

Capítulo 1

Introducción

En este trabajo se presenta la investigación e implementación de un modelo híbrido de series temporales utilizado para hacer pronósticos del precio de criptomonedas (Bitcoin y Litecoin). El desarrollo del proyecto implica la implementación de dos modelos (ARIMA y LSTM) que se han utilizado actualmente para hacer los pronósticos y se evalúan los resultados obtenidos por dichos modelos contra un tercer modelo híbrido ARIMA-LSTM que a su vez se implementa.

1.1 Contexto del Problema

Las criptomonedas son monedas digitales que operan de forma descentralizada, permitiendo que los usuarios lleven a cabo transacciones sin que éstas dependan de un intermediario. Las criptomonedas, y en especial Bitcoin que es la más popular, atraen cada vez más la atención haciendo que muchas personas opten por invertir en ellas y que ciertos negocios comiencen a aceptarlas como forma de pago [1].

Las criptomonedas se describen como rápidas, convenientes, libres de impuesto, e incluso revolucionarias [1], sin embargo, también son conocidas por su alta volatilidad debido a distintos factores como el tener un mercado relativamente pequeño, baja liquidez, gran cantidad de especulaciones respecto a su futuro, entre otros. Esta naturaleza volátil, hace que predecir los precios de criptomonedas sea una tarea complicada [2].

1.2 Motivación

La alta volatilidad de las criptomonedas hace que sean una inversión bastante riesgosa, sin embargo, actualmente se llevan a cabo más de 250, 000 transacciones de criptomonedas al día [3]. El desplome del valor de Bitcoin más reciente ocurrió el 12 de Marzo, reportando pérdidas de prácticamente el 50% para los inversionistas en tan solo un día [4]. A partir de datos como éstos, nace la motivación para realizar este trabajo, que busca explorar la posibilidad de acercarnos a pronósticos más precisos y así reducir el riesgo al invertir en criptomonedas.

1.3 Objetivo del trabajo

El objetivo de este trabajo es determinar si los pronósticos del precio de criptomonedas hechas por el modelo híbrido ARIMA-LSTM son más certeros que las de los modelos ya hechos anteriormente.

1.4 Estructura del trabajo

Este trabajo está conformado por 7 capítulos que detallan la investigación hecha y posteriormente la implementación de los modelos.

En el capítulo 2 se habla del Estado del Arte. Se explican conceptos relevantes para el proyecto en general, así como otras soluciones propuestas para la problemática.

En el capítulo 3 se describe más a detalle la problemática actual y del porqué es importante seguir buscando soluciones alternas.

En el capítulo 4 se habla a profundidad de la propuesta de solución, explicando tanto los aspectos y conceptos relevantes para su desarrollo, como el funcionamiento de los modelos a implementar.

En el capítulo 5 se define la metodología de evaluación que se utiliza, así como la justificación para la elección de la misma.

En el capítulo 6 se muestran los resultados obtenidos al evaluar cada modelo. Se aprecian visualmente los pronósticos de precios hechos por cada modelo y los precios reales para compararlos.

En el capítulo 7 se presentan las conclusiones basadas en los resultados obtenidos, se retoman puntos importantes del trabajo y se plantean posibles trabajos futuros teniendo ya una base.

Capítulo 2

Estado del arte

En este capítulo se describen conceptos relevantes en el área **maybe tales como o no sé**. Posteriormente se mencionan los modelos ARIMA y LSTM, que se han implementado para hacer pronósticos del precio de las criptomonedas, así como los resultados que obtuvieron. Analizar estos trabajos permite encontrar un área de oportunidad a la cual se puede contribuir.

2.1 Contexto del área

2.1.1 Series de tiempo

2.1.1.1 Definición

Una serie de tiempo es un conjunto secuencial de datos medidos típicamente en tiempos sucesivos. Las medidas que se toman durante cierto periodo se organizan en orden cronológico [5]. Se define matemáticamente una serie de tiempo como:

$$\{X_1, X_2, \dots, X(t)\}, t = 1, 2, 3, \dots$$

t: Representa el tiempo transcurrido

X(t): Es el valor del dato medido durante un tiempo en específico.

Las series de tiempo se denominan univariantes si los registros que se tienen a lo largo del tiempo son de una sola variable, pero si los registros son de más de una variable entonces se dice que es multivariante. [5, 6] Un ejemplo de serie de tiempo univariante es el medir la temperatura de alguna máquina de forma periódica cada minuto, por otro lado si se midiera tanto la temperatura como la vibración y las revoluciones, se estaría hablando de una serie de tiempo multivariante.

2.1.1.2 Visualización

Para visualizar los datos de una serie de tiempo, generalmente éstos se grafican trazando las observaciones con respecto al tiempo. Esto permite tener una primer aproximación y entendimiento sobre el cómo se comporta la variable a lo largo del tiempo.

La figura tal muestra un ejemplo de shalala, el tiempo va en el eje x los valores de la variable registrada en el eje y

2.1.1.2 Análisis y Pronóstico de series de tiempo

El pronóstico de series de tiempo consiste en recolectar datos históricos de la variable de interés para posteriormente ser analizados y desarrollar un modelo que comprenda las relaciones que existen en los datos y que posiblemente no sean notorias. El modelo después es utilizado para hacer pronósticos del comportamiento que se espera que tenga la variable recolectada. [7, 8]

2.2 Trabajos Relevantes

ARIMA

LSTM

ARIMA-LSTM

Ya para la metodología

$x(t)$ hace referencia al valor que tiene una criptomoneda para un día en particular.

Capítulo 7

Conclusiones

TRABAJOS FUTUROS:

Yo use el closing Price porque es como el estado del arte

Maybe un EDA para checar si es chido usar algo mas que el closing price

Bibliografía

[1] M. Rahouti, K. Xiong and N. Ghani, "Bitcoin Concepts, Threats, and Machine-Learning Security Solutions," in *IEEE Access*, vol. 6, pp. 67189-67205, 2018.

[2] Asante, S. Are Bitcoins price predictable? Evidence from machine learning techniques using technical indicators. 2019. pp. 2.

[3] Beck et al. Sensing Social Media Signals for Cryptocurrency News . 2019. pp. 1-2.

[4] Bovaird, C. (2020). *Bitcoin Lost Roughly 50% Of Its Value In A Day*. Mayo 4, 2020, de Forbes Sitio web: <https://www.forbes.com/sites/cbovaird/2020/03/12/bitcoin-lost-roughly-50-of-its-value-in-a-day>

[5] Adhikari, R; Agrawal R. An Introductory Study on Time Series Modeling and Forecasting. pp. 12-13.

[6] John H. Cochrane, "Time Series for Macroeconomics and Finance", Graduate School of Business, University of Chicago, spring 1997.

[7] G.P. Zhang, "A neural network ensemble method with jittered training data for time series forecasting", *Information Sciences* 177 (2007), pp: 5329–5346.

[8] G.P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model", *Neurocomputing* 50 (2003), pp: 159–175.