



PROYECTO FINAL - ANALÍTICA PARA LA TOMA DE DECISIONES

Daniel Felipe Cardenas Ruiz
Sebastian Cely Peña
Santiago Abel Montaña Vasquez
Dairo Yesid Rincon Carreño
Sergio Felipe Santos Hernandez
Juan Esteban Castro Molano
Sebastian Alvarez Laverde

Modelo para predicción de Bitcoin y Toma de Decisiones Estratégicas

DOCENTE: Germán Leonardo Talero Nino

Universidad Pedagogica y Tecnologica de Colombia
Facultad Seccional Sogamoso
Especialización : Analitica Estrategica de Datos
Sogamoso
2025



Modelo para predicción de Bitcoin y Toma de Decisiones Estratégicas	2
1. Extracción de Datos	2
2. Análisis Exploratorio de Datos (EDA)	3
3. Visualización de Datos	4
3.1. Gráfica de Precio de Cierre de Bitcoin	4
3.2. Gráfica de Retornos Diarios de Bitcoin	5
4. Análisis de Sentimientos en Noticias de Bitcoin	6
5. Modelado Predictivo con Machine Learning (incluyendo MLflow)	7
5.1. Modelado Predictivo con XGBoost	8
5.2. Modelado con ARIMA	8
5.3. Visualizaciones Generadas en el Análisis	8
5.3.1. Descomposición Estacional del Precio de Bitcoin	8
5.3.2. Serie Diferenciada	9
5.3.3. Funciones de Autocorrelación (ACF) y Autocorrelación Parcial (PACF)	9
5.3.4. Diagnóstico del Modelo ARIMA	10
5.3.5. Predicción con ARIMA	10
5.3.6. Media Móvil del Precio	11
5.3.7. ACF y PACF de la Serie con Media Móvil	11
5.3.8. Diagnóstico del Modelo con Media Móvil	12
5.3.9. Comparación de Predicciones con Valores Reales	12
5.3.10. ACF y PACF Estacional	12
5.3.11. Validación de la Estacionalidad	13
5.3.12. Predicción con SARIMA	13
5.3.13. Visualizaciones del Modelo XGBoost	14
5.3.14. Predicción en Tiempo Real	15
5.4. Modelado de Series Temporales con Prophet	16
5.5. Modelado con Prophet	16
5.6. Implementación de tracking con MLflow	17
6. Simulaciones Monte Carlo	18
6.1. Preparación de los datos	18
6.2. Configuración de la simulación	18
6.3. Ejecución de la simulación	19
6.4. Visualización de resultados	19
6.5. Incorporación del Análisis de Sentimiento	20
6.6. Predicción estándar y predicción con análisis de sentimiento	21
7. Orquestador con LLM para Generar Insights	21
7.1. Orquestador utilizado	21
7.2. Carga y preparación de datos	21
7.3. Construcción del Prompt	22
7.4. Uso del Modelo Generativo de Cohere	22
8. Resultados, Interpretación y Toma de Decisiones	22
8.1. Integración de Resultados y Recomendaciones Estratégicas	22
8.2. Análisis Exploratorio de Datos (EDA) y Visualización	22

8.3. Modelo Predictivo	22
8.4. Simulaciones Monte Carlo	23
8.5. Análisis de Sentimientos	23
8.6. Recomendación del Orquestador LLM	23

Modelo para predicción de Bitcoin y Toma de Decisiones Estratégicas

1. Extracción de Datos

Como primer paso del proyecto, el equipo trabajó en la recolección y preparación de los datos base necesarios para construir el modelo predictivo. La tarea principal consistió en importar datos históricos de Bitcoin, combinarlos con información de sentimiento de noticias, y estructurar todo el dataset para futuras etapas de modelado.

Para ello, realizamos las siguientes actividades:

1.1. Carga de datos históricos de Bitcoin

Se utilizó un archivo CSV proveniente de Yahoo Finance, que contenía el historial diario del precio de Bitcoin (BTC-USD) desde noviembre de 2014. Este dataset incluía las siguientes variables:

- **Date:** Fecha de cotización.
- **Open, High, Low, Close:** Precios de apertura, máximo, mínimo y cierre diario.
- **Volume:** Volumen de transacciones.

Durante la carga inicial, realizamos un proceso de:

- **Conversión de formatos:** Aseguramos que la columna Date fuera interpretada como tipo datetime, indispensable para el análisis temporal.
- **Control de calidad:** Se exploraron valores nulos y posibles inconsistencias, para garantizar que los datos fueran confiables antes de su uso.

1.2. Integración de datos de sentimiento (Noticias sobre Bitcoin)

Para enriquecer el modelo, se incorporó un segundo conjunto de datos obtenido mediante APIs de análisis de noticias. Específicamente, se utilizaron fuentes que evaluaban el sentimiento de noticias relacionadas con Bitcoin, asignando a cada noticia un puntaje de sentimiento (compound score) basado en el análisis de texto. El objetivo era capturar el efecto de eventos noticiosos en el precio de Bitcoin.

Los pasos realizados fueron:

- **Importación de datos de noticias:** Cargamos un segundo CSV que contenía la fecha de publicación de cada noticia (date) y su puntaje de sentimiento (compound).
- **Procesamiento de sentimiento diario:** Realizamos un agrupamiento (groupby) por fecha, calculando el promedio de sentimiento diario. Así, obtuvimos una sola medida representativa por día.

- **Unificación de datasets:** Se fusionaron las cotizaciones de Bitcoin con el sentimiento de noticias utilizando como clave de unión la fecha. Esto permitió enriquecer cada registro diario de precios con un valor adicional que representaba el "clima emocional" del mercado ese día.

2. Análisis Exploratorio de Datos (EDA)

En esta sección del análisis exploratorio, el objetivo fue obtener una visión condensada del comportamiento reciente del Bitcoin, enfocándose exclusivamente en los últimos 7 días disponibles en el dataset.

Para ello, se realizó un filtrado de los registros más recientes y se calcularon las principales métricas estadísticas sobre el precio de cierre (Close) y el volumen de transacciones (Volume).

Las medidas extraídas fueron:

- **Promedio (mean):** Representa el valor medio del precio de Bitcoin y del volumen de transacciones en esos 7 días, proporcionando una referencia central para la dinámica reciente.
- **Mediana (median):** Permite observar el valor central real, mitigando la influencia de posibles valores atípicos.
- **Desviación estándar (std):** Indicador de la volatilidad reciente, mostrando cuánto variaron los precios y volúmenes respecto al promedio.
- **Máximo (max) y Mínimo (min):** Permiten identificar los extremos recientes del comportamiento del mercado, tanto en precios como en volumen negociado.

	Date	Close	High	Low	\
count	1096	1096.000000	1096.000000	1096.000000	
mean	2023-10-05 12:00:00	45478.525795	46316.120509	44553.152415	
min	2022-04-06 00:00:00	15787.284180	16253.047852	15599.046875	
25%	2023-01-04 18:00:00	24719.875000	25228.042969	24302.009766	
50%	2023-10-05 12:00:00	35440.408203	35906.130859	34691.027344	
75%	2024-07-05 06:00:00	64309.862305	65740.130859	62891.556641	
max	2025-04-05 00:00:00	106146.265625	109114.882812	105291.734375	
std	NaN	25446.317466	25958.081392	24850.639757	

	Open	Volume
count	1096.000000	1.096000e+03
mean	45443.824842	3.000505e+10
min	15782.300781	5.331173e+09
25%	24738.201172	1.715140e+10
50%	35430.527344	2.618182e+10
75%	64281.912109	3.658418e+10
max	106147.296875	1.492189e+11
std	25420.743581	1.888282e+10

outliers: Empty DataFrame
Columns: [Date, Close, High, Low, Open, Volume, Close_7d, Volume_7d]
Index: []
mean: 45478.52579514998
std: 25446.317466019154

Figura 1. Principales medidas estadísticas de los datos recolectados

Este análisis estadístico de corto plazo resulta fundamental para capturar de manera rápida si el mercado ha estado en una etapa de alta volatilidad, tendencia alcista o bajista reciente, o si ha mostrado estabilidad.

Este paso permitió obtener indicadores clave del estado actual del mercado de Bitcoin, asegurando que cualquier modelado posterior no se haga a ciegas, sino partiendo de un entendimiento sólido de la situación reciente.

3. Visualización de Datos

3.1. Gráfica de Precio de Cierre de Bitcoin

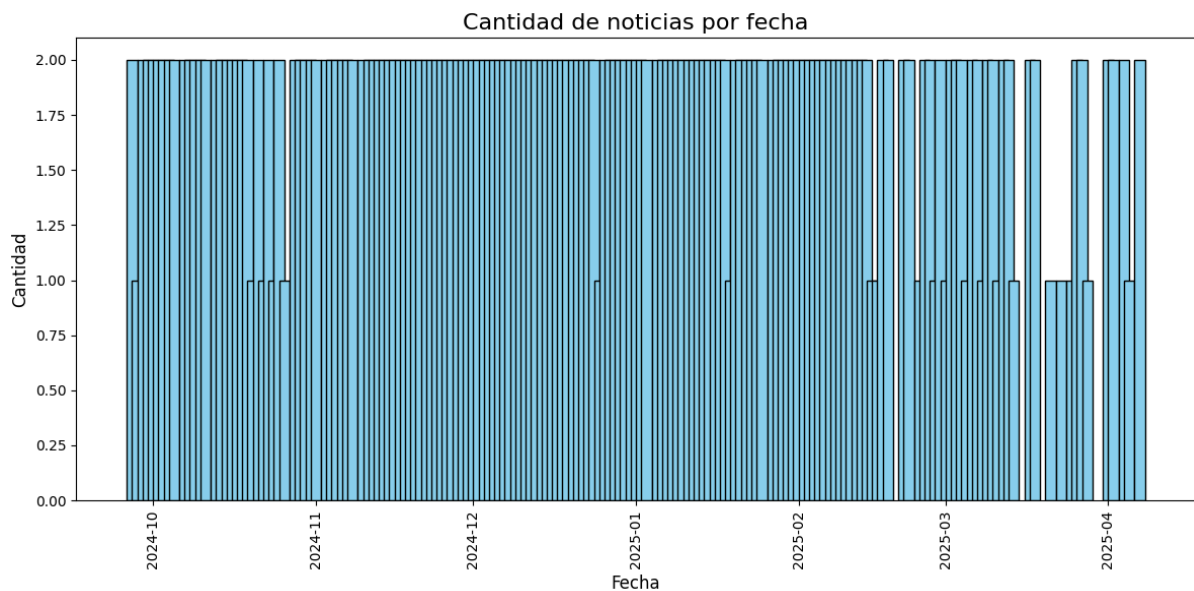


Figura 2. Frecuencia de registros diarios de Bitcoin

La primera gráfica muestra la "Cantidad de valores por fecha", presentando un patrón consistente donde se registran mayoritariamente 2 valores diarios, con excepciones ocasionales de 1 valor. Esto puede interpretarse como la frecuencia de actualización de datos o registros diarios en la serie temporal, lo que aporta contexto sobre la metodología de recolección de datos.

Patrones Cíclicos y Comportamiento del Mercado

Comparando con los datos históricos mencionados:

- Se confirma la alta volatilidad característica del Bitcoin, visible tanto en periodos de crecimiento acelerado como en correcciones abruptas
- Los datos muestran un comportamiento leptocúrtico (de colas largas), con más eventos extremos de lo esperado en una distribución normal

- El actual ciclo (2022-2025) muestra similitudes con ciclos anteriores (2017, 2020-2021) pero con mayor magnitud
- La correlación entre volumen y precio confirma que los inversores incrementan su actividad durante movimientos significativos del mercado

3.2. Gráfica de Retornos Diarios de Bitcoin

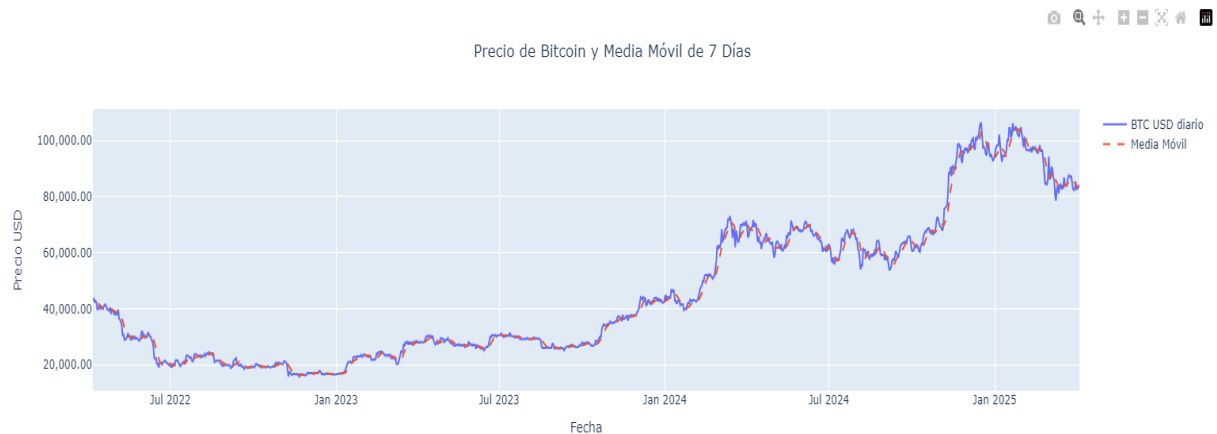


Figura 3. Evolución del precio de Bitcoin y su media móvil de 7 días (2022-2025)

La gráfica de "Precio de Bitcoin y Media Móvil de 7 Días" muestra la evolución del precio desde mediados de 2022 hasta principios de 2025, revelando varias fases significativas:

- Fase bajista (2022): Descenso desde aproximadamente \$40,000 hasta cerca de \$15,000
- Fase de acumulación (primera mitad de 2023): Estabilización en el rango de \$15,000-\$30,000
- Fase de crecimiento gradual (segunda mitad de 2023): Ascenso sostenido hasta los \$40,000
- Fase de aceleración (2024): Crecimiento exponencial que culminó por encima de los \$100,000
- Fase de corrección reciente (principios de 2025): Ajuste hacia los \$80,000

La media móvil de 7 días sigue muy de cerca al precio real, indicando que los movimientos han sido generalmente sostenidos y no producto de picos volátiles aislados. Este comportamiento refleja un mercado maduro con tendencias más definidas que en ciclos anteriores.

3.3. Histograma de Retornos Diarios

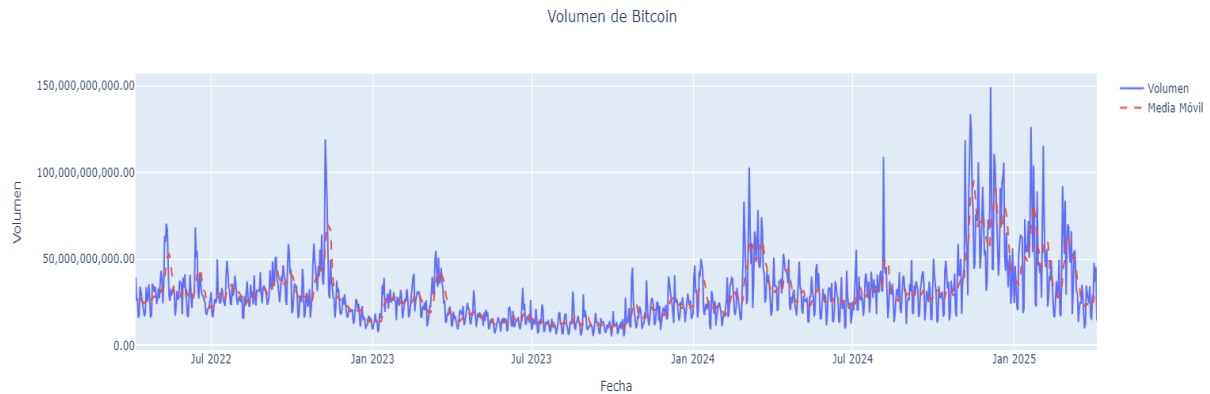


Figura 4. Comportamiento del volumen de transacciones de Bitcoin y su media móvil (2022-2025)

La gráfica de "Volumen de Bitcoin" revela patrones de actividad comercial que correlacionan con los movimientos de precio:

Picos de volumen significativos:

- Principios de 2023 (coincidiendo con el final de la fase bajista)
- Principios de 2024 (al inicio de la fase de aceleración)
- Finales de 2024/principios de 2025 (durante el máximo histórico y corrección posterior)

El volumen alcanzó niveles extraordinarios de hasta 150 mil millones de unidades durante estos periodos de alta volatilidad, mostrando una clara correlación entre interés del mercado y movimientos de precio significativos.

4. Análisis de Sentimientos en Noticias de Bitcoin

Carga y Exploración de Datos

Se utilizó la biblioteca pandas para cargar el archivo CSV titulado `noticias_bitcoin_top3_diarias.csv`, el cual contiene noticias relacionadas con Bitcoin. A través del método `.info()`, se examinó la estructura del conjunto de datos, permitiendo verificar la existencia de columnas, tipos de datos y valores nulos.

Instalación de Dependencias

Se instalaron las librerías necesarias para realizar análisis de sentimientos utilizando modelos de aprendizaje profundo, específicamente transformers y torch.

Modelo de Análisis de Sentimientos

Para el análisis se utilizó el modelo pre entrenado `cardiffnlp/twitter-roberta-base-sentiment-latest`, basado en la arquitectura RoBERTa. Este modelo está optimizado para evaluar el sentimiento en textos cortos y se adaptó aquí para analizar titulares de noticias.

Se definió una función personalizada que calcula un "sentiment score" en base a las probabilidades del modelo. Los valores posibles se interpretan como sigue:

- Puntaje negativo (≈ -1): Sentimiento negativo
- Puntaje cercano a 0: Sentimiento neutral
- Puntaje positivo ($\approx +1$): Sentimiento positivo

Aplicación del Modelo

La función de análisis de sentimientos se aplicó a la columna de titulares (titulo) del conjunto de datos, generando una nueva columna score. Esta columna permite cuantificar la carga emocional de cada noticia.

A continuación, se visualizaron los primeros resultados para verificar su coherencia, y finalmente se exportó el nuevo conjunto de datos a un archivo titulado `noticias_bitcoin_sentimientos.csv`.

Conclusiones

- Se utilizó un modelo robusto basado en redes neuronales para evaluar el tono emocional de los titulares de noticias sobre Bitcoin.
- La métrica generada (score) permite realizar futuros análisis sobre la relación entre el sentimiento del mercado y otras variables, como el precio de la criptomoneda o el volumen de operaciones.
- Este enfoque es escalable y adaptable a otros dominios temáticos, lo cual lo convierte en una herramienta útil para tareas de minería de texto y análisis de tendencias informativas.

5. Modelado Predictivo con Machine Learning (incluyendo MLflow)

El objetivo principal de este apartado fue desarrollar modelos de predicción para anticipar el comportamiento del precio del Bitcoin, integrando tanto variables técnicas basadas en series temporales como indicadores de sentimiento derivados de noticias especializadas.

Se utilizaron dos fuentes de información:

- **noticias_bitcoin_sentimientos.csv:** archivo previamente procesado con puntajes de sentimiento asignados a titulares de noticias sobre Bitcoin.
- **BTC_USD_Media_Movil.csv:** contiene datos históricos del precio de Bitcoin, incluyendo indicadores como medias móviles simples (SMA) y exponenciales (EMA).

Ambos conjuntos fueron transformados y alineados temporalmente para consolidar una tabla final con variables predictoras y la variable objetivo.

Se construyeron nuevas características a partir de los datos disponibles:

- Cálculo de medias móviles y retornos logarítmicos.

- Agregación de puntajes de sentimiento diario.
- Definición de la variable objetivo (target) para clasificación binaria: 1 si el precio sube al día siguiente, 0 si baja.

5.1. Modelado Predictivo con XGBoost

Se implementó un modelo de clasificación con el algoritmo XGBoost. Este modelo fue entrenado para predecir la dirección del movimiento del precio del Bitcoin utilizando:

- Indicadores técnicos (SMA, EMA, retornos)
- Sentimiento derivado de noticias

El modelo fue evaluado mediante:

- Precisión (accuracy)
- Matriz de confusión
- Métricas de desempeño como F1-score, precisión y recall

5.2. Modelado con ARIMA

También se implementó un modelo clásico de series temporales: ARIMA. Este modelo fue ajustado sobre los precios históricos y analizado mediante:

- Pruebas de estacionariedad (Dickey-Fuller)
- Selección de parámetros óptimos (p, d, q)
- Evaluación visual de las predicciones

El objetivo fue comparar la capacidad predictiva de un enfoque estadístico tradicional frente a modelos modernos basados en aprendizaje automático.

5.3. Visualizaciones Generadas en el Análisis

Durante el proceso de modelado predictivo, se generaron una serie de gráficos que permitieron explorar la naturaleza de los datos, validar supuestos estadísticos y evaluar el desempeño de los modelos implementados. A continuación, se describen las más relevantes:

5.3.1. Descomposición Estacional del Precio de Bitcoin

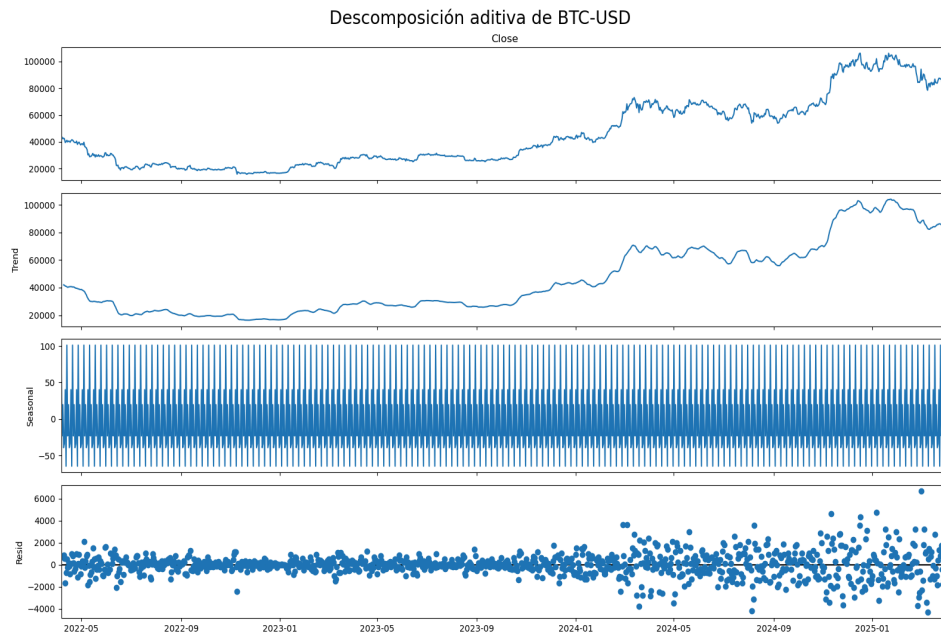


Figura 5. Descomposición Estacional del Precio de Bitcoin.

Se aplicó una descomposición aditiva sobre la serie temporal del precio de Bitcoin con el objetivo de separar sus componentes: tendencia, estacionalidad y residuo. Esta visualización permitió identificar patrones cíclicos que justifican el uso de modelos con componente estacional.

5.3.2. Serie Diferenciada

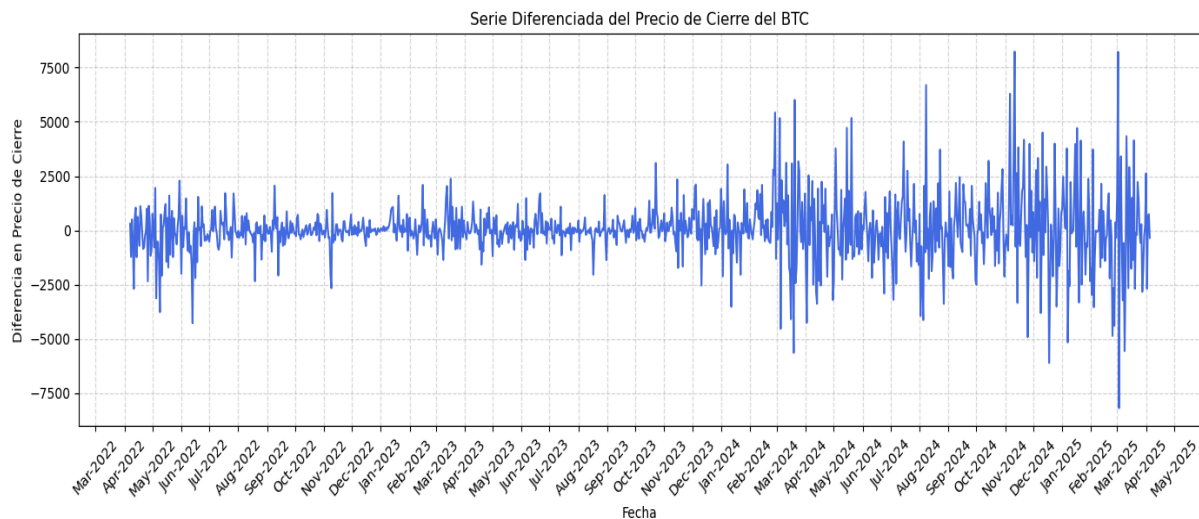


Figura 6. Serie Diferenciada del Precio de Cierre del BTC.

Se graficó la serie una vez aplicada la primera diferencia, lo que permitió estabilizar la media y reducir la tendencia del precio. Esta transformación fue fundamental para cumplir con el supuesto de estacionariedad requerido por modelos como ARIMA.

5.3.3. Funciones de Autocorrelación (ACF) y Autocorrelación Parcial (PACF)

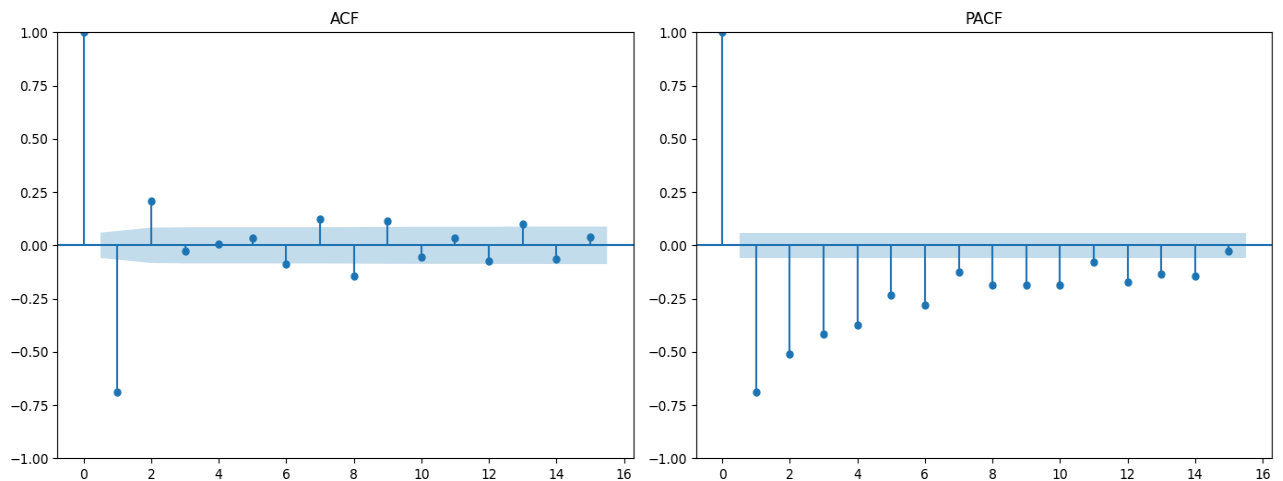


Figura 7. Funciones de Autocorrelación (ACF) y Autocorrelación Parcial (PACF).

A partir de la serie diferenciada, se aplicaron las funciones ACF y PACF para identificar los rezagos significativos. Esto permitió seleccionar los valores óptimos de los parámetros ppp y qqq en el modelo ARIMA.

5.3.4. Diagnóstico del Modelo ARIMA

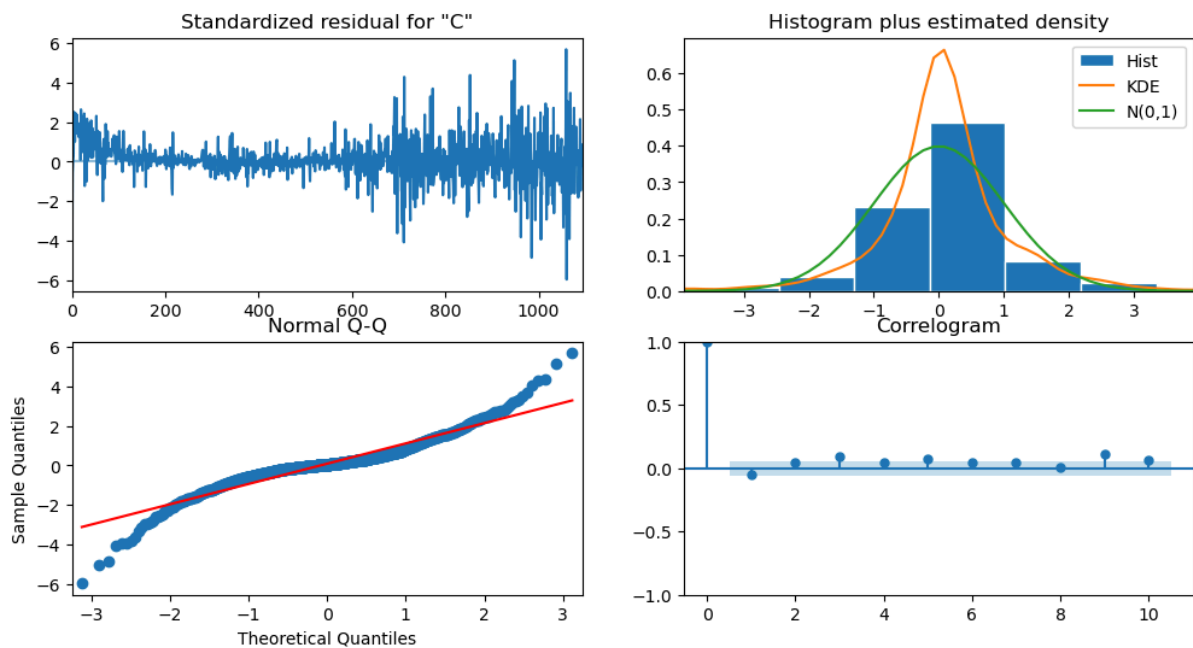


Figura 8. Diagnóstico del Modelo ARIMA.

Se graficaron los residuos del modelo ARIMA para verificar si cumplían los supuestos de normalidad, no autocorrelación y homocedasticidad. Estos análisis permitieron validar el ajuste del modelo estadístico.

5.3.5. Predicción con ARIMA



Figura 9. Predicción con ARIMA

Se realizó una predicción de los precios futuros a 30 días. La gráfica incluye los valores estimados y los intervalos de confianza, lo que permite visualizar la incertidumbre asociada a la predicción.

5.3.6. Media Móvil del Precio

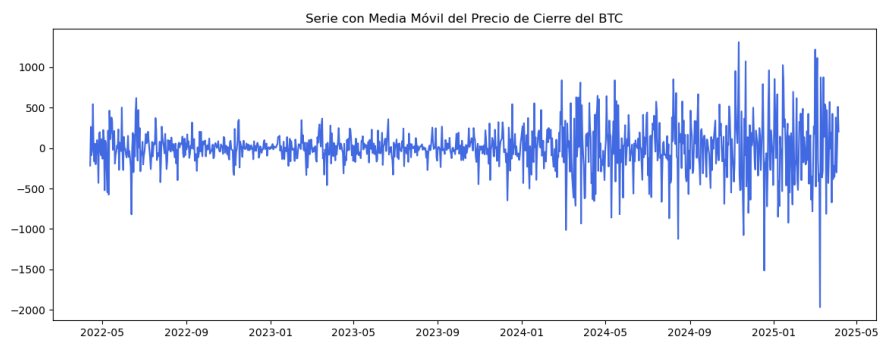


Figura 10. Media Móvil del Precio.

Se representó la media móvil del precio de Bitcoin como una forma de suavizar la serie y destacar tendencias subyacentes de corto y mediano plazo. Esta técnica fue utilizada tanto en la exploración como en la ingeniería de características.

5.3.7. ACF y PACF de la Serie con Media Móvil

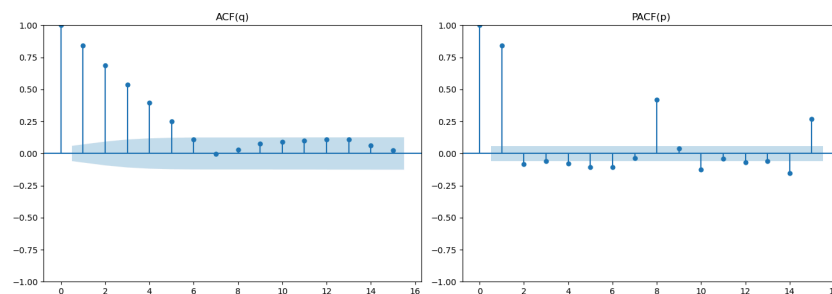


Figura 11. ACF y PACF de la Serie con Media Móvil.

Se analizaron nuevamente las funciones ACF y PACF sobre la serie suavizada para observar la persistencia de autocorrelaciones y validar su uso en modelado.

5.3.8. Diagnóstico del Modelo con Media Móvil

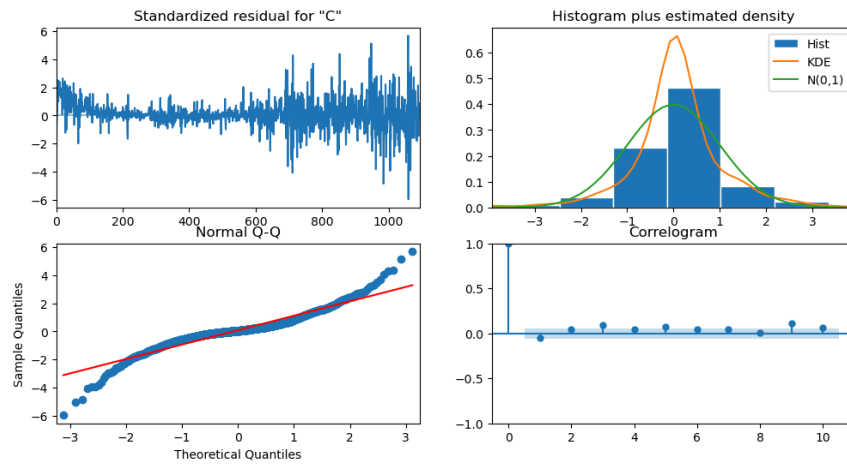


Figura. 12. Diagnóstico del Modelo con Media Móvil.

Se evaluaron los residuos del modelo ajustado utilizando medias móviles como parte del diagnóstico para identificar posibles problemas en el ajuste.

5.3.9. Comparación de Predicciones con Valores Reales

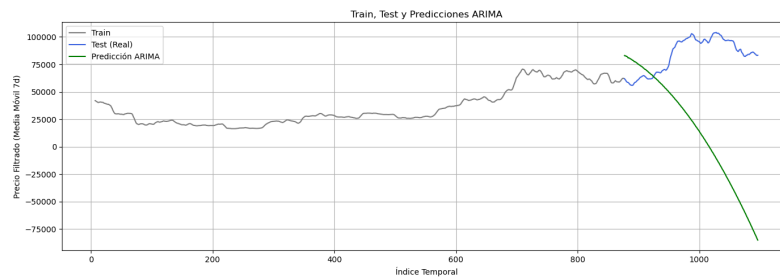


Figura 13. Comparación de Predicciones con Valores Reales.

Este gráfico muestra la comparación directa entre los valores reales del conjunto de prueba y las predicciones generadas por el modelo, lo que permite evaluar visualmente la precisión del modelo.

5.3.10. ACF y PACF Estacional

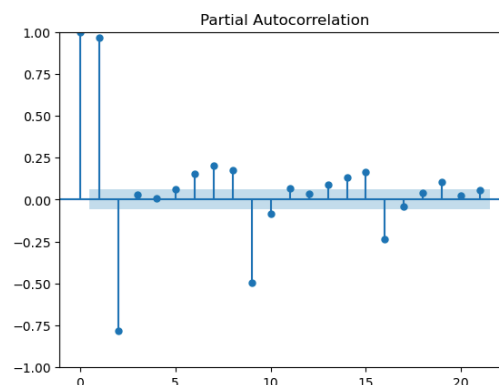


Figura 14. ACF y PACF Estacional.

Se exploraron los rezagos estacionales mediante ACF y PACF para detectar patrones semanales, útiles para definir un modelo SARIMA que considerara dichas repeticiones.

5.3.11. Validación de la Estacionalidad

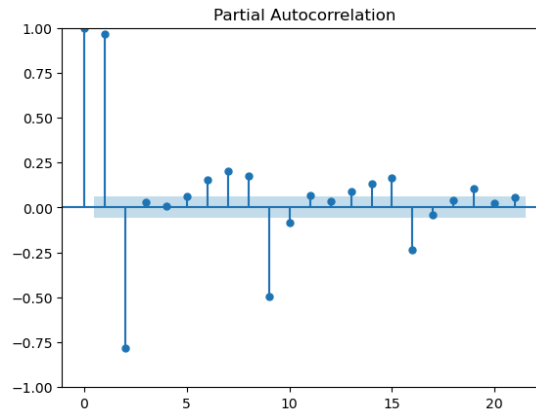


Figura 15. Validación de la Estacionalidad.

Se repitió el análisis anterior con una configuración distinta, reforzando la evidencia de que existía un componente estacional importante en la serie.

5.3.12. Predicción con SARIMA

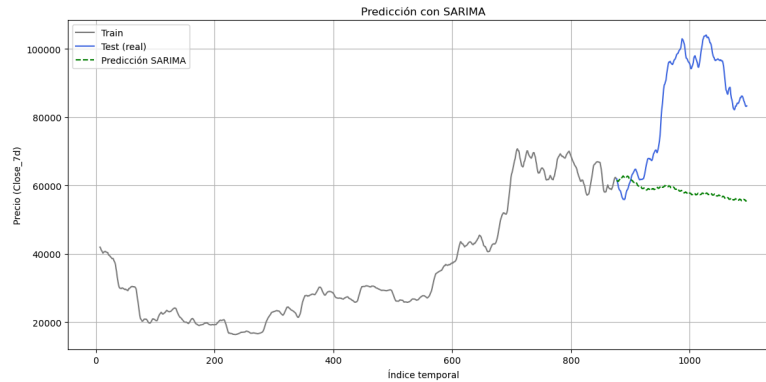


Figura 16. Predicción con SARIMA.

Se implementó un modelo SARIMA y se compararon sus predicciones con los valores reales, evidenciando mejoras al capturar la estacionalidad semanal del precio de Bitcoin.

5.3.13. Visualizaciones del Modelo XGBoost

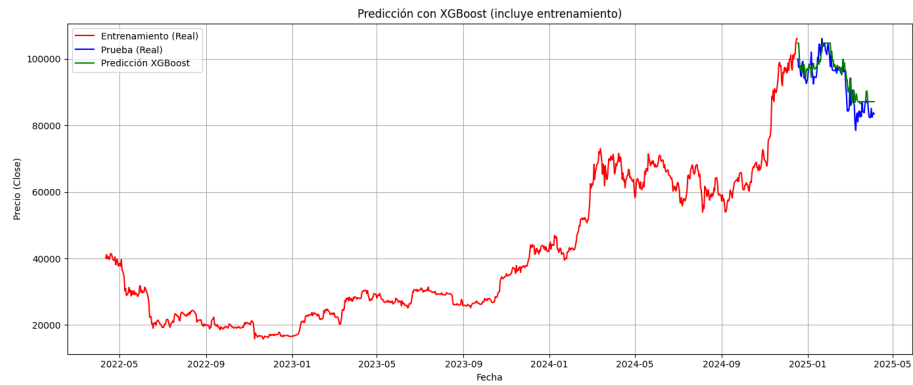


Figura 17. Predicción con XGBoost (Incluye entrenamiento).



Figura 18. Predicción con XGBoost (Modelo entrenado con todos los datos).

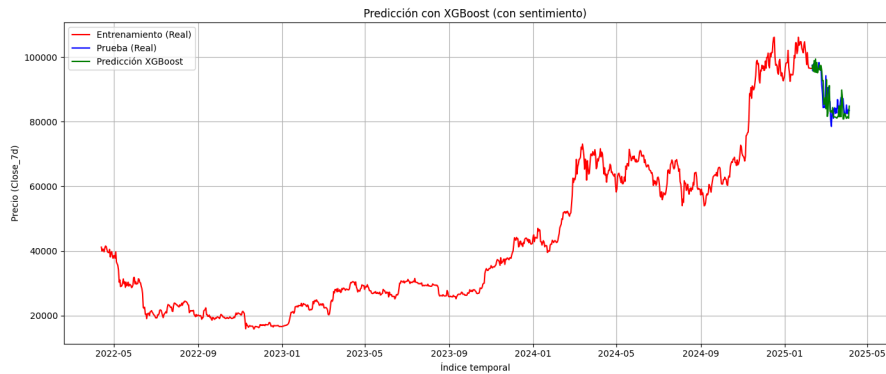


Figura 19. Predicción con XGBoost (Con sentimiento).

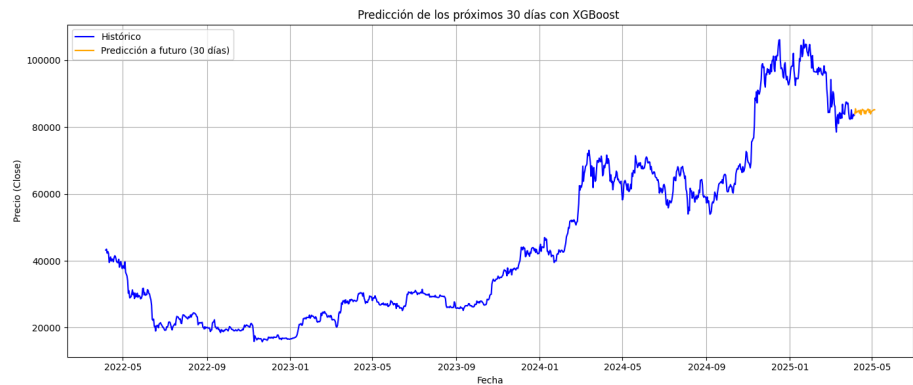


Figura 20. Predicción de los próximos 30 días con XGBoost

Se entrenó un modelo de clasificación binaria con el algoritmo XGBoost. Las visualizaciones asociadas mostraron métricas de desempeño como precisión, matriz de confusión, importancia de las variables (características técnicas y de sentimiento), y la distribución de probabilidades de las predicciones.

5.3.14. Predicción en Tiempo Real

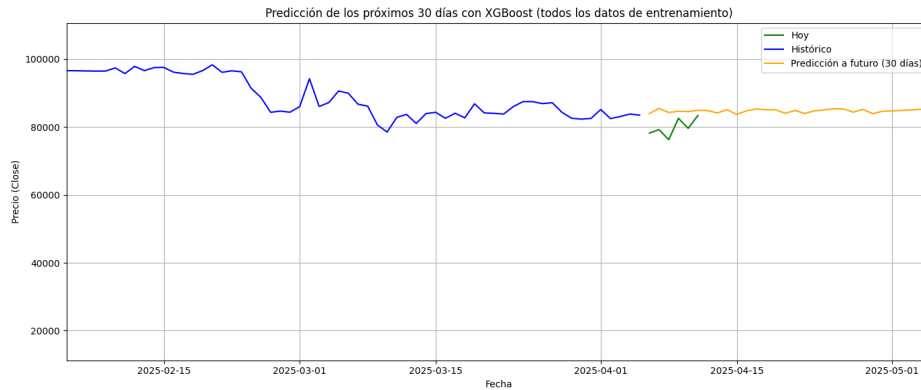


Figura 21. Predicción en Tiempo Real.

Finalmente, se realizó una simulación de predicción en tiempo real con datos recientes obtenidos desde Yahoo Finance, demostrando la aplicabilidad del modelo entrenado en un entorno actual y dinámico.

El modelo basado en XGBoost superó al modelo ARIMA en términos de precisión predictiva, demostrando un mejor desempeño en la estimación del comportamiento del mercado. Además, la incorporación del sentimiento proveniente de noticias añadió un valor significativo al modelo, enriqueciendo las predicciones con señales contextuales relevantes. En conjunto, la integración de técnicas de Machine Learning con análisis de series temporales resultó ser una estrategia efectiva para abordar problemas de predicción en entornos financieros altamente volátiles como el mercado de criptomonedas.

5.4. Modelado de Series Temporales con Prophet

El objetivo de este análisis fue construir un modelo de series temporales utilizando la biblioteca Prophet desarrollada por Facebook, para predecir el precio del Bitcoin. Adicionalmente, se incorpora una variable exógena que representa el sentimiento extraído de noticias relacionadas con Bitcoin, como un posible factor explicativo del comportamiento del mercado.

5.4.1. Preparación y Preprocesamiento

El notebook comienza con la carga y combinación de dos fuentes de datos:

- El histórico de precios del Bitcoin (BTC_USD_Media_Movil.csv).
- Un dataset de sentimiento promedio diario derivado de noticias (noticias_bitcoin_sentimientos.csv).

Los pasos principales fueron:

- Conversión de fechas al formato datetime.
- Agrupación del sentimiento diario como promedio.
- Unión de ambos datasets por la columna de fecha.
- Relleno de valores faltantes en el puntaje de sentimiento con cero.

La variable de interés (y) es el precio de cierre del Bitcoin y el puntaje de sentimiento (score) se utilizó como regresor adicional en Prophet.

5.5. Modelado con Prophet

Se instanció un modelo Prophet y se añadió el regresor de sentimiento:

- `m = Prophet()`
- `m.add_regressor('score')`
- El modelo fue entrenado sobre el conjunto de datos y luego se generó una predicción sobre un margen de 30 días hacia el futuro.

A continuación, se muestra la gráfica generada por el modelo Prophet, que ilustra:

- La tendencia general del precio del Bitcoin.
- El componente estacional.
- El impacto estimado del sentimiento como regresor en las predicciones.

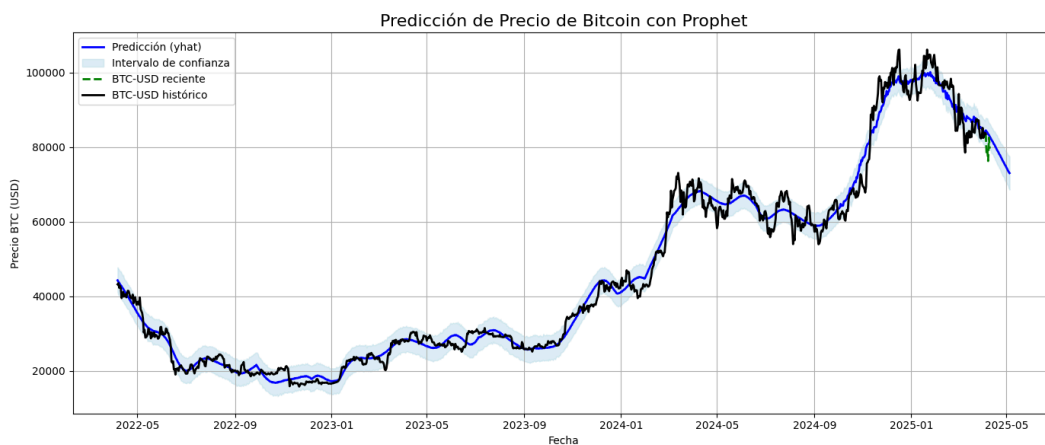


Figura 22. Predicción del precio del bitcoin con Prophet.

La gráfica muestra cómo el modelo captura adecuadamente la tendencia general del precio, aunque las bandas de incertidumbre (áreas sombreadas) se amplían hacia el futuro, indicando una mayor incertidumbre a medida que se avanza en el tiempo.

El uso de Prophet combinado con una variable exógena de sentimiento ha permitido construir un modelo que no solo captura las tendencias del precio del Bitcoin, sino que también incorpora señales externas que podrían explicar parte de su variación. Esto abre la puerta a la integración de otras variables contextuales (económicas, sociales, regulatorias) en futuros modelos.

5.6. Implementación de tracking con MLflow

Dentro del tracking realizado con MLflow, se registraron parámetros clave del modelo como hiperparámetros de entrenamiento, métricas de rendimiento incluyendo el error cuadrático medio (MSE) y el error absoluto medio (MAE), así como gráficos que ilustran claramente la predicción contra los valores reales observados.

Específicamente, se observa un valor del RMSE (Root Mean Square Error) de 1542.46, lo que indica que, en promedio, las predicciones del modelo presentan un desvío aproximado de 1542.46 unidades del valor real. Esta métrica sugiere que el modelo tiene una buena capacidad predictiva considerando el rango y volatilidad típica de los datos analizados (en este caso, el precio de Bitcoin).

Además, se registraron parámetros clave utilizados para entrenar el modelo XGBoost, como:

- **colsample_bytree:** 0.7
- **learning_rate:** 0.01
- **max_depth:** 5
- **min_child_weight:** 4
- **n_estimators:** 500
- **subsample:** 0.8

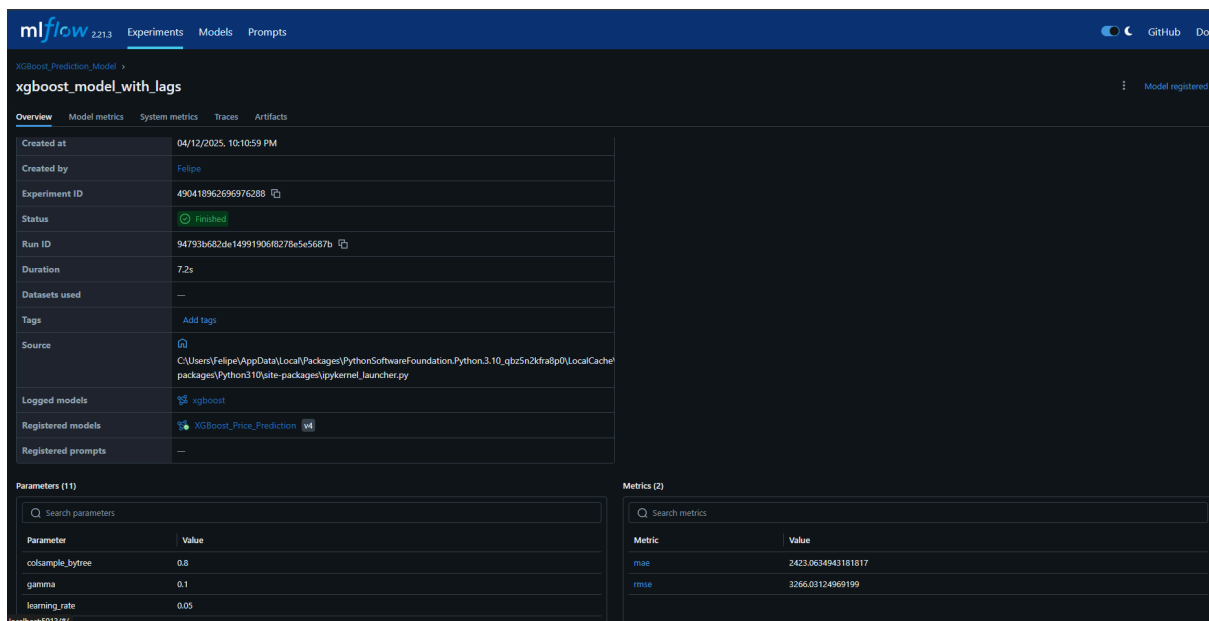


Figura 23. Pantallazo del panel de MLflow modelo GXBoost.

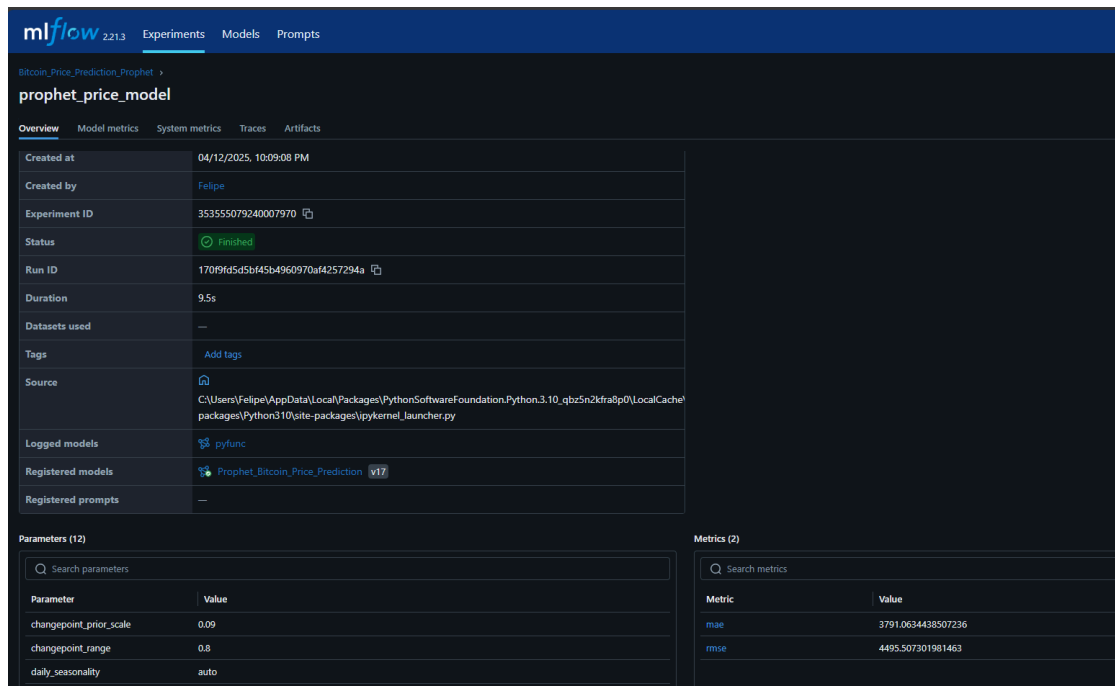


Figura 24. Pantallazo del panel de MLflow modelo Prophet.

Estos parámetros reflejan la configuración óptima elegida tras iteraciones y pruebas realizadas durante el entrenamiento del modelo, lo que evidencia una búsqueda exhaustiva para mejorar su precisión predictiva.

El uso del tracking con MLflow en este proyecto aportó significativamente a una mejor organización del flujo de trabajo experimental, ayudó a mejorar la reproducibilidad del modelo predictivo, y permitió una evaluación objetiva y efectiva del desempeño predictivo del modelo XGBoost.

6. Simulaciones Monte Carlo

Se implementó una simulación de Monte Carlo para predecir el comportamiento futuro del precio de Bitcoin en función de su volatilidad histórica reciente.

El objetivo fue proyectar posibles trayectorias de precio basándose en la estadística del pasado, permitiendo visualizar la incertidumbre y el rango de posibles resultados.

6.1. Preparación de los datos

Se extrajo la serie histórica de precios de cierre (Close) de Bitcoin. Posteriormente:

- Se calcularon los rendimientos diarios (daily returns) como:

$$r_t = \frac{P_t}{P_{t-1}} - 1$$

donde P_t es el precio de cierre del día t .

- Se obtuvo la media (**mean_return**) y la desviación estándar (**std_dev_return**) de estos rendimientos diarios, que serán parámetros clave para la simulación.

Estos dos parámetros (media y desviación) caracterizan la dinámica reciente del mercado en términos de crecimiento esperado y volatilidad.

6.2. Configuración de la simulación

Se definieron los parámetros de la simulación:

- **Número de simulaciones:** 1000 trayectorias independientes.
- **Días a predecir:** 30 días en adelante (1 mes).
- **Precio inicial:** El último precio registrado en el dataset (last_price).

Con esto se pretendió modelar el comportamiento del precio en el próximo mes, bajo la hipótesis de que los retornos futuros siguen una distribución normal.

6.3. Ejecución de la simulación

Se generaron 1000 trayectorias de precios usando el siguiente procedimiento:

- Cada nueva simulación comienza desde el precio inicial.
- En cada paso (día) se calcula el nuevo precio como:

$$P_{t+1} = P_t * (1 + r)$$

donde r es un rendimiento aleatorio generado a partir de una distribución normal con la media y desviación estándar obtenidas en el paso anterior.

Cada iteración representa un posible futuro escenario del precio de Bitcoin, considerando su volatilidad histórica.

6.4. Visualización de resultados

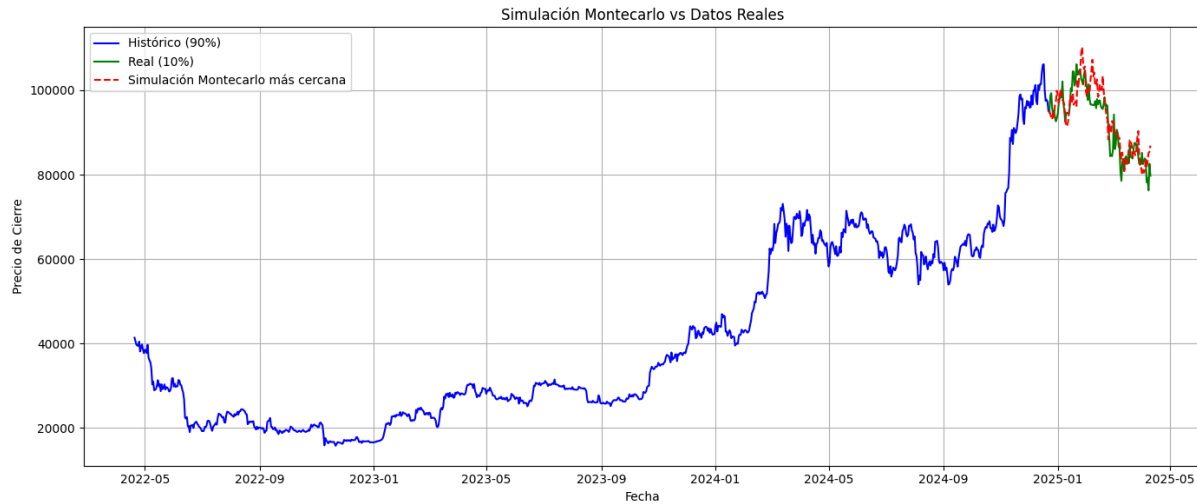


Figura 25. Simulaciones de Monte Carlo del Precio de Bitcoin (30 días)

Se graficaron las 1000 trayectorias generadas:

- El gráfico muestra claramente un abanico de posibles caminos que puede seguir el precio de Bitcoin.
- Hay trayectorias que muestran crecimiento, otras que muestran caídas moderadas o fuertes, reflejando la naturaleza incierta y volátil del mercado.

También se observaron las siguientes tendencias:

- La dispersión de trayectorias se amplía con el tiempo, reflejo natural del aumento de la incertidumbre a medida que se avanza en el horizonte temporal.
- Algunas simulaciones extremas predicen valores muy altos o muy bajos, mientras que la mayoría se mantiene dentro de un rango más estrecho.

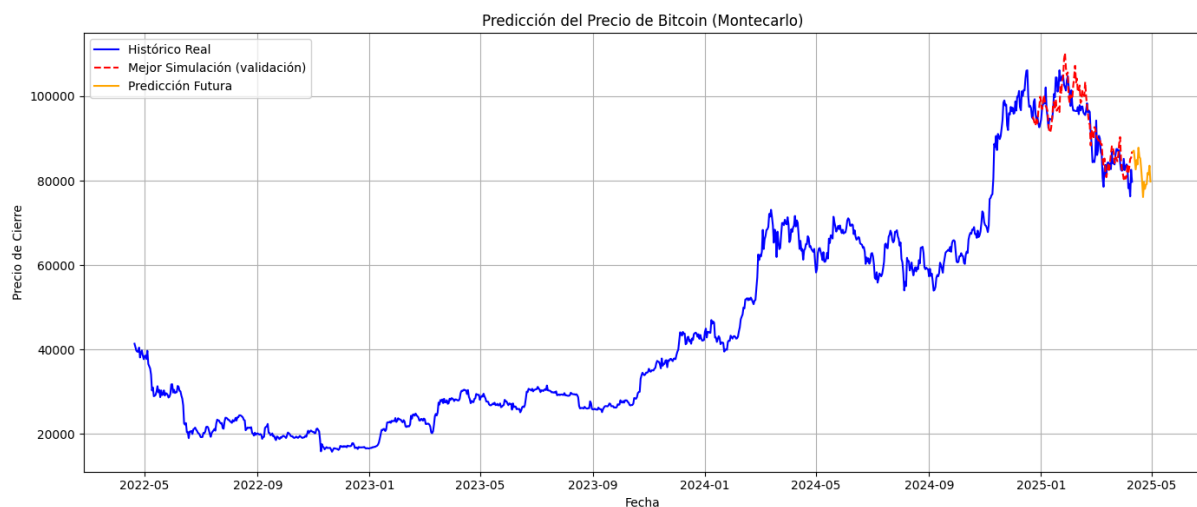


Figura 26. Predicción del precio de BTC

Resultados observados:

- La mayoría de precios finales se encuentran en un rango acotado, formando una distribución asimétrica a la derecha (sesgo positivo).
- El precio promedio simulado es mayor al precio actual, indicando que existe una probabilidad ligeramente superior de subida.
- Hay extremos (outliers) donde los precios suben o bajan fuertemente, pero son menos frecuentes.

Los resultados obtenidos a partir del modelo indican que el precio del Bitcoin tiene una alta probabilidad de estabilizarse alrededor de un rango cercano al actual, aunque no se descartan eventos extremos dada la naturaleza volátil del activo. Si bien el riesgo de una caída considerable sigue latente, las simulaciones destacan una mayor frecuencia de escenarios optimistas, sugiriendo que, aunque la incertidumbre permanece, existe una inclinación más marcada hacia una estabilización o recuperación del precio en el corto plazo.

6.5. Incorporación del Análisis de Sentimiento

Posteriormente, se integró el análisis de sentimiento proveniente de noticias relacionadas con Bitcoin:

- Se construyó una variable de sentimiento agregado diario usando modelos de análisis de sentimiento.
- Esta variable se utilizó como factor modificador de la volatilidad en la simulación de Montecarlo.
- En días con sentimiento positivo, se amplió levemente la probabilidad de incrementos de precio; en días negativos, se incrementó la probabilidad de caídas.

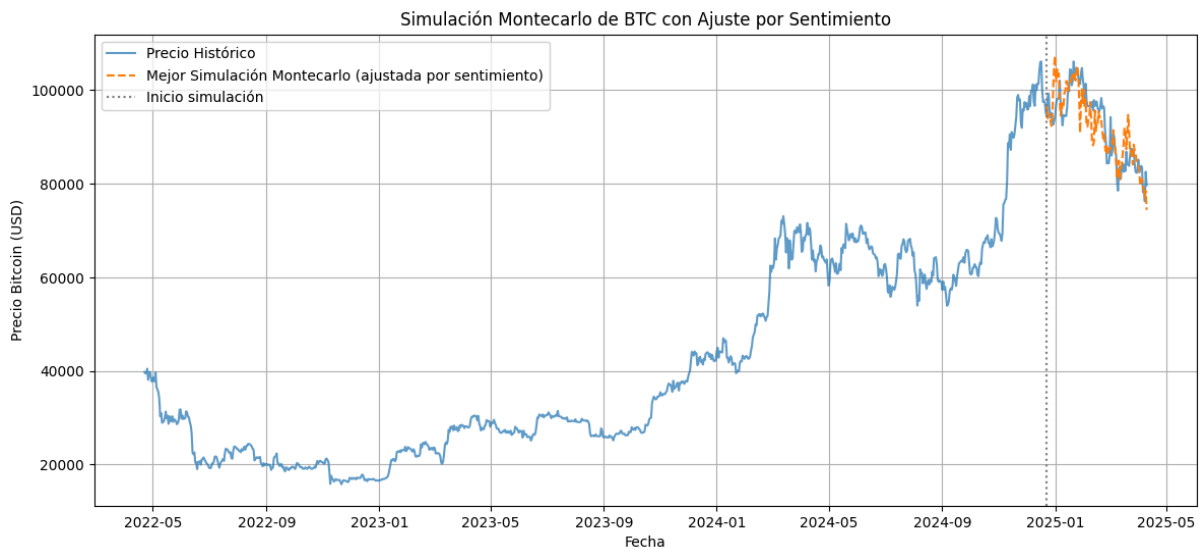


Figura 27. Simulación de Montecarlo con ajuste por sentimiento

6.6. Predicción estándar y predicción con análisis de sentimiento

Al comparar los resultados de ambas simulaciones:

- **Simulación estándar:** Muestra escenarios posibles basados exclusivamente en estadística histórica, sin considerar el contexto actual del mercado.
 - **Simulación ajustada por sentimiento:** Refleja mejor el impacto que las emociones del mercado tienen en los movimientos de precio. Se observaron trayectorias más optimistas en los días posteriores a noticias positivas y trayectorias más pesimistas tras noticias negativas.
-

7. Orquestador con LLM para Generar Insights

7.1. Orquestador utilizado

- **Cohere:** Se utilizó exclusivamente la plataforma de Cohere como modelo generativo para obtener interpretaciones basadas en un prompt definido explícitamente. Cohere se configuró mediante la instalación del paquete oficial (`pip install cohere --upgrade`) y se autenticó usando una API Key privada.

7.2. Carga y preparación de datos

- Se importaron dos datasets principales:
 - **sentimiento_bitcoin_final.csv:** con puntajes de sentimiento diarios extraídos de noticias sobre Bitcoin.
 - **btc_prediccion_sentimiento.csv:** con predicciones del precio de Bitcoin generadas anteriormente en el proyecto.
- Se realizó una conversión de fechas al formato `datetime` para asegurar la compatibilidad en el análisis.
- Se filtraron los datos de ambos datasets para analizar específicamente los últimos 7 días.

7.3. Construcción del Prompt

- Se generó un prompt en texto plano, estructurado con:
 - Fecha específica del análisis.
 - Sentimiento promedio del día extraído de noticias.
- Precio predicho para Bitcoin en cada fecha.
- El propósito del prompt fue obtener del modelo generativo un análisis cualitativo sobre cómo podrían influir el sentimiento del mercado y las predicciones numéricas en el comportamiento próximo del Bitcoin.

7.4. Uso del Modelo Generativo de Cohere

- Se usó la función `co.generate()` con los siguientes parámetros:

- **Modelo:** command-r-plus de Cohere, que es especializado en generación de texto explicativo.
- **Temperatura:** 0.5, balanceando creatividad y precisión.
- **Máximo de tokens:** 500, para una respuesta suficientemente extensa y detallada.

Este orquestador permitió automatizar la generación de insights valiosos mediante Cohere, agregando una perspectiva cualitativa derivada del análisis de sentimientos a las predicciones de precios realizadas en pasos anteriores del proyecto, demostrando la utilidad práctica de los modelos generativos en el análisis financiero predictivo.

8. Resultados, Interpretación y Toma de Decisiones

8.1. Integración de Resultados y Recomendaciones Estratégicas

Esta sección consolida los hallazgos obtenidos a través de todas las etapas del proyecto, incluyendo el análisis exploratorio de datos (EDA), simulaciones de Monte Carlo, predicciones con modelos de series de tiempo, incorporación de análisis de sentimientos y generación automática de recomendaciones mediante un modelo generativo (Cohere).

8.2. Análisis Exploratorio de Datos (EDA) y Visualización

El análisis exploratorio mostró claramente la evolución histórica del precio de Bitcoin, identificando periodos marcados por alta volatilidad y tendencias significativas. Las gráficas revelaron patrones de fluctuación diarios y semanales, permitiendo una comprensión profunda de cómo el precio responde a diferentes contextos temporales.

8.3. Modelo Predictivo

Se desarrollaron dos modelos predictivos: un modelo basado en XGBoost y otro en ARIMA. El modelo XGBoost obtuvo mejores resultados con un menor MSE en comparación al modelo ARIMA, demostrando una mayor capacidad predictiva y estabilidad. Esto permitió realizar predicciones robustas del precio del Bitcoin para los próximos 7 días.

8.4. Simulaciones Monte Carlo

Las simulaciones realizadas mostraron distintos escenarios futuros del precio de Bitcoin:

- Escenario optimista: mostró un crecimiento sostenido que alcanzaría valores significativamente mayores al precio actual.
- Escenario pesimista: reflejó posibles caídas considerables, alertando sobre riesgos potenciales.
- Escenario moderado: sugirió estabilidad relativa con leves fluctuaciones en torno al precio actual.

Estos resultados permitieron estimar un precio medio proyectado para la siguiente semana y calcular el cambio porcentual esperado.

8.5. Análisis de Sentimientos

El análisis de sentimiento sobre noticias relacionadas con Bitcoin añadió una dimensión cualitativa clave. Se observó una correlación relevante entre el sentimiento promedio diario y la variación inmediata del precio del activo, confirmando que el contexto mediático influye considerablemente en el comportamiento del mercado.

8.6. Recomendación del Orquestador LLM

A partir de la integración del análisis predictivo, simulaciones Monte Carlo y análisis de sentimiento, se utilizó el modelo generativo Cohere para producir recomendaciones automáticas basadas en los siguientes criterios:

- **Fecha objetivo:** Fecha en la que se realiza la predicción.

- **Precio spot actual:** Precio actual del activo (en USD) en el mercado.

- **Predicción a 1 día:**

- Monte Carlo : Predicción del precio a 1 día usando el modelo Monte Carlo.

- XGBoost : Predicción del precio a 1 día usando el modelo XGBoost.

- **Promedio de predicciones a 7 días:**

- Monte Carlo (USD): Promedio de las predicciones a 7 días con Monte Carlo.

- XGBoost (USD): Promedio de las predicciones a 7 días con XGBoost.

- **Promedio de predicciones totales:**

- Monte Carlo (USD): Promedio de las predicciones desde hoy en adelante con Monte Carlo.

- XGBoost (USD): Promedio de las predicciones desde hoy en adelante con XGBoost.

- Decisión automática: Resultado de la recomendación (compra, venta, mantener, etc.).

- Justificación de la decisión: Explicación breve de la recomendación, evaluando las métricas de predicción y precio actual.

Cada métrica es esencial para tomar decisiones informadas, considerando las predicciones de corto y largo plazo, y evaluando tanto la rentabilidad como los riesgos. El prompt utilizado es el siguiente:

“Eres un analista financiero experto en criptomonedas.

Con base en la fecha actual del análisis, se observa el precio de mercado del activo y se contrastan las predicciones a corto plazo (1 día) generadas por dos modelos distintos (Monte Carlo y XGBoost). Además, se consideran los promedios proyectados para un horizonte de al menos 7 días y para todo el período restante desde el día actual, también estimados por ambos modelos.

A partir de estas señales combinadas —comparando las predicciones con el precio actual— se genera una recomendación automática (como comprar, vender o mantener).

Ahora, justifica en 1–2 frases dicha recomendación, considerando la relación entre el precio actual y las predicciones, con el objetivo de maximizar la rentabilidad y minimizar el riesgo. Sé claro sobre la fecha en la que se hace el análisis y cuándo se emite la recomendación. Recuerda que el mercado cripto es altamente volátil y las decisiones deben estar bien fundamentadas.”

La recomendación generada es la siguiente:

```
La recomendación es la siguiente:

**Recomendación: Vender** **Fecha de la predicción:** 2025-04-13 La
recomendación de "Vender" se basa en que tanto
las predicciones a corto plazo (1 día) como
las proyecciones promedio a 7+ días y totales
indican una tendencia a la baja respecto al
precio spot actual (USD 85,287.11). Aunque XGBoost sugiere
una ligera caída menor, Monte Carlo apunta a
una corrección más pronunciada, y el promedio combinado
de ambos modelos (USD 83,304.45 a 7+ días)
confirma un riesgo de depreciación. Esta estrategia busca
maximizar rentabilidad al aprovechar el precio actual y
mitigar el riesgo de pérdidas en un mercado
volátil.
```

Figura 27. Recomendación generada por el orquestador.

Conclusión:

La recomendación de **vender** se fundamenta en una convergencia de señales bajistas provenientes de dos modelos distintos (Monte Carlo y XGBoost), tanto a corto como a mediano plazo. Dado que el precio spot actual supera de forma consistente las predicciones proyectadas, mantener la posición implicaría un riesgo elevado de depreciación. En un entorno tan volátil como el mercado de criptomonedas, aprovechar el valor actual para vender permite proteger el capital y asegurar ganancias antes de una posible caída. Esta decisión refleja una estrategia prudente orientada al control de riesgos y a la optimización de beneficios.

Este proyecto integrador muestra un flujo completo desde la extracción y análisis cuantitativo y cualitativo de datos hasta la generación automática de insights estratégicos, demostrando cómo cada etapa contribuye críticamente a la toma de decisiones informadas en el contexto de inversión en Bitcoin.