



THE UNIVERSITY OF CHICAGO

DEEP LEARNING FOR DYNAMIC NFT VALUATION

By
Mingxuan He

May 2024

A paper submitted in partial fulfillment of the requirements for
the Master of Arts degree in the Master of Arts in
Computational Social Science

Faculty Advisor: Dr. Anthony Lee Zhang
Preceptor: Dr. Gina Pieters

Abstract

I study the price dynamics of non-fungible tokens (NFTs) and propose a deep learning framework for dynamic valuation of NFTs. I use data from the Ethereum blockchain and OpenSea to train a deep learning model on historical trades, market trends, and traits/rarity features of Bored Ape Yacht Club NFTs. After hyperparameter tuning, the model is able to predict the price of NFTs with high accuracy. I propose an application framework for this model using zero-knowledge machine learning (zkML) and discuss its potential use cases in the context of decentralized finance (DeFi) applications.

Keywords: NFTs; deep learning; onchain data; zero-knowledge machine learning

1 Introduction

A non-fungible token (NFT) is a digital asset stored on a blockchain. The uniqueness of each NFT asset is verifiable through the unique identification assigned by the blockchain on which they exist. NFTs can represent photos, videos, audio, and other types of digital files. NFTs are most commonly bought and sold with cryptocurrencies via digital marketplaces, with the Ethereum blockchain being the most popular market.

Recently, picture NFTs have emerged as a popular investment vehicle. For example, the NFT of a digital artwork by Beeple was sold for \$69 million in March 2021. The NFT market has grown rapidly since 2019, with the total trading volume exceeding \$12 billion in Q1 2022 alone.

Similar to their fungible counterpart (cryptocurrencies), a picture NFT can vary from a few dollars to millions of dollars. The price of a picture NFT is determined by a variety of factors, most notably the possession of “rare traits”. However, similar to the traditional arts market, most NFT prices are also heavily dependent on market taste and sentiment, such as the current popularity of the collection and/or the artist.

In this paper, I apply machine learning methods to value the price of picture NFTs. Specifically, I train models on market prices and rarity data obtained from the Ethereum blockchain. The goal is to establish a model framework that provides accurate estimates of NFT prices. Such a model has various applications, including but not limited to onchain trading and bidding contracts, risk management for crypto portfolios, and fair valuation of staked/collateralized NFTs for decentralized finance (DeFi) applications. In addition, the model can be used to price large batches of NFTs quickly, which is useful for AI-generated NFTs.

The rest of this paper is structured as follows. Section 2 discusses the existing literature on this topic and the novelty of this paper. Section 3 provides an overview of the data used in this project. Section 4 outlines the general methodology and the machine learning models used in training. Section 5 presents the results. Section 7 draws conclusions.

2 Literature Review

The fast-growing market of NFTs has attracted interest from various academic disciplines. In the economics and finance literature, various authors have discussed the price dynamics of the NFT market. A comprehensive work by Nadini et al. (2021) mapped out important statistical features of NFT markets, revealing that mean prices, sales per asset, sales per collection all follow power-law distributions. Ante (2022) found that NFT sales are triggered by price shocks in Bitcoin (BTC) and active NFT wallets are reduced by price shocks in Ether (ETH). Similarly, Dowling (2022) found co-movement between cryptocurrencies and NFTs through wavelet coherence analysis. On the buyer side, Kong and Lin (2021) and Oh et al. (2022) found that well-connected and experienced investors generally earn higher returns on NFT trades. von Wachter et al. (2022) estimated that from January 2018 to November 2021, nearly all collections were contaminated by suspicious transactions related to NFT wash trading. These results support the general belief that the NFT market is not fully efficient, and there is room for improvement in pricing NFTs.

The literature agrees on that rarity is one of, if not the most important factors in determining the price of NFTs. Using data from the CryptoPunks collection, Kong and Lin (2021) built a hedonic regression model and highlighted rarity as a key determinant of price premium in the cross-section. Mekacher et al. (2022a) used a custom-built rarity score and showed that rarer NFTs sell for higher prices, are traded less frequently, guarantee higher returns, and are less risky. Taking a network clustering approach, Nadini et al. (2021) extracted vector representations of the visual features of NFTs and analyzed their cosine distance network using the pre-trained convolutional neural network AlexNet. These results inspired the inclusion of trait features and the OpenRarity rarity rank as a key set of features in my model.

In the applied machine learning literature, there have been notable attempts to train predictive models on visual features of NFTs. Nadini et al. (2021) built a linear regression model with extracted principal components of the NFT's image, features of the trading network, and market history. Seyhan and Sefer (2023) and Costa et al. (2023) used multi-modal data including NFT images and text characteristics while excluding financial features in an effort to predict both primary and secondary sale prices. The literature also identified other features such as search trends (Jain et al., 2022; Kaneko, 2021), social media influence

e.g. Twitter (Kapoor et al., 2022). My proposed framework contributes to the literature by incorporating a large set of tabular features and build a deep learning model tailored to each collection. This approach is more scalable and can be applied to a wide range of collections for secondary sale price prediction and valuation, especially in fully decentralized applications where image data is costly to store and transfer onchain.

3 Background & Data

3.1 Institutional background

The NFT market is an emerging market with unique aspects. In this section, I provide a brief overview of the background of NFT markets, including relevant terminology.

- **NFT Marketplaces:** NFT Marketplaces are platforms built with automated blockchain contracts (known as “smart contracts”) to facilitate transactions using cryptocurrencies. Most popular marketplaces include OpenSea, Blur, Rarible, SuperRare, etc.
- **NFT Collections:** Many NFTs are created in collections, in which all NFTs share a common theme. For example, the CryptoPunks collection consists of 10,000 unique pictures of pixelated faces.
- **Traits:** Each NFT in the collection has a unique combination of traits, including color, background, face/expression, and accessories. The traits typically differ by rarity. For example, there are more than 2,000 CryptoPunks with the trait “Earring”, but only 44 with the trait “Beanie”.
- **Rarity Score & Rarity Rank:** The current industry standard for calculating the rarity of an NFT within a collection is the OpenRarity Standard ¹, where the rarity of an NFT is evaluated on the rarity of its traits. The calculations for this metric is outlined in Appendix 7. By comparing the rarity score of all NFTs in a collection, we can rank the rarity of each NFT, with 1 being the rarest.

In this paper, I use the Bored Ape Yacht Club (BAYC) collection as an example. The Bored Ape Yacht Club is a well-known collection of 10,000 unique Bored Ape NFTs. The collection was created by Yuga Labs in April 2021. BAYC has a total trading volume of 1.4 million ETH (\$2.2 billion USD) as of December 2023. BAYC is favored by many celebrities and has a large community of collectors.

¹<https://www.openrarity.dev/>

3.2 Data sources

I obtained data from two sources: Dune Analytics and OpenSea.

3.2.1 Dune Analytics

Dune Analytics is a platform for querying public databases from the Ethereum blockchain. For each collection, I query the NFT trades database and gather information on all NFT transactions from issuance to Sep 30, 2023.

3.2.2 OpenSea

OpenSea is the largest NFT marketplace focusing on NFTs based on Ethereum and Ethereum's Layer-2 ecosystem. The data from OpenSea's public API include data on NFT collections. In particular, I extracted data on the traits and rarity of each unique NFT in the collections. The set of available traits varies by collection, but generally include color, background, face/expression, and accessories.

3.3 Data description

As of September 30, 2023, there has been 43,383 BAYC trades recorded in the Dune `nft.trades` database. It is worth noting that trades conducted outside of marketplaces, such as through private over-the-counter agreements, are not recorded. Here I outline the features and target variables used in this project. The features are divided into three categories: market, traits & rarity, and last trade. The features are described in Table 1.

Table 1: Feature Columns Description (Bored Ape Yacht Club)

Feature Category	Feature	Description
Market	volume_eth_lag1	daily market volume (ETH)
	price_p5_eth_lag1	daily 5-percentile price (ETH)
	price_max_eth_lag1	daily highest price (ETH)
	price_min_eth_lag1	daily minimum (floor) price (ETH)
Traits & Rarity	rarity_rank	rarity rank measured by OpenRarity
	Background_count	number of items with the same <i>Background</i> trait
	Mouth_count	number of items with the same <i>Mouth</i> trait
	Eyes_count	number of items with the same <i>Eyes</i> trait
	Fur_count	number of items with the same <i>Fur</i> trait
	Clothes_count	number of items with the same <i>Clothes</i> trait
	Earring_count	number of items with the same <i>Earring</i> trait
Last Trade	last_trade_timediff	time since the last time this item was traded
	last_trade_price	price at the last time this item was traded

Note that all market data (volume and prices) are at daily frequency and lagged by 1 day. The traits and rarity data are static and do not change over time. The last trade data is dynamic and changes with each trade. For each trait, I use the count of items in the collection which shares the same trait (e.g. `Fur="tan"` corresponds to `Fur_count=626`), with lower counts signaling higher rarity of the specific trait. The outcome variable is the trade price of the NFT (in ETH).

3.4 Exploratory analysis

3.4.1 Price history

Figure 1 plots the price history of all BAYC trades, with each entry representing an individual trade. Since the BAYC collection was launched in 2021, its price has increased significantly with a peak in mid 2022, matching the bullish sentiment in the NFT and broader blockchain economy. Since then, the market price has been relatively stable, with a slight downward trend in average price.

Figure 1: Trade Price by Date

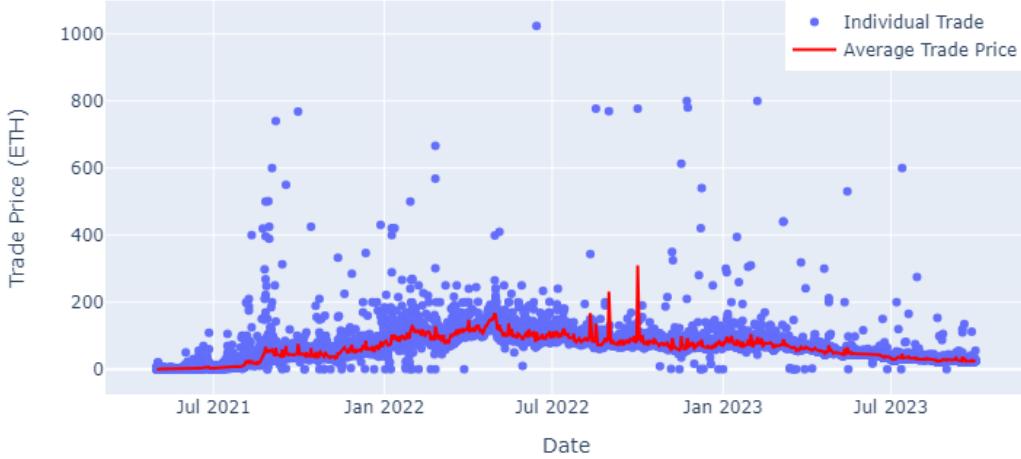
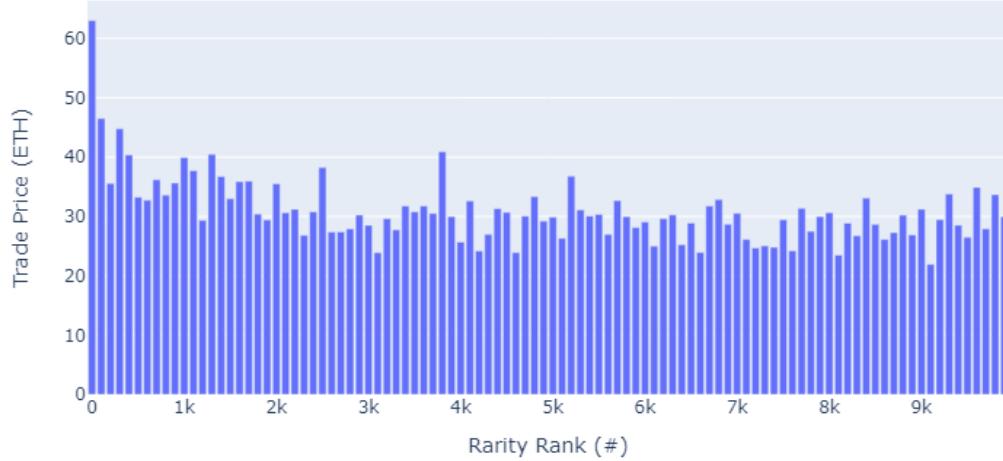


Figure 2 plots the average trade price (in ETH) of BAYCs grouped by their rarity rank. The distribution shows a high rarity premium for the rarest NFTs in the collection (top 100) with an average trading price above 60 ETH, almost twice as much as the collection average. There's also a significant premium for the top 2,000 rarest NFTs. However, the relationship between rarity and price is less clear for NFTs below the 2,000 rank threshold. This suggests that factors other than rarity, such as feature aesthetics and market sentiment, play a significant role in determining the price of the majority of BAYC NFTs.

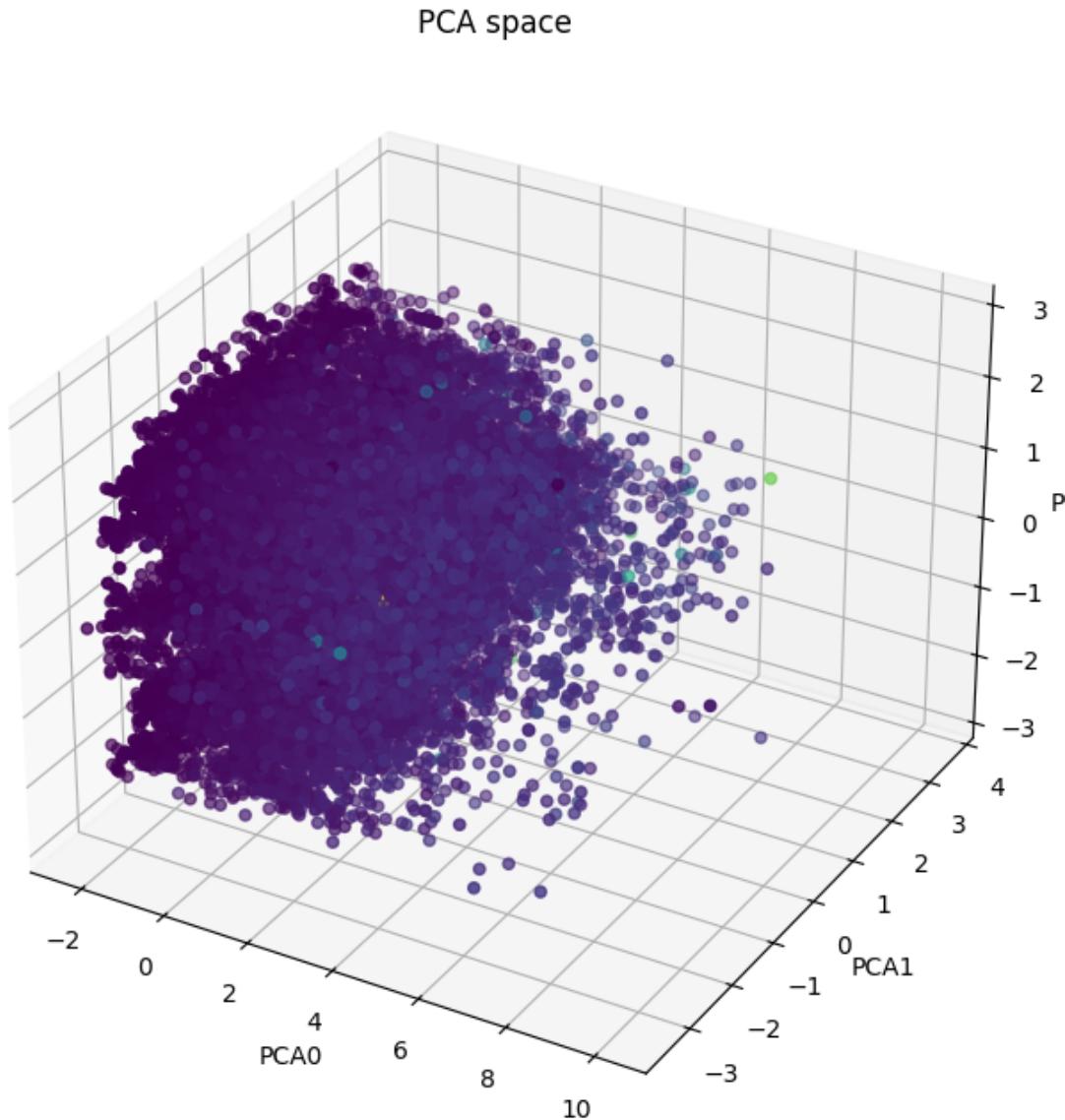
Figure 2: Trade Price by Rarity Rank



3.4.2 Principal component analysis

Figure 3 plots the first three principal components of features listed in table 1. The level of clustering indicates no clear separation between price levels in the linear space, which motivates the use of non-linear models such as tree-based models and deep neural networks.

Figure 3: PCA(3) plot of NFT prices (Bored Ape Yacht Club)



4 Methodology

4.1 Feature selection and benchmark models

To find the set of features with the highest predictive power, I define three feature sets and examine the predictive outcome use linear and tree-based models.

Approximately 20% of the trades dataset feature first-time trades, which lack the last_trade

features. This implies a skewed distribution of trading frequency among BAYCs. Therefore one major goal of this step is to examine the choice of including the last_trade variables at the cost of dropping that portion of training data.

I define three feature sets, each containing a combination of feature categories:

- Set 1: market, traits & rarity, last trade
- Set 2: market, traits & rarity
- Set 3: traits & rarity, last trade

For each feature set, I fit three types of linear models (OLS, Lasso, Ridge) and two types of tree-based ensemble models (random forest, gradient boosting), selecting the feature set that yields the best outcome metric.

As a robustness check, I use the same model-feature set pairs to predict an alternative outcome variable: the log-transformed trade price of each trade.

4.2 Deep learning models

I define a class of convolutional neural network (CNN) and use hyperparameter tuning to find the architecture that results in minimum loss. The choice for CNN is motivated by the cross-sectional nature of the data. Convolutional layers are designed to capture high-order interactions between the traits features and market movements. The CNN model is tuned using random search over the hyperparameters described in Section 5.2.2.

5 Results

5.1 Benchmarking

Benchmarking results using linear and tree-based models are reported in Table 2

Table 2: MSE Loss Metric for Benchmark Models and Explanatory Variable Sets

(a) Y = Trade price (ETH)			
Model	X1	X2	X3
OLS	353.02	319.01	762.74
Lasso	374.11	327.36	772.64
Ridge	353.02	319.01	762.74
Random Forest	154.97	200.96	254.85
Gradient Boosting	143.20	200.50	262.19

(b) Y = log (Trade price (ETH))			
Model	X1	X2	X3
OLS	0.97	1.63	1.52
Lasso	2.07	2.73	2.73
Ridge	0.97	1.63	1.52
Random Forest	0.20	0.43	0.27
Gradient Boosting	0.22	0.43	0.28

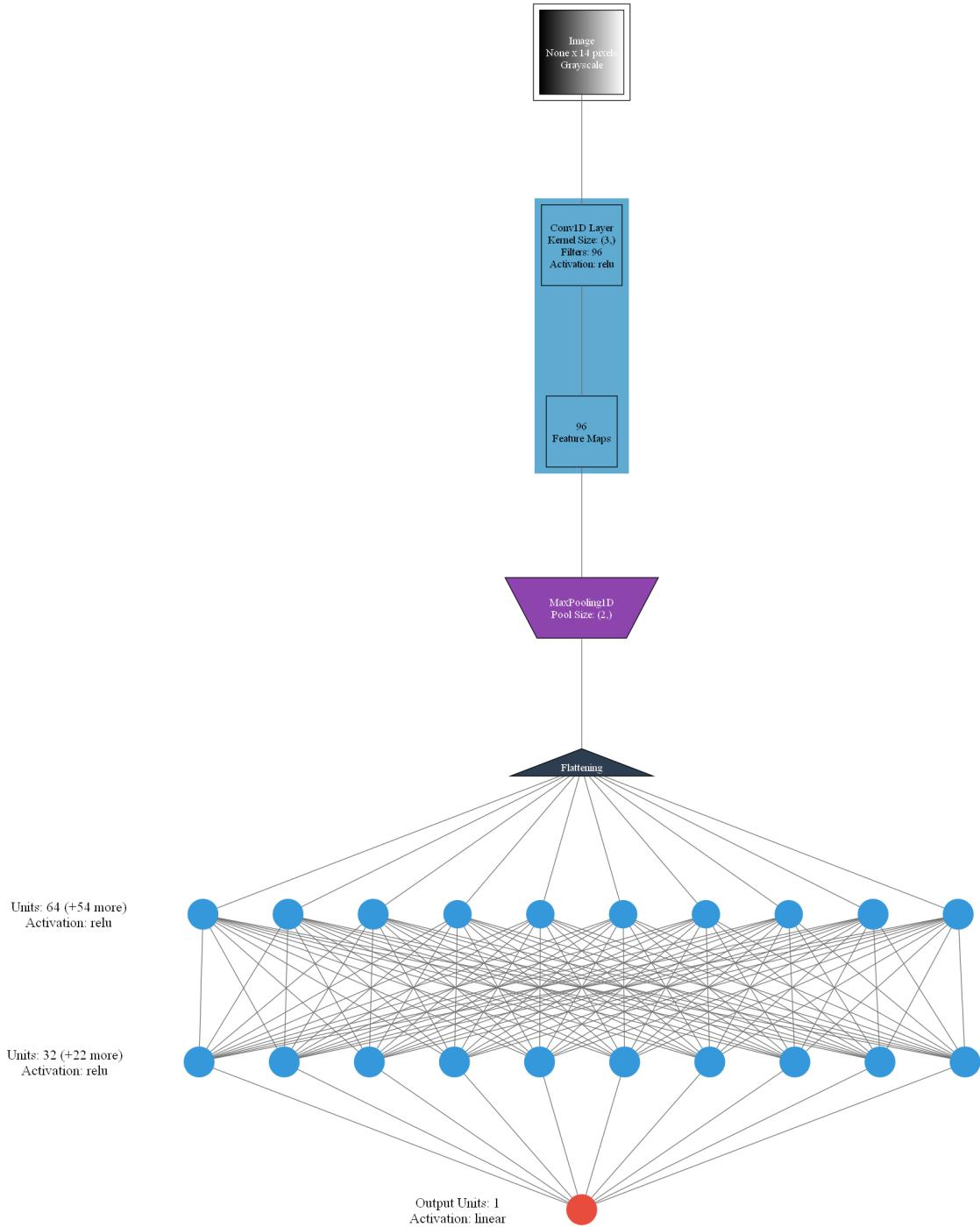
(c) Y = Trade price (USD)			
Model	X1	X2	X3
OLS	5.01e+09	4.37e+09	8.29e+09
Lasso	5.01e+09	4.37e+09	8.29e+09
Ridge	5.01e+09	4.37e+09	8.29e+09
Random Forest	1.29e+09	1.27e+09	2.91e+09
Gradient Boosting	2.03e+09	2.29e+09	3.23e+09

5.2 Deep Learning Models

5.2.1 Model architecture

The deep learning models are trained using the `tensorflow.keras` package in Python 3.9 under an `WSL-2 Ubuntu` operating system. Figure 4 illustrates the model architecture, without the optional dropout layer.

Figure 4: Example CNN Architecture. The actual architecture is tuned using `keras_tuner` package.



5.2.2 Hyperparameter tuning

The tunable hyperparameters include: number of filters and kernel size in the convolutional layer, number of units in the first dense layer, a boolean to include the dropout layer, and the learning rate for Adam optimizer. The feature importance and correlation metrics are computed post-hoc using the weights & biases package.

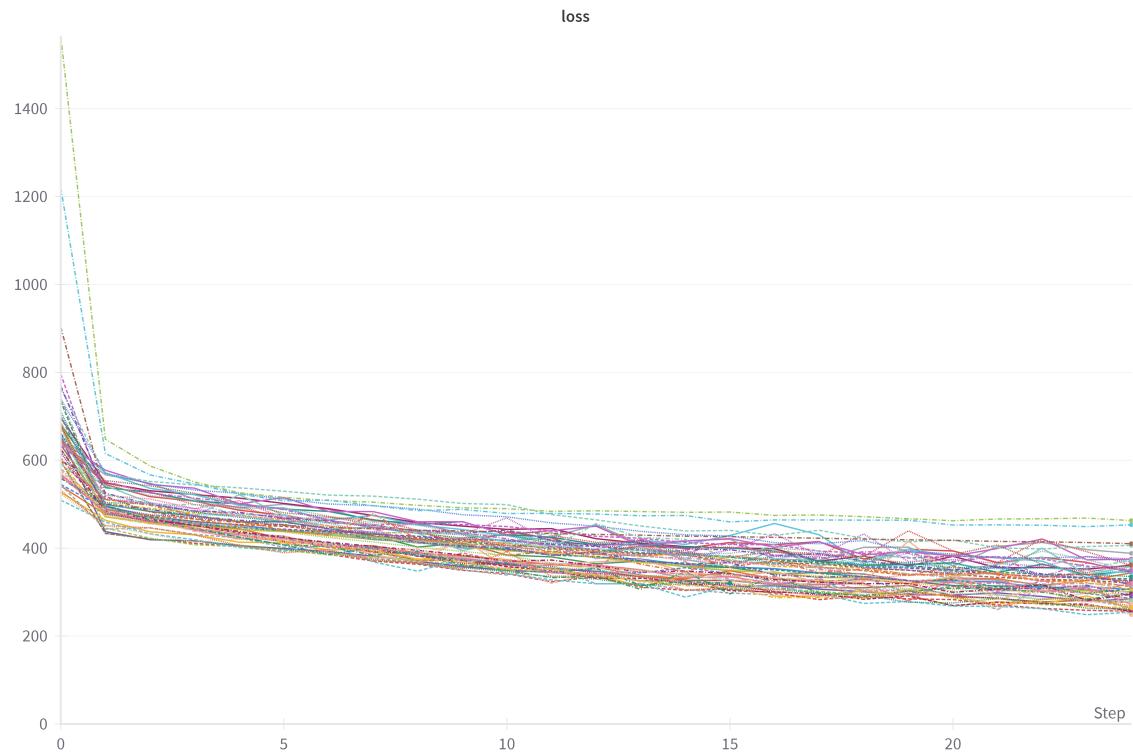
Table 3: List of Hyperparameters and Importance with Respect to Validation Loss.

Config Parameter	Importance	Correlation
kernel size	.427	-.102
learning rate	.356	.055
filters	.100	.008
units	.087	-.091
dropout	.030	-.098

I tune these hyperparameters using the `RandomSearch` method in `keras_tuner` package. During training, the number of epochs is capped at 25, with a batch size of 32. The training and validation loss by epoch is shown in Figure 5. The searched hyperparameters and resulting validation loss (MSE) are shown in Figure 6.

Figure 5: CNN Training and Validation Loss by Epoch. Each line represents a trial with different hyperparameters

(a) Training loss by epoch



(b) Validation loss by epoch

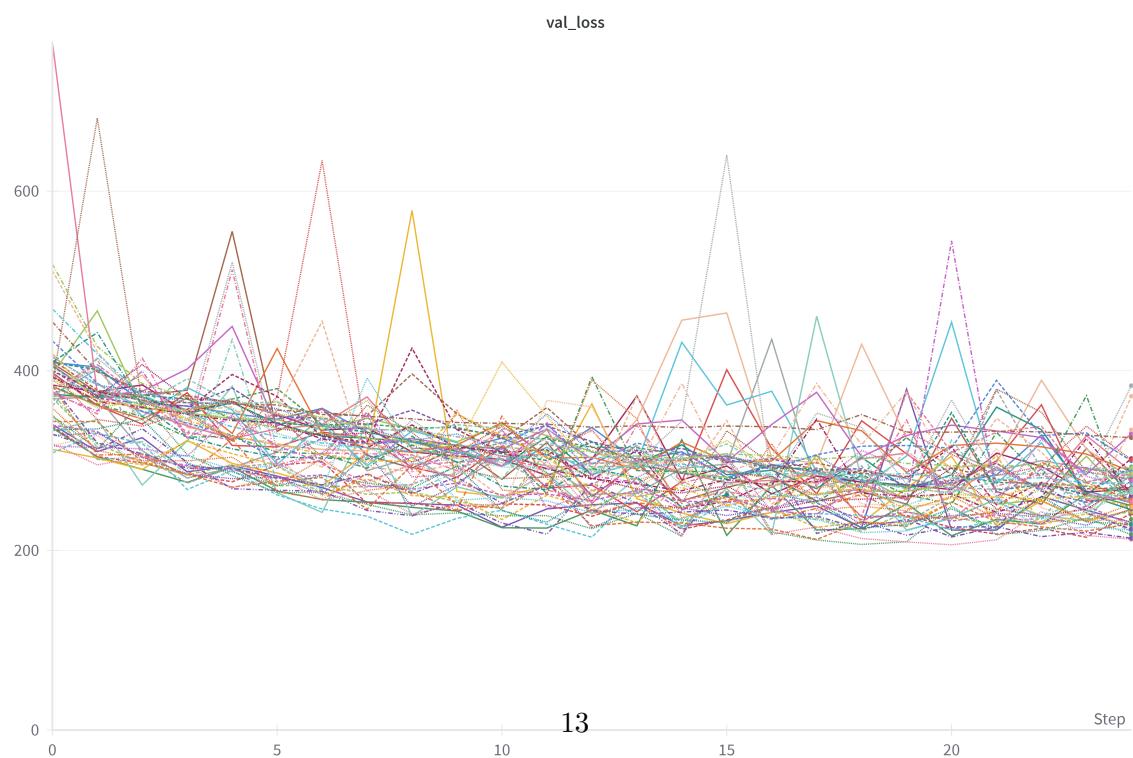
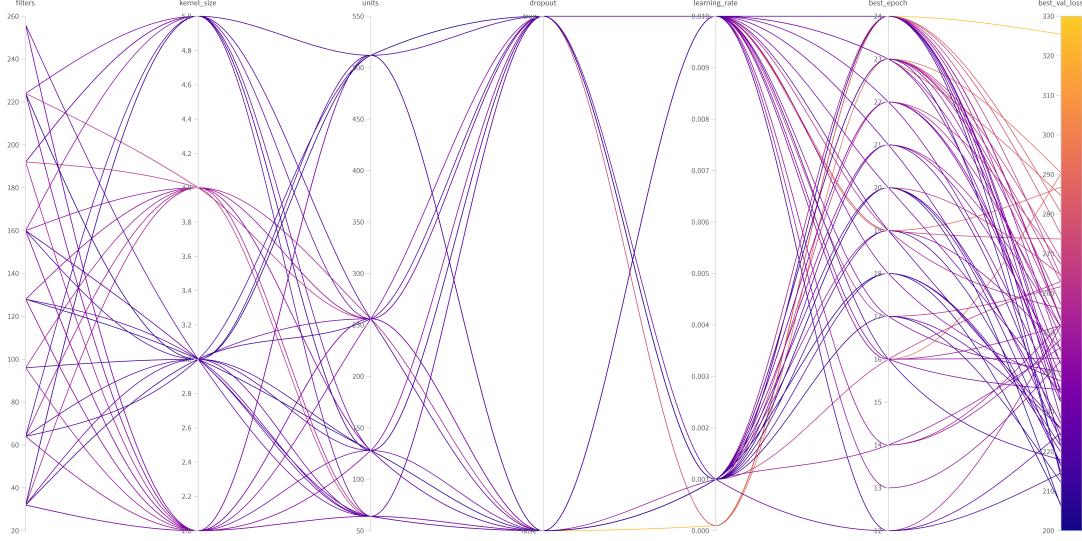


Figure 6: CNN Hyperparameters Searched using `keras_tuner.RandomSearch` method



6 Application

In this section, I propose an application framework for this model using zero-knowledge machine learning (zkML) and discuss its potential use cases in the context of decentralized applications (DApps).

The goal of the framework is to provide a dynamic and trustless valuation model for NFTs while can be easily integrated with onchain NFT marketplaces and decentralized finance (DeFi) applications.

6.1 Dynamic valuation

Dynamic valuation empowers the use of NFTs as collateral for DeFi applications, a recent trend in the NFT market. Borrowing against NFTs is currently limited and risky due to the highly volatile market (e.g. “pump-and-dumps”) and the lack of a reliable valuation model for volatile NFT prices. Most existing implementations rely on human input such as user-defined offers, direct negotiations, and Dutch auctions. The incorporation of a reliable valuation model can greatly improve the efficiency of these applications, and allows for dynamic adjustment of borrowing rates and auto-liquidations to protect lenders.

Valuation for an NFT is conducted as follows. For each NFT in a collection (e.g. BAYC #8817), a collection-specific model is able to output a real-time valuation in ETH based on 1. the last trade price and time since last trade 2. the traits and rarity of the NFT 3. the current market conditions (e.g. volume, price floor, etc.) Due to the dynamic nature of

the model, the valuation of each NFT can be updated in real-time as new trades occur and market conditions change.

6.2 Trustless valuation

In the traditional banking and loan sector, valuation models are considered top level enterprise secrets and are never published. Hence all users must rely on trust in the enterprise to fairly value their collateralized property, such as a house.

In decentralized applications, however, such level of trust is difficult to establish. This proposed framework solves the issue of trustless valuation by using zero-knowledge machine learning. zkML allows for verifiable proof that a price valuation is indeed calculated by the model, instead of e.g. a manipulated price oracle. Meanwhile, the proof does not reveal the parameters of the model itself. This solves the problem of trustless valuation while preventing exploitation. Recent works by Liu et al. (2021) and Zhang et al. (2020) have demonstrated the feasibility of zkML in ensemble and deep learning models.

In addition, the design choice of using tabular data for all model inputs generates great benefits for fully onchain applications, since storing and transferring image data onchain is costly.

6.3 Technical implementation

The technical implementation of this framework consists of three components: a data pipeline, a family of pre-trained prediction models, and a smart contract. The data pipeline is responsible for collecting and processing data from the Ethereum blockchain. The prediction model is responsible for generating the real-time valuation of each NFT. The smart contract is responsible for emitting the valuation of each NFT and providing zero-knowledge proofs onchain.

7 Conclusion

This research provides insights into valuation of Non-Fungible Tokens (NFTs), particularly in the context of dynamic markets. I provide an innovative approach to the dynamic valuation of NFTs, using deep learning models and a rich cross-sectional dataset from the Ethereum blockchain and OpenSea. Using the Bored Ape Yacht Club NFTs as a demonstration, the model integrates historical trades, market trends, and traits/rarity features to achieve high accuracy in price prediction.

In real-world applications, the incorporation of zero-knowledge machine learning (zkML) can further enhances the model's utility, opening up new possibilities in decentralized finance

(DeFi) applications. This work not only contributes to the field of digital asset valuation but also sets a foundation for future research in dynamic and trustless NFT pricing within the evolving landscape of blockchain technologies and decentralized markets.

Data and Code Availability Statement

The code and data used in this project are available at <https://github.com/mingxuan-he/NFT-pred>. The Dune Analytics query can be viewed at <https://dune.com/queries/3124232>.

References

- Alessandretti, L., ElBahrawy, A., Aiello, L. M., & Baronchelli, A. (2018). Anticipating cryptocurrency prices using machine learning. *Complexity, 2018*, 1–16.
- Ante, L. (2022). The non-fungible token (nft) market and its relationship with bitcoin and ethereum. *FinTech, 1*(3), 216–224.
- Branny, J., Dornberger, R., & Hanne, T. (2022). Non-fungible token price prediction with multivariate lstm neural networks. *2022 9th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, 56–61.
- Costa, D., La Cava, L., & Tagarelli, A. (2023). Show me your nft and i tell you how it will perform: Multimodal representation learning for nft selling price prediction. *Proceedings of the ACM Web Conference 2023*, 1875–1885.
- Dowling, M. (2022). Is non-fungible token pricing driven by cryptocurrencies? *Finance Research Letters, 44*, 102097.
- Henriques, I., & Sadorsky, P. (2023). Forecasting nft coin prices using machine learning: Insights into feature significance and portfolio strategies. *Global Finance Journal, 100904*.
- Horky, F., Rachel, C., & Fidrmuc, J. (2022). Price determinants of non-fungible tokens in the digital art market. *Finance Research Letters, 48*, 103007.
- Jain, S., Bruckmann, C., & McDougall, C. (2022). Nft appraisal prediction: Utilizing search trends, public market data, linear regression and recurrent neural networks. *arXiv preprint arXiv:2204.12932*.
- Kaneko, Y. (2021). A time-series analysis of how google trends searches affect cryptocurrency prices for decentralized finance and non-fungible tokens. *2021 International Conference on Data Mining Workshops (ICDMW)*, 222–227.
- Kapoor, A., Guhathakurta, D., Mathur, M., Yadav, R., Gupta, M., & Kumaraguru, P. (2022). Tweetboost: Influence of social media on nft valuation. *Companion Proceedings of the Web Conference 2022*, 621–629.
- Kong, D.-R., & Lin, T.-C. (2021). Alternative investments in the fintech era: The risk and return of non-fungible token (nft). *Available at SSRN 3914085*.
- Liu, T., Xie, X., & Zhang, Y. (2021). Zkcnn: Zero knowledge proofs for convolutional neural network predictions and accuracy. *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, 2968–2985.
- Loi, F. L., Tang, K., & You, Y. (2023). Cultural price premium: Evidence from cryptopunks. *Available at SSRN 4202168*.

- Mekacher, A., Bracci, A., Nadini, M., Martino, M., Alessandretti, L., Aiello, L. M., & Baronchelli, A. (2022a). Heterogeneous rarity patterns drive price dynamics in nft collections. *Scientific reports*, 12(1), 13890.
- Mekacher, A., Bracci, A., Nadini, M., Martino, M., Alessandretti, L., Aiello, L. M., & Baronchelli, A. (2022b). How rarity shapes the nft market. *arXiv preprint arXiv:2204.10243*, 9.
- Nadini, M., Alessandretti, L., Di Giacinto, F., Martino, M., Aiello, L. M., & Baronchelli, A. (2021). Mapping the nft revolution: Market trends, trade networks, and visual features. *Scientific reports*, 11(1), 20902.
- Oh, S., Rosen, S., & Zhang, A. L. (2022). Investor experience matters: Evidence from generative art collections on the blockchain. *Available at SSRN 4042901*.
- Seyhan, B., & Sefer, E. (2023). Nft primary sale price and secondary sale prediction via deep learning. *4th ACM International Conference on AI in Finance*, 116–123.
- von Wachter, V., Jensen, J. R., Regner, F., & Ross, O. (2022). Nft wash trading: Quantifying suspicious behaviour in nft markets. *International Conference on Financial Cryptography and Data Security*, 299–311.
- Zhang, J., Fang, Z., Zhang, Y., & Song, D. (2020). Zero knowledge proofs for decision tree predictions and accuracy. *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 2039–2053.

Appendix

Rarity Score and Rarity Rank

I use OpenSea's OpenRarity standard to calculate the rarity of each NFT within its collection. The rarity score is defined as follows:

For an NFT x with traits $i \dots n$, its rarity score is

$$R(x) = \frac{I(x)}{\mathbb{E}[I(x)]}, \text{ where } I(x) = \sum_{i=1}^n -\log_2 P(\text{trait}_i).$$