

Instituto Tecnológico y de Estudios Superiores de Occidente

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial 15018, publicado en el Diario Oficial de la Federación del 29 de noviembre de 1976.

Departamento de Matemáticas y Física
Maestría en Ciencia de Datos



Análisis de oportunidades de arbitraje entre criptomonedas aplicando Reinforcement Learning.

TESIS que para obtener el **GRADO** de
MAESTRO EN CIENCIA DE DATOS

Presenta: **OSCAR NAVA ABARCA**

Director: **LUIS RAÚL RODRÍGUEZ REYES**

Tlaquepaque, Jalisco. 15 de noviembre de 2022.

RESUMEN

Se presenta una herramienta de análisis de una técnica financiera que ha sido estudiada en el diseño de estrategias de inversión, el arbitraje en el intercambio de divisas, sin embargo, el enfoque de este proyecto es llevar a cabo un ejercicio similar en el mercado de las criptomonedas utilizando un modelo de machine learning que permita detectar oportunidades para operar el arbitraje al realizar el intercambio entre una cantidad variable de criptomonedas.

Se planteó el problema como uno de tipo optimización combinatoria y se utilizó un algoritmo de reinforcement learning debido a que diferentes estudios han demostrado que esta técnica es capaz de arrojar buenos resultados para este tipo de problemas.

Los resultados obtenidos muestran que es posible encontrar rutas de intercambios que representen una oportunidad de arbitraje en un tiempo relativamente corto y recursos computacionales no tan sofisticados como los empleados en la actualidad por las grandes instituciones financieras que dedican una gran cantidad de recursos en este ámbito.

TABLA DE CONTENIDO

MAESTRÍA EN CIENCIA DE DATOS.....	1
RESUMEN	3
TABLA DE CONTENIDO	4
LISTA DE FIGURAS	6
LISTA DE TABLAS.....	7
LISTA DE ACRÓNIMOS Y ABREVIATURAS	8
1.INTRODUCCIÓN	9
1.1. ANTECEDENTES	10
1.2. JUSTIFICACIÓN	10
1.3. PROBLEMA.....	12
1.4. HIPÓTESIS	12
1.5. OBJETIVOS	12
1.5.1. Objetivo General:	12
1.5.2. Objetivos Específicos:.....	12
1.6. NOVEDAD CIENTÍFICA, TECNOLÓGICA O APORTACIÓN	13
2. ESTADO DEL ARTE O DE LA TÉCNICA	14
2.1. CARACTERIZACIÓN DE ARBITRAJE FINANCIERO EN EL MERCADO DE DIVISAS	15
2.2. TRADING AND ARBITRAGE IN CRYPTOCURRENCY MARKETS	15
2.3. EFFICIENCY IN THE MARKETS OF CRYPTO-CURRENCIES	16
2.4. APRENDIZAJE POR REFUERZO ELEMENTOS BÁSICOS Y ALGORITMOS	16
2.5. REINFORCEMENT LEARNING FOR THE TRAVELING SALESMAN PROBLEM WITH REFUELING	17
3. MARCO TEÓRICO/CONCEPTUAL	18
3.1. BLOCKCHAIN	19
3.2. CRIPTOMONEDAS	19
3.3. OPTIMIZACIÓN COMBINATORIA	20
3.3.1. TRAVELING SALESMAN PROBLEM.....	20
3.4. MACHINE LEARNING	20
3.4.1. REINFORCEMENT LEARNING.....	20
3.4.1.1. Q-LEARNING.....	21
3.4.1.2. E-GREEDY.....	21
4. DESARROLLO METODOLÓGICO	22
4.1. OBTENCIÓN DE DATOS	23
4.2. FORMULACIÓN MATEMÁTICA DEL PROBLEMA	25
4.3. PREPARACIÓN DE LOS DATOS	25
4.4. SELECCIÓN DEL MODELO	26
4.5. EJECUCIÓN DEL MODELO	27
4.5.1. PARÁMETROS	27
4.5.2. ENTRENAMIENTO	28
4.5.3. OBTENCIÓN DE RUTA.....	28

5.	RESULTADOS Y DISCUSIÓN	30
5.1.	RESULTADOS	31
5.2.	DISCUSIÓN	32
6.	CONCLUSIONES	33
6.1.	CONCLUSIONES	34
6.2.	TRABAJO FUTURO	34
BIBLIOGRAFÍA		36

LISTA DE FIGURAS

Figura 1. Rutas posibles de intercambio en Bitso. Fuente: Elaboración propia.	24
Figura 2. Estructura del objeto Ticker de Bitso. Fuente: Elaboración propia.	25

LISTA DE TABLAS

Tabla 1. Precios criptomonedas seleccionadas. Fuente: Elaboración propia.	26
Tabla 2. Ejemplo de matriz Q actualizada. Fuente: Elaboración propia.	29
Tabla 3. Resultados de rutas encontradas con utilidad por iteración. Fuente: elaboración propia.	31

LISTA DE ACRÓNIMOS Y ABREVIATURAS

CM	Criptomoneda
ML	Machine Learning
RL	Reinforcement Learning
QL	Q-learning

1. INTRODUCCIÓN

1.1. Antecedentes

El arbitraje es una herramienta de inversión que ha sido estudiada ampliamente en diversos mercados financieros, desde el accionario, pasando por índices bursátiles hasta materias primas. En el caso de mercados de criptomonedas los estudios relacionados con arbitraje se han enfocado en la eficiencia como el caso del trabajo de Tran y Leirvik [1] en donde se analizaron los cinco mercados de criptomonedas más grandes y se demostró que los niveles de eficiencia varían con el tiempo siendo un mercado mayormente ineficiente y sensible a eventos puntuales que suceden alrededor de estos activos. Sin embargo, el objeto de estudio de este trabajo es el arbitraje en el intercambio de criptomonedas en un mismo mercado local, por tal motivo se analizó el mercado de divisas en el cual es común encontrar estrategias de arbitraje triangular muy parecidas a lo que se pretende estudiar. En ese sentido Luna [2] plantea la utilización de modelos de aprendizaje automático para identificar oportunidades de compra y venta y el tiempo que están disponibles para operar el arbitraje en el par NZDUSD que es poco común.

Por otra parte, Markarov y Schoar [3] realizaron un análisis entre diferentes plataformas de intercambios de criptomonedas y su influencia en la eficiencia, se indica que las oportunidades de arbitraje se dan en mayor medida entre países pues no existen medidas de regulación de los mercados, haciendo que los inversionistas sean los que establezcan los precios con base en la oferta y demanda.

El enfoque de este trabajo es establecer un mecanismo que sea capaz de detectar oportunidades de arbitraje al realizar intercambios entre diferentes criptomonedas en una misma plataforma digital de intercambio en el menor tiempo posible por lo que se analizó utilizar modelos de aprendizaje reforzado para maximizar la utilidad en diferentes combinaciones de intercambios. Partiendo de esta idea se estudió el trabajo de Lascorz [4] que hace una revisión de las bases del aprendizaje por refuerzo y los diferentes algoritmos que se utilizan. Siguiendo en esa línea, el trabajo de Ottoni, et al. [5] propone la utilización de este método de aprendizaje automático para resolver un problema de optimización combinatoria, aun cuando el problema tratado es relacionado a distancias entre ciudades, se pudo replantear de tal manera que se adaptó al problema de arbitraje de este trabajo.

1.2. Justificación

El arbitraje es una técnica financiera en la cual se busca aprovechar la diferencia del precio de un activo en diferentes mercados y en la que teóricamente no existe riesgo [6], sin embargo, en la práctica pueden surgir algunos como el riesgo base o el riesgo de operación en el cual el mercado se puede mover en el momento en que se está ejecutando. Por ejemplo, debido al uso de sistemas informáticos por parte grandes instituciones, que operan a gran velocidad, las brechas entre los mercados se cierran rápidamente por lo tanto el riesgo asociado con este tipo de operaciones es el de ejecución pues existe la posibilidad de perder la oportunidad debido al cambio rápido en los mercados.

En términos de teoría financiera, la idea detrás del arbitraje es aprovechar las ineficiencias de información que se llegan a presentar en el mercado. De acuerdo con la Hipótesis de Mercados Eficientes de Fama [7], la eficiencia del mercado se refiere al grado en que el precio de un activo refleja el verdadero valor de éste, basado en toda la información relevante disponible para todos los actores, dicho de otro modo, cuando un mercado es eficiente en información existente acerca de un activo se ve incorporada en el precio. Por esta razón, en un mercado completamente eficiente es teóricamente imposible que un inversionista pueda obtener retornos extraordinarios del mercado, ya que cualquier información que posea será del conocimiento de todos y estará reflejada en el precio. Existen tres grados de eficiencia de mercado: débil, semi-fuerte y fuerte, que miden el nivel en que impacta nueva información en el precio. Por ejemplo, en un mercado eficiente en el sentido débil, no se puede predecir el precio usando el historial, por lo tanto, los precios futuros serán debido a nueva información que vaya surgiendo; en un semi-fuerte los precios se ajustarán rápidamente conforme se genere nueva información y únicamente se podría aprovechar información privada que no esté disponible para todos los actores del mercado. Finalmente, en un mercado en el cual la eficiencia de información es fuerte, aún la disponibilidad de información privada no representa una ventaja respecto al resto de los inversionistas.

Existe evidencia que en el mercado de las criptomonedas la eficiencia de información pudiera ser intermitente, permitiendo que el uso del arbitraje financiero pudiera generar ganancias extraordinarias. Por ejemplo, Tran y Leirvik [1] proponen que la eficiencia de este mercado varía con el tiempo, mostrando que hay periodos con ineficiencia y otros donde se vuelven más eficientes, sin embargo, encontraron que el mercado de criptomonedas es en gran parte ineficiente. Analizaron cómo el creciente interés y volumen de transacciones de algunas criptomonedas afecta la eficiencia y encontraron que estos mercados son sensibles a varios factores.

Las criptomonedas por su naturaleza cuentan con ciertas características que las hacen diferentes a las monedas tradicionales como el hecho de que no existen de forma física, no están reguladas ni respaldadas por ninguna institución financiera, están basadas en Blockchain por lo que además se consideran descentralizadas y su valor se determina con base en el mercado, es decir, por la oferta y la demanda. La principal y primer criptomoneda es Bitcoin y según su creador, Nakamoto [8], tiene la intención de crear un sistema de pagos electrónicos entre iguales y sin intermediarios como instituciones financieras basado en pruebas criptográficas. En octubre de 2021 alcanzó el mayor nivel de capitalización con un valor estimado de 1.15 billones de dólares, después de ese periodo ha ido decreciendo y en enero de 2022 su valor era de 816 mil millones de dólares.

Actualmente existen una gran cantidad de proyectos basados en Blockchain que han dado origen a otras criptomonedas llegando a ser incluso miles de éstas. La importancia de las criptomonedas en los mercados financieros es tal que se han convertido en parte activa de portafolios de inversión respaldados por importantes entidades financieras.

Por todo esto y por el poco tiempo de existencia de estos activos (en comparación con los valores tradicionales) es que los mercados de criptomonedas son diferentes y generan un campo de estudio.

1.3. Problema

La existencia de oportunidades de arbitraje dentro del mercado de criptomonedas indica que el mercado no está siendo eficiente y llevarlo a cabo mejora los niveles de eficiencia [9]. Para realizar esta actividad es necesario hacer operaciones de compra y venta de un activo de forma simultánea [6] en diferentes mercados para aprovechar la discrepancia de precios, como ocurre, por ejemplo, en el mercado de divisas. En este caso, se pretende hacer un ejercicio que permita detectar ineficiencias en el mercado de criptomonedas similar al arbitraje triangular de divisas con la diferencia de que se aplique para n puntos, se busca la mejor combinación o ruta de intercambio entre aquellas criptodivisas que al ejecutarlo genere una utilidad monetaria, es decir un flujo monetario positivo. Para encontrar dicha ruta es necesario hacerlo en el menor tiempo posible pues las ventanas de oportunidad que permiten ejecutar el arbitraje son muy cortas.

1.4. Hipótesis

Detectar oportunidades de arbitraje entre n criptodivisas es posible si se plantea como un problema de optimización combinatoria, haciendo uso de algoritmos y modelos de aprendizaje automático que permitan encontrar la combinación o ruta de intercambios que maximice la utilidad expresada como flujo monetario.

1.5. Objetivos

1.5.1. Objetivo General:

Encontrar, entre los posibles intercambios de criptomonedas, la mejor ruta de varios puntos que permita realizar operaciones de arbitraje en un periodo de tiempo, planteando el problema de arbitraje como un problema de optimización combinatoria, utilizando un algoritmo de Aprendizaje Reforzado (Reinforcement Learning) que optimice la utilidad.

1.5.2. Objetivos Específicos:

1. Definir ex-ante los posibles intercambios entre las criptomonedas que se van a utilizar.
2. Obtener datos de precios y del mercado cada cierto tiempo de las criptomonedas seleccionadas.
3. Entrenar un modelo de aprendizaje por refuerzo que permita identificar la ruta que genera más oportunidades de arbitraje.
4. Definir cuál es la mejor ruta para realizar arbitraje.

1.6. Novedad científica, tecnológica o aportación

En este proyecto se plantea la posibilidad de llevar a cabo un ejercicio financiero, como el arbitraje, a través del uso de herramientas de ciencia de datos, como un algoritmo de aprendizaje reforzado y técnicas de ingeniería de datos. Si bien lo que se plantea en este proyecto no es, per se, una herramienta para generar utilidades monetarias, sino una herramienta de optimización, los resultados pueden ser implementados para crear una estrategia de inversión por arbitraje con base en datos del mercado de criptomonedas. En otras palabras, la aportación de este proyecto es una aplicación de ciencia de datos en un sector financiero.

Para lograr el objetivo del proyecto se utilizó un modelo de aprendizaje por refuerzo utilizando el algoritmo Q-learning, con base en lo planteado por Ottoni et al. (2020) en el cual proponen una solución al problema del agente viajero haciendo uso de este algoritmo y el cual se adaptó para plantear el problema del intercambio entre criptomonedas.

2. ESTADO DEL ARTE o de la TÉCNICA

2.1. *Caracterización de arbitraje financiero en el mercado de divisas*

Luna [2] aborda el arbitraje del mercado de divisas o FOREX como estrategia de inversión prediciendo patrones. Se estudia el arbitraje triangular, pero sobre todo propone encontrar las características del comportamiento para encontrar oportunidades que permitan operarlo. Se plantea que el mercado de divisas es más grande que el de acciones, opciones y futuros y los cambios estructurales que se han dado con el tiempo, debido principalmente al desarrollo de la tecnología y de las instituciones financieras, han impactado en aspectos del mercado como el tamaño, la competencia e incluso en la eficiencia. El cambio principal es la evolución de un mercado que ha permitido la libre oferta y demanda haciendo que se ofrezcan más productos y esto a su vez que aumente el volumen, cuando esto sucede se presentan oportunidades potenciales de operar el arbitraje.

El planteamiento es que el arbitraje en el mercado de divisas consiste en realizar compra y venta simultáneas de pares de divisas, y las oportunidades de operarlo surgen cuando se pierde el equilibrio en los tipos de cambio haciendo que los precios sean inconsistentes en el cambio cruzado. El tipo de arbitraje más estudiado es triangular del par de divisa EURUSD pues se sugiere que tiene menos errores sin embargo el autor escogió un par menos común, el NZDUSD buscando primero las propiedades más importantes del arbitraje triangular para determinar la existencia de oportunidades de operar ya que normalmente éstas se presentan en ciertos periodos del día y tienen una duración muy breve.

Se identificó que existían pocas oportunidades de compra y venta sin embargo la duración de éstas era de hasta 5 minutos, otros artículos afirman que solamente duran segundos. Además, se identificó que al usar un par menos convencional las oportunidades son mayores, esto respecto al par EURUSD (el más estudiado). En cuanto a las características, los indicadores de volatilidad y tendencia fueron predominan en los diferentes experimentos y serían factores para tomar en cuenta al momento de identificar la existencia de posible arbitraje y para determinar si se debe realizar una compra o venta.

2.2. *Trading and Arbitrage in Cryptocurrency Markets*

Artículo escrito por Makarov y Schoar [3] enfocado en realizar un análisis sistemático de los intercambios y la eficiencia de los mercados de criptomonedas en donde se estudia el tema del arbitraje entre diferentes plataformas digitales de intercambio también conocidas como *exchanges*. Los autores indican que los mercados muestran grandes periodos de oportunidades de arbitraje entre países, menores dentro del mismo país y aún más pequeños entre criptodivisas.

Alrededor del mundo existe una gran cantidad de plataformas de intercambio en la que la mayoría funcionan como un mercado tradicional donde los inversionistas colocan órdenes de compra o venta y la plataforma las lleva a cabo basadas en un libro centralizado de órdenes, sin embargo, no existen mecanismos para asegurar que serán ejecutadas al mejor

precio, esta es una de las razones principales por las que potencialmente se pueden segmentar los mercados dando pie al arbitraje entre mercados, a diferencia de las monedas respaldadas por un banco central en donde las transferencias de capital entre países están reguladas para inversionistas, aunque suelen ser evitadas por grandes instituciones, aunque por otra parte esta falta de regulación frena la adopción de criptomonedas por parte de las instituciones públicas haciendo que el capital por arbitraje disminuya.

Midiendo la diferencia en el comportamiento de precios entre mercados se demostró que existe un movimiento significativo similar en el margen de arbitraje entre países. El margen en el resto del mundo está correlacionado positivamente con Estados Unidos y Europa y aumenta cuando el bitcoin tiene periodos de apreciación.

2.3. *Efficiency in the Markets of Crypto-Currencies*

Tran y Leirvik [1] muestran a través de este estudio que la eficiencia de los cinco mercados de criptomonedas más grandes es variable con el tiempo, especialmente antes de 2017, debido al gran crecimiento de estos mercados que han hecho que crecieran exponencialmente los intercambios, liquidez y precios. Después de ese año los mercados mostraron ser más eficientes en las herramientas de medición utilizadas para determinar si la eficiencia era significativa, pero habían sido ineficientes en el pasado. Encontraron que a pesar de que la eficiencia creció significativamente durante 2017, mientras incrementa el volumen y se vuelven menos volátiles, los mercados son sensibles a eventos puntuales, un ejemplo señalado es cuando un hackeo en 2016 derivó en la separación de Ethereum en Ethereum y Ethereum Classic, este evento hizo que creciera la incertidumbre y disminuyera la eficiencia.

Este proyecto hace una revisión de diferentes trabajos relacionados con la eficiencia, desde la Hipótesis de Mercados Eficiente de Fama (1991), el trabajo de Lo y MacKinlay (1989) donde proponen un método para probar si un mercado es eficiente o la Hipótesis del Mercado Adaptable de Lo (2004) entre otros que se derivan de este último. Por lo tanto, este estudio aborda ideas como la segmentación de la eficiencia, que los mercados financieros son en gran medida fuertemente eficientes y que la información disponible se refleja en el precio, que la eficiencia varía con el tiempo.

2.4. *Aprendizaje por Refuerzo Elementos básicos y algoritmos*

Reinforcement Learning o Aprendizaje por Refuerzo es un área de estudio de las matemáticas y la computación en la cual se combinan sistemas y algoritmos con el objetivo de generar una herramienta que permita resolver un problema del cual no se tiene suficiente conocimiento. Lascorz [4] define los fundamentos del campo del aprendizaje automático, en específico del aprendizaje por refuerzo, así como el algoritmo Q-learning el cual es utilizado en modelos de este tipo, además muestra ejemplos de aplicación.

La idea del aprendizaje automático está fundamentada en replicar la psicología humana y su capacidad para aprender, se intenta imitar el proceso didáctico de prueba y error, por ejemplo, en el Aprendizaje por Refuerzo el problema se estructura como un proceso de decisión de Markov y en este un agente interactúa con un ambiente conociendo únicamente el objetivo final y recibe una señal de recompensa o castigo dependiendo de si se tiene éxito o se falla alcanzando el objetivo. También existen otras alternativas dependiendo de la información que se tenga del problema como el Aprendizaje Supervisado o Aprendizaje No Supervisado, y el verdadero valor se encuentra en la capacidad de aprender de los datos teniendo como fundamento algoritmos de aprendizaje para encontrar patrones y tomar decisiones dependiendo de la situación.

2.5. Reinforcement Learning for the traveling salesman problem with refueling

Otoni et al. [5] en su trabajo proponen principalmente una solución al problema del agente viajero con recarga de combustible utilizando RL, también se enfoca en analizar los parámetros del modelo y la influencia de la recarga de combustible en la optimización de costos. El problema del agente viajero (TSP) es uno de los más estudiados en el tema de optimización combinatoria, aquí un agente debe optimizar la ruta entre un grupo de ciudades o nodos de manera que visite cada uno y regrese al punto de partida. La realización de este trabajo parte de la investigación en el área de vehículos autónomos siguiendo la línea de atacar el problema de recarga de combustible para asegurar que el vehículo se quede sin él.

El objetivo principal es definir una ruta que minimice el consumo de combustible y se revisaron diferentes métodos heurísticos y encontraron que RL había mostrado ser bueno en problemas de optimización combinatoria. Uno de los principales temas de investigación relacionados con esta herramienta es la estimación de parámetros que, como se analiza en este trabajo, puede influir directamente en un buen aprendizaje de la ruta que mejor optimiza, por esta razón los autores proponen una metodología con enfoque en encontrar estos parámetros mediante experimentos que involucran algoritmos de aprendizaje por refuerzo y métodos estadísticos para la estimación de parámetros.

3. MARCO TEÓRICO/CONCEPTUAL

3.1. Blockchain

A raíz del surgimiento del bitcoin en el 2008 se presentó un concepto revolucionario del cuál se ha escrito y estudiado por diferentes instituciones alrededor del mundo, siendo blockchain una tecnología con un potencial aún por explotar.

Esta tecnología se basa en el paradigma de sistemas distribuidos que en términos computacionales se trata de dos o más nodos realizando tareas paralelas de manera coordinada para alcanzar un fin común, mientras tanto el usuario lo percibe como una sola plataforma. Blockchain es en esencia una base de datos distribuida entre usuarios que está protegida por métodos criptográficos, en la cual solo se pueden agregar datos, es inmutable, y actualizar por medio del consenso de los pares [9].

En 2008 Satoshi Nakamoto publicó el paper Bitcoin: A Peer-to-Peer Electronic Cash System en donde presentó por primera vez el término chain of blocks, término que después evolucionó a blockchain. Entre los beneficios de este sistema distribuido es la descentralización, transparencia y eficiencia.

3.2. Criptomonedas

Bitcoin es la primera aplicación de la tecnología de blockchain la cual comenzó una revolución al presentar un nuevo sistema descentralizado de moneda digital que ha demostrado ser segura y estable. Sin embargo, no es el primer proyecto que trata de abordar la creación de un sistema seguro de dinero digital.

En la década de los 80 aparecieron propuestas como la de David Chaum en la cual se establecía un esquema para usar firmas para crear una divisa digital que no fuera trazable, sin embargo, estaba limitada por la necesidad de que un banco llevara un registro de todas las firmas, se trataba de un sistema centralizado. Después se actualizó a una versión llamada e-cash en donde además era necesario una forma de identificación privada que era enviada como mensaje al banco central.

En 1997 Adam Back presentó hashcash como una forma de combatir el spam por email, el proyecto consistía en resolver un problema computacional que fuera fácil de verificar, pero difícil de calcular.

Wei Dai propuso utilizar Proof of Work como una forma de crear dinero digital, la principal desventaja de este sistema era que los usuarios con mayor poder computacional podían generar dinero sin permitir a la red ajustar el nivel de dificultad.

En 2008 Satoshi Nakamoto presentó un sistema de efectivo digital entre pares, el cual necesitaba una entidad financiera central como intermediario para realizar pagos y transferencias entre los usuarios. Representa la unión de ideas y conceptos formados por

años de investigación en criptografía, pues los proyectos mencionados funcionan como los fundamentos en los que se basa esta nueva tecnología [9].

3.3. Optimización combinatoria

Optimización combinatoria es un campo de las matemáticas en donde se emplean diferentes técnicas y algoritmos con el objetivo de resolver un problema de optimización de estructuras discretas [9], se enfoca en encontrar el mejor resultado entre un conjunto de combinaciones.

Las estructuras discretas pueden ser un grupo de conjuntos normalmente asociados a un valor que puede representar costos, por ejemplo. El problema se mide en términos de la longitud de la codificación de la estructura más el tamaño del conjunto de números, este dato es importante ya que de este depende el rendimiento de un algoritmo, sobre todo en instancias grandes.

3.3.1. Traveling Salesman Problem

Este problema es uno de los más estudiados en el campo de la optimización combinatoria y recibe este nombre (problema del agente de ventas viajero) porque tiene como objetivo encontrar el circuito más simple que un agente de ventas tendría que recorrer si visitara todos los nodos, de una lista de ciudades, de tal manera que la suma de la distancia recorrida sea mínima. Una de las características de este problema es que la mayoría de los métodos para resolverlo no funcionan bien hablando de la relación entre la calidad del resultado y tiempo de ejecución del método [10].

3.4. Machine Learning

Machine learning hace referencia a los métodos computacionales que permiten a un sistema aprender con base en datos para mejorar su rendimiento, así como los seres humanos aprenden a través de la experiencia. La tarea principal de ML es desarrollar algoritmos para generar modelos que se alimentan de datos, de esta manera se pueden hacer predicciones de nuevas observaciones [11].

Al proceso de buscar patrones en los datos se le llama aprendizaje o entrenamiento, para lo cual se usa un conjunto de datos que se denominan de entrenamiento, con el objetivo de ajustar un modelo sujeto a las reglas implícitas en la información con la que se entrena. Cuando ya se ha aprendido un modelo, el siguiente paso es hacer predicciones, para lo cual se utilizan datos que permitan probarlo, llamados datos de prueba.

La función principal de ML es aprender modelos que puedan arrojar buenos resultados de nuevas muestras, más allá de los datos de entrenamiento [11].

3.4.1. Reinforcement learning

Reinforcement learning es un área de machine learning en el cual se utilizan métodos de entrenamiento en forma de una secuencia de decisiones basadas en recompensas y castigos dependiendo del comportamiento deseado. Esta técnica está basada en el modelo de los

procesos de decisión de Markov y se compone de un conjunto finito de acciones, estados y recompensas [5]. El objetivo es hacer que un agente aprenda mediante la interacción con un ambiente que representa el mundo en el cual se mueve y ejecuta acciones por las cuales recibe una recompensa, al final se obtiene una política, o *policy* (π) que maximiza el valor de la recompensa.

3.4.1.1. Q-learning

Es un algoritmo que basa el aprendizaje en una diferencia temporal, es decir, las actualizaciones se dan a través un número de pasos sucesivos en la toma de decisiones [5]. Sin embargo, este algoritmo no depende de la siguiente acción para realizar la actualización. La ecuación del algoritmo es la siguiente:

$$Q_{t+1} = Q_t(s, a) + \alpha \left[r(s, a) + \gamma \max_{a'} Q(s', a') - Q_t(s, a) \right]$$

En donde t es el instante actual y s el estado a la acción a tomar, por lo tanto s' y a' son el estado y la acción en el siguiente instante. Por otra parte $r(s, a)$ representa la recompensa que de tomar la acción en estado actual. Los parámetros α es la tasa de aprendizaje, γ el factor de descuento que pueden tener valores entre 0 y 1.

3.4.1.2. E-greedy

Épsilon greedy o $\epsilon - greedy$ es un método utilizado ampliamente en modelos de RL que sirve para seleccionar la acción, en el cual el parámetro ϵ ($0 < \epsilon < 1$) define la policy (π). $\epsilon - greedy$ selecciona la acción máxima estimada con probabilidad $1 - \epsilon$ y una acción aleatoria con probabilidad ϵ . En algunos estudios se ha demostrado que este algoritmo tiene un resultado superior como método de aprendizaje para otros como Q-learning en comparación a otros [12].

4. DESARROLLO METODOLÓGICO

4.1. *Obtención de datos*

Para la realización de este proyecto se necesita contar con la información de los precios en el mercado de las criptomonedas que representan los diferentes posibles cruces de intercambio, obtenidos de tickers, un instrumento que contiene la última información relevante de un activo financiero, por lo que se considera que se utilizan datos en tiempo real y no información histórica para la ejecución del modelo. Sin embargo realizaron consultas de tickers en diferentes periodos de tiempo que permitieron generar una base de datos de precios en distintos momentos para la ejecución de diferentes pruebas.

Los datos para el proyecto son obtenidos de Bitso.com que, según la propia página web, se define como: “Una plataforma digital de intercambio (Exchange) NO de inversión, que proporciona seguridad y permite a los usuarios Comprar y Vender Divisas Digitales entre sí”. La obtención de información es mediante la API de Bitso que es una herramienta que permite a través de código tener acceso a la fuente de datos, en este caso el exchange, y del cual se obtiene un objeto de tipo Ticker que contiene la última información acerca del precio y transacciones de un libro determinado. Los datos que incluye el Ticker y que a su vez representan las columnas que componen la base de datos son las siguientes:

- Ask: Orden de venta más baja.
- Bid: Orden de compra más alta.
- Book: Símbolo del libro de compra.
- Date time: Marca de tiempo cuando el ticker fue generado.
- High: Precio más alto en las últimas 24 horas.
- Last: Precio de la última transacción de compra o venta ejecutada.
- Low: Precio más bajo en las últimas 24 horas.
- Volume: Volumen de las últimas 24 horas.
- Vwap: Precio medio ponderado por volumen (volume weighted average price) de las últimas 24 horas.

Las variables que se utilizan principalmente serán Ask y Bid, el motivo es que éstas representan la intención del mercado y dan sentido hacia donde se mueve la oferta y demanda ya que Ask indica el precio mínimo que un vendedor estaría dispuesto a aceptar por su activo y Bid el precio máximo que un comprador estaría dispuesto a pagar por el mismo, en este caso la criptomoneda que se indica en Book. La diferencia entre el precio de Ask y el de Bid es conocida como spread, el cual puede ser un indicador de la liquidez y volatilidad de un mercado. El constante cambio de precios es lo que propicia la presencia de oportunidades de arbitraje, al menos de forma operativa o en la práctica, por esta razón se vuelve muy importante monitorear estos precios constantemente.

Como ya se ha mencionado anteriormente, el objetivo del proyecto es encontrar la mejor ruta de intercambio entre diferentes criptomonedas de tal manera que al realizar dicho

intercambio se genere una utilidad o en su defecto la menor pérdida. De todas las opciones que con las que cuenta la plataforma para la compra y venta, se determina utilizar las siguientes criptomonedas:

- MXN. Peso mexicano, único activo que no se trata de criptomoneda sino una moneda respaldada por un banco central (Banco de México).
- USD. Stablecoins , bajo este símbolo se engloban diferentes criptomonedas que buscan replicar el precio del dólar norteamericano y cada unidad de stablecoin está respaldada por un dólar real.
- BTC. Bitcoin.
- ETH. Ethereum.
- XRP. Ripple.
- LTC. Litecoin
- MANA. Decentraland.

Por lo tanto, estas serán las criptomonedas en las cuáles se analizarán los diferentes posibles intercambios. Es importante mencionar que MXN, USD y BTC representan mercados en Bitso, es decir, que sirven como monedas base o la moneda en la que se asigna el precio, también se denominan “minor” o menor de acuerdo con la notación de la documentación de la API. Por ejemplo, con BTC es posible comprar cualquiera de la otras que no sean minor. Sin embargo, no es posible comprar ninguna moneda con ETH, XRP, LTC o MANA como base, ya que no existen los mercados al menos en la plataforma de Bitso. Por lo tanto, existen ciertas restricciones entre los posibles intercambios, como el hecho de que es imposible pasar de MANA a XRP directamente, por mencionar un caso. En el siguiente diagrama se detallan las formas en las que es posible realizar intercambios directos en Bitso entre las criptodivisas seleccionadas anteriormente:

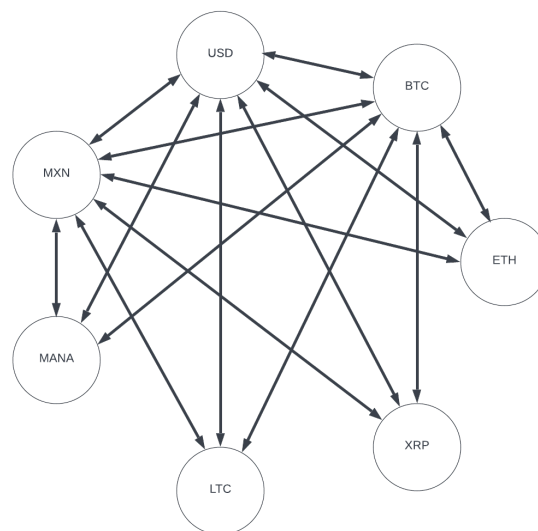


Figura 1. Rutas posibles de intercambio en Bitso. Fuente: Elaboración propia.

De acuerdo con lo que se muestra en la Figura 1, una posible ruta podría ser USD - ETH - BTC - MANA - USD, pero una que no podría existir sería USD - LTC - XRP - USD, solo por mencionar algunos ejemplos.

4.2. Formulación matemática del problema

Partiendo de las adaptaciones mencionadas al TSP, el planteamiento matemático del problema pretende encontrar la ruta que maximice la utilidad en el intercambio entre criptomonedas queda de la siguiente manera:

$$\begin{aligned} & \max \sum_{i=1}^N \sum_{j=1}^N u_{ij} c_{ij} \\ & \text{subject to:} \\ & \sum_{i=1}^N c_{ij} = 1 \quad j = 1, \dots, N \quad (1) \\ & \sum_{j=1}^N c_{ij} = 1 \quad i = 1, \dots, N \quad (2) \\ & c_{ij} \in \{0,1\} \quad (3) \end{aligned}$$

Donde N representa el conjunto de las criptomonedas como nodos. La variable u_{ij} representa la utilidad obtenida de realizar el cambio desde la criptomoneda i hacia el nodo j . Por su parte c_{ij} es una variable de control que asume 1 si el intercambio entre las criptomonedas i y j forman parte de la solución, 0 de caso contrario. La restricciones (1) y (2) indican que cada nodo se visita una única vez. La tercera (3) restricción indica que la variable c_{ij} es binaria.

4.3. Preparación de los datos

Como ya se mencionó, de Bitso se obtienen objetos Tickers de cada una de las criptomonedas seleccionadas y para cada una de las monedas menor (bid y ask), en los cuales viene contenida la información de los precios, para poder hacer uso de estos datos se creó una tabla de precios que se llena recorriendo los Tickers.

El objeto Ticker es de tipo `bitso.models.Ticker` y así es como se ve cuando se accede a él:

```
Ticker(book=usd_mxn,ask=20.179,bid=20.169, high=20.224, last=20.169, low=20.051,
volume=2214911.90676307, created_at=2022-09-30 23:10:32+00:00, vwap=20.1297208337)
```

Figura 2. Estructura del objeto Ticker de Bitso. Fuente: Elaboración propia.

Por otra parte en la Tabla 1 es posible observar los precios en el formato que servirá como información de entrada para el modelo y que representa los datos de entrenamiento del mismo.

index	mxn_bid	mxn_ask	usd_bid	usd_ask	btc_bid	btc_ask
☒	☒	☒	☒	☒	☒	☒
mxn	0	0	0	0	0	0
usd	20.088	20.093	0	0	0	0
btc	399630	399670	19864	19885	0	0
eth	27229	27249	1354.4	1355.4	0.068101	0.0683
xrp	9.6025	9.6068	0.4779	0.4781	0.00002392	0.0000241
ltc	1095.8	1098.8	54.491	54.549	0.002745	0.0027474
mana	14.22	14.24	0.7071	0.70959	0.00003556	0.00003576

Tabla 1. Precios criptomonedas seleccionadas. Fuente: Elaboración propia.

4.4. Selección del modelo

Tomando como base el trabajo de Ottoni et al. [5] *Reinforcement Learning for the traveling salesman problem with refueling* se adaptó el problema con la intención de poder utilizar un modelo de aprendizaje reforzado llamado Q-learning, de tal manera que en lugar un agente viajero que recorre ciudades se utilizaron las criptomonedas que son recorridas en los intercambios, los costos de desplazamiento fueron remplazados por el cálculo de la utilidad que generaría hacer el intercambio de dos nodos utilizando la tabla de precios obtenidos de la API de Bitso, esta utilidad es la que se busca maximizar. Por otra parte, se omitieron algunos costos relacionados con la recarga de combustible como lo plantea el trabajo original.

Para la implementación del modelo Q-learning es importante mencionar la terminología utilizada particularmente en esta adaptación, ya que estos conceptos ayudarán a entender mejor el funcionamiento:

- Q-Matrix. Matriz de aprendizaje, este instrumento es generado por el modelo, es donde se refleja el aprendizaje obtenido, de aquí se obtiene la mejor acción a tomar en cada estado generando la ruta óptima. Se actualiza de manera iterativa siguiendo la siguiente ecuación:

$$Q(s, a) = Q(s, a) + \alpha \left[r(s, a) + \gamma \left(\max_{a'} Q(s', a') \right) - Q(s, a) \right]$$

- Agent. En modelos de RL se dice que un agente es el que recorre un ambiente y toma decisiones dependiendo el estado en el que se encuentre, para cada decisión obtendrá una recompensa o penalización.

- States (Estados). Representan en dónde se encuentra ubicado el agente y dependiendo de cada estado se determinan las acciones que puede tomar. En este caso las criptomonedas representan el espacio de estados donde por ejemplo si se intercambié USD (estado inicial) por BTC (estado actual) se evaluaría a qué otra criptomoneda es conveniente pasar tomando en cuenta los precios de bitcoin.
- Actions (Acciones). Para este proyecto en particular las acciones son igual a los estados ya que dependiendo el estado se decidirá a dónde moverse, es decir, a qué criptomoneda pasar. Siguiendo con el ejemplo anterior si el agente se encuentra en el estado BTC y decide cambiar a MXN ésta representa la acción y se convertirá en estado hasta que se realice el movimiento.
- Reward (Recompensa). Es la recompensa que se otorga por realizar un movimiento de intercambio entre dos nodos, aquí esta recompensa viene dada por la utilidad a la que además se le aplicó una función cúbica para hacer más notables las diferencias, dicho de otra forma, para recompensar más a los intercambios con mayor utilidad:

$$r = u_{sa}^3$$

donde s es el estado y a la acción
y donde $u = monto_{final} - monto_{inicial}$

Cabe mencionar que para el cálculo de esta utilidad se utiliza 1 unidad con base en la moneda inicial del recorrido, determinada previamente, y tiene que ser una de las monedas menor, es decir, la utilidad de pasar de la moneda inicial hacia el estado s (en caso de que s sea diferente de inicial) y por último regresar a la inicial. Por ejemplo, si se toma MXN como base y se quiere obtener la utilidad de pasar de USD(s) a BTC(a), la recompensa sería:

$$r = \left(\frac{1 * precio_{usdmxn_{ask}} * precio_{btcusd_{ask}}}{precio_{btcmxn_{bid}}} - 1 \right)^3$$

- Route. Son los pasos recorridos entre estados y acciones o dicho de otra forma los intercambios realizados.

4.5. Ejecución del modelo

4.5.1. Parámetros

Una vez que se ha obtenido la tabla de precios es necesario seleccionar los parámetros del modelo estos son: Alpha (índice de aprendizaje) que establece la capacidad de aprendizaje del algoritmo, un valor de 0 hará que el agente no aprenda y un valor de 1 hará que ignore la información previa y solo tome en cuenta la información reciente; gamma (factor de descuento) que determina el impacto de las futuras recompensas, valores cercanos a 0 hará

que el agente considere únicamente las recompensas actuales mientras que el valor sea cercano a 1 hará que sea mayor el peso de las recompensas de las últimas etapas. Además, se establece un parámetro ϵ que se utiliza en el método e-greedy como política para elegir la acción hacia donde se moverá el agente durante el entrenamiento, aquí ϵ funciona como la probabilidad de que el agente explore el espacio de acciones, cuando el valor es bajo la prioridad será explotar el conocimiento adquirido previamente. Por una parte, la exploración permite mejorar el conocimiento a la larga descubriendo nuevos caminos sin embargo se recompensará en menor cantidad a las mejores opciones, por otro lado, la explotación elige a las opciones más óptimas y las recompensa más, a cambio podría no ser capaz de identificar la ruta que genera mayor recompensa y entregar una solución subóptima, por esta razón es importante encontrar un equilibrio que permita estas dos opciones.

$$\pi(s) = \begin{cases} \max_a Q_s(a) & \text{con probabilidad } 1 - \epsilon \\ \text{random } a & \text{con probabilidad } \epsilon \end{cases}$$

4.5.2. *Entrenamiento*

Ya establecidos los parámetros es posible realizar el entrenamiento del modelo, que consiste en inicializar la matriz Q con valores en cero y después actualizarla de manera iterativa por el agente que explora las posibles rutas y se encarga de calcular la recompensa a las acciones escogida por el método e-greedy. El entrenamiento consta de un procedimiento en donde se inicializa el estado inicial después en un bucle el agente selecciona la acción, calcula la recompensa y actualiza la matriz Q para cada uno de los nodos sin repetir los que ya se recorrieron, este procedimiento se engloba dentro de otro bucle que tiene como criterio de paro cumplir N repeticiones o episodios.

4.5.3. *Obtención de ruta*

Por último y ya teniendo la matriz Q actualizada es posible determinar la mejor ruta, aquella que maximiza la utilidad, para esto se escoge la acción con el valor más alto para cada estado de manera sucesiva, tomando como siguiente estado la acción seleccionada previamente, siempre partiendo de una criptomoneda base que esté en el grupo de las menor. Por ejemplo, partiendo desde MXN, si en la matriz el valor más alto para el estado es BTC entonces sería el siguiente estado a observar para saber hacia dónde moverse, buscando el valor más alto, así sucesivamente hasta encontrar la ruta completa ya sea moviéndose a más acciones o regresando al estado inicial, suponiendo que el caso fuera USD y el valor más alto de este estuviera en MXN entonces la ruta quedaría de la siguiente manera: MXN – BTC – USD – MXN.

	0	1	2	3	4	5	6
0	-4955.037...	49.957628...	49.95823...	49.93906...	49.94469...	49.93457...	49.913780...
1	49.957917...	-4955.037...	49.950315...	49.934018...	49.931398...	49.93562...	49.90504...
2	49.957832...	49.95868...	-4955.037...	49.93600...	49.918321...	49.927503...	49.90678...
3	49.947217...	49.949178...	49.93944...	-4955.045...	-4955.043...	-4955.048...	-4955.054...
4	49.951112...	49.93496...	49.907729...	-4955.045...	-4955.043...	-4955.048...	-4955.054...
5	49.94578...	49.942727...	49.93488...	-4955.045...	-4955.043...	-4955.048...	-4955.054...
6	49.93020...	49.939991...	49.910545...	-4955.045...	-4955.043...	-4955.048...	-4955.054...

Tabla 2. Ejemplo de matriz Q actualizada. Fuente: Elaboración propia.

Cabe mencionar que a pesar de que gracias a la matriz es posible encontrar una ruta óptima no significa necesariamente que se trata de una oportunidad de arbitraje, pues al hacer la evaluación de la ruta es posible que la utilidad sea negativa como es el caso del anterior ejemplo donde $u = -0.00022$, en este caso se encontró la ruta con menor pérdida. Por otra parte, en ocasiones puede existir una ruta con mayor utilidad, aquí entra en juego ajustar los parámetros como ϵ que permite priorizar la exploración sobre la explotación y viceversa para encontrar mejores resultados.

5. RESULTADOS Y DISCUSIÓN

5.1. Resultados

Se realizaron diferentes ejercicios de ejecución del modelo con precios recabados en diferentes periodos, como ya se mencionó anteriormente, generando una base de datos de 20 muestras de precios. Se utilizó una base de datos no estructurada de manera que fue posible almacenar los precios en el formato de tabla generada a partir de los tickers, es decir, cada tabla como un registro de la base de datos. Las consultas se realizaron en diferentes horas durante dos días, el criterio para seleccionar estos periodos fue aleatorio. Los parámetros del modelo utilizados en cada ejercicio fueron los siguientes: episodes=1000, alpha=0.75, gamma=0.9, epsilon=0.5.

De los 20 ejercicios realizados se encontró que en 10 ocasiones se detectó una oportunidad de arbitraje, es decir, la utilidad de la ruta encontrada era mayor al monto inicial que consiste en 1 unidad de peso mexicano (MXN) pues se estableció esta divisa como la base inicial de la evaluación en los ejercicios. En la tabla X se puede consultar la ruta aprendida, la utilidad que genera y la mejor utilidad encontrada durante todo el proceso iterativo, existe la posibilidad de que esta última pueda ser mejor incluso que la de la ruta aprendida debido a que en el entrenamiento se recorren rutas de manera aleatoria, que podrían ser mejores que la aprendida. Es importante mencionar que la utilidad es calculada con los mismos precios con los que fue entrenado el modelo. A pesar de que la utilidad está definida en la metodología como $u = monto_{final} - monto_{inicial}$, para fines de facilitar la lectura de los resultados se toma como utilidad el monto final, de manera que en este apartado se considera oportunidad de arbitraje cuando $u = monto_{final} > monto_{inicial} = 1$.

Ejercicio	Ruta Encontrada	Utilidad	Mejor utilidad recorrida
1	mxn,usd,mxn	0.9993549988	0.9995606237
2	mxn,usd,xrp,mxn	1.0004346371	1.0004346371
3	mxn,usd,xrp,mxn	1.0002448732	1.0002448732
4	mxn,usd,xrp,mxn	1.0006422118	1.0006422118
5	mxn,eth,usd,mxn	0.9997016601	0.99989995
6	mxn,usd,xrp,mxn	1.0128658413	1.0169093845
7	mxn,usd,xrp,mxn	1.0031771244	1.0031771244
8	mxn,usd,mxn	0.9996505068	0.9996505068
9	mxn,btc,mxn	0.9999488111	0.9999488111
10	mxn,usd,btc,mxn	0.9995820404	0.9999743944
11	mxn,usd,btc,mxn	0.999891173	1.0001557311
12	mxn,usd,xrp,mxn	1.0011324596	1.0011324596
13	mxn,usd,mxn	0.9995998199	0.9996678844
14	mxn,usd,xrp,mxn	0.9999202805	0.9999202805
15	mxn,eth,usd,mxn	1.0003866034	1.0003866034
16	mxn,usd,xrp,mxn	1.0005085577	1.0005085577
17	mxn,usd,xrp,mxn	1.0705855496	1.0718325316
18	mxn,usd,xrp,mxn	1.0017516132	1.0017516132
19	mxn,btc,usd,mxn	0.9998392737	0.9999003041
20	mxn,usd,xrp,mxn	0.9999561841	0.9999891541

Tabla 3. Resultados de rutas encontradas con utilidad por iteración. Fuente: elaboración propia.

Los resultados presentados en la tabla X muestran que en 11 ocasiones la ruta aprendida fue la mejor de las mil recorridas durante el aprendizaje (episodos), esto se podría interpretar como una precisión o accuracy del 55%. Por otra parte, tomando en cuenta los 10 casos en los que se presentaron oportunidades de arbitraje, el 80% de las veces sí se encontró la mejor ruta, esto es al menos en las pruebas realizadas.

5.2. Discusión

El modelo fue exitoso en el sentido de que fue capaz de encontrar rutas que representan oportunidad de arbitraje, además, cuando se presentaron estas ocasiones, en la mayoría de los casos la utilidad fue la más alta, por lo que es un buen instrumento para maximizar la utilidad. Si bien hubo casos en los que, debido al parámetro de aleatoriedad que permite recorrer rutas fuera del aprendizaje, se recorrieron rutas con mayor utilidad la ruta aprendida no estaba lejos de la óptima, siendo en la mayoría de los casos la segunda con mejores resultados.

Es importante señalar que en los términos en los que fueron realizados los ejercicios para determinar la utilidad no se contemplaron costos de operación o comisiones pues se trata de un ejercicio en donde el arbitraje se plantea con base en la discrepancia entre precios principalmente. Sin embargo, estos resultados no difieren tanto de una posible operación real en la plataforma de Bitso, pues el esquema de comisiones es variable dependiendo de la frecuencia de operaciones de los usuarios, por lo que en la teoría es posible operar en la plataforma con comisiones cercanas al 0% incluso alcanzando este valor en algunos casos, dependiendo del mercado.

Debido a la naturaleza del proyecto e incluso a la misma del propio arbitraje, lo planteado en este proyecto no representa una herramienta de inversión que pueda generar grandes volúmenes de ganancias monetarias, ya que en la gran mayoría de casos la oportunidad de arbitraje representa un retorno de inversión muy pequeño en relación con la cantidad a invertir. Sin embargo, la utilización del modelo puede ser relevante en la definición de otras estrategias de inversión en donde se analice el comportamiento e incluso la eficiencia de los mercados, pues como se estudió en la literatura, la presencia de oportunidades de arbitraje es indicio de ineficiencia; incluso el contar con un modelo de este tipo podría ayudar a obtener índices económicos de los mercados estudiados. Lo más relevante es que debido a que el modelo tiene como principal entrada una tabla de precios se podría adaptar a cualquier activo financiero, siempre y cuando se tenga acceso a esta información.

6. CONCLUSIONES

6.1. Conclusiones

En la actualidad en donde los mercados financieros son operados principalmente por sistemas informáticos cada vez más desarrollados, el uso de nuevas tecnologías se hace indispensable en la creación de métodos que permitan aprovechar las oportunidades que se presentan al operar en los diferentes mercados. El análisis de la información que se genera con los sistemas financieros siempre ha sido una parte fundamental para detectar patrones, tendencias u otros indicadores que permiten a los inversionistas diseñar estrategias de inversión, con el creciente desarrollo de disciplinas como la ciencia de datos que permite mediante modelos matemáticos y uso de poder computacional para realizar esta tarea, es natural que se estudie como aprovechar estas herramientas en los mercados financieros.

En ese sentido, lo que se propone mediante este trabajo es justamente replantear un problema que permita abordarlo desde el uso de la ciencia de datos. El enfoque es en el arbitraje de diversos puntos en donde se busca una utilidad a través del intercambio, en este caso aplicado a los mercados de criptomonedas, que debido a su origen generan una gran cantidad de datos y su naturaleza hace que sean objeto de estudio. Se planteó como un problema de optimización combinatoria para encontrar de manera eficiente precisamente las combinaciones o rutas que representaran una oportunidad de arbitraje de este tipo, para esto se utilizó un algoritmo de aprendizaje por refuerzo. La información se obtuvo de la plataforma de intercambios o *exchange* Bitso gracias al uso de la API que permitió acceder a los datos necesarios. Se realizaron diferentes pruebas corriendo el algoritmo y se obtuvieron resultados que permiten concluir que se cumple la hipótesis de plantear de esta forma este ejercicio de arbitraje para detectar oportunidades mediante un proceso iterativo en pocas iteraciones.

6.2. Trabajo Futuro

Como ya se ha mencionado, dependiendo de la situación, este proyecto no representa una herramienta para generar ganancias monetarias por sí mismo, sino que representa una oportunidad de analizar las situaciones que definen los comportamientos de los mercados de criptomonedas, en términos de eficiencia, para diseñar estrategias de inversión con base en los resultados obtenidos por el modelo de reinforcement learning.

Por otra parte, al ser el modelo dependiente de los precios, es posible adaptarlo a otro tipo de mercados, siempre y cuando se tenga esta información y dependiendo de la inmediatez de la disponibilidad, es posible detectar con relativa facilidad y rapidez los niveles de eficiencia del mercado en que se aplica, por lo tanto, podría ser una herramienta útil en el desarrollo de un sistema que permita alertar de manera temprana sobre problemas económicos que deriven de la volatilidad por incertidumbre de un mercado a estudiar, es

decir, saber qué significa que se estén dando cierta cantidad de oportunidades de arbitraje en un mercado y qué podría reflejar en temas económicos más profundos.

En términos del proyecto presentado, uno de los temas a estudiar a futuro es la mejora del modelo que permita obtener mejores resultados para situaciones más generalizadas, la recomendación es hacer un enfoque en el análisis más profundo sobre los parámetros del algoritmo de tal manera que se puedan definir de manera correcta para llevar a cabo el entrenamiento del modelo.

BIBLIOGRAFÍA

- [1] V. Tran y T. Leirvik, Efficiency in the Markets of Crypto-Currencies, vol. 35, Finance Research Letters, 2019, p. 101382.
- [2] H. Luna, "Caracterización de arbitraje financiero en el mercado de divisas", Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California, 2021.
- [3] I. Makarov y A. Schoar, "Trading and arbitrage in cryptocurrency markets", vol. 135, Journal of Financial Economics, 2020, pp. 293-319.
- [4] L. Lascorz, "Aprendizaje por Refuerzo Elementos básicos y algoritmos", Universidad Zaragoza, 2018.
- [5] A. L. Ottoni, E. G. Nepomuceno, M. S. de Oliveira y D. C. R. de Oliveira, Reinforcement learning for the traveling salesman problem with refueling, vol. 8, Complex & Intelligent Systems, 2021, pp. 2001-2015.
- [6] Z. Cui y S. Taylor, Arbitrage Detection Using Max Plus Product Iteration on Foreign Exchange Rate Graphs, vol. 35, Finance Research Letters, 2019.
- [7] E. F. Fama, Efficient market hypothesis, University of Chicago , 1960.
- [8] S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, Decentralized Business Review, 2008, p. 21260.
- [9] A. Shleifer y R. W. Vishny, «The Limits of Arbitrage,» *The Journal of Finance*, vol. 52, nº 1, pp. 35-55, 1997.
- [10] I. Bashir, Mastering Blockchain, 2020.
- [11] C. H. Papadimitriou y K. Steiglitz, Combinatorial Optimization: Algorithms and Complexity, Courier Corporation, 1998.
- [12] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank y A. Schrijver, Combinatorial Optimization Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, 2011.
- [13] Z.-H. Zhou, Machine Learning, Springer Nature, 2021.
- [14] M. Wunder, M. L. Littman y M. Babes, «Classes of Multiagent Q-learning Dynamics with epsilon-greedy Exploration,» *ICML*, pp. 1167-1174, 2010.

APÉNDICE A. Título

El código puede ser consultado en la siguiente liga:

<https://github.com/oscarina/TOG>