UNIVERSITY OF WARSAW

Faculty of Economic Sciences

WORKING PAPERS
No. 27/2022 (403)

WHAT MAKES PUNKS WORTHY? VALUATION OF NON-FUNGIBLE TOKENS BASED ON THE CRYPTOPUNKS COLLECTION USING THE HEDONIC PRICING METHOD.

EWELINA PLACHIMOWICZ
PIOTR WÓJCIK

WARSAW 2022

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

What makes Punks worthy? Valuation of Non-Fungible Tokens based on the CryptoPunks collection using the hedonic pricing method.

Ewelina Plachimowicz, Piotr Wójcik

Faculty of Economic Sciences, University of Warsaw

Corresponding author: pwojcik@wne.uw.edu.pl

Abstract: This article focuses on an attempt to value Non-Fungible Tokens from the CryptoPunks collection. Based on the data from January 2021 to July 2021, a hedonic pricing model was built, based on the transaction history and characteristics of a given NFT, as well as external markets variables - cryptocurrency prices (Bitcoin and Ethereum), natural gas prices and the popularity of the collection in social media (Twitter). According to the literature, we decided to build three regression models: Ordinary Least Squares model, XGBoost algorithm and bidirectional Long Short-Term Memory Model. Based on the results, we were able to prove that such complex issues as NFT valuation require more advanced methods than the classical regression model. In addition, we proved that one of the most important categories of variables in the case of NFT valuation is the history of token sales and its characteristics, indicating a particular rarity. Moreover, we have shown that the cryptocurrency and natural gas market is not an important factor in the NFT valuation. Finally, we proved that the increase in the popularity of tokens in social media translates into an increase in NFT prices, and this is an important element when trying to valuate tokens.

Keywords: NFT, Non-Fungible Token, pricing, valuation, luxury goods, cryptocurrencies

JEL codes: Z11, G14, G12

Introduction

Non-Fungible Tokens (NFTs) have become one of the most popular topics on the fintech market in the last two years. NFT is a type of cryptographic token built in the blockchain ecosystem. Unlike cryptocurrencies, NFTs are indivisible and non-interchangeable - there is no conversion factor to how much NFT can be worth, as is the case with cryptocurrencies or regular currencies. Due to its structure, each NFT is unique and can only be owned by one person at a time. Additionally, the transfer between buyer and seller is easy and secured, and the transaction history of a given NFT is easy to trace. They are sold using cryptocurrencies, on specially created marketplaces, such as OpenSea, AtomicMarket, Rarible or Mintable. The first mentions of NFT appeared in 2012, but they only gained popularity in 2017 thanks to the game CryptoKitties. The peak of interest in NFTs was reached around 2021 due to the staggering sums that collectors were able to pay for them. The most expensive tokens include The Merge ($ 91.8 million), The First 5000 Days ($ 69.3 million), Clock ($ 52.7 million), or avatars available in the CryptoPunks collection (prices ranging from several hundred thousand dollars to even 23 million dollars). NFT come in all kinds of guises - art, video, games, metaverse, music, and collectibles like Twitter CEO's first tweet. However, NFT is not only about virtual storage on your computer - buying some of them, for example from The Bored Ape Yacht Club collection, will grant you online community membership, which gives access to private chat rooms or social events. In the last two years, many celebrities, athletes, sports teams or business people have become involved in the NFT market, which has translated into a drastic increase in the popularity of tokens. According to the 'NFT Quarterly Report - Q1 2021', the turnover on the NFT market in the first quarter of 2021 was over $ 2 billion, which, compared to the first quarter of 2020, is an increase of 13,118% (in the first quarter of 2020, the turnover was over $ 15 million).

In this article, we will analyze one of the most popular NFT collections - CryptoPunks. Created in 2017, the collection focuses on punk-style avatars. CryptoPunks, along with collections such as CryptoKitties or The Bored Ape Yacht Club, is associated with the explosion of popularity on the NFT in 2021. Due to its popularity and exclusivity, it is one of the most expensive collections in existence.

Due to their young age and characteristics, NFTs are particularly difficult to value. If you look at the price distribution that NFTs achieve, they have a particularly long right tail - token prices start from a dozen or so dollars and end at tens of millions of dollars. Of course, there are some basic aspects to consider when pricing an NFT: market demand, investment potential, the popularity of the creator, the uniqueness of a given NFT, usability, belonging to a given collection and ownership history. However, NFTs do not have any upper or lower price limit - their value is based on the personal perception of buyers and creators. Even if we rely on the history of a given NFT, it is difficult to indicate whether the next buyer will be willing to pay the appropriate amount for the token. For the NFT, no Generally Accepted Accounting Principles (GAAP) standards have been developed that would be able to indicate the best digital asset valuation model. Of course, we can rely on similar models as in the case of art, luxury goods, or cryptocurrencies, but NFTs require completely new metrics. Due to the payment system, NFTs are strongly associated with cryptocurrencies (especially Bitcoin and Ethereum), which makes them extremely unstable - suffice to mention that in 2019 Bitcoin's price was around $ 20,000, at the end of 2021 it increased to about $ 245,000, and in in mid-June 2022, it fell to $ 100,000. Moreover, the cryptocurrency market, and hence the NFT, is linked to gas prices, which are currently showing a dramatic increase, inter alia, due to the war in Ukraine. Finally, in many countries there are still discussions on legal and tax issues in the case of NFT and cryptocurrencies. The results of these talks may significantly change both markets, which will ultimately lead to drastic changes in valuations.

The main aim of this article is an attempt to value NFTs from the CryptoPunks collection using classic econometric methods, as well as machine learning and deep learning models. In this work, we want to focus on popular and well-known regression methods from various levels of advancement. Based on the available literature, we will try to build a hedonic pricing method for the valuation of NFT prices. Various regression models will be applied, starting from traditional Ordinary Least Squares (OLS) regression (benchmark), to non-linear machine learning algorithms including XGBoost and bidirectional Long Short-Term Memory (LSTM). Additional Explainable Artificial Intelligence (XAI) tools will be used to understand the relationships. We will use Permutation Feature Importance (PFI) to identify the most important variables. Additionally, to unhide the shape of identified relationships we will apply Partial Dependence Plots (PDP).

There are several innovative elements in this article. It focuses on a completely new topic in the financial world for which many scientific studies have not yet been written. What is more, the token valuation attempt will be made on a complex and unique dataset containing not only the history of NFT sales and its characteristics, but also information from related markets and social media. In addition, valuation focuses not only on classical regression methods that often appear in scientific studies on art valuation, but also on new machine learning and deep learning algorithms. Finally, we will also explain the complex relationships between the dependent variable and the explanatory variables using the XAI methods.

There are 4 research hypotheses verified in the paper. Firstly, due to the complicated process and structure, NFT validation requires more advanced methods than basic regression models, therefore we claim that machine learning algorithms, i.e. xgboost and neural network will show better predictive performance. Secondly, we assume that the main factors influencing NFT price are its characteristics - sales history, as well as the uniqueness of a given token. Third, we state that NFT valuation accuracy improves when additional predictors that affect the market - such as the prices of cryptocurrencies and gas, are considered. The final research hypothesis is that social media play an important role in the valuation of NFT and the inclusion of social media variables (such as the number of mentions of NFT on social media) improves the predictive accuracy of the model.

The structure of the article is as follows. In the first part, the available literature on the valuation of NFT, art and luxury goods are discussed, along with the distinction of individual approaches to this type of research. In the second part, we discuss the individual components of the approach to NFT validation applied in the article and the assessment of the significance of the variables in the model. In the third part, we summarize the process of collecting and cleaning the data used in modeling. The fourth part contains a summary of empirical research, model performance, as well as the verification of the hypotheses set out at the beginning of the article. The last part of the article summarizes the paper and indicates possible extensions of the study.

1. Literature review

As mentioned earlier, NFT is a relatively new topic in the world of technology and finance. For this reason, the pool of research and literature that can be used is quite small - research on this topic has been repeated practically every month, as academic papers are just beginning to appear. However, there are some patterns that can be seen and research on similar topics can be followed. Based on the nature of NFT, it can be concluded that when valuing them, researchers may use proven methods in the valuation of art and luxury goods. What's more, research conducted on the cryptocurrency market is also important. In the case of NFT, in addition to the research method, it is important to identify individual components that can significantly affect their valuation - these are not only the features of a given NFT, but also the impact of the cryptocurrency market, gas market and social media.

Renneboog and Spaenjers (2013) used a hedonic regression analysis in their article on paintings valuation. Based on data containing information on over a million sales of paintings, they constructed OLS and repeat-sales regressions based on the characteristics of the paintings, their authors and auction houses where works of art are sold. As a result, not only the variables relating to the painting itself turned out to be important, but also the variables relating to the author and the sale itself (the auction house). Moreover, thanks to the repeat-sales regression, they proved the increase in the investment potential of art over the years. Finally, they found that measures of high-income consumer confidence and art market sentiment predict art price trends.

Horky et al. (2022) used data from one of the most popular NFT marketplaces - SuperRare. They used a similar approach as Renneboog and Spaenjers (2013) - by combining classic econometric methods with machine learning models, they considered not only information about a given NFT, such as its size or file type, but also information on gas prices, market size and the Ethereum-to-dollar ratio. To select the variables used in the model, they used machine learning methods: neural networks, decision trees and clustering. They then modelled the NFT price using LASSO regression with adjusted R squared value equal to 0.265. Finally, they obtained results proving that a hedonic approach to this type of valuation is appropriate. The model showed the importance of not only the variables related to a given NFT, but also the variables related to the NFT market and the ancillary markets. Moreover, the model proved that NFT cannot be considered a simple derivative of cryptocurrencies.

Campos and Barbosa (2009) dealt with the study of sales rates, prices, and returns in Latin American art auctions. Using the hedonic regression method, they proved that attributes such as the reputation of the author or provenance of the artwork are more important than the basic values of a painting (such as size or technique). Similar conclusions as well as the method of valuing works of art from Romania were adopted by Dinu et al. (2020).

On the other hand, in recent years, additional methods have been developed that can value assets with much higher accuracy than the classical econometric methods. Aubry et al. (2019) valued over a million of paintings using convolutional neural networks (CNNs). They found that based on the non-visual and visual aspects of the paintings, out-of-sample neural network valuations predict results much better than the same exercise performed with a standard hedonic valuation model (OLS regression). What's more, machine learning methods helped to explain the price levels and sale probabilities after conditioning on auctioneers' pre-sale estimates. Importantly, using the neural network, the authors were able to overcome the experts' systematic biases in expectations formation.

Liu (2022) made another attempt to use neural networks for the valuation of works of art. The author used the advantages of LSTM in problems with modelling time series. The analysis was based on a combination of one-way LSTM and two-way LSTM for art price prediction, taking into account the correlation between the time series. Considering the reverse dependence of the time series, the bidirectional LSTM layer is used to obtain the bidirectional time correlation between historical data. Moreover, two-way LSTM is used to correlate the potential contextual information of the historical data of the artwork price stream. According to the author, this method brought much better results than other machine learning and deep learning models used as a benchmark (Support Vector Regression, Contextual LSTM, Random Forest, Sparse Autoencoder).

Worth (2020) applied the valuation of works of art in a similar way. He combined two neural networks - CNN and LSTM. Siamese[1] CNN was used to estimate the similarity between the newly added paintings to the database and those already in it. The LSTM was used to predict the prices of works of art based on their characteristics and market data. The K-Nearest Neighbours algorithm was used as a benchmark for LSTM. The model composed of CNN and LSTM turned out to be significantly better than the model composed only of CNN, and better than the combination of CNN and KNN.

---

[1] A Siamese neural network uses the same weights while working in tandem on two different input vectors to compute comparable output vectors. It is usually used to image classification, text classification or voice classification.

A similar approach was used by Li and Liu (2021), who focused on the valuation of works of art using the Bidirectional Encoder Representations from Transformers (BERT) model for tabular data analysis and the CNN for image analysis. They proved that the valuation of works of art by means of the analysis of tabular data gives better results than the analysis of the visual aspects alone. BERT performed much better than CNN - both with and without the Siamese architecture. Moreover, they proved that the results of the analysis of images with Siamese CNN complement the results of the BERT analysis, thus giving the best results in the valuation of works of art.

Most of the more advanced articles relate to art or luxury goods valuation analysis. However, in 2021, Anyfty (NFT Bank driven by DAO[2]) published an attempt to value NFT. In their study, they used a wide dataset from OpenSea API and Binance / Uniswap API. The data was based on the characteristics of a given NFT, transaction history, market data (such as Bitcoin and Ethereum rates) and data on the popularity of NFT in social media. The authors then used several models for price prediction: Gradient Boosting, Linear Regression, Linear Regression with LASSO, ElasticNet (best model) and Catboost. The results showed that analysing the transaction history of a given NFT gives better results than analysing the set as a whole. Additionally, the authors drew attention to the fact that the division of individual NFTs into clusters could increase the accuracy of the models.

The above examples prove that if we treat NFT as a kind of art, the approach to their valuation should be relatively flexible. In addition to the method that will be used for their valuation, it is worth considering the individual components of the model. As evidenced by the above-mentioned research, the basis in this case are the characteristics of a given NFT, as well as its uniqueness, popularity, or author. Schaar and Kampakis (2022) focused on the analysis of the CryptoPunks collection as an alternative investment. Based on the hedonistic regression method, they proved that one of the main NFT price drivers is the rarity of their individual attributes, such as - Alien, Beanie, Pilot Helmet, Orange Side or Choker.

Nadini et al. (2021) analyzed over 6 million transactions of approximately 5 million individual NFTs. To estimate the prices of the first and second sales of a given NFT, they applied a linear

[2] Decentralized Autonomous Organization.

regression. They proved that the sales history is the most important aspect of the valuation, while the visual aspects of a given NFT are in second place - such as hair colour and clothing accessories.

Another important aspect in the NFT world are cryptocurrencies - due to the fact that the payment process for each token is based on them, a frequent topic of research is the relationship between NFT and cryptocurrencies. Dowling (2022) examines the hypothesis whether NFT pricing is related to the pricing of cryptocurrencies. The spillover index only indicated limited volatility transmission effects between the two financial assets. On the other hand, wavelet coherence analysis showed some co-movements between the two markets. Overall, this analysis suggests that cryptocurrency behaviour and analysis may be a benefit in terms of NFT price behaviour, however low volatility transmissions indicate that NFT can be analysed as a separate financial asset from cryptocurrencies.

On the other hand, Ante (2022), using a vector autoregressive (VAR) model based on the daily data from January 2018 to April 2021, proved that Bitcoin price shocks affect the increase in NFT sales. In addition, Ethereum's price shocks reduce the number of active NFT wallets. These results suggest that movements in the most significant cryptocurrency markets are driving changes in the NFT market, but without the opposite effect.

A similar analysis and results were obtained by Pinto-Gutierrez et al. (2022). Using VAR models, they showed that Bitcoin returns significantly contributed to the increase in NFT popularity next week (measured by the number of queries on Google). What's more, they obtained similar results using wavelet coherence analysis - Bitcoin and Ethereum returns significantly drove NFT's popularity next week. To sum up, significant spikes in cryptocurrency prices affect the popularity boom around the NFT, and thus - it can be suspected that it also has a significant impact on their sales growth.

Apart from the variables typical for NFT, it is worth considering social media. Because it was celebrities and influencers who promoted the NFT, the sentiment and the number of statements about tokens on popular social networking sites may have a lot of added value in the NFT valuation. Kapoor et al. (2022) analyzed the impact of Twitter on NFT available on OpenSea. In order to value the asset, they created several models - binary classification (XGBoost), multiclass classification (XGBoost) and image analysis models (CNN). Both the binary classification and multiclassification models achieve approximately 6 p.p. improvement in accuracy when the Twitter based variables are added, compared to the baseline models with variables based on the characteristics of a given NFT. Regardless of the architecture used in the CNN model, models based on table data (Twitter and data from OpenSea) turned out to be better than models based on image analysis. The most important variables from Twitter turned out to be the count of user membership lists, as well as number of likes and retweets. This indicates that social media can be considered a potentially significant variable in the NFT valuation.

The subject of this article is fresh and therefore not sufficiently researched. Based on the available literature, we are able to identify several research gaps that should be analyzed. First, relatively few papers contain a full attempt at NFT valuation - these are mostly articles dealing with NFT's relationships with other markets, such as social media or cryptocurrencies. We would like to conduct a full valuation test - based on the sales history, attributes of a given NFT, market data (cryptocurrencies, gas prices) and data from social media (Twitter, Reddit, Google Trends). In addition, most articles are based on basic econometric models, such as a linear regression or LASSO. We would like to extend the study with additional machine learning methods and in addition apply selected tools of XAI to better understand the results of these models. We believe that thanks to full analysis and more advanced models, we will be able to predict NFT prices with relatively high accuracy. During the analysis, we will rely on data obtained from The Blockchain Research Center (BRC) - research community established by the University of Zurich and the Humboldt University of Berlin. The data includes 40,800 NFT transactions from the popular CryptoPunks collection. Additionally, market and social media data will be attached.

2. Methodology

The main goal of our work is the valuation of NFT from the CryptoPunks collection, as well as the analysis of the relationship between the price and NFT characteristics, transaction history, market data, and social media traffic data. Our valuation approach is based on machine learning methods. Our goal is to create a hedonic pricing model - this method is often used when valuing art or real estate prices (Hill, 2013). The approach assumes that the price is determined both by the internal characteristics of the good and the external factors affecting it. In the case of real estate prices, houses features as well as external features, such as neighbourhood (e.g., crime rate) or the environment, are considered. It perfectly reflects consumers' willingness to pay for given asset and has many advantages - including the ability to estimate value based on specific consumer choices, but also flexibility in adapting to market conditions and external factors (Horky et al. - 2022, Fedderke and Li - 2020, Chanel et al. - 1996). One of the most important disadvantages of the approach are taking into account only the point of view of consumers (consumer's willingness to pay) and the lack of taking into account legal regulations, such as taxes. However, in any model, there are certain market and data limitations to consider. In the case of NFT, the legal and tax situation is still in the phase of dynamic development and discussion, so this problem is not as important as in the case of very well-developed real estate law. Based on the hedonic valuation method, several models and algorithms will be applied: OLS regression model, XGBoost algorithm, and a neural network. For all models used in the analysis (except OLS regression), we are going to perform cross-validation process to find optimal values of the hyperparameters. We will compare the results using the classic metrics for regression models - Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). To examine the importance of variables in the model, we will use the Permutation Feature Importance algorithm. Additionally, in order to obtain information about the relationship between the NFT price and independent variables in all models, we will use Partial Dependence Plots.

Taking into account the literature and approach discussed earlier, we consider various groups of factors that potentially influence the price of NFT:

- NFT Features,
- Sales History,
- Cryptocurrency Market Features,
- Natural Gas Market Features,
- Social Media Market Features.

Linear regression is one of the most popular and simple regression methods. In our work, we will focus on the Ordinary Least Squares (Hayashi, 2000), which, assuming a linear relationship between the independent variables and the dependent variable, is designed to minimize the sum of squares of residuals. The advantages of a linear regression include ease of implementation and interpretation, and relatively quick training time. OLS regression is our benchmark model, due to the fact that it is the simplest regression method, which at the same time has many disadvantages - it is sensitive to outliers, and also has a lot of strict assumptions - such as the linear relationship between variables, which is hard to achieve on real-life datasets as well as data independence (which makes it not immune to multicollinearity between variables).

Gradient boosting is an approach that sequentially uses newly constructed models to predict the residuals or errors of previous models and then adds them all to make the final predictions. XGBoost (Chen and Guestrin, 2016) is one of the most popular machine-learning models based on the gradient boosted trees algorithm. Unlike traditional models, it does not train the best possible model on the whole train data. XGBoost creates a sequence of models, each based on a different subset of the training dataset and the final prediction in regression approach is the weighted average of predictions from all models. Its most important features include parallelization, regularization (to avoiding overfitting) and the ability to detect and learn from non-linear data patterns. XGBoost is a highly flexible algorithm that deals with various datasets and is relatively fast. However, it does not do well on sparse and unstructured data. Moreover, it is very sensitive to outliers.

Traditional econometric models suffer from one problem above all - they do not consider the processed data as sequences, and they assume all inputs as independent parts. Recurrent Neural Networks (RNNs) were supposed address this issue (Rojas, 1996). However, this type of model suffers from relatively short memory - if the data sequence is long enough, they may have trouble carrying information from the previous steps to the next. RNNs suffer from a vanishing gradient problem (Bengio et al., 1994). In RNNs, gradients are used to update the weights of neural networks. During modelling, gradient shrinks as it back propagates through time. In the case of a vanishing gradient problem, the weights cannot be updated, and therefore the neural network is unable to learn. The result is a decrease in the performance of the model. Finally, Hochreiter and Schmidhuber (1997) introduced Long Short-Term Memory (LSTM) networks, an advanced type of RNN capable of learning long-term dependencies. The LSTM concept is based on cell state and various gates. Cell state is something like a memory that carries information during the processing of the model to the next steps. This information is added to or removed from cell state by gates. It allows the network to decide which information is useful to it and which is not, thus preventing information loss between steps and avoiding the long-term dependency problem. In our work, we will focus on bidirectional LSTMs (Graves and Schmidhuber, 2005). In these networks, each training sequence is performed back and forth, and the sequences are linked to one and the same output layer. This allows LSTMs to have complete information about each point in the sequence, both before and after. It is a significant improvement over conventional Recurrent Neural Networks, which are only able to analyse information before a given point in the data. Thanks to their bidirectional structure, LSTMs are especially useful in forecasting models with longer sequences of time series data. One of their biggest drawbacks is the relatively long computation time.

The data will be broken down into a training sample (80% of the data) and a test sample (20% of the most recent data). The training sample will also be used to determine the best hyperparameters for all models used in the analysis, omitting OLS regression. Due to the fact that the data we will use for the analysis is in the form of a time series, the use of a typical cross-validation (eg K-fold cross-validation) would pose several problems. First, if the data were randomized into the training part and the validation part, there would be a risk that the validation data would occur before the training data. Moreover, if the sample for validation was selected from the middle of the data, we would have a problem with data leakage and gaps in the time series (the training sample in the timeline would appear both before and after the validation sample). For this reason, we decided to use TimeSeriesSplit which generates folds across a sliding window over time. In this solution, the training set grows with successive iterations until the training set and validation set are equal to the entire sample taken for model training and tuning. Due to the time-consuming nature of the calculations, we decided to use the number of splits equal to 5. The hyperparameters for the XGBoost model will be selected based on the Randomized Search algorithm with the number of combinations equal to 1000 and the MAE metric as scoring. The Randomized Search algorithm is based on the selection of random n combinations of hyperparameters from the possibilities provided by the user (in our case, n is equal to 1000), and then the selection of the model with the best performance using cross-validation and the selected metric. In the case of neural networks, hyperparameters will be selected manually based on the analysis of averaged metric results (RMSE, MAE, MAPE) for training and validation sets.

All regression models will be compared using three metrics – MAE, RMSE and MAPE. MAE measures the mean of the absolute error values, i.e., the mean of the absolute values of the difference between the predictions and the real values. This metric ignores the direction of these differences, and the differences are equally weighted. The advantages of MAE include ease of interpretation and implementation. However, this metric is not able to indicate whether the model tends to over-estimate or under-estimate, due to the fact that the direction of differences is levelled by the absolute value. Moreover, this metric can be insensitive to large outliers. RMSE is as popular a metric as MAE and is defined as the square root of the mean of the squared errors in the model. Because errors are squared, the RMSE pays particular attention to large errors, while at the same time making small errors less important. Due to its structure, the RMSE seems to be a bit more difficult to interpret than the MAE. MAPE is defined as the sum of the individual absolute errors divided by the demand for each period separately. The biggest disadvantage of MAPE is its skewness – because each error is individually divided by the demand, high errors during low-demand periods will significantly impact this metric. On the other hand, a big advantage of MAPE is its value expressed as a percentage, and thus - easy to interpret and compare individual results between subsamples. All three metrics are negatively oriented, which means that the lower the values, the better the performance of the model. RMSE will always be greater than or equal to MAE - the greater the difference between them, the greater the variance of individual errors in the sample. Considering that the MAE, RMSE and MAPE pay attention to different aspects of errors, they are a great complement.

To better understand the results of all models – some of which are "black-box" machine learning algorithms, we will apply selected XAI (Gunning et al. – 2019, Arietta et al. – 2019) tools.

Permutation Feature Importance (Breiman – 2001, Fisher et al. – 2018) is one of the algorithms used to assess the significance of independent variables in a given model. It measures the predictive value of a given variable by examining how the prediction error increases when information from that variable is not available in the model. The values of the particular variable are randomly shuffled, then the change in prediction error of the model is measured. After repeating several times, the mean for that variable is taken. The algorithm works like this for each variable in the dataset, and then ranks the most important variables according to their influence on the model's score. Any metric can be used to compare the results - for example in case of regression, MAE, RMSE, MAPE or R2. But the technique can also be applied for classification with appropriate metrics of accuracy. It is especially useful for nonlinear or opaque estimators. The algorithm shows the relative importance of the variables. In the case of algorithms based on the assessment of the importance of a variable through a decrease in the average impurity (in tree-based models), there is a risk that a variable that does not contribute to the improvement of the model prediction will be considered important due to overfitting (because these algorithms are calculated on the training dataset). Feature Permutation Importance avoids this problem as it is calculated on unseen data. Additionally, the importance of variables based on tree-based models is heavily biased and tends to favour high cardinality features (typically numerical features) and ignore the importance of low cardinality features (categorical variables). In the case of Feature Permutation Importance, this issue does not occur. The disadvantages of this method include the calculation time, but also the poor performance in the case of multicollinearity.

The Partial Dependence Plot (Friedman 2001) shows the marginal effect that one or two of the variables in the model have on the predicted dependent variable. Thanks to this method, we can obtain an answer to the question about the nature of the relationship between variables (e.g., linear).

3. Data description and overview

The NFT data was obtained from The Blockchain Research Center. It contains 40,800 rows and 35 columns and presents information about one of the most popular NFT collections - CryptoPunks. For the analysis, however, we take into account only the lines marked with the transaction type 'successful' and with information about the value of the transaction and the time of its execution. We also decided to remove columns that do not contribute anything for analysis (like the image URL) and columns that have a significant amount of missing values (over 50%). Finally, the columns with the Ethereum price and the NFT price in USD were removed due to the fact that the entire database has a single numerical value for Ethereum. We decided to supplement the NFT price in dollars and the ETH price using financial data available on Yahoo. The outbreak of NFT's popularity dates back to the beginning of 2021, which makes the prices in 2019 and 2020 significantly different from the prices in 2021. In 2019-2020, they reached the value of several thousand dollars, while in 2021 they were from several dozen thousand to several dozen million dollars. Due to such a huge difference, we decided to analyze only the data from 2021, which is about 65% of the data after cleaning. In the end, our table with transactions consists of 6,389 rows and 15 columns. The data covers the time period from January 2021 to July 2021. This data was then decoded - each row had a *trait_value* column that contained information about the characteristics of a given NFT - such as hair color, gender, and accessories. The column was decoded into dummy variables to be able to determine the individual set of features of each NFT (0 - if it did not have the attribute, and 1 - if it did). In addition, we created variables regarding the market history of a given NFT - the number of unique previous owners at the time of the transaction (if first owner - then 0), the number of previous sales of the asset at the time of the transaction (if first transaction – then 0), the time between the current and previous sale in seconds (if first transaction – then 0), as well as the characteristics of the previous and first sale of a given NFT - the price in tokens, the token price at that time and the NFT price in dollars. If this was the first sale, the values for these columns were 0.

According to the literature, we decided to consider cryptocurrency variables. For this purpose, we used Yahoo Finance databases and obtained daily quotations of Bitcoin-USD and Ethereum-USD. For the calculations, we decided to use the closing values. Using this data, we created additional variables for our database - Bitcoin price in USD at the trade date, Ethereum price in USD at the trade date, Bitcoin to Ethereum ratio at the trade date, Bitcoin to Ethereum ratio at the first sale, and the percentage change in price week over week (the difference between the price on the day of sale and the price 7 days earlier, divided by the price from 7 days ago) and month over month (the difference between the price on the day of sale and the price 30 days earlier, divided by the price from 30 days ago) - for both Bitcoin and Ethereum. Thanks to the daily Ethereum-USD data, we were also able to calculate the price of a given NFT in dollars for a specific day.

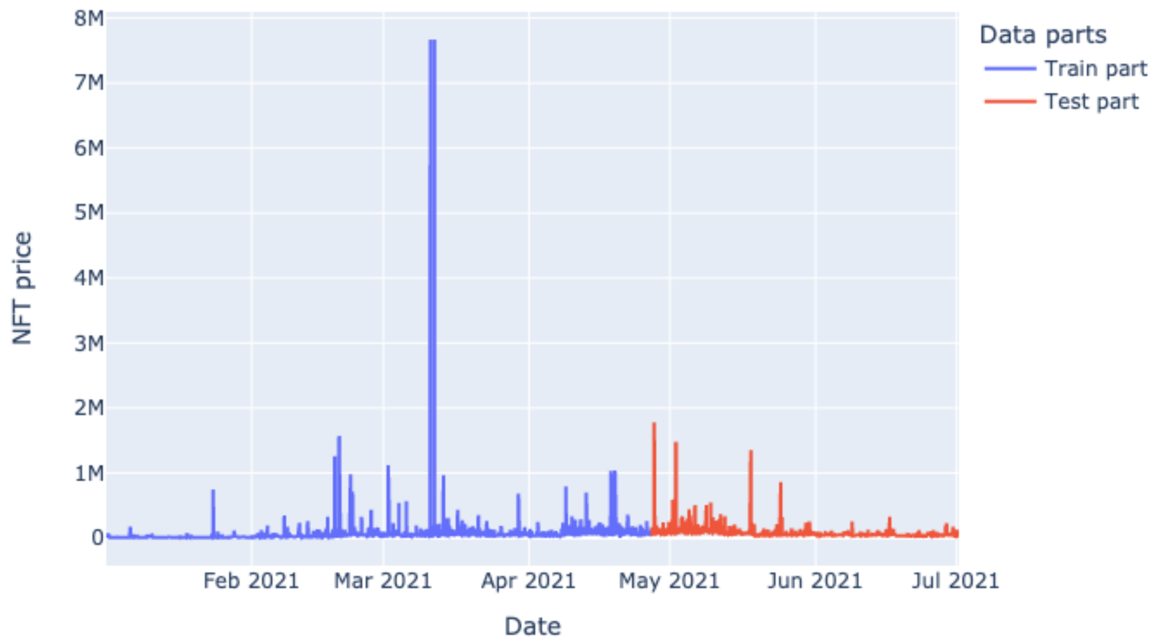


Figure 1. Distribution of the NFT price in USD over time for the analyzed period (split into train and test samples)

We used a similar strategy when calculating the variables for natural gas. Data for the examined period was obtained from the Nasdaq website. There are some gaps in gas prices - for example, prices are not quoted for weekends. Therefore, we have adopted the rule to supplement the missing data with the last available value before the missing value for a given date (e.g. for Saturday and Sunday it is the gas price on the Friday preceding a given weekend). We have attached to the table the price of gas on the day of a given transaction, the price of gas on the day of the first sale of a given NFT, as well as the change in gas price week over week (the difference between the price on the day of sale and the price 7 days earlier, divided by the price from 7 days ago) and month over month (the difference between the price on the day of sale and the price 30 days earlier, divided by the price from 30 days ago).

Finally, using the Twitter API, we supplemented the database with information on social media traffic around the NFT. According to the basic billing plan, Twitter allows one to scrape its resources up to 7 days back. Thanks to the 'Academic Research' option, we were able to access all historical data in the API. For the purposes of the analysis, for each day in the analyzed period, we obtained the number of tweets containing the hashtag 'cryptopunks' or 'cryptopunk'. These tweets were filtered according to two additional rules:

a. the tweet could not be a retweet,
b. the tweet had to be in English.

This search is case insensitive and insensitive to accents, so in the case of the 'cryptopunks' hashtag, 'CryptoPunks', 'Cryptopunks' or 'CRYPTOPUNKS' were also found. This provided us with a comprehensive analysis of the movement around the collection in the period under study.

Finally, we tested all potential independent variables for multicollinearity between them. Variables with a high VIF (greater than or equal to 6) have been removed from the database. Moreover, we analyzed the Kendall correlation index for numeric variables and removed highly correlated variables.

In Table 1, we have summarized the basic descriptive statistics for the numerical variables that we finally used in the model. For categorical variables, due to their number, after decoding the appearance elements of each NFT, we did not include such a table.

	Total USD price	**Prev. total USD price**	**ETH WoW Change**	**ETH MoM Change**	**BTC WoW Change**	**BTC MoM Change**
Count	6389.00	6389.00	6389.00	6389.00	6389.00	6389.00
Mean	56659.63	19486.72	0.07	0.33	0.05	0.22
Std	151811.33	32298.84	0.15	0.32	0.12	0.29
Min	316.14	0.00	-0.41	-0.41	-0.25	-0.43
25%	29546.45	0.00	-0.02	0.08	-0.03	-0.01
50%	41939.33	1772.25	0.06	0.36	0.05	0.20
75%	62292.69	35251.65	0.16	0.54	0.16	0.41
Max	7670018.77	712797.17	0.68	1.31	0.39	1.20
	Time diff. from prev. sale	**Gas WoW Change**	**Gas MoM Change**	**Tweets count**	**Tweets WoW Change**	**Tweets MoM Change**
Count	6389.00	6389.00	6389.00	6389.00	6389.00	6389.00
Mean	4356375.98	0.01	0.06	884.72	0.85	4.55
Std	9926026.80	0.05	0.12	579.13	1.45	6.09

Min	0.00	-0.13	-0.21	18.00	-0.80	-0.75
25%	0.00	-0.03	-0.05	399.00	-0.03	0.63
50%	288942.00	0.02	0.06	798.00	0.36	1.91
75%	3103088.00	0.05	0.14	1381.00	0.99	6.20
Max	66550625.00	0.13	0.25	2319.00	6.25	22.76

Table 1. Numerical features descriptive statistics

4. Empirical research

The aim of our article is to identify the factors that influence the price of NFT. Related to that, we test 4 research hypotheses. First, due to the complicated process and structure, NFT requires a more advanced valuation method than classical regression methods (e.g. OLS). We will be able to verify this hypothesis by comparing the metrics (RMSE, MAE, MAPE) for each of the models for the training set and the test set. Secondly, we assumed that their characteristic elements and sales history play an important role in the valuation of NFTs. Third, we stated that NFT valuation requires additional, external variables that affect the prices of non-fungible tokens, such as cryptocurrency and gas prices. Fourth, we established that social media play an important role in the valuation of NFTs and creating their popularity, and thus - driving their prices. We will be able to verify all these hypotheses by analyzing the importance of variables and Partial Dependence Plots.

We decided to standardise each column individually to the range from 0 to 1. Thanks to this, we solve the problem of variables in different scales and units.

In order to evaluate the NFT, we used three models recommended in the discussed literature - OLS, XGBoost and Bidirectional LSTM. The OLS model was tested on 80% of the available data, and then predictions were made on the test data set. The comparison of predictions for the training set with the true values of the dependent variable for the OLS model is shown in Appendix B, in Figure 2. A similar comparison for the test data for the OLS model is shown in Appendix B, Figure 3. In the case of the XGBoost model, using the 5-split cross-validation and the Randomized Search algorithm, we established the values of 6 considered hyperparameters. The best hyperparameters for the model are presented in Table 2, while the list of considered values for each of the hyperparameters is shown in Appendix A. Similarly to the OLS model, the comparison of the predictions with the real values for the XGBoost model is shown in Figures 4 and 5 (for training and test data, respectively).

Hyperparameter	Value
Number of gradient boosted trees	200
Maximum tree depth	8
Boosting learning rate	0.1
λ	2
α	0.1
γ	None

Table 2. Best hyperparameters for XGBoost model

The last model used in our analysis is the Bidirectional LSTM. Due to the fact that the input data for this type of model must be in three-dimensional form, we assumed the time step equal to 1. To optimize the training of neural networks process, we used the Early Stopping algorithm with the patience parameter equal to 10 and the loss metric. This means that the training of the neural network is stopped after 10 epochs without improving the loss. In the case of neural networks, hyperparameters were selected manually, based on the analysis of averaged metric values (MAE, MAPE, RMSE) for training and validation sets. The data was trained and validated for 8 hyperparameters, the full list of which is included in Appendix A. The best hyperparameters for the model are shown in Table 3. We used an adoptive learning rate optimizer 'Adam'. This algorithm updates any parameter with an individual learning rate. Thanks to this, each parameter in the neural network has a specific learning rate. For this reason, neural networks with a rigid learning rate and decay rate value turned out to be worse than those for which the values were selected by the optimizer.

Hyperparameter	Value
Activation function	'relu'
Number of hidden layers	6
Number of hidden neurons	16
Number of epochs	200
Batch size	128
Dropout layer value	None
Learning rate	-
Decay rate	-

Table 3. Best hyperparameters for Bidirectional LSTM

The first research hypothesis in our article was based on the question of whether simple linear regression models are able to predict values for such a complex financial asset as NFT as well as more advanced machine learning and deep learning models. In Table 4 and Table 5, we have included all three metrics for each of the analyzed models, for training and testing data, respectively. In the case of the training set, the OLS model turns out to be the worst model of the three. All three metrics indicate worse predictions than in the case of machine learning and deep learning models. MAPE on the training set for OLS is 0.68, while in the case of XGBoost and Bidirectional LSTM it is relatively similar to one another - 0.27 and 0.33, respectively. This result, compared to the results discussed in the literature review, is a very good result, but still requires improvement and deeper research. MAE for XGBoost and Bidirectional LSTM models in the training set is about 2 times smaller than in the OLS model. In the case of the test set, overfitting is noticeable. The OLS model shows lower values of the MAPE and RMSE metrics, while the R2 for this model has fallen more than twice, and the MAE has increased by about 18,000. In the case of the XGBoost and LSTM models, all three metrics indicate a worse prediction quality for the test set than in the training data. MAPE for both models is 0.45, while MAE is around 34,000 for XGBoost and around 45,000 for LSTM. These results can certainly be improved by further tuning the hyperparameters, but also by refining the dataset taken for the prediction. In the case of the LSTM model, we cannot say that its performance is better than that of the OLS model, while the XGBoost turned out to be the better model for each of the metrics. Based on the prediction graphs for all three models included in Appendix B, we can conclude that the OLS model tends to overestimate the prediction values, while the LSTM underestimates them. Summarizing the above analysis, Tables 4 and 5, as well as the results available in Appendix B, we can conclude that our first hypothesis that NFT pricing requires more advanced methods than classical regression methods is confirmed, but requires a deeper analysis in the case of selecting hyperparameters for models and the selection of data used in the model.

Metric	OLS	XGBoost	Bidirectional LSTM
MAE	20138.31	9951.07	11854.22
MAPE	0.68	0.27	0.33
RMSE	90079.31	32403.87	61969.29
R2	0.69	0.96	0.87

Table 4. Metrics comparison for training data

Metric	OLS	XGBoost	Bidirectional LSTM
MAE	38594	34396.33	45507.28
MAPE	0.54	0.45	0.45
RMSE	78178.65	71293.87	110669.07
R2	0.3	0.42	-0.4

Table 5. Metrics comparison for test data

We decided to verify the next three hypotheses by analyzing the importance of variables obtained using the Permutation Feature Importance method and based on the Partial Dependence Plots analysis. The analysis of this type of characteristics in the case of neural networks is much more difficult[3]. In addition, the XGBoost model was found to have the best performance for all three metrics for both the training and test sample (Tables 4 and 5). As mentioned above, the OLS model performs worse than the XGBoost model on both sub-samples, while in the case of the Bidirectional LSTM model, the issue of further parameterization causes significant overfitting on the test sample. Based on the performance of the XGBoost model and the previously mentioned difficulties in assessing the significance of variables for the neural network model, we decided to apply the PFI and PDP analysis on the XGBoost model. Moreover, for each type of variable (variables related to cryptocurrencies, natural gas, and Twitter), we decided to recalculate the regression rating metrics for all three models while maintaining the same hyperparameters as for the main regressions. This will allow us to thoroughly analyze the change in the value of the metrics in the absence of one of the three categories of external variables, which is particularly important when verifying the third and fourth hypotheses.

Figure 2 shows the PFI analysis for the test set for the 15 most important variables in the model. In Figure 3, we present the PDP analysis for 4 variables that stand out in the PFI analysis - the previous NFT price in USD, dummy variable, whether the NFT is presented as a zombie, time since the last sale of a given NFT in seconds and the number of tweets about the CryptoPunks collection on the date of sale of the NFT. Additionally, in Appendix C, we present the results of the PFI analysis, taking into account the 15 most important variables in the model for training

[3] E.g. in python popular libraries such as *scikit-learn* do not offer 3D data analysis, while the *shap* library is incompatible with the tensorflow library above version 2.X.X.

dataset. In Appendix D we present the results of the PDP analysis for other numerical variables from the model.

Figure 2. Permutation Feature Importance based on xgboost model – test data

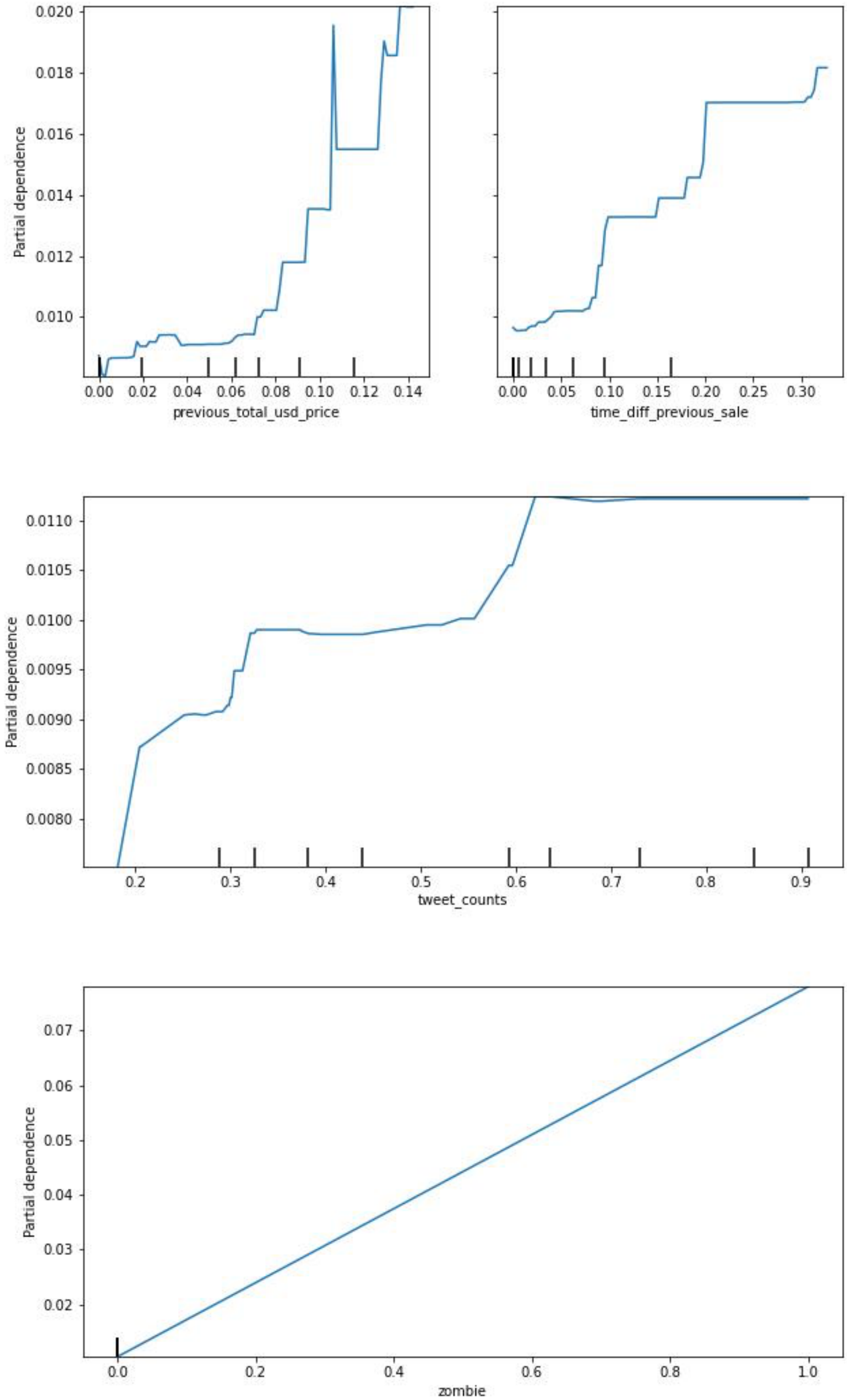


Figure 3. Partial Dependence Plots for four most significant variables

Both analyzes allowed us to verify the other three hypotheses. First, based on Figure 2, we can see that the most important regressors in the NFT valuation model turned out to be: the previous NFT price in USD, the time since the last sale of a given NFT, or the most characteristic avatar elements, such as alien, zombie, beanie or hoodie. It is noteworthy that NFT prices are especially driven when there are rare appearance elements, such as a zombie avatar or avatar with hoodie, beanie, or pilot helmet. This indicates that NFT is subject to similar mechanisms as art or luxury goods - their rarity particularly affects the price of a given good, which is consistent with the literature discussed earlier. Hence, we can conclude that the main variables influencing the valuation of NFTs are their characteristic elements and sales history, which confirms our second hypothesis.

Secondly, only the variables related to the BTC week over week price change and the week over week natural gas price change were indicated among fifteen most important variables in the model (Figure 2), but they were classified in 9th and 11th place. The ETH price fluctuations have not been identified as ones of the most important predictors. By analyzing Tables 6 and 7, where we have included the regression results without variables related to the cryptocurrency market (Table 6) or without variables related to the natural gas market (Table 7), we can conclude that the absence of these variables only minimally affects the deterioration of the regression assessment metrics in the case of XGBoost and Bidirectional LSTM models. In the case of the OLS model, some of the metrics even improve, especially in the absence of variables related to the cryptocurrency market. Therefore, we cannot state unequivocally that these variables are crucial in the valuation of NFT, especially when comparing their impact on the performance of models with the impact of variables related to social media. Thus, we are unable to confirm the third hypothesis posed in this article.

Metric	OLS	XGBoost	Bidirectional LSTM
MAE	27667.59	36022.21	48940.94
MAPE	0.33	0.5	0.49
RMSE	71042.33	71644.6	107323.93
R2	0.42	0.41	-0.32

Table 6. Metrics comparison for test data without cryptocurrencies related variables

Metric	OLS	XGBoost	Bidirectional LSTM
MAE	36118.28	37100.44	52614.83
MAPE	0.44	0.49	0.58
RMSE	78028.57	76346.65	109138.32
R2	0.3	0.33	-0.36

Table 7. Metrics comparison for test data without natural gas related variables

Finally, based on Figure 2, we are able to indicate that the number of tweets about the CryptoPunks collection on a given day is the fourth most important variable in the NFT valuation model (*tweet_counts* feature). Moreover, with the same parameters as in the case of the main regression, for each of the considered models, we recalculated the regression rating metrics for the test sets, but without taking into account Twitter-related variables - the number of tweets on a given day, week over week change in the number of tweets, as well as month over month change in the number of tweets. The results are presented in Table 8. All three metrics for each model are much worse than for the main regression and its results in Table 5. Based on Figure 2 and 3, where the PDP analysis for *tweet_counts* variable is visible, as well as results in Table 8, we are able to conclude that the number of tweets has a significant impact on the NFT price. It confirms our fourth hypothesis, which assumed an important role of social media in creating token prices.

Metric	OLS	XGBoost	Bidirectional LSTM
MAE	52174.98	42120.17	61766.32
MAPE	0.79	0.54	0.66
RMSE	86419.18	84609.56	114160.41
R2	0.14	0.18	-0.5

Table 8. Metrics comparison for test data without Twitter related variables

Conclusions

The main focus of our article was the valuation of non-fungible tokens. For this purpose, we have collected NFT sales data from the CryptoPunks collection, including transaction history as well as the characteristics of individual NFTs. In addition, we supplemented the database with daily data on cryptocurrencies (Bitcoin and Ethereum), daily natural gas prices, as well as the number of tweets about the collection during the analyzed period. The final data sample used in the analysis covered the period from January 2021 to July 2021. Based on the available literature, we used three regression models in the analysis: Ordinary Least Squares regression, XGBoost model and Bidirectional Long Short-Term Memory. In the article, we examined four research hypotheses:

1. Due to the complicated process and structure, NFT validation requires more advanced methods than basic regression models, eg machine learning or deep learning algorithms.
2. The main factors influencing NFT are its characteristics - sales history, as well as the uniqueness of a given token.
3. NFT valuation requires additional external variables that affect the market - such as the prices of cryptocurrencies and gas.
4. Social media play an important role in the valuation of NFT and the inclusion of social media variables (such as the number of mentions of NFT on social media) will improve the predictive accuracy of the model.

We verified the first hypothesis by comparing classical regression metrics for all three models on both samples (training and testing). In this way, we were able to judge the difference in the model performance. We verified the next three hypotheses using the Permutation Feature Importance values and the Partial Dependence Plots analysis for the XGBoost model. Also, we recalculated all three models and their metrics on the test data again in three different cases - without taking into account the variables related to the cryptocurrency market, without taking into account the variables related to the natural gas market and without taking into account social media variables. We compared them with the metrics for the models built on all variables, while maintaining the same hyperparameters.

Firstly, taking into account the analyzed metrics for assessing the performance of regression models (MAE, MAPE, RMSE), we could conclude that the XGBoost model performs better than the OLS model for each of the analyzed metrics, both for the training and test set. In the case of the Bidirectional LSTM model, the model on the training data had much better predictions than the OLS model, while in the case of the test data it showed worse performance. This allows us to confirm our first hypothesis, while both machine learning and deep learning models require further hyperparameter tuning and better data selection for training.

Secondly, with the help of Permutation Feature Importance and Partial Dependence Plots, we were able to prove that the sales history as well as unusual, unique characteristics largely drive the value of NFT, which confirms the second research hypothesis of the article.

Third, based on PFI and PDP analyzes, as well as recalculated regression metrics without individual market variables, we rejected the hypothesis that the variables related to the cryptocurrency market or the natural gas market are crucial for the valuation of NFT.

Last but not least, the number of tweets about the collection turned out to be one of the most important variables in the evaluation process (PFI and PDP analyses). Recalculation of models with the same hyperparameters, but without Twitter variables, produced significantly worse results for all three models. This confirms the fourth hypothesis put forward in the article. Social media significantly influence the popularity of NFTs, and thus play an important role in shaping their prices.

Our study is one of the few studies on the overall NFT valuation, taking into account external market variables. Therefore, we see several possible extensions of the study in order to improve the performance of the models and the possibility of more accurate price prediction. First of all, expanding the sample with new observations. The data could be extended not only with a larger time span (the second half of 2021 and 2022), but also with other NFT collections, e.g. CryptoKitties or Bored Ape Yacht Club. Additionally, it is worth considering other independent variables in the model. As the knowledge of NFT is being constantly expanded, a greater analysis of external variables could be of great benefit. What's more, research can be extended with new models, as well as deeper hyperparameter tuning. Finally, an important element of the analysis is the interpretation of the importance of the variables. It would be worth exploring this issue in the case of results available for neural networks.

The NFTs are a fresh topic that is still being explored. Therefore, an attempt to further expand our research with new observations, variables and methods may bring about a significant improvement in the NFT valuation.

References

Ante, L. (2022). The Non-Fungible Token (NFT) Market and Its Relationship with Bitcoin and Ethereum. FinTech 1, no. 3: 216-224.

Anyfty. (2021). Ayfty's first results in NFT pricing prediction. Medium.com. August 9, 2021.

Aubry, M., & Kraeussl, R., & Manso, G., & Spaenjers, C. (2020). Machines and Masterpieces: Predicting Prices in the Art Auction Market. Working Papers hal-02896049, HAL.

Campos, N. F., & Barbosa, R. L. (2009). Paintings and Numbers: An Econometric Investigation of Sales Rates, Prices, and Returns in Latin American Art Auctions. Oxford Economic Papers, 61(1), 28–51.

Dinu, V., Dinca, V.M., Dima, A.M., Andronie, M. (2020). "The Price of Artworks in 2020: Examining the Elements with Impact on the Appraisement of Art Pieces", Transformations in Business & Economics, Vol. 19, No 2 (50), pp.283-295.

Dowling, M. (2022). Is non-fungible token pricing driven by cryptocurrencies?. Finance Research Letters. Volume 44. 102097. ISSN 1544-6123.

Horky, F., Rachel, C., Fidrmuc, J. (2022). Price determinants of non-fungible tokens in the digital art market. Finance Research Letters. Volume 48. 103007. ISSN 1544-6123.

Kaczynski, S., and Kominers, S. (2021). How NFTs Create Value. Harvard Business Review. November 10, 2021.

Kapoor, A., & Guhathakurta, D., & Mathur, M., & Yadav, R., & Gupta, M., & Kumaraguru, P. (2022). TweetBoost: Influence of Social Media on NFT Valuation.

Li, B., and Liu, T. (2021). An Analysis of Multi-Modal Deep Learning for Art Price Appraisal. 2021 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom). pp. 1509-1513.

Liu, C. (2022). Prediction and Analysis of Artwork Price Based on Deep Neural Network. Scientific Programming. 1-10. 10.1155/2022/7133910.

Murphy, M. (2021). NFTs come with big valuation challenges. Journal of Accountancy. July 15, 2021.

Nadini, M., Alessandretti, L., Di Giacinto, F. et al. (2021). Mapping the NFT revolution: market trends, trade networks, and visual features. Sci Rep 11, 20902.

NonFungible.com. (2021). Non-Fungible Tokens Quarterly Report Q1 – 2021.

Pinto-Gutiérrez, C., Gaitán, S., Jaramillo, D., and Velasquez, S. (2022). The NFT Hype: What Draws Attention to Non-Fungible Tokens?. Mathematics 10, no. 3: 335.

Renneboog, L., & Spaenjers, C. (2013). Buying Beauty: On Prices and Returns in the Art Market. Management Science, 59(1), 36–53.

Schaar, L., and Kampakis, S. (2022). Non-Fungible Tokens as an Alternative Investment: Evidence from CryptoPunks. The Journal of The British Blockchain Association, January.

Wang, Q., & Li, R., & Wang, Q., & Chen, S. (2021). Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges.

Worth, T. (2021). Painting2Auction: Art Price Prediction with a Siamese CNN and LSTM.

Appendix A

XGBoost hyperparameters grid:

1. Number of gradient boosted trees: [100, 200, 300, 400, 500, 1000]
2. Maximum tree depth: [3, 4, 5, 6, 7, 8]
3. Boosting learning rate: [0.01, 0.05, 0.1, 0.5, 0.6]
4. λ: [0, 0.5, 1, 1.5, 2]
5. α: [0, 0.001, 0.05, 0.1]
6. γ: [None, 0.1, 0.2, 0.3, 0.4]

LSTM hyperparameters grid:

1. Activation function: [‘tanh’, ‘relu’, ‘sigmoid’]
2. Number of hidden layers: [2, 3, 4, 5, 6]
3. Number of hidden neurons: [10, 16, 32, 64, 128]
4. Number of epochs: [100, 200, 300, 500, 1000, 1500]
5. Batch size: [16, 32, 64, 128, 256]
6. Dropout layer value: [0.2, 0.3, 0.4, 0.5]
7. Learning rate: [0.000001, 0.00001, 0.0001, 0.001, 0.01, 0.1]
8. Decay rate: [0.0001, 0.001, 0.01, 0.1]

Appendix B

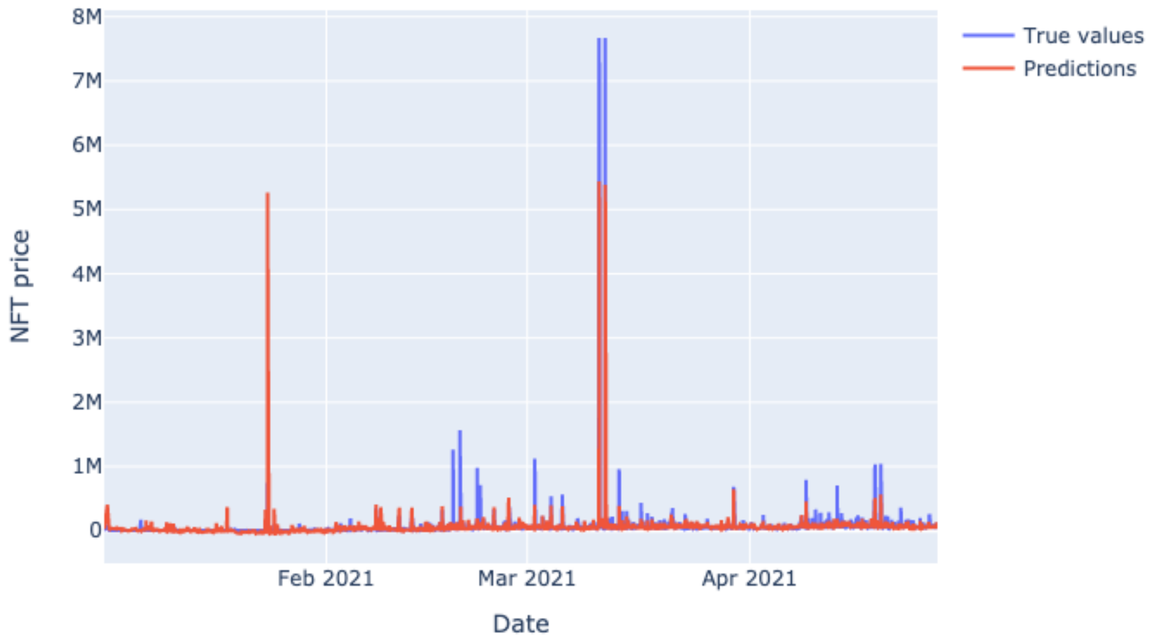


Figure 4. OLS model predictions – train data

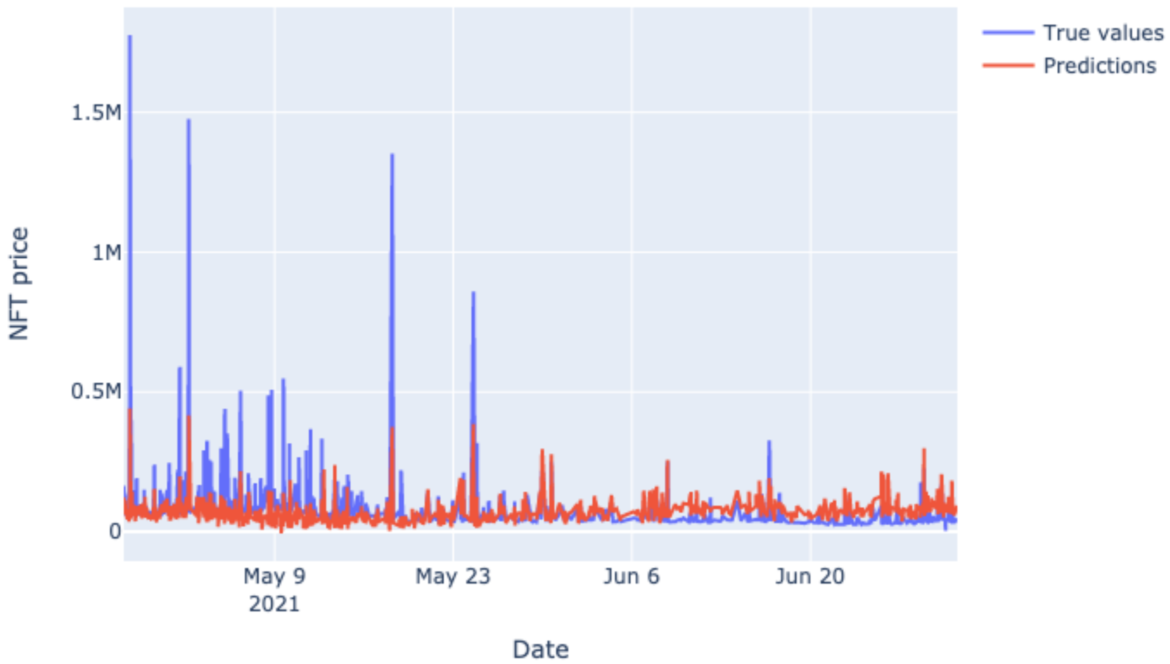


Figure 5. OLS model predictions – test data

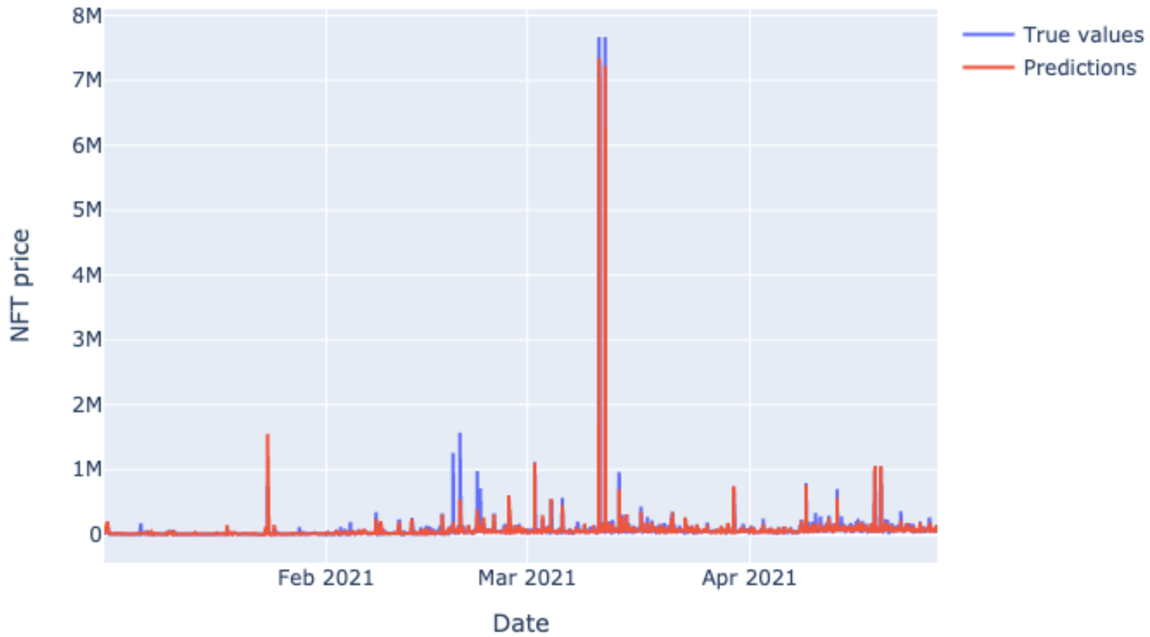


Figure 6. XGBoost model predictions – train data

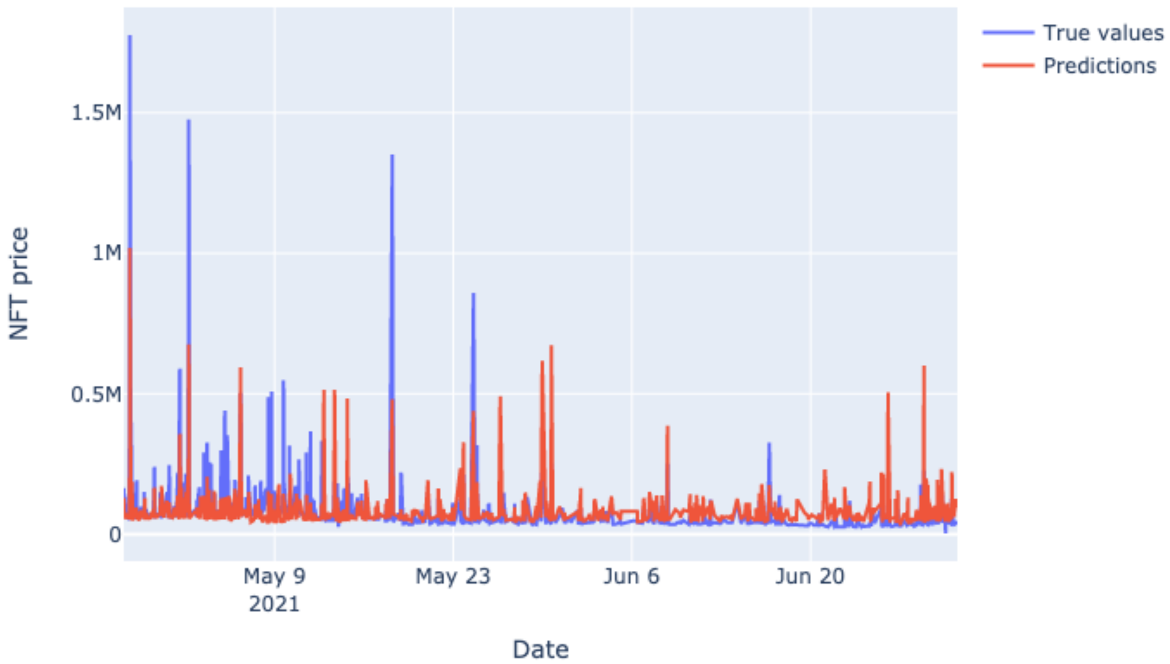


Figure 7. XGBoost model predictions – test data

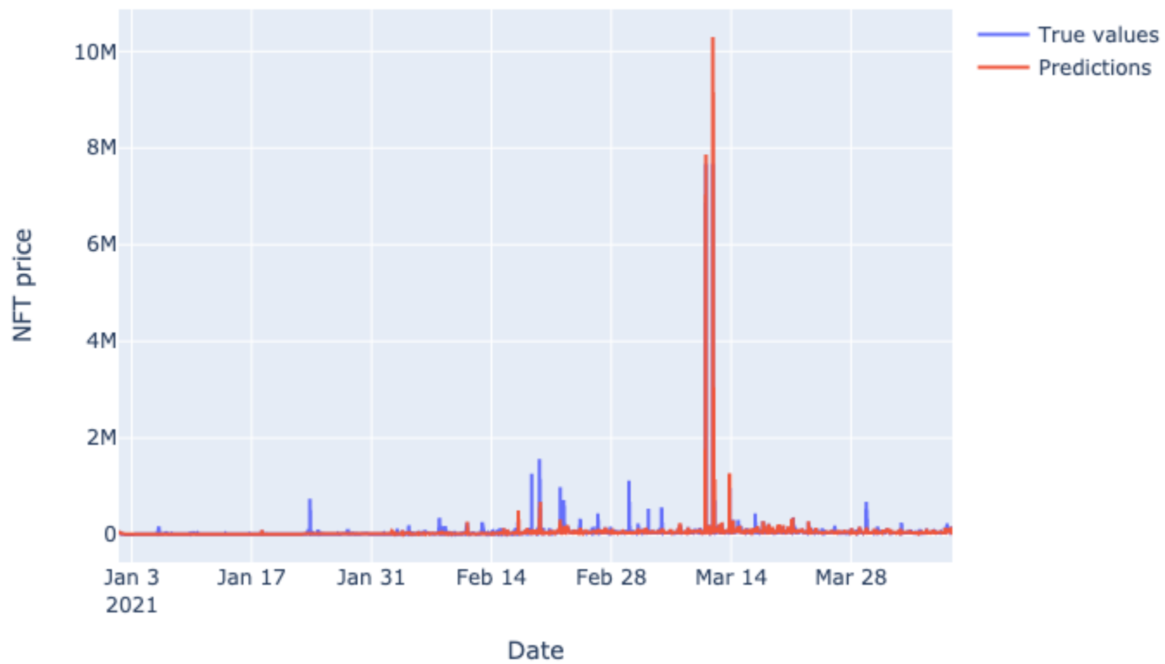


Figure 8. Bidirectional LSTM model predictions – train data

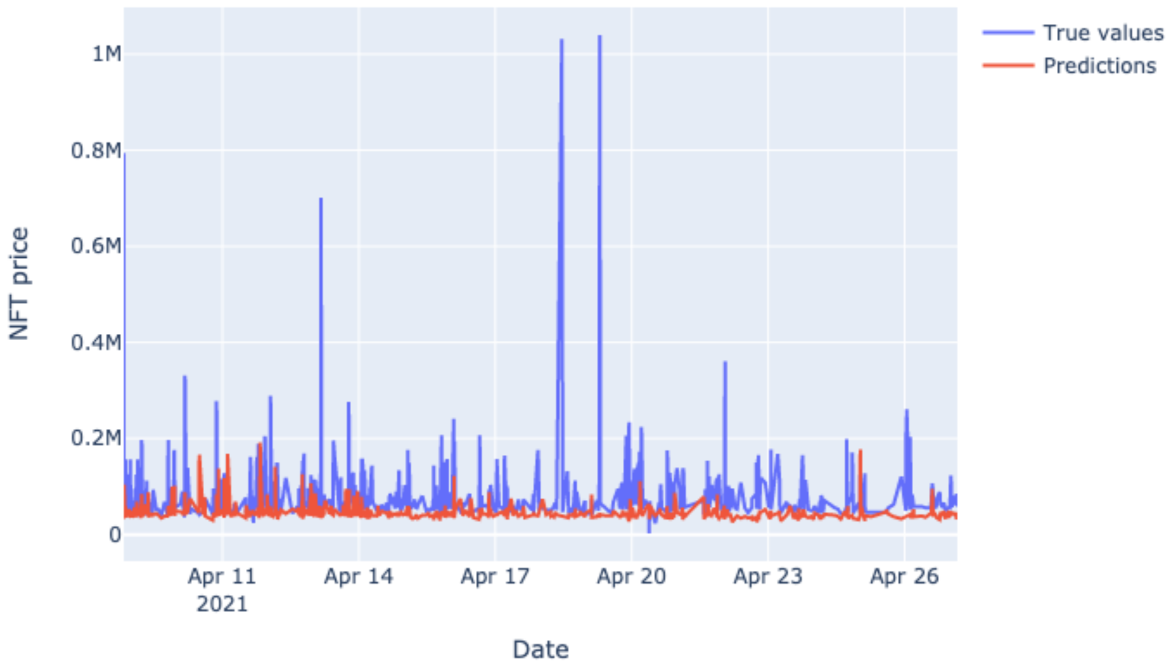


Figure 9. Bidirectional LSTM model predictions – validation data

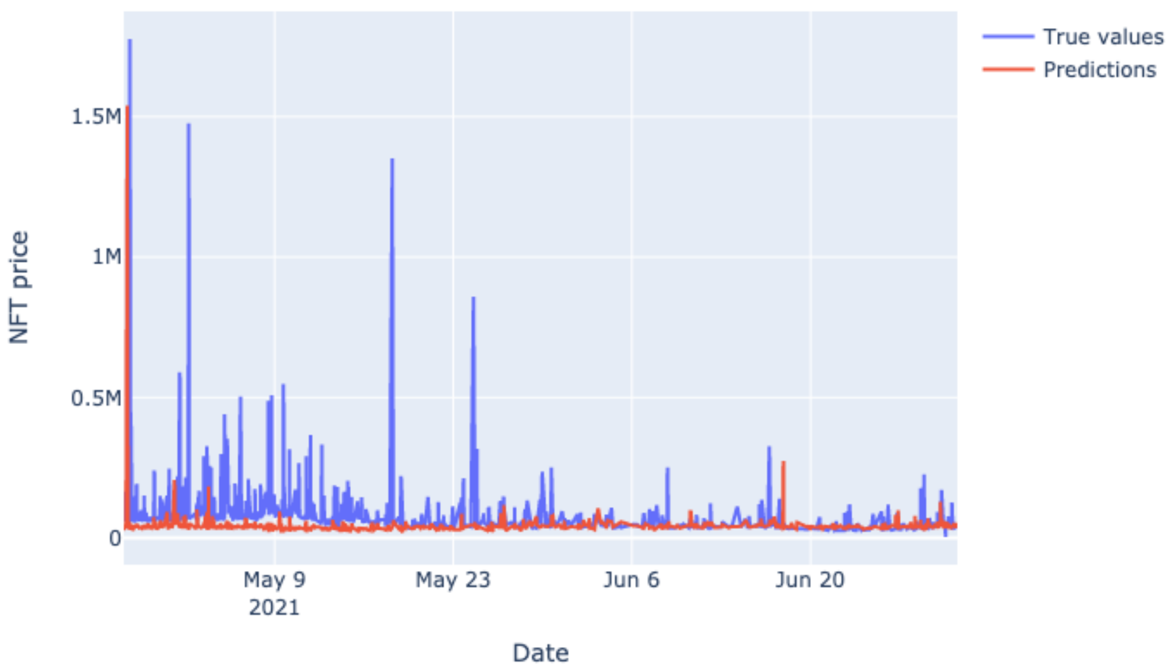


Figure 10. Bidirectional LSTM model predictions – test data

Appendix C

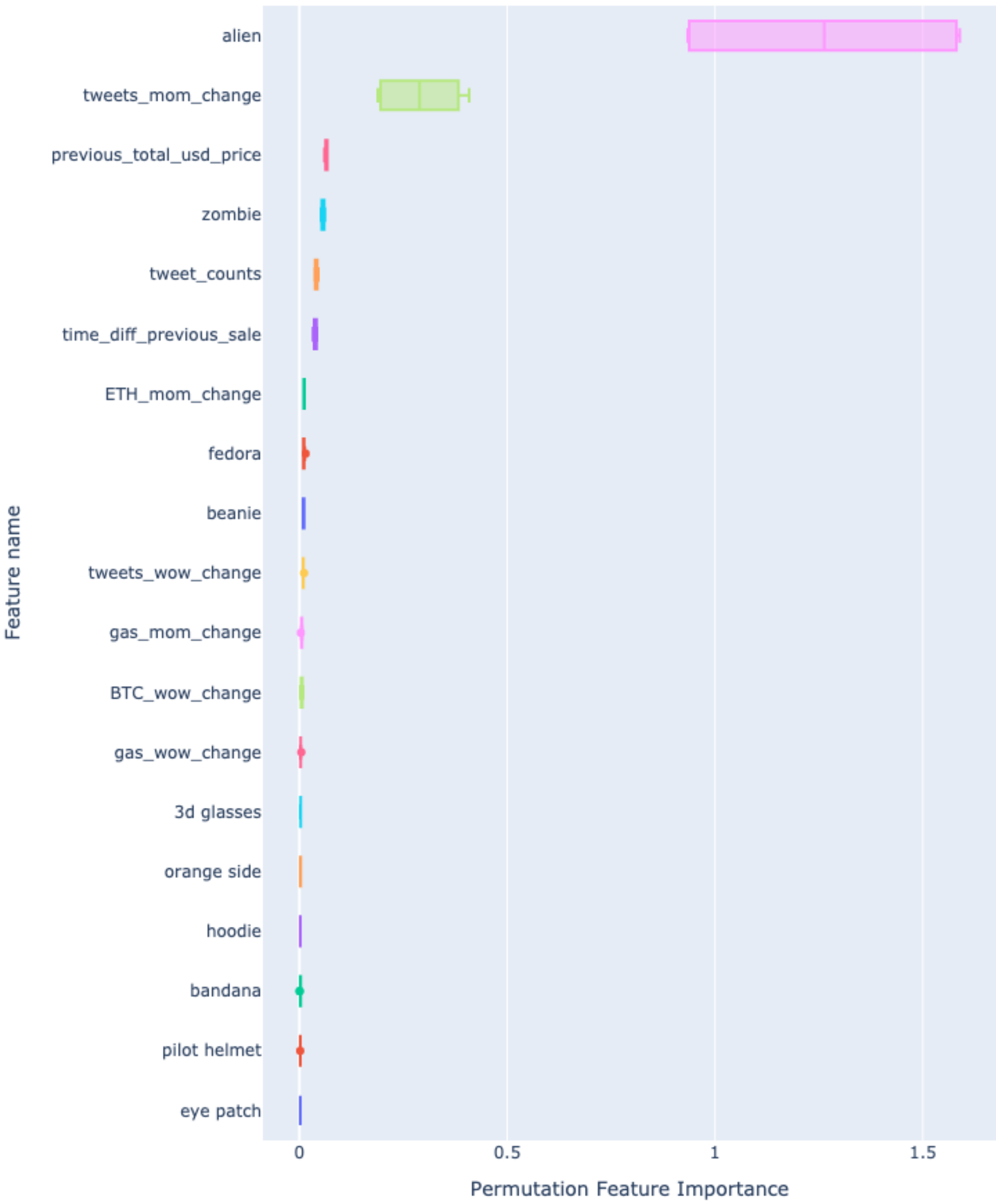
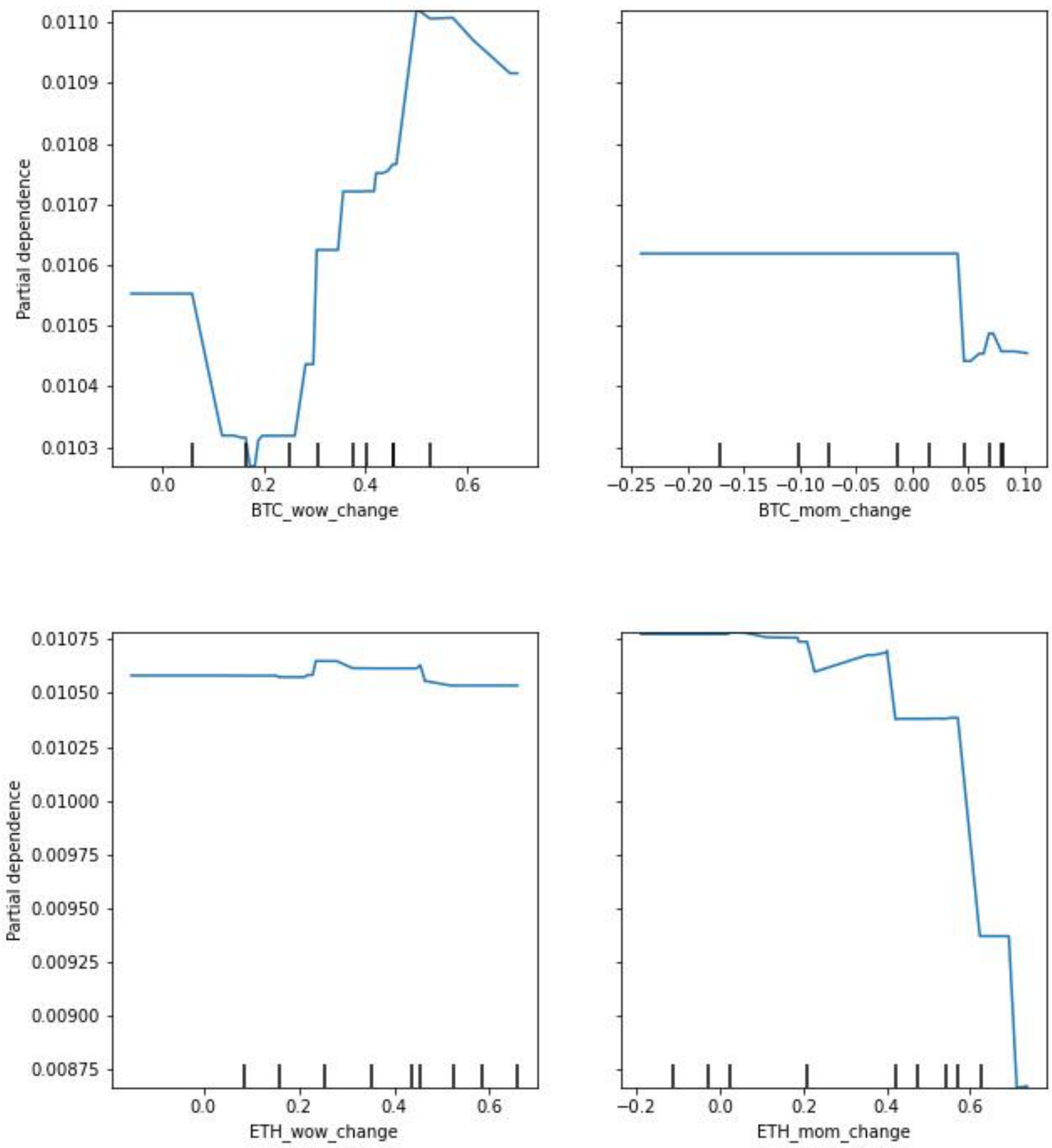
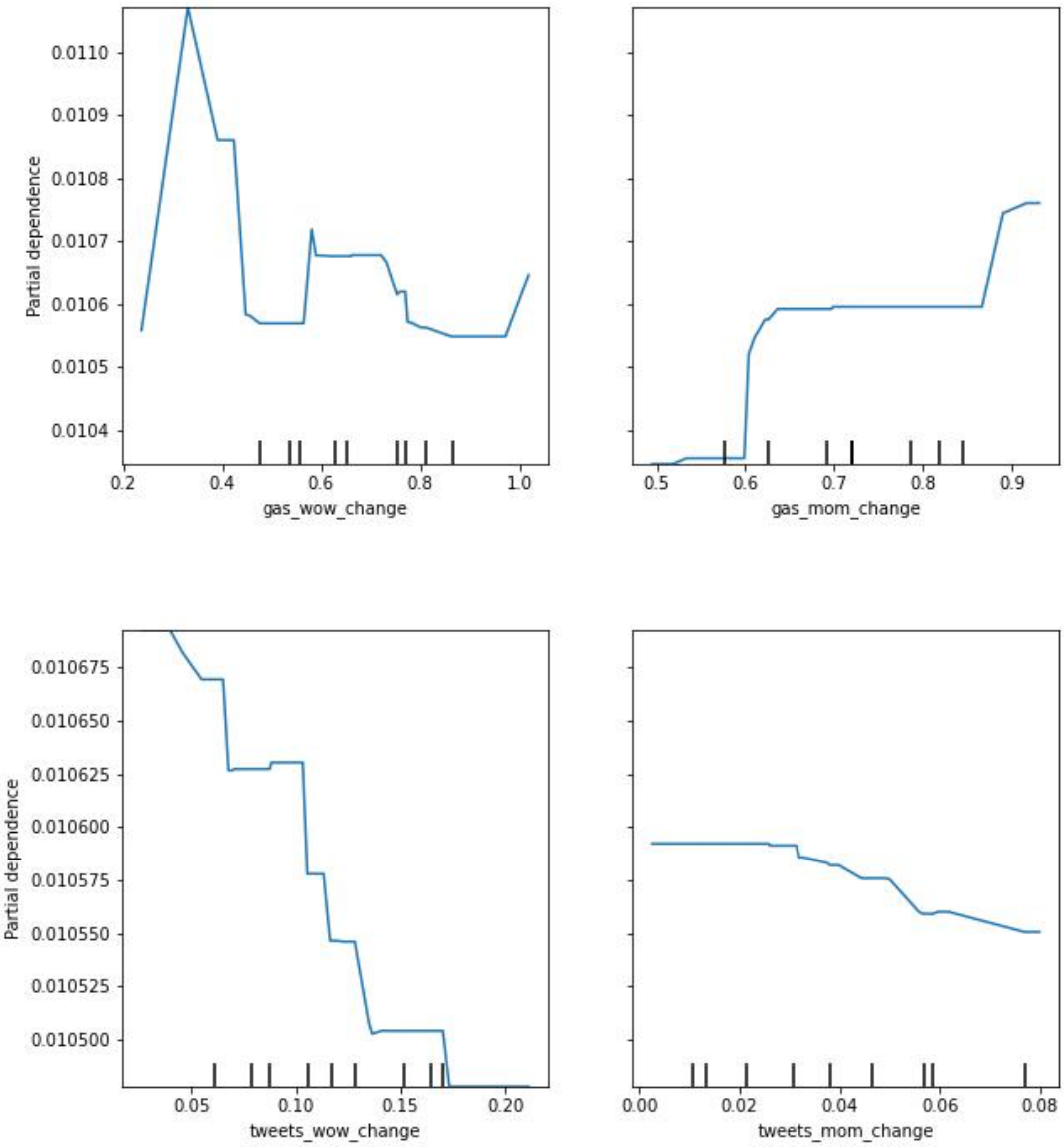


Figure 11. Permutation Feature Importance – train data

Appendix D


Partial dependence
BTC_wow_change
BTC_mom_change
Partial dependence
ETH_wow_change
ETH_mom_change


Partial dependence
0.0110
0.0109
0.0108
0.0107
0.0106
0.0105
0.0104
0.2
0.4
0.6
0.8
1.0
gas_wow_change
0.5
0.6
0.7
0.8
0.9
gas_mom_change
0.010675
0.010650
0.010625
0.010600
0.010575
0.010550
0.010525
0.010500
0.05
0.10
0.15
0.20
tweets_wow_change
0.00
0.02
0.04
0.06
0.08
tweets_mom_change

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl