

¿Que tanto vale el arte digital?

Predicción del precio de los NFTs usando RNCs

Integrantes:

Fernando Castrillón, Andres Chaparro, Sergio Pinilla & Juan
Diego Valencia

27 de abril de 2022



Contenido

- 1 Problema
- 2 Literatura y contribución
- 3 Datos
- 4 Metodología
- 5 Hipótesis y evaluación

¿Qué es un NFT?

- Los Non-Fungible Tokens (NFTs) son activos financieros digitales representados en arte, objetos coleccionables virtuales de videojuegos, entre otros, que se tranzan en un mercado descentralizado utilizando un sistema de blockchain.

¿Qué es un NFT?

- Los Non-Fungible Tokens (NFTs) son activos financieros digitales representados en arte, objetos coleccionables virtuales de videojuegos, entre otros, que se tranzan en un mercado descentralizado utilizando un sistema de blockchain.
- A diferencia de las criptomonedas, cada NFT se constituye como un **objeto digital único** y no es homologable con otro.

¿Se puede predecir el precio de los NFT's, en formato imagen, utilizando redes neuronales?

Motivación

- El mercado de NFTs es **nuevo y de rápido crecimiento**, y, junto con el mercado de las criptomonedas, cada vez está más interconectado con los mercados financieros.

Motivación

- El mercado de NFTs es **nuevo y de rápido crecimiento**, y, junto con el mercado de las criptomonedas, cada vez está más interconectado con los mercados financieros.
- Los NFTs son **altamente volátiles** y gran parte de las variaciones en los precios de este mercado todavía permanecen inexplícadas (Borri, Liu & Tsyvinski, 2022).

Motivación

- El mercado de NFTs es **nuevo y de rápido crecimiento**, y, junto con el mercado de las criptomonedas, cada vez está más interconectado con los mercados financieros.
- Los NFTs son **altamente volátiles** y gran parte de las variaciones en los precios de este mercado todavía permanecen inexplicadas (Borri, Liu & Tsyvinski, 2022).
- Hoy en día existe una gran variedad de investigaciones que buscan entender mejor el comportamiento de los NFTs, dada la **incertidumbre** actual sobre el tema.

Motivación

- El mercado de NFTs es **nuevo y de rápido crecimiento**, y, junto con el mercado de las criptomonedas, cada vez esta mas interconectado con los mercados financieros.
- Los NFTs son **altamente volátiles** y gran parte de las variaciones en los precios de este mercado todavía permanecen inexplicadas (Borri, Liu & Tsyvinski, 2022).
- Hoy en dia existe una gran variedad de investigaciones que buscan entender mejor el comportamiento de los NFTs, dada la **incertidumbre** actual sobre el tema.
- Los tradicionales modelos de series de tiempo son **poco efectivos** para predecir el comportamiento de variables financieras, mas aun cuando hay elevada volatilidad (Diebold, 2017).

Relevancia: Un mercado en auge

La cripto más famosa involucrada con NFTs es Ethereum. En la actualidad, OpenSea es el mayor mercado de NFTs, avaluado en USD 13,3 billones. En marzo del 2022, el volumen de mercado fue de USD 2,5 billones.

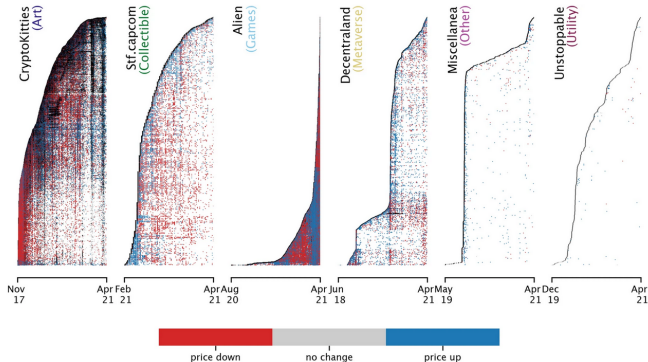


Capitalización de mercado de principales proyectos NFT (Young, 2021)

Relevancia: NFT Market

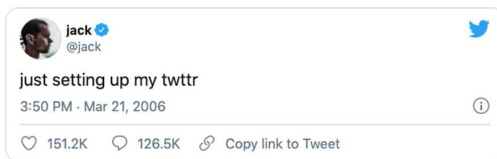
En julio de 2020 (A inicios de la Pandemia del COVID 19) se empezó a popularizar este mercado.

Cambio en los precios de algunos NFTs (Borri, Liu & Tsyvinski, 2022)



Relevancia: NFT Market

Es un mercado **altamente especulativo**. El NFT del primer Tweet se vendió a \$2.9 Millones.



Want to buy this tweet?

The highest offer is \$2500000 by [@sinaEstavi](#)

\$0.00

(💎 0.0000)

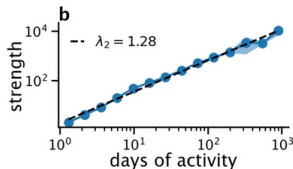
OFFER

Counter-offer must be a minimum increase of \$1 or 10%, whichever is more.

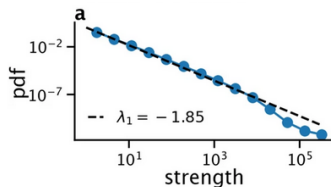
Fuente: Semana (2021)

Características de los usuarios de NFT's en la red de NFT's

Entiéndase 'fuerza' como la cantidad de transacciones que hace un usuario en la red de NFT's.



Fuerza del Comerciante en función los días de actividad

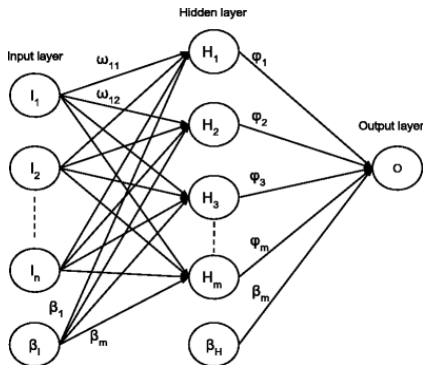


Función de densidad de probabilidad de la Fuerza

Fuente: Nadini et al. (2021)

¿Cual es la solución?

Utilizar **redes neuronales** para clasificar imagenes con el objetivo de predecir el precio de los NFTs.



Fuente: Faris et al., 2016

Redes neuronales utilizando imágenes

De acuerdo con Saad, Tareq Abed & Saad (2017) y O'Shea & Nash (2015):

- Las redes neuronales artificiales (RNA) pueden ser utilizadas para **clasificar** conjuntos de imágenes, sin embargo, estas requieren de **mucha capacidad computacional** para detectar patrones visuales específicos.

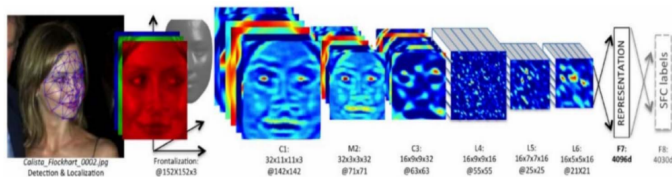
Redes neuronales utilizando imágenes

De acuerdo con Saad, Tareq Abed & Saad (2017) y O'Shea & Nash (2015):

- Las redes neuronales artificiales (RNA) pueden ser utilizadas para **clasificar** conjuntos de imágenes, sin embargo, estas requieren de **mucha capacidad computacional** para detectar patrones visuales específicos.
- El desarrollo de las redes neuronales convolucionales (RNC) mejoró la parsimonia y **simplificó** el proceso de reconocimiento de patrones en imágenes mediante redes neuronales.

Redes neuronales utilizando imágenes

Las RNC poseen una capas adicionales, las **capas convolucionales**, que facilitan la abstracción de patrones visuales complejos sin caer en un problema de *over-fitting*, lo cual incrementa su eficiencia relativa y mejora sus pronósticos fuera de muestra (Saad, Tareq Abed & Saad, 2017; O'Shea & Nash, 2015).



Fuente: Saad, Tareq Abed & Saad, 2017

Casos de uso de las RNC: Misceláneos

En la actualidad, RNC corresponde a una metodología estandarizada para realizar predicciones a partir de la información de imágenes:

Casos de uso de las RNC: Misceláneos

En la actualidad, RNC corresponde a una metodología estandarizada para realizar predicciones a partir de la información de imágenes:

- Se han combinado las RNC con modelos de procesamiento de lenguaje para predecir el **riesgo de abuso de sustancias** (Hasanpour et al.,2018).

Casos de uso de las RNC: Misceláneos

En la actualidad, RNC corresponde a una metodología estandarizada para realizar predicciones a partir de la información de imágenes:

- Se han combinado las RNC con modelos de procesamiento de lenguaje para predecir el **riesgo de abuso de sustancias** (Hasanpour et al.,2018).
- Utilizando imagenes e información de los perfiles de redes sociales, se obtiene un modelo cuya identificación del riesgo es **mejor que el azar**, con resultados estadísticamente significativos para el **uso del alcohol** (precision=68.6 %, recall=76.6 %)

Casos de uso de las RNC: Misceláneos

- También, se emplearon RNC para predecir la orientación política de personas utilizando reconocimiento facial (Kolinski, 2021).

Casos de uso de las RNC: Misceláneos

- También, se emplearon RNC para predecir la orientación política de personas utilizando reconocimiento facial (Kolinski, 2021).
- Los resultados muestran que el modelo **clasifica correctamente al 72 %** de las personas entre liberales o conservadoras, lo cual supera a la clasificación aleatoria (50 %), humana (55 %) y de un test de personalidad (66 %)

Casos de uso de las RNC: Precios

- Chen et al. (2018) por otro lado utilizan un modelo de *Machine Learning* para analizar el precio y las características que lo afectan para dos bienes: bicicletas y carros.

Casos de uso de las RNC: Precios

- Chen et al. (2018) por otro lado utilizan un modelo de *Machine Learning* para analizar el precio y las características que lo afectan para dos bienes: bicicletas y carros.
- Utilizando una base de datos de imágenes, logran encontrar: (i) las características mas relevantes para determinar el precio, y (ii) **las RNC son mas precisas para predecir el precio respecto a otros modelos** ($F_1=0,80$; bicicletas) ($F_1=0,82$; carros), como clasificadores lineales ($F_1=0,43$; bicicletas) ($F_1=0,82$; carros).

Casos de uso de las RNC: Precios



Fuente: Chen et al. (2018)

Casos de uso de las RNC: Activos financieros

Adicionalmente, durante los últimos años, ha surgido literatura que busca estudiar si las RNC pueden predecir el precio y el comportamiento de activos financieros, a partir de imágenes.

Casos de uso de las RNC: Activos financieros

Adicionalmente, durante los últimos años, ha surgido literatura que busca estudiar si las RNC pueden predecir el precio y el comportamiento de activos financieros, a partir de imágenes.

- Sim, Kim & Ahn (2019), contrastan el desempeño de un RNC, un RNA y un MVS en la predicción de una imagen de una serie de tiempo del S&P 500, y encuentran que el RNC posee **ganancias significativas** en la capacidad predictiva, relativo a los otros modelos (llega a un *hit ratio* del 68%).

Casos de uso de las RNC: Activos financieros

- Jin & Kwon (2021) complementan esta literatura al encontrar que, efectivamente, para imágenes de series de tiempo financieras (KOSPI Index) los RNC pueden alcanzar un *accuracy* del 64 %, **superando a metodologías de pronóstico tradicionales** como los MLP. Su desempeño mejora cuando se procesan imágenes limpias con colores mas saturados.

Casos de uso de las RNC: Activos financieros

- Jin & Kwon (2021) complementan esta literatura al encontrar que, efectivamente, para imágenes de series de tiempo financieras (KOSPI Index) los RNC pueden alcanzar un *accuracy* del 64 %, **superando a metodologías de pronóstico tradicionales** como los MLP. Su desempeño mejora cuando se procesan imágenes limpias con colores mas saturados.
- Worth (2020), encuentra que se puede predecir de forma eficaz (*accuracy* de 86 %) el precio de subasta de una pintura utilizando un RNC que incluya tanto fotos de las obras de arte como información histórica acerca de su precio y su valuación.

Contribución

- Primer ejercicio académico en el cual se va a intentar predecir, mediante RNC, el valor de un NFT de acuerdo con sus características visuales, así como se ha intentado hacer **previamente con las obras de arte.**

Contribución

- Primer ejercicio académico en el cual se va a intentar predecir, mediante RNC, el valor de un NFT de acuerdo con sus características visuales, así como se ha intentado hacer **previamente con las obras de arte**.
- Se expandirá la literatura acerca de **valoración de activos financieros** el mediante uso de imágenes y redes neuronales.

Contribución

- Primer ejercicio académico en el cual se va a intentar predecir, mediante RNC, el valor de un NFT de acuerdo con sus características visuales, así como se ha intentado hacer **previamente con las obras de arte**.
- Se expandirá la literatura acerca de **valoración de activos financieros** el mediante uso de imágenes y redes neuronales.
- Puede servir como un **benchmark**, tanto para compradores como para vendedores, de la valoración en el mercado de los NFTs. Lo cual, puede servir como un insumo de **toma de decisiones** dentro de las subastas de estos activos.

Descripción de los datos

Cuadro: Tabla de estadísticas descriptivas

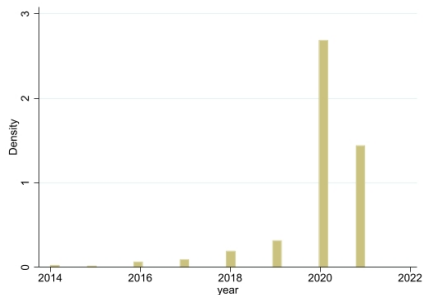
VARIABLES	(1) Tamaño	(2) media	(3) ds	(4) max	(5) min	(6) p75	(7) p90	(8) p99
Precio (Hive)	2,934	767.832	16025.08	822222.2	2.5	150	350	9999
Likes	2,934	0.4829	0.9166	7	0	1	2	4

La base fue descargada de Kaggle.com, posee datos provenientes de una variedad de *showrooms* para NFT creados entre 2014 y 2021, y el precio corresponde a su valoración en HIVE en Abril de 2021.

La criptomoneda HIVE es la moneda usada para estandarizar el valor de los NFT presentes en la base de datos.

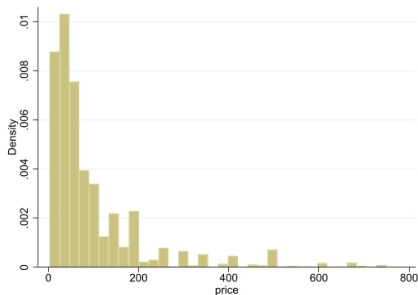
Datos

Fecha de creación del
NFT



Fuente: Cálculos propios

Distribucion de los pre-
cios de los NFT (HIVE)



Fuente: Cálculos propios

Metodología: Entrenamiento vs Prueba

- De las 2934 imágenes que tenemos en nuestra muestra, usaremos un 70 % de las imágenes para el entrenamiento del modelo y las restantes serán usadas para el testeo del modelo.

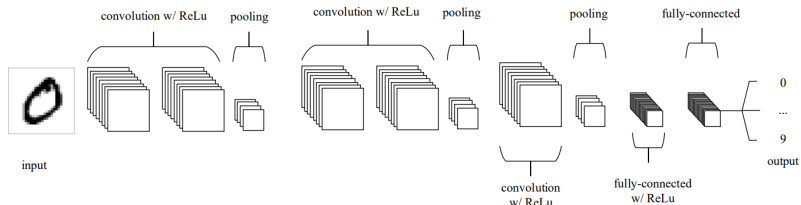
Metodología: Entrenamiento vs Prueba

- De las 2934 imágenes que tenemos en nuestra muestra, usaremos un 70 % de las imágenes para el entrenamiento del modelo y las restantes serán usadas para el testeo del modelo.
- En principio, usaremos datos atípicos dentro de nuestra muestra de entrenamiento y de testeo.

Metodología: Entrenamiento vs Prueba

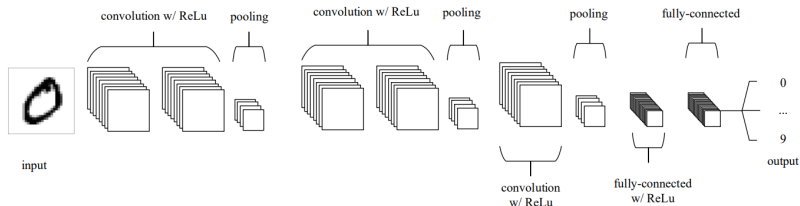
- De las 2934 imágenes que tenemos en nuestra muestra, usaremos un 70 % de las imágenes para el entrenamiento del modelo y las restantes serán usadas para el testeo del modelo.
- En principio, usaremos datos atípicos dentro de nuestra muestra de entrenamiento y de testeo.
- Luego haremos el ejercicio sin tener en cuenta estos datos atípicos. Los datos atípicos son aquellos para los cuales hay niveles de precios exorbitantes (ej. un precio de 8,222,222.2 HIVE por un NFT).

Red neuronal convolucional (RNC)



Fuente: O'Shea & Nash (2015)

Red neuronal convolucional (RNC)



Fuente: O'Shea & Nash (2015)

Dentro del procesamiento con RNC se utilizan cuatro tipos de capas diferentes: primero, una capa de *input* con la información de color por cada pixel de cada NFT; segundo, capas de convolución; tercero, capas de *pooling*; y cuarto, capas *fully-connected* que corresponden a las capas tradicionales de una RNA, con sus respectivas funciones de activación.

Convolutional y Pooling Layers

En una **capa convolucional** se seleccionan subgrupos locales de píxeles de una imagen del *input* y mediante un producto punto se les aplica un **filtro matricial**, el cual permite captar características específicas de la imagen y sirve para reducir ligeramente su tamaño:

Convolutional y Pooling Layers

En una **capa convolucional** se seleccionan subgrupos locales de píxeles de una imagen del *input* y mediante un producto punto se les aplica un **filtro matricial**, el cual permite captar características específicas de la imagen y sirve para reducir ligeramente su tamaño:

- Facilita la identificación de patrones gráficos en las imágenes.
- Reduce el trabajo de las neuronas de las capas sencillas.
- Mitiga el *over-fitting* y mejora el ajuste por fuera de muestra.

Convolutional y Pooling Layers

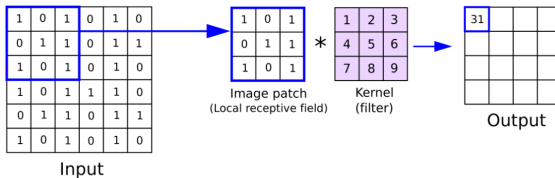
En una **capa convolucional** se seleccionan subgrupos locales de píxeles de una imagen del *input* y mediante un producto punto se les aplica un **filtro matricial**, el cual permite captar características específicas de la imagen y sirve para reducir ligeramente su tamaño:

- Facilita la identificación de patrones gráficos en las imágenes.
- Reduce el trabajo de las neuronas de las capas sencillas.
- Mitiga el *over-fitting* y mejora el ajuste por fuera de muestra.

En una **capa pooling** se reduce la dimensionalidad de la matriz (de la muestra) imputada, al colapsar la información por subgrupos de píxeles. Lo cual, disminuye la complejidad de la información a procesar y aligera los requerimientos de capacidad computacional.

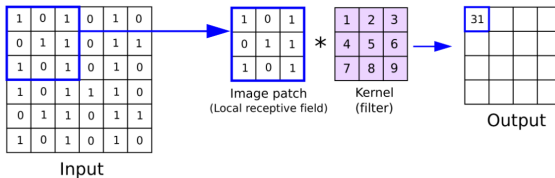
Convolutional y Pooling Layers

Capa convolucional (Reynolds, 2019)

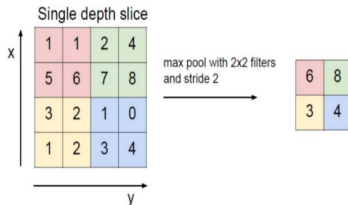


Convolutional y Pooling Layers

Capa convolucional (Reynolds, 2019)



Capa *pooling* (Saad, Tareq Abed & Saad, 2017)



Hipotesis y resultados esperados

Hipótesis: Existen características del arte en la imagen que hacen que los usuarios las valoren más que otras.

Criterios de Evaluación: Accuracy

Se utilizará como medida de evaluación el *accuracy* del modelo, es decir, el porcentaje de predicciones correctas del precio de los NFTs

Criterios de Evaluación: Accuracy

Se utilizará como medida de evaluación el *accuracy* del modelo, es decir, el porcentaje de predicciones correctas del precio de los NFTs

- Como la variable de resultado es **continua**, se dividirá la distribución de precios en deciles.

Criterios de Evaluación: Accuracy

Se utilizará como medida de evaluación el *accuracy* del modelo, es decir, el porcentaje de predicciones correctas del precio de los NFTs

- Como la variable de resultado es **continua**, se dividirá la distribución de precios en deciles.
- Una predicción será **correcta** si logra ubicar el precio predicho en el **mismo decil** que el precio observado. De lo contrario, se considerará la predicción como incorrecta.

Criterios de Evaluación: RMSE y MAE

Posteriormente, calcularemos los precios predichos de cada NFT y obtenemos el error de predicción como la diferencia entre entre el valor predicho y el valor observado: $\epsilon_i = Y_i - \hat{Y}_i$

Criterios de Evaluación: RMSE y MAE

Posteriormente, calcularemos los precios predichos de cada NFT y obtenemos el error de predicción como la diferencia entre el valor predicho y el valor observado: $\epsilon_i = Y_i - \hat{Y}_i$

- **RMSE:** Calculamos: \sqrt{MSE} donde $MSE = \frac{1}{N} \sum_{i=1}^N \epsilon_i^2$

Criterios de Evaluación: RMSE y MAE

Posteriormente, calcularemos los precios predichos de cada NFT y obtenemos el error de predicción como la diferencia entre el valor predicho y el valor observado: $\epsilon_i = Y_i - \hat{Y}_i$

- **RMSE:** Calculamos: \sqrt{MSE} donde $MSE = \frac{1}{N} \sum_{i=1}^N \epsilon_i^2$
- **MAE:** Calculamos: $\frac{1}{N} \sum_{i=1}^N |\epsilon_i|$

Criterios de Evaluación

Estos estadísticos los compararemos con los mismos criterios de información extraídos de otros dos modelos:

- Modelo Logístico.
- Distribución Aleatoria no Paramétrica.

Criterios de Evaluación

Estos estadísticos los compararemos con los mismos criterios de información extraídos de otros dos modelos:

- Modelo Logístico.
- Distribución Aleatoria no Paramétrica.

El modelo con el menor criterio de información sobre la muestra de prueba corresponderá a aquel con el mejor ajuste *out-of-sample*.

Criterios de Evaluación

Estos estadísticos los compararemos con los mismos criterios de información extraídos de otros dos modelos:

- Modelo Logístico.
- Distribución Aleatoria no Paramétrica.

El modelo con el menor criterio de información sobre la muestra de prueba corresponderá a aquel con el mejor ajuste *out-of-sample*.

Resultados esperados:

Criterios de Evaluación

Estos estadísticos los compararemos con los mismos criterios de información extraídos de otros dos modelos:

- Modelo Logístico.
- Distribución Aleatoria no Paramétrica.

El modelo con el menor criterio de información sobre la muestra de prueba corresponderá a aquel con el mejor ajuste *out-of-sample*.

Resultados esperados:

1. El diseño del modelo permita predecir el precio usando las características de las imágenes.

Criterios de Evaluación

Estos estadísticos los compararemos con los mismos criterios de información extraídos de otros dos modelos:

- Modelo Logístico.
- Distribución Aleatoria no Paramétrica.

El modelo con el menor criterio de información sobre la muestra de prueba corresponderá a aquel con el mejor ajuste *out-of-sample*.

Resultados esperados:

1. El diseño del modelo permita predecir el precio usando las características de las imágenes.
2. El modelo de RNC logra una predicción **mas acertada** del precio de los NFTs que el modelo logístico y que una clasificación por elección aleatoria.

Bibliografía

- [1] Borri, N., Liu, Y. and Tsyvinski, A., Aleh, The Economics of Non-Fungible Tokens (March 7, 2022). Available at SSRN: <https://ssrn.com/abstract=4052045> or <http://dx.doi.org/10.2139/ssrn.4052045>
- [2] Chen, Steven Chou, Edward Yang, Richard. (2018). The Price is Right: Predicting Prices with Product Images.
- [3] Diebold, F.X. (2017), Forecasting, Department of Economics, University of Pennsylvania, <http://www.ssc.upenn.edu/fdiebold/Textbooks.html>.
- [4] Faris, H., Aljarah, I. Mirjalili, S. Training feedforward neural networks using multi-verse optimizer for binary classification problems. Appl Intell 45, 322–332 (2016). <https://doi.org/10.1007/s10489-016-0767-1>

- [5] Hassanpour, S., Tomita, N., DeLise, T. et al. Identifying substance use risk based on deep neural networks and Instagram social media data. *Neuropsychopharmacol* 44, 487–494 (2019).
<https://doi.org/10.1038/s41386-018-0247-x>
- [6] Jin, G., Kwon, O. (2021). Impact of chart image characteristics on stock price prediction with a convolutional neural network. *PloS one*, 16(6), e0253121.
<https://doi.org/10.1371/journal.pone.0253121>
- [7] Kosinski, M. Facial recognition technology can expose political orientation from naturalistic facial images. *Sci Rep* 11, 100 (2021).
<https://doi.org/10.1038/s41598-020-79310-1>
- [8] Matney, L. (4 de enero de 2022). NFT kingpin OpenSea lands monster \$13.3B valuation in new raise. URL:
<https://techcrunch.com/2022/01/04/nft-kingpin-opensea-lands-13-3b-valuation-in-300m-raise-from-paradigm-and-coatue/>

- [9] Mason, B. (1 de abril de 2022). OpenSea NFT Trading Volumes Weaker for a Second Consecutive Month. Yahoo Finance. URL: <https://finance.yahoo.com/news/opensea-nft-trading-volumes-weaker-111217549.html>
- [10] Nadini, M., Alessandretti, L., Di Giacinto, F. et al. Mapping the NFT revolution: market trends, trade 3networks, and visual features. Sci Rep 11, 20902 (2021). <https://doi.org/10.1038/s41598-021-00053-8>
- [11] O'Shea, K., Nash, R. (2015). An Introduction to Convolutional Neural Networks. ArXiv, abs/1511.08458.
- [12] Reynolds, Anh (2019). Convolutional Neural Networks (CNNs). Recuperado de: <https://anhreynolds.com/blogs/cnn.html>

- [13] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.
- [14] Semana. (8 de marzo de 2021). Un tuit de 2,5 millones de dólares: la puja histórica por el primer mensaje de Jack Dorsey en Twitter. URL: <https://www.semana.com/economia/articulo/un-tuit-de-25-millones-de-dolares-la-puja-historica-por-el-primer-mensaje-de-jack-dorsey-en-twitter/202122/>
- [15] Sim, H. S., Kim, H. I., Ahn, J. J. (2019). Is deep learning for image recognition applicable to stock market prediction?. Complexity, 2019.
- [16] Worth, T. (2020). Painting2Auction: Art Price Prediction with a Siamese CNN and LSTM.

- [17] Young, J (2021). NFT Market Rages On: NFTs Market Cap Grow 1,785 % In 2021 As Demand Explodes. Forbes. Recuperado de: <https://www.forbes.com/sites/youngjoseph/2021/03/29/nft-market-rages-on-nfts-market-cap-grow-1785-in-2021-as-demand-explodes/?sh=60abd91d7fdc>
- [18] De Haan, A. (2021). NFT art collection 2021. Kaggle. URL: <https://www.kaggle.com/datasets/vepnar/nft-art-dataset>

Gracias por su atención!