

UNIVERSIDAD PERUANA DE CIENCIAS
APLICADAS FACULTAD DE INGENIERÍA

CARRERA PROFESIONAL DE CIENCIAS DE LA
COMPUTACIÓN



Asignatura:

CC57- MACHINE LEARNING

Sección: CC57

PRACTICA CALIFICADA 2

Código	Nombre y Apellido
U202019115	Carlos Daniel Llanos Llamoca
U201922773	Alejandro Olaf Lopez Flores
U202014068	Andrea Fabiana Garcia Napuri

Lima, setiembre 2024

ÍNDICE:

I. Descripción de la situación/contexto real.....	3
II. Origen de los datos	4
III. Elaboración del Modelo.....	5
IV. Descripción de las características y origen de los datos.....	7
V. Despliegue de la red	9
VI. Referencias Bibliográficas.....	10

I. Descripción de la situación/contexto real

El mercado de criptomonedas experimenta una evolución notable, como un ecosistema financiero complejo que abarca una variedad de sectores tecnológicos emergentes, como la Inteligencia Artificial, los videojuegos, los activos del mundo real, los tokens de memes, entre otros. Cada una de estas categorías desempeña un papel crucial en la diversificación y expansión del criptomercado, al ofrecer soluciones innovadoras y atraer a una amplia gama de inversores y usuarios (Antonopoulos, 2017). Las crecientes interdependencias entre estos sectores, impulsadas por los avances tecnológicos y el acceso globalizado a las criptomonedas, han intensificado la interconexión entre los diversos actores del mercado, generando un entorno colaborativo y dinámico (Casey & Vigna, 2018).

En estudios previos, se ha caracterizado el criptomercado como un sistema dinámico que involucra a múltiples actores operando a través de redes financieras descentralizadas. En estas redes, proyectos, plataformas y usuarios interactúan mediante flujos de valor que eliminan la necesidad de intermediarios tradicionales, facilitando un entorno donde los activos y la información se intercambian de forma ágil y segura, favoreciendo la eficiencia del sistema (Tapscott & Tapscott, 2016).

Con el fin de comprender mejor las interacciones dentro de este ecosistema, en esta práctica se adopta un enfoque de red compleja para modelar dichas relaciones. Esto implica la creación de un grafo, donde los nodos representan proyectos criptográficos específicos dentro de las categorías mencionadas, y las aristas representan las conexiones derivadas de asociaciones comerciales, integraciones tecnológicas o flujos de valor compartido. Este modelo no solo facilita la visualización de la dinámica interna del criptomercado, sino que también ofrece una herramienta analítica que permite identificar cómo las interacciones entre los distintos sectores contribuyen al crecimiento, la estabilidad y el comportamiento colectivo del sistema (Boucher, Nascimento, & Kritikos, 2017).

Las fuentes de datos clave, como CoinMarketCap y CoinGecko, proporcionan métricas esenciales para ponderar el comportamiento de los proyectos considerados, como el valor del activo, la capitalización de mercado, el volumen de operaciones y la participación en redes sociales. Este análisis detallado permitirá profundizar en la dinámica del mercado, proporcionando información valiosa para la adquisición de datos y el posterior entrenamiento de algoritmos de Machine Learning, con el objetivo de predecir tendencias y comportamientos futuros en el ecosistema criptográfico (Mougayar, 2016).

II. Origen de los datos

Para la adquisición de los datos del mercado de criptomonedas, se ha utilizado como fuente principal la API de CoinMarketCap. Esta API es una de las plataformas más reconocidas y utilizadas para obtener información actualizada y precisa sobre criptomonedas. La elección de esta fuente se debe a su amplio conjunto de datos y a la confiabilidad que ofrece a los desarrolladores e investigadores del ecosistema cripto.

1. CoinMarketCap API

La API de CoinMarketCap es una herramienta que proporciona acceso a datos detallados sobre criptomonedas en tiempo real. A través de su endpoint `listings/latest`, se puede obtener información clave sobre los activos digitales más recientes, con la posibilidad de personalizar los resultados mediante parámetros como el rango de criptomonedas, la cantidad de datos a obtener y la moneda de conversión, en este caso, el dólar estadounidense (USD). De la **API de CoinMarketCap**, se están obteniendo los siguientes datos clave de cada criptomoneda:

1. **Rank (Rango):** La posición de la criptomoneda en el mercado global, según su capitalización de mercado. Este ranking indica la importancia relativa de cada criptomoneda en comparación con otras.
2. **Name (Nombre):** El nombre completo de la criptomoneda (por ejemplo, Bitcoin, Ethereum, etc.).
3. **Symbol (Símbolo):** El símbolo abreviado de la criptomoneda (por ejemplo, BTC para Bitcoin, ETH para Ethereum), que se utiliza comúnmente en las plataformas de trading y en el análisis de precios.
4. **Price (USD) (Precio en USD):** El valor actual de la criptomoneda en dólares estadounidenses. Este es el precio por unidad de la criptomoneda en el mercado.
5. **Market Cap (Capitalización de mercado):** El valor total de mercado de la criptomoneda, calculado multiplicando el precio actual por la cantidad total de monedas en circulación. Esta métrica proporciona una indicación del tamaño y la relevancia del proyecto dentro del mercado de criptomonedas.
6. **Volume (24h) (Volumen en 24 horas):** El valor total de las transacciones realizadas en las últimas 24 horas para esa criptomoneda, expresado en dólares estadounidenses. El volumen de 24 horas es un indicador clave de la liquidez y la actividad de trading.

```
1 #carga del dataset y muestreo de las primeras filas
2 import pandas as pd
3 df=pd.read_csv('cryptocurrency_data.csv')
4 print(df.head())
```

	Rank	Name	Symbol	Price (USD)	Market Cap	\
0	1500	Public Masterpiece Token	PMT	0.055709	2.306157e+06	
1	1501	Reality Metaverse	RMV	0.009212	2.302069e+06	
2	1502	NEST Protocol	NEST	0.000788	2.295190e+06	
3	1503	Rubic	RBC	0.013858	2.290438e+06	
4	1504	Success Kid	SKID	0.026236	2.275806e+06	

	Volume (24h)
0	2.047934e+05
1	2.154848e+05
2	1.588184e+05
3	1.092145e+06
4	1.695152e+05

III. Elaboración del Modelo

El modelo lógico planteado para este estudio se fundamenta en la creación de un grafo en el que las criptomonedas se representan como nodos, y las conexiones entre ellas, como aristas, están determinadas por sus capitalizaciones de mercado. El propósito es mostrar cómo las criptomonedas con capitalizaciones similares se agrupan y cómo el tamaño de cada una en el mercado afecta sus relaciones. La métrica principal utilizada es la capitalización de mercado, ya que permite calcular el valor total de una criptomoneda al multiplicar su precio por la cantidad en circulación. Este enfoque facilita el análisis comparativo dentro del mercado cripto.

- **Nodos** **(Actores):**
Cada nodo en el grafo representa una criptomoneda específica, con atributos asociados como su Name, Rank, Market Cap, y Volume (24h). Estos atributos permiten añadir contexto adicional a los nodos, enriqueciendo la visualización y proporcionando oportunidades para análisis futuros. Aunque el tamaño de los nodos es uniforme en la versión inicial del grafo, podría ajustarse para reflejar visualmente el Market Cap o el volumen de operaciones diario.
- **Aristas** **(Relaciones):**
Las aristas representan relaciones entre dos criptomonedas cuya capitalización de mercado tiene una diferencia menor o igual a un umbral predeterminado de 10 mil millones de USD. Esto sugiere que las criptomonedas conectadas tienen una capitalización de mercado comparable, lo que las hace relevantes entre sí en términos de influencia, popularidad y atractivo para los inversores. Las aristas no tienen una ponderación numérica explícita en este modelo inicial, pero el umbral podría ajustarse para realizar análisis más detallados.

Interpretación del Criptomercado

En este modelo, el mercado de criptomonedas se percibe como un entorno altamente competitivo donde las criptomonedas se organizan en diferentes categorías según su capitalización de mercado. Criptomonedas como SingularityNET (AGIX) y Fetch.ai (FET), que pertenecen a la categoría de inteligencia artificial (IA), forman un grupo que representa criptomonedas especializadas en este campo. Mientras tanto, criptomonedas más pequeñas dentro de la misma categoría, como Ocean Protocol (OCEAN) o Cortex (CTXC), también crean sus propios grupos, permitiendo distinguir distintos niveles de relevancia dentro del sector de IA.

Este agrupamiento es útil para encontrar competidores directos en cada segmento. Por ejemplo, criptomonedas con capitalizaciones similares podrían estar compitiendo por los mismos inversores o tener usos o aplicaciones parecidas.

Cálculo de la Ponderación de las Relaciones

En el modelo actual, las conexiones entre criptomonedas se basan en un umbral de capitalización de mercado de 10 mil millones de USD. No se ha asignado una ponderación numérica específica a las aristas, ya que las conexiones son binarias: las criptomonedas están dentro del umbral o no. Sin embargo, un análisis más profundo podría asignar una ponderación a las aristas según la diferencia exacta en la capitalización de las criptomonedas conectadas. Esto permitiría captar mejor las variaciones en el tamaño relativo de las criptomonedas y añadir más detalles a las relaciones observadas

Actores a Representar

El modelo actual es unipartito, ya que únicamente incluye criptomonedas como nodos y establece relaciones directas entre ellas en función de su capitalización de mercado. No obstante, se puede ampliar el análisis incorporando redes bipartitas o tripartitas para representar otros actores, como:

- **Inversores o grupos de inversores:** Vincular nodos de criptomonedas con actores clave del mercado, como fondos de inversión o exchanges, permitiría identificar influencias y patrones de inversión.
- **Plataformas de uso:** Las criptomonedas podrían conectarse a las plataformas donde se utilizan o intercambian. Esta red bipartita mostraría qué plataformas tienen mayor relevancia en ciertos grupos de criptomonedas.
- **Sectores de aplicación:** Representar criptomonedas junto con sus principales casos de uso (DeFi, NFT, almacenamiento de valor, etc.) generaría una red tripartita, facilitando la visualización de cómo se agrupan según su tamaño y su rol en el ecosistema.

IV. Descripción de las características y origen de los datos

Características de los Datos:

- **Variables seleccionadas:** Para cada criptomoneda se han extraído los siguientes datos:
 - Rank: Posición de la criptomoneda en el mercado según capitalización.
 - Name (Nombre): Nombre completo del proyecto.
 - Symbol (Símbolo): Abreviación del proyecto.
 - Price (USD): Precio actual en USD.
 - Market Cap: Capitalización de mercado (valor total de todas las monedas en circulación).
 - Volume (24h): Volumen de transacciones en las últimas 24 horas.

```
for crypto in data:
    rank = crypto['cmc_rank']
    name = crypto['name']
    symbol = crypto['symbol']
    price = crypto['quote']['USD']['price']
    market_cap = crypto['quote']['USD']['market_cap']
    volume_24h = crypto['quote']['USD']['volume_24h']

    writer.writerow([rank, name, symbol, price, market_cap, volume_24h])

print(f"Datos guardados correctamente en {os.path.abspath(csv_file)}")
else:
    print(f"Error en la solicitud: {response.status_code}")
```

- **Origen de los Datos:**
 - Los datos han sido adquiridos mediante la **API de CoinMarketCap**, una fuente confiable de datos en tiempo real sobre el mercado de criptomonedas.
 - Se realizó una consulta específica para obtener datos de criptomonedas clasificadas en las categorías mencionadas.

Técnicas utilizadas para la adquisición de los datos

- **Uso de la API de CoinMarketCap:** A través de una solicitud HTTP GET, se accedió a datos sobre criptomonedas, limitando los resultados a las categorías mencionadas. El código de Python se encargó de obtener la información relevante y guardarla en un archivo CSV.
- **Procesamiento de Datos:**
 - **Eliminación de duplicados:** Como parte del preprocesamiento, se eliminaron filas duplicadas.
 - **Manejo de valores atípicos (outliers):** Se utilizó la técnica de Winsorización para ajustar valores que estuvieran fuera de un rango típico, basándose en los cuartiles Q1 y Q3 de la distribución de cada variable.

```

1 import numpy as np
2 # manejo de valores atipicos
3 def normalizar_outliers(column):
4     Q1 = column.quantile(0.25)
5     Q3 = column.quantile(0.75)
6     RIC = Q3 - Q1
7
8     lower_bound = Q1 - 1.5 * RIC
9     upper_bound = Q3 + 1.5 * RIC
10
11     # Winsorización (ajustar outliers a los límites)
12     column = np.where(column < lower_bound, lower_bound, column)
13     column = np.where(column > upper_bound, upper_bound, column)
14
15     return column

```

Visualización mediante grafos

Para representar las relaciones entre las criptomonedas, se utilizó la librería NetworkX para construir un grafo complejo donde:

- Los nodos representan proyectos individuales (criptomonedas).
- Las aristas representan relaciones basadas en la similitud de la capitalización de mercado entre proyectos (es decir, si la diferencia de la capitalización está dentro de un rango especificado).
- El grafo se ha visualizado utilizando matplotlib, y se han creado subgrafos para representar diferentes conjuntos de proyectos según la categoría (IA, Videojuegos, RWA, Memes).

```

1 import networkx as nx
2
3 G = nx.Graph()
4
5 #agrega nodos
6 for index, row in dfnor.iterrows():
7     crypto_name = row['Name']
8     G.add_node(crypto_name, rank=row['Rank'], market_cap=row['Market Cap'], volume=row['Volume (24h)'])
9
10 #agregar aristas (relaciones basadas en similitud en la capitalización de mercado)
11 for i, row_i in dfnor.iterrows():
12     for j, row_j in dfnor.iterrows():
13         if i != j:
14             #relaciona criptomonedas si su capitalización de mercado está dentro de un rango similar
15             if abs(row_i['Market Cap'] - row_j['Market Cap']) < 1e10: #rango de 10 mil millones de USD
16                 G.add_edge(row_i['Name'], row_j['Name'])
17
18 plt.figure(figsize=(20, 20))
19 pos = nx.spring_layout(G, k=0.90, iterations=100)
20 nx.draw(G, pos, with_labels=True, node_size=500, font_size=8, node_color='lightblue', edge_color='gray')
21 plt.title('Grafo de Criptomonedas basado en Capitalización de Mercado')
22 plt.show()

```

Número de nodos y representación

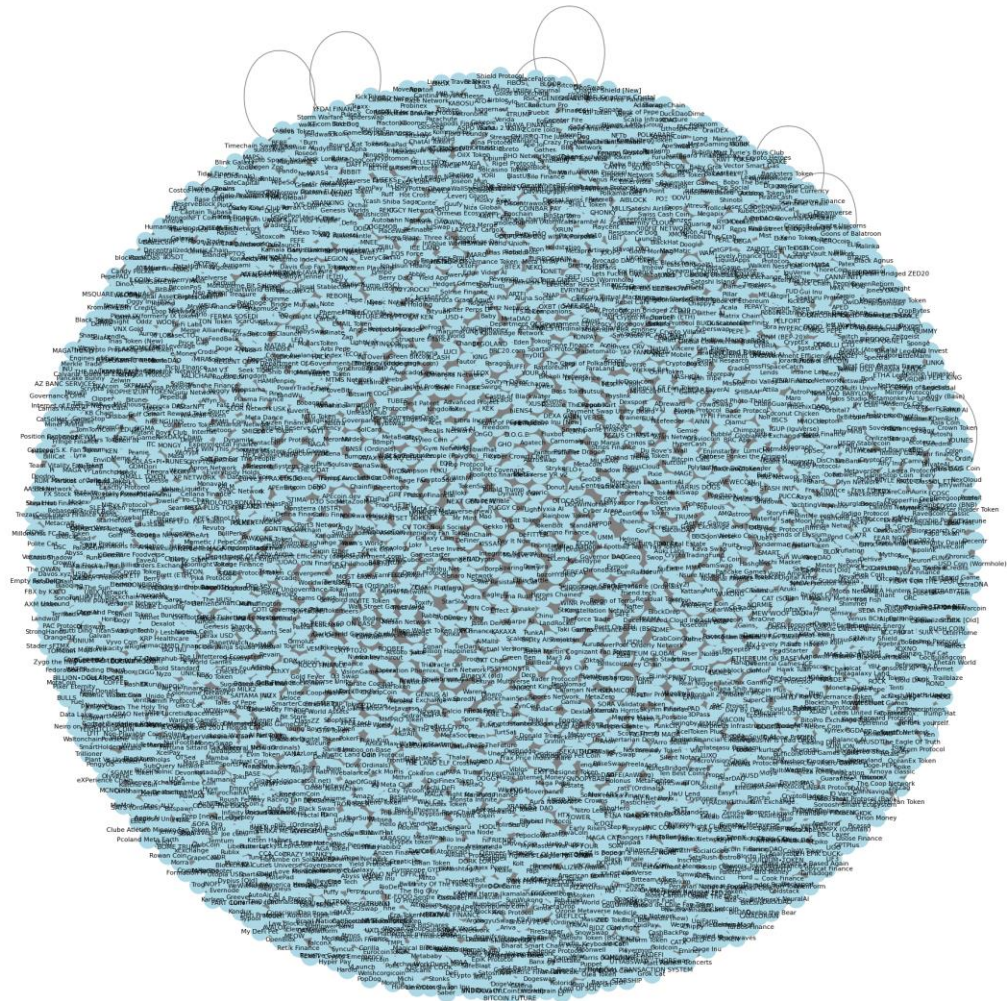
En el análisis, se han incluido al menos 1000 nodos para representar la variedad de criptomonedas. El criterio de selección fue basar la creación del grafo en la capitalización de mercado y su clasificación en una de las categorías mencionadas. Se han creado subgrafos para explorar las relaciones internas dentro de cada categoría y entre categorías.

V. Despliegue de la red

Grafo de criptomonedas según su capitalización de mercado

Se crea un grafo con los datos del dataset ya procesado, este une las criptomonedas que tengan un rango similar en la capitalización de mercado utilizando los datos 'Name' y 'Market Cap' del dataset.

Grafo de Criptomonedas basado en Capitalización de Mercado



link del github del trabajo: https://github.com/olafLopezF/Machine_learning_TB2

link de colab:

<https://colab.research.google.com/drive/1S3ZVHoCHSZl0R2vVwm3BzI7y6R5ctqbY?usp=sharing>

VI. Referencias Bibliográficas

- Antonopoulos, A. M. (2017). *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*. O'Reilly Media. Recuperado de: <https://unglueit-files.s3.amazonaws.com/ebf/05db7df4f31840f0a873d6ea14dcc28d.pdf>
- Boucher, P., Nascimento, S., & Kritikos, M. (2017). *How Blockchain Technology Could Change Our Lives*. European Parliamentary Research Service. Recuperado de: [https://www.europarl.europa.eu/RegData/etudes/IDAN/2017/581948/EPRS_IDA\(2017\)581948_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2017/581948/EPRS_IDA(2017)581948_EN.pdf)
- Casey, M. J., & Vigna, P. (2018). *The Truth Machine: The Blockchain and the Future of Everything*. St. Martin's Press. Recuperado de: <https://ilp.mit.edu/sites/default/files/2020-01/Casey.2019.Vienna.pdf>
- Mougayar, W. (2016). *The Business Blockchain: Promise, Practice, and Application of the Next Internet Technology*. Wiley. Recuperado de: <http://www.scielo.org.pe/pdf/derecho/n83/a03n83.pdf>
- Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Recuperado de: <https://bitcoin.org/bitcoin.pdf>
- CoinMarketCap. (2024.). Most trusted Cryptocurrency Market Data API | CoinMarketCap. coinmarketcap.com. <https://coinmarketcap.com/api/>