

¿Que tanto vale el arte digital?

Predicción del precio de los NFTs usando RNCs

Fernando Castrillón (201730699), Andrés Chaparro (201813157),
Sergio David Pinilla (201814755) & Juan Diego Valencia (201815561)

código: <https://github.com/sdpinilla18/Predicting-NFTs-using-CNN.git>

1. Introducción

Los Non-Fungible Tokens (NFTs) son unos de los activos financieros mas populares en el internet. Estos se componen principalmente de arte digital, objetos coleccionables virtuales de videojuegos, entre otros. Se tranzan en un mercado descentralizado utilizando los sistemas de Blockchain y, a diferencia de otros criptoactivos como las Criptomonedas, los NFTs son objetos digitales únicos no homologables con otros, son no *fungibles*. En los últimos años es que el mercado de los NFTs ha explotado, teniendo un rápido crecimiento desde inicios de la pandemia del Coronavirus, y hoy en día la plataforma OpenSea (el mayor mercado global de NTFs) esta avaluada en USD 13.3 billones y su volumen mensual de transacciones ronda los USD 2.5 billones.

Adicionalmente, no solo este mercado es de rápido crecimiento, sino que la misma literatura ha identificado que el comportamiento del precio de estos activos es muy volátil (Borri, Liu Tsyvinski, 2022). Lo cual dificulta la posibilidad de realizar buenos ejercicios predictivos, dado que las técnicas tradicionales de series de tiempo son poco robustas ante la presencia de series muy volátiles (Diebold, 2017). Por lo tanto, se vuelve relevante la necesidad de investigar si existen otras aproximaciones metodológicas que permitan predecir de forma mas efectiva el comportamiento de estos criptoactivos. Consecuentemente, en este trabajo se va evaluar si se puede predecir el precio de los NTFs, en formato imagen, utilizando redes neuronales convolucionales (RNC). Además, para cumplir con este objetivo, se va a partir del supuesto de que existen características gráficas y/o artísticas en cada imagen que influyen su valoración.

Este documento se divide en seis (6) secciones. La primera sección (1) presentó una introducción al mercado de NFTs, la relevancia del tema y la pregunta de investigación del trabajo. La segunda sección (2) presenta una revisión de literatura de los documentos relevantes en el uso de redes neuronales para la clasificación de imágenes. La tercera sección (3) realiza una descripción de los datos utilizados. La cuarta sección (4) especifica la metodología empleada. Resultados y evaluación son presentados en la sección cinco (5). Finalmente, en la sección seis (6) se provee una conclusión del ejercicio.

2. Revisión de literatura

Nuestro trabajo se enmarca en la aplicación de las redes neuronales para el análisis y la predicción de características relacionadas con imágenes. Los estudios de este tipo utilizaron primero las redes neuronales artificiales (ANN), una metodología de aprendizaje supervisado que emplea operaciones no lineales a lo largo de tres o mas capas que en su conjunto, permiten elaborar predicciones sobre variables de resultado. La primera capa recibe la información de las variables

que se usarán para el pronóstico. Luego, se combina linealmente dicha información a partir de unos pesos que se obtienen de un proceso de optimización. Esta combinación pasa por una o varias capas “ocultas” que aplican funciones de activación no lineales. El resultado de este proceso entra finalmente a la capa de *output*, la cual a partir de unos pesos combina los resultados de las neuronas de las capas ocultas y provee una predicción en términos de los *outputs* del modelo.

Sin embargo, existen dos problemas al aplicar *ANN* a las imágenes. El primero es que el procesamiento es bastante exigente en términos de capacidad computacional. El segundo es que, debido a la magnitud de la información que entra a la primera capa, a menudo la aplicación de las neuronas genera un problema de *overfitting* o sobreajuste del modelo (O’Shea & Nash, 2015). Por tanto, el estándar actual en el procesamiento de imágenes son las redes neuronales convolucionales (*CNN*), una metodología que innova las *ANN* al complementar la red con capas adicionales denominadas *capas convolucionales*. La ventaja de estas es que simplifican los procesos al capturar subconjuntos de las imágenes que permiten identificar, luego de aplicar una serie de filtros matriciales, patrones y aspectos notables en las imágenes. Esta reducción de dimensionalidad del problema genera que el cálculo de los pesos y las activaciones no lineales sea menos exigente, lo cual soluciona los dos problemas de las *ANN*. Por un lado, disminuye los requerimientos de capacidad computacional, pero también evita el problema de *overfitting* gracias a que disminuye el número de parámetros (pesos) que se deben calcular (Saad, Tareq Abed & Saad, 2017; O’Shea & Nash, 2015).

Por otro lado, existen trabajos que aplican la misma estrategia propuesta en este documento: predecir el precio de bienes mediante *CNNs* basándose únicamente en imágenes. Por ejemplo, utilizando una base de datos que contiene fotografías de bicicletas y carros, los autores (Chen et al., 2018) emplean distintos modelos de clasificación lineal, o basados en redes neuronales y otros utilizando *CNNs* con el objetivo de identificar no solo el precio de los artículos, sino las características que aparentan afectarlo. Los resultados indican que las *CNNs* proveen las mejores predicciones en comparación con los demás modelos, como para el precio de las bicicletas [*CNNs* $F_1=0,80$; segundo mejor modelo $F_1=0,43$]. En últimas, se concluye que existen características intrínsecas a las imágenes de los bienes que pueden determinar el precio de mercado de estos.

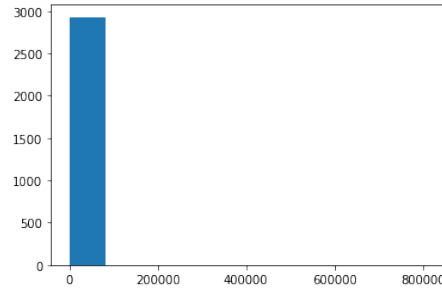
Por último, (Worth, 2020) emplea una estrategia similar a la planteada en nuestro trabajo, pues encuentra que se puede predecir de forma eficaz (accuracy de 86 %) el precio de subasta de una pintura utilizando las imágenes y la información histórica de su precio, soportando la hipótesis de que existen características visuales que pueden determinar el precio de un bien. Sin embargo, dentro de los límites de nuestro conocimiento, no existe ningún trabajo que realice una predicción del precio de los NFTs utilizando únicamente sus imágenes. Por tanto, la contribución de este artículo recae sobre dos aspectos. Primero, se expande la literatura relacionada con la valoración de activos financieros utilizando características visuales procesadas mediante redes neuronales. Segundo, con la red neuronal se provee un marco de referencia respecto al precio de los NFTs que puede servir como insumo para guiar las decisiones de compradores y vendedores en las subastas de estos activos.

3. Datos

Para los datos se utilizó la base de dominio público “*NFT art collection 2021*” de la página Kaggle.com, en la cual se presentaban NFTs desde el año 2014 para archivos como imágenes, videos y Gifs con sus respectivos precios en la criptomoneda HIVE. De estos formatos se utilizaron únicamente las imágenes (jpg, png, jpeg), lo cual resultó en una muestra primaria de 2934 imágenes, la mayoría creadas a partir del año 2020.

La figura 1 muestra la distribución de los precios para la base de datos. Se puede observar una alta concentración de los precios alrededor del 0, aunque la existencia de valores atípicos puede distorsionar la escala. El cuadro 1 muestra las estadísticas descriptivas del precio de los NFTs. El precio promedio de salida fue de 765.52 HIVE (alrededor de \$USD 459.312). La desviación estándar es de 16,024 HIVE (\$USD 9,614.4). El precio máximo es de 822,222 HIVE (\$USD 493,333), lo cual contrasta con el valor de corte del percentil 75, que es de 150 HIVE o \$USD 90. Esto infla la varianza, debilitando el poder predictivo del futuro modelo.

Figura 1: Distribución de los Precios pre-refinamiento



Fuente: Elaboración propia con la base de datos de *Kaggle.com*

Cuadro 1: Estadísticas descriptivas

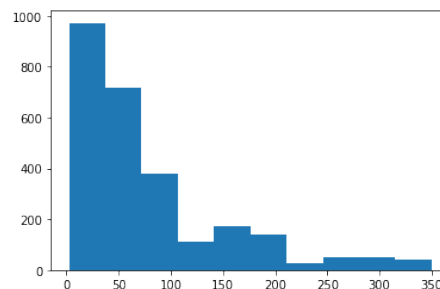
Estad.	Obs.	Media	Dev. Est.	Min	p25	p50	p75	Max
Precio	2934	765.52	16024.51	2.5	30	55	150	822222

Fuente: Elaboración propia con la base de datos de *Kaggle.com*

Para solucionar este problema de datos atípicos, se utilizó la parte de la base de datos que fuera igual o menor al percentil 90 del precio: todos aquellos NFTs que cuesten igual o menos que 350 HIVE. La figura 2 muestra la distribución de los datos luego de realizar este recorte. Se puede observar una distribución sesgada hacia la izquierda, así como una concentración alrededor del 0. Esto significa que mayoría de NFTs tienen un valor entre los 2 y 100 HIVE.

Luego, el cuadro 2 presenta las estadísticas descriptivas después de tomar los datos hasta el percentil 90. El máximo ahora es de 350. La media es de 77.95 HIVE (\$USD 46.77), mientras que la desviación estándar es de 73.95 HIVE (\$USD 44.37), que es menor a los datos de la base completa. En ultimas, se puede observar que la media y la desviación estándar ya no están infladas, por lo que el poder predictivo del futuro modelo será mayor.

Figura 2: Distribución de los precios tomando hasta el percentil 90



Fuente: Elaboración propia con la base de datos de *Kaggle.com*

Además, dentro de la base de datos de NFTs las imágenes importadas vienen en diferentes formatos y resoluciones. Entonces, para facilitar el uso de esos insumos dentro de modelos de

Cuadro 2: Estadísticas descriptivas tomando hasta percentil 90

Estad.	Obs.	Media	Dev. Est.	Min	p25	p50	p75	Max
Precio	2658	77.95	73.95	2.5	25	50	100	350

Fuente: Elaboración propia con la base de datos de *Kaggle.com*

predicción, se realizó un ejercicio de pre-procesamiento donde todos los archivos se ajustaron a dimensiones de 256x256 píxeles, y aquellas imágenes que estaban en formatos de color de uno o cuatro canales se reajustaron a un formato RGB (tres canales).

Con el fin de analizar las características de las imágenes de la base de datos, se calculó la varianza de estas como una medida acerca de la dispersión existente entre los colores de los píxeles. Si una imagen comprende un solo color, su varianza será 0. Por otro lado, una imagen con un alto contraste de colores tendrá una varianza alta. Lo anterior es importante en tanto ofrece una guía para la calibración de la red neuronal.

Para ello, primero se realiza la división de la base de datos en dos, una que servirá para el entrenamiento del modelo (70 % de la base original) y la otra que servirá para el testeo (30 % de la base original). El cuadro 3 presenta las estadísticas descriptivas de la varianza de las imágenes por cada nueva base. En la base de entrenamiento se encuentran 1,859 imágenes, mientras que en la base de testeo se cuentan con 797 imágenes. La media de las varianzas es relativamente similar, pues la de entrenamiento solo se incrementa un 3.31 % respecto a la de testeo.

Cuadro 3: Estadísticas descriptivas. Varianza de las imágenes

Estad.	Obs.	Media	Dev. Est.	Min	p25	p50	p75	Max
Varianza Entrenamiento	1859	4.93^6	3.41^6	2.05^5	2.43^6	3.98^6	6.48^6	2.70^7
Varianza Testeo	797	4.78^6	3.44^6	4.01^5	2.37^6	3.94^6	6.25^6	2.44^7

Fuente: Elaboración propia con la base de datos de *Kaggle.com*.

4. Metodología

Para cumplir con el objetivo de predecir el valor de los NFTs a partir de sus características visuales y en consistencia con la literatura revisada, en este trabajo se va a hacer uso de una *CNN* como metodología de estimación del precio de estos criptoactivos. Lo anterior implica que, en adición a las capas *densas* que usualmente se utilizan dentro de una *ANN*, se van a incluir **cuatro** nuevos tipos de capas al inicio de la red. Primero, unas capas *convolucionales* que facilitan la identificación de patrones gráficos mediante el entrenamiento de filtros matriciales (*kernels*). Luego, unas capas *pooling* que permiten reducir el tamaño de los *inputs* mientras conservan su información mas relevante. Posteriormente, para mitigar posibles problemas de sobreajuste, las capas *dropout* transforman aleatoriamente a cero una proporción de los *inputs*. Por ultimo, se adicionan capas de *batch normalization* que sirven para reestandarizar (reajustar) los *outputs* antes de que estos entren a las capas *densas* tradicionales del final de la red.

Antes de implementar este modelo, también fue necesario realizar un ejercicio de estandarización de la información. Todos los precios del *output* se dividieron sobre su valor máximo (350 HIVE) y cada uno de los canales de los píxeles de cada imagen del *input* fue dividido sobre el máximo valor que puede tomar un canal de color en RGB (255). De esta forma, cada valor de

cada elemento tanto del *input* como del *output* quedo acotado ente cero y uno.

En particular, para este trabajo se utilizó una *CNN* que, mediante aprendizaje supervisado, se encontró que era las mas efectiva (respecto a otras posibles especificaciones) a la hora de predecir el precio en HIVE de imágenes con un tamaño de 256x256 píxeles y con tres canales de color (formato RGB). La red se constituyo con: (i) cuatro capas *convolucionales*, con *kernels* de tamaño 3x3 inicializados con distribuciones normales, y con funciones de activación tipo *relu*; (ii) tres capas intermedias tipo *max-pooling* con matrices 3x3; (iii) tres capas intermedias tipo *dropout* con un *dropout rate* del 10%; (iv) cuatro capas de *batch normalization* para cerrar cada proceso de convolución; (v) una capa *densa* al final, con función de activación tipo *relu* e inicializada con una distribución normal; y, por ultimo (vi) una capa de *output* con función de activación *sigmoide* e inicializada con una distribución normal.

Para evaluar el ajuste por fuera de muestra del modelo se utilizaron tres criterios de información, la raíz del error cuadrático medio (RMSE), el error absoluto medio (MAE) y un *accuracy* por deciles. Además, dado que estas son medidas de evaluación relativas, el ajuste de la *CNN* estimada se va a contrastar con el rendimiento por fuera de muestra de otros cuatro modelos de predicción: una asignación aleatoria de precios (consistente con la distribución muestral) creada a partir de 10,000 repeticiones de Monte-Carlo; un modelo de regresión lineal; un modelo clasificador de regresión logística; y una regresión LASSO, que incorpora procesos de selección y reducción de parámetros. No obstante, diferentes criterios de evaluación operan con variables de distinto tipo. Por un lado, se estimará una *CNN* de variable dependiente continua para contrastarla en términos del RMSE y el MAE con otros modelos de variable continua, como la asignación aleatoria, la regresión lineal y el LASSO. Y, por otro lado, se estimará una *CNN* de variable categórica (por deciles) y se va a contrastar en términos de *accuracy* con otros modelos de variable categórica, como la asignación aleatoria, la regresión logística y el LASSO.

5. Resultados y evaluación

El cuadro 4 presenta los resultados de implementar y evaluar los modelos de predicción. La asignación aleatoria realizada con 10,000 repeticiones se aplicó para los *outputs* de testeo tanto del modelo continuo como del modelo categórico. La regresión lineal se aplicó únicamente al modelo continuo, por lo que no se presenta el *accuracy*. En su lugar, se aplicó una regresión logística al modelo categórico. La regresión LASSO se aplicó para ambos tipos de variable.

Cuadro 4: Criterios de evaluación. Modelos de predicción

Modelo	RMSE	MAE	Accuracy
Asignación Aleatoria	107.71	77.19	17.74 %
Regresión Lineal	91.76	69.14	- - - -
Regresión Logistica	- - - -	- - - -	23.43 %
LASSO	76.31	57.14	14.54 %
CNN	82.46	59.54	22.31 %

Fuente: Elaboración propia con la base de dgatos de *Kaggle.com*.
Las unidades del RMSE y del MAE se encuentran en HIVE.

Los resultados sugieren que la *CNN* provee un RMSE de 82.46 HIVE y un MAE de 59.54 HIVE, equivalentes a 1.05 veces y 0.763 veces la media de la distribución del precio. Estas medidas del error de predicción indican que la *CNN* posee un ajuste mayor que la regresión lineal [RMSE=91.76,MAE=69.14] y que la asignación aleatoria [RMSE=107.7,MAE=77.1]. Sin em-

bargo, el modelo LASSO [RMSE=76.31,MAE=57.14] supera a la *CNN* y es el de mejor ajuste. La estimación del modelo categórico indica que una asignación aleatoria que replica la distribución de los precios logra acertar el 17.74 % de las veces. El modelo LASSO posee un *accuracy* del 14.54 %, luego su desempeño en términos relativos es menor al que presentó en el modelo continuo. El modelo logístico logra un *accuracy* del 23.43 %, siendo el modelo que mejor se ajusta a los datos de testeo. Por ultimo, la *CNN* acierta el 22.32 % de sus predicciones, lo cual supera a la asignación aleatoria, al modelo LASSO, pero no al logístico.

A primera vista, el rendimiento predictivo *out-of-sample* de la *CNN* aparenta ser desalentador. En términos absolutos su nivel de *accuracy* difícilmente supera el 20 %, lo cual solo sería atractivo para inversionistas amantes al riesgo. Además, en términos relativos esta red se ve superada en todos los criterios de información por otros modelos mas sencillos o mas fáciles de implementar como lo son la regresión logística o el LASSO. De hecho, el *accuracy* de la *CNN* solo logra ser 4.57 puntos porcentuales superior al de una asignación aleatoria.

Sin embargo, al revisar en mas detalle, hay dos características de estos resultados que si favorecen a la *CNN*. Primero, por la misma naturaleza del activo que se esta pronosticando, ninguno de los métodos utilizados, ni siquiera los mas sofisticados como el LASSO, presentan un ajuste ostentosamente superior al de la asignación aleatoria. Y, segundo, **la *CNN* es el modelo que mejor afronta el *trade-off* sesgo-varianza** dentro de este ejercicio predictivo. Ciertos modelos, como las regresiones lineal y logística, buscan eliminar el sesgo en la estimación (mayor *accuracy*), pero sacrifican la varianza (elevados RMSE y MAE), mientras que otros modelos como el LASSO, que prefieren una varianza pequeña (bajo RMSE y MAE), sacrifican su nivel de sesgo (bajo *accuracy*). La *CNN* queda bien ubicada tanto en términos de sesgo como de varianza, quedando como la segunda mejor en todos los criterios de información, sin sacrificar significativamente a ninguno.

6. Conclusión

En el presente articulo se utilizaron *CNN* que identifican los patrones mas notables de las imágenes de los NFTs, con el fin de predecir su precio. Para tal fin, se empleó una base de datos que contenía 2,658 imágenes de NFTs. Luego de dividir la base en entrenamiento y testeo, preprocesar las imágenes y analizarlas, se implementó una *CNN* para dos tipos de modelados del precio: uno categórico y otro continuo. Los resultados se contrastaron con cuatro (4) modelos: una asignación aleatoria, una regresión lineal, una logística y otro utilizando LASSO.

Se concluye entonces que el rendimiento predictivo *out-of-sample* de la red convolucional no es el mejor. Contrario a la hipótesis inicial, es probable que existan factores mucho mas influyentes que expliquen el precio además de las características inherentes al diseño de los NFTs, como la especulación, la difusión que hayan tenido en la plataforma e, incluso, la popularidad de sus creadores. Además, es probable que el tamaño de la muestra haya afectado el rendimiento de la red en tanto pudo requerirse de un mayor numero de imágenes para que su entrenamiento permita alcanzar un mayor ajuste fuera de muestra. No obstante, si se busca un balance entre la desviación de las predicciones frente a los datos reales y el sesgo que pueda presentar cada una de estas, las *CNN* son la mejor opción dentro de los modelos analizados.

Referencias

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>

Chen, Steven Chou, Edward Yang, Richard. (2018). The Price is Right: Predicting Prices with Product Images.

Diebold, F.X. (2017), Forecasting, Department of Economics, University of Pennsylvania, <http://www.ssc.upenn.edu/fdiebold/Textbooks.html>.

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>

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/>

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>

O'Shea, K., Nash, R. (2015). An Introduction to Convolutional Neural Networks. ArXiv, abs/1511.08458.

Reynolds, Anh (2019). Convolutional Neural Networks (CNNs). Recuperado de: <https://anhreynolds.com/blogs/cnn.html>

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.

Worth, T. (2020). Painting2Auction: Art Price Prediction with a Siamese CNN and LSTM. De Haan, A. (2021). NFT art collection 2021. Kaggle. URL: <https://www.kaggle.com/datasets/vepnar/nft-art-dataset>