272524
weighted avg	1.00	1.00	1.00	1272524





# IMPLEMENTACIÓN

Usamos Streamlit para poder desarrollar una api que nos permita implementar nuestro modelo.

## Detección de Fraude en Transacciones

Esta aplicación predice si una transacción es fraudulenta o no usando un modelo de Gradient Boosting.

Introduce el valor para step:

1.00 - +

Selecciona el tipo de transacción (type):

CASH\_IN v

Introduce el valor para amount:

100225.00 - +

Introduce el valor para oldbalanceOrg:

100.00 - +

Introduce el valor para newbalanceOrig:

10000.00 - +

Introduce el valor para oldbalanceDest:

15485.00 - +

Introduce el valor para newbalanceDest:

100.00 - +

Predcir