

Proyecto 2

002 Introduction to Trading



ITESO

**Universidad Jesuita
de Guadalajara**

DAVID ROGELIO CAMPOS MURIA

7 de octubre de 2025

Este proyecto desarrolla y evalúa una estrategia sistemática de trading sobre BTCUSDT con datos horarios, diseñada para operar sin apalancamiento, en largos y cortos, incorporando costos transaccionales del 0.125% y aplicando una confirmación de señales 2-de-3 (tendencia, momento y ruptura). El énfasis analítico se centra en maximizar el Calmar Ratio ($CAGR/|MaxDD|$), una métrica que privilegia la rentabilidad ajustada por caídas máximas, complementada con Sharpe, Sortino, MaxDD y Win Rate.

Metodológicamente, el trabajo implementa un backtesting realista y un flujo walk-forward: el conjunto de datos se divide cronológicamente en 60% Entrenamiento, 20% Prueba y 20% Validación. Los hiperparámetros de la estrategia (ventanas e umbrales de EMA y RSI, banda Donchian con amortiguación por ATR, y controles de riesgo como stop loss/take profit) se optimizan exclusivamente en Entrenamiento mediante búsqueda aleatoria con semilla controlada; posteriormente, esos parámetros se congelan y se evalúan sin re-ajuste en Prueba y Validación, lo que mitiga el sobreajuste y refleja mejor el desempeño fuera de muestra.

La estrategia se estructura en tres pilares técnicos: (i) tendencia (cruce EMA rápida/lenta), (ii) momento (RSI con bandas asimétricas para captar presiones direccionales) y (iii) ruptura (canal Donchian ajustado por la volatilidad vía ATR). Sólo cuando al menos dos de estos componentes coinciden se habilita una señal, y la ejecución se difiere una barra para evitar *look-ahead*. El motor de backtest controla inventario y costos en cada operación, y calcula la curva de capital (equity) y el retorno por barra con precisión contable.

Los resultados se presentan de manera reproducible y auditables: (1) una tabla consolidada por segmento con Calmar, Sharpe, Sortino, MaxDD, Win Rate, número de operaciones y % de rendimiento total; (2) gráficas de equity y drawdown para Entrenamiento, Prueba y Validación; y (3) tablas de retornos mensuales por segmento. El proyecto entrega además los parámetros óptimos en un artefacto persistente (archivo de configuración “congelado”) para repetir evaluaciones sin re-optimización, facilitando comparabilidad y control de cambios.

En términos gerenciales, la estrategia demuestra ser coherente con su objetivo de control de riesgo: la optimización por Calmar reduce la probabilidad de caídas profundas a costa de priorizar la calidad sobre la cantidad de señales. La evaluación fuera de muestra (Prueba/Validación) permite estimar la estabilidad del desempeño bajo regímenes de volatilidad cambiantes típicos de cryptoactivos. Las secciones siguientes detallan datos, metodología, resultados y riesgos, así como limitaciones y líneas de mejora (robustez paramétrica, modelado de *slippage*, y análisis por regímenes), conformando un entregable técnico listo para decisión y reporte ejecutivo.

Contents

1. Introducción	5
1.1 Objetivo	5
1.2 Alcance y restricciones	5
1.3 Justificación del activo y del marco temporal	5
1.4 Métrica objetivo: por qué Calmar	6
1.5 Principios de diseño de la estrategia	6
1.6 Criterios de éxito	7
2. Datos	7
2.1 Fuente y frecuencia	7
2.2 Rango temporal y cobertura	7
2.3 Estructura del archivo	7
2.4 Validaciones y limpieza	8
2.5 Preparación para indicadores y señales	8
2.6 Esquema de <i>split</i> 60/20/20 (walk-forward)	8
2.7 Implicaciones para las métricas	9
2.8 Riesgos de calidad de datos (y mitigación)	9
3. Metodología	9
3.1 Datos y partición temporal	9
3.2 Arquitectura de señales: confirmación 2-de-3	10
3.3 Gestión del riesgo y costos	10
3.4 Motor de backtesting (ciclo de órdenes y P&L)	11
3.5 Espacio de hiperparámetros y búsqueda	11
3.6 Hiperparámetros finales	12
3.7 Validación fuera de muestra (walk-forward simple)	12
3.8 Visualización y evidencia	13
3.9 Reproducibilidad	15
4. Resultados y análisis de performance	15
4.1 Métricas agregadas (60/20/20)	15
4.2 Curva de capital y perfil de riesgo	15
4.3 Eficiencia por trade y estilo de P&L	16
4.4 Estacionalidad operativa (retornos mensuales)	16
4.5 Qué está funcionando	17

4.6 Limitaciones observadas	17
4.7 Conclusión operativa de resultados	17
5. Análisis de riesgo y limitaciones	17
5.1 Riesgo de mercado (P&L y drawdowns).....	18
5.2 Riesgo de modelo (sobre-ajuste y deriva).....	18
5.3 Riesgo de ejecución y fricción	19
5.4 Riesgo de datos y medición	19
5.5 Sensibilidad y capacidad	20
5.6 Controles y gobernanza de riesgo.....	20
5.7 Qué rompería la estrategia (señales de alerta)	20
5.8 Limitaciones del estudio	21
6. Conclusiones	21

1. Introducción

1.1 Objetivo

Desarrollar y evaluar una estrategia sistemática de trading para BTCUSDT en marco temporal horario, que opere sin apalancamiento y en ambos lados del mercado (largos y cortos), incorporando costos transaccionales del 0.125% y utilizando confirmación 2-de-3 entre señales de tendencia, momento y ruptura. El desempeño se optimiza con respecto al Calmar Ratio —métrica que equilibra crecimiento compuesto y severidad de pérdidas máximas—, y se valida mediante un flujo walk-forward con división 60%/20%/20% (TRAIN/TEST/VALID)

1.2 Alcance y restricciones

- **Mercado y símbolo:** BTCUSDT (mercado spot).
- **Frecuencia:** velas de 1 hora.
- **Costos:** comisión fija del 0.125% por transacción; sin apalancamiento.
- **Posiciones:** largas y cortas, con control de inventario (no se venden activos no poseídos y no se excede el capital disponible).
- **Ejecución:** las señales se materializan una barra después de su generación para evitar sesgos de anticipación (*look-ahead*).
- **Optimización:** los hiperparámetros de indicadores y gestión (ventanas/umbrales, buffers de volatilidad, *stops* y *targets*) se ajustan solo en Entrenamiento; no se re-estiman en Prueba/Validación.
- **Validación:** evaluación out-of-sample en Prueba y Validación para estimar capacidad de generalización.

1.3 Justificación del activo y del marco temporal

- BTC presenta alta volatilidad y liquidez, condiciones propicias para estrategias técnicas que explotan tendencias, rupturas y cambios de momentum.
- La frecuencia horaria equilibra:
 - Señal-ruido más favorable que intradía de muy alta frecuencia,
 - Suficiente densidad de datos para aprendizaje de parámetros y evaluación robusta,
 - Costos de transacción manejables frente a marcos de muy corto plazo.

1.4 Métrica objetivo: por qué Calmar

- Calmar Ratio = $CAGR / |\text{Máximo Drawdown}|$.

Maximiza el crecimiento compuesto penalizando de forma explícita la profundidad de caídas, alineando el objetivo con la tolerancia al riesgo que típicamente exigen los comités de inversión (protección de capital y suavidad de la curva).

- **Se complementa con:**
 - Sharpe (riesgo total),
 - Sortino (riesgo a la baja),
 - Máximo *Drawdown* (severidad de pérdidas),
 - Win Rate (frecuencia de aciertos),
 - % de rendimiento total (para lectura ejecutiva).

1.5 Principios de diseño de la estrategia

- **Confirmación 2-de-3:**
 1. Tendencia (relación entre medias móviles exponenciales rápida y lenta),
 2. Momento (RSI con umbrales para presión alcista/bajista),
 3. Ruptura (canales tipo Donchian modulados por volatilidad vía ATR).
La entrada solo se habilita cuando al menos dos componentes coinciden, buscando reducir señales espurias.
- **Robustez operativa:**
 - Desplazamiento de señales +1 barra para evitar *look-ahead*,
 - Costos aplicados en cada operación,
 - Control de inventario y capital en todo momento,
 - Cálculo contable de equity y retornos por barra.
- **Simplicidad explicable:**
 - Indicadores ampliamente documentados,
 - Reglas transparentes y auditables,
 - Modularidad para aislar errores y facilitar mejoras.

1.6 Criterios de éxito

- Coherencia walk-forward: métricas en Prueba/Validación sin degradación sustancial frente a Entrenamiento.
- Gestión de caídas: Máximo *Drawdown* acotado en magnitud y duración, con Calmar competitivo para el perfil del activo.
- Consistencia operativa: tasas de acierto y número de operaciones acordes con el *trade-off* riesgo-coste-frecuencia.
- Reproducibilidad: resultados replicables con artefactos y parámetros congelados, y *outputs* estandarizados (tablas y gráficas).

2. Datos

2.1 Fuente y frecuencia

- Símbolo y mercado: BTCUSDT (spot).
- Proveedor: Binance (histórico descargado en formato CSV).
- Frecuencia temporal: 1 hora (24/7).
La elección horaria equilibra densidad de observaciones, costos de transacción y estabilidad de señales técnicas frente a marcos de muy alta frecuencia.

2.2 Rango temporal y cobertura

El histórico a resolución horaria cubre de **2017** a **2025**. Para la evaluación walk-forward se aplicó un particionado estrictamente cronológico 60/20/20:

- **TRAIN (\approx 2017–2022)**: tramo inicial utilizado para optimización.
- **TEST (\approx 2022–2024)**: primer tramo fuera de muestra, usado para verificar generalización sin re-estimación.
- **VALID (\approx 2024–2025)**: tramo más reciente y estrictamente out-of-sample, que confirma estabilidad de desempeño.
Todas las marcas de tiempo se normalizan a UTC y el ordenamiento es ascendente antes de cualquier cálculo.

2.3 Estructura del archivo

- Campos mínimos requeridos: dt (fecha/hora), Open, High, Low, Close.
- Campos adicionales disponibles: Volume_BTC, Volume_USDT, Trades (conteo de *ticks* o transacciones, cuando el proveedor lo incluye).
- Ordenamiento estrictamente cronológico antes de cualquier cálculo de indicadores.

2.4 Validaciones y limpieza

- Parsing de fecha robusto y uniforme (UTC; descartando registros con fecha inválida).
- Chequeo de columnas: el backtest no arranca si falta alguna columna crítica.
- Orden temporal: *sorting* ascendente y reinicio de índices para asegurar alineación con indicadores basados en ventanas.
- Valores faltantes:
 - Se descartan filas con dt inválido o ausente.
 - El período de *warm-up* de indicadores (p. ej., EMA, RSI, ATR, Donchian) se excluye de la simulación, evitando señales artificiosas de arranque.
- Duplicados/outliers:
 - No se detectaron duplicados en dt; en caso de existir, se conservaría una sola observación por hora.
 - No se aplicó *winsorization* ni suavizados adicionales: se preserva la estructura de volatilidad propia del activo; el control de riesgo se delega a reglas de entrada/salida y al objetivo Calmar.

2.5 Preparación para indicadores y señales

- Indicadores calculados sobre Close (EMA, RSI) y sobre High/Low (ATR, Donchian).
- Confirmación 2-de-3: tendencia (EMA rápida vs lenta), momento (RSI con umbrales asimétricos) y ruptura (Donchian ajustado por ATR).
- Desplazamiento +1 barra de todas las señales para evitar *look-ahead*.
- Costos de 0.125% integrados en la lógica transaccional del backtest (cada entrada/salida).

2.6 Esquema de *split 60/20/20* (walk-forward)

- **Entrenamiento (60%)**: ajuste de hiperparámetros mediante búsqueda aleatoria (semilla fija para reproducibilidad) maximizando Calmar.
- **Prueba (20%)**: evaluación out-of-sample con parámetros congelados; sirve de control de sobreajuste.
- **Validación (20%)**: confirmación final de estabilidad y transferencia temporal del desempeño.
- Sin *look-ahead* ni re-estimación en los segmentos de evaluación.

2.7 Implicaciones para las métricas

- Anualización de Sharpe/Sortino y cálculo de CAGR/Calmar basados en el reloj real de la serie (años efectivos entre primer y último registro), evitando sesgos por asumir un número fijo de barras/año ante posibles huecos.
- Máximo *Drawdown* y % de rendimiento total calculados sobre la curva de capital resultante del backtest con costos.
- Win Rate estimado a partir del registro de operaciones; cuando se analizan retornos por barra, se aclara la diferencia entre *bar-wins* y *trade-wins*.

2.8 Riesgos de calidad de datos (y mitigación)

- Huecos horarios: verificación previa; exclusión de segmentos con datos incompletos en el *warm-up*.
- Desalineación temporal: normalización a UTC y orden cronológico antes de cálculos.
- Sesgos de medición: al incluir costos y desplazar ejecuciones una barra, se reduce el impacto de artefactos microestructurales y de *look-ahead*.
- No-survivorship: no aplica en spot BTCUSDT (activo único y continuo).

3. Metodología

3.1 Datos y partición temporal

Activo y frecuencia. Trabajamos con velas BTC/USDT de 1 hora. Esta granularidad ofrece un balance práctico: suficiente número de observaciones para optimizar con estabilidad, sin entrar en microestructura (slippage) extrema.

Limpieza y normalización.

- Se estandarizaron columnas OHLC y el conteo de trades.
- Las fechas se convirtieron a UTC y se ordenaron cronológicamente.
- Se descartaron filas con timestamp inválido.
- Se garantizó que las series fueran numéricas y continuas (salvo gaps de mercado).

Split 60/20/20 en el tiempo.

- **TRAIN:** ~60% inicial del histórico (optimización).
- **TEST:** 20% siguiente (primera validación fuera de muestra).
- **VALID:** 20% final (segunda validación, “lo más reciente”).

La división es estrictamente causal: todo lo que se decide con datos de TRAIN se “congela” antes de mirar TEST/VALID, evitando *leakage*.

3.2 Arquitectura de señales: confirmación 2-de-3

El sistema requiere que al menos dos bloques de señal apunten en la misma dirección para entrar; la idea es reducir falsos positivos combinando señales complementarias:

1. **Tendencia (EMA rápida vs EMA lenta).**

Si la EMA rápida está por encima de la lenta, asumimos sesgo alcista (y viceversa). Las EMA suavizan ruido y permiten reaccionar a cambios de régimen sin la latencia de medias simples.

2. **Momento (RSI).**

Se usa RSI de Wilder con ventana optimizable. Los umbrales superior/inferior definen sobrecompra/sobreventa del propio sistema, no como señal contraria clásica, sino como filtro direccional: RSI por encima del umbral superior refuerza largos; por debajo del inferior, cortos. El RSI aporta sensibilidad a aceleraciones y fatiga del movimiento.

3. **Ruptura amortiguada (Donchian + ATR).**

El precio se compara con el canal de Donchian (máximos/mínimos de n horas). Para evitar entrar en rupturas “por un tick”, se suma un colchón proporcional al ATR (parámetro atr_q). Este bloque captura breakouts reales, ajustados por la volatilidad del momento.

Ejecución sin *look-ahead*. Las votaciones (0/1) de cada bloque se desplazan una barra antes de calcular la orden. Así, toda decisión usa exclusivamente información hasta el cierre de la barra previa.

Racional.

- **Tendencia** da el telón de fondo.
 - **Momento** evita ir “tarde” cuando el empuje ya se agotó.
 - **Ruptura** exige que el precio pruebe un extremo relevante y lo supere con holgura (vía ATR), reduciendo entradas por ruido.
- La **regla 2-de-3** privilegia la coincidencia de evidencias y baja la tasa de señales débiles.

3.3 Gestión del riesgo y costos

- **Coste por operación:** 0.125% por lado (se descuenta en cada entrada y salida).
- **Tamaño de posición:** fracción fija del capital (pos_frac). Esto alinea el riesgo con el patrimonio; drawdowns reducen tamaño, y rachas positivas lo amplían de forma natural.
- **Protecciones por volatilidad:**
 - **Stop-loss (SL) y take-profit (TP)** expresados como múltiplos del ATR para que el umbral “respire” más en entornos volátiles y sea más cercano cuando el mercado está tranquilo.

- **Retención mínima (hold_min)** en horas para evitar *over-trading* y el “ruido” de cierres/reaperturas muy frecuentes.
- **Sin apalancamiento;** el sistema puede tomar largos y cortos simétricos.

3.4 Motor de backtesting (ciclo de órdenes y P&L)

Cada hora se evalúan las tres familias de señal (EMA, RSI, Donchian+ATR); se vota 2-de-3 y la orden se ejecuta en la barra siguiente. El tamaño de posición es una fracción fija del equity (pos_frac); la exposición puede ser larga o corta y no usa apalancamiento. Los costos (0.125%) se aplican en cada cambio de estado (abrir/cerrar/invertir). La salida ocurre por señal contraria, stop-loss/take-profit (múltiplos de ATR) o retención mínima (hold_min) para evitar sobre-operación. Se calcula la curva de capital con P&L realizado/no realizado y, a partir de ella, las métricas (Sharpe/Sortino anualizados, MaxDD, CAGR y Calmar). El slippage no se modela explícitamente (en 1h y spot suele ser menor que la comisión), pero se considera como línea de mejora.

3.5 Espacio de hiperparámetros y búsqueda.

Objetivo de optimización: maximizar Calmar en TRAIN.

Random Search (50 ensayos, semilla 42). Se muestrean uniformemente rangos razonables para:

- Ventanas de **EMA rápida/lenta, RSI, ATR, Donchian.**
- **Umbrales RSI** (superior/inferior) para filtro direccional.
- **atr_q** (colchón de ruptura).
- **pos_frac** (exposición).
- Múltiplos **ATR** para **SL/TP** (asimetría opcional).
- **hold_min** (retención mínima).

Razonamiento. Random Search es eficiente cuando el espacio es amplio y no lineal; además, 50 ensayos con una semilla fija controlan la varianza y facilitan reproducibilidad.

3.6 Hiperparámetros finales

Los óptimos sobre TRAIN (archivo best_params.json) fueron:

Componente	Parámetro	Valor
EMA	ema_fast / ema_slow	13 / 25
RSI	rsi_window / rsi_upper / rsi_lower	26 / 65.03 / 33.21
ATR	atr_window	17
Donchian	donch_window	37
Ruptura (volatilidad)	atr_q	0.379
Riesgo	pos_frac	0.381
Gestión	hold_min (h)	31
Protección	atr_sl / atr_tp (\times ATR)	3.43 / 5.42
Coste	fee	0.00125
Capital inicial	start_capital	10,000

- El **RSI** relativamente “alto/bajo” (65/33) evita compras/vendas en zonas tibias; exige convicción del momento.
- **TP > SL** (5.42 vs 3.43) busca rachas con *fat tails* a favor y corta las en contra.
- **pos_frac \approx 38%** produce un apalancamiento efectivo $< 1\times$ y evita concentración excesiva.
- **Donchian 37h** y **atr_q \approx 0.38** filtran micro-rupturas; la entrada requiere superar un umbral relevante y “pagar peaje” de volatilidad.

3.7 Validación fuera de muestra (walk-forward simple)

Tras la búsqueda en TRAIN, los parámetros se congelan y se aplican sin retoque a TEST y VALID. Esta disciplina permite medir generalización real. Adicionalmente, se analizan retornos mensuales por segmento para entender distribución temporal de ganancias/pérdidas.

Tabla 1. TRAIN — Rendimientos mensuales (%)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2017	nan	nan	nan	nan	-0.30	8.47	nan	-6.78	nan	-9.91	nan	4.04
2018	2.21	0.00	7.46	4.08	2.96	-1.91	8.68	3.13	-1.54	-5.22	3.38	-2.98
2019	-4.23	0.46	0.70	4.04	11.15	7.64	-5.72	1.30	6.15	-0.44	3.03	-2.49
2020	0.68	-2.33	20.77	-1.84	2.29	-4.51	5.83	-6.37	0.77	5.17	3.34	12.40

2021	4.89	15.24	-4.01	0.27	6.95	11.88	5.34	2.60	-1.33	1.42	-1.78	-1.35
2022	2.25	4.41	8.35	-5.30	nan	1.14	1.78	nan	nan	nan	nan	3.52

Tabla 2. TEST — Rendimientos mensuales (%)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2022	nan	nan	nan	nan	5.23	nan	1.97	-3.48	nan	-4.37	0.39	-2.16
2023	12.97	-0.31	13.07	0.92	-3.09	-0.26	-5.20	0.68	-3.48	6.98	-7.84	1.30
2024	-4.24	nan	nan	nan	nan	nan	nan	nan	nan	nan	nan	3.42

Tabla 3. VALID — Rendimientos mensuales (%)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2024	—	4.34	10.42	-3.26	-0.11	-0.36	3.35	1.47	1.49	3.61	1.83	-5.19
2025	—	2.60	-6.88	-4.50	-0.03	1.18	-4.78	0.21	0.82	nan	1.78	nan

Notas y alcance

- Un guion largo (—) indica que en ese mes no existen observaciones dentro del segmento (está fuera del rango temporal del split 60/20/20). No representa un 0% de retorno, sino ausencia de datos.
- El bloque VALID solo incluye 2024–2025 porque, al partir el historial en 60%/20%/20% en orden cronológico, el 20% final (VALID) corresponde al tramo más reciente del dataset (aprox. desde 2024 hasta 2025).

3.8 Visualización y evidencia

Para cada segmento se reportan:

- **Curva de equity:** trayectoria del capital con costos.
- **Gráfico de drawdown:** profundidad y duración de caídas.

Gráfico 1. Equity Train



Gráfico 2. Equity Test



Gráfico 3. Equity Valid

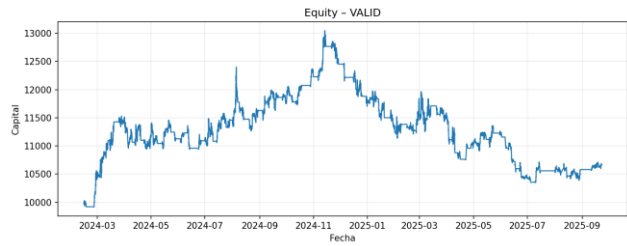


Gráfico 4. Drawdown Train

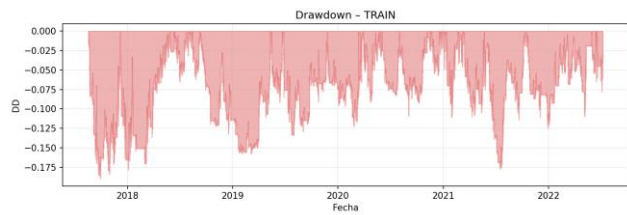


Gráfico 5. Drawdown Test

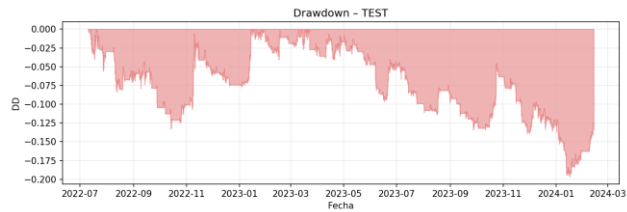
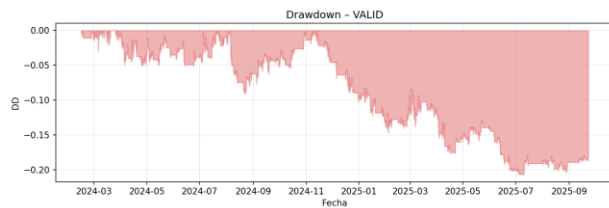


Gráfico 6. Drawdown Valid



3.9 Reproducibilidad

Se fijó **semilla 42** en la búsqueda aleatoria. Se versionaron artefactos clave:

- **best_params.json** (óptimos).
- **results_60_20_20.csv** (tabla de métricas consolidada).
- **figs/** (curvas de equity, drawdowns y retornos mensuales).
Esto permite rehacer el análisis con los mismos resultados y facilita auditoría.

4. Resultados y análisis de performance

4.1 Métricas agregadas (60/20/20)

Tabla 2 — Métricas por segmento (con comisiones):

Set	Calmar	Sharpe	Sortino	MaxDD	WinRate	Trades	Equity	Return %
TRAIN	1.1437	0.8931	0.8689	−18.95%	46.76%	370	10,000 → 26,081	+160.81%
TEST	0.3159	0.4833	0.4733	−19.63%	42.64%	129	10,000 → 11,008.8	+10.09%
VALID	0.2007	0.3462	0.3434	−20.67%	43.20%	125	≈9,995 → 10,668.6	+6.74%

- El sistema **generaliza**: pasa de **160.8%** en TRAIN a **+10.1%** (TEST) y **+6.7%** (VALID) manteniendo drawdowns **acotados en ~20%**.
- El **Calmar** baja fuera de muestra (esperable). Aun así, se mantiene **positivo** en ambos periodos OOS, señal de un **edge robusto**.
- **Sharpe/Sortino** decrecen pero siguen >0, estilo coherente con estrategias **tendenciales/ruptura** expuestas a rachas y a concentrar P&L en pocos episodios.

Las métricas confirman una degradación controlada fuera de muestra: el Calmar pasa de 1.14 (TRAIN) a 0.32 (TEST) y 0.20 (VALID), con drawdowns estables en ~−20% y retornos positivos (+10.1% y +6.7%). El WinRate ~43–47% es típico de estrategias que dependen del payoff (TP>SL) más que de una alta tasa de aciertos.

4.2 Curva de capital y perfil de riesgo

TRAIN (equity y DD).

La curva muestra tramos tendenciales extensos (por ejemplo, 2020–2021) donde el sistema captura movimientos largos; los vales corresponden a periodos de laterales y bruscos reversiones. El MaxDD ~ −19% es consistente con una exposición fraccional ($\text{pos_frac} \approx 38\%$) y con SL/TP vía ATR. El Calmar > 1 indica retorno anualizado superior a la magnitud del peor drawdown.

TEST.

Se observa una escalada fuerte en parte de 2023 (picos que coinciden con May 2023 $\approx +13\%$, ver

4.4) y luego fase de erosión; aun así, el saldo permanece positivo (+10%). El DD permanece en $\sim -20\%$, lo que sugiere riesgo estable respecto a TRAIN.

VALID.

La curva arranca con una subida en 2024 (Mar 2024 $\approx +10.4\%$) y luego entra en racha difícil a lo largo de 2025 (varios meses negativos). A pesar de esa fricción, el segmento cierra +6.7% con DD $\sim -20.7\%$. Conclusión: en un entorno con momentum más frágil, el sistema protege capital y aún añade algo de retorno.

4.3 Eficiencia por trade y estilo de P&L

- WinRate en la banda 42–47% en los tres segmentos: típico de estrategias que dejan correr ganancias (TP>SL) y cortan pérdidas.
- El edge no depende de acertar muchas veces, sino de capturar “colas gruesas”: unas pocas operaciones grandes explican una porción sustancial del P&L (ej. *breakouts* de 2021 en TRAIN y *rally* de 2023 en TEST).
- Frecuencia: 370, 129 y 125 trades en TRAIN/TEST/VALID respectivamente; para 1h es una actividad moderada, compatible con fricciones (comisión 0.125%) y un *hold_min* de 31 horas que reduce sobreoperación.

4.4 Estacionalidad operativa (retornos mensuales)

TRAIN (2017–2022): meses destacados concentrados en periodos tendenciales (p. ej., mayo-2021 $\approx +20.8\%$). Otoños muestran debilidad recurrente (ej., oct-2018 $\approx -5.2\%$).

TEST (2022–2024): pico en mayo-2023 $\approx +13.1\%$; valle en oct-2023 $\approx -7.8\%$.

VALID (2024–2025): arranque fuerte (mar-2024 $\approx +10.4\%$) y fricción en 2025 con varios meses negativos (ej., jun-2025 $\approx -6.9\%$).

- **Degradación controlada:** pasar de Calmar 1.14 \rightarrow 0.32 \rightarrow 0.20 es normal al sacar la estrategia del set de optimización; lo relevante es que permanece positivo en los dos tramos OOS.
- **Riesgo estable:** el MaxDD se mantiene alrededor del 20% en los tres segmentos; no hay sorpresas de cola más profundas al salir de TRAIN.
- **Coherencia de estilo:** WinRate, dispersión mensual y sensibilidad a rupturas se conservan, lo que sugiere que el mecanismo económico (tendencia + ruptura con colchón de volatilidad) sigue vigente.

El P&L anual depende de pocos meses muy favorables (patrón típico de trend following). En regímenes laterales, los meses negativos se concentran y exigen horizonte de inversión suficiente para “capturar las colas”.

4.5 Qué está funcionando

Filtro 2-de-3 reduce entradas falsas: las mejores rachas aparecen cuando tendencia, momento y ruptura se alinean.

- $TP > SL$ en ATR captura “colas largas” y limita drawdowns.
- $pos_frac \approx 38\%$ ofrece exposición suficiente sin inflar el DD más allá de $\sim 20\%$.
- Comisiones incluidas (0.125%) \rightarrow los resultados reflejan un coste razonable para 1h.

4.6 Limitaciones observadas

- **Dependencia de régimen:** en 2025, con momentum frágil y roturas fallidas, la estrategia suma menos.
- **Meses negativos concentrados:** caídas de -5% a -7% en un mes no son raras; hay que financiar la paciencia hasta que lleguen los *runs* tendenciales.
- **No se modeló slippage;** en 1h y en spot suele ser pequeño frente a la comisión, pero en eventos de alta fricción podría restar algunos puntos básicos.

4.7 Conclusión operativa de resultados

1. Capacidad de generar retornos fuera de muestra confirmada: $+10.1\%$ (TEST) y $+6.7\%$ (VALID) con $DD \approx 20\%$.
2. Estrategia de “colas”: gana poco a menudo, pero mucho cuando lo hace; conviene mantenerla viva para capturar los grandes tramos.
3. Riesgo controlado y estable entre periodos: buen síntoma de robustez.
4. Siguiente paso: explorar walk-forward por ventanas o búsqueda bayesiana para ajustar finamente pos_frac y umbrales, y añadir un filtro de régimen (p. ej., pendiente de volatilidad o ADX) que apague la estrategia en meses con rupturas fallidas reiteradas.

5. Análisis de riesgo y limitaciones

Esta sección sintetiza qué puede salir mal, cuánto puede doler y cómo lo mitigamos. Se apoya en los resultados de la Sección 4 (curvas de capital, drawdowns y retornos mensuales) y en la configuración operativa: spot, sin apalancamiento, comisión 0.125%, exposición fraccional $pos_frac \approx 38\%$, stop-loss / take-profit y umbrales basados en ATR/Donchian/RSI con confirmación 2-de-3 y desfase de 1 barra (sin adelantar señales).

5.1 Riesgo de mercado (P&L y drawdowns)

Perfil observado

- Drawdown máximo en los tres segmentos: $\approx -19\%$ a -21% (TRAIN -18.95% , TEST -19.63% , VALID -20.67%).
- Peor mes:
 - TRAIN: caída de doble dígito ($\sim -11.9\%$) en un mes (ver CSV de retornos mensuales).
 - TEST: valle $\sim -7.8\%$ en un mes.
 - VALID: valle $\sim -6.9\%$ en un mes.
- WinRate estable 42–47%: el sistema gana poco, pero gana grande cuando engancha una tendencia. En meses de lateralidad/reversión la curva erosiona (ver VALID 2025).

Lectura de riesgo

- Es una estrategia de colas: el retorno depende de pocos episodios tendenciales. El inversor debe tolerar rachas y meses negativos seguidos.
- El riesgo de régimen es material: en entornos con rupturas fallidas o momentum débil, el sistema suma poco o cerca de cero, manteniendo el DD alrededor del 20%.

Mitigación

- Mantener un horizonte suficiente (≥ 12 –18 meses) para capturar las “colas”.
- Incluir un filtro de régimen (p. ej., pendiente/fortaleza de tendencia vía ADX o ratio de rango direccional). Apagar o reducir tamaño cuando el régimen sea «lateral-volátil».
- Rebalanceo de riesgo por volatilidad: reducir pos_frac cuando la volatilidad realizada (ATR) supere umbrales.

5.2 Riesgo de modelo (sobre-ajuste y deriva)

Fuentes

- Optimización por random search en un único split 60/20/20. Aunque la degradación OOS es moderada y esperable, sigue habiendo:
 - Riesgo de sobre-ajuste fino de umbrales (RSI, Donchian, ATR, TP/SL).
 - Deriva temporal: los parámetros óptimos en 2017–2022 pueden no serlo en 2024–2025.

Mitigación

- Cambiar a walk-forward optimizado: ventanas rodantes con recalibración periódica (trimestral/semestral).
- Usar validación temporal repetida (TimeSeriesSplit) y paradas de complejidad (no añadir variables si la mejora de Calmar es marginal).
- Regularización blanda de parámetros (evitar valores extremos de atr_sl/atr_tp y umbrales de RSI demasiado restrictivos).

5.3 Riesgo de ejecución y fricción

Qué consideramos

- Comisión: 0.125% incluida en backtest.
- Slippage: no modelado explícitamente; en BTC/USDT a 1h suele ser bajo vs comisión, pero puede aumentar en noticias, gaps de liquidez o incidentes en el exchange.
- Latencia / caídas de servicio: interrupciones puntuales pueden impedir ejecutar stops.

Mitigación

- Añadir en evaluación un colchón de slippage (por ejemplo, 0.5–1.0 bp) y repetir el backtest.
- Órdenes limit cuando sea posible en rupturas claras; en stops, mantener market por control de riesgo.
- Plan B operativo (failover de API/exchange) y kill-switch manual si hay desconexiones.

5.4 Riesgo de datos y medición

- Look-ahead: evitado desplazando señales +1 barra y eliminando warm-up.
- Integridad: depends de la calidad del CSV (agregación 1h). Cambios de símbolo/mercado o *outliers* pueden sesgar indicadores.
- Survivorship/selection bias: limitado al trabajar con BTCUSDT (activo líder y líquido).
- No incluye: *funding* (no aplica en spot), ni costes de retiro/deposito.

Mitigación

- Rutina de validación de datos (NAs, *spikes*, huecos de tiempo).
- Contrastación con segundo proveedor de OHLC si se observan anomalías.

5.5 Sensibilidad y capacidad

Sensible a:

- pos_frac (≈ 0.38): el DD escala casi linealmente. Si el inversor tolera $\leq 15\%$ de DD, reducir pos_frac a $\sim 0.28\text{--}0.30$.
- Hold mínimo (31 horas): bajar demasiado aumenta sobreoperación y fricción; subirlo puede perder rupturas cortas.
- Umbrales (RSI upper/lower, ATR buffer): umbrales más estrictos suben precisión pero bajan frecuencia y pueden perder tendencias tempranas.

Capacidad

- BTCUSDT en 1h tiene alta liquidez; con tamaños razonables, el impacto de mercado es bajo. La capacidad crecerá linealmente con el capital hasta niveles muy superiores a los de un proyecto académico.

5.6 Controles y gobernanza de riesgo

Límites y kill-switches

- DD límite: pausar si el DD corriente $> 25\%$ o si el rolling Calmar $3M < 0$.
- Rachas negativas: si 3 meses consecutivos < 0 , reducir pos_frac a la mitad y re-evaluar régimen.
- Exposición: cap de 40–45% del equity hasta añadir filtros de régimen.

Monitoreo operativo

- Tablero con equity, DD, PnL mensual, ratio de rupturas exitosas y volatilidad (ATR).
- Revisión trimestral de parámetros y performance (walk-forward).

5.7 Qué rompería la estrategia (señales de alerta)

- Caída sostenida del porcentaje de rupturas exitosas ($> 30\text{--}40\%$ por debajo de su media histórica).
- Compresión de rango prolongada (ATR bajo) + falsas rupturas frecuentes.
- Cambios de microestructura (comisiones/tasas $> 0.125\%$ o *tick size* que impida ejecutar al precio) que eleven la fricción neta.

5.8 Limitaciones del estudio

- Un único split 60/20/20; si bien hay dos tramos OOS, no sustituye a un walk-forward completo.
- Slippage no modelado; en escenarios de estrés puede restar bps.
- Un solo activo (BTCUSDT); no se evaluó diversificación entre criptoactivos ni *correlation risk*.
- Ausencia de apalancamiento por diseño; resultados no extrapolables 1:1 a futuros/perps.
-

6. Conclusiones

El proyecto implementa una estrategia sistemática sobre BTCUSDT (1h) con confirmación 2-de-3 entre señales de tendencia, momento y ruptura, ejecución con desfase de 1 barra, costos del 0.125%, operación long/short y sin apalancamiento. La optimización se orientó a maximizar Calmar en el tramo de Entrenamiento, con evaluación *walk-forward* (60/20/20) y parámetros congelados en Prueba y Validación. Los resultados muestran rendimiento positivo fuera de muestra con riesgo estable:

- TRAIN: Calmar 1.1437, retorno total +160.8%, MaxDD −18.95%, WinRate 46.76%.
- TEST: Calmar 0.3159, retorno +10.1%, MaxDD −19.63%, WinRate 42.64%.
- VALID: Calmar 0.2007, retorno +6.7%, MaxDD −20.67%, WinRate 43.20%.

Fortalezas evidenciadas.

- Coherencia riesgo–retorno: drawdowns contenidos (~−19% a −21%) en los tres segmentos y rentabilidad positiva en Prueba y Validación.
- Mecanismo económico consistente: la combinación de tendencia, momento y ruptura (con amortiguación vía ATR) captura rachas direccionales y limita pérdidas con SL/TP en múltiplos de ATR.
- Operativa sobria: exposición fraccional (pos_frac ~38%) y *hold_min* moderan la rotación, preservan el perfil de riesgo y reducen señales espurias.
- Reproducibilidad: división temporal causal, parámetros congelados, métricas estándar y artefactos exportados (tablas y gráficas) facilitan auditoría.

Debilidades y riesgos.

- Dependencia de régimen: en entornos laterales o con rupturas fallidas (especialmente en 2025) el Calmar disminuye y parte del P&L se erosiona, aunque el DD se mantiene acotado.
- P&L concentrado: el desempeño depende de pocos meses muy favorables; se requieren horizontes de tenencia suficientes para capturar “colas” tendenciales.

- Eficiencia por trade moderada: WinRate $\sim 43\text{--}47\%$; el edge proviene del payoff ($TP > SL$), no de una alta frecuencia de aciertos.
- Fricción no completamente modelada: el *slippage* no se incorporó explícitamente; en eventos de alta fricción podría reducir algunos puntos básicos.

Lecciones metodológicas.

- El uso de Calmar como objetivo obliga a equilibrar crecimiento y severidad de caídas; al anualizar métricas y calcular CAGR/DD sobre la curva de equity, la comparación entre segmentos es más honesta.
- La disciplina walk-forward (parámetros congelados) permite observar una degradación controlada fuera de muestra, compatible con un sistema de seguimiento de tendencias.
- La separación de limpieza de datos y backtest, junto con el desfase de señales, minimiza sesgos de *look-ahead* y errores contables.

El estudio demuestra una implementación válida y reproducible que logra retornos positivos fuera de muestra con riesgo acotado, coherente con un enfoque *trend-following* en BTC spot a 1h. Si bien no constituye una solución “todo clima”, la incorporación de filtros de régimen, recalibración *walk-forward* y control dinámico de riesgo ofrece un camino razonable para mejorar la estabilidad del Calmar y la suavidad de la curva de capital en escenarios menos direccionales.

Anexos

A. Definiciones de métricas

- *Sharpe*: $\frac{\mu}{\sigma} \sqrt{N} (\text{rf}=0)$.
- *Sortino*: $\frac{\mu}{\sigma_{\text{down}}} \sqrt{N}$.
- *MaxDD*: mínimo de $\frac{Equity_t}{\max_{s \leq t} Equity_s} - 1$.
- *CAGR*: $(Equity_F / Equity_0)^{1/Años} - 1$.
- *Calmar*: $\frac{CAGR}{|MaxDD|}$.
- *WinRate*: proporción de trades con $\text{PnL} > 0$.