

## Integrantes

- Victoria Arias
- Alejandro Gretter
- Juan Molina

### 1. Descripción de los datos y variables de entrada

Los datos fueron obtenidos desde Yahoo Finance utilizando la librería `yfinance`, bajo el ticker BTC-USD. Se trabajó con datos diarios de precio de Bitcoin, tomando como referencia la variable de salida: precio de cierre ajustado.

#### Variables de entrada seleccionadas:

- Precio de apertura, máximo, mínimo y volumen.
- Retornos diarios.
- Ventanas móviles (media y desviación estándar de 7, 14 y 30 días)
- Desplazamientos temporales (lags de 1 a 7 días).

### 2. Preprocesamiento

- Eliminación de valores nulos.
- Normalización de las variables numéricas mediante escalado MinMax.
- División del dataset en entrenamiento (70%), validación (15%) y prueba (15%), preservando la estructura temporal.
- Construcción de un formato supervisado multisalida para predecir los próximos 7 días.

### 3. Modelos evaluados y búsqueda de hiperparámetros

Se evaluaron tres modelos principales, con búsqueda de hiperparámetros mediante `GridSearchCV` y `RandomizedSearchCV`:

- Regresión Ridge (lineal regularizada): parámetro `alpha` ajustado.
- Random Forest Regressor: parámetros ajustados `n_estimators`, `max_depth`, `min_samples_split`.
- LightGBM Regressor (Multioutput): parámetros ajustados `num_leaves`, `learning_rate`, `n_estimators`.

#### 4. Estrategia de predicción de 7 días

Se implementó un modelo multisalida que, dado el día  $T$ , genera directamente la predicción de los 7 días siguientes ( $T+1 \dots T+7$ ). Esto evita el error acumulado de estrategias autorregresivas y mantiene consistencia temporal. El script `predict_7days.py` permite generar estas predicciones de forma automática para cualquier día de entrada.

#### 5. Evaluación y error esperable

Se evaluaron los modelos con MAE y RMSE:

- Ridge: mayor error, usado como baseline.
- Random Forest: error intermedio, mejora significativa sobre Ridge.
- LightGBM: menor error, modelo elegido para producción.
- El error esperable se encuentra en torno al 3–5% del precio diario en horizontes de 1–3 días, y aumenta hacia el día 7.

#### 6. Conclusiones

Se construyó un sistema capaz de predecir el precio de cierre de Bitcoin para los próximos 7 días, cumpliendo las restricciones de la consigna. El modelo más efectivo fue LightGBM multisalida. El trabajo evidenció la importancia del preprocesamiento de series temporales, la inclusión de variables derivadas y la comparación sistemática de modelos con ajuste de hiperparámetros. La precisión no es el foco central del trabajo, sino el diseño metodológico y la correcta implementación del flujo de machine learning.