



Universidad Central de Venezuela  
Facultad de Ciencias  
Escuela de Matemáticas

## *Estudio de Criptomonedas Mediante Series de Tiempo y Redes Neuronales*

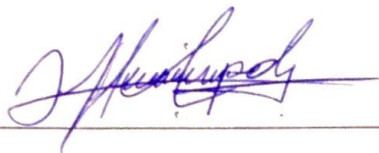
Trabajo Especial de Grado presentado ante  
la ilustre Universidad Central de Venezuela  
por el **Br. Juan Carlos Acosta Valderrama** para optar al  
título de Licenciado en Matemática.

**Tutor: Dr. José Benito Hernández.**

**Co-tutor: Dra. Mairene Colina.**

Caracas, Octubre del 2018

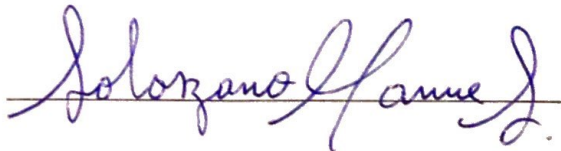
Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado "Estudio de Criptomonedas Mediante Series de Tiempo y Redes Neuronales", presentado por el Br. Juan Carlos Acosta Valderrama, titular de la Cédula de Identidad 21.438.684, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de Licenciado en Matemática.



Dr. José Benito Hernández  
Tutor



Dra. Mairene Colina  
Co-tutor



Lic. Manuel Solorzano  
Jurado

# Índice general

0.1. Resumen . . . . .	II
0.2. Palabras Claves: . . . . .	II
<b>Resumen</b>	<b>II</b>
<b>Introducción</b>	<b>1</b>
<b>1. Criptomonedas</b>	<b>3</b>
1.1. Blockchain . . . . .	3
1.2. ¿Cómo funciona una red de Blockchain? . . . . .	4
1.3. Características principales de Blockchain . . . . .	4
1.4. Beneficios y casos de uso de Blockchain . . . . .	5
1.5. Bitcoin BTC . . . . .	5
1.6. Litecoin LTC . . . . .	6
1.7. Bitcoin Cash BCH . . . . .	7
1.8. Ether ETH . . . . .	8
1.9. Ripple XRP . . . . .	8
<b>2. Series de Tiempo</b>	<b>10</b>
2.1. Series de Tiempo . . . . .	10
2.2. Modelos Estadísticos para Series de Tiempo . . . . .	11
2.3. Medidas de Dependencia: Autocorrelación y Correlación Cruzada . . . . .	12
2.4. Series de Tiempo Estacionarias . . . . .	13

2.5. Estimadores de Correlación . . . . .	14
2.6. Modelos ARMA . . . . .	15
2.6.1. Los Modelos Autoregresivos AR . . . . .	15
2.6.2. Los modelos de promedio móvil MA . . . . .	16
2.6.3. Modelos Autoregresivos de Promedio Móvil: ARMA . . . . .	17
2.7. Modelos ARIMA . . . . .	18
2.8. Funciones de Autocorrelación (ACF) y Autocorrelación Parcial (PACF) . . . . .	20
2.8.1. Función de Autocorrelación (ACF) . . . . .	20
2.8.2. Función de Autocorrelación Parcial (PACF) . . . . .	21
2.9. Modelos SARIMA . . . . .	22
<b>3. Aprendizaje Supervisado y Redes Neuronales . . . . .</b>	<b>24</b>
3.1. Aprendizaje Supervisado . . . . .	24
3.1.1. Aprendizaje Supervisado . . . . .	24
3.1.2. Tipos de Aprendizajes Supervisados . . . . .	25
3.2. Redes Neuronales . . . . .	25
3.2.1. Perceptrones . . . . .	26
3.2.2. Sigmoide . . . . .	26
3.2.3. MLP (Multi-Layer Perceptron) . . . . .	27
3.2.4. Descenso del Gradiente . . . . .	28
3.2.5. Red Neuronal Convolutacional CNN . . . . .	28
<b>4. Análisis de Resultados y Conclusiones . . . . .</b>	<b>30</b>
4.1. R (lenguaje de programación) . . . . .	30
4.2. Python . . . . .	31
4.3. Análisis de Series de Tiempo . . . . .	31
4.3.1. Estudio y Predicción del Bitcoin . . . . .	32
4.3.2. Estudio y Predicción Bitcoin Cash . . . . .	38
4.3.3. Estudio y Predicción Ether ETH . . . . .	43
4.4. Análisis de la Red Neuronal . . . . .	48
4.4.1. Red CNN para el Bitcoin BTC . . . . .	51
4.4.2. Red CNN para el Ether ETH . . . . .	53

---

4.4.3. Red CNN para el Litecoin LTC . . . . .	55
4.4.4. Red CNN para el Ripple XRP . . . . .	57
4.4.5. Red CNN Bitcoin Cash BCH . . . . .	59
4.5. Conclusiones . . . . .	61
4.5.1. Trabajos futuros . . . . .	62
<b>Bibliografía</b>	<b>63</b>

## 0.1. Resumen

La idea de este trabajo es realizar un estudio a las criptomonedas mediante series de tiempo y redes neuronales, para tratar de generar un modelo de análisis que permita predecir y clasificar el precio de las mismas. Para esto implementaremos técnicas de series de tiempo, redes neuronales y un modelo de pronóstico, previo a estos análisis estudiaremos los conceptos básicos de criptomonedas, blockchain, series de tiempo, aprendizaje supervisado y redes neuronales, que nos permitirán realizar los posteriores estudios a los datos de las 5 criptomonedas con mayor capitalización en el mercado las cuales son el Bitcoin, Litecoin, Bitcoin Cash, Ether y Ripple.

## 0.2. Palabras Claves:

Criptomoneda; Bitcoin; Litecoin; Bitcoin Cash; Ether; Ripple; Blockchain; Series de Tiempo; Redes Neuronales; Modelos SARIMA; Redes Neuronales Convolucionales CNN; Aprendizaje Supervisado; Truncamiento; Predicción; Función de Costo; Accuracy; Matriz de Confusión; Clasificación.

---

Dedicado a Dios, a la virgen María Auxiliadora y a mis familiares.

---

Primero que nada a Dios y a la virgen María, por guiarme, darme la fortaleza y sabiduria para continuar y llegar hasta este logro.

A mis Padres Luz Celeste y Juan Carlos por nunca haber dejado de confiar en mi, por todo su apoyo, consejos y esfuerzo que realizaron por mi.

A mis hermanos Alejandra y Juan Pablo, por ser como son y ayudarme a sobrellevar las cosas.

A mis abuelos por todos transmitirme toda su sabiduría, mi abuelo Acosta que aunque ya no está con nosotros se que se sentiria muy orgulloso.

A mis compañeros con los que estuve a lo largo de este trayecto en particular a mis amigos Alfredo Quintana y Miguel Porro.

A mi tutor José Benito Hernández y co-tutora Mairene Colina por su apoyo incondicional en este último tramo de la carrera.

---



En los últimos años ha venido en crecimiento una opción que ha atraído el interés de los inversores y fondos que desean apostar en ella, ésta es las criptomonedas. Las también conocidas como criptodivisas son, esencialmente, software y funcionan a través de redes P2P (de igual a igual, peer-to-peer) en Internet.

Podemos decir que es un nuevo concepto de dinero que funcionaría privadamente entre usuarios que lo poseen e intercambian entre sí. Pueden comprarse y cambiarse por dinero "tradicional" y cotizan en mercados donde se especula con su valor, como ocurre con las divisas comunes.

Este estudio se busca implementar técnicas de series de tiempo, redes neuronales y un modelo de pronóstico para la predicción de precios de las 5 criptomonedas de mayor capitalización en el mercado, hasta enero del 2018, de esta forma poder crear un modelo que nos ayude en la toma de decisiones, ya sea para el trading, el cual se basa en comprar y vender entre las distintas criptomonedas a estudiar o para especular con el precio y generar rentabilidad, en donde determinar el momento idóneo para vender es crucial, y así lograr optimizar la ganancia.

Una criptomoneda, criptodivisa es un medio digital de intercambio, su nombre proviene del prefijo "Cripto" que hace referencia a términos de codificación y encriptación matemática, que es a través de los sistemas criptográficos que se crean, almacenan y transfieren éstas. Creadas en su mayoría sobre la base de un sistema llamado Blockchain, el cual no es más que una base de datos distribuida creada por cadena de bloques para impedir su modificación una vez que un dato esté ha sido publicado usando un sellado de tiempo confiable o token por el bloque anterior.

Los primeros intentos de integración entre el dinero y la criptografía fueron hechos por David Chaum, con ECASH y DIGICASH, hoy en día la criptomoneda más conocida es el Bitcoin, creado en 2009, por Craig Wright, alias Satoshi Nakamoto. Las características principales y por lo que ha llamado la atención del mundo son la confianza, la rapidez y la seguridad, esto hace que los usuarios puedan realizar transacciones desde diferentes países con monedas y leyes diferentes en cuestión de minutos, en lugar de las vías comunes a través de las entidades bancarias en las cuales dichas transacciones pueden durar días, a su vez se elimina la necesidad de usar agentes financieros lo que reduce las comisiones por transferencia de los bancos. El mayor atractivo sin duda es su seguridad, En sus casi 8 años de existencia, han ido ganando paulatinamente la atención del público general como también el de grandes empresas que han dado un paso adelante a invertir y aceptar pagos en criptomonedas. Desde 2011, el interés ha aumentado rápidamente, en especial, durante el ascenso vertiginoso

---

de Bitcoin, en abril de 2013.

Una plataforma de intercambio de criptomonedas (cryptocurrency exchange) es un sitio web en el que se puede comprar, vender o cambiar criptomonedas por otra moneda digital o dinero fiduciario (USD, EUR, etc.). Dependiendo del intercambio, pueden operar como una bolsa de valores o bien como una casa de cambio de divisas, unas de las más utilizadas son Coinbase, Bittrex, Kucoin, entre otras, que son muy efectivas y seguras para los usuarios.

Éste trabajo busca predecir el valor de las criptomonedas a través de redes neuronales y modelos SARIMA de series de tiempo. Para poder realizar el modelo de análisis se procedera a ajustar un modelo SARIMA a los registros históricos de criptomonedas. Aplicar algoritmos de redes neuronales a los datos de las distintas criptomonedas. Identificar los parámetros que mejor se ajustan para cada modelo, para luego mostrar los resultados de los análisis realizados en el software estadístico R y Python tanto en redes neuronales como en series de tiempo.

---

# Criptomonedas

En los últimos años ha venido en crecimiento una opción que ha atraído el interés de los inversores y fondos que desean apostar en ella, ésta es las criptomonedas. Una criptomoneda no es más que un medio digital de intercambio de valor utilizada para enviar y recibir pagos a través de ordenadores conectados entre sí, su nombre proviene del prefijo "Cripto" que hace referencia a términos de codificación y encriptación matemática, que es a través de los sistemas criptográficos que se crean, almacenan y transfieren éstas, creadas en su mayoría sobre la base de un sistema llamado Blockchain, el cual no es más que una base de datos distribuida creada por cadenas de bloques para impedir su modificación una vez que algún dato esté y haya sido publicado, usando un sellado de tiempo confiable por el bloque anterior. Las también conocidas como criptodivisas son, esencialmente, software y funcionan a través de redes P2P (de igual a igual, peer-to-peer) en Internet. Podemos decir que es un nuevo concepto de dinero que funcionaría privadamente entre usuarios que lo poseen e intercambian entre sí. Pueden comprarse y cambiarse por dinero "tradicional" y cotizan en mercados donde se especula con su valor, como ocurre con las divisas comunes.

## 1.1. Blockchain

El Blockchain es una tecnología que permite la transferencia de datos de una manera completamente segura gracias a una codificación muy sofisticada. Se suele comparar con un libro de contabilidad de una empresa donde están registradas todas las entradas y salidas de dinero. Claro que, en este caso, hablaríamos de un libro de acontecimientos digitales.

## 1.2. ¿Cómo funciona una red de Blockchain?

Hasta ahora siempre habíamos necesitado de un tercero en el que las dos partes confiaran para garantizar la autenticidad de las transacciones ya sea un banco, un auditor, un notario o Paypal, por ejemplo, que tuviese un registro o sello de veracidad. La Blockchain soluciona este problema de doble gasto al combinar la tecnología P2P de intercambio entre pares con la criptografía y así crear una nueva forma de comunicación e intercambio digital. Las transacciones incluidas en los bloques son creadas por los integrantes del sistema. Todas las transacciones son registradas y transmitidas a todos los nodos (Un nodo es un ordenador conectado a la red que utiliza un software para almacenar y distribuir una copia actualizada en tiempo real de la cadena de bloques) de la red. Así, todos los integrantes tiene la información constantemente actualizada con todas las transacciones.

## 1.3. Características principales de Blockchain

1. Es un sistema seguro dado que su tecnología se basa en la criptografía de datos.
2. Las transacciones se concentran en bloques, en dichos bloques la información se almacena cronológicamente.
3. Una vez aceptada, la información no se puede borrar ni modificar, y puede ser consultada en cualquier momento.
4. Una Blockchain puede ser pública o privada, inclusive puede permitir ciertas consultas con un permiso especial.
5. La confianza, dada por la descentralización y no disponer de un nodo central donde se aglutine toda la información, hace que las partes intervinientes confíen plenamente en esta tecnología.
6. Transacciones en tiempo real: Su inmediatez minimiza el riesgo de contra-parte producido en otro tipo de transacciones, donde el compromiso de pago tarda varios días en ejecutarse (evitando riesgos de impago, concursos de acreedores y fraudes durante el proceso de pago).

## 1.4. Beneficios y casos de uso de Blockchain

Blockchain tiene el potencial de hacer el intercambio de moneda e información más seguro en variedad de industrias, consiguiendo transferencias de datos más sencillas entre entidades. Podemos valorar su uso en varios sectores o industrias.

**En el sector Financiero:** Hoy en día las transacciones interbancarias pueden tardar en muchos casos días en ser aprobadas y finalmente ingresadas, especialmente fuera del horario de oficina. Con Blockchain son procesadas a cualquier hora, y el tiempo en ser completadas puede reducirse a minutos.

**Firmas y contratos:** Una de las principales problemáticas asociadas a la firma de contratos es la posible manipulación de los documentos para intentar alguna estafa o fraude legal. Sin embargo, estos temores a una potencial conspiración en nuestra contra se disipan de base al incorporar el blockchain a la ecuación, ya que la información que se guarde en la cadena de bloques no se puede alterar de ninguna forma.

**Sanidad:** Gracias a los beneficios de los dispositivos médicos conectados, la cadena de bloques puede ayudar a la industria de la salud resolviendo el desperdicio de tiempo, dinero y duplicación de procesos que la distribución actual en distintas instalaciones y proveedores supone.

Uno de los países más proactivos en la materia en este momento es Estonia, donde su gobierno trabaja de la mano con Bitnation, para permitirle a sus residentes con identidad digital, registrar sus uniones matrimoniales y partidas de nacimiento en la Blockchain Horizon.

## 1.5. Bitcoin BTC

La definición que Bitcoin hace del mismo es corta y sencilla, pero sólo si sabemos qué significan los conceptos con los que se define (Bitcoin Foundation):

“Bitcoin es una red consensuada que permite un nuevo sistema de pago y una moneda completamente digital. Es la primera red entre pares de pago descentralizado impulsado por sus usuarios sin una autoridad central o intermediarios. Desde un punto de vista de usuario, Bitcoin es como dinero para Internet. Bitcoin puede ser el único sistema de contabilidad triple existente.”

En noviembre de 2008, se registró un artículo en una lista de distribución de criptografía firmado con el nombre Satoshi Nakamoto y titulado Bitcoin: Un sistema de efectivo electrónico de Peer-to-Peer. En este documento se detalla cómo usar una red peer-to-peer para crear (un sistema para transacciones electrónicas que no dependa de la confianza).

El 4 de enero de 2009, Satoshi Nakamoto registra el primer bloque, creando así la red Bitcoin y la emisión de los primeros bitcoins.

Desde 2011, el interés ha aumentado rápidamente, en especial, durante el ascenso vertiginoso de Bitcoin, en abril de 2013. Una plataforma de intercambio de criptomonedas (cryptocurrency exchange) es un sitio web en el que se puede comprar, vender o cambiar criptomonedas por otra moneda digital o dinero fiduciario (USD, EUR, etc.). Dependiendo del intercambio, pueden operar como una bolsa de valores o bien como una casa de cambio de divisas, unas de las más utilizadas son Coinbase, Bittrex, Kucoin, entre otras, que son muy efectivas y seguras para los usuarios.

### **Generación de Bitcoins**

El sitio oficial en Internet de Bitcoin afirma que la generación de bitcoins se lleva a cabo mediante un proceso competitivo y descentralizado llamado "minería", el cual hace referencia a la actividad de "minar" metales preciosos.

#### **¿Qué es un Satoshi?**

Los Satoshis son otra unidad de medida para fraccionar un Bitcoin. La equivalencia es  $100.000.000 = 1$  Bitcoin, Un Satoshi es la mínima unidad en que la que puede dividirse un Bitcoin, y se lo expresa como 0.00000001.

El Bitcoin es una criptomoneda, con una cantidad finita de 21 millones de unidades, por eso es una de las razones de lo cotizada que es.

## **1.6. Litecoin LTC**

Litecoin (LTC) nació en 2011 de la mano de Charlie Lee. Antiguo trabajador de Google y exdirector de ingeniería en Coinbase (principal plataforma de cambio de monedas digitales).

Litecoin es una Criptomoneda sustentada por la red P2P, y un proyecto de software de código abierto publicado bajo la licencia MIT. Inspirada y prácticamente idéntica en su aspecto técnico a Bitcoin (BTC), la creación y transferencia de Litecoin se basa en un protocolo criptográfico de código abierto que no es administrado por ninguna autoridad central. Litecoin fue pensada para ser una moneda electrónica alternativa de Bitcoin y ofrece tres diferencias principales. En primer lugar, la red Litecoin realiza el procesamiento de un bloque cada 2,5 minutos en vez de cada 10 minutos, lo cual permite una confirmación más rápida de transacciones.

En segundo lugar, la red Litecoin producirá aproximadamente 4 veces más unidades que Bitcoin, o cerca de 84 millones de Litecoins. En tercer lugar, Litecoin utiliza la función scrypt en su algoritmo de prueba de trabajo: una función secuencial de memoria dura concebida por primera vez por Colin Percival, facilitando la minería, ya que no necesita de equipamiento sofisticado como en el caso de Bitcoin. Cada litecoin es fraccionado en 100.000.000 unidades más pequeñas, definidas por ocho decimales.

## 1.7. Bitcoin Cash BCH

Bitcoin Cash (BCH) es un sistema de efectivo electrónico peer-to-peer. Es una criptomoneda descentralizada sin permisos que no requiere de terceros de confianza ni de un banco central.

El primer software de bitcoin se lanzó en enero de 2009 y funcionó durante varios años con tarifas bajas y transacciones rápidas y confiables. Para 2016, debido al gran interés, la cadena de bloques de bitcoin comenzó a tener problemas de escala. Esta congestión de la red conduce a un aumento sustancial en las tarifas de transacción. La comunidad, incapaz de llegar a un consenso, que se conoce como "debate sobre el tamaño del bloque", forzó una nueva criptomoneda llamada Bitcoin Cash el 1 de agosto de 2017, a la altura del bloque 478558, Bitcoin y Bitcoin Cash comparten un historial idéntico hasta el bloque 478558.

### Similitudes y diferencias entre BTC y BCH:

- \* El BTC fue lanzado en Enero del año 2009 y el BCH en Agosto del año 2017.
- \* Mientras BTC es una criptomoneda original, BCH nació a partir de un hard fork o bifurcación dura de BTC.
- \* BTC tiene bloques de 1 MB, aunque puede llegar a 2 MB en caso de aplicarse el mecanismo de "Testigo Segregado", llamado SegWit2x, mientras BCH implementó los bloques de 8 MB, además de permitir seguir escalando dicho límite hasta 32 MB, permitiendo entre 23 y 92 transacciones por segundo frente a las 3-7 transacciones por segundo (transacciones por segundo) que permite BTC.
- \* BTC prioriza mantener una red pública, sin permisos y altamente resistente a la censura, mientras que BCH prioriza mantener una red que permita pagos rápidos y baratos.

## 1.8. Ether ETH

Ethereum busca ser un “ordenador mundial” mediante el cual toda persona conectada a Internet pueda interactuar; añadiendo y utilizando sus programas o aplicaciones.

Al igual que Bitcoin, es una plataforma descentralizada es decir, no controlada por nadie en concreto y que opera sobre una Blockchain o “cadena de bloques”.

Su plataforma soporta, entre otros, Contratos Inteligentes, DAOs y Dapps y a su vez utiliza el Ether, su criptomoneda, como combustible para poder interactuar con ella.

Creada en el 2015 por Vitalik Buterin, el ether, a diferencia del bitcoin, es ilimitado. Es decir, su creación puede ser infinita mientras que el bitcoin tiene como límite 21 millones de unidades, a partir de ahí no se pueden crear más.

Unas de sus principales características son:

- \* La descentralización, al no haber un banco central que las controle.
- \* El anonimato, ya que las direcciones para enviarlas y recibirlas son series numéricas aleatorias.
- \* Permiten realizar transacciones directas de forma mucho más rápida y segura que las que se producen con dinero tradicional “fiat”.

Los Ethers, como criptomoneda, se pueden comprar en diferentes portales web. El más popular es Coinbase, a través del cual puedes crear de forma rápida y sencilla un billetero de Ethereum y empezar a operar con Ethers. En función de las personas que quieran comprarlo o venderlo, el valor del Ether sube o baja de valor. El precio del Ether cambia constantemente al igual que el precio de las acciones, de las monedas como el euro o el dólar o las propiedades inmobiliarias.

## 1.9. Ripple XRP

Considerada la criptomoneda de los bancos, Ripple es un sistema virtual de pagos en tiempo real basado en la tecnología blockchain, que es utilizada por instituciones financieras como una forma más rápida y económica de enviar y recibir dinero, así como de liquidar transacciones a una velocidad de entre 5 y 10 segundos. También permite la conversión entre distintas divisas y criptomonedas. La abreviatura de la divisa Ripple se conoce como XRP.

Ripple fue concedida por Ryan Fugger en 2004 cuando trabajaba en un “sistema de intercambio local” en Vancouver. Por otro lado, en 2011 Jed McCaleb comienza a desarrollar su propia moneda digital, una moneda



cuyo valor se fijara por consenso, en lugar de por minería como en el caso del Bitcoin. En 2012, Fugger transfiere el control de Ripple a Jed McCaleb y a su compañero, Chris Larsen y crean la organización OpenCoin, que posteriormente pasarán a llamar Ripple Labs. En 2014, Ripple se convierte en la criptomoneda más grande del mundo, superando al Litecoin y comienza a ser adoptado y financiado por grandes bancos como el Fidor Bank alemán o el banco Santander. Desde entonces se acentúa el crecimiento de esta criptomoneda, alcanzando en octubre de 2017 una capitalización de mercado de 10 mil millones de dólares, siendo reconocida por los mayores bancos del mundo como la única criptomoneda para medio de pago. En diciembre de 2017, el Ripple rompe su máximo histórico en 1,60\$ por moneda, creciendo más del 1000 % en un solo mes.

Unas de sus principales características son:

- \* Hace uso de la tecnología de ondulación, es decir, no depende de entidades bancarias ni tarjetas de crédito, por lo que puede suponer una ventaja para realizar transferencias a nivel internacional, gracias a su mayor flexibilidad frente a otras criptomonedas.
- \* El Ripple es un protocolo de pago que funciona de forma similar a un sistema de pago, red de remesas y cambio de divisas. Funciona como criptomoneda, moneda fiat y materia prima.
- \* Una de sus principales ventajas es la reducción de los costos de transacción.
- \* Es posible realizar liquidaciones y transferencias al instante, reduciendo el tiempo de cada operación.

# Series de Tiempo

En este capítulo, vamos a definir conceptos fundamentales con los que vamos a trabajar, como lo son las series de tiempo, las componentes de una serie de tiempo y qué es un modelo de serie de tiempo. El análisis de series de tiempo desempeña un papel importante en el análisis requerido para el pronóstico de eventos futuros. Existen varias formas o métodos de calcular cual va a ser la tendencia del comportamiento del proceso en estudio.

## 2.1. Series de Tiempo

Una serie tiempo es una secuencia de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y, espaciados entre sí de manera uniforme, así los datos usualmente son dependientes entre sí. El principal objetivo de una serie de tiempo es su análisis para hacer pronóstico. Formalmente se tiene la siguiente definición.

**Definición 2.1.1** *Una serie de tiempo es un conjunto de observaciones  $X_t$ , cada una registrada a un tiempo específico  $t$ .*

El análisis clásico de las series de tiempo se basa en la suposición de que los valores que toma la variable de observación es la consecuencia de tres componentes, cuya actuación conjunta da como resultado los valores medidos, estos componentes son:

1. **Tendencia:** Se puede definir como un cambio a largo plazo que se produce en la relación al nivel medio, o el cambio a largo plazo de la media. La tendencia se identifica con un movimiento suave de la serie a largo plazo.

2. **Estacional:** Muchas series de tiempo presentan cierta periodicidad o dicho de otro modo, variación de cierto período (semestral, mensual, etc.). Por ejemplo las ventas al detalle en Puerto Rico aumentan por los meses de noviembre y diciembre por las festividades navideñas. Estos efectos son fáciles de entender y se pueden medir explícitamente o incluso se pueden eliminar de la serie de datos, a este proceso se le llama desestacionalización de la serie.
3. **Aleatoria:** Esta componente no responde a ningún patrón de comportamiento, sino que es el resultado de factores fortuitos o aleatorios que inciden de forma aislada en una serie de tiempo.

De los tres componentes anteriores los dos primeros son componentes determinísticos, mientras que la última es aleatoria. Así, se puede denotar la serie de tiempo como: Así, se puede denotar la serie de tiempo como

$$X_t = T_t + E_t + \epsilon_t$$

donde  $T_t$  es la tendencia,  $E_t$  es la componente estacional y  $\epsilon_t$  es la componente aleatoria.

## 2.2. Modelos Estadísticos para Series de Tiempo

El principal objetivo en el análisis de Series de Tiempo es desarrollar modelos matemáticos que provean una descripción apropiada para los datos muestrales, en principio utilizaremos la definición de Series de Tiempo.

**Definición 2.2.1** *Un proceso estocástico es una familia de variables aleatorias indexadas  $x(\omega, t)$  ó  $x_t(\omega)$  donde  $t$  pertenece a un conjunto de índices  $T$  y  $\omega$  pertenece a un espacio muestral  $\Omega$ . Si  $t = t^*$  fijo,  $x(\omega, t^*)$  es una variable aleatoria. Si  $\omega = \omega^*$  fijo,  $x(\omega^*, t)$  es una función de  $t$ , y se llama una realización del proceso. Una **serie de tiempo** es la realización de un proceso estocástico.*

Veamos unos ejemplos:

### Ejemplo 2.2.1 Ruido Blanco.

Una manera sencilla de generar series de tiempo, puede ser considerando una sucesión de variables aleatorias no-correlacionadas,  $w_t$  con media 0 y varianza  $\sigma_w^2$ . Las series de tiempo generadas de esta manera son usadas como modelos para ruido en aplicaciones de ingeniería, donde ellas son llamadas ruido blanco, denotaremos este proceso como  $w_t \sim wn(0, \sigma_w^2)$ . La designación blanco se origina de la analogía con luz blanca e indica que todos los posibles períodos de oscilación están presente con igual intensidad.

**Ejemplo 2.2.2 Promedio móvil.**

Un promedio móvil se construye sustituyendo cada valor de una serie por la media obtenida con esa observación y algunos de los valores inmediatamente anteriores y posteriores.

$$v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1}), \quad (2.1)$$

**2.3. Medidas de Dependencia: Autocorrelación y Correlación Cruzada**

**Definición 2.3.1** La función de media está definida como

$$\mu_{x_t} = \mathbb{E}(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx, \quad (2.2)$$

en caso de que exista,  $\mathbb{E}$  denota el operador usual de esperanza. Ahora escribiremos  $\mu_{x_t}$  como  $\mu_t$ . Es relevante entender que  $\mu_t$  consiste en que es una media teórica para la serie de tiempo en un punto particular, en donde la media se calcula sobre todos los posibles eventos que podrían haber producido  $X_t$ .

**Definición 2.3.2** La función de autocovarianza es definida como producto del segundo momento

$$\gamma_x(s, t) = \mathbb{E}[(x_s - \mu_s)(x_t - \mu_t)], \quad (2.3)$$

para todo  $t$  y  $s$ . cuando no haya confusión en la existencia sobre a que serie nos referimos, escribiremos  $\gamma_x(s, t) = \gamma(s, t)$ .

**Definición 2.3.3** La función de autocorrelación (ACF) se define como

$$\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}}. \quad (2.4)$$

La ACF mide la predictibilidad lineal de una serie de tiempo en tiempo  $t$ , digamos  $x_t$  usando solo el valor  $x_s$ , si podemos predecir  $x_t$  exactamente de  $x_s$  a través de la relación lineal  $x_t = \beta_0 + \beta_1 x_s$  entonces la correlación será 1 cuando  $\beta_1 > 0$  y  $-1$  cuando  $\beta_1 < 0$ .

## 2.4. Series de Tiempo Estacionarias

Las series temporales se pueden clasificar en:

1. **Estacionarias:** Una serie es estacionaria cuando es estable a lo largo del tiempo, es decir, cuando la media y varianza son constantes en el tiempo. Esto se refleja gráficamente en que los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también permanece constante en el tiempo.
2. **No estacionarias:** Son series en las cuales la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante.

**Definición 2.4.1** Una serie de tiempo *estrictamente estacionaria* es una serie para la cual el comportamiento probabilístico de cada sucesión de valores

$$\{x_{t_1}, x_{t_2}, \dots, x_{t_k}\}$$

es idéntico a la serie trasladada en el tiempo

$$\{x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}\}.$$

Esto es,

$$P\{x_{t_1} \leq c_1, \dots, x_{t_k} \leq c_k\} = P\{x_{t_1+h} \leq c_1, \dots, x_{t_k+h} \leq c_k\} \quad (2.5)$$

para todo  $k = 1, 2, \dots$ , todo puntos de tiempos  $t_1, t_2, \dots, t_k$  y números reales  $c_1, c_2, \dots, c_k$  y todo salto  $h = 0, \pm 1, \pm 2, \dots$

Esta versión de estacionaridad es muy fuerte para la mayoría de aplicaciones. Más aún, es difícil conseguir estricta estacionaridad en un conjunto sencillo de datos.

**Definición 2.4.2** Una serie de tiempo *débilmente estacionaria*  $X_t$ , es un proceso de varianza finita tal que

1. la función de media  $\mu_t$ , definida en (2.2) es constante y no depende del tiempo  $t$ , y
2. la función de covarianza,  $\gamma(s, t)$ , definida en (2.3) depende solo de las diferencias de  $s$  y  $t$ ,  $|s - t|$ .

Por consiguiente, usaremos el término **estacionaridad** para referirnos a estacionaridad débil; si un proceso es estacionario en el sentido estricto usaremos el término estrictamente estacionario.

Es claro de la definición de estrictamente estacionario que una una serie de tiempo estrictamente estacionaria con varianza finita, también es una serie estacionaria.

**Definición 2.4.3** La función de autocovarianza de una serie de tiempo estacionaria se escribirá como

$$\gamma(h) = \mathbb{E}[(x_{t+h} - \mu)(x_t - \mu)]. \quad (2.6)$$

**Definición 2.4.4** La función de autocorrelación (ACF) de una serie de tiempo estacionaria será escrita, usando (2.4) como

$$\rho(h) = \frac{\gamma(t+h, t)}{\sqrt{\gamma(t+h, t+h)\gamma(t, t)}} = \frac{\gamma(h)}{\gamma(0)}. \quad (2.7)$$

Por la desigualdad de Cauchy-Schwartz tenemos que  $-1 \leq \rho(h) \leq 1$  para todo  $h$ .

## 2.5. Estimadores de Correlación

**Definición 2.5.1** Sea  $x_1, x_2, \dots, x_n$  una muestra de una serie de tiempo. La **media muestral** de  $x_1, x_2, \dots, x_n$  es

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t. \quad (2.8)$$

**Definición 2.5.2** La función de autocovarianza muestral se define como

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(x_t - \bar{x}) \quad (2.9)$$

con  $\hat{\gamma}(-h) = \hat{\gamma}(h)$  para  $h = 0, 1, \dots, n-1$ .

**Definición 2.5.3** La función de autocorrelación muestral se define como

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}. \quad (2.10)$$

## 2.6. Modelos ARMA

Los modelos ARMA deben su nombre a sus siglas en inglés, Auto Regressive Moving Average, ó Modelos Autoregresivos de Promedio Móvil.

### 2.6.1. Los Modelos Autoregresivos AR

Los modelos autoregresivos están basados en la idea de que el valor actual de la serie  $x_t$  se puede explicar como una función de  $p$  valores pasados  $x_{t-1}, x_{t-2}, \dots, x_{t-p}$  donde  $p$  determina el número de pasos en necesarios para predecir el valor actual.

**Definición 2.6.1** *Un modelo autoregresivo de orden  $p$ , abreviado  $AR(p)$ , es de la forma*

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t \quad (2.11)$$

donde  $x_t$  es estacionario,  $\phi_1, \phi_2, \dots, \phi_p$  son constantes ( $\phi_p \neq 0$ ). A menos que se declare lo contrario, se asume que  $w_t$  es un ruido blanco gaussiano de media cero y varianza  $\sigma_w^2$ . La media de  $x_t$  en (2.11) es cero. Si la media  $\mu$  de  $x_t$  no es cero, reemplazamos  $x_t$  por  $x_t - \mu$  en (2.11), es decir

$$x_t - \mu = \phi_1 (x_{t-1} - \mu) + \phi_2 (x_{t-2} - \mu) + \dots + \phi_p (x_{t-p} - \mu) + w_t,$$

o escribimos

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t \quad (2.12)$$

donde  $\alpha = \mu(1 - \phi_1 + \dots + \phi_p)$ .

**Definición 2.6.2** *El operador autoregresivo se define como*

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p. \quad (2.13)$$

Iniciaremos el estudio de los modelos AR considerando el modelo de primer orden  $AR(1)$ , el cual es dado por  $x_t = \phi x_{t-1} + w_t$ . Iterando el operador de cambio  $k$  veces, obtenemos

$$\begin{aligned} x_t &= \phi x_{t-1} + w_t = \phi(\phi x_{t-2} + w_{t-1}) + w_t \\ &= \phi^2 x_{t-2} + \phi w_{t-1} + w_t \\ &\vdots \\ &= \phi^k x_{t-k} + \sum_{j=0}^{\infty} \phi^j w_{t-j}. \end{aligned}$$

Este método sugiere que por iteración continua del operador de cambio, siempre que  $|\phi| < 1$  y  $x_t$  sea estacionario, podemos representar un modelo AR(1) como un proceso lineal dado por <sup>1</sup>

$$x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j}. \quad (2.14)$$

El proceso AR(1) definido en (2.14) es estacionario con media

$$\mathbb{E}(x_t) = \sum_{j=0}^{\infty} \phi^j \mathbb{E}(w_{t-j}) = 0,$$

y función de autocovarianza

$$\begin{aligned} \gamma(h) &= \text{cov}(x_{t+h}, x_t) = \mathbb{E} \left[ \left( \sum_{j=0}^{\infty} \phi^j w_{t+h-j} \right) \left( \sum_{k=0}^{\infty} \phi^k w_{t-k} \right) \right] \\ &= \sigma_w^2 \sum_{j=0}^{\infty} \phi^j \phi^{j+h} = \sigma_w^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} = \frac{\sigma_w^2 \phi^h}{1 - \phi^2}, \quad h \geq 0. \end{aligned} \quad (2.15)$$

Recuerde que  $\gamma(h) = \gamma(-h)$  de modo que sólo presentamos la función de autocovarianza para  $h \geq 0$ .

De (2.15) la **ACF de un modelo AR(1)** es

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h, \quad h \geq 0, \quad (2.16)$$

y  $\rho(h)$  satisface la recursión

$$\rho(h) = \phi \rho(h-1), \quad \text{con } h = 1, 2, \dots \quad (2.17)$$

## 2.6.2. Los modelos de promedio móvil MA

Como una alternativa a la representación autoregresiva en la cual  $x_t$  del lado izquierdo de la ecuación se asume como una combinación lineal, en los modelos de promedio móvil de orden  $q$  abreviados  $MA(q)$  asumimos el ruido blanco  $w_t$  del lado derecho de la ecuación que los define como una combinación lineal de los datos observados

**Definición 2.6.3** *El modelo de promedio móvil de orden  $q$  o modelo  $MA(q)$ , se define como*

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \dots + \theta_q w_{t-q}, \quad (2.18)$$

donde hay  $q$  pasos en el promedio móvil y  $\theta_1, \theta_2, \dots, \theta_q$  ( $\theta_q \neq 0$ ) son parámetros.

<sup>1</sup>Note que  $\lim_{k \rightarrow \infty} \mathbb{E} \left( x_t - \sum_{j=0}^{\infty} \phi^j w_{t-j} \right)^2 = \lim_{k \rightarrow \infty} \phi^{2k} \mathbb{E}(x_{t-k}^2) = 0$ , de modo que (2.14) existe en el sentido de media cuadrado.



El ruido  $w_t$  se asume como un ruido blanco gaussiano.

Podemos también escribir el proceso MA(q) en la forma equivalente

$$x_t = \theta(B)w_t, \quad (2.19)$$

usando la siguiente definición.

**Definición 2.6.4** *El operador de promedio móvil es*

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q. \quad (2.20)$$

En contraste con el proceso autorregresivo, el proceso de promedio móvil es estacionario para cada valor de los parámetros  $\theta_1, \theta_2, \dots, \theta_q$ .

### 2.6.3. Modelos Autorregresivos de Promedio Móvil: ARMA

Dada una serie temporal de datos  $X_t$ , el modelo ARMA es una herramienta para entender y, aún más, para predecir futuros valores de la serie. El modelo está formado por dos partes, una parte autorregresiva (AR) y otra de media móvil (MA). El modelo se conoce con el nombre de modelo  $ARMA(p, q)$ , donde  $p$  es el orden de la parte autorregresiva y  $q$  es el orden de la parte de media móvil.

**Definición 2.6.5** *Una serie de tiempo  $\{x_t; t = 0, \pm 1, \pm 2, \dots\}$  es un proceso autorregresivo de promedio móvil, denotado  $ARMA(p, q)$ , si es estacionario y*

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q} \quad (2.21)$$

con  $\phi_p \neq 0, \theta_q \neq 0$  y  $\sigma_w^2 > 0$ . Los parámetros  $p$  y  $q$  son llamados ordenes autorregresivos y de promedio móvil respectivamente. Si  $x_t$  tiene media  $\mu$  distinta de cero, hacemos  $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$  y escribimos el modelo como

$$x_t = \alpha + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}. \quad (2.22)$$

A menos que se declare lo contrario,  $\{w_t; t = 0, \pm 1, \pm 2, \dots\}$  es una sucesión de ruido blanco gaussiano.

Como ayuda en la investigación de los modelos ARMA, será útil escribir estos usando el operador AR (2.13) y el operador MA (2.20). En particular el modelo  $ARMA(p, q)$  en (2.21) se puede escribir en forma concisa como

$$\phi(B)x_t = \theta(B)w_t. \quad (2.23)$$

**Definición 2.6.6** Los *Polinomios AR y MA* se definen como:

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p, \phi_p \neq 0. \quad (2.24)$$

y

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_p z^p, \theta_p \neq 0. \quad (2.25)$$

respectivamente, donde  $z$  es un número complejo.

**Definición 2.6.7** Un modelo  $ARMA(p,q)$ ,  $\phi(B)x_t = \theta(B)w_t$ , se dice que es **causal** si la serie de tiempo  $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$  se puede escribir como un proceso lineal de un lado

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \psi(B)w_t \quad (2.26)$$

donde  $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$  y  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ ; haciendo  $\psi_0 = 1$ .

**Definición 2.6.8** Un modelo  $ARMA(p,q)$ ,  $\phi(B)x_t = \theta(B)w_t$  se dice **invertible** si la serie de tiempo  $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$  se puede escribir como

$$\pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = w_t \quad (2.27)$$

donde  $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$  y  $\sum_{j=0}^{\infty} |\pi_j| < \infty$ ; hacemos  $\pi_0 = 1$ .

## 2.7. Modelos ARIMA

Los modelos econométricos contemplan de forma explícita la información que aportan las variables causales del fenómeno de interés, de acuerdo con una determinada teoría económica. Una ventaja de este modelo consiste en que los resultados que se generan son más eficientes y poseen mayor poder explicativo que los modelos univariados. Sin embargo, en estos modelos, cuando se desea realizar predicciones, el desconocimiento de los valores de las variables explicativas en el futuro, determina la necesidad de utilizar predicciones para éstas, lo cual incrementa el nivel de incertidumbre con que se realiza la predicción econométrica. Por otra parte, cuando el futuro puede suponer una alteración de tendencias de comportamiento respecto al pasado reciente, es recomendable utilizar estos modelos para predecir a mediano plazo (1 a 5 años), por lo cual nos resultarán de gran ayuda.

Ya hemos discutido la importancia de los modelos ARMA para representar series estacionarias. Una generalización de estos modelos, que incorporan un amplio rango de series no estacionarias, es proporcionado por los procesos ARIMA, es decir, procesos que se reducen a procesos ARMA cuando se los diferencia infinitamente muchas veces. terminológicamente hablando: decimos que  $Y_t$  está integrado de orden 1,  $I(1)$ , porque tiene que ser diferenciado una vez para obtener una serie temporal estacionaria. En general, una serie puede ser  $I(d)$ , si debe ser diferenciada "d" veces para obtener una serie estacionaria.

**Definición 2.7.1** Si  $d$  es un entero no negativo, entonces  $X_t$  es un proceso ARIMA  $(p, d, q)$  si

$$Y_t := (1 - B)^d X_t, \quad (2.28)$$

es un proceso ARMA causal  $(p, q)$ .

Esta definición significa que  $X_t$  satisface una ecuación en diferencia de la forma

$$\phi^*(B)X_t \equiv \phi(B)(1 - B)^d X_t = \theta(B)Z_t, \quad (2.29)$$

donde  $Z_t \sim WN(0, \sigma^2)$ ,  $\phi(z)$  y  $\theta(z)$  son polinomios de grados  $p$  y  $q$ , respectivamente, y  $\phi(z) \neq 0$  para  $|z| \leq 1$ . El polinomio  $\phi^*(z)$  tiene un cero de orden  $d$  en  $z = 1$ . El proceso  $X_t$  es estacionario si y solo si  $d = 0$ , en cuyo caso se reduce a un proceso ARMA  $(p, q)$ . Tenga en cuenta que si  $d \geq 1$ , podemos agregar una tendencia polinómica arbitraria de grado  $(d-1)$  a  $X_t$  sin violar la ecuación de diferencia (2.32). Los modelos ARIMA son, por lo tanto, útiles para representar datos con tendencia.

Sin embargo, los procesos ARIMA también pueden ser apropiados para modelar series sin tendencia.

NOTA: Para trabajar con los modelos ARMA y ARIMA nuestra serie de tiempo necesariamente debe ser estacionaria.

**Definición 2.7.2** *Suavizado Logarítmico: Consta de tomarle el Logaritmo en base 10 o logaritmo neperiano a los datos para alisar o suavizar la serie de tiempo.*

Las desviaciones de la estacionariedad pueden ser sugeridas por el gráfico de la serie misma o por la función de autocorrelación de muestra o ambos.

#### **Prueba de Dickey-Fuller:**

La Prueba de Dickey-Fuller busca determinar la existencia o no de raíces unitarias en una serie de tiempo. La hipótesis nula de esta prueba es que existe una raíz unitaria en la serie.

Esta prueba contrasta la hipótesis nula de existencia de una raíz unitaria contra la alternativa de que no existen raíces unitarias. Para verificar la existencia de una raíz unitaria en un proceso AR(p), se lleva a cabo la siguiente prueba:

$$H_0 : \rho = 1 \quad \text{versus} \quad H_a : \rho < 1 \quad (2.30)$$

aplicando la regresión

$$X_t = c_t + \rho X_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta X_{t-i} + \epsilon_t, \quad (2.31)$$

donde  $c_t$  es una función determinista del tiempo  $t$  y  $\Delta X_j = X_j - X_{j-1}$  es la serie de diferencias de  $X_t$  usada para aproximar la estructura ARMA de los errores, y el valor de  $\rho$  se fija de modo que el error  $\epsilon_t$  sea correlacionado serialmente. En la práctica,  $c_t$  puede ser cero o una constante o bien  $c_t = w_0 + w_1 t$ . El ADF se basa en las estimaciones de mínimos cuadrados de (1.34) y está dado por

$$ADF - Test = \frac{\hat{\rho} - 1}{std(\hat{\rho})}, \quad (2.32)$$

donde  $\hat{\rho}$  denota el valor estimado por mínimos cuadrados de  $\rho$  y  $std(\hat{\rho})$  su desviación estandar.

Fundamento econométrico de la prueba: En un simple modelo autorregresivo de orden (1):

$$X_t = \rho X_{t-1} + \epsilon_t. \quad (2.33)$$

Donde  $X_t$  es la variable de interés,  $t$  es el índice de tiempo,  $\rho$  es un coeficiente, y  $\epsilon_t$  es el término de error. La raíz unitaria está presente si  $\rho = 1$ . En este caso, el modelo no sería estacionario.

Mientras el valor de  $\rho$  sea más cercano a cero se considerará más estacionaria la serie.

## 2.8. Funciones de Autocorrelación (ACF) y Autocorrelación Parcial (PACF)

### 2.8.1. Función de Autocorrelación (ACF)

Iniciemos mostrando la ACF de un proceso MA(q)  $x_t = \theta(B)w_t$ , donde  $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ . Dado que  $x_t$  es una combinación lineal de términos de ruido blanco, el proceso es estacionario con media

$$\mathbb{E}(x_t) = \sum_{j=0}^q \theta_j \mathbb{E}(w_{t-j}) = 0,$$

donde podemos escribir  $\theta_0 = 1$ , y la función de autocovarianza es

$$\begin{aligned}\gamma(h) = \text{cov}(x_{t+h}, x_t) &= \mathbb{E} \left[ \left( \sum_{j=0}^q \theta_j w_{t+h-j} \right) \left( \sum_{k=0}^q \theta_k w_{t-k} \right) \right] \\ &= \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & \text{si } 0 \leq h \leq q \\ 0, & \text{si } h > q. \end{cases}\end{aligned}\quad (2.34)$$

La ACF de un MA(q):

$$\rho(h) = \begin{cases} \frac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{1 + \theta_1^2 + \dots + \theta_q^2}, & \text{si } 1 \leq h \leq q \\ 0, & \text{si } h > q. \end{cases}\quad (2.35)$$

Para un modelo ARMA(p,q) causal  $\phi(B)x_t = \theta(B)w_t$ , donde los ceros de  $\phi(z)$  están fuera del círculo unitario, podemos escribir

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}.$$

Se sigue inmediatamente que  $\mathbb{E}(x_t) = 0$ . También, la función de autocovarianza de  $x_t$  se puede escribir como

$$\gamma(h) = \text{cov}(x_{t-h}, x_t) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}, \quad h \geq 0. \quad (2.36)$$

Una ecuación general homogénea para la ACF de un proceso ARMA causal:

$$\gamma(h) - \phi_1 \gamma(h-1) - \dots - \phi_p \gamma(h-p) = 0, \quad h \geq \max(p, q+1) \quad (2.37)$$

con condiciones iniciales

$$\gamma(h) - \sum_{j=1}^p \phi_j \gamma(h-j) = \sigma_w^2 \sum_{j=h}^q \theta_j \psi_{j-h}, \quad 0 \leq h < \max(p, q+1). \quad (2.38)$$

Dividiendo (2.37) y (2.38) por  $\gamma(0)$  nos permite resolver la ACF  $\rho(h) = \gamma(h)/\gamma(0)$ .

## 2.8.2. Función de Autocorrelación Parcial (PACF)

Vimos en (2.35) que para un modelo MA(q) La ACF será cero para pasos mayores que  $q$ . Más aún, dado que  $\theta_q \neq 0$ , la ACF no es cero en paso  $q$ . Por lo tanto, la ACF proporciona una cantidad considerable de información sobre el orden de dependencia, cuando el proceso es un proceso de promedio móvil. Si el proceso, sin embargo, es ARMA o AR, la ACF solo nos dice un poco sobre el orden de dependencia. Por lo tanto, vale la pena buscar una función que se comporte como la ACF de los modelos MA, pero para los modelos AR, esta será la función de autocorrelación parcial.

**Definición 2.8.1** La función de autocorrelación parcial (PACF) de un proceso estacionario  $x_t$  denotada  $\phi_{hh}$ , para  $h = 1, 2, \dots$ , es

$$\phi_{11} = \text{corr}(x_1, x_0) = \rho(1) \tag{2.39}$$

y

$$\phi_{hh} = \text{corr}(x_h - x_h^{h-1}, x_0 - x_0^{h-1}) \text{ para } h \geq 2. \tag{2.40}$$

Tanto  $(x_h - x_h^{h-1})$  como  $(x_0 - x_0^{h-1})$  son no-correlacionados con  $\{x_1, x_2, \dots, x_{h-1}\}$ .

Por estacionaridad, la PACF  $\phi_{hh}$  es la correlación entre  $x_t$  y  $x_{t-h}$  con la dependencia lineal  $\{x_{t-1}, \dots, x_{t-(h-1)}\}$  removida en cada uno.

Si el proceso  $x_t$  es gaussiano, entonces  $\phi_{hh} = \text{corr}(x_t, x_{t-h} | x_{t-1}, \dots, x_{t-(h-1)})$ . Esto es,  $\phi_{hh}$  es el coeficiente de correlación entre  $x_t$  y  $x_{t-h}$  en la distribución bivariada de  $(x_t, x_{t-h})$  condicionada por  $\{x_{t-1}, \dots, x_{t-(h-1)}\}$ .

En la siguiente sección discutiremos los métodos para calcular la PACF. La PACF para los modelos MA se comporta como el ACF para los modelos AR. También, la PACF para modelos AR se comporta como la ACF para modelos MA. Debido a que un modelo ARMA invertible tiene una representación AR infinita, la PACF no tendrá corte. Resumimos estos resultados en el Cuadro 2.1

	AR(p)	MA(q)	ARMA(p,q)
ACF	Disminución gradual	Corte después de paso $q$	Disminución gradual
PACF	Corte después de paso $q$	Disminución gradual	Disminución gradual

Tabla 2.1: Comportamiento de la ACF y la PACF para modelos ARMA causal e invertible

## 2.9. Modelos SARIMA

La media móvil integrada autorregresiva, o ARIMA, es uno de los métodos de pronóstico más utilizados para la predicción de datos de series de tiempo univariadas. Aunque el método puede manejar datos con una tendencia, no admite series de tiempo con un componente estacional. Una extensión de ARIMA que admite el modelado directo del componente estacional de la serie se llama SARIMA (Seasonal Autoregressive Integrated Moving Average) o ARIMA estacional (seasonal ARIMA).

**Definición 2.9.1** Si  $d$  y  $D$  son enteros no negativos, entonces  $X_t$  es una  $SARIMA(p, d, q) \times (P, D, Q)_s$  proceso con el período  $s$  si las series diferenciadas  $Y_t = (1 - B)^d(1 - B^s)^D X_t$  Es un proceso causal ARMA definido por

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t, \quad Z_t \sim WN(0, \sigma^2), \quad (2.41)$$

donde  $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ ,  $\Phi(z) = 1 - \Phi_1 z - \dots - \Phi_p z^p$ ,  $\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$  y  $\Theta(z) = 1 + \Theta_1 z + \dots + \Theta_q z^q$

### Elementos de tendencia

Hay tres elementos de tendencia que requieren configuración. Son los mismos que el modelo ARIMA; específicamente:

- \*  $p$  : orden de autorregresión de la tendencia.
- \*  $d$  : orden de diferencia de tendencia.
- \*  $q$  : Tendencia media móvil de orden.

### Elementos estacionales

Hay cuatro elementos estacionales que no forman parte de ARIMA que deben configurarse; son:

- \*  $P$  : Orden autorregresivo estacional.
- \*  $D$  : orden de diferencia estacional.
- \*  $Q$  : orden de media móvil estacional
- \*  $s$  : El número de pasos de tiempo para un solo período estacional.

# Aprendizaje Supervisado y Redes Neuronales

## 3.1. Aprendizaje Supervisado

El Aprendizaje Automático o mejor conocido como **Machine Learning** es un término que se escucha cada vez más con mayor frecuencia, dado que proporciona valor al negocio y arroja un nuevo uso y sentido a los datos con que cuentan las organizaciones.

A pesar de ser un término recientemente muy sonado y en ocasiones relacionado con el concepto de inteligencia artificial, lo cual de hecho, es bastante cierto, pues está catalogado como un subconjunto del mismo, sus inicios tienen raíces en los años 50 en el MIT. Hoy en día el aprendizaje automático está presente en muchas de las actividades que nos rodean, los asistentes de voz de Siri y Alexa, el reconocimiento facial de Facebook y Microsoft, las recomendaciones de Amazon y Netflix y la tecnología que evita que los vehículos autónomos se estrellen, son el resultado de los avances en el aprendizaje automático.

### 3.1.1. Aprendizaje Supervisado

**Definición 3.1.1** *Se genera una función que establece una correspondencia entre las entradas y las salidas deseadas del sistema, donde la base de conocimientos del sistema está formada por ejemplos ya etiquetados (es decir, ejemplos de los que sabemos su clasificación correcta). Este consta de dos fases, una de entrenamiento y una de prueba;*



1. En la fase de entrenamiento se cuenta con un conjunto de datos que por lo general está entre el 60 % o 70 % del total de los datos disponibles, que son con los que se entrena al algoritmo para encontrar los patrones y relaciones en el conjunto de datos.
2. Posteriormente en la fase de prueba, se cuenta con un conjunto de datos de prueba entre el 40 % o 30 % del total de los datos disponibles, el cual sirve para validar el rendimiento del algoritmo.

### 3.1.2. Tipos de Aprendizajes Supervisados

1. En un problema de **Regresión**, tratamos de predecir los resultados con una salida continua, lo que significa que estamos tratando de asignar variables de entrada para una función continua. Por ejemplo si se tienen los datos de los precios de alquiler y las características de las viviendas (superficie, año, etc.) se puede predecir el precio. Si la función precio en función de las características tuviera una salida de datos continua
2. En un problema de **Clasificación**, se trata por tanto de predecir los resultados de una salida/output discreta, es decir, tratamos de asignar variables de entrada en categorías discreta. Por ejemplo, clasificar si una determinada noticia es de deportes, entretenimiento, política, etc. Si sólo se permiten 2 posibles clases, entonces se llama clasificación binaria; si se permiten más de 2 clases, estamos hablando de clasificación multiclase. Como identificar si un tumor es maligno o benigno, dado su tamaño y edad del paciente.

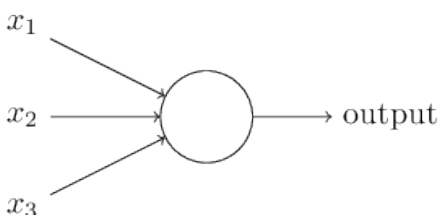
## 3.2. Redes Neuronales

Últimamente las redes neuronales están volviendo a la actualidad por los logros que están consiguiendo. Por ejemplo, Google ha logrado derrotar a su propio reCAPTCHA con redes neuronales, en Stanford han conseguido generar pies de fotos automáticamente. Metas bastante impresionantes y que cada vez se acercan más a esa idea original de reproducir el funcionamiento del cerebro humano en un ordenador.

El nombre, viene de la idea de imitar el funcionamiento de las redes neuronales de los organismos vivos, un conjunto de neuronas conectadas entre sí y que trabajan en conjunto, sin que haya una tarea concreta para cada una. Con la experiencia, las neuronas van creando y reforzando ciertas conexiones para "aprender" algo que se queda fijo en el tejido.

### 3.2.1. Perceptrones

Un perceptrón toma varias entradas binarias,  $x_1, x_2, \dots$ , y produce una salida binaria única:



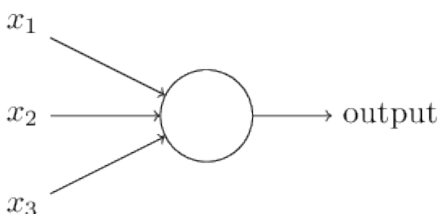
En el ejemplo mostrado, el perceptrón tiene tres entradas,  $x_1, x_2, x_3$ . En general podría tener más o menos entradas.

Rosenblatt propuso una regla simple para calcular la salida. Introdujo pesos,  $w_1, w_2, \dots$ , números reales que expresan la importancia de las entradas respectivas a la salida. La salida se determina si la suma ponderada  $\sum_j w_j X_j$  es menor o mayor que algún valor de umbral. Al igual que los pesos, el umbral es un número real que es un parámetro de la neurona. Para ponerlo en términos algebraicos más precisos:

$$\text{salida} = \begin{cases} 0 & \text{si } \sum_j w_j X_j \leq \text{umbral} \\ 1 & \text{si } \sum_j w_j X_j > \text{umbral} \end{cases} \quad (3.1)$$

### 3.2.2. Sigmoide

Las neuronas sigmoides son similares a los perceptrones, pero se modifican de modo que pequeños cambios en sus pesos y sesgos causan solo un pequeño cambio en su producción. Ese es el hecho crucial que permitirá que una red de neuronas sigmoideas aprendan. Representaremos las neuronas sigmoideas de la misma manera que describimos los perceptrones



Al igual que un perceptrón, la neurona sigmoide tiene entradas,  $x_1, x_2, \dots$ . Pero en lugar de ser solo 0 o 1, estas entradas también pueden tomar cualquier valor entre 0 y 1. Entonces, por ejemplo, 0.638. Es una entrada

válida para una neurona sigmoide. También al igual que un perceptrón, la neurona sigmoide tiene pesos para cada entrada,  $w_1, w_2, \dots$ , y un sesgo general,  $b$ . Pero la salida no es 0 o 1. En su lugar, es  $\sigma(w \cdot x + b)$ , donde  $\sigma$  se llama la función sigmoidea y se define por:

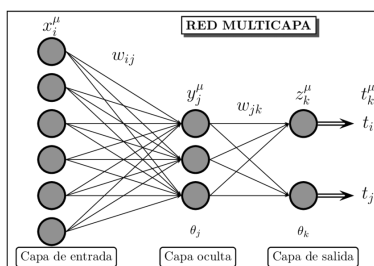
$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (3.2)$$

Para ponerlo todo un poco más explícitamente, la salida de una neurona sigmoide con entradas  $x_1, x_2, \dots$ , peso  $w_1, w_2, \dots$  y sesgo  $b$  es

$$\frac{1}{1 + e^{-\sum_j w_j X_j - b}} \quad (3.3)$$

### 3.2.3. MLP (Multi-Layer Perceptron)

El perceptrón multicapa es una red neuronal artificial formada por múltiples capas, esto le permite resolver problemas que no son linealmente separables, lo cual es la principal limitación del perceptrón. El perceptrón multicapa puede ser totalmente o localmente conectado. En el primer caso cada salida de una neurona de la capa "i" es entrada de todas las neuronas de la capa "i+1", mientras que en el segundo cada neurona de la capa "i" es entrada de una serie de neuronas (región) de la capa "i+1".



Las capas pueden clasificarse en tres tipos:

- \* Capa de entrada: Constituida por aquellas neuronas que introducen los patrones de entrada en la red. En estas neuronas no se produce procesamiento.
- \* Capas ocultas: Formada por aquellas neuronas cuyas entradas provienen de capas anteriores y cuyas salidas pasan a neuronas de capas posteriores.
- \* Capa de salida: Neuronas cuyos valores de salida se corresponden con las salidas de toda la red.

La propagación hacia atrás (también conocido como retropropagación del error o regla delta generalizada), es un algoritmo utilizado en el entrenamiento de estas redes, por ello, el perceptrón multicapa también es conocido como red de retropropagación.

### 3.2.4. Descenso del Gradiente

El Descenso del Gradiente es el algoritmo de entrenamiento más simple y también el más extendido y conocido. Solo hace uso del vector gradiente, y por ello se dice que es un método de primer orden.

Este método para construir el punto  $w_{i+1}$  a partir de  $w_i$  se traslada este punto en la dirección de entrenamiento  $d_i = -g_i$ . Es decir:

$$w_{i+1} = w_i - g_i v_i$$

Donde el parámetro  $v$  se denomina tasa de entrenamiento, que puede fijarse a priori o calcularse mediante un proceso de optimización unidimensional a lo largo de la dirección de entrenamiento para cada uno de los pasos (aunque esta última opción es preferible, a menudo se usa un valor fijo,  $v_i = v$  con el fin de simplificar el proceso).

Aunque es muy sencillo, este algoritmo tiene el gran inconveniente de que, para funciones de error con estructuras con valles largos y estrechos, requiere muchas iteraciones. Se debe a que, aunque la dirección elegida es en la que la función de error disminuye más rápidamente, esto no significa que necesariamente produzca la convergencia más rápida.

Por ello, es el algoritmo recomendado cuando tenemos redes neuronales muy grandes, con muchos miles de parámetros, ya que sólo almacena el vector gradiente (de tamaño  $n$ ), pero no hace uso de la Hessiana (de tamaño  $n^2$ ).

### 3.2.5. Red Neuronal Convolutiva CNN

Una red neuronal convolutiva (CNN) es una capa artificial de múltiples capas. Red neuronal que ha tenido éxito reconociendo patrones visuales. La arquitectura más común de la CNN es una pila de tres tipos de capas múltiples: Capa convolutiva, capa de submuestreo y capa totalmente conectada. Convencionalmente, un CNN consiste en capas alternativas de convolutiva capas y sub-muestreo de capas en la parte inferior y varias capas totalmente conectadas siguiéndolas.

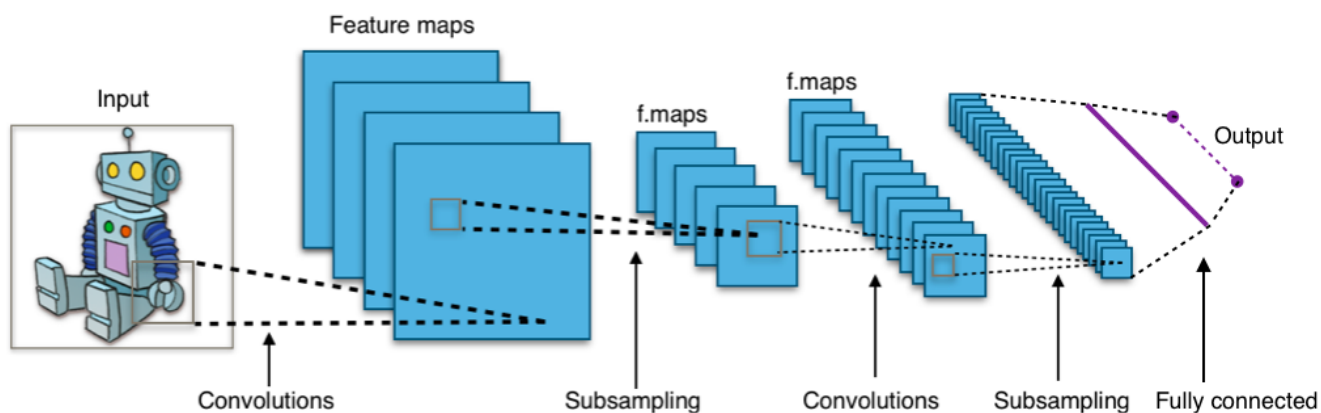
Primero, una unidad de una capa convolutiva recibe entradas de un conjunto de nodos vecinos de la capa anterior. Los pesos locales de las capas convolutivas se comparten con los nodos en la misma capa. Tales cálculos locales en la capa reducen la carga de memoria y mejoran el rendimiento de clasificación.

Capa no lineal descendente, que es el segundo tipo de capas CNN, es otra característica importante de CNNs. La idea del submuestreo local es que una vez que una característica se ha detectado, su ubicación en sí no es tan importante como su ubicación relativa con otras características. Reduciendo la dimensionalidad, reduce la sensibilidad local de la red y complejidad computacional (LeCun & Bengio, 1995; LeCun et al., 1998).

El último tipo de capas es la capa totalmente conectada. Calcula un cálculo matricial completo con todas las activaciones, y nodos iguales a las redes neuronales regulares. Después de las capas convolucionales y sub-muestreo, las capas totalmente conectadas implementan el razonamiento y dan la salida real. Entonces el modelo se entrena en la forma de minimizar el error entre la salida real del modelo y los valores de salida de destino por el método de propagación hacia atrás.

Las CNN son muy efectivas para la resolución de ciertos problemas como: problemas de visión, como la clasificación, detección de objetos y segmentación semántica. También se aplica a otros problemas como el lenguaje natural. Recientemente, se aplican variantes de CNN al análisis de varios tipos de series de tiempo como valores de sensores y EEG. Señales (electroencefalograma).

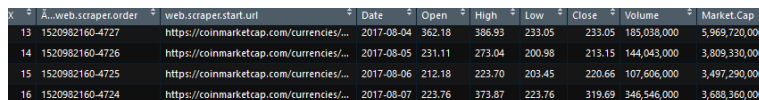
La siguiente imagen nos muestra la arquitectura de las CNN :



# Capítulo 4

## Análisis de Resultados y Conclusiones

Para la realización de ésta etapa resulta indispensable la obtención de los datos históricos de las diferentes criptomonedas a estudiar, que son el Bitcoin, Litcoin, Bitcoin Cash, Ether y Ripple, extraídos en la página "<https://coinmarketcap.com/>", estructurados de la siguiente manera



	web.scra...order	web.scra...start.url	Date	Open	High	Low	Close	Volume	Market.Cap
13	1520982160-4727	https://coinmarketcap.com/currencies/...	2017-08-04	362.18	386.93	233.05	233.05	185,038,000	5,969,720,000
14	1520982160-4726	https://coinmarketcap.com/currencies/...	2017-08-05	231.11	273.04	200.98	213.15	144,043,000	3,809,330,000
15	1520982160-4725	https://coinmarketcap.com/currencies/...	2017-08-06	212.18	223.70	203.45	220.66	107,606,000	3,497,290,000
16	1520982160-4724	https://coinmarketcap.com/currencies/...	2017-08-07	223.76	373.87	223.76	319.69	346,546,000	3,688,360,000

Figura 4.1: Estructura de los datos

Para la realización de los modelos SARIMA las columnas que se utilizaran son, Date (la fecha en la que ocurrió cada observación) y Close (nos indica el precio de cierre registrado por la criptomoneda el día de la observación), luego para las Redes Neuronales se usaran, Open (nos indica el precio inicial de la criptomoneda en el día de la observación), High (el precio máximo alcanzado por la criptomoneda el día de la observación), Low (el precio más bajo que alcanzo la criptomoneda el día de la observación) y Close. Los cuáles nos permitirían hacer los análisis correspondientes. Para los estudios de dichos datos utilizaremos el software estadístico R, en su versión Rstudio y Phyton.

### 4.1. R (lenguaje de programación)

R es un entorno y lenguaje de programación con un enfoque al análisis estadístico. R es una implementación de software libre del lenguaje S pero con soporte de alcance estático. Se trata de uno de los lenguajes

más utilizados en investigación por la comunidad estadística, siendo además muy popular en el campo de la minería de datos, la investigación biomédica, la bioinformática y las matemáticas financieras. A esto contribuye la posibilidad de cargar diferentes bibliotecas o paquetes con funcionalidades de cálculo y gráficas. R es parte del sistema GNU y se distribuye bajo la licencia GNU GPL. Está disponible para los sistemas operativos Windows, Macintosh, Unix y GNU/Linux.

## 4.2. Python

Python es un lenguaje de programación interpretado cuya filosofía hace hincapié en una sintaxis que favorezca un código legible. Se trata de un lenguaje de programación multiparadigma, ya que soporta orientación a objetos, programación imperativa y en menor medida, programación funcional. Es un lenguaje interpretado, usa tipado dinámico y es multiplataforma. Es administrado por la Python Software Foundation. Posee una licencia de código abierto, denominada Python Software Foundation License que es compatible con la Licencia pública general de GNU a partir de la versión 2.1.1, e incompatible en ciertas versiones anteriores.

## 4.3. Análisis de Series de Tiempo

A continuación veremos el análisis de las series de tiempo de las diferentes criptomonedas, a estas series les serán ajustados los modelos SARIMA arrojados por la función "auto.arima" del lenguaje de programación R.

### 4.3.1. Estudio y Predicción del Bitcoin

A continuación se presenta el comportamiento de los precios de cierre de la Criptomoneda Bitcoin desde el 28 de Abril de 2013 hasta 11 de Marzo de 2018.

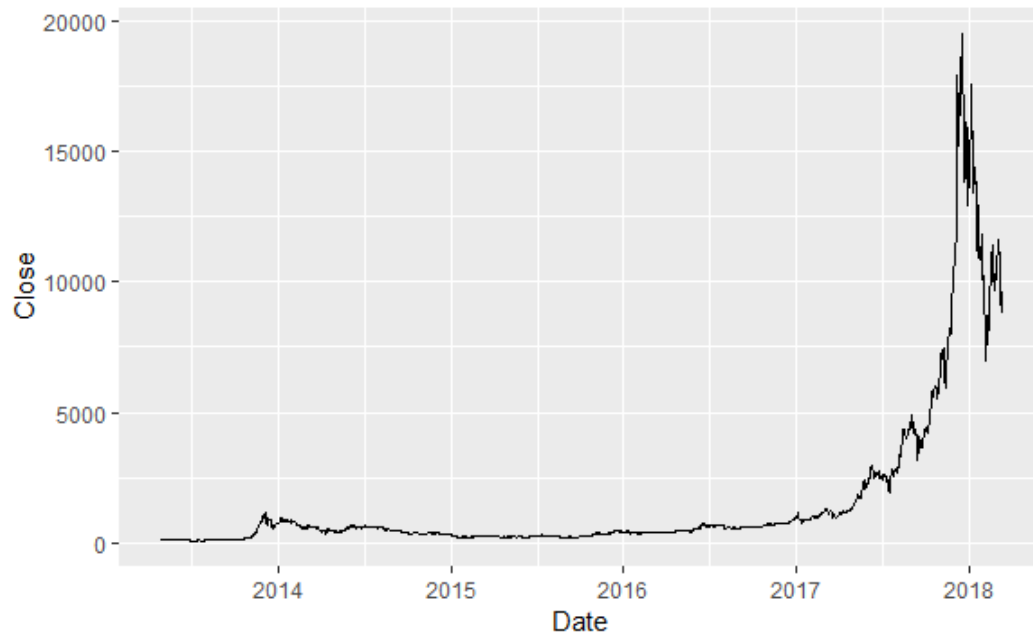


Figura 4.2: Serie tiempo de los precios del Bitcoin desde 2013

La serie de los precios del Bitcoin alcanzó su mínimo el 05/07/2013 en este día el valor de cierre fue de \$68.43, además tal y como se puede apreciar en la figura 4.2, entre 28/04/2013 hasta aproximadamente el primer trimestre del año 2017 la serie no muestra mayores fluctuaciones. Sin embargo a partir de este punto se comienza a apreciar una variabilidad en los precios. El 16/02/2017 la serie alcanza su máximo valor, este día el precio del Bitcoin llegó a \$19497.40.



Con la finalidad de entender la variabilidad de los precios de esta Criptomoneda presentada en los últimos años truncamos la serie de los precios y consideramos para el análisis el periodo comprendido entre el 01/01/2017 hasta el 11/03/2018. El gráfico de esta serie truncada lo podemos apreciar en la figura 4.3.

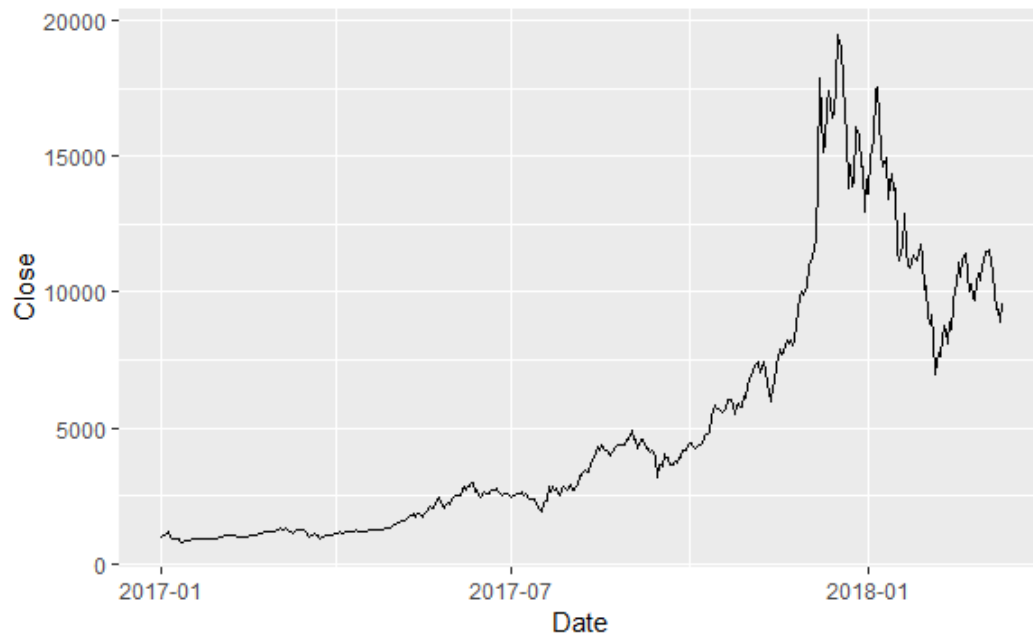


Figura 4.3: Serie tiempo de los precios del Bitcoin desde 01/01/2017

Este procedimiento además garantiza que el modelo ajustado tenga un mejor desempeño pues este no se ve alterado, en su aprendizaje, por valores básicamente constantes.

En la búsqueda de patrones y características globales de la serie se procedió a realizar una segmentación de la serie por los meses del año. El gráfico de estas segmentaciones se puede apreciar a continuación.

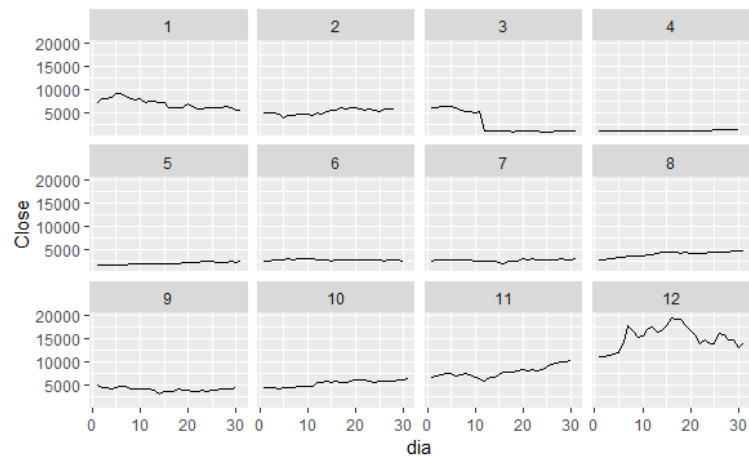


Figura 4.4: Gráficas para verificar la tendencia

La figura 4.4 se divide en doce cuadros que representan cada mes del año. Al examinar este gráfico notamos que no hay una tendencia o patrón persistente en los datos .

Dado que la serie tiene una escala muy amplia de valores, se aplica un suavizado logarítmico a los datos, esto proporciona datos menos distorsionados al comportamiento real de misma. En la figura 4.5 se aprecia como quedó la serie suavizada.

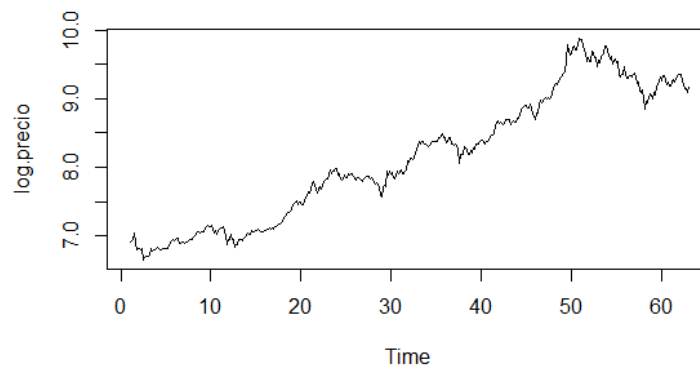


Figura 4.5: Aplicando un suavizado logarítmico a la serie a partir de 01/01/2017

Una vez suavizada la serie se procedió a extraer la tendencia y el componente estacional de la misma. Estos patrones se pueden apreciar en la figura 4.6.

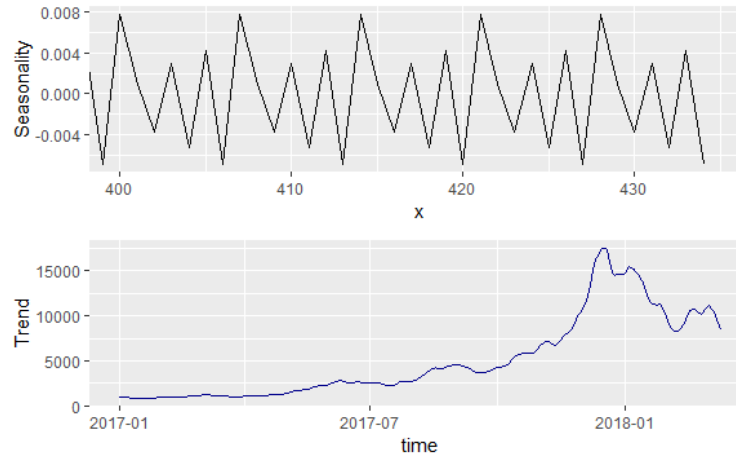


Figura 4.6: Gráfica de la tendencia y la estacionaridad

Podemos observar en la figura anterior que la serie tiene una tendencia creciente además, notamos una estacionaridad que se repite cada 7 días.

Posteriormente se aplicó el comando "auto.arima", de R, a la serie de los logaritmos, para definir cuales eran los parametros de nuestro modelo SARIMA, este arrojó como resultado un  $ARIMA(0, 1, 0)(2, 0, 0)[7]$ , podemos observar que el modelo hace una diferenciación para eliminar la tendencia y propone una estacionaridad cada 7 días, lo cual concuerda con los patrones que fueron determinados en los pasos previos y que son expuestos en la figura 4.6. Dada esta concordancia se acepta este modelo propuesto por la función "auto.arima", que además se caracteriza por ser el de menor AIC.

El modelo es entrenado con los datos desde el 01/01/2017 hasta 11/03/2018, para luego realizar el pronóstico, el cual sera de 15 días. Posteriormente se realiza el cálculo del error de estimación mediante la siguiente ecuación la cual es llamada porcentaje de error

$$PorcentajeError = \frac{ValorReal - ValorPronosticado}{ValorReal}. \quad (4.1)$$

Donde el valor real representa el precio real de cierre de la criptomneda y el valor pronosticado representa el valor del precio estimado mediante el modelo SARIMA. El porcentaje de error medio fue de un 18 % lo que nos indica que el modelo tiene un 82 % eficacia.

A continuación presentamos la gráfica de los precios de cierre del Bitcoin seguido de 15 días de predicción obtenidos mediante el modelo SARIMA.

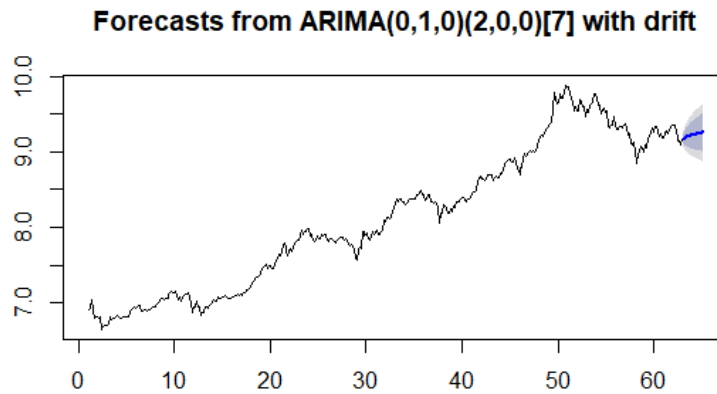


Figura 4.7: Pronostico del modelo

En la figura 4.8 se puede apreciar el gráfico de los valores reales de los precios del Bitcoin y el valor del precio que predijo el modelo. Notamos que el valor de predicción queda por encima del valor real. Esto puede deberse a que la serie de entrenamiento tiene una tendencia de pendiente pronunciada en casi todo el periodo.

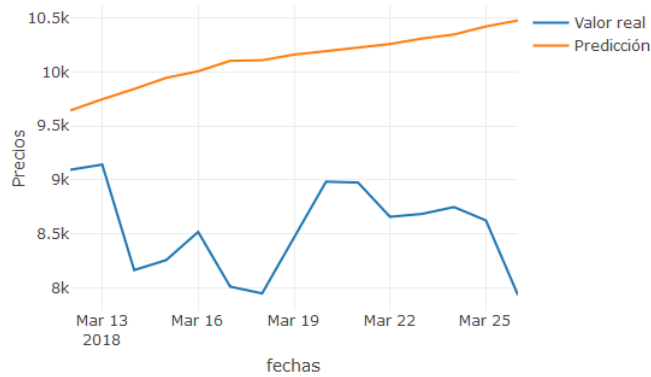


Figura 4.8: Valores reales y predicción del Bitcoin

---

Valor Real	Valor Pronosticado
9094.55	9643.384
9142.76	9747.654
8165.33	9842.976
8257.94	9945.862
8519.09	10006.989
8012.56	10103.247
7949.53	10109.145
8471.54	10161.305
8984.42	10193.162
8976.22	10224.007
8658.83	10259.149
8685.86	10309.771
8749.37	10347.455
8625.36	10422.235
7935.02	10477.427

Tabla 4.1: Tabla precios reales y precios pronosticados del Bitcoin.

En la tabla anterior comparamos el valor real del precio con el pronostico. En particular en los primeros 3 días de predicción se aprecia una proximidad con el valor real.

### 4.3.2. Estudio y Predicción Bitcoin Cash

En lo que sigue analizaremos el comportamiento de cierre de la criptomoneda Bitcoin Cash desde 23 Julio de 2017 hasta 12 de Marzo de 2018

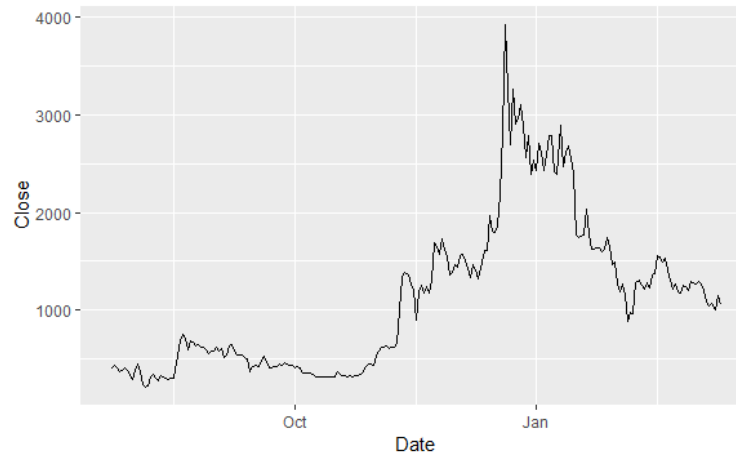


Figura 4.9: Serie tiempo de los precios del Bitcoin Cash desde 23/07/2017

La figura 4.9 nos muestra la serie de los precios de cierre del Bitcoin Cash, alcanzó su valor máximo el día 20/12/2017 el cual fue de \$3923.1 y el 05/08/2017 la serie alcanza su valor mínimo, este día el precio de cierre fue de \$213.2.

Para tratar de encontrar patrones y características en las serie se realizó una segmentación de la serie por los meses de año.

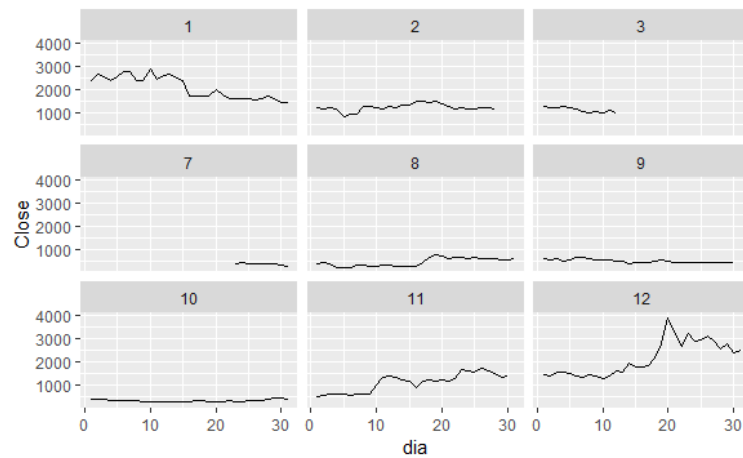


Figura 4.10: Gráficas para verificar la tendencia

En figura 4.10 no podemos notar una tendencia o patrón persistente en los datos al dividirlo en los meses del año. Por otra parte el mes de Marzo (3) y Julio (7) vemos que no están los días completos ya que para estos días los precios de esta criptomoneda no fueron registrados.

En vista de que la serie tiene una escala muy grande de valores, se le realiza un suavizado logarítmico a los datos. En la siguiente gráfica podemos observar como queda la serie luego del suavizado.

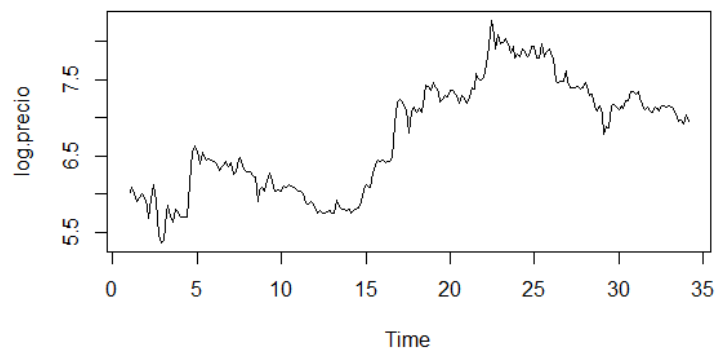


Figura 4.11: Aplicando un suavizado logarítmico a la serie

Luego de suavizar la serie extraemos la tendencia y la componente estacional de la misma.

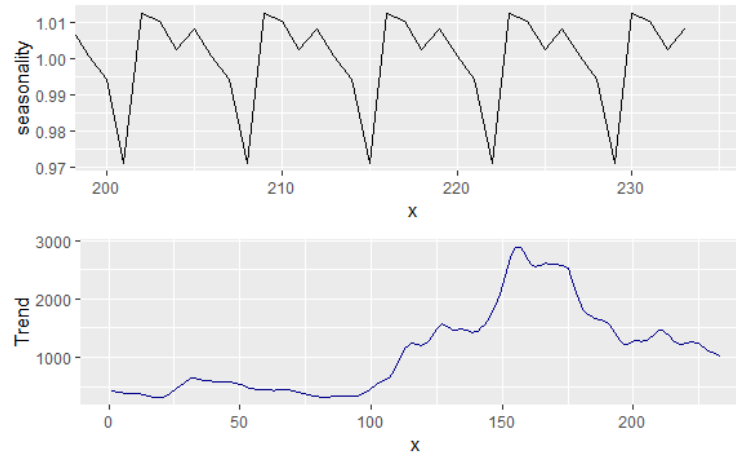


Figura 4.12: Gráfica de la tendencia y la estacionalidad

Podemos observar en la figura 4.12 que la serie presenta una tendencia creciente por otra parte vemos una estacionaridad bien marcada que se repite cada 7 días.

Procedemos a realizar el modelo con el comando "auto.arima" aplicándolo a la serie luego de realizarle el suavizado logarítmico, dicha instrucción propone que se aplique a los datos un modelo SARIMA con los siguientes parámetros  $ARIMA(2, 1, 2)(2, 0, 0)[7]$ , vemos que el modelo hace una diferenciación para eliminar la tendencia y propone una estacionaridad cada 7 días, lo cual se asemeja con los patrones que ya fueron vistos anteriormente y que se pueden observar en la figura 4.12. Dada con esta correspondencia se acepta el modelo propuesto por la función "auto.arima", que se caracteriza por ser el de menor AIC.

El modelo es entrenado con todos los datos es decir desde el 23/07/2017 hasta el día 12/03/2018 y con este se generó una predicción para los siguientes 15 días, la cual se puede apreciar en el siguiente gráfico.



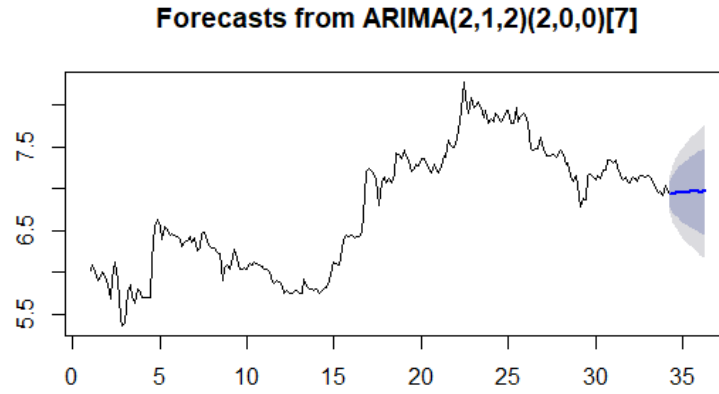


Figura 4.13: Pronostico del modelo

En la figura 4.14 vemos la gráfica de los valores reales del Bitcoin Cash y el valor del precio que ha predicho el modelo.

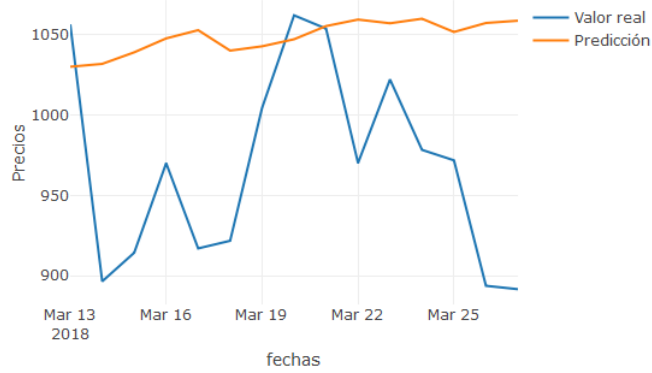


Figura 4.14: Valores reales y predicción del Bitcoin Cash

---

Valor Real	Valor Pronosticado
1056.38	1030.020
896.53	1031.809
914.34	1038.993
970.03	1047.652
917.03	1052.778
921.78	1040.067
1004.48	1042.733
1062.01	1047.116
1056.73	1055.320
970.17	1059.372
1022.12	1057.054
978.33	1059.894
971.81	1051.643
893.78	1057.246
891.67	1058.658

Tabla 4.2: Tabla precios reales y precios pronosticados del Bitcoin Cash.

En la tabla 4.2 comparamos el valor real del precio con el pronostico. Tanto en la tabla como en el gráfico anterior podemos apreciar como la predicción dada por el modelo se ajusta bastante bien en varios de los días. Estas apreciaciones se corroboran al calcular el porcentaje de error, dado en la ecuación (4.1), pues para este caso el porcentaje de error medio fue de 9 % lo que nos indica que el modelo tiene un 91 % de eficiencia.

### 4.3.3. Estudio y Predicción Ether ETH

Ahora veremos el comportamiento de cierre de la última criptomoneda a estudiar el Ether desde 07 de Agosto de 2015 hasta 11 de Marzo de 2018.

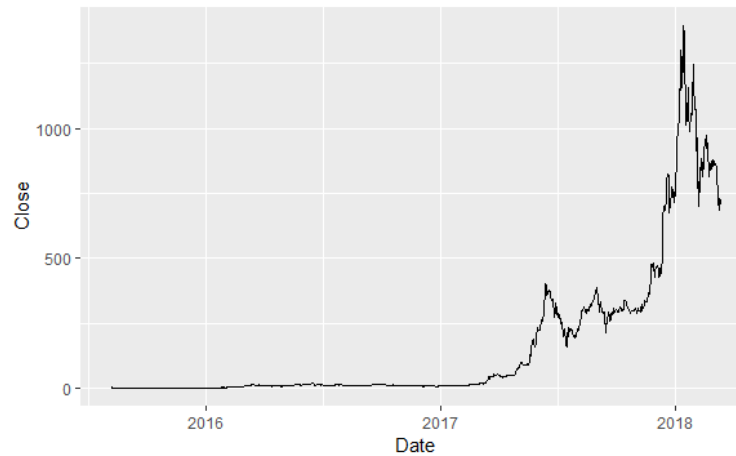


Figura 4.15: Serie tiempo de los precios del Ether ETH desde 2015

En la figura 4.15, vemos la serie de los precios de cierre del Ether, alcanzó su valor máximo el día 13/01/2018 el cual fue de \$1396.42 y el 20/10/2015 alcanzó su valor mínimo el cual fue de \$0.43. Como se puede observar en la gráfica anterior desde el 07/08/2015 hasta aproximadamente mayo del 2017 la serie no presenta muchos cambios pero a partir de este punto comienza a apreciar fluctuaciones en el cierre de los precios.

Para comprender mejor la variabilidad de los precios del Ether truncamos la serie de los precios para el análisis entre 01/01/2017 hasta el 11/03/2018. En la figura 4.16 podemos ver el gráfico de la serie truncada.

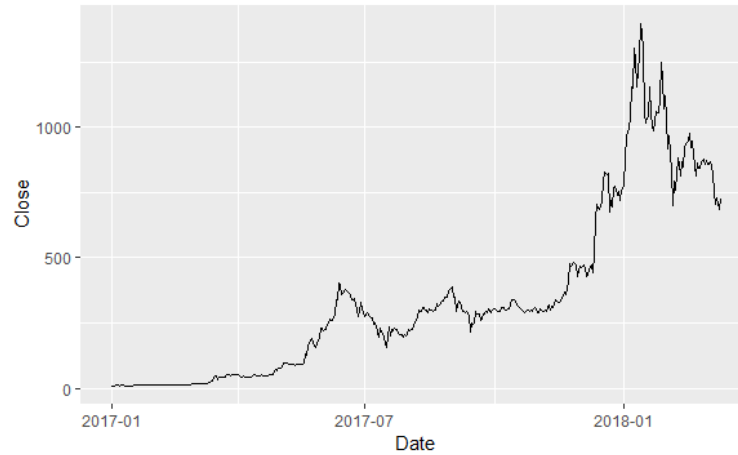


Figura 4.16: Serie tiempo de los precios del Ether desde 01/01/2017

Este truncamiento de la serie garantiza que el modelo ajustado tenga un mejor desempeño pues este no se ve alterado en su aprendizaje por valores básicamente constantes.

Con la finalidad de conseguir patrones y características se hizo una segmentación de la serie por los meses del año. En la siguiente gráfica podemos apreciar estas segmentaciones.

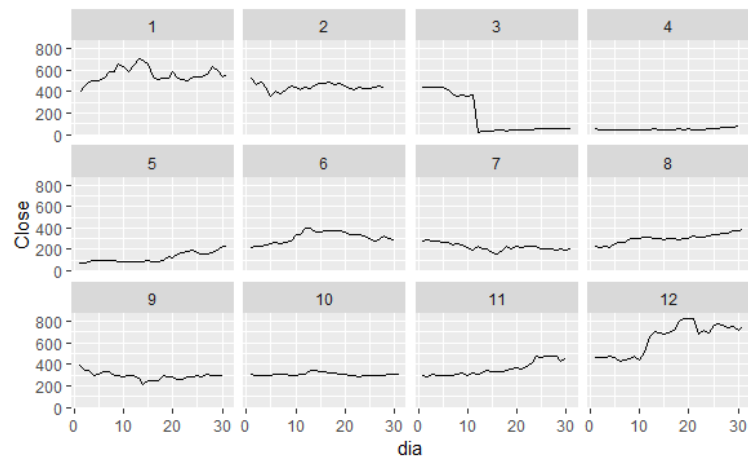


Figura 4.17: Gráficas para verificar la tendencia

En la figura 4.17 no se nota una tendencia o patrón obvio en los datos al dividirlo en intervalos mensuales.

Como la serie tiene una escala muy amplia de valores, se le realiza un suavizado logarítmico a los datos. En la figura 4.18, podemos observar la serie con el suavizado.

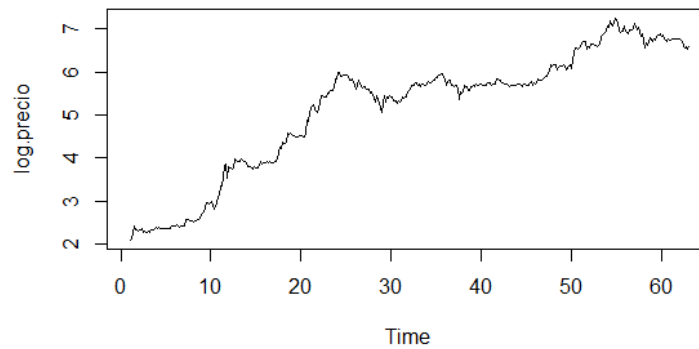


Figura 4.18: Aplicando un suavizado logarítmico a la serie a partir de 01/01/2017

Luego del suavizado se procede a extraer la tendencia y el componente estacional de la serie, que se pueden observar en la siguiente gráfica.

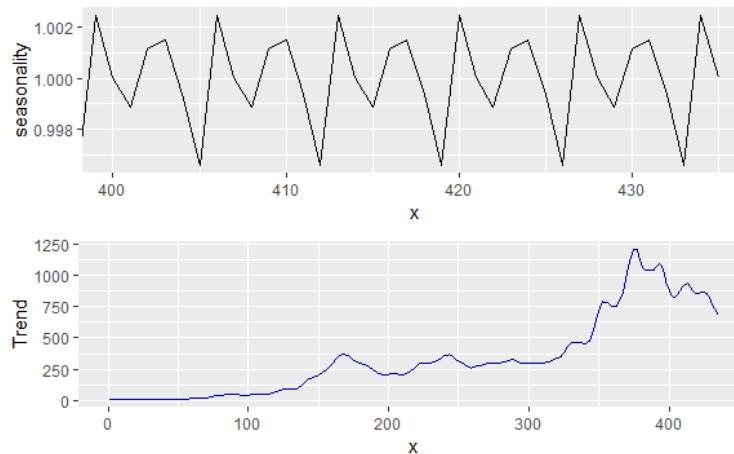


Figura 4.19: Gráfica de la tendencia y la estacionalidad

Podemos observar en la gráfica anterior que la serie tiene una tendencia creciente, por otra parte notamos estacionaridad de 7 días.

Creamos el modelo con el comando "auto.arima" de R, a la serie suavizada, el que nos arroja como resultado un modelo SARIMA con los siguientes parámetros  $ARIMA(0, 1, 0)(1, 0, 0)[7]$ , podemos observar que el modelo hace una diferenciación para eliminar la tendencia y propone una estacionaridad cada 7 días, lo cual concuerda con los patrones que fueron determinados en los pasos anteriores y que son expuestos en la figura 4.19. Dada con esta correspondencia se acepta el modelo propuesto por la función "auto.arima", que se caracteriza por ser el de menor AIC.

El modelo es entrenado con los datos desde el 01/01/2017 hasta 11/03/2018, para luego realizar el pronóstico, el cual será de 15 días. Podemos observar el gráfico del pronóstico en la figura 4.20.

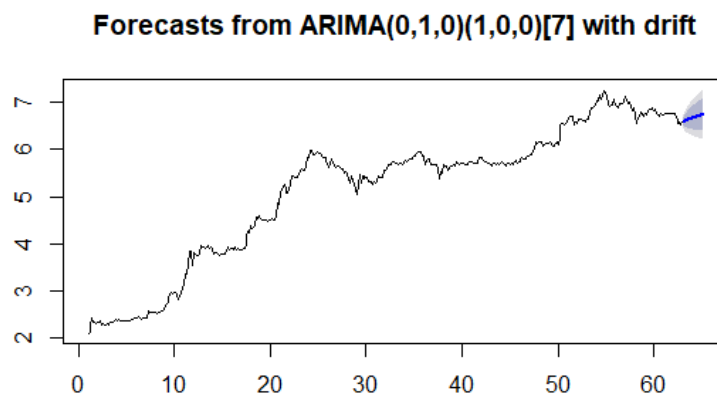


Figura 4.20: Pronostico del modelo

En la figura 4.21 se puede apreciar el gráfico de los valores reales de los precios del Ether y el valor del precio que predijo el modelo. Notamos que el valor de predicción queda por encima del valor real. Esto puede deberse a que la serie de entrenamiento tiene una tendencia de pendiente pronunciada en casi todo el periodo.

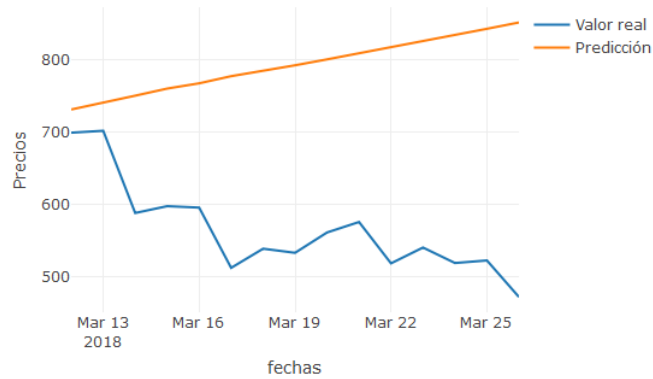


Figura 4.21: Valores reales y predicción del ETH

Valor Real	Valor Pronosticado
699.44	731.5078
702.09	740.5156
588.32	750.6256
597.76	760.4603
595.80	767.7283
512.41	777.6014
538.91	784.5451
533.22	792.6729
561.38	800.8566
575.88	809.0874
518.63	817.4184
540.54	825.9364
519.11	834.4478
522.65	843.1616
472.13	851.9242

Tabla 4.3: Tabla precios reales y precios pronosticados del Ether.

En la tabla 4.3 comparamos el valor real del precio con el valor pronosticado por nuestro modelo. En los primeros 2 días de pronóstico podemos observar una proximidad con el valor real de la criptomoneda. Calculando el porcentaje de error con la ecuación (4.1) tenemos un porcentaje de error medio de 42 % lo que nos dice que nuestro modelo tiene un 58 % de eficiencia.

#### 4.4. Análisis de la Red Neuronal

Uno de los enfoques más usados en el tema del estudio de instrumento financiero es el enfoque de clasificación, cuyo objetivo en general es etiquetar los datos según sus cambios o interacciones de un instante de tiempo a otro. En nuestro caso de estudio este objetivo se adapta a la variación del precio de cierre de la criptomoneda de un día de transacción a otro, específicamente lo que hacemos es comparar el valor del día en curso con el valor del precio de cierre del día anterior, si el precio actual es mayor al precio de cierre del día anterior se coloca la etiqueta o clase [1,0] y en caso contrario se asigna [0,1].



Para realizar un análisis de múltiples variables de las series de tiempo de las distintas criptomonedas, el modelo utilizado fue una red neuronal convolucional "CNN". Para poder aplicar este modelo se construyó un arreglo numpy en python donde se consideran 4 variables a saber "Open", "High", "Low" y "Close", además se tiene otro arreglo donde estan las etiquetas de las observaciones referentes a la subida o bajada del precio mencionadas anteriormente. El modelo para predecir y clasificar un nuevo valor usa la información de los 30 días precedentes.

Se requirió implementar la librería "TensorFlow" de python, se utilizó un modelo secuencial ya que una red de tiempo no es más que una sucesión de puntos o valores a través del tiempo. La arquitectura de la red neuronal CNN que se utilizó, básicamente se conforma por dos capas convolucionales, entre ellas hay varias capas de normalización y de regularización que permiten mejorar los resultados y desempeño de la red y finalmente una última capa de salida donde se usó la función de activación "softmax", esto permite tener una capa de salida con diferentes probabilidades para cada nodo. La suma de estas probabilidades o pesos es igual a uno o equivalentemente al 100% y está dada por la siguiente expresión.

$$\sigma(z_j) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, \quad (4.2)$$

donde  $z$  es un vector de las entradas a la capa de salida,  $j$  indexa las unidades salidas siendo  $j = 1, 2, \dots, k$ .

La función de pérdida o de costo utilizada fue la categorical cross entropy (entropía cruzada), de finida por

$$H(p, q) = - \sum_x p(x) \log q(x) \quad (4.3)$$

donde  $p(x)$  es la probabilidad deseada y  $q(x)$  la probabilidad real.

Es importante en este tipo de análisis poder determinar si se está realizando una buena clasificación de los datos y para esto juega un papel fundamental la llamada matriz de confusión.

		<b>Predicción</b>	
		<b>Positivos</b>	<b>Negativos</b>
<b>Observación</b>	<b>Positivos</b>	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	<b>Negativos</b>	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Figura 4.22: Matriz de Confusión

1. **VP** es la cantidad de positivos que fueron clasificados correctamente como positivos por el modelo.
2. **VN** es la cantidad de negativos que fueron clasificados correctamente como negativos por el modelo.
3. **FN** es la cantidad de positivos que fueron clasificados incorrectamente como negativos.
4. **FP** es la cantidad de negativos que fueron clasificados incorrectamente como positivos.

Otro factor que se debe evaluar para determinar la calidad de ajuste del modelo es el rendimiento del mismo. En este caso este es medido mediante la métrica llamada accuracy que está dada en términos de la matriz de confusión, de la siguiente manera,

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN}. \quad (4.4)$$

Hay que destacar que a dicha red no se le puede pasar la tabla de observaciones así como se tiene debido a la estructura convolucional de la misma. Por ello se debe construir un arreglo, que se puede ver como una matriz de matrices, el cual contiene el número de observaciones de los datos. Donde cada matriz del arreglo tiene asociada dos dimensiones, una asociada a las variables que en este caso es 4 y otra asociada al tamaño de la ventana o "Window", que en este caso es 30, la cual se utiliza para predecir el siguiente valor. Usamos jupyter notebook de python para copilar los modelos.

En cada iteración el modelo se entrena y actualiza los pesos de las neuronas utilizando el descenso del gradiente, mientras más iteraciones se realizan mayor aprendizaje adquiere la red. La idea es realizar iteraciones

hasta obtener un equilibrio entre tiempo de ejecución y precisión en el ajuste. En este estudio ese punto se alcanzó con 100 iteraciones como veremos a continuación.

#### 4.4.1. Red CNN para el Bitcoin BTC

Como se explicó previamente se considera la serie indicadora de subida y bajada del valor del precio al momento del cierre. El desempeño del modelo a través de las 100 iteraciones que realiza se pueden observar en las siguientes gráficas.

En la figura 4.23 podemos observar la gráfica de la función de costo del modelo, anteriormente mencionamos que es la categorical cross entropy (entropía cruzada).

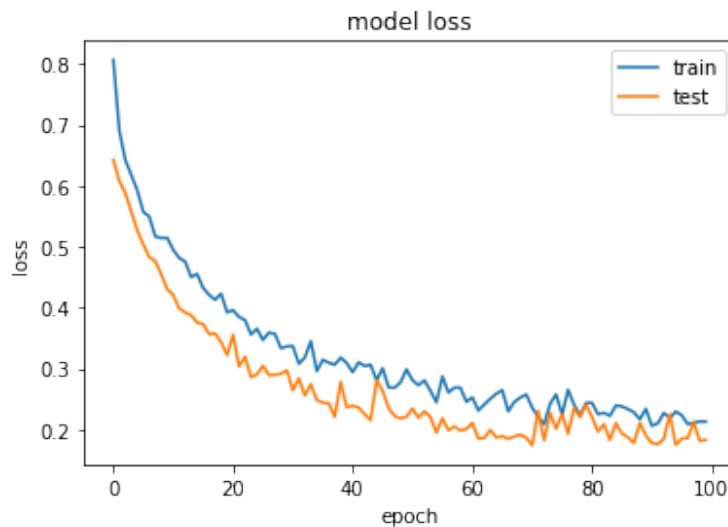


Figura 4.23: Función de costo

En este gráfico se percibe la evolución de la función de costo a través de las iteraciones, se nota que al avanzar el tiempo va decreciendo, sin embargo no lo hizo de manera constante tal y como se puede apreciar en la figura 4.23. Lo cual es ideal ya que este tipo de algoritmos busca minimizar los valores de dicha función, mientras más cercano a cero sea la función de costo más efectivo es el modelo. En la última iteración es decir la número 100 la función de costo tiene un valor de 0,17433.

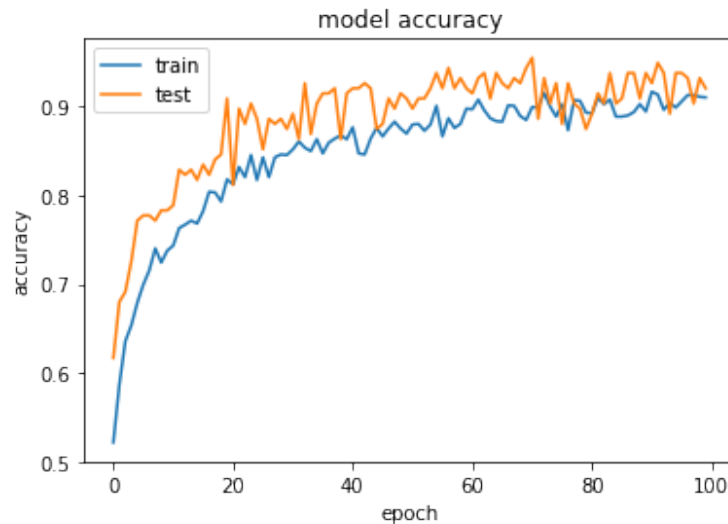


Figura 4.24: Accuracy

Para analizar la efectividad del modelo tanto sobre el conjunto de entrenamiento como sobre el conjunto de prueba se estudia el comportamiento de la trayectoria a través de las iteraciones de la función accuracy lo cual se puede apreciar en la figura 4.24. Notamos que la precisión del modelo aumenta con las iteraciones y que ya al cabo de 100 iteraciones se ha obtenido una precisión superior al 90 %.

La calidad del ajuste observado en la figura anterior se corrobora al estudiar la matriz de confusión arrojada por el modelo la cual se presenta a continuación.

$$\begin{pmatrix} 0,93421053 & 0,05050505 \\ 0,03947368 & 0,96969697 \end{pmatrix}$$

En ella se presentan los valores VP, VN, FP y FN de manera porcentual. Al calcular el promedio de la diagonal principal, lo cual es la metodología para determinar la efectividad del modelo, se obtiene que la precisión del mismo es de 95 %.

El modelo nos está diciendo que esta clasificando de manera acertiva si el precio del Bitcoin sube o baja con un 95 % de exactitud.

#### 4.4.2. Red CNN para el Ether ETH

Al igual que en el caso de la criptomoneda Bitcon se analiza la serie indicadora de subida y bajada de los precios de la criptomoneda Ether mediante un modelo de redes neuronales convolucionales el cual es capacitado con iteraciones. A continuación se muestra la gráfica de la función de costo del modelo para el Ether.

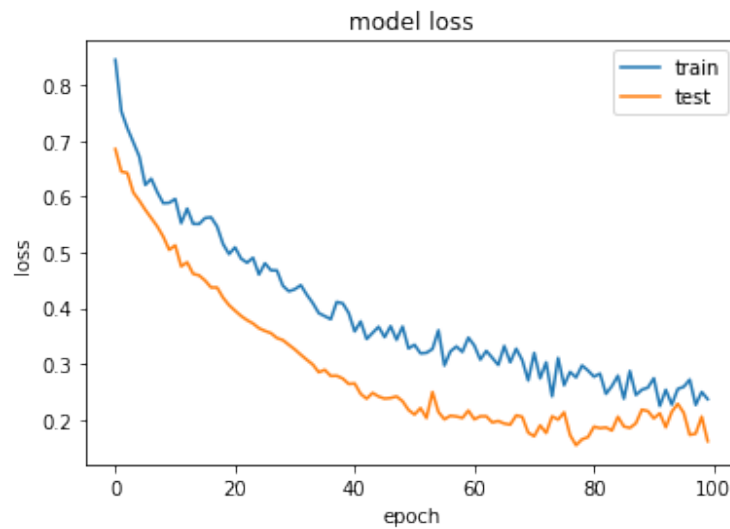


Figura 4.25: Función de costo

Notamos como se va generando la función de costo a lo largo de las iteraciones, podemos observar que a lo largo del tiempo va disminuyendo que es lo buscando, minimizar el valor de esta función de costo. En la última iteración la función de costo nos da un valor de 0,15391.

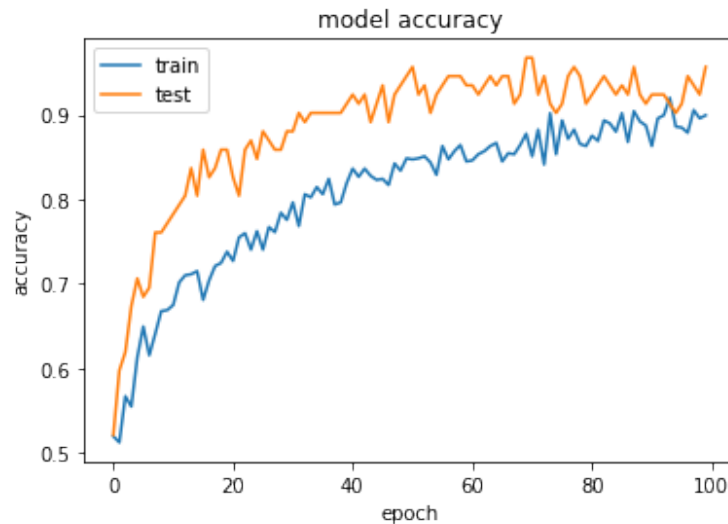


Figura 4.26: Accuracy

Para analizar la efectividad del modelo se analiza el comportamiento de la trayectoria a través de las iteraciones de la función accuracy. Observamos que la precisión del modelo aumenta con las iteraciones y que ya al cabo de 100 iteraciones se ha obtenido una precisión superior al 90%.

La calidad del ajuste mostrado en la figura 4.26 se confirma al analizar la matriz de confusión que nos da el modelo la cual se presenta a continuación.

$$\begin{pmatrix} 0,95744681 & 0,04444444 \\ 0,04255319 & 0,95555556 \end{pmatrix}$$

En ella se presentan los valores son de manera porcentual. Al calcular el promedio de la diagonal principal, lo cual es la metodología para determinar la efectividad del modelo, se obtiene que la precisión del mismo es de 95%.

El modelo nos está mostrando que esta clasificando de manera acertiva si el precio del Ether sube o baja con un 95% de exactitud.

### 4.4.3. Red CNN para el Litecoin LTC

En lo que sigue analizaremos la serie indicadora de subida y bajada del valor del precio al momento del cierre de la criptomoneda Litecoin y estudiaremos el desempeño de la red ajustada luego de 100 iteraciones.

En la figura 4.27 podemos observar la gráfica de la función de costo del modelo, que es la categorical cross entropy (entropía cruzada).

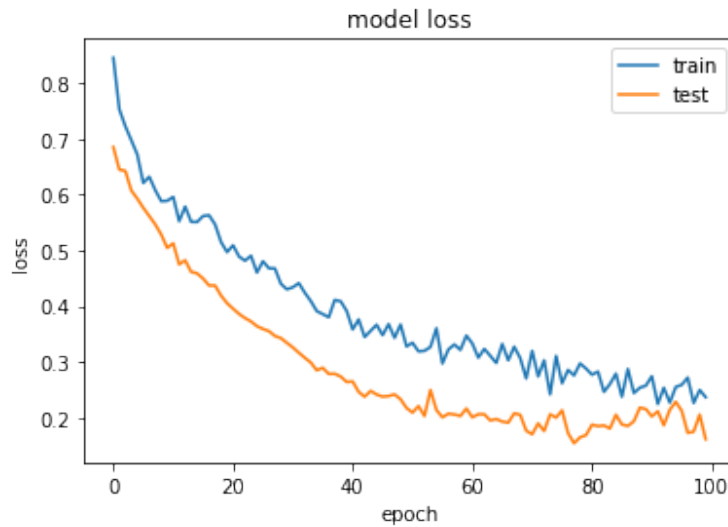


Figura 4.27: Función de costo

Vemos al igual que en las anteriores criptomonedas que con el paso de las iteraciones la función de costo tiende a cero. Lo cual es ideal ya que este tipo de algoritmos busca minimizar los valores de dicha función, mientras más cercano a cero sea la función de costo más efectivo es el modelo. En la última iteración es decir la número 100 la función de costo tiene un valor de 0,21847.

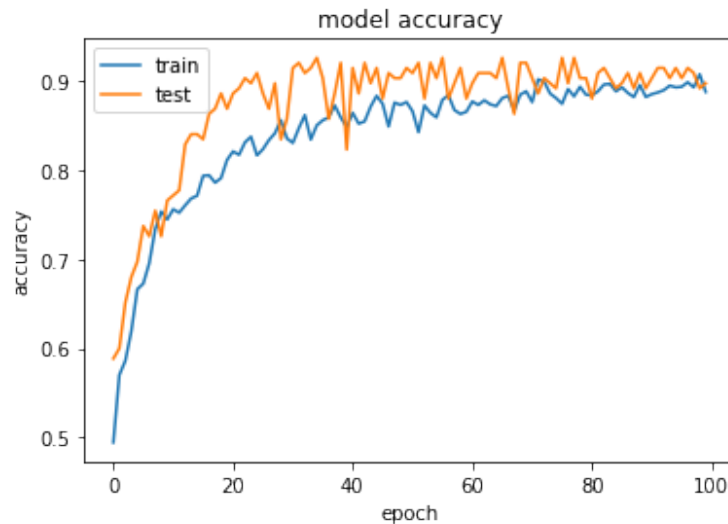


Figura 4.28: Accuracy

Notamos que la precisión del modelo aumenta con las iteraciones y que ya al cabo de 100 iteraciones se ha obtenido una precisión superior al 90 %.

La calidad del ajuste observado en la figura anterior se corrobora al estudiar la matriz de confusión arrojada por el modelo la cual podemos ver en lo que sigue.

$$\begin{pmatrix} 0,87179487 & 0,10309278 \\ 0,06410256 & 0,94845361 \end{pmatrix}$$

Nuevamente los valores de la matriz están en porcentajes. Al calcular el promedio de la diagonal principal, lo cual es la metodología para determinar la efectividad del modelo, se obtiene que la precisión del mismo es de 91 %.

El modelo nos está diciendo que está clasificando de manera acertiva si el precio del Litecoin sube o baja con un 91 % de exactitud.



#### 4.4.4. Red CNN para el Ripple XRP

Consideremos la serie indicadora de subida y bajada del valor del precio al momento del cierre de la moneda Ripple. A este conjunto de datos se le ajusta un modelo de red Neuronal convolucional y posteriormente se analiza el rendimiento del mismo luego de iteraciones. Como primer análisis se presenta el gráfico de la función costo

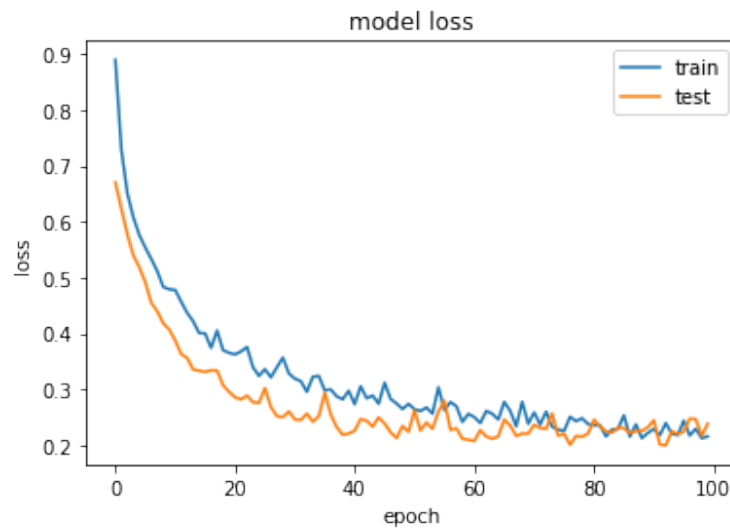


Figura 4.29: Función de costo

Notamos un comportamiento similar al de las otras funciones de costo que hemos analizado, esta tiende a disminuir. En la última iteración la función de costo tiene un valor de 0,19987.

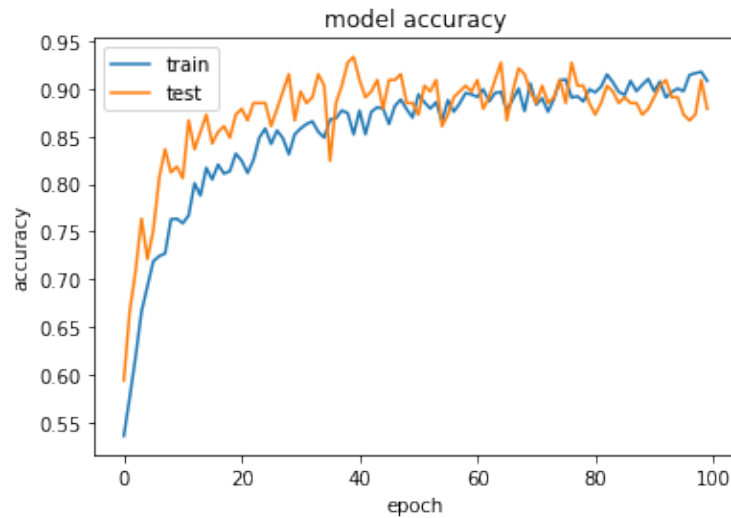


Figura 4.30: Accuracy

Observamos un comportamiento análogo a las otras criptomonedas. Vemos que la precisión del modelo aumenta con las iteraciones y que ya al cabo de 100 iteraciones se ha obtenido una precisión superior al 85%.

La calidad del ajuste observado en la figura 4.30 se reafirma al estudiar la matriz de confusión arrojada por el modelo la cual se presenta a continuación.

$$\begin{pmatrix} 0,8666667 & 0,1111111 \\ 0,0666667 & 0,9444444 \end{pmatrix}$$

Las entradas de la matriz están dadas de manera porcentual. Al calcular el promedio de la diagonal principal, se obtiene que la precisión del mismo es de 90%.

El modelo nos está diciendo que está clasificando de manera acertiva si el precio del Ripple sube o baja con un 90% de exactitud.

#### 4.4.5. Red CNN Bitcoin Cash BCH

Finalmente analizaremos la serie indicadora de bajada y subida del valor del precio al momento del cierre de la criptomoneda Bitcoin Cash.

En la figura 4.31 podemos observar la gráfica de la función de costo del modelo, al igual que en todas las criptomonedas analizadas se usó la categorical cross entropy (entropía cruzada).

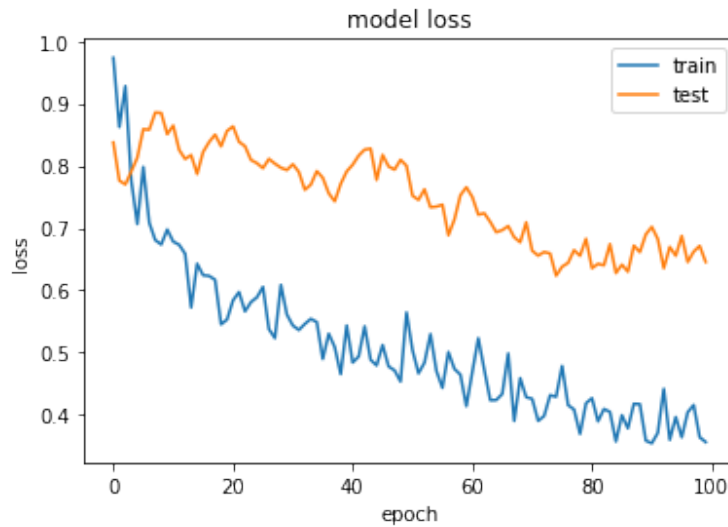


Figura 4.31: Función de costo

Observamos como se va generando la función de costo a lo largo de las iteraciones, al igual que las anteriores criptomonedas notamos que tiende a disminuir pero en este caso no es tan eficiente. En la última iteración la función de costo tiene un valor de 0,62349.

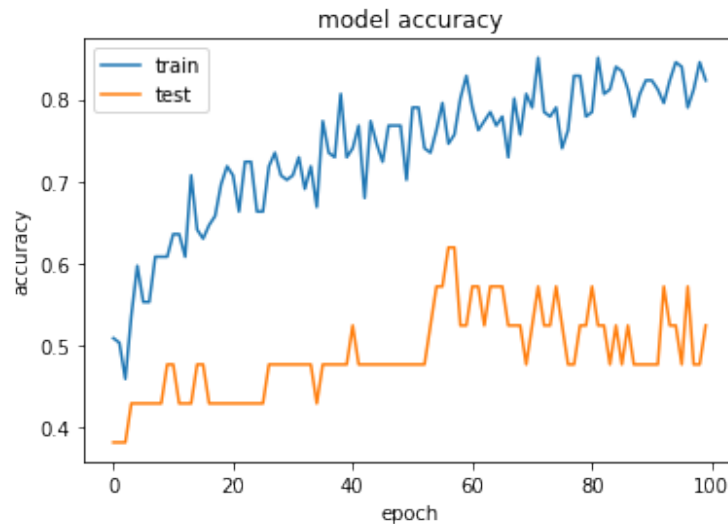


Figura 4.32: Accuracy

Análogamente se analizó la efectividad del modelo, estudiando el comportamiento de la trayectoria a través de las iteraciones de la función accuracy lo cual se puede apreciar en la figura 4.32. Notamos que la precisión del modelo aumenta, pero de una manera no tan efectiva, con las iteraciones en este caso al cabo de 100 iteraciones solo se ha obtenido una precisión superior al 55 %

La calidad del ajuste observado en la figura anterior se confirma al estudiar la matriz de confusión arrojada por el modelo la cual se presenta a continuación.

$$\begin{pmatrix} 0,75 & 0,15384615 \\ 0,875 & 0,46153846 \end{pmatrix}$$

Como en todas las anteriores criptomonedas los valores de la matriz están dados de manera porcentual. Al calcular el promedio de la diagonal principal, lo cual es la metodología para determinar la efectividad del modelo, se obtiene que la precisión del mismo es de 60 %.

Estos resultados del BCH nos dan diferente a los de las otras criptomonedas, notamos una gran disminución en la eficiencia del modelo, esto puede estar asociado a que este conjunto de datos es el que tiene menos observaciones con 233, debido a que es la criptomoneda más nueva de nuestro estudio. Estos modelos de redes neuronales son más precisos a medida que son entrenados con un número mayor de observaciones.

## 4.5. Conclusiones

Se logró construir un modelo de análisis de criptomonedas, para predecir y clasificar el valor del precio mediante modelos SARIMA y Redes Neuronales CNN.

Por la parte de los Modelos SARIMA de series de tiempo, hicimos el estudio de tres criptomonedas, las cuales fueron el Bitcoin, Bitcoin Cash y Ether con pronósticos favorables de 82 %, 91 % y 58 % respectivamente, como ya hemos visto en el trabajo realizado por Miguel Porro "Modelos de análisis de criptomonedas basado en aprendizaje automático y series temporales", los modelos ARIMA arrojan una mejor predicción para el Bitcoin, Ether con un 94 % para ambas y vemos que para el Bitcoin Cash son muy similares los modelos ARIMA y SARIMA con un 91 % para ambos modelos. Una de las posibles razones por la cual los dos modelos ARIMA y SARIMA dieron diferentes es el intervalo de tiempo de la predicción. En el caso del Bitcoin, la moneda más importante del mercado en cuanto a la capitalización, es también la criptomoneda más sensible al comportamiento humano, anuncios, noticias o divulgaciones afectan positiva o negativamente la confianza de los inversionistas, comerciantes o tenedores de criptomonedas, haciéndola así una de las más sensibles.

Luego tenemos el análisis de clasificación de las redes neuronales CNN para las criptomonedas Bitcoin, Ether, Litecoin, Ripple y Bitcoin Cash mostrando un accuracy o efectividad del 95 %, 95 %, 91 %, 90 %, 60 % respectivamente, pudimos observar que para 4 de las 5 criptomonedas los modelos son muy buenos en sus resultados y para el Bitcoin Cash es de menor eficiencia se debe a que es la criptomoneda que tiene menos observaciones, debido a que es la más reciente con datos a partir del 27/07/17, estos modelos de redes neuronales cuanto mayor son las observaciones más precisos son.

Por último pudimos observar que es posible predecir el comportamiento de los precios de las diferentes criptomonedas, implementando los estudios realizados en este trabajo.

### 4.5.1. Trabajos futuros

Partiendo del hecho de que cada día se crean nuevas criptomonedas basadas en cadenas de bloques y la adopción de esta tecnología en muchos sectores, no sólo el financiero se propone lo siguiente como trabajos futuros para continuar con el desarrollo de este trabajo :

- \* Adopción de una herramienta como Apache Hadoop para la implementación de estos análisis en gran escala, es decir, tomando mas datos.
- \* Trabajar con otros modelos de redes neuronales.
- \* Implementar estos modelos como bases para el manejo de portafolio de inversión

# Bibliografía

- [1] Brockwell, P.J., & Davis, R.A., (1996). *Introduction to Time Series and Forecasting* Springer-Verlag, New York Inc, New York.
- [2] Charu C. Aggarwal (1998). *Data Mining: The Textbook*. Springer, New York.
- [3] Colonescu C. (2016) *Principles of Econometrics with R*.
- [4] Crye Jonathan D. & Chan Kun-Sik , (2008). *Time Series Analyses With Applications in R*, 2nd edition. Springer.
- [5] Gupta, M. (2017). *Blockchain for dummies*. John Wiley and Sons, Inc.
- [6] Hassoun Mohamad, (2003). *Fundamentals of Artificial Neural Networks*
- [7] Haykin Simon, (2008) *Neural Networks and Learning Machines* . 3rd edition.
- [8] Hernández José B. y Colina Mairene. *Características de las Series de Tiempo*. Guía de clases.
- [9] Hernández José B. y Colina Mairene. *Modelos ARMA*. Guía de clases.
- [10] Nakamoto Satoshi (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*.
- [11] Olson David L. & Durson Delen (2008) *Advanced Data Mining Techniques*. Springer-Verlag. Berlin Heidelberg.
- [12] Ordinas, M. (2017) *Las criptomonedas: Oportunidad o Burbújas?* . Informe mensual de estrategias de Banca-March.

- 
- [13] Porro, Miguel. (2018) *Modelos de análisis de criptomonedas basado en aprendizaje automático y series temporales*. Trabajo Especial de Grado de Licenciatura. Universidad Central de Venezuela.
- [14] Ruey S. Tsay (2002). *Analysis of Financial Time Series*, John Wiley & sons, Inc.
- [15] Sanchez Gil A. & Terán Varela, O. (2018) *Criptomonedas como oportunidad de negocio de empresas del Sector Turístico en la Zona Sur Oriente Del estado de México*. Revista global de negocios v.6 (1), pp 93-104.
- [16] Shumway, R.H & Stoffer, D.S. (2006). *Time Series Analysis and Its Applications with R examples*, 2nd edition. Springer.