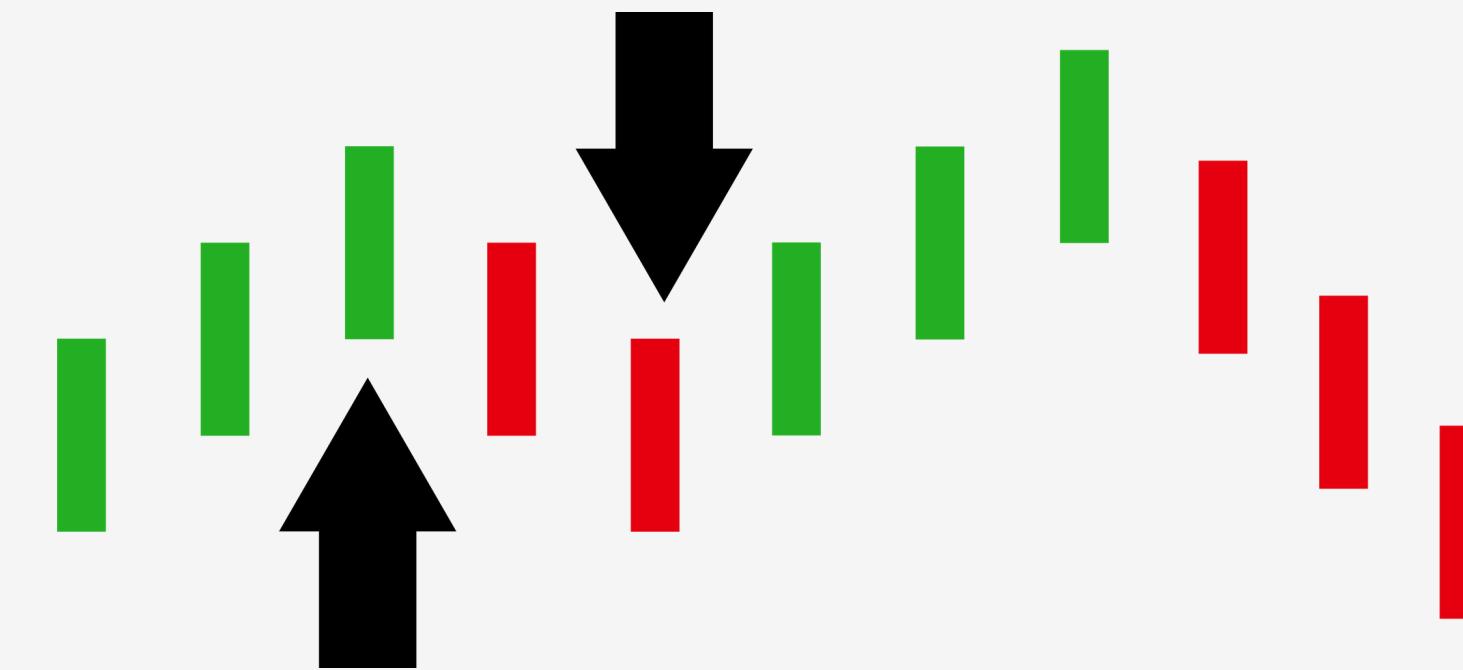


TRADING AND BOOSTING



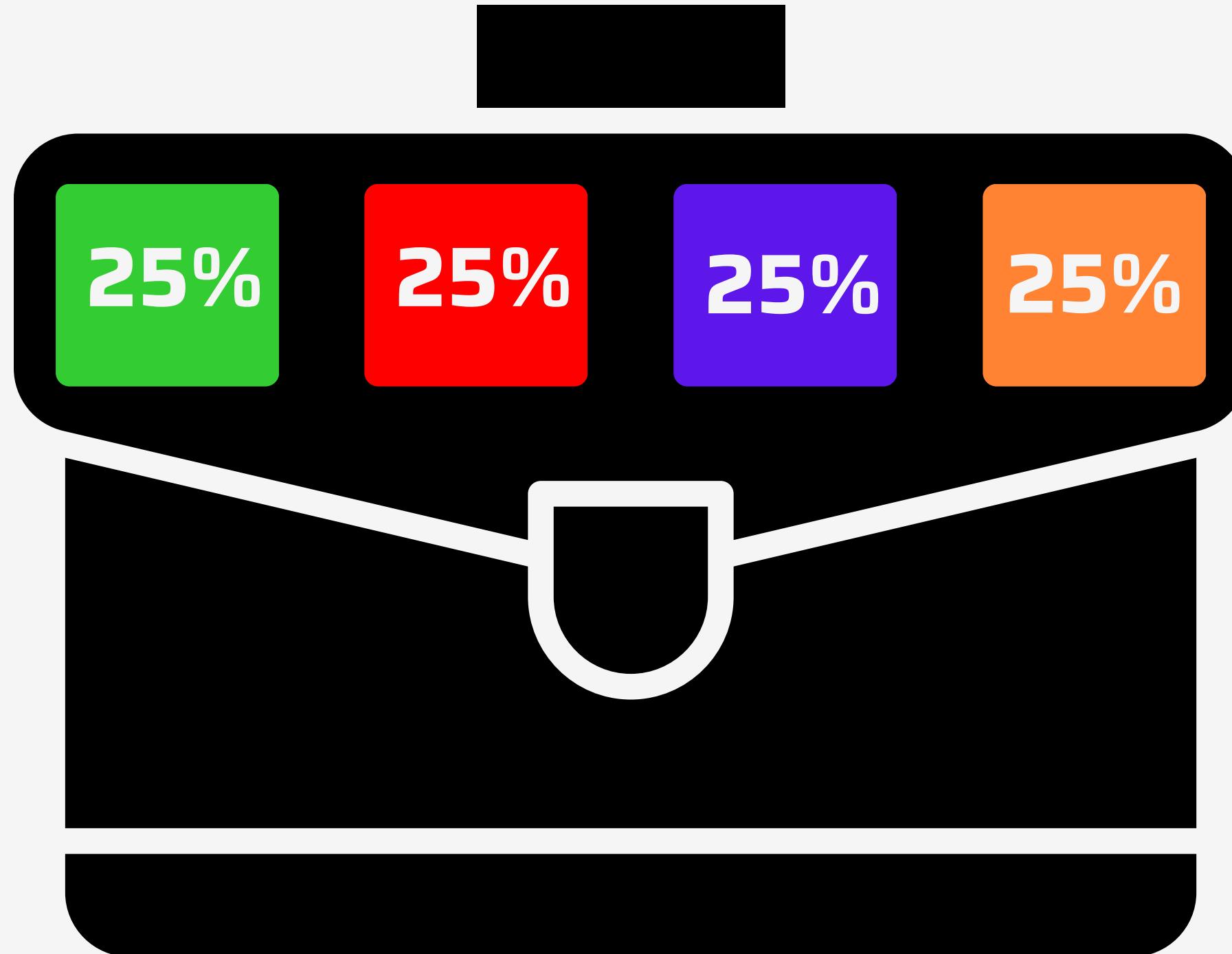
MELANIE ALVAREZ

MARIE CUCALÓN

MA. EMILIA RIVADENEIRA

ROBERTH LARA

Portafolio



SPLV

Invesco S&P 500 Low Volatility ETF

TLT

iShares 20+ Year Treasury Bond ETF

JNJ

Johnson & Johnson

BTC

Grayscale Bitcoin Mini Trust ETF

TLT

- T-Bonds, o Bonos del Tesoro, son títulos de deuda de renta fija.
- Tienen un plazo de 20 o 30 años y pagan una tasa de interés cada seis meses hasta su vencimiento.
- Básicamente, comprar un Bono del Tesoro es prestar dinero al gobierno



El Gobierno necesita dinero

"Necesito **financiar** un proyecto, pediré un **préstamo de \$100M**"



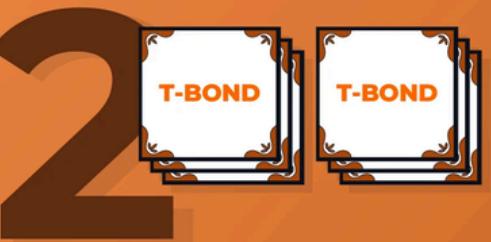
Subasta de Bonos

El Tesoro subasta 1,000 bonos a \$1,000 c/u



Pago de Intereses

Los inversionistas reciben un cupón de \$20 cada 6 meses, \$40 al año, por 20 años



El Tesoro de EE.UU. emite bonos

"Te ayudaré a conseguir \$100M emitiendo bonos"



El Bono



El Bono Madura

Pasan **20 años**, el período del préstamo acaba. **El valor del bono** (\$1,000) es **regresado** al inversionista

Feature Engineering



Técnicas

- **return_lag_{n}**
- **sma_5**: media móvil simple del Close.
- **ema_5**: media móvil exponencial, da más peso a los valores recientes.
- **rsi_5**: índice de fuerza relativa, mide el momentum del precio.
- **macd**: diferencia entre dos medias móviles, detecta tendencias.
- **price_diff**



Temporales

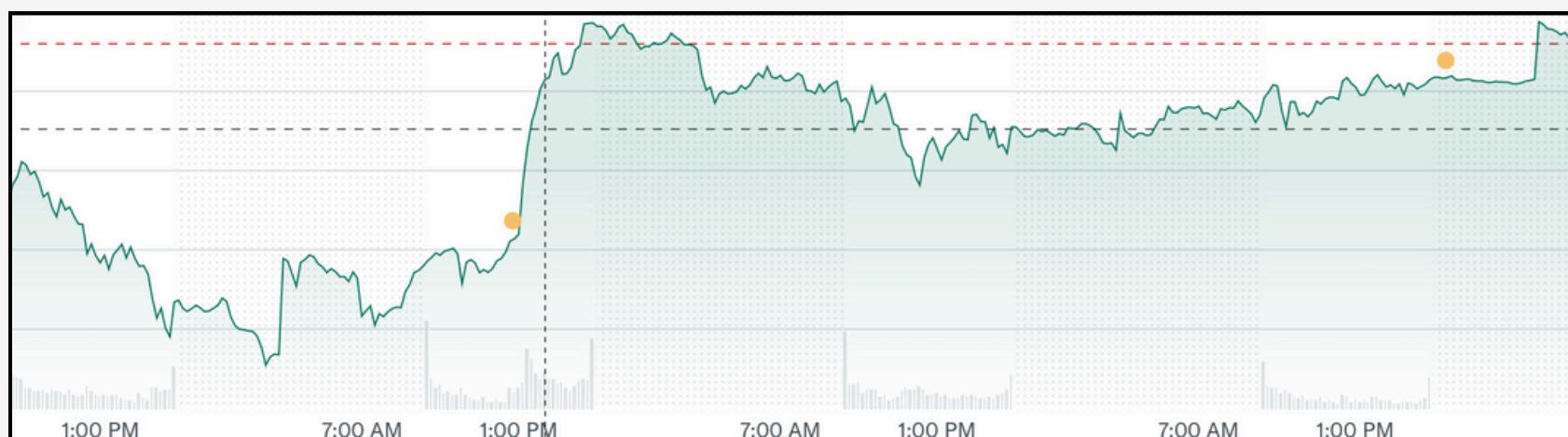
- **day_of_week**
- **is_month_end**
- **month**



Macroeconómicas

- **us10y_yield**: tasa de interés a 10 años.
- **fed_rate**: tasa de interés fijada por la Reserva Federal.
- **CPI**: índice de precios al consumidor, mide la inflación.
- **unemployment_rate**: porcentaje de personas desempleadas, indicador de salud del mercado laboral.

Series temporales financieras

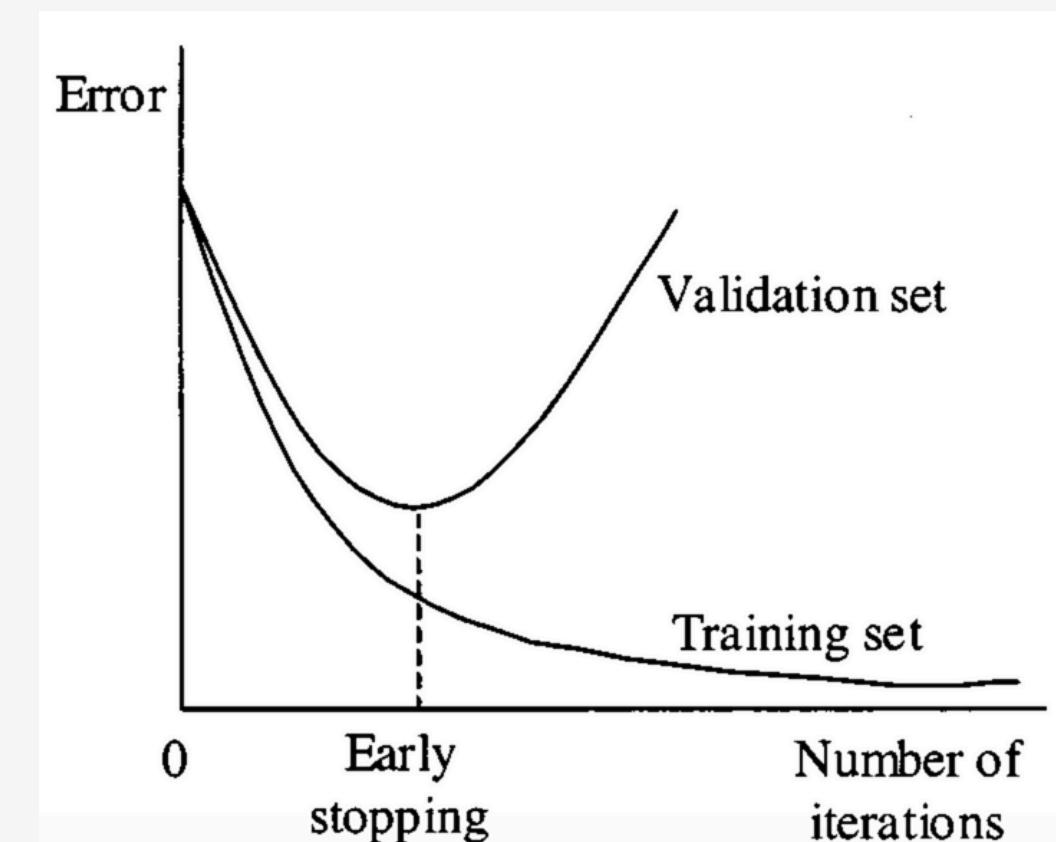


Respetar el orden cronológico de los datos para evitar 'data leakage'

Time Series Split

Split 1:	Training set	Test set
Split 2:	Training set	Test set
Split 3:	Training set	Test set
Split 4:	Training set	

Early Stopping



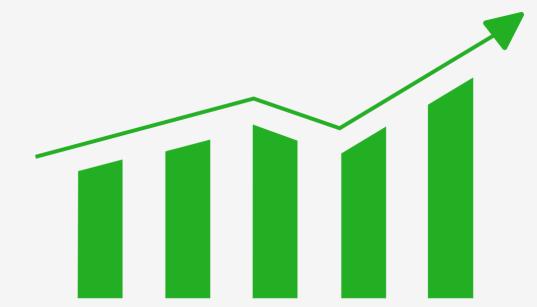
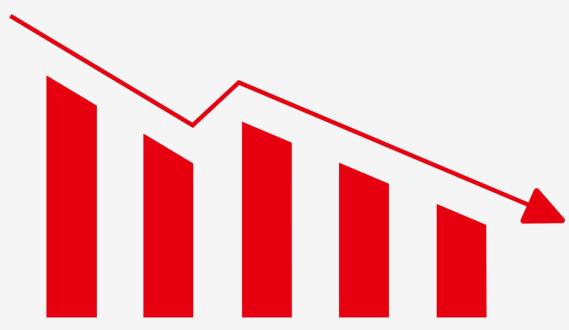
	Modelo	Threshold	Precisión Macro	Recall Macro	F1-score Macro
	XGBoost	0.60	0.8269	0.6538	0.6306
	GradientBoosting	0.30	0.8269	0.6538	0.6306
	LightGBM	0.60	0.7931	0.5385	0.4410
	RandomForest	0.70	0.7931	0.5385	0.4410
	DecisionTree	0.55	0.7200	0.6244	0.6032
	AdaBoost	0.45	0.5556	0.5204	0.4000
	CatBoost	0.50	0.3891	0.3891	0.3891

SELL**BUY****SELL****BUY**

Backtesting

Modelo	Threshold	Trades realizados	Ganancia Total (\$)	Capital Final (\$)	Precision (1)
XGBoost	0.6	4	9.70	1009.70	1.000
RandomForest	0.6	5	6.14	1006.14	0.600
XGBoost	0.7	1	2.17	1002.17	1.000
GradientBoosting	0.7	1	2.17	1002.17	1.000
GradientBoosting	0.6	1	2.17	1002.17	1.000
GradientBoosting	0.5	1	2.17	1002.17	1.000
RandomForest	0.7	1	2.17	1002.17	1.000
DecisionTree	0.6	4	1.00	1001.00	0.750
DecisionTree	0.7	4	1.00	1001.00	0.750
RandomForest	0.5	10	0.69	1000.69	0.600
LightGBM	0.5	13	0.38	1000.38	0.615
AdaBoost	0.6	0	0.00	1000.00	0.000
AdaBoost	0.7	0	0.00	1000.00	0.000
CatBoost	0.7	0	0.00	1000.00	0.000
CatBoost	0.6	0	0.00	1000.00	0.000
LightGBM	0.7	0	0.00	1000.00	0.000
LightGBM	0.6	0	0.00	1000.00	0.000
XGBoost	0.5	13	-5.51	994.49	0.385
AdaBoost	0.5	12	-8.13	991.87	0.333
DecisionTree	0.5	23	-13.47	986.53	0.435
CatBoost	0.5	13	-13.59	986.41	0.308

**Mayor precisión no significa
mayor ganancia**





¿Qué es el ETF SPLV?

- ETF que replica las 100 acciones con menor volatilidad del S&P 500.
- Se seleccionan aquellas con menor volatilidad observada en los últimos 12 meses.
- Su cartera está 100% compuesta por acciones de baja volatilidad, principalmente de sectores defensivos como servicios públicos, consumo básico y salud.
- Este ETF busca ofrecer exposición al mercado accionario con menor riesgo relativo, siendo ideal para inversionistas conservadores o en períodos de incertidumbre, ya que tiende a caer menos.

Ingeniería de Características y Modelado

Lo que se utilizó:

- Técnicos: RSI, MACD, Bollinger Bands, medias móviles.
- Estadísticos: retornos diarios y rezagos (return_lag).
- Velas japonesas: gap_up, lower_shadow, etc.
- Variables de calendario (día de semana, fin de mes, etc.)

Modelos

- Árbol de Decisión (baseline)
- RandomForest, AdaBoost, GradientBoosting
- LightGBM y XGBoost (boosting avanzado)



En este proyecto, nos interesa especialmente evitar falsas compras (cuando el modelo dice que subirá y no sube).

Usar precisión macro nos permite evaluar si el modelo realmente acierta cuando dice “compra”, incluso si eso ocurre pocas veces.

Modelo	Accuracy	Precisión Macro	F1 Score	AUC	Comentario
XGBoost	0.68	0.75	0.59	0.64	Buen balance entre precisión y riesgo
RandomForest	0.68	0.67	0.62	0.63	Consistente, pero con más operaciones
GradientBoosting	0.68	0.83	0.55	0.59	Conservador, pocas compras pero efectivas
Árbol de Decisión	0.64	0.63	0.63	0.63	Baseline confiable
LightGBM	0.54	0.54	0.53	0.62	Alta cantidad de falsos positivos
AdaBoost	0.61	0.30	0.37	0.50	No logró identificar subidas

Evaluación y Resultados de los Modelos

- XGBoost y RandomForest: balance entre precisión y rentabilidad.
- AdaBoost y LightGBM: pobre rendimiento predictivo o financiero.
- GradientBoosting es útil en escenarios conservadores: genera pocas señales, pero todas fueron exitosas.

Threshold ajustado (0.55):

- Mejora precisión sin comprometer recall de forma drástica.
- Reduce operaciones erróneas.

SPLV

RESULTADOS

Modelo	Profit (\$)	Final Capital	Trades	Wins	Losses
✓ RandomForest	+15.75	1015.75	6	4	2
✓ XGBoost	+14.78	1014.78	4	3	1
GradientBoosting	+3.27	1003.27	2	2	0
✗ LightGBM	-10.12	989.88	14	6	8
✗ DecisionTree	-12.45	987.55	11	6	4
✗ AdaBoost	0.00	1000.00	0	0	0

Capital inicial: \$1000 – se invirtió el 25% del capital en cada señal de compra (predicción positiva del modelo).

SPLV

GANANCIA: +14.78

XGBoost es el modelo más equilibrado:

- Alta precisión en predicciones positivas (subidas reales).
- Solo 1 operación perdedora.
- Bajo volumen de operaciones → menos comisiones y menos riesgo operativo.
- Aunque RandomForest ganó un poco más, generó más operaciones y más errores → mayor exposición al riesgo.
- GradientBoosting fue muy conservador (solo 2 trades), útil en mercados volátiles pero con poca rentabilidad.
-

XGBoost con umbral 0.55 es la mejor opción si el objetivo es tomar decisiones de compra rentables pero con bajo riesgo.

Prioriza precisión sin entrar en muchas operaciones. Ideal para estrategias conservadoras en un ETF de baja volatilidad como SPLV.



¿Qué es Johnson & Johnson (JNJ)?

- Johnson & Johnson (JNJ) es una de las empresas más grandes y reconocidas del mundo en el sector de la salud.
- Opera globalmente en tres segmentos clave: productos farmacéuticos, dispositivos médicos y productos de consumo (como cuidado personal y salud del hogar).
- Forma parte del índice S&P 500, lo que la vuelve relevante y representativa del mercado estadounidense.
- JNJ es considerada una acción defensiva (menos volátil y más estable en comparación con empresas de sectores cíclicos).
- Esta acción fue escogida debido a que el grupo esta compuesto por inversores conservadores, y estamos en épocas de incertidumbre económica

Feature Engineering



Diferencias de precio y retorno históricos

- **price_diff** y **pct_diff**: diferencias absolutas y relativas entre precios de apertura y cierre.
- **return_daily**: retornos diarios.
- **return_lag_n**: rezagos.



Estadísticas Móviles e Indicadores Técnicos

- **rolling_std_return_5** y **sma_5**: promedios móviles y desviaciones estándar
- **RSI**: mide la fuerza del movimiento de precios.
- **MACD y su señal**: capturan cambios en la fuerza, dirección y duración de una tendencia.
- **Bollinger Bands**: limitan rangos de precios en función de la volatilidad reciente.



Banderas Binarias

- **volume_spike**: identifica días con volumen de negociación inusualmente alto.
- **price_above_SMA50**: señala si el precio se encuentra por encima de su media móvil de 50 días.
- **RSI_overbought**: indica si el RSI supera el umbral de 70.
- **MACD_above_signal**: marca cuando el MACD cruza por encima de su línea de señal

Modelo	Accuracy	Precision Macro	Recall Macro	F1 Macro
CatBoost	0.6000	0.6250	0.5804	0.5500
XGBoost	0.6000	0.6027	0.6027	0.6000
LightGBM	0.6000	0.5982	0.5982	0.5982
RandomForest	0.5333	0.5313	0.5313	0.5313
GradientBoosting	0.5000	0.5045	0.5045	0.4994
DecisionTree	0.4667	0.4750	0.4777	0.4570
AdaBoost	0.4333	0.4367	0.4375	0.4327

Evaluación y Resultados de los Modelos

- LightGBM: modelo más equilibrado (accuracy, precisión, recall y f1 macro de 0.60)
- XGBoost: modelo más confiable desde el punto de vista generalista (mejor F1 macro (0.60) y el recall más equilibrado entre clases (0.64 en bajadas, 0.56 en subidas)).
- CatBoost: modelo más eficaz para detectar subidas en el precio (alta precisión macro (0.625) y un recall clase 1 de 0.88)

JNJ

RESULTADOS - BACKTESTING

Modelo	Mejor Threshold	Ganancia Total	Capital Final
LightGBM	0.5	\$11.93	\$1011.93
XGBoost	0.5	\$11.88	\$1011.88
LightGBM	0.7	\$10.45	\$1010.45
LightGBM	0.4	\$8.60	\$1008.60
GradientBoosting	0.5	\$3.58	\$1003.58
CatBoost	0.5	-\$1.29	\$998.71
DecisionTree	0.4	-\$31.72	\$968.28

Capital inicial: \$1000 – se invirtió el 25% del capital en cada señal de compra (predicción positiva del modelo).

JNJ

MEJOR MODELO: LIGHTGBM

- Con threshold 0.5 es el modelo que ofrece el mejor balance entre rendimiento estadístico y retorno financiero.
- Se alcanzó una suma de \$1011.93, producto de 15 operaciones de compra, de las cuales 10 fueron exitosas.
- Este resultado es coherente con las métricas de clasificación previamente evaluadas, en donde tanto LightGBM como XGBoost obtuvieron los valores más altos de F1 macro (~0.60), mostrando equilibrio entre precisión y recall.
- Se evidencia que la capacidad de adaptación lo hacen adecuado para escenarios de corto plazo como el planteado.
- Cabe recalcar que XGBoost también es una alternativa muy buena, y de hecho, su rendimiento en la simulación fue prácticamente indistinguible del de LightGBM.



Bitcoin BTC-USD

- ANALIZAMOS DATOS DIARIOS DE BITCOIN DESDE 2021 HASTA 2025.
- OBJETIVO: PREDECIR SI EL PRECIO SUBIRÁ AL DÍA SIGUIENTE → FORMULADO COMO UN PROBLEMA DE CLASIFICACIÓN BINARIA.
- ¿POR QUÉ BITCOIN? ALTA VOLATILIDAD, COMPORTAMIENTO NO LINEAL, IDEAL PARA MODELOS DE BOOSTING.



Feature Engineering



- **SMA / EMA:** Promedios móviles (suave y rápida) → Detectan tendencias.
- **MACD:** Diferencia entre EMAs → Señales de compra o venta.
- **RSI:** Fuerza del precio → Sobrecompra (>70) o sobreventa (<30).
- **ROC:** Velocidad del cambio → Mide momentum.
- **Bollinger Bands:** Rango de precio → Detecta volatilidad.
- **OBV:** Volumen acumulado → Confirma fuerza de la tendencia.



Temporales

- day_of_week
- is_month_end
- month

Usamos técnicas de ingeniería de características para capturar momentum, tendencia, volatilidad y contexto temporal.

Luego aplicamos SelectKBest para quedarnos con las 20 más informativas.



Indicadores de precio

- Retornos
- Diferencias
- Lags.

Modelo	Accuracy	ROC AUC	Precision	Retorno Final	Sharpe Ratio	Win Rate	Nº Ops
Decision Tree	0.5484	0.4679	0.4615	11.20	0.4447	51.61%	31
Random Forest	0.6129	0.6880	0.5294	175.87	3.1308	65.38%	26
AdaBoost	0.5161	0.6303	0.4500	139.19	2.6224	54.17%	24
Gradient Boosting	0.5161	0.5598	0.4444	189.07	2.8800	50.00%	30
LightGBM	0.6452	0.7051	0.5625	352.81	4.9717	66.67%	30
XGBoost	0.6452	0.7436	0.5625	154.94	2.4379	66.67%	30

Evaluación y Resultados de los Modelos

PROBAMOS 6 MODELOS PARA PREDECIR SI EL PRECIO SUBIRÁ.

MEJORES RESULTADOS:

-  **LIGHTGBM:** MAYOR PRECISIÓN Y RETORNO.
-  **XGBOOST:** MUY CERCANO, ALTA PRECISIÓN Y BUEN RETORNO.

AMBOS GENERARON SEÑALES DE COMPRA MÁS ACERTADAS Y RENTABLES.



Muchas Gracias

LINK DE
**GITHUB: HTTPS://GITHUB.COM
/TITUSR099/PROJECT_TRADIN
G_DT.GIT**

