



ANÁLISIS Y PREDICCIÓN DE LOS PRECIOS DEL BITCOIN

Gómez Ávila A.S, Arriaga Gómez J.J, Eleazar Escobedo H, Martínez Ávila D.J



INTRODUCCIÓN

El prominente mercado de las criptomonedas, caracterizado por un alto nivel especulativo y de gran volatilidad, plantea un novedoso y desafiante escenario para la aplicación de métodos de pronósticos sobre series temporales. Los pronósticos se realizan con el fin de orientar las decisiones en muchas áreas del mundo como los mercados, el mercado de las criptomonedas no es la excepción.

En este contexto, Bitcoin se destaca por abarcar la mayor parte de la capitalización total del mercado, así como del volumen total de transacciones diarias, por lo que esta criptomoneda resulta muy atractiva para los inversionistas que buscan activos de alto rendimiento.

Debido a esto, nos pareció interesante conocer que tan rentable será el Bitcoin en el 2021.



Figura 1. Bitcoin, moneda digital descentralizada

OBJETIVO PRINCIPAL

Predecir el precio del bitcoin mediante un modelo de series de tiempo.

OBJETIVOS SECUNDARIOS

- Identificar tendencias del precio del bitcoin.
- Analizar los precios del bitcoin en los últimos años.

METODOLOGÍA

La base de datos utilizada contaba con los precios promedio, precios de cierre, precios de apertura, volumen del bitcoin y el volumen de la unidad monetaria, para el periodo del 2012 hasta principios del 2021, para la limpieza de la base rellenamos las celdas que no tenían operaciones en distintos momentos del día y agrupamos los datos por promedio diario. No se eliminaron columnas

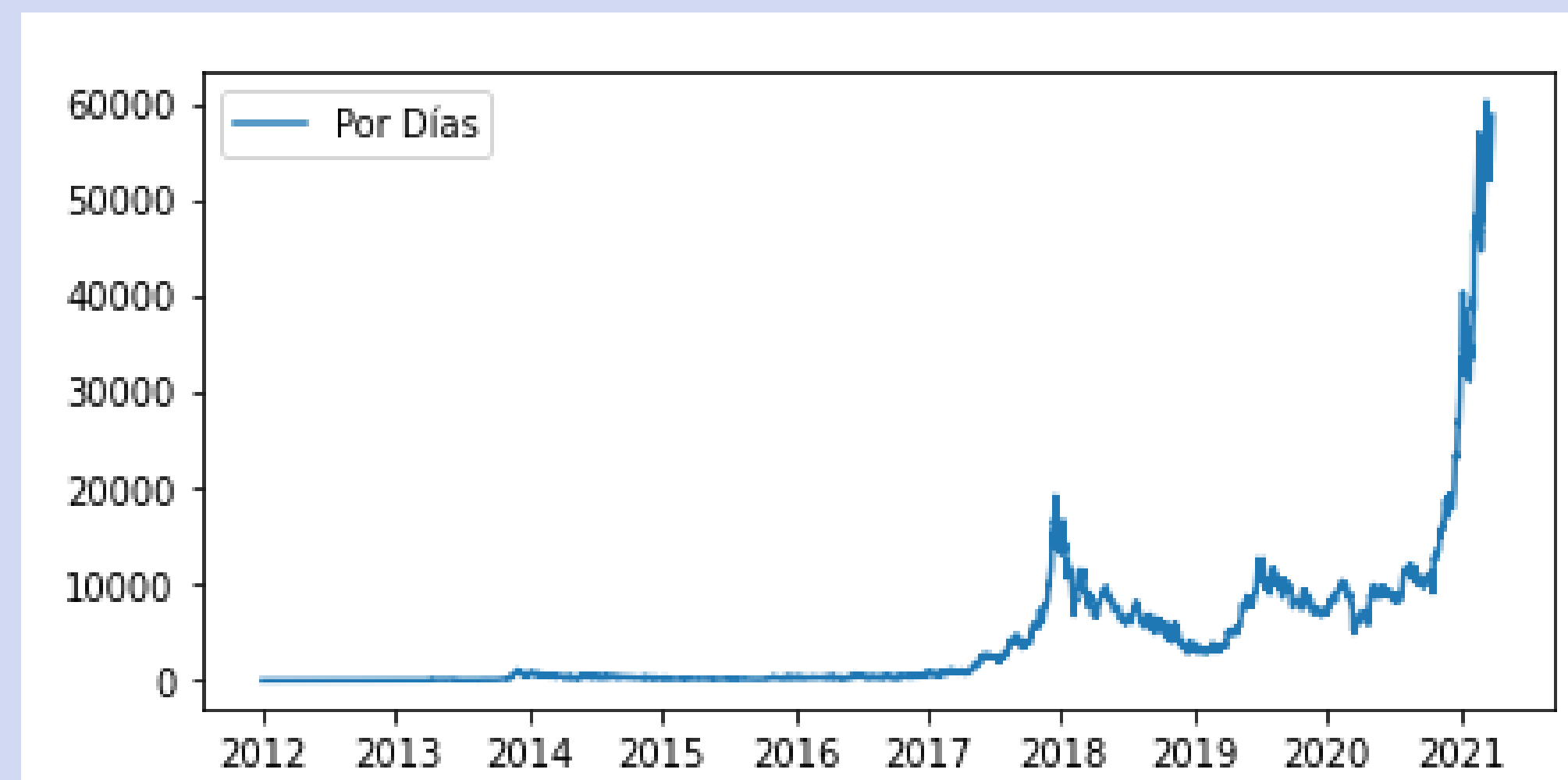


Figura 2. Precios del Bitcoin frecuencia diaria (2012-2020)

Realizando un análisis más detallado para el periodo de 2017-2020, pudimos notar que el bitcoin es muy volátil, es decir, el precio del bitcoin puede crecer a un ritmo acelerado, como se observa en la figura 3 donde se identifica una tendencia creciente a finales del 2017, pero poco después su precio se desploma a inicios del 2018. A principios de 2020 el precio del bitcoin volvió a crecer de manera acelerada alcanzando nuevos máximos históricos. En base a esto se decidió trabajar con modelos de series de tiempo, la clave de estos modelos está en considerar la información pasada de los precios y su volatilidad observada como factor altamente explicativo de su comportamiento presente y por extensión lógica, de su futuro predecible.

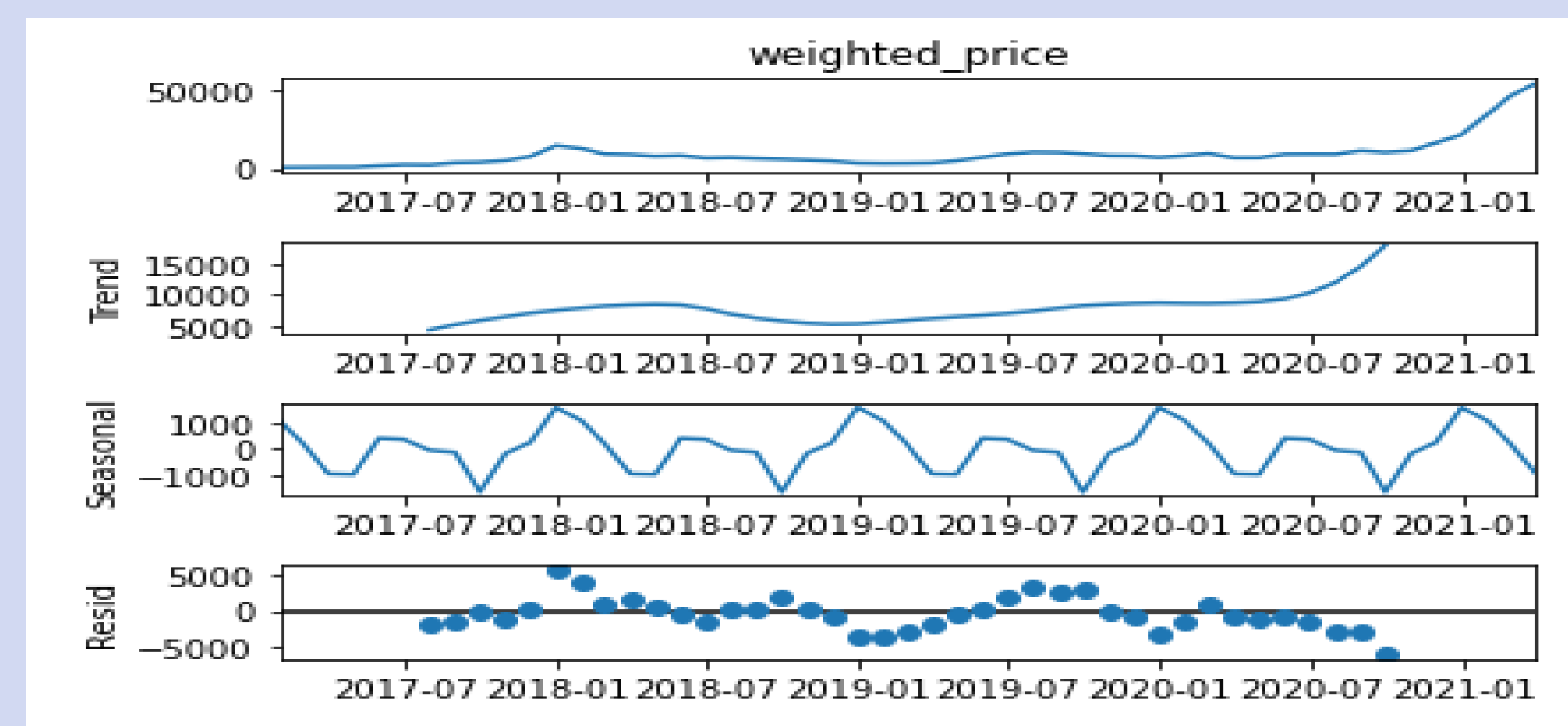


Figura 4. Valores observados, tendencia, estacional, Residuales

La evaluación de los modelos se realizó utilizando la métrica del error cuadrático medio(1), el cual calcula la diferencia que hay entre el valor estimado por el modelo y el valor real, el coeficiente de determinación(2) que es una medida estadística de la bondad del ajuste o fiabilidad del modelo estimado a los datos y se utilizó el criterio de desempeño aic(3) para evaluar la combinación de parámetros que generaban el mejor modelo.



Figura 3. Precios promedio, apertura y cierre del Bitcoin, recuencia diaria (2017-2020)

Para la construcción de los modelos de series de tiempo, primero revisamos la estacionariedad de la serie y mediante la prueba de Deckey Fuller se comprobó que la serie no era estacionaria, se puede corroborar observando la tendencia en la figura 4. Por lo tanto, fue necesario aplicar varias transformaciones para llegar a la estacionariedad de la serie. Estas transformaciones fueron: Time shift, Box cox, Logaritmica.

$$(1) \quad MSE = \frac{1}{M} \sum_{i=1}^M (real_i - estimado_i)^2$$

$$(2) \quad AIC = 2k - 2 \times \ln(L)$$

$$(3) \quad R^2 = \frac{\sum_{i=1}^M (estimado_i - promedio)^2}{\sum_{i=1}^M (real_i - promedio)^2}$$

RESULTADOS

Construimos tres modelos de series temporales los cuales fueron: AR, MA y ARIMA, las medidas de bondad de ajuste del Error cuadrático medio (Mean Squared Error, MSE) del activo objeto de estudio, se evidencian en la tabla 1.

MODELO	MSE
ARIMA	0.022
MA	0.036
AR	0.025

Tabla 1. Comparación del error cuadrático medio de los modelos

Para la selección del modelo que utilizamos para realizar las predicciones, elegimos el modelo que tuvo un error cuadrático medio más bajo, el cuál fue el modelo ARIMA que tuvo un Error cuadrático medio de 0.022.

Cabe mencionar que para la construcción de este modelo se tomo en consideración los parámetros que generaban un valor aic más bajo. Evaluamos la calidad de ajuste del modelo ARIMA y obtuvimos un coeficiente de determinación de 80.66% , lo que indica que el modelo es relativamente eficiente en el pronóstico de precios durante el periodo de estudio.

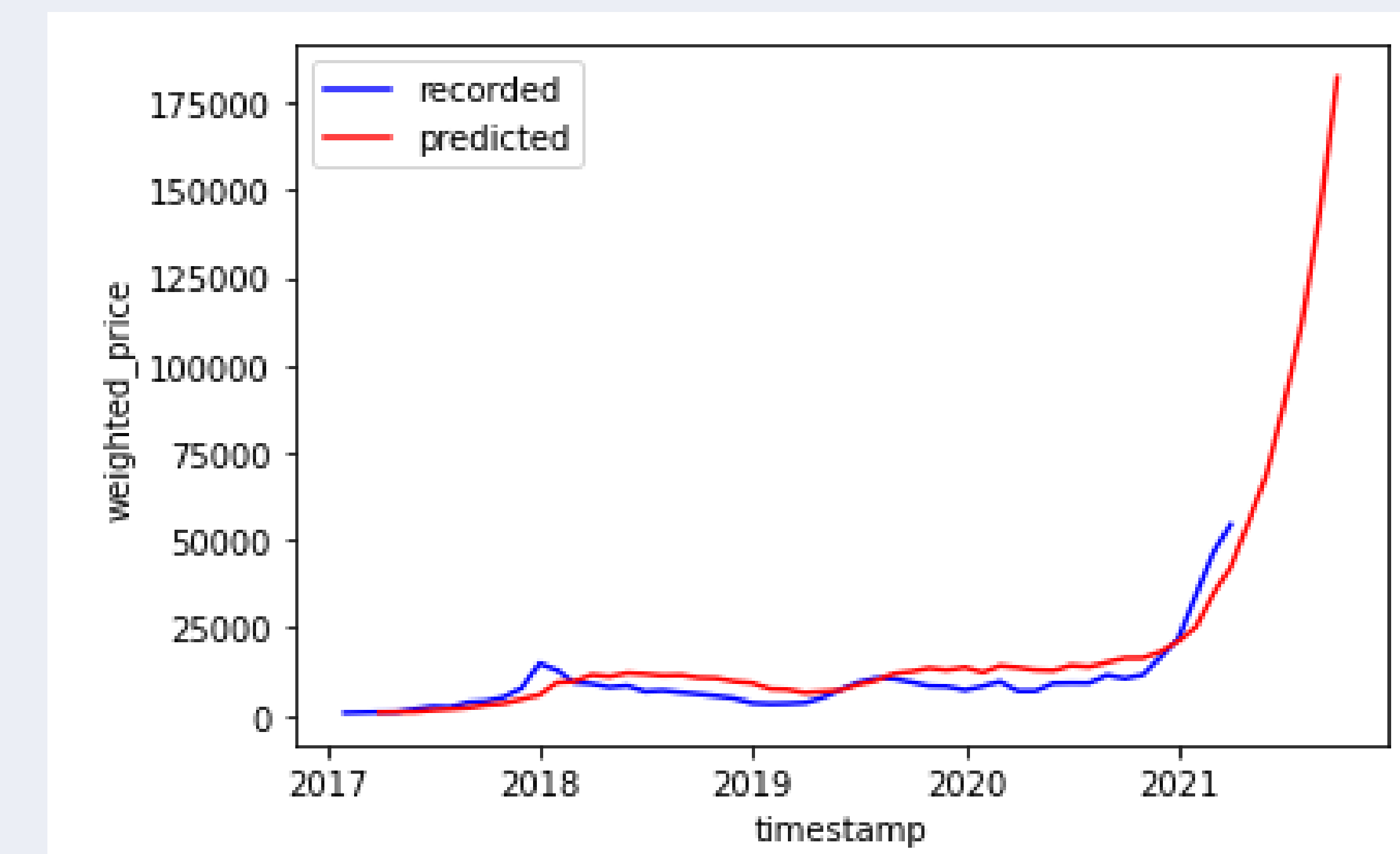


Figura 5. Comparación del valor real y pronosticado del precio del bitcoin para el año 2021

Finalmente, analizando el comportamiento del valor real del bitcoin en la figura 5, se observa que el pronóstico es cercano al valor real, tal como lo corrobora las medidas de evaluación observadas en la tabla 1 y el coeficiente de determinación.

CONCLUSIONES Y TRABAJO

A FUTURO

Es posible construir un modelo de pronóstico de los precios del bitcoin bajo el modelo ARIMA utilizando parámetros estadísticos para su estimación, realizando simulaciones para los valores pronosticados a partir de los valores de series periódicas, como lo son las series diarias del activo objeto de estudio. En este sentido, es importante destacar que dicho modelo no genera un pronóstico altamente certero para los precios debido a que los resultados de la calidad de ajuste generó valores que entran en el rango considerado como aceptable, sin embargo, el resultado de los pronósticos podría ser mejor.

Por lo tanto, se concluye que si bien es cierto que el modelo sirve para tener una aproximación a los posibles valores mínimos y máximos que puede tomar el activo, sus resultados carecen de la suficiente precisión para inducir la compra certera de este tipo de criptomoneda. Aunque en el modelo incorpora el factor aleatorio, aplicando las simulaciones, no alcanza a interpretar todos los eventos que se producen durante el periodo, por lo que al ser un activo muy volátil, se le genera menor capacidad predictiva.

Consecuencia de lo anterior, se recomienda para próximas investigaciones la aplicación de modelos como LSTM que son un tipo especial de redes recurrentes. y modelos de la familia Arch y Garch que generan mayor capacidad de predicción dado que permiten mejor modelamiento de su volatilidad.

RECURSOS

Base de datos

Código

