



Universidad Tecnológica de Panamá

Modelos Predictivos

Profesor: Juan Marcos Castillo, PhD

ANÁLISIS DEL PRECIO DEL ORO CON SERIES DE TIEMPO

Autor: Jaime Acosta

Julio 2025

Introducción

El oro, considerado desde tiempos antiguos como un metal precioso de gran valor, ha desempeñado un papel fundamental en la economía global.

El presente proyecto tiene como objetivo realizar un análisis descriptivo y predictivo del precio del oro, combinando técnicas estadísticas tradicionales y modelos computacionales de aprendizaje automático. En particular, se emplearán modelos clásicos y modernos de series de tiempo, como ARIMA, Holt-Winters y Prophet, así como enfoques basados en regresión, con el fin de entender la evolución histórica del precio y generar predicciones que puedan tener aplicación práctica en la toma de decisiones.

Para este análisis se utilizaron datos históricos del precio del oro correspondientes al período comprendido entre enero de 2014 y enero de 2025, lo que permite observar tanto tendencias de largo plazo como fluctuaciones más recientes.

Justificación

La motivación principal de este proyecto proviene del interés personal por profundizar en el análisis de series de tiempo, una de las técnicas fundamentales dentro de la caja de herramientas para crear modelos predictivos. Cabe mencionar que este trabajo representa mi primera aproximación formal a la aplicación de modelos de series de tiempo, lo que lo convierte en una experiencia clave para afianzar conocimientos tanto teóricos como prácticos en el manejo de datos temporales.

Para ello, se eligió el precio del oro como variable de estudio, dada su naturaleza dinámica y su relevancia económica global. El oro ha sido históricamente considerado un activo de resguardo frente a la incertidumbre financiera, mostrando comportamientos característicos ante eventos macroeconómicos y geopolíticos globales.

De esta manera, con este proyecto no solo se busca un objetivo académico, al aplicar y contrastar modelos de series de tiempo, sino que también busca tener una utilidad práctica al ofrecer una herramienta básica para la comprensión del comportamiento de un activo de alta importancia actual.

Antecedentes

El análisis del precio del oro ha sido objeto de numerosos estudios económicos y financieros debido a su papel como activo de refugio ante escenarios de incertidumbre.

Existen numerosos estudios e investigaciones que abordan la predicción del precio del oro a través de modelos estadísticos y de machine learning, con modelos como ARIMA y LSTM o incluso aproximaciones con modelos híbridos. Y es que en general, se ha encontrado que factores como la tasa de interés, el valor del dólar y eventos globales afectan significativamente el comportamiento del oro, por lo que su variabilidad está influenciada por muchos factores externos, sin embargo, por razones de tiempo y simplicidad dichos factores están fuera del alcance de este proyecto, nos enfocaremos solamente en el precio en sí.

Definición del problema

Comprender el comportamiento del precio del a lo largo del tiempo representa un desafío interesante desde el punto de vista analítico y predictivo.

El problema que se plantea en este proyecto es cómo modelar de forma adecuada la evolución temporal del precio del oro utilizando modelos de series de tiempo, con el objetivo de realizar predicciones con un grado razonable de precisión. Más que obtener un modelo perfectamente ajustado, el interés principal es aplicar técnicas estadísticas apropiadas para analizar cómo varía el precio con el tiempo y comprobar si los modelos aplicados permiten realizar predicciones razonables.

Es importante resaltar que, por razones de alcance y tiempo, el análisis se basa únicamente en la variable temporal del precio del oro, sin considerar otras variables externas que podrían influir en su comportamiento.

Análisis Predictivo

Determinación de la base de datos

La base de datos “**Gold Price**” contiene registros diarios del comportamiento del precio del oro en el mercado internacional. Los datos originales provienen de MCX Market. Esta base de datos fue obtenida de la plataforma Kaggle e incluye los precios diarios del oro desde enero de 2014 hasta enero de 2025.

La moneda de los datos es INR (rupias de la india). Las observaciones incluyen valores desde el año 2025 e integran las siguientes variables:

- **Date:** fecha de la observación (formato AAAA-MM-DD)
- **Price:** precio de cierre diario del oro (INR por 10 gramos de oro 24K)
- **Open:** precio de apertura diario del oro (INR por 10 gramos de oro 24K)
- **High:** precio máximo alcanzado durante el día (INR por 10 gramos de oro 24K)
- **Low:** precio mínimo alcanzado durante el día (INR por 10 gramos de oro 24K)
- **Volume:** volumen de transacciones realizadas en el día (número de contratos o unidades negociadas)
- **Chg%:** variación porcentual del precio con respecto al día anterior (%)

Preprocesamiento y limpieza

Primero se estableció la columna fecha como índice en los datos.

Luego el campo fecha se transformó de tipo object a tipo datetime, esto para que las librerías puedan reconocerla como un dato temporal.

Se estableció la frecuencia de la serie a días hábiles. Se encontraron 49 fechas con valores nulos, para estas fechas se rellenó el precio con el valor del día anterior.

El precio que está en rupias se transformó a dólares estadounidenses usando una tasa de conversión de 85.96, que es el valor aproximado de las rupias en dólares en julio de 2025.

Análisis descriptivo

Iniciamos el análisis graficando el precio del oro en nuestros datos.



De este gráfico podemos ver que la tendencia del precio del oro es al alza. Se notan algunos periodos en los que el precio parece subir muy rápidamente. A simple vista no se notan patrones claros de estacionalidad y no parece haber cambios grandes en la tendencia.

Algo que también se puede apreciar es que existen puntos a partir de los cuales los precios no vuelven a descender a estos puntos los llamaremos mínimos irreversibles.

Utilizando la función `find_peaks` podemos encontrar los puntos mínimos en la gráfica para luego obtener de esos puntos aquellos que sean mínimos irreversibles.

Al marcar cada uno de estos puntos en la serie de tiempo obtenemos este gráfico:



Estos puntos sugieren una línea, por lo que trazamos una línea de regresión lineal y polinomial de grado dos solo con los puntos de los mínimos irreversibles.



Ahora verificamos si existe alguna anomalía en el precio, para eso obtenemos la desviación estándar y buscamos aquellos puntos que estén 3 puntos por arriba o por debajo de la desviación estándar. De este análisis notamos que de acuerdo con nuestra definición no obtenemos puntos anómalos.

Descomposición temporal

La descomposición temporal es una técnica que permite separar una serie de tiempo en componentes más simples para entender mejor su comportamiento. Se divide en tres partes:

1. **Tendencia o Trend:** Muestra la dirección general de la serie a lo largo del tiempo, en nuestro caso si los precios del oro tienen a subir o a bajar.
2. **Estacionalidad o Seasonality:** Representa patrones repetitivos en intervalos regulares, estos intervalos pueden ser de cualquier cantidad, semanal, mensual, anual o cualquier otro intervalo.
3. **Residual o Noise:** Es lo que queda después de remover la tendencia y la estacionalidad; incluye variaciones impredecibles o ruido.

Compararemos tres métodos para hacer la descomposición, el aditivo, el multiplicativo y STL.

El **método aditivo** asume que cada componente contribuye de manera independiente y lineal al valor total. Se expresa como:

$$\text{Serie} = \text{Tendencia} + \text{Estacionalidad} + \text{Residual}$$

Este método es útil cuando la estacionalidad tiene aproximadamente la misma amplitud a lo largo del tiempo y cuando las variaciones son constantes, sin importar si los valores de la serie son altos o bajos

El **método multiplicativo** considera que los componentes interactúan de forma proporcional.

Se expresa como:

$$\text{Serie} = \text{Tendencia} \times \text{Estacionalidad} \times \text{Residual}$$

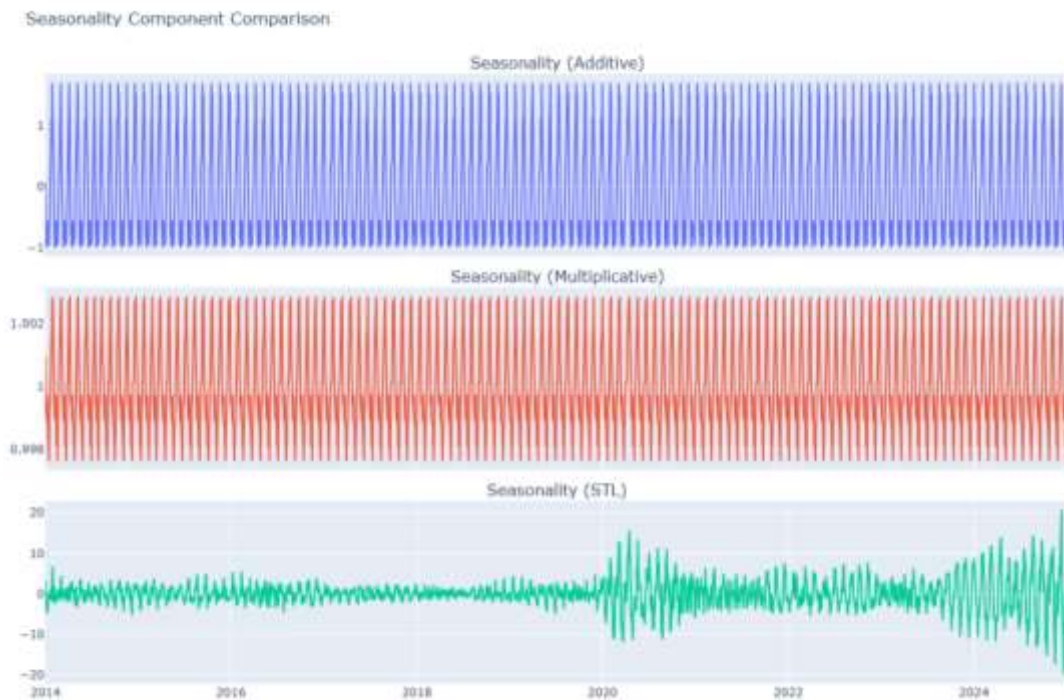
Este método resulta especialmente útil en los casos en los que La estacionalidad aumenta o disminuye proporcionalmente con el nivel de la serie o cuando las variaciones dependen del valor total.

El método **STL** (Seasonal-Trend decomposition using Loess) es una técnica más robusta que utiliza un método de suavizamiento local (LOESS). A diferencia de los métodos anteriores que aplican una misma función a toda la serie, STL toma pequeñas porciones de los datos alrededor de cada punto y hace un ajuste local. Este proceso se repite a lo largo de la serie y como resultado, le permite estimar con mayor flexibilidad los componentes de la serie temporal.

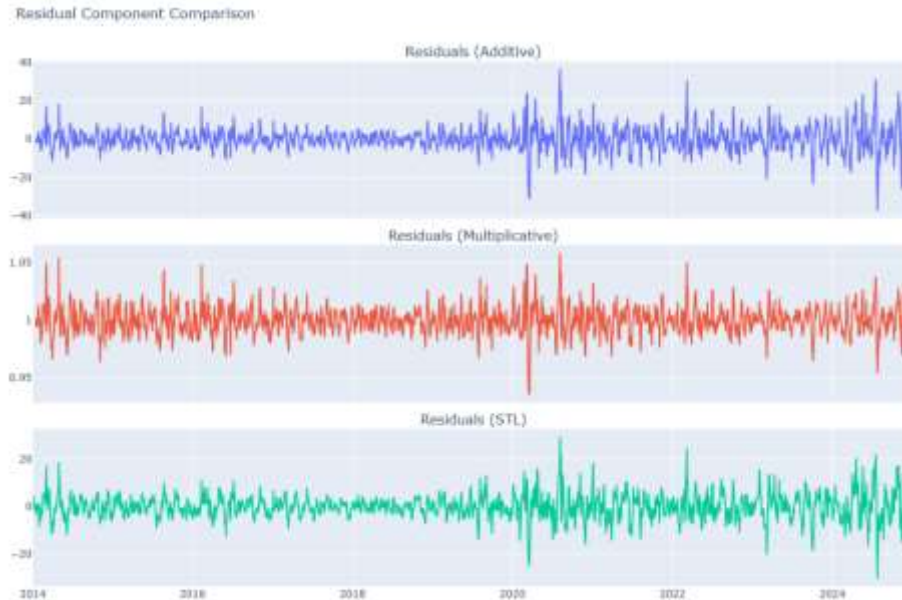
Presentamos los resultados de cada componente de la serie temporal utilizando los tres métodos anteriormente descritos.



En la tendencia vemos que los tres métodos nos dan un resultado muy similar, notamos que la tendencia es al crecimiento.



Con la estacionalidad vemos que tanto el modelo aditivo como el multiplicativo parecen mostrar alguna clase de estacionalidad, sin embargo, el método STL muestra un resultado muy diferente en donde no parece haber una estacionalidad real.



Con el residuo vemos que el resultado es prácticamente el mismo para cada uno de los modelos. La misma forma con ligeras variaciones en la amplitud en diferentes segmentos.

Pruebas estadísticas

La prueba ADF es una herramienta estadística que se utiliza para verificar si una serie de tiempo es estacionaria. En este punto es importante recordar la diferencia entre estacionariedad y estacionalidad, que, aunque son términos parecidos, son dos conceptos diferentes. Cuando decimos que una serie de tiempo es estacionaria, significa que sus propiedades estadísticas como la media, la varianza y la autocorrelación se mantienen constantes a lo largo del tiempo. Esta propiedad es importante porque modelos clásicos como ARIMA requieren que los datos sean estacionarios, si la serie no es estacionaria, normalmente se aplica una transformación para que lo sea. En cambio, la estacionalidad, como mencionamos anteriormente, es un patrón que se repite de forma regular en el tiempo.

La prueba ADF busca determinar si la serie es o no estacionaria.

- Hipótesis nula (H_0): La serie no es estacionaria.
- Hipótesis alternativa (H_1): La serie es estacionaria.

Con la prueba ADF obtenemos un p-valor de 0.0765, dado que es mayor a 0.05 no podemos rechazar la hipótesis nula de que la serie no es estacionaria.

Como la serie es no estacionaria hacemos una diferenciación, luego de lo cual aplicamos la prueba ADF y tenemos un p-valor de 0.0 logrando que la serie sea estacionaria.

Ahora aplicaremos las funciones de autocorrelación ACF y PACF.

La función de autocorrelación (ACF) mide cuánto se relaciona un valor de la serie con sus valores anteriores, es decir sus lags.

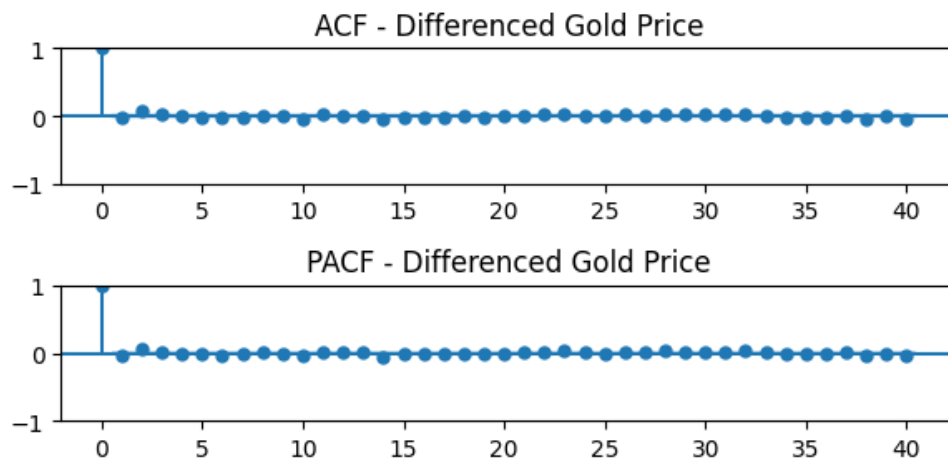
Si hay una alta autocorrelación en los primeros lags, significa que los valores recientes tienen influencia directa sobre los actuales.

Un patrón que disminuye gradualmente puede indicar presencia de tendencia o estacionalidad.

Con ella identificamos la parte MA (Media Móvil) en modelos como ARIMA.

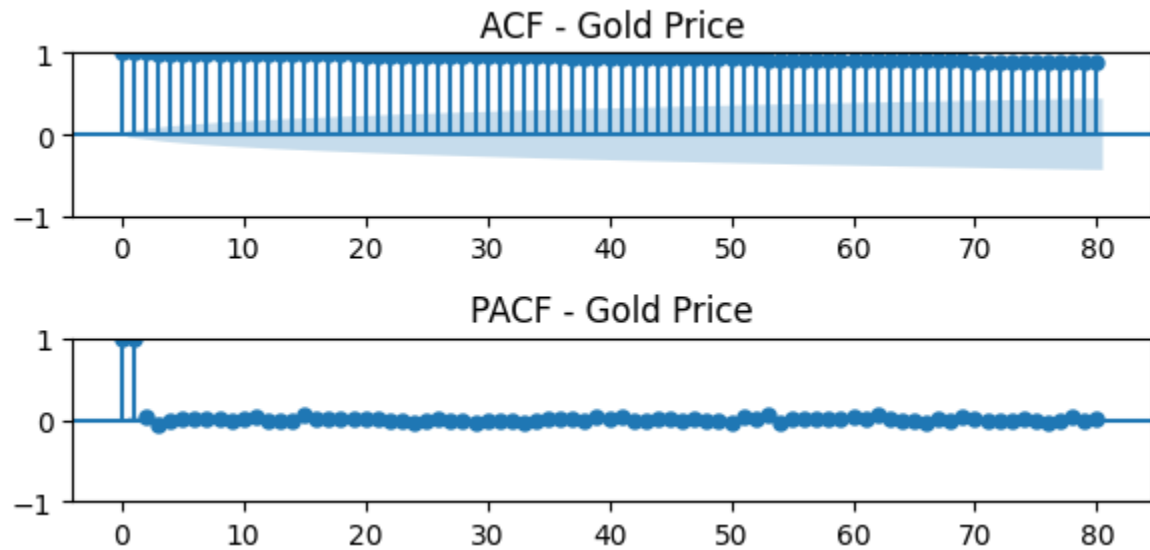
La PACF mide la correlación entre el valor actual y un valor pasado, pero eliminando la influencia de los valores intermedios.

Nos ayuda a identificar la parte AR (AutoRegresiva) del modelo.



Del gráfico de ACF y PACF podemos observar que luego del lag 0, los valores de ACF y PACF caen por debajo de 0.25, lo que sugiere que no hay correlación significativa en los lags posteriores. Esto quiere decir que en la serie solo tenemos ruido y no hay patrones significativos que podamos modelar de momento.

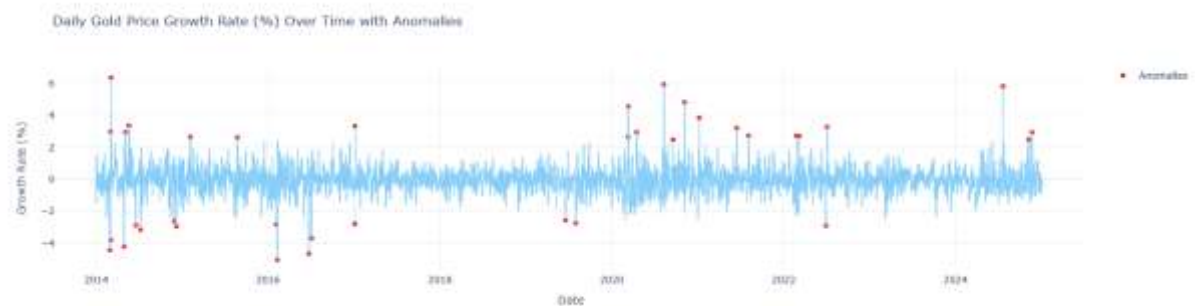
Intentaremos con la serie no diferenciada ya que la serie no diferenciada estaba dentro del intervalo de confianza del 90%.



La función de autocorrelación confirma que la serie es no estacionaria, pues la autocorrelación disminuye muy lentamente.

Cambio de precio

Se realizó también un breve análisis sobre la serie de tiempo del cambio del precio.



En este gráfico se observa la variabilidad en el precio del oro a lo largo del tiempo. Los puntos rojos representan momentos en que el precio del experimentó cambios extremos, definidos como los valores que se encuentran a más de tres veces desviaciones estándar por encima o por debajo de la media. Estos eventos reflejan periodos de alta inestabilidad, los cuales ocurren con frecuencia a lo largo de la serie.

Selección de variables

Al tratarse de modelos univariados, la única variable considerada es el precio.

Selección de Modelos

Para los modelos, compararemos los rendimientos de ARIMA, Holt-Winters y Prophet.

ARIMA (Autoregressive Integrated Moving Average)

El modelo ARIMA permite realizar pronósticos a partir de una serie de tiempo utilizando solo su comportamiento pasado. Se compone de tres elementos: una parte autorregresiva (AR), que incorpora los valores pasados; una parte de diferenciación (I), que transforma la serie en estacionaria eliminando tendencias; y una parte de promedio móvil (MA), que considera los errores pasados.

Holt-Winters (Suavizamiento exponencial estacional)

El modelo Holt-Winters es una extensión del suavizamiento exponencial simple, que es una técnica basada en la idea de que los datos más recientes son más relevantes que los antiguos. A cada valor observado se le asigna un peso, y estos pesos disminuyen exponencialmente a medida que se retrocede en el tiempo.

Además del suavizamiento exponencial, Holt-Winters incorpora componentes de tendencia y estacionalidad, lo que lo hace útil para series que presentan patrones repetitivos a lo largo del tiempo. Holt-Winters es especialmente útil cuando se busca capturar de forma sencilla tanto la evolución general de la serie como sus fluctuaciones periódicas.

Prophet

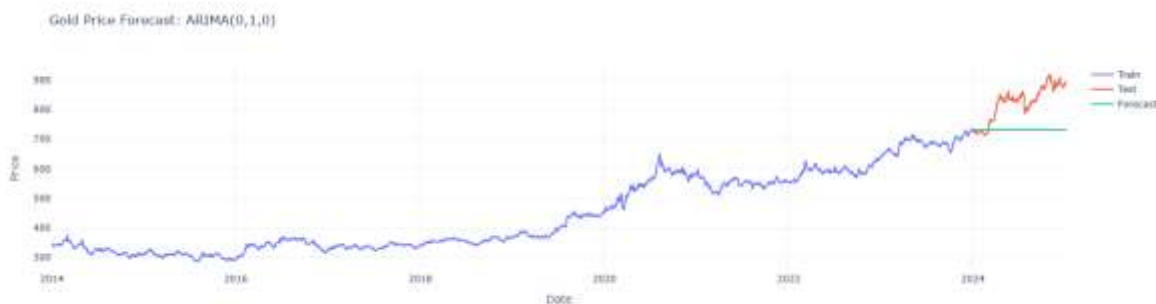
Desarrollado por Facebook, Prophet es un modelo diseñado para facilitar el pronóstico de series temporales con patrones complejos de estacionalidad y eventos especiales. Se basa en una estructura aditiva que combina una tendencia no lineal, estacionalidades múltiples y la posibilidad de incluir efectos externos como días festivos. Es flexible, fácil de ajustar y especialmente útil en situaciones en donde la interpretabilidad y la personalización son importantes.

Estos fueron los resultados de los modelos:

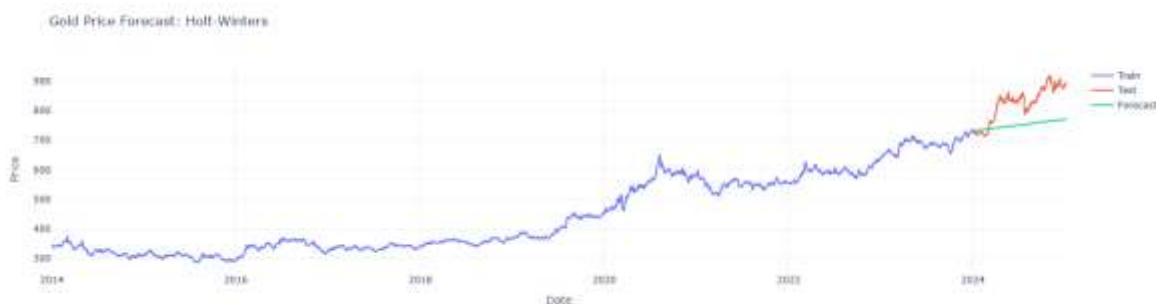
	<i>MAE</i>	<i>MSE</i>	<i>RMSE</i>	<i>MAPE</i>
<i>ARIMA</i>	95.03	11791.24	108.59	11.12
<i>Holt-Winters</i>	76.48	7539.97	86.83	8.97
<i>Prophet</i>	49.65	3387.99	58.21	6.02

Podemos notar rápidamente que el mejor modelo fue el de Prophet.

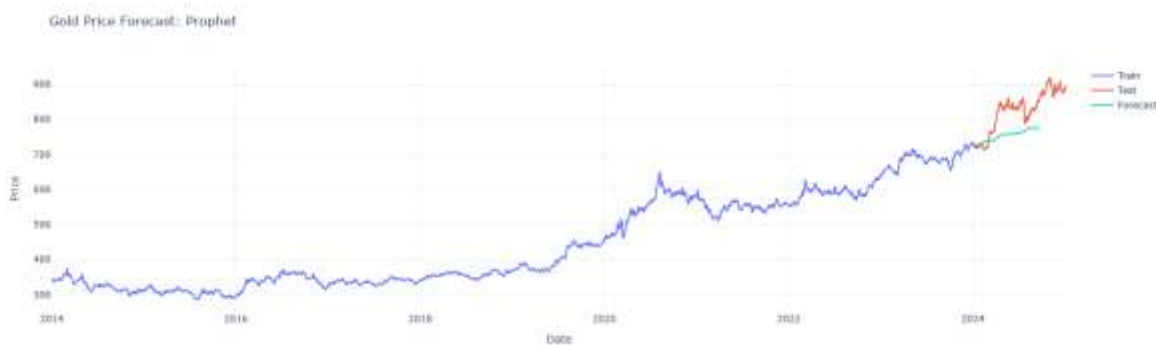
Veamos los gráficos de cada uno.



Tal como esperabamos, el modelo ARIMA no logra capturar la tendencia de la serie de tiempo. Esto se debe a que ARIMA es un modelo que requiere una serie estacionaria, y la serie de precios del oro no es estacionaria.



El modelo Holt-Winters logra capturar la tendencia de la serie de tiempo, pero no logra capturar la estacionalidad. Esto se debe a que la serie de precios del oro no presenta una estacionalidad clara o consistente.



Con Prophet podemos ver que modelo sigue la tendencia general bien, pero tiene dificultades para capturar los movimientos extremos, cuando el precio es más volátil.

Además de estos modelos, se trabajaron modelos de regresión para los puntos de los mínimos irreversibles, considerando que su tendencia parecía la de una curva suave.

Para esta sección se utilizó un modelo de regresión lineal, regresión polinomial de grado 2 y una regresión polinomial de grado dos con regularización L2 (Ridge).

Estos fueron los resultados:

Modelo	R^2	MAE	MSE	MAPE
Linear Regression	-1.7017	107.36	14461.60	13.37
Polynomial Regression (deg=2)	0.4573	38.90	2905.08	4.64
Polynomial Ridge (deg=2)	0.4273	39.68	3065.36	4.73

Podemos observar que el modelo polinomial sin regularización tuvo el mejor rendimiento, para el modelo de Ridge se probó con diferentes alfas de regularización, pero no superaron al modelo polinomial sin regularización.



Podemos ver que prácticamente no hay diferencia entre los modelos polinomiales con y sin regularización. A pesar de ser superior al modelo lineal, con un R^2 de 0.45 el este modelo no sería práctico de usar en escenarios reales.

Conclusiones

A lo largo del proyecto se logró el objetivo de adquirir conocimientos sobre la aplicación de modelos de series de tiempo y los análisis que estos conllevan como pruebas de estacionariedad y autocorrelación, también se pudo entender con mayor profundidad valor del análisis visual detallado.

Se aplicaron modelos de series de tiempo y se evaluaron sus resultados.

Los modelos ARIMA y Holt-Winters no resultaron tan buenos prediciendo el precio, debido a que la serie no era estacionaria y que no había una estacionalidad clara. Por otro lado, el modelo de Prophet logró capturar la tendencia, pero falló al capturar la volatilidad del precio, esto es de esperarse dado que la volatilidad se ve fuertemente influenciada por factores externos.

Por otro parte, los mínimos irreversibles presentan un caso de estudio interesante con posibles aplicaciones en decisiones financieras.

Recomendaciones y futuros estudios

Para trabajos futuros deben agregarse variables externas que influyen en el precio como tasas de interés, tipo de cambio, o índices de volatilidad.

Utilizar técnicas de feature engineering con el fin de aprovechar al máximo las nuevas variables incorporadas al modelo.

Luego de agregar estas variables externas a modelos que puedan procesarlos, como Prophet o ARIMAX, o modelos de aprendizaje profundo como LSTM o modelos híbridos.

Bibliografía

Kakkar, S., & Chitrao, P. V. (2023). *Forecasting of gold prices using ARIMA model*. European Economic Letters, 13(3). <https://eelet.org.uk/index.php/journal/article/view/417>

Rashidi, M., & Modarres, M. (2025). *Predicting the price of gold in the financial markets using hybrid models*. arXiv. <https://arxiv.org/pdf/2505.01402>

U, J., Lu, P. Y., Kim, C. S., Ryu, U., & Pak, K. S. (2020). *A new LSTM based reversal point prediction method using upward/downward reversal point feature sets*. Chaos, Solitons & Fractals, 132, 109559. <https://doi.org/10.1016/j.chaos.2019.109559>

Anexo

Dataset original: [Daily Gold Price \(2015-2025\) Time Series](#)

Repositorio: [jariel17/time-series-analysis-gold](#)