



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería Mecánica Eléctrica

**DISEÑO DE INVESTIGACIÓN PARA LA COMPARACIÓN DE ALGORITMOS
DE *MACHINE LEARNING* APLICADOS A LA PREDICCIÓN DEL VALOR DEL PRECIO
DE CRIPTOMONEDAS, A TRAVÉS DE PRUEBAS ESTADÍSTICAS DE CONTRASTE
Y *POST HOC*, PARA SELECCIONAR AQUELLOS CON EL MEJOR DESEMPEÑO**

Kevin Adolfo Duarte Chamalé

Asesorado por Mtro. Ing. Luis Carlos Leonardo Bolaños Méndez

Guatemala, mayo de 2023

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**DISEÑO DE INVESTIGACIÓN PARA LA COMPARACIÓN DE ALGORITMOS
DE *MACHINE LEARNING* APLICADOS A LA PREDICCIÓN DEL VALOR DEL PRECIO
DE CRIPTOMONEDAS, A TRAVÉS DE PRUEBAS ESTADÍSTICAS DE CONTRASTE
Y *POST HOC*, PARA SELECCIONAR AQUELLOS CON EL MEJOR DESEMPEÑO**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

KEVIN ADOLFO DUARTE CHAMALÉ

ASESORADO POR MTRO. ING. LUIS CARLOS LEONARDO BOLAÑOS MÉNDEZ

AL CONFERÍRSELE EL TÍTULO DE

INGENIERO EN ELECTRÓNICA

GUATEMALA, MAYO DE 2023

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANA	Inga. Aurelia Anabela Cordova Estrada
VOCAL I	Ing. José Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martinez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Kevin Bladimir Cruz Lorente
VOCAL V	Br. Fernando José Paz Gonzalez
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANA	Inga. Aurelia Anabela Cordova Estrada
EXAMINADOR	Ing. José Aníbal Silva de los Angeles
EXAMINADOR	Ing. Helmunt Federico Chicol Cabrera
EXAMINADOR	Ing. Marvin Marino Hernández Fernández
SECRETARIO	Inga. Hugo Humberto Rivera Pérez

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

DISEÑO DE INVESTIGACIÓN PARA LA COMPARACIÓN DE ALGORITMOS DE *MACHINE LEARNING* APLICADOS A LA PREDICCIÓN DEL VALOR DEL PRECIO DE CRIPTOMONEDAS, A TRAVÉS DE PRUEBAS ESTADÍSTICAS DE CONTRASTE Y *POST HOC*, PARA SELECCIONAR AQUELLOS CON EL MEJOR DESEMPEÑO

Tema que me fuera asignado por la Dirección de la Escuela de Estudios de Posgrado, con fecha 8 de noviembre de 2022.

Kevin Adolfo Duarte Chamalé



EEPFI-PP-1807-2022

Guatemala, 10 de noviembre de 2022

Director
Armando Alonso Rivera Carrillo
Escuela De Ingenieria Mecanica Electrica
Presente.

Estimado Ing. Rivera

Reciba un cordial saludo de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería.

El propósito de la presente es para informarle que se ha revisado y aprobado el Diseño de Investigación titulado: **COMPARACIÓN DE ALGORITMOS DE MACHINE LEARNING APLICADOS A LA PREDICCIÓN DEL VALOR DEL PRECIO DE CRIPTOMONEDAS, A TRAVÉS DE PRUEBAS ESTADÍSTICAS DE CONTRASTE Y POST HOC, PARA SELECCIONAR AQUELLOS CON EL MEJOR DESEMPEÑO**, el cual se enmarca en la línea de investigación: **Todas las áreas - Métodos no paramétricos**, presentado por el estudiante **Kevin Adolfo Duarte Chamalé** carné número **201503960**, quien optó por la modalidad del "PROCESO DE GRADUACIÓN DE LOS ESTUDIANTES DE LA FACULTAD DE INGENIERÍA OPCIÓN ESTUDIOS DE POSTGRADO". Previo a culminar sus estudios en la Maestría en ARTES en Estadística Aplicada.

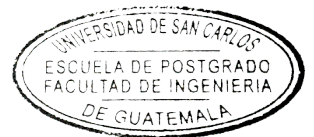
Y habiendo cumplido y aprobado con los requisitos establecidos en el normativo de este Proceso de Graduación en el Punto 6.2, aprobado por la Junta Directiva de la Facultad de Ingeniería en el Punto Décimo, Inciso 10.2 del Acta 28-2011 de fecha 19 de septiembre de 2011, firmo y sello la presente para el trámite correspondiente de graduación de Pregrado.

Atentamente,

"Id y Enseñad a Todos"

Mtro. Luis Carlos Leonardo Bolaños Méndez
Asesor(a)

Mtro. Edwin Adalberto Bracamonte Orozco
Coordinador(a) de Maestría



Mtro. Edgar Darío Álvarez Cotí
Director
Escuela de Estudios de Postgrado
Facultad de Ingeniería





EEP-EIME-1457-2022

El Director de la Escuela De Ingenieria Mecanica Electrica de la Facultad de Ingenieria de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del Asesor, el visto bueno del Coordinador y Director de la Escuela de Estudios de Postgrado, del Diseño de Investigación en la modalidad Estudios de Pregrado y Postgrado titulado: **COMPARACIÓN DE ALGORITMOS DE MACHINE LEARNING APLICADOS A LA PREDICCIÓN DEL VALOR DEL PRECIO DE CRIPTOMONEDAS, A TRAVÉS DE PRUEBAS ESTADÍSTICAS DE CONTRASTE Y POST HOC, PARA SELECCIONAR AQUELLOS CON EL MEJOR DESEMPEÑO**, presentado por el estudiante universitario **Kevin Adolfo Duarte Chamalé**, procedo con el Aval del mismo, ya que cumple con los requisitos normados por la Facultad de Ingenieria en esta modalidad.

ID Y ENSEÑAD A TODOS

Ing. Armando Alonso Rivera Carrillo
Director
Escuela De Ingenieria Mecanica Electrica

Guatemala, noviembre de 2022

Decanato
Facultad de Ingeniería
24189101- 24189102
secretariadecanato@ingenieria.usac.edu.gt

LNG.DECANATO.OI.433.2023

La Decana de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería Mecánica Eléctrica, al Trabajo de Graduación titulado: **DISEÑO DE INVESTIGACIÓN PARA LA COMPARACIÓN DE ALGORITMOS DE MACHINE LEARNING APLICADOS A LA PREDICCIÓN DEL VALOR DEL PRECIO DE CRIPTOMONEDAS, A TRAVÉS DE PRUEBAS ESTADÍSTICAS DE CONTRASTE Y POST HOC, PARA SELECCIONAR AQUELLOS CON EL MEJOR DESEMPEÑO**, presentado por: **Kevin Adolfo Duarte Chamalé**, después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:

Inga. Aurelia Anabela Cordova Estrada

Decana



Guatemala, mayo de 2023

AACE/gaoc

ACTO QUE DEDICO A:

Dios		Por darme la vida y guiarme siempre.
Mis padres		Gustavo Duarte y Olga Chamalé, por brindarme su amor y su apoyo incondicional, por ser mi ejemplo a seguir y por ser mi inspiración.
Mi hermana		Rocío Duarte, por siempre estar ahí para darme ánimos y su apoyo.
Mi familia		A mis abuelas, tíos, tías, primos y primas, por ser una familia muy unida, de la cual me siento muy feliz de ser parte.
Mi abuelo		José Duarte (q. e. p. d.), por todas sus enseñanzas y consejos.
Departamento de Matemática	de	Al ingeniero Samayoa, por haberme dado la oportunidad de ser auxiliar de matemáticas, permitiéndome vivir una experiencia única y conocer personas maravillosas.

Departamento de Física

Al ingeniero Walter Álvarez, por haberme dado la oportunidad de ser auxiliar de física, lo que me permitió aprender mucho, y crecer personal y profesionalmente.

Mis compañeros de carrera

Por haberme compartido sus conocimientos, brindarme su apoyo y amistad; a Juan Valdez, José González, Jeffrey Hipp, Erwin Lemus, Eduardo López, Enrique Coloch, Cynthia Hernández, Renato Cifuentes, Fernando Xitumul y a todos con quienes pude compartir cursos.

Mis amigos

Marco Gómez, Erick Mendoza, Josué Juárez, Mitchel Cano, Pedro Chamalé, Christian Estrada, Diego Orellana, Max Morales, Andrés Divas, Alejandro López, Gilberto Arauz, Jairo García y a todos los demás amigos que conocí en mi etapa universitaria, por todos los buenos e inolvidables momentos que hemos vivido juntos.

AGRADECIMIENTOS A:

Universidad de San Carlos de Guatemala	Por ser mi casa de estudios y darme la oportunidad de acceder a una formación académica de nivel superior.
Facultad de Ingeniería	Por darme la oportunidad de estudiar la carrera de Ingeniería Electrónica, y permitirme crecer académica y profesionalmente.
Mi asesor	Al Mtro. Ing. Luis Bolaños, por su tiempo y su apoyo al asesorarme a lo largo de este trabajo.
Bam	Al banco y a todo mi equipo de trabajo, por darme la oportunidad de aprender y desarrollarme profesionalmente junto a compañeros y amigos increíbles.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
LISTA DE SÍMBOLOS	VII
GLOSARIO	IX
RESUMEN.....	XI
1. INTRODUCCIÓN	1
2. ANTECEDENTES	5
3. PLANTEAMIENTO DEL PROBLEMA	11
3.1. Contexto general	11
3.2. Descripción del problema	11
3.3. Formulación del problema	12
3.4. Delimitación del problema	12
4. JUSTIFICACIÓN	13
5. OBJETIVOS	15
5.1. General.....	15
5.2. Específicos	15
6. NECESIDADES A CUBRIR Y ESQUEMA DE SOLUCIÓN.....	17
7. MARCO TEÓRICO.....	19
7.1. Fundamentos estadísticos.....	19

7.1.1.	Estadística.....	19
7.1.1.1.	Estadística descriptiva.....	19
7.1.1.2.	Estadística inferencial	20
7.1.1.2.1.	Métodos paramétricos...20	
7.1.1.2.2.	Métodos no paramétricos	21
7.1.2.	Pruebas de normalidad	21
7.1.2.1.	Análisis de curtosis y sesgo	22
7.1.2.2.	Prueba de Kolmogórov-Smirnov	24
7.1.2.3.	Prueba de Anderson-Darling	25
7.1.2.4.	Prueba de Shapiro-Wilk.....	25
7.1.3.	Diseño de experimentos.....	26
7.1.3.1.	Paramétricos	26
7.1.3.1.1.	Análisis de varianza de un factor (ANOVA)	26
7.1.3.2.	No paramétricos	27
7.1.3.2.1.	Prueba de Friedman	27
7.1.3.2.2.	Prueba de Kruskal-Wallis	29
7.1.4.	Pruebas post hoc	31
7.1.4.1.	Paramétricas	31
7.1.4.1.1.	Método de Tukey	31
7.2.	Criptomonedas y machine learning.....	32
7.2.1.	Criptomonedas	32
7.2.1.1.	Blockchain	33
7.2.1.2.	Criptomonedas con la mayor capitalización del mercado	33
7.2.1.2.1.	Bitcoin (BTC).....	33
7.2.1.2.2.	Ethereum (ETH)	34

	7.2.1.2.3.	Tether (USDT)	34	
	7.2.1.2.4.	BNB (BNB)	34	
	7.2.1.2.5.	USDC (USDC).....	34	
	7.2.1.2.6.	Ripple (XRP).....	35	
	7.2.1.2.7.	BUSD (BUSD)	35	
	7.2.1.2.8.	Cardano (ADA).....	35	
	7.2.1.2.9.	Bitcoin cash (BCH)	35	
	7.2.1.2.10.	Litecoin (LTC).....	36	
	7.2.1.2.11.	Stellar (XLM).....	36	
	7.2.1.2.12.	Monero (XMR)	36	
7.2.2.	Machine learning		36	
	7.2.2.1.	Aprendizaje supervisado	38	
	7.2.2.2.	Aprendizaje no supervisado.....	40	
	7.2.2.3.	Aprendizaje semisupervisado	42	
	7.2.2.4.	Aprendizaje reforzado.....	42	
7.2.3.	Redes neuronales.....		42	
8.	PROPUESTA DE ÍNDICE DE CONTENIDOS		45	
9.	METODOLOGÍA.....		49	
	9.1.	Características del estudio	49	
	9.2.	Unidad de análisis	49	
	9.3.	Variables.....	50	
	9.4.	Fases del estudio	51	
		9.4.1.	Fase uno: revisión bibliográfica	51
		9.4.2.	Fase dos: recolección y limpieza de los datos.....	51
		9.4.3.	Fase tres: desarrollo de los modelos	51
		9.4.4.	Fase cuatro: cálculo de resultados	53
		9.4.5.	Fase cinco: interpretación de resultados	54

9.4.6.	Fase seis: redacción del informe final	54
10.	TÉCNICAS DE ANÁLISIS DE INFORMACIÓN	55
11.	CRONOGRAMA	57
12.	FACTIBILIDAD DEL ESTUDIO	59
12.1.	Recurso humano	59
12.2.	Recursos financieros	59
12.3.	Recursos tecnológicos	60
12.4.	Acceso a información y permisos	60
12.5.	Equipo e infraestructura	60
13.	REFERENCIAS	61
14.	APÉNDICE	67

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Diagrama algoritmo de clasificación de aprendizaje supervisado	38
2.	Diagrama algoritmo de regresión de aprendizaje supervisado.....	39
3.	Visualización de algoritmos de aprendizaje no supervisado (PCA).....	41
4.	Diagrama red neuronal.....	43
5.	Diagrama de la fase dos y tres del estudio.....	52
6.	Diagrama de la fase cuatro del estudio	53
7.	Flujograma del estudio	54
8.	Cronograma de actividades.....	57

TABLAS

I.	Diseño de una prueba de Kruskall-Wallis.....	30
II.	Variables del estudio	50
III.	Presupuesto	59

LISTA DE SÍMBOLOS

Símbolo	Significado
\$	Dólar
Gb	Gigabyte
Mb	Megabyte

GLOSARIO

Algoritmo	Conjunto ordenado y finito de operaciones que permite hallar la solución de un problema.
Aprendizaje de máquina	Subcampo de las ciencias de la computación y una rama de la inteligencia artificial, cuyo objetivo es desarrollar técnicas para que las computadoras aprendan.
Criptomoneda	También denominada criptodivisa; es un tipo de moneda digital que utiliza la criptografía para proporcionar un sistema de pagos seguro.
Redes neuronales	Una red neuronal es un método de la inteligencia artificial, que enseña a las computadoras a procesar datos de una manera que está inspirada en la forma en que lo hace el cerebro humano.

RESUMEN

Este diseño de investigación presenta una propuesta para el proceso de comparación de múltiples algoritmos de *machine learning*, aplicados a la predicción del precio de las principales criptomonedas con la mayor capitalización en el mercado. El objetivo es emplear un proceso estadístico riguroso, que permita o no hallar diferencias significativas en el desempeño de estos algoritmos.

Por lo tanto, en este diseño se detallan las criptomonedas que se utilizarán, los algoritmos que se van a comparar y especialmente las pruebas estadísticas que se llevarán a cabo en el proceso. Estas pruebas se aplicarán a las métricas de precisión y eficiencia de los algoritmos seleccionados, los cuales se entrenarán y probarán con datos recolectados de las criptomonedas que se detallan en el trabajo.

Finalmente, con los resultados de estas pruebas, se concluirá la existencia o no de diferencias significativas en el desempeño de los algoritmos. Lo cual permitirá seleccionar aquel o aquellos con el mejor desempeño entre los demás, bajo criterios estadísticos.

1. INTRODUCCIÓN

El presente trabajo de investigación consiste en la sistematización del proceso de comparación estadística de algoritmos de aprendizaje automático, así como de redes neuronales aplicadas a la predicción del valor de criptomonedas.

La predicción de series temporales es un proceso complicado, sobre todo en aquellos fenómenos de alta variabilidad y muy volátiles, como lo es el valor de las criptomonedas en el mercado bursátil. Para este proceso existen varios métodos, desde métodos clásicos hasta métodos más complejos como lo son los algoritmos de aprendizaje automático y redes neuronales. En este trabajo se busca comparar el desempeño de estos últimos, seleccionando aquellos algoritmos que hayan demostrado desempeños superiores en estudios previos, cuantificando y comparando su eficiencia y precisión con un proceso de alto rigor estadístico.

La popularidad de las criptomonedas ha ido en crecimiento en los últimos años, su uso se ha extendido incluso en la economía de varios países. Por lo tanto, el desarrollo de modelos para su predicción ya no solamente es un tema de interés académico, sino también un tema de interés económico y laboral. En este trabajo, entonces, se busca describir cuáles son los algoritmos de aprendizaje automático y redes neuronales que presenten desempeños superiores con estos datos, así como detallar el proceso estadístico adecuado para su comparación.

La metodología de esta investigación es de enfoque cuantitativo de tipo descriptivo, con un diseño cuasiexperimental y un alcance descriptivo-correlacional transversal. La razón es porque se utilizarán datos históricos del valor de precio de las criptomonedas y número de transacciones, entre los años de 2018 y 2022, pero para la comparación de los algoritmos se deben controlar distintas fuentes de variación aleatoria.

Los resultados esperados con esta investigación consisten en una descripción, cuantificación y comparación del desempeño de algoritmos de aprendizaje automático y redes neuronales, aplicados a datos del mercado de criptomonedas. Se espera que la información obtenida sirva como base para futuros trabajos, que busquen desarrollar y optimizar un modelo de predicción con datos de esta índole, empleando algoritmos de aprendizaje automático y redes neuronales, sin la necesidad de rehacer el proceso de comparación de algoritmos, para ahorrar tiempo y recursos.

Esta investigación consta de las fases que se describen a continuación. La primera fase consiste en la revisión bibliográfica que servirá de referencia, utilizando trabajos similares, y la recolección del fundamento estadístico necesario. La segunda fase se centra en la recolección y limpieza de los datos a utilizar. En la tercera fase se desarrollarán los modelos de aprendizaje automático y redes neuronales que se van a comparar. En la cuarta fase se hará el cálculo de resultados, que se refiere a todas las pruebas estadísticas planteadas. En la quinta fase se interpretarán los resultados mencionados. Y en la última fase, la sexta, se harán las conclusiones y recomendaciones, y redacción del informe final.

Este trabajo es factible porque se cuenta con todos los recursos necesarios para desarrollar las fases del estudio.

La investigación consistirá en cuatro capítulos divididos de la siguiente forma:

En el primer capítulo se presentará el marco referencial, que consistirá en el estudio de diversos trabajos relacionados, que sirvan de referencia y base para el desarrollo de la investigación. En este caso, se estudiarán investigaciones en las que se hayan desarrollado y comparado modelos de aprendizaje automático y redes neuronales con datos del valor de criptomonedas.

El segundo capítulo, el marco teórico, se trata de un compendio de la teoría que justificará todas las pruebas y procedimientos que se realizarán en la investigación. Esta tendrá dos partes: la parte estadística y la parte específica del tema. En la parte estadística se incluyen pruebas de bondad de ajuste, análisis de varianza (ANOVA), pruebas de contraste de muestra de tipo no paramétrico y pruebas *post hoc*. En la parte específica del tema de investigación se incluirá información acerca de modelos de aprendizaje automático y redes neuronales, así como información acerca de las criptomonedas.

En el tercer capítulo se presentarán los resultados, mostrando la cuantificación del desempeño de los modelos, las pruebas de normalidad y las pruebas de contraste y *post hoc* adecuadas a la distribución de los datos.

En el cuarto capítulo se hará la discusión de resultados, donde se resaltarán aquellas diferencias significativas entre los algoritmos comparados.

Finalmente, se presentarán las conclusiones y recomendaciones del estudio, basado en si se encuentran o no diferencias significativas en el desempeño de los algoritmos estudiados.

2. ANTECEDENTES

Para la comparación de algoritmos de *machine learning*, antes de establecer las pruebas estadísticas a utilizar, es necesario definir todas aquellas pautas y consideraciones que se debe tener en cuenta para la comparación que se desea efectuar, y que garanticen una aplicación correcta de dichas pruebas. Una vez definidas estas, también es necesario estudiar qué algoritmos son los más adecuados para emplear en este ámbito, y las métricas principales con las que se estudia su precisión.

Dietterich (1998) menciona que, para el diseño y evaluación de pruebas estadísticas, uno de los primeros pasos es identificar las fuentes de variación que deben ser controladas para cada prueba. En el caso de aplicaciones con algoritmos de redes neuronales, se asumen cuatro fuentes de variación importante. La primera es la aleatoriedad en la selección del *dataset* de prueba (*test dataset*) utilizado para evaluar los algoritmos; esto es un problema sobre todo cuando se tienen *datasets* pequeños. La segunda fuente de variación viene de la selección del *dataset* de entrenamiento (*train dataset*), esto debido a que la menor alteración en los datos puede llegar a cambiar grandemente los resultados de algunos algoritmos. La tercera fuente de variación es la aleatoriedad interna presente en los algoritmos de aprendizaje; por ejemplo, en algunos es usual iniciar con ciertos parámetros aleatorios, que luego se van optimizando, por lo que un inicio diferente puede cambiar también los resultados. La última fuente de variación se encuentra en errores aleatorios que pueden existir en los datos (refiriéndose a valores incorrectos), en este caso por problemas a la hora de registrar los datos, entre otros.

Hay que diferenciar la comparación de modelos y la comparación de algoritmos de *machine learning*. Raschka (2018) indica que, si se desea comparar modelos, no se debe tomar en cuenta la variación aleatoria del *dataset* de entrenamiento, cosa que hay que considerar para comparar algoritmos, y que puede ser un problema para un *dataset* pequeño. Esto, sobre todo, si se utilizan algoritmos de aprendizaje que son susceptibles a las perturbaciones en el *dataset* de entrenamiento. También menciona que para comparar algoritmos se debe tener un conjunto de modelos que han sido optimizados, utilizando diferentes algoritmos de aprendizaje sobre varios datos de entrenamiento independientes. La comparación de algoritmos es deseable en muchos casos, por ejemplo, cuando estamos comparando un nuevo algoritmo o cuando queremos decidir cuál es el mejor para nuestra aplicación. En estos casos, el objetivo puede ser encontrar qué tan diferente es el desempeño de los algoritmos, utilizando varios conjuntos de datos provenientes de un dominio similar.

Trawinski, Smętek, Telec y Lasota (2012) mencionan que cuando un investigador busca comparar una nueva técnica (o algoritmo) con un número de algoritmos conocidos, o elegir el mejor entre un grupo de algoritmos, las pruebas que implican un contraste por pares no son adecuadas. En tales situaciones, el investigador pierde control sobre la llamada tasa de error de familia, debido a un error acumulado que viene de la combinación de las comparaciones por pares. Por lo tanto, en estos casos es adecuado utilizar pruebas de comparaciones múltiples combinadas con pruebas *post hoc*, ya sea para comparar 1 algoritmo frente a varios (comparación 1xN) o evaluar todas las combinaciones (comparación NxN). Entre las principales pruebas que se pueden aplicar están la prueba de Friedman, y también su derivación más potente, la prueba de Imán y Davenport. Estas pruebas tienen la característica que ranquean nuestros algoritmos del mejor al peor desempeño.

Almqvist (2019) utiliza dos métricas de precisión para la comparación de los algoritmos. La primera, el RMSE, la cual presenta un mayor peso en los *outliers* que el MAE; por lo que un valor bajo de RMSE sería un mejor indicador que el MAE, para revelar una estabilidad predictiva. La segunda métrica es el MAPE, el cual muestra un promedio porcentual del error, ya que es una herramienta práctica y útil, pues es fácil de entender e interpretar.

Mahir, Muhammad, Fawwad, Khurum y Ali (2021) utilizaron los algoritmos ARIMA, FBProphet y XGBoost para entrenar los modelos para predecir el precio de mercado y la estabilidad del bitcoin. Las métricas de evaluación que utilizaron fueron el RMSE, MAE y R^2 . Sus resultados indican que el modelo entrenado con el algoritmo ARIMA fue el mejor.

Hitam e Ismail (2018) compararon los algoritmos basados en SVM (*Support Vector Machines*), ANN (*Artificial Neural Networks*) y *Deep Learning Forecasting* para seis diferentes criptomonedas: bitcoin, ethereum, litecoin, nem, ripple y stellar. Para la comparación utilizaron el MAPE. Los resultados muestran que los modelos de algoritmos basados en SVM tienen los valores más bajos del MAPE, haciéndolo el mejor algoritmo para la predicción de las criptomonedas mencionadas.

Mendes (2019), en su investigación, compara modelos de predicción del precio del bitcoin, utilizando dos algoritmos: ARIMA y LSTM. Comparándolos a través del RMSE y el MAE, descubrió que LSTM supera el desempeño de ARIMA.

Derbentsev, Datsenko, Stepanenko y Bezkorovainyi (2019) compararon modelos basados en BART, que es una combinación del clásico algoritmo C&RT y el modelo ARIMA, y usaron el método ARIMA-ARFIMA. Los modelos

fueron realizados con un alcance de corto plazo (5 a 30 días) para las tres criptomonedas más importantes: bitcoin, ethereum y ripple. El método propuesto con BART obtuvo los resultados más precisos, con un RMSE para 14, 21 y 30 días, de 4 %, 6 % y 8 %, respectivamente. Asimismo, demostró una mejor eficiencia para construir estimadores predictivos, en comparación con técnicas tradicionales de series temporales.

Luengo, García y Herrera (2009) utilizaron varias pruebas para estudiar si las muestras obtenidas, a partir de múltiples corridas de algoritmos convencionales de redes neuronales artificiales y SVM, poseían las condiciones necesarias para analizarlas a través de pruebas paramétricas. Para ello, consideraron tres fuentes de variación en los experimentos de clasificación: variación aleatoria en el *dataset* de prueba (test), variación aleatoria en el *dataset* de entrenamiento (*train*) y la variación aleatoria de los propios algoritmos. Encontraron que cumplir con las condiciones paramétricas depende del problema de estudio y está indefinido, lo que justifica el uso de pruebas no paramétricas para el análisis experimental de estos problemas.

McNally, Roche y Caton (2018) estudiaron con qué precisión se puede predecir la dirección del cambio en el precio del bitcoin. Para ello, implementaron dos modelos: una red neuronal recurrente con optimización bayesiana (RNN) y una red de *Long Short Term Memory* (LSTM). La red LSTM obtuvo los mejores resultados, con una precisión de clasificación del 52 % y un RMSE del 8 %. También se compararon los resultados con el modelo ARIMA para la predicción de series temporales. Como se esperaba, encontraron que los modelos de *Deep Learning* superaron el método clásico. También compararon el tiempo de entrenamiento en una GPU y en una CPU, lo cual reveló que en la primera era un 67.7 % más eficiente.

Los estudios previos se utilizarán para la selección de las métricas de precisión, adecuadas para la comparación de los algoritmos. Las métricas de mayor uso en estudios previos son las siguientes: RMSE, MAE y MAPE. Otro aspecto que se encontró en estudios previos revela que inicialmente es necesario un estudio exhaustivo de las condiciones de normalidad de los datos, para poder seleccionar las pruebas de contraste y *post hoc* adecuadas. Se observa también que, como mínimo, es necesario tomar en cuenta tres fuentes de variación principales: datos de entrenamiento, datos de prueba y variación aleatoria interna de los algoritmos.

Además, estudios previos demuestran que los algoritmos de redes neuronales son los más utilizados para predecir el valor de criptomonedas; sus resultados superan a métodos tradicionales de análisis y predicción de series temporales. Esta información será útil para delimitar y seleccionar los algoritmos de aprendizaje automático que se utilizarán.

3. PLANTEAMIENTO DEL PROBLEMA

3.1. Contexto general

En el desarrollo de los proyectos de *machine learning* es usual llegar a tener a nuestra disposición varios algoritmos de aprendizaje. En una parte del proyecto, es necesario elegir que algoritmos son más adecuados para el fenómeno en estudio. Esta decisión depende del tipo de datos que se utilicen, el dominio de aplicación, además de otros factores como por ejemplo los recursos computacionales, el tiempo que se disponga, entre otros.

3.2. Descripción del problema

En el caso de la predicción del valor de las criptomonedas, existe una variedad de algoritmos que pueden ser utilizados y cada uno genera modelos con determinados valores de desempeño. Para comparar su desempeño (de los algoritmos) en un dominio de aplicación (en este caso, predicción del precio de las criptomonedas), se debe tomar en cuenta que hay múltiples fuentes de variación aleatoria presentes, ya sea en los datos, por temas de muestreo o en el mismo proceso de aprendizaje de los algoritmos. Por lo tanto, para compararlos adecuadamente, se desconoce el conjunto de pruebas estadísticas que nos permitan encontrar diferencias significativas entre el desempeño de estos, así como el grado de estas diferencias.

3.3. Formulación del problema

- Pregunta central

¿Cuál es la diferencia entre el desempeño de los algoritmos de *machine learning* aplicados a la predicción del valor de las criptomonedas?

- Preguntas auxiliares
 - ¿Cuáles son las métricas del desempeño de los modelos de prueba que deben utilizarse para la comparación?
 - ¿Qué distribución de probabilidad se ajusta mejor al comportamiento de las métricas de desempeño?
 - ¿Qué diferencias existen entre el desempeño de los algoritmos?

3.4. Delimitación del problema

Las criptomonedas más populares por su volumen de mercado son: BTC, ETH, USDT, BNB, USDC, XRP, BUSD, ADA, XLM, XMR. Segmentadas por hora durante el periodo de 2018 a 2022. Los tipos de algoritmos para comparar son: FNN, RNN, LSTMs, CNNs, SVMs y *gradient boosting*.

4. JUSTIFICACIÓN

En este estudio se realizarán pruebas de hipótesis de contraste y *post hoc* para muestras múltiples, por lo que su línea de investigación principal serán los métodos no paramétricos. Sin embargo, se consideran a los métodos paramétricos como línea de investigación secundaria, pues con base en la distribución de los datos que se encuentren, se utilizarán pruebas correspondientes a una u otra línea. Así también, se utilizarán conceptos del diseño de experimentos para poder controlar las fuentes de variación presentes en las pruebas que se hagan.

Esta investigación permitirá hallar, con rigor estadístico, la existencia de diferencias significativas entre un conjunto de algoritmos de aprendizaje automático y redes neuronales que se han encontrado, con base en investigaciones previas consultadas, que tienen buenos resultados para el desarrollo de modelos de predicción con datos de criptomonedas. Esto permitirá a futuras investigaciones, que busquen desarrollar modelos de predicción en el tema de criptomonedas, tener una referencia para la elección del algoritmo óptimo. Así mismo, se pretende dejar evidencia detallada del proceso de comparación seguido, para que pueda ser replicado con datos de otro dominio.

La motivación para desarrollar esta investigación proviene del interés en aplicar los métodos estadísticos de pruebas de hipótesis, en el ámbito del aprendizaje de máquina y redes neuronales. Además, del interés en el estudio de las criptomonedas, pues es un tema muy relevante y ha adquirido especial importancia económica.

Los aportes del estudio serán todo el conjunto de pruebas que se llevarán a cabo, entre los desempeños de los diferentes algoritmos de aprendizaje de máquina y redes neuronales seleccionado. Además, el detalle de todo el proceso seguido, para su aplicación con otros conjuntos de datos.

Entre los beneficiarios de este trabajo, se pueden considerar a todas aquellas personas que busquen desarrollar modelos de predicción para el precio de criptomonedas. Así también, se pueden contemplar a investigadores que busquen comparar de forma estadística un conjunto de algoritmos con datos de un dominio o tema específico.

La relevancia social de este trabajo en el ámbito laboral radica en la importancia económica que ha adquirido el tema de criptomonedas. Pues el desarrollo de modelos de aprendizaje de máquina ya no es un tema solo de interés académico, sino también de interés económico y laboral.

5. OBJETIVOS

5.1. General

Comparar algoritmos de *machine learning* aplicados a la predicción del valor del precio de criptomonedas, a través de pruebas estadísticas de contraste y *post hoc*, para seleccionar aquellos con el mejor desempeño.

5.2. Específicos

- Estimar el desempeño de los algoritmos, calculando su precisión y eficiencia en varios conjuntos de datos de precios de criptomonedas, para realizar las comparaciones correspondientes.
- Identificar la distribución que siguen las métricas de desempeño calculadas, utilizando análisis de bondad de ajuste, para inferir el tipo de pruebas que deben ser aplicadas en la comparación de los algoritmos.
- Cuantificar las diferencias en el desempeño de los algoritmos, utilizando las pruebas de contraste y *post hoc* adecuadas, para distinguir los algoritmos que muestran estadísticamente un desempeño superior.

6. NECESIDADES A CUBRIR Y ESQUEMA DE SOLUCIÓN

La comparación estadística del desempeño de los algoritmos de *machine learning* permitirá obtener evidencia empírica sobre los algoritmos más idóneos, para ser utilizados en el dominio de aplicación de predicción de series temporales con datos del mercado de criptomonedas. Este generará una referencia para futuros proyectos o investigaciones, enfocados en este ámbito.

Para la solución, inicialmente se deben definir de forma concreta los algoritmos que se van a comparar, a través de una búsqueda exhaustiva; se identificarán los algoritmos de *machine learning* más novedosos y populares aplicados en materia de predicción de series temporales. La recolección de los datos se hará a través de la consulta a las bases de datos en la plataforma Binance, que contienen la información histórica de las transacciones de criptomonedas desde el año 2017 hasta la actualidad.

Una vez definido lo anterior, se procederá a la solución estadística. Para esto, en primer lugar se debe buscar minimizar las fuentes de variación aleatoria. Una de las principales acciones para lograrlo es la selección de un *dataset* suficientemente grande, que nos permita mitigar en lo posible el error generado en la división entre el *train/test dataset*. A continuación, se debe efectuar una optimización simple de los modelos para cada uno de los algoritmos y datos seleccionados. Una vez obtenidos, se procede a la evaluación de la precisión, y para ello se utilizarán el RMSE, MAE y MAPE. También se utilizarán las métricas de eficiencia, entre ellas el tiempo de ejecución de los algoritmos.

Una vez obtenidas las métricas de desempeño (precisión y eficiencia), se procede a aplicar las pruebas de bondad de ajuste a estas: Shapiro-Wilkes, Kolmogórov-Smirnov y Anderson-Darling. Finalmente, con los resultados anteriores se procederá a realizar las pruebas de contraste necesarias, ANOVA o Kruskal-Wallis. Si se rechaza la hipótesis nula de estas, se aplicarán las pruebas *post hoc* de Duncan o Nemenyi, para hallar el grado de las diferencias en el desempeño de los algoritmos.

7. MARCO TEÓRICO

7.1. Fundamentos estadísticos

En los siguientes incisos se describen los fundamentos estadísticos necesarios para el desarrollo de este trabajo de investigación.

7.1.1. Estadística

La estadística es “la ciencia que recoge, organiza, presenta, analiza e interpreta datos con el fin de propiciar una toma de decisiones más eficaz” (Lind, Wathen y Marchal, 2011, p. 5). Y en términos generales podemos dividir a la estadística en dos grandes ramas: estadística descriptiva y estadística inferencial.

7.1.1.1. Estadística descriptiva

Como indica Devore (2010), un investigador usualmente desea resumir y describir las características importantes de cierto grupo de datos. Esto implica métodos de cálculo como medidas de tendencia central y de variabilidad, así como otros métodos más gráficos, tales como histogramas, diagramas de caja, entre otros. Todos estos son parte de lo que conocemos como estadística descriptiva, la cual consiste, según Lind, *et al.* (2011), en “... métodos para organizar, resumir y presentar datos de manera informativa” (p. 7).

7.1.1.2. Estadística inferencial

Aguilar, Altamira y García (2009) mencionan que “el objetivo de la inferencia estadística es construir estimaciones y pruebas de hipótesis acerca de las características de una población mediante la información de una muestra” (p. 3). Asimismo, Devore (2010) indica que aquellos métodos que implican el generalizar conclusiones de una población, basados en una muestra (subgrupo de la población), forman lo que conocemos como estadística inferencial.

Estos métodos, pertenecientes a la estadística inferencial, también los podemos dividir en dos ramas, basados en lo que se asume de las poblaciones, siendo estos: métodos paramétricos (se asume, entre varias cosas, la normalidad de los datos) y no paramétricos (no se asume la normalidad).

7.1.1.2.1. Métodos paramétricos

Cómo indica Walpole, Myers y Myers (2012), se conoce como métodos paramétricos a aquellos que se basan en la suposición de que la muestra se selecciona aleatoriamente de una población normal. Estos métodos suelen ser lo suficientemente potentes para seguir siendo confiables, incluso en situaciones en los que se tienen ligeras desviaciones de la normal, sobre todo si la muestra es lo suficientemente grande.

Asimismo, Corder y Foreman (2014) indican que en los métodos paramétricos se debe asumir lo siguiente con respecto a la muestra:

- Se toman aleatoriamente de una población distribuida de forma normal

- Consisten en observaciones independientes, a excepción de los datos pareados
- Consiste en valores de escala de intervalo o de razón
- Las respectivas poblaciones tienen varianzas aproximadamente iguales
- Son adecuadamente grandes
- Se aproximan a una distribución normal

7.1.1.2.2. Métodos no paramétricos

Walpole *et al.* (2012) indican: “La mayoría de los analistas consideran que los cálculos involucrados en los métodos no paramétricos son muy atractivos e intuitivos” (p. 655). Esto debido a que en su mayoría los métodos no paramétricos se basan en el estudio de rangos.

Ahora bien, es necesario definir a que nos referimos con métodos no paramétricos o también conocido como métodos de distribución libre. Como se mencionó en secciones anteriores, tradicionalmente se les conoce como métodos paramétricos a aquellas pruebas en las que se asume la normalidad de las muestras y la homogeneidad de las varianzas. Si bien estos métodos pueden llegar a ser usados en casos en los que la distribución se desvíe ligeramente de la normal, en las situaciones en las que claramente se violan estas condiciones, se deben aplicar los métodos no paramétricos.

7.1.2. Pruebas de normalidad

En la investigación es usual identificar una población que se desea estudiar y tomar varias mediciones aleatorias e independientes de una variable particular de esta. A este conjunto de mediciones se le suele llamar muestra. Si nuestra técnica experimental es buena y nuestra muestra representa

adecuadamente la población de estudio, esta se puede utilizar para hacer diferentes inferencias sobre nuestra población. (Corder y Foreman, 2014)

En algunos métodos estadísticos (estadística paramétrica) se asume que la población de la que se toma la muestra es normal, por lo tanto, para su aplicación es necesario un análisis que nos permita establecer con cierta confianza o probabilidad, si nuestra muestra se asemeja a una distribución normal. Esto se puede hacer a través de pruebas conocidas como pruebas de bondad de ajuste, las cuales nos brindan el ajuste de una muestra a una distribución específica, en este caso, con la normal.

7.1.2.1. Análisis de curtosis y sesgo

De acuerdo con Corder y Foreman (2014), una distribución de frecuencias que se parezca a una curva normal es aproximadamente normal. Sin embargo, no todas las distribuciones se aproximan a una normal. Las curvas pueden estar concentradas en el centro o sus datos pueden estar dispersos a los lados, o pueden ser simétricas respecto al centro o tener concentrados sus datos de un lado u otro; a estas condiciones se les conoce como curtosis y sesgo, respectivamente.

La curtosis es una medida que identifica que tan plana o picuda es la distribución con respecto a la distribución normal; dicho de otra forma, la curtosis se refiere a que tan concentrados están los valores en el centro de la distribución. Su cálculo se puede efectuar de la siguiente forma:

$$K = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4 \right] - \frac{3(n-1)^2}{(n-2)(n-3)} \quad (\text{Ec. 01})$$

$$SE_K = \sqrt{\frac{24n(n-1)^2}{(n-2)(n+1)(n+5)(n+3)}} \quad (\text{Ec. 02})$$

$$Z_K = \frac{K}{SE_K} \quad (\text{Ec. 03})$$

Donde:

K : curtosis

SE_K : error estándar de curtosis

Z_K : z-score para la curtosis

El sesgo es una medida de la simetría horizontal, que tiene una distribución con respecto a la distribución normal. Su cálculo se puede efectuar de la siguiente forma:

$$S_k = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3 \quad (\text{Ec. 04})$$

$$SE_{S_k} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}} \quad (\text{Ec. 05})$$

$$Z_{S_k} = \frac{S_k}{SE_{S_k}} \quad (\text{Ec. 06})$$

Donde:

S_k : sesgo

SE_{S_k} : error estándar del sesgo

Z_{S_k} : z-score para el sesgo

Para saber si nuestra distribución cumple con el esperado de normalidad, se deben comparar los valores obtenidos por las ecuaciones 05 y 06, y compararlos con el valor Z de la distribución normal a una confianza dada.

7.1.2.2. Prueba de Kolmogórov-Smirnov

La prueba de Kolmogórov-Smirnov (K-S) es una prueba de bondad de ajuste. Según indican Kvam, Vidakovic y Kim (2022), la prueba K-S tiene ventajas sobre la de chi-cuadrado, que depende de un tamaño adecuado de muestra y de una asignación apropiada de intervalos para que la aproximación sea válida. No obstante, la prueba de K-S también tiene sus propias limitaciones. Una de estas limitaciones es que en teoría la prueba solo es válida para distribuciones continuas. Además, la prueba tiende a ser más sensible en el centro de la distribución que en las colas; en otras palabras, es muy susceptible a los valores atípicos. Otra limitación muy seria de la prueba de K-S la encontramos en que la distribución esperada debe ser completamente especificada.

Corder y Foreman (2014) indican que la prueba de K-S para una muestra compara las frecuencias acumuladas de dos distribuciones. Luego, utiliza la mayor divergencia para estimar con una prueba a dos colas, si ambas muestras son estadísticamente similares o diferentes.

Según Gutiérrez y Vladimirovna (2016), la prueba de K-S pretende probar si las observaciones de una muestra provienen de una función de distribución acumulada F_o , y para ello utiliza el siguiente estadístico:

$$D = \max\{D^+, D^-\} \quad (\text{Ec. 08})$$

$$D^+ = \max\left\{\frac{i}{n} - F_o(\bar{x}_i)\right\} \quad (\text{Ec. 09})$$

$$D^- = \max\left\{F_o(\bar{x}_i) - \frac{i}{n}\right\} \quad (\text{Ec. 10})$$

Su implementación en el lenguaje de programación Python se puede hacer a través del paquete Scipy. Asimismo, su implementación en el lenguaje de programación R, es a través del paquete base y con el comando *ks.test*.

7.1.2.3. Prueba de Anderson-Darling

La prueba de Anderson-Darling (A-D) es una prueba especializada de normalidad. Según indican Kvam, Vidakovic y Kim (2022), esta fue creada con base en la prueba de Kolmogórov-Smirnov, pero con el objetivo de modificarla para mejorar su desempeño con la distribución de interés, en este caso la normal. A diferencia de la prueba K-S, la prueba de A-D presta mayor atención a las colas. La ventaja de esto es que se hace más exacta la prueba, con la desventaja de que los valores críticos deben ser calculados para cada distribución probada.

Según Gutiérrez y Vladimirovna (2016), la prueba de A-D debe probar si las observaciones provienen de una función de distribución acumulada propuesta F_o , y para ello se puede utilizar el siguiente estadístico:

$$AD = -n - \sum_{i=1}^n \frac{(2i-1)}{n} \{\ln[F_o(\bar{x}_i)] + \ln[1 - F_o(\bar{x}_{n+1-i})]\} \quad (\text{Ec. 11})$$

7.1.2.4. Prueba de Shapiro-Wilk

La prueba de Shapiro-Wilk es una prueba especializada para estimar la normalidad de una muestra. Por tal motivo, según indican Kvam, Vidakovic y Kim (2022), la prueba de Shapiro-Wilk supera por mucho a la prueba de Kolmogórov-Smirnov e incluso a otras pruebas, cuando se están estudiando datos normales.

Esta prueba se basa en el cálculo del estadístico:

$$W = \frac{(\sum_{i=1}^n a_i X_{i:n})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (\text{Ec. 12})$$

$$a = M'V^{-1}((M'V^{-1})(V^{-1}M))^{-1/2} \quad (\text{Ec. 13})$$

7.1.3. Diseño de experimentos

Para el diseño del experimento se necesitan algunos parámetros que se presentan a continuación.

7.1.3.1. Paramétricos

Para el diseño se iniciará realizando el análisis de varianza de un factor, tal como se describe en el inciso 7.1.3.1.1.

7.1.3.1.1. Análisis de varianza de un factor (ANOVA)

Como indican Herzog, Francis y Clarke (2019), hay situaciones en las que se encuentra en la necesidad de comparar más de dos muestras; sin embargo, al utilizar las pruebas de contraste para dos muestras en pares, se tiene el riesgo de aumentar el error Tipo I, mientras el número de comparaciones aumenta. En estos casos, se puede utilizar el Análisis de Varianza o ANOVA. Para evitar esto, el ANOVA simplifica la hipótesis alternativa en analizar si se tiene como mínimo una muestra que difiere de las otras. Y al igual que en los métodos paramétricos de comparación de dos

muestras, el ANOVA asume la normalidad y homogeneidad de varianzas en las muestras.

Tal como indica Devore (2010), el ANOVA de un solo factor se utiliza para comparar más de dos poblaciones o medias de tratamientos y utiliza las siguientes hipótesis:

- $H_o: \mu_1 = \mu_2 = \dots = \mu_I$
- H_a : Una o más medias son diferentes

Según Walpole *et al.* (2012), cada observación del ANOVA puede modelarse con la siguiente ecuación:

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad (\text{Ec. 14})$$

Donde:

ϵ_{ij} : desviación que tiene la j-ésima de la i-ésima muestra

μ : media general para todas las μ_i

α_i : es el efecto del i-ésimo tratamiento.

7.1.3.2. No paramétricos

Para definir los datos no paramétricos se realizará la prueba de Friedman.

7.1.3.2.1. Prueba de Friedman

Esta es una prueba de tipo no paramétrico de contraste para muestras pareadas. Kvan y Vidakovic (2007) mencionan que la prueba de Friedman es

una alternativa no paramétrica para un diseño de bloques aleatorizados (RBD por sus siglas en inglés) en un ANOVA regular y que esta reemplaza el RBD cuando las asunciones de normalidad están en cuestión o cuando las varianzas posiblemente difieren de población en población.

Corder y Foreman (2014) indican que para calcular el estadístico F_r en la prueba de Friedman, cuando no hay empates, se debe utilizar la siguiente fórmula:

$$F_r = \left[\frac{12}{nk(k+1)} \sum_{i=1}^k R_i^2 \right] - 3n(k+1) \quad (\text{Ec. 15})$$

Donde:

n : número de filas, o sujetos.

k : es el número de columnas o condiciones.

R_i : suma de los rangos por columna o condición

Además, agregan que en el caso que hallan empates, se debe hacer la siguiente corrección:

$$F_r = \frac{n(k-1) \left[\sum_{i=1}^k \frac{R_i^2}{n} - C_F \right]}{\sum r_{ij}^2 - C_F} \quad (\text{Ec. 16})$$

$$C_F = \frac{1}{4} nk(k+1)^2 \quad (\text{Ec. 17})$$

Donde:

C_F : es la corrección por empates

r_{ij} : rango correspondiente al sujeto j en la columna i

Finalmente, mencionan que el resultado se debe comparar con la distribución chi-cuadrado, para comprobar la prueba de hipótesis.

7.1.3.2.2. Prueba de Kruskal-Wallis

La prueba de Kruskal-Wallis, también conocida como prueba H, es una prueba no paramétrica de contraste para muestras independientes. Esta "... utiliza rangos de datos de tres o más muestras aleatorias simples independientes para someter a prueba la hipótesis nula de que las muestras provienen de poblaciones con medianas iguales" (Triola, 2012, p. 686). Esta busca probar la hipótesis alternativa de que las poblaciones tienen medianas que no son iguales.

Como indican Kvam y Vidakovic (2007): "*The Kruskal-Wallis (KW) test is a logical extension of the Wilcoxon-Mann-Whitney test. It is a nonparametric test used to compare three or more samples*" (p. 141). Lo anterior se traduce como: La prueba de Kruskal-Wallis (KW) es una extensión lógica de la prueba de Wilcoxon-Mann-Whitney. Es una prueba no paramétrica utilizada para comparar tres o más muestras.

Esta prueba es un análogo de la prueba F utilizada en el ANOVA, y utiliza un diseño de la siguiente forma:

Tabla I. **Diseño de una prueba de Kruskal-Wallis**

Muestra 1	$X_{11,}$	$X_{12,}$...	$X_{1, n1}$
Muestra 2	$X_{21,}$	$X_{22,}$...	$X_{2, n2}$
.
.
Muestra k - 1	$X_{k-1,1,}$	$X_{k-1,2,}$...	$X_{k-1, nk-1}$
Muestra k	X_{k1}	X_{k2}	...	$X_{k, nk}$

Fuente: elaboración propia, hecho con Microsoft Excel 365.

De acuerdo con Corder y Foreman (2014), para evaluar la prueba primero es necesario combinar las muestras y hallar sus rangos en conjunto, para después calcular el siguiente estadístico:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1) \quad (\text{Ec. 18})$$

Donde:

N : número total de valores de las muestras combinadas

R_i : suma de los rangos de una muestra en particular

n_i : número de valores de la correspondiente suma de rangos

Además, el estadístico H de la prueba Kruskal Wallis “se aproxima muy bien mediante una distribución chi cuadrada ... cuando H_0 es verdadera, siempre y cuando cada muestra conste de al menos 5 observaciones” (Walpole *et al.* 2012, p. 668). Cabe resaltar también que los grados de libertad en la prueba están definidos por el número de grupos k , siendo estos:

$$df = k - 1 \quad (\text{Ec. 19})$$

Corder y Foreman (2014) indican que cuando los valores tienen empates, se debe aplicar una corrección y hallar un nuevo estadístico, esta corrección está dada por la siguiente ecuación.

$$C_H = 1 - \frac{\sum(T^3 - T)}{N^3 - N} \quad (\text{Ec. 20})$$

Donde:

N : número total de valores de las muestras combinadas

T : número de empates

7.1.4. Pruebas post hoc

Son pruebas que se hacen después de rechazar la hipótesis nula en una prueba de hipótesis de comparación de múltiples muestras.

7.1.4.1. Paramétricas

Cuando se rechaza la hipótesis nula en un ANOVA se debe aplicar una prueba *post hoc* para hallar el nivel de diferencias que hay entre las muestras. Existen diferentes métodos para ellos, entre los que se puede mencionar el método de Tukey.

7.1.4.1.1. Método de Tukey

Gutiérrez y De la Vara (2008) indican que este método realiza el siguiente cálculo:

$$T_\alpha = q_\alpha(k, N - k) \sqrt{CM_E/n_i} \quad (\text{Ec. 21})$$

Donde:

CM_E : cuadrado medio del error

n : número de observaciones por tratamiento

k : número de tratamientos

$N - k$: grados de libertad

α : nivel de significancia deseado

7.2. Criptomonedas y machine learning

En los siguientes incisos se realiza una breve descripción acerca de las criptomonedas y *machine learning*.

7.2.1. Criptomonedas

Las criptomonedas actuales se basan en la tecnología del *blockchain*. Como lo indica Campbell-Verduyn (2019), en su esencia esta tecnología son secuencias numéricas digitales codificadas en programas informáticos que permiten el intercambio, el registro y la difusión segura de las transacciones individuales de los usuarios desde cualquier lugar del mundo con acceso a internet.

La tecnología del *blockchain* utiliza la encriptación digital que, en grados variables, ocultan tanto la identidad de los usuarios individuales como el contenido preciso transferido. Los algoritmos, series precodificadas de instrucciones paso a paso, también se movilizan para resolver complejas ecuaciones matemáticas y llegar a un consenso sobre la validez de las transacciones dentro de las redes de usuarios. Las tecnologías de sellado de tiempo agrupan periódicamente las transacciones verificadas en conjuntos de datos o bloques. Estos bloques, unidos de forma secuencial, forman cadenas

que constituyen bases de datos de transacciones más amplias, llamadas *blockchain*.

7.2.1.1. Blockchain

Snyers y Houben (2018) mencionan que el *blockchain* es un subtipo de tecnología basada en DLT (*Distributed Ledger Technology*), que es una forma de grabar, compartir, mantener y controlar información a través de múltiples nodos de una red. Esta emplea algoritmos matemáticos de encriptación para crear y verificar un continuo crecimiento en las estructuras de datos, las cuales solo pueden ser creados, pero no removidos.

7.2.1.2. Criptomonedas con la mayor capitalización del mercado

Entre algunas de las principales criptomonedas con la mayor capitalización de mercado, así como la mayor importancia, se pueden mencionar las siguientes: BTC, ETH, USDT, BNB, USDC, XRP, BUSD, ADA, XLM, XMR.

7.2.1.2.1. Bitcoin (BTC)

Como indican Snyers y Houben (2018), el bitcoin es una moneda virtual y descentralizada, que utiliza un mecanismo llamado *mining* (minado). En este mecanismo, una persona pone a disposición voluntariamente su computadora a la red de bitcoin, para que esta resuelva problemas matemáticos complejos (necesarios para generar los bloques utilizados por esta criptomoneda) y recibir bitcoins a cambio de la resolución de estos problemas.

7.2.1.2.2. Ethereum (ETH)

Como indican Snyers y Houben (2018), el ethereum es una plataforma descentralizada lanzada en julio de 2015 y que utiliza contratos inteligentes. En teoría, la plataforma no es una criptomoneda, sin embargo, esta requiere una forma de incentivo para validar las transacciones dentro de la red, y aquí es donde la criptomoneda nativa de la plataforma, el ether, entra en juego. Esta criptomoneda permite construir los contratos inteligentes de la plataforma, y también funciona como medio de intercambio.

7.2.1.2.3. Tether (USDT)

Es una criptomoneda que funcionan a través de la tecnología *blockchain* como otras monedas digitales, sin embargo, estos están ligados 1 a 1 al precio del dólar, lo que la hace más estables que otras criptomonedas. Fueron lanzados en 2014, siendo pioneros en el modelo de monedas estables.

7.2.1.2.4. BNB (BNB)

BNB es la criptomoneda nativa utilizada en la cadena del ecosistema BNB de la empresa Binance. Esta criptomoneda fue lanzada el 26 de junio de 2017 y permite efectuar transacciones en la red BNB, así como el pago en línea de ciertos bienes y servicios. (Binance Academy, 2018)

7.2.1.2.5. USDC (USDC)

La USDC es una criptomoneda *open-source* respaldada por el precio del dólar, pues tiene un valor 1 a 1. Por lo tanto, se considera como una moneda de precio estable; no obstante, es sensible a la inflación del dólar. (Picardo, 2022)

7.2.1.2.6. Ripple (XRP)

Ripple, lanzada en 2012, es una plataforma P2P, *open-source* y descentralizada, creada para pagos digitales, que permite las transferencias instantáneas independientemente de la moneda de uso. La criptomoneda utilizada por esta plataforma es el XRP, la cual fue construida como una moneda puente, que permite a las instituciones financieras ejecutar pagos internacionales más barato y rápido. (Snyers y Houben, 2018)

7.2.1.2.7. BUSD (BUSD)

El BUSD es una criptomoneda del grupo de las monedas de precio estable, y cuyo valor está ligado 1 a 1 con el precio del dólar. Fue creada por las compañías Paxos y Binance.

7.2.1.2.8. Cardano (ADA)

Cardano es una plataforma de contratos inteligentes y aplicaciones descentralizadas, comenzó en 2015 y fue lanzada oficialmente en septiembre de 2017. Utiliza la criptomoneda ADA. (Snyers y Houben, 2018)

7.2.1.2.9. Bitcoin cash (BCH)

Bitcoin cash, es una moneda digital P2P descentralizada, creada el 1 de agosto de 2017, nació para reducir los costes de las transacciones y mejorar los tiempos de confirmación. (Snyers y Houben, 2018)

7.2.1.2.10. Litecoin (LTC)

Litecoin (LTC) es una criptomoneda P2P descentralizada y *open-source*, lanzada en octubre de 2011. Se diferencia con el bitcoin en que ofrece velocidades de transacciones más rápidas, y que tiene un límite de suministro mayor, 84 millones de monedas comparadas con los 21 millones del bitcoin. (Snyers y Houben, 2018)

7.2.1.2.11. Stellar (XLM)

Stellar es una plataforma *open-source* que brinda una infraestructura distribuida para pagos, fue creada en 2014. La criptomoneda utilizada en la red de stellar es el lumen (XLM), y que en teoría permite mover dinero alrededor del mundo y efectuar transacciones entre diferentes monedas de manera rápida y segura. (Snyers y Houben, 2018)

7.2.1.2.12. Monero (XMR)

Monero (XMR) es una criptomoneda *open-source* P2P, que se enfoca principalmente en la completa anonimidad de las transacciones. (Snyers y Houben, 2018)

7.2.2. Machine learning

Machine learning o aprendizaje automático, como indica Géron (2019), es la ciencia que se encarga del estudio del proceso de programar computadoras (máquinas), para que estas puedan aprender de los datos. Es decir, que es el proceso de extracción de patrones de los datos, para construir modelos usualmente utilizados en aplicaciones de predicción de datos.

Kelleher, Namee y D'Arcy (2015) mencionan que los algoritmos de *machine learning* funcionan buscando a través de un conjunto de posibles modelos de predicción, aquel modelo que mejor capture la relación entre las características de las variables explicativas y la característica o variable respuesta de estos.

Los tipos de aprendizaje automático, según las características que se describan, son los siguientes:

- Según la supervisión humana:
 - Aprendizaje supervisado
 - Aprendizaje no supervisado
- Según la continuidad del aprendizaje:
 - *Online*
 - *Batch*
- Según las comparaciones que hagan:
 - Basados en instancias
 - Basados en modelos

Géron (2019) indica que en general, y dependiendo de la cantidad de supervisión que los algoritmos reciben durante el entrenamiento, estos también se pueden clasificar en cuatro tipos:

- Supervisado
- No supervisado
- Semi supervisado
- Reforzado

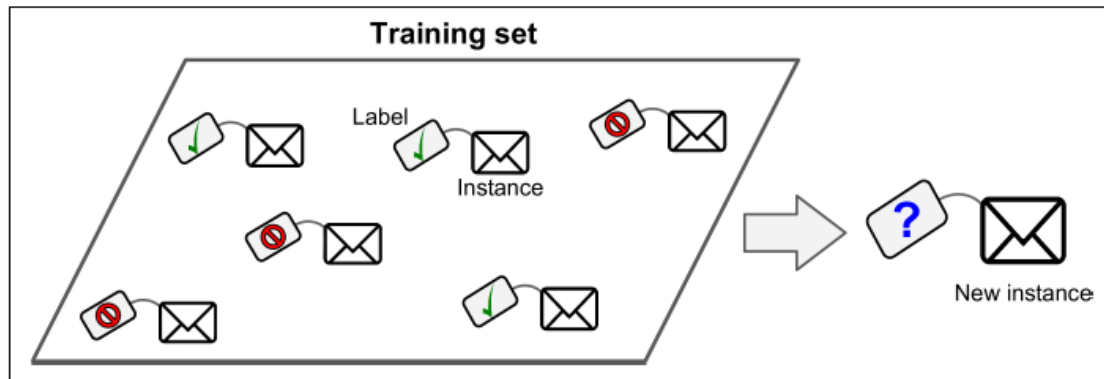
7.2.2.1. Aprendizaje supervisado

Consoli, Recupero y Saisana (2021) indican que estos algoritmos emplean un conjunto de valores y respuestas como entradas, referidos como *dataset* de entrenamiento, y un conjunto valores y respuestas para validar las predicciones, conocidos como *dataset* de prueba. Estos datos son usualmente contruidos a través de muestreos aleatorios de un grupo de datos inicial.

Géron (2019) menciona que se pueden tener dos tipos de tareas en el aprendizaje supervisado:

- Clasificación: que se obtienen etiquetas para los datos. Un claro ejemplo es la clasificación de correos como correo no deseado.

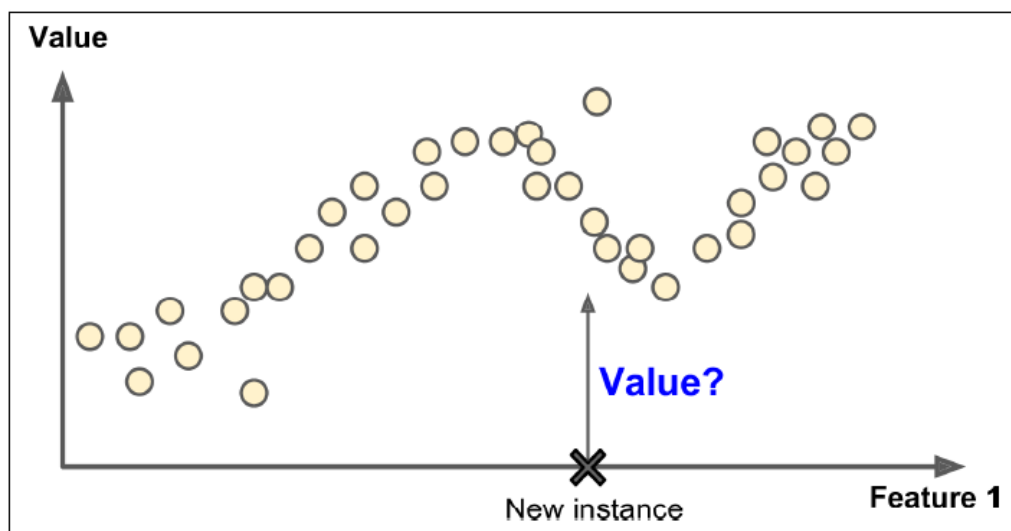
Figura 1. Diagrama algoritmo de clasificación de aprendizaje supervisado



Fuente: Géron (2019). *Hands-on machine learning with scikit-learn, keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems.*

- Regresión: esta consiste predecir un valor utilizando un grupo de valores llamados predictores. Para poder entrenar a estos es necesario tener tanto los valores predictores como su valor. Un ejemplo es el cálculo del precio de un carro basado en sus características.

Figura 2. **Diagrama algoritmo de regresión de aprendizaje supervisado**



Fuente: Géron (2019). *Hands-on machine learning with scikit-learn, keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems.*

Entre algunos de los algoritmos más importantes de aprendizaje supervisado se pueden mencionar. (Géron, 2019)

- k-Nearest Neighbors
- Regresión lineal
- Regresión logística
- Support Vector Machines (SVMs)
- Árboles de decisión y bosques aleatorios

- Redes neuronales

7.2.2.2. Aprendizaje no supervisado

Consoli y Recupero y Saisana (2021) indican que los algoritmos de aprendizaje no supervisado tienen como objetivos inferir las propiedades de un conjunto de datos, sin necesidad de tener las respuestas correctas o el error en cada observación. James, Witten, Hastie y Tibshirani (2013) hablan sobre los retos del aprendizaje no supervisado, y mencionan que:

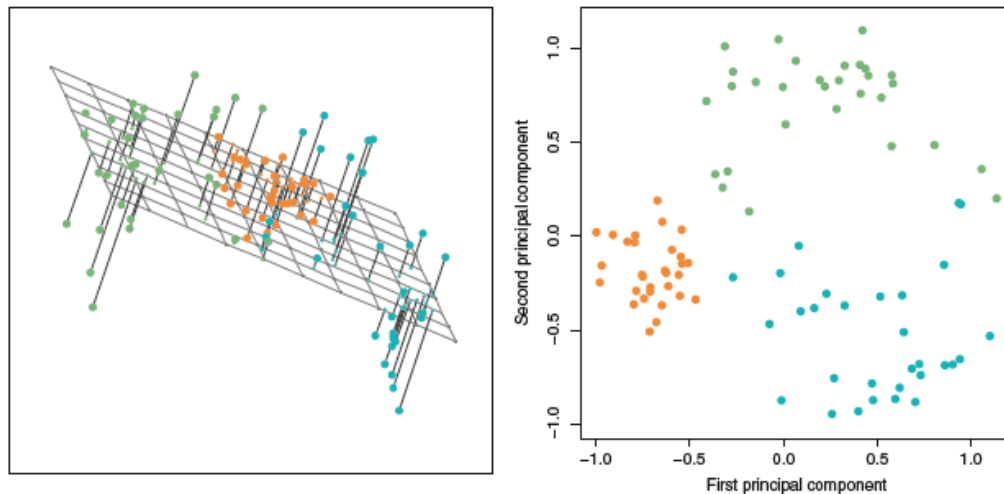
Unsupervised learning is often performed as part of an exploratory data analysis. Furthermore, it can be hard to assess the results obtained from unsupervised learning methods, since there is no universally accepted mechanism for performing cross validation or validating results on an independent data set. (p. 374). Traduciéndolo se lee: El aprendizaje no supervisado es usualmente parte de un análisis exploratorio de los datos. Por consiguiente, puede ser difícil evaluar los resultados obtenidos a través de algoritmos de aprendizaje no supervisados, dado que no hay un mecanismo universal aceptado para realizar validación cruzada o validación de los resultados en un conjunto de datos independiente.

Entre algunos de los algoritmos más importantes de aprendizaje no supervisado se pueden mencionar. (Géron, 2019)

- *Clustering*
 - *k-means*
 - DBSCAN
 - Análisis jerárquico de clúster (HCA)
- Detección anómala y detección novedosa

- SVM de una clase
- *Isolation Forest*
- Visualización y reducción de dimensión
 - Análisis de componentes principales (PCA)
 - Kernel PCA
 - *Locally-Linear Embedding* (LLE)
 - *t-distributed Stochastic Neighbor Embedding* (t-SNE)
- Algoritmos de regla de asociación
 - A priori
 - Eclat

Figura 3. **Visualización de algoritmos de aprendizaje no supervisado (PCA)**



Fuente: James, Witten, Hastie y Tibshirani (2013). *An introduction to statistical learning: With applications in R*.

7.2.2.3. Aprendizaje semisupervisado

Estos se utilizan para lidiar con datos que en su mayoría no poseen etiqueta o valor respuesta. Y usualmente, estos algoritmos son combinaciones de algoritmos supervisados y no supervisados. (Géron, 2019)

7.2.2.4. Aprendizaje reforzado

Géron (2019) indica lo siguiente sobre el sistema que utiliza el aprendizaje reforzado: *“an agent in this context, can observe the environment, select and perform actions, and get rewards in return. It must then learn by itself what is the best strategy, called a policy, to get the most reward over time”*. (p. 8-9).

La traducción anterior indica: Un agente, en este contexto, es capaz de observar el ambiente, seleccionar y realizar acciones para obtener recompensas. Y después, aprender por sí mismo, la mejor estrategia, llamada política, para obtener las mayores recompensas en el tiempo.

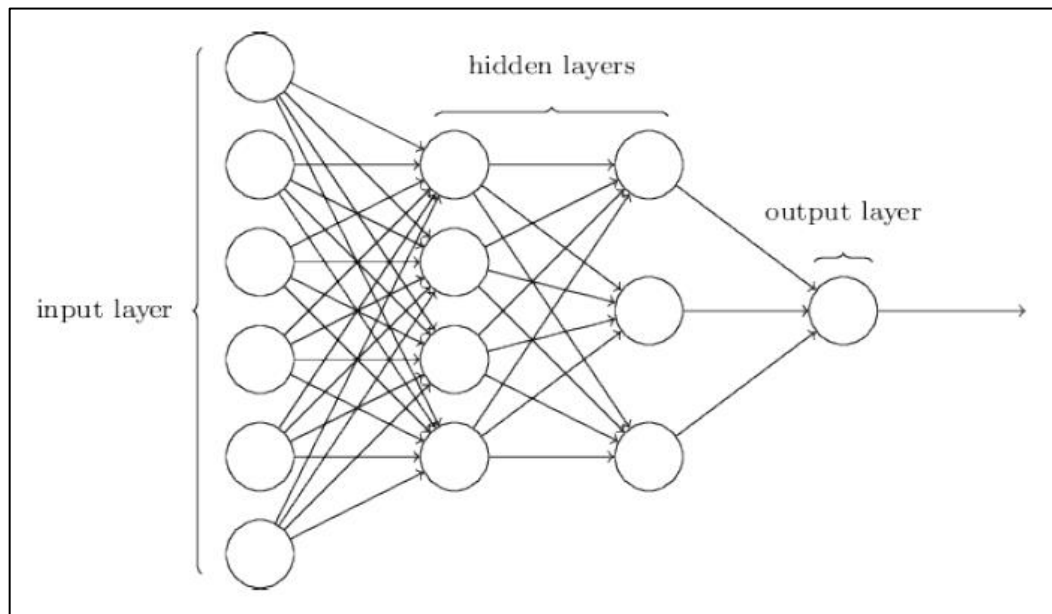
7.2.3. Redes neuronales

Haykin (2009) indica que *“a neural network is a massively parallel distributed processor made up of simple processing units that has a natural propensity for storing experiential knowledge and making it available for use”*. (p. 28).

Traduciéndolo dice que: una red neuronal es un procesador paralelo distribuido formado por unidades simples de procesamiento que pueden almacenar conocimiento y hacerlo disponible para su uso. También menciona

que la red neuronal se asemeja al cerebro, en que esta adquiere el conocimiento a través de un algoritmo, y que el peso de sus conexiones es utilizado para almacenar el conocimiento adquirido.

Figura 4. **Diagrama red neuronal**



Fuente: Nielsen (2018). *Neural Networks and Deep Learning*.

8. PROPUESTA DE ÍNDICE DE CONTENIDOS

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES

LISTA DE SÍMBOLOS

GLOSARIO

RESUMEN

PLANTEAMIENTO DEL PROBLEMA

OBJETIVOS

MARCO METODOLÓGICO

INTRODUCCIÓN

1. MARCO REFERENCIAL

2. MARCO TEÓRICO

2.1. Fundamentos estadísticos

2.1.1. Estadística

2.1.1.1. Estadística descriptiva

2.1.1.2. Estadística inferencial

2.1.1.2.1. Métodos paramétricos

2.1.1.2.2. Métodos no paramétricos

2.1.2. Pruebas de normalidad

2.1.2.1. Análisis de curtosis y sesgo

2.1.2.2. Prueba de Kolmogórov-Smirnov

2.1.2.3. Prueba de Anderson-Darling

2.1.2.4. Prueba de Shapiro-Wilk

2.1.3. Diseño de experimentos

- 2.1.3.1. Paramétricos
 - 2.1.3.1.1. Análisis de varianza de un factor (ANOVA)
- 2.1.3.2. No paramétricos
 - 2.1.3.2.1. Prueba de Friedman
 - 2.1.3.2.2. Prueba de Kruskal-Wallis
- 2.1.4. Pruebas *post hoc*
 - 2.1.4.1. Paramétricas
 - 2.1.4.1.1. Método de Tukey
- 2.2. Criptomonedas y *machine learning*
 - 2.2.1. Criptomonedas
 - 2.2.1.1. *Blockchain*
 - 2.2.1.2. Criptomonedas con la mayor capitalización del mercado
 - 2.2.1.2.1. Bitcoin (BTC)
 - 2.2.1.2.2. Ethereum (ETH)
 - 2.2.1.2.3. Tether (USDT)
 - 2.2.1.2.4. BNB (BNB)
 - 2.2.1.2.5. USDC (USDC)
 - 2.2.1.2.6. Ripple (XRP)
 - 2.2.1.2.7. BUSD (BUSD)
 - 2.2.1.2.8. Cardano (ADA)
 - 2.2.1.2.9. Bitcoin Cash (BCH)
 - 2.2.1.2.10. Litecoin (LTC)
 - 2.2.1.2.11. Stellar (XLM)
 - 2.2.1.2.12. Monero (XMR)
 - 2.2.2. *Machine learning*
 - 2.2.2.1. Aprendizaje supervisado
 - 2.2.2.2. Aprendizaje no supervisado

2.2.2.3. Aprendizaje semisupervisado

2.2.2.4. Aprendizaje reforzado

2.2.3. Redes neuronales

3. RESULTADOS

4. DISCUSIÓN DE RESULTADOS

CONCLUSIONES

RECOMENDACIONES

REFERENCIAS

ANEXOS

9. METODOLOGÍA

9.1. Características del estudio

El enfoque del estudio propuesto es cuantitativo. Esto debido a que se analizará el desempeño, cuantificado la precisión y eficiencia de la predicción de un conjunto de algoritmos de *machine learning*.

El diseño de la investigación es cuasiexperimental, pues se analizará la información histórica del precio de criptomonedas. Por lo tanto, no se influirá ni manipulará su valor. No obstante, si se debe garantizar el control de ciertas fuentes de variación aleatoria en la comparación de los algoritmos. La investigación, además es transversal debido a que solamente se tomará la información histórica de 2018 a 2022.

El alcance es descriptivo-correlacional, pues en la investigación se busca describir el desempeño de los modelos predictivos propuestos, y a la vez comparar su comportamiento simultáneo para encontrar diferencias significativas entre ellos.

9.2. Unidad de análisis

La población de estudio corresponde a algoritmos de *machine learning*, aplicados a la predicción del precio de criptomonedas con datos de los años entre 2018 y 2022, de los cuales se busca estudiar, describir y comparar su desempeño cuantificado a través de su precisión y su eficiencia.

9.3. Variables

En la tabla II se describen las variables del estudio.

Tabla II. **VARIABLES DEL ESTUDIO**

Variable	Definición teórica	Definición operativa	Escala
Precio (P_o)	Valor de mercado de la criptomoneda al inicio de un intervalo de tiempo analizado.	Es el valor real en dólares que tiene la criptomoneda al inicio de un intervalo de tiempo determinado.	De razón
Precio (P_f)	Valor de mercado de la criptomoneda al final de un intervalo de tiempo analizado.	Es el valor real en dólares que tiene la criptomoneda al final de un intervalo de tiempo determinado.	De razón
Precio (P_{min})	Valor mínimo de mercado de la criptomoneda en un intervalo de tiempo analizado.	Es el valor real mínimo en dólares que tiene la criptomoneda en un intervalo de tiempo determinado.	De razón
Precio (P_{max})	Valor máximo de mercado de la criptomoneda en un intervalo de tiempo analizado.	Es el valor real máximo en dólares que tiene la criptomoneda en un intervalo de tiempo determinado.	De razón
Fecha de análisis (Fa)	Ubicación temporal de análisis	Es el valor de la fecha y hora en la que se analiza el precio y volumen de transacciones.	De razón
Volumen de Transacciones (V)	Es el número de operaciones financieras de una criptomoneda en un tiempo específico.	Es el número de transacciones llevadas a cabo en una criptomoneda en una fecha de análisis determinada.	De razón
Raíz del Error cuadrático medio ($RMSE$)	Es la raíz de la media cuadrática de los valores predichos por un modelo y los valores observados.	Es una medida de precisión para comparar modelos en un conjunto particular de datos.	De razón
Error absoluto medio (MAE)	Es el promedio aritmético del valor absoluto de los valores predichos de un modelo y los valores observados.	Es una medida de error entre dos observaciones relacionadas a un mismo fenómeno.	De razón
Error Porcentual Absoluto Medio ($MAPE$)	Es el promedio de las diferencias absolutas de un valor predicho frente al valor observado dividido por el valor observado.	Es una medida de precisión relativa de predicción.	De razón

Fuente: elaboración propia, hecho con Microsoft Excel 365.

9.4. Fases del estudio

El estudio consistirá en las siguientes fases:

9.4.1. Fase uno: revisión bibliográfica

En este paso se obtienen los antecedentes y fuentes bibliográficas que sustentarán todos los cálculos y procesos que se llevarán a cabo con la información, las cuales son: pruebas de normalidad, pruebas de contrastes y *post hoc* paramétricas, pruebas de contraste y *post hoc* no paramétricas, para la parte estadística; algoritmos y modelos de *machine learning*, modelos predictivos para criptomonedas, en la parte específica.

9.4.2. Fase dos: recolección y limpieza de los datos

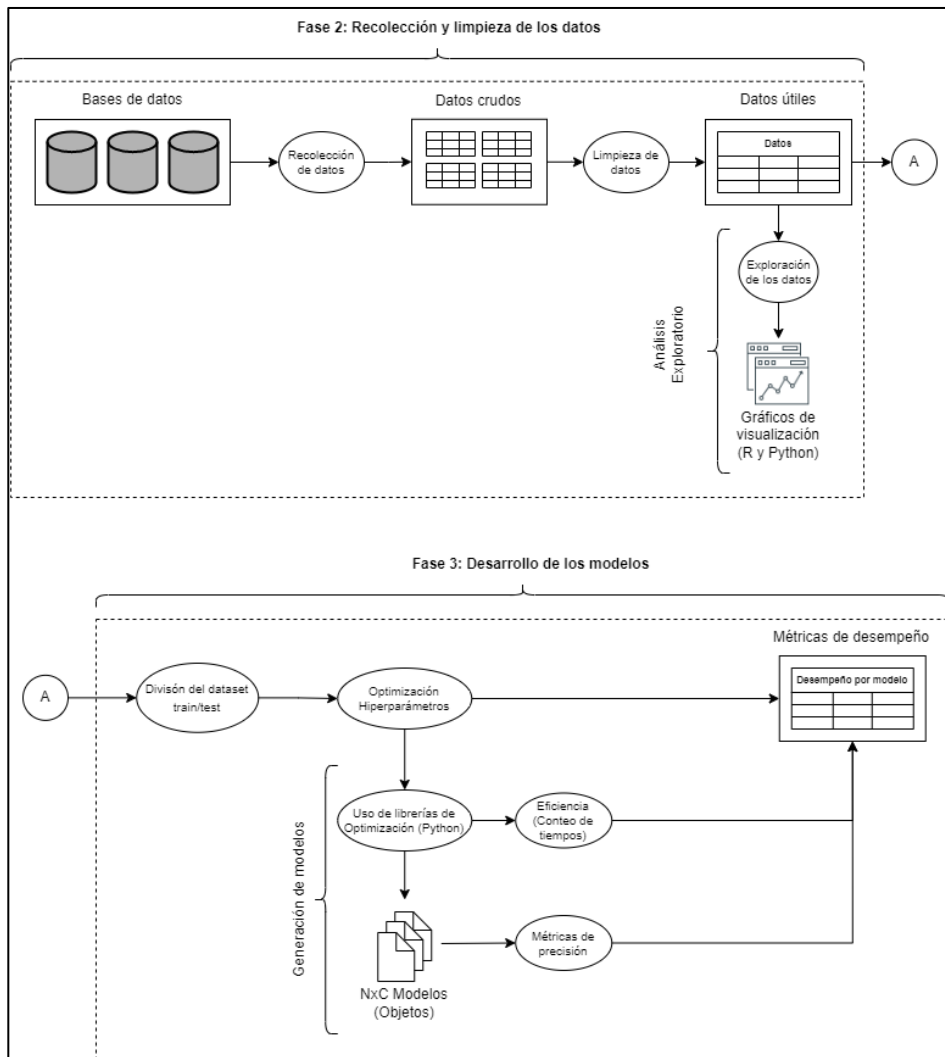
Este paso consistirá en la recolección de los datos a través de la consulta en las diferentes bases de datos públicas, proporcionadas por las empresas que brindan soporte para las transacciones con criptomonedas. Una vez recolectada la información, esta se debe limpiar y dar formato para llevar a cabo los cálculos correspondientes.

9.4.3. Fase tres: desarrollo de los modelos

Una vez se obtengan los datos, se procede a optimizar los hiperparámetros de cada uno de los algoritmos de aprendizaje seleccionados; esto se hará para obtener un modelo por cada uno de los algoritmos (N) y por cada una de las criptomonedas (C) seleccionadas. Estos modelos se entrenarán utilizando datos de 2017 a 2021, y se probarán utilizando datos del 2022. Por lo tanto, se tendrán $N \times C$ modelos, que servirán como elementos

para la comparación entre los algoritmos. Finalmente, se cuantificará el desempeño de estos modelos, al utilizar las métricas de precisión seleccionadas, y se medirán sus tiempos de entrenamiento y ejecución, lo cual se hará en el mismo *script* que se desarrollará para entrenar los modelos, con el uso del lenguaje de programación *Python*.

Figura 5. Diagrama de la fase dos y tres del estudio

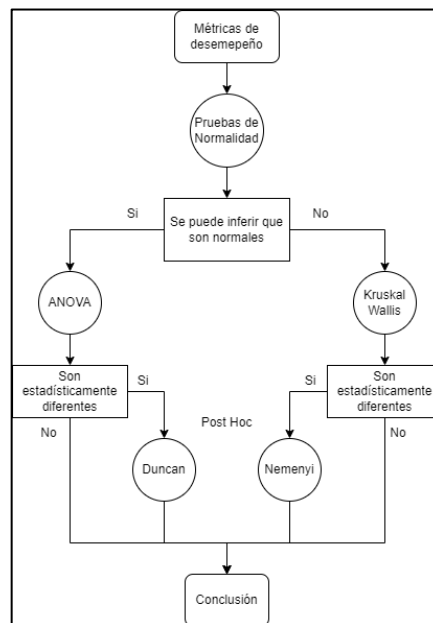


Fuente: elaboración propia, hecho con Microsoft Visio 365.

9.4.4. Fase cuatro: cálculo de resultados

Este paso comenzará con el análisis exhaustivo de la normalidad de las métricas de desempeño obtenidas, que inicia con un análisis de curtosis y sesgo, seguido de las pruebas de Kolmogórov-Smirnov, Anderson-Darling y Shapiro-Wilk, siendo esta última la que tendrá mayor peso para concluir la existencia de la normalidad. Con base en estos resultados, se procederá a efectuar las pruebas de contraste necesarias; en el caso de que se concluya la existencia de la normalidad (caso paramétrico) se efectuará un ANOVA, en caso contrario (caso no paramétrico), se realizará una prueba de Kruskal-Wallis. Finalmente, si se concluye que existen diferencias significativas, se realizará una prueba *post hoc* de Duncan o de Nemenyi (caso paramétrico y no paramétrico respectivamente). Por lo tanto, se seguirá el siguiente flujo:

Figura 6. Diagrama de la fase cuatro del estudio



Fuente: elaboración propia, hecho con Microsoft Visio 365.

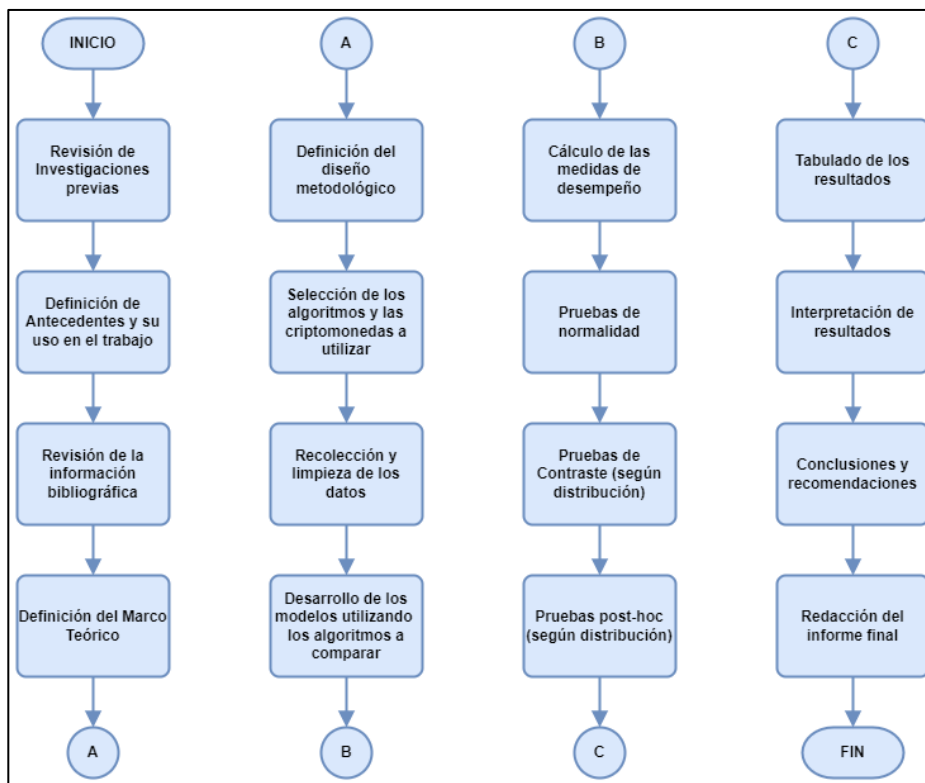
9.4.5. Fase cinco: interpretación de resultados

Con base en los resultados obtenidos por las pruebas de contraste y *post hoc*, se interpreta la existencia de diferencias significativas entre los modelos comparados y se ordenan los algoritmos según su desempeño.

9.4.6. Fase seis: redacción del informe final

Se redacta el informe final para presentar los resultados y conclusiones. A continuación, se presenta el flujograma del estudio.

Figura 7. **Flujograma del estudio**



Fuente: elaboración propia, hecho con Microsoft Visio 365.

10. TÉCNICAS DE ANÁLISIS DE INFORMACIÓN

Las técnicas estadísticas que se utilizarán para el análisis de la información se detallan a continuación:

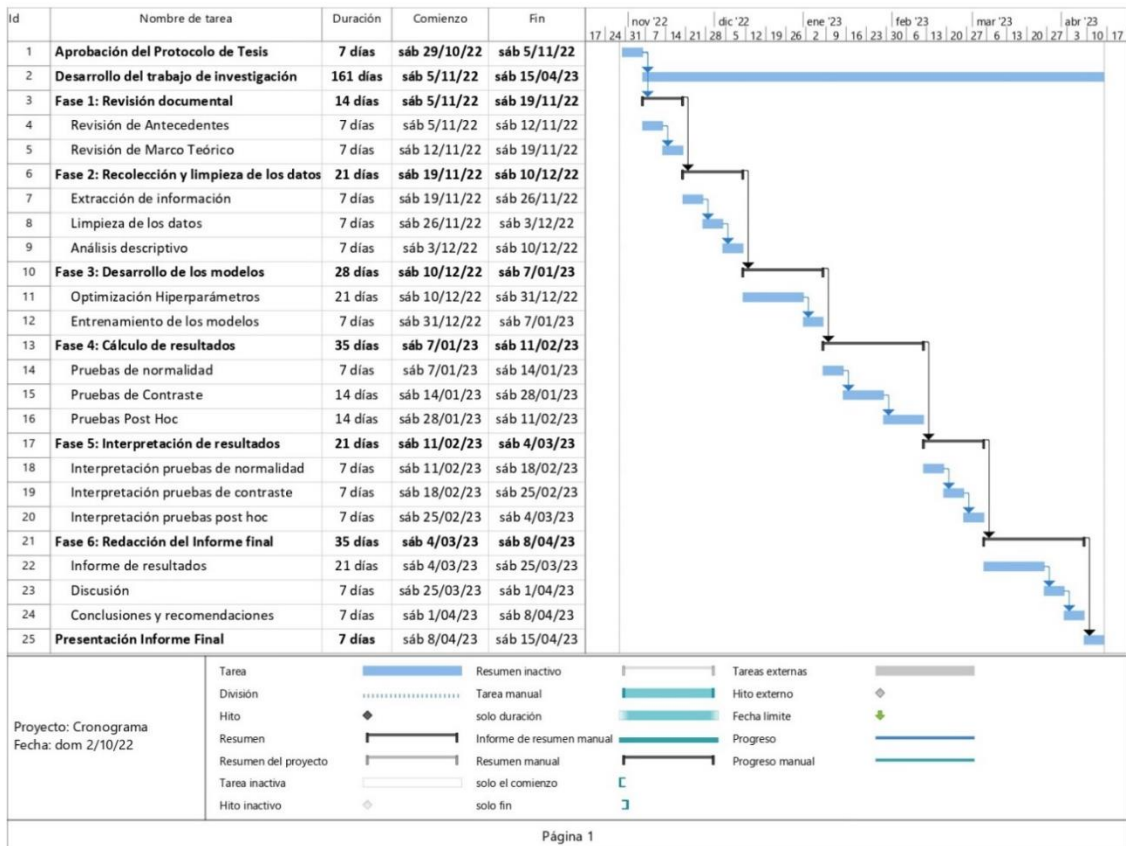
- Pruebas de normalidad: se utilizarán las siguientes pruebas de normalidad: análisis de sesgo y curtosis, Kolmogórov-Smirnov, Anderson-Darling y Shapiro-Wilk. Estas pruebas se llevarán a cabo sobre las métricas de desempeño de los diferentes algoritmos a comparar, para inferir la distribución de las que provienen las muestras. Con base en ello, se tomarán decisiones sobre la aplicación de pruebas de contraste y *post hoc* de tipo paramétrico o no paramétrico, para explorar las diferencias significativas entre el desempeño de los diferentes algoritmos estudiados.

- Pruebas de contraste: de acuerdo con el análisis de la normalidad de los datos, se utilizarán las siguientes pruebas:
 - Datos normales: se utilizará un Análisis de Varianza (ANOVA)
 - Datos no normales: se utilizarán las pruebas de Friedman

- Pruebas *post hoc*: de acuerdo con el análisis de la normalidad de los datos, se utilizarán las siguientes pruebas:
 - Datos normales: se utilizará una prueba de Duncan
 - Datos no normales: se utilizará una prueba de Nemenyi

11. CRONOGRAMA

Figura 8. Cronograma de actividades



Fuente: elaboración propia, hecho con Microsoft Project 365.

12. FACTIBILIDAD DEL ESTUDIO

Los recursos necesarios para desarrollar la investigación se describen a continuación.

12.1. Recurso humano

El recurso humano consistirá en el estudiante de la maestría y el asesor de la investigación; este último rol será *ad honorem*.

12.2. Recursos financieros

A continuación, se detalla el presupuesto de la investigación, el cual será cubierto por el estudiante de la maestría.

Tabla III. Presupuesto

No.	Tipo de recurso	Descripción del gasto	Cantidad	Monto Unitario	Subtotal
1	Humano	Tiempo del investigador	1	Q 25,000.00	Q25,000.00
2	Equipo	Resma de papel	3	Q 40.00	Q 120.00
3	Equipo	Impresora	1	Q 1,000.00	Q 1,000.00
4	Equipo	Cartuchos de tinta	6	Q 150.00	Q 900.00
5	Equipo	Computadora	1	Q 9,000.00	Q 9,000.00
6	Infraestructura	Electricidad	6	Q 315.00	Q 1,890.00
7	Infraestructura	Internet	6	Q 350.00	Q 2,100.00
8	Tecnológico	Licencia de <i>Microsoft Office</i>	1	Q 1,000.00	Q 1,000.00
9	Otros	Imprevistos	1	Q 1,000.00	Q 1,000.00
Total					Q42,010.00

Fuente: elaboración propia, hecho con Microsoft Excel 365.

12.3. Recursos tecnológicos

Los recursos tecnológicos que se emplearán en la investigación se refieren a los programas que se utilizarán, los cuales son: los lenguajes de programación Python y R, a través del *software* Visual Studio Code, para la extracción de datos, el procesamiento de la información y elaboración de gráficas (son de licencia libre); Microsoft Word y Excel (licencia de pago) para la elaboración del informe de la investigación.

12.4. Acceso a información y permisos

En este estudio se utilizarán bases de datos de carácter público, que contienen el historial de precios y transacciones de las criptomonedas en el periodo de 2018 a 2022. Estas bases serán consultadas a través de la API de la aplicación de la empresa Binance.

12.5. Equipo e infraestructura

En el equipo se considera la computadora que se utilizará para realizar todo el procesamiento de la información y la elaboración del informe. También se considera como parte del equipo a los implementos que se utilizarán en el informe, siendo estos: papel, impresora y tinta. Finalmente, en la parte de infraestructura se considera la electricidad, que será utilizada por los equipos (computadora e impresora) y el servicio de internet.

13. REFERENCIAS

1. Aguilar, A., Altamira, J. y García, O. (2010). *Introducción a la inferencia estadística*. Naucalpan de Juárez, México: Pearson Educación de México.
2. Almqvist, O. (2019). *A comparative study between algorithms for time series forecasting on customer prediction: An investigation into the performance of ARIMA, RNN, LSTM, TCN and HMM* (tesis doctoral). University of Skovde, Suecia. Recuperdo de <http://www.divaportal.org/smash/get/diva2:1321224/FULLTEXT01.pdf>
3. Binance Academy (28 de noviembre, 2018). What Is BNB? [Mensaje en un blog]. Recuperado de <https://academy.binance.com/en/articles/what-is-bnb>.
4. Campbell-Verduyn, M. (2019). *Bitcoin and Beyond: Cryptocurrencies, Blockchains, and Global Governance*. Londres, Inglaterra: Routledge.
5. Consoli, S., Recupero, D. y Saisana, M. (2021). *Data Science for Economics and Finance: Methodologies and Applications*. Italia: Springer.
6. Corder, G. y Foreman, D. (2014). *Nonparametric statistics: A step-by-step approach*. Nashville, TN: John Wiley & Sons.

7. Derbentsev, V., Datsenko, N., Stepanenko, O. y Bezkorovainyi, V. (septiembre, 2019). Forecasting cryptocurrency prices time series using machine learning approach. *SHS Web Conf.*, 65, 320-334. Recuperado de <https://ceur-ws.org/Vol-2422/paper26.pdf>.
8. Devore, J. L. (2010). *Probability and statistics for engineering and the sciences*. Florence, AL, Estados Unidos de América: Cengage Learning.
9. Dietterich, T. (enero, 1998). Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Computation*, 10(7), 1895-1923. Recuperado de <https://sci2s.ugr.es/keel/pdf/algorithm/articulo/dietterich1998.pdf>.
10. Géron, A. (2019). *Hands-on machine learning with scikit-learn, keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. Sebastopol, Ucrania: O'Reilly Media.
11. Gutiérrez, E. y Vladimirovna, O. (2016). *Estadística Inferencial 1 para ingeniería y ciencias*. Ciudad de México: Grupo Editorial Patria.
12. Gutiérrez, H. y De la Vara, R. (2008). *Análisis y diseño de experimentos*. México: McGraw-Hill Interamericana Editores.
13. Herzog, M., Francis, G. y Clarke, A. (2019). *Understanding Statistics and Experimental Design: How to not lie with statistics*. Basel, Suiza: Springer International Publishing.

14. Hitam, N., y Ismail, A. (septiembre, 2018). Comparative Performance of Machine Learning Algorithms for Cryptocurrency Forecasting. *Indonesian Journal of Electrical Engineering and Computer Science* 11(3), 1121-1128. Recuperado de <https://ijeecs.iaescore.com/index.php/IJEECS/article/view/13469/9222>.
15. James, G., Witten, D., Hastie, T. y Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. New York: Springer.
16. Kelleher, J., Namee, B. y D'Arcy, A. (2015). *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*. Londres, Inglaterra: MIT Press.
17. Kvam, P., Vidakovic, B. y Kim, S. (2022). *Nonparametric statistics with applications to science and engineering with R*. Estados Unidos: John Wiley & Sons, Inc.
18. Lind, D., Wathen, S. y Marchal, W. (2011). *Estadística aplicada a los negocios y la economía*. Álvaro Obregón, México: McGraw-Hill Interamericana Editores.
19. Luengo, J., García, S. y Herrera, F. (mayo, 2009). A study on the use of statistical tests for experimentation with neural networks: Analysis of parametric test conditions and non-parametric tests. *Expert Systems with Applications*, 36(4), 7798-7808. Recuperado de <https://www.sciencedirect.com/science/article/abs/pii/S0957417408008646>.

20. Mahir, I., Muhammad S., Fawwad H., Khurum I. Ali, H. (julio, 2021). Time-Series Prediction of Cryptocurrency Market using Machine Learning Techniques. *EAI Endorsed Transactions on Creative Technologies*, 8(28), 1-8. Recuperado de <https://eudl.eu/pdf/10.4108/eai.7-7-2021.170286>.
21. McNally, S., Roche, J. y Caton, S. (marzo, 2018). Predicting the Price of Bitcoin Using Machine Learning. In *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, 339-343. DOI 10.1109/PDP2018.2018.00060.
22. Mendes, J. (2019). *Forecasting Bitcoin prices: ARIMA vs LSTM* (tesis de maestría). ISCTE Business School, Lisboa. Recuperado de https://repositorio.iscteuiul.pt/bitstream/10071/19724/1/Master_Joao_Batista_Mendes.pdf
23. Nielsen, M. (2018). *Neural Networks and Deep Learning*. Australia: Determination Press.
24. Picardo, E. (27 de septiembre, 2022). USD Coin [Mensaje en un blog]. Recuperado de <https://www.investopedia.com/usd-coin-5210435>
25. Raschka, S. (noviembre, 2018). Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. *ArXiv*, *abs/1811.12808*. Recuperado de <https://doi.org/10.48550/ARXIV.1811.12808>.
26. Snyers, A. y Houben, R. (2018). *Cryptocurrencies and blockchain: legal context and implications for financial crime, money laundering and*

tax evasion. Bélgica: European Parliament, Directorate-General for Internal Policies of the Union. Recuperado de <https://op.europa.eu/en/publication-detail/-/publication/631f847c-b4aa-11e8-99ee-01aa75ed71a1/language-en>

27. Trawinski, B., Smętek, M., Telec, Z. y Lasota, T. (diciembre, 2012). Nonparametric statistical analysis for multiple comparison of machine learning regression algorithms. *International Journal of Applied Mathematics and Computer Science*, 22(4), 867-881. Recuperado de <https://dl.acm.org/doi/abs/10.2478/v10006-012-0064-z>
28. Triola, M. (2012). *Estadística*. Naucalpan de Juárez, México: Pearson Educación.
29. Walpole, R., Myers, R. y Myers, S. (2012). *Probabilidad y estadística para ingeniería y ciencias*. México: Pearson Education.

14. APÉNDICE

Apéndice 1. Matriz de Coherencia

ELEMENTOS	PROBLEMA DE INVESTIGACIÓN	PREGUNTAS DE INVESTIGACIÓN	OBJETIVOS	PROCEDIMIENTO Y TÉCNICAS ESTADÍSTICAS	RESULTADOS A PRESENTAR	METODOLOGÍA
GENERAL O CENTRAL	No se conoce la comparación del desempeño de algoritmos de <i>machine learning</i> aplicados a la predicción del valor de las criptomonedas.	¿Cuál es la diferencia entre el desempeño de modelos de <i>machine learning</i> aplicados a la predicción del valor de las criptomonedas?	Comparar algoritmos de <i>machine learning</i> aplicados a la predicción del valor del precio de las criptomonedas, a través de pruebas estadísticas de contraste y post hoc, para seleccionar aquellos con el mejor desempeño.	<ul style="list-style-type: none"> • Pruebas de normalidad • Pruebas de contraste • Pruebas post hoc 	<ul style="list-style-type: none"> • Modelos analíticos • Métricas de desempeño. • Resultados de las pruebas de normalidad. • Pruebas de contraste • Pruebas post hoc • Análisis de la selección de los modelos con mejor desempeño. 	<p>El enfoque del estudio propuesto es cuantitativo. Esto debido a que se analizará el desempeño de algoritmos de aprendizaje de máquina y redes neuronales.</p> <p>El tipo de investigación es descriptivo, ya que se busca especificar la existencia de diferencias significativas en el desempeño de los algoritmos seleccionados.</p>
ESPECÍFICOS O AUXILIARES	01. No se conocen las métricas que se van a utilizar para evaluar el desempeño de estos.	01. ¿Cuáles son las métricas del desempeño de los modelos de prueba que deben utilizarse para la comparación de desempeño y cuáles son los valores para los algoritmos?	01. Estimar el desempeño de los modelos de prueba, calculando su precisión y eficiencia en varios conjuntos de datos de precios de criptomonedas, para realizar comparaciones correspondientes.	<ul style="list-style-type: none"> • RMSE • MAE • MAPE • Promedio del tiempo de ejecución 	<ul style="list-style-type: none"> • Modelos analíticos • Conjunto de métricas de desempeño. 	<p>El diseño adoptado será cuasiexperimental, pues se analizará la información histórica del precio de criptomonedas entre 2018 y 2022. Por lo tanto, no se influirá ni manipulará su valor. No obstante, si se debe garantizar el control de ciertas fuentes de variación aleatoria a la hora de la comparación que se hará.</p> <p>El alcance es descriptivo-correlacional pues en la investigación se busca caracterizar el desempeño de los modelos predictivos propuestos, y a la vez comparar su comportamiento simultáneo para encontrar diferencias significativas entre ellos.</p>
	02. No se conoce la distribución de los datos.	02. ¿Cuál es la distribución de probabilidad que se ajusta mejor al comportamiento de las criptomonedas para la comparación de modelos?	02. Identificar la distribución que siguen las métricas de desempeño calculadas, utilizando análisis de bondad de ajuste, para inferir el tipo de pruebas que deben ser aplicadas en la comparación de los modelos.	<p>Pruebas de bondad de ajuste:</p> <ul style="list-style-type: none"> • Análisis de Curtosis y sesgo • Chi cuadrado • Kolmogórov-Smirnov • Anderson-Darling • Shapiro-Wilk 	<ul style="list-style-type: none"> • Normalidad de los datos. • Gráficos 	
	03. No se sabe si existen diferencias significativas en el desempeño de los modelos ni su grado de diferencia.	03. ¿Cuáles son las diferencias en el desempeño de los modelos?	03. Cuantificar las diferencias en el desempeño de los modelos, utilizando las pruebas de contraste y post hoc adecuadas, para distinguir los modelos que muestran estadísticamente un desempeño superior.	<p>Pruebas de contraste:</p> <ul style="list-style-type: none"> • ANOVA • Friedman • Kruskal-Wallis <p>Pruebas post-hoc:</p> <ul style="list-style-type: none"> • Tukey • Nemenyi 	<ul style="list-style-type: none"> • Tablas con los resultados de los estadísticos de las pruebas de contraste. • Tablas con los resultados de los estadísticos de las pruebas post hoc. 	

Fuente: elaboración propia, hecho con Microsoft Excel 365.

