NFT ART MARKET PREDICTOR

C Lanston Davis
20BCE1613

Chaitanya Kamasani
20BRS1037

Abstract – Non-Fungible Tokens (NFTs) are digital assets that represent objects like art, collectible, and in-game items. They are traded online, often with cryptocurrency, and are generally encoded within smart contracts on a blockchain.
Our proposed project predicts this NFT art using ARIMA model. The main advantage of predicting using ARIMA model is because it makes use of lagged moving averages to smooth time series.

Key Words : NFT's, Blockchain, ARIMA,

INTRODUCTION

Non-fungible tokens (NFTs) are a highly nascent and emerging phenomenon revolutionizing how digital assets are traded. NFTs embody immutable rights to unique digital assets such as digital art and collectibles and are represented as digital tokens that can be traded across marketplaces utilizing blockchain technologies. NFTs engender new ways to organize, consume, move, program, and store digital information and have experienced a rapid rise in various adaptations across art, sports, broadcasting, content creation, and tech-crypto businesses.

From the top five collections in the NFT market, the most exchanged NFTs belong to the categories Games, Collectibles, and Art, which account for 44%, 38%, and 10% respectively of transactions. In terms of market volume, the Art category has dominated, contributing ~71% of the total transaction volume. With regards to the relationship between traders, the top 10% of traders (measured by their number of purchases and sales) perform 85% of all transactions. Furthermore, traders specialized in a collection tend to buy and sell NFTs with other traders specialized in the same collection.

Here we use ARIMA model, methodology is a statistical method for analysing and building a forecasting model which best represents a time series by modelling the correlations in the data. Owing to purely statistical approaches, ARIMA models only need the historical data of a time series to generalize the forecast and manage to increase prediction accuracy while keeping the model parsimonious.

Additionally, this data was passed through various machine learning models, each with varying goals. A linear regression model was used to identify the features most strongly correlated with an NFT's valuation as well as to identify redundancy among features; furthermore, various iterations were investigated including : ARIMA, XBoost, Random Forest.

LITERATURE SURVEY

Paper-1 : **Characterizing the OpenSea NFT Marketplace**

Existing paper work:

They use the OpenSea (nft market place) for getting the sales data of the nft's.

Advantages :

- The data is easily available we can download it anytime
- The data is high quality it is consistent, reliable, accurate and complete

Disadvantages :

- We have to manually download the data everytime to run the algorithm
- It takes some days for the real time sales data to get updated in the website.

Overcome the problem :

We have used real time sales data which is fetched directly from the market place through COVALENTHQ API instead of doing it manually.

Paper-2 : Understanding Security Issues in the NFT Ecosystem

Existing Paper Work : Most academic research has focused on attacks against decentralized finance (DeFi) protocols and automated techniques to detect smart-contract vulnerabilities .They focused on the market dynamics and security issues of the NFT Ecosytem.

Advantages :

- We can identify which marketplace is autherized .
- We can also collect a large amount of asset and event data pertaining to the NFTs being traded in the examined marketplaces

Disadvantages :

- Danger to artist identity

Overcome the problem :

We just use the smart contract address of the nft and don't reveal the identity of the artist their identity remain anynomous.

Paper-3 : **Mapping the NFT revolution: market trends, trade networks, and visual features.**

Existing Paper Work :

This paper aims to demystify the overall structure of the Non Fungible Token (NFT) market and provides a framework for quantifying its evolution.

Advantages :

- It shows network of interactions between traders (linked by buyer and seller) and a clustering of objects by visual features and collections.
- AlexNet, a convolutional neural network, was used to produce dense vector representations of images, and principal component analysis (PCA) was used to study the similarity between 1.25 million different NFTs.

Disadvantages :

- Several limitations of this study include data gathered from NFT marketplaces instead of directly from Ethereum or WAX blockchains, causing independent NFT producers to be left unaccounted for.

Overcome the problem :

The data we gather through API is directly from the ethereum block chain so independent nft producers are not left to be unaccounted for.

Paper-4 : The NFT Hype: What Draws Attention to Non-Fungible Tokens?

Existing Paper Work : This paper focuses on utilizing vector autoregressive models (VARs) to show that core cryptocurrencies, namely Bitcoin (BTC) and Ether (ETH) draw the most attention towards predicting future NFT price.

Advantages :

- This team utilizes the S&P 500, google search trends, and the prices of cryptocurrencies as indicators for future price of an NFT .
- This team uses wavelet coherence techniques to investigate co-movement between cryptocurrency returns and NFT levels of attention

Disadvantages :

- The results of this paper show that there is no significant relationship between Ether returns and attention to NFTs

Overcome the problem :

We show that there is a relationship between Ether returns and the prediction of an NFT.

Paper-5 : TweetBoost: Influence of Social Media on NFT Valuation

Existing Paper Work : This paper aims to answer two main questions: a) What is the relationship between user activity on Twitter and price on OpenSea? b) Can we predict NFT value using signals obtained from Twitter and OpenSea, and identify which features have the greatest impact on the prediction?.

Advantages :

- create one of the first NFT datasets consisting of both OpenSea and Twitter data.
- Using both a Binary and Multi-classification model to first predict whether or not the NFT will be profitable and then classifying the profitable NFTs into varying price brackets.

Disadvantages :

- The resulting accuracy of the models are not that high.

Overcome the problem :

We use the XGBoost algorithm which gives more accuracy than the models they have used.

Paper-6 : NFT Market Prediction Using Linear Regression

Existing Paper : In exiting paper NFT art predictions using linear regression and some previous data of price.

Advantage :

- It uses simple linear regression, which is basic and understandable to all.

Disadvantage :

- This model uses previous data which is not latest and which cannot be relied on it.

Overcome Solution :
- We use latest information by taking its smart contract of nft, where it collects information from yesterday’s market value also.

Paper-7 : **Predicting NFT Marketplace Growth Using Frequency of Tweets regarding Safety Concerns**

Existing paper :
Its a regression analysis was conducted by monitoring Twitter posts regarding NFT safety and using this information to try and predict price movement.

Advantage :

- Advantage of this prediction is latest Twitter posts where it acts as latest information for predicting the price
- Its helpful, which bridges the information gap by investigating one of potential driving forces of this market, namely public perception over its safety.

Disadvantage :

- In this twitter posts creates negative impact, if there is any negativity news on nft.
- In this model was able to predict with a very low accuracy, indicating that a weal relation may exist.

Overcome Solution :
We have used XBoost algorithm which gives more accuracy.

Paper-8 : **NFT Appraisal Prediction : Utilizing Search Trends Public, Market Data, Linear Regression and Recurrent Neutral Networks**

Existing Paper :
In this paper, they investigate the corelation between NFT valuations and various features from public market data and social trends data. These data sources were chosen such that they draw connections to more traditional investment classes as well as make an effort to quantify data.

Advantage:
- This paper uses NFT metadata and social trends data such that it will be latest data for predicting accurately.

Disadvantage :

- Eventually this model uses only linear regression and RNN model, where RNN model is so complex.

Overcome Solution :

We use ARIMA model which produces lower error values that LSTM model in monthly and weekly series.

Paper-9 : Research on Artwork Pricing of China

Existing Paper :
To provide reference for market participants and policy makers, this paper introduced the present situation of art auction market, summarized the main methods of Chinese artwork pricing, analyzed the existing problems of artwork pricing in China and put forward suggestions to perfect present pricing mechanism of Chinese art market

Advantage :

- Research shows that external conditions, scarcity of artworks and characteristics of auction are the important factors to influence the price of art. Influencing factors analyzing, comparable comparison pricing, arithmetic average pricing and the Hedonic method are the main methods of artwork pricing in China.
- Completing relevant laws and regulations, cultivating professions for artwork pricing and completing electronic database of art market could help to complete Chinese artwork pricing mechanism.

Disadvantage :

- The results of this paper show only about popular trends and considers only its own people sentiments and thoughts and gives bias for some than equality for all

Overcome Solution :

Here we remove he bias and apply the given algorithm on whole of art market without considering their prices and popularity and then only we will predict price

Paper 10: Analysis of Non-Fungible Token Pricing Factors with Machine Learning

Existing Paper Work : Not all rare NFTs are associated with a higher price, especially for play-toearn gaming NFTs. In this paper, they studied the top-ranked play-to-earn gaming NFTs on Axie Infinity
Advantages :

- They found that, in addition to rarity, utility is also a significant factor influencing the price. Furthermore, we use utility as a predictor to predict the price of Axies using the XGBoost regressor.
- Their results reveal that, compared to using rarity-based predictors only, leveraging utility based predictors can improve the prediction accuracy, thus highlighting utility as a price determinant for play-to-earn gaming NFTs

Disadvantages :

- Did not consider market trends and social media impression which have the highest factor in supply demand

Overcome Solution :

We use the XGBoost algorithm which gives more accuracy which they have used, and also included real time analysis using the real time price from second to second.

PROPOSED WORK

NFT Data

Today, there are more NFTs on OpenSea and other NFT marketplaces than there were websites in 2010 [30]. As such, querying and searching data related to NFTs has become a challenging problem due to the volume and variation in data types.

While it may seem trivial to index historical data on the Ethereum blockchain related to digital assets, we found this to be a challenge with existing APIs as compared to traditional finance market APIs (i.e. Yahoo finance). Ethereum nodes today do not natively store transactions by wallet address, and this must be done with an indexer. Currently, the indexers that have been built to query the Ethereum blockchain have high developer friction, are fragmented, or are prohibitively expensive for this project.

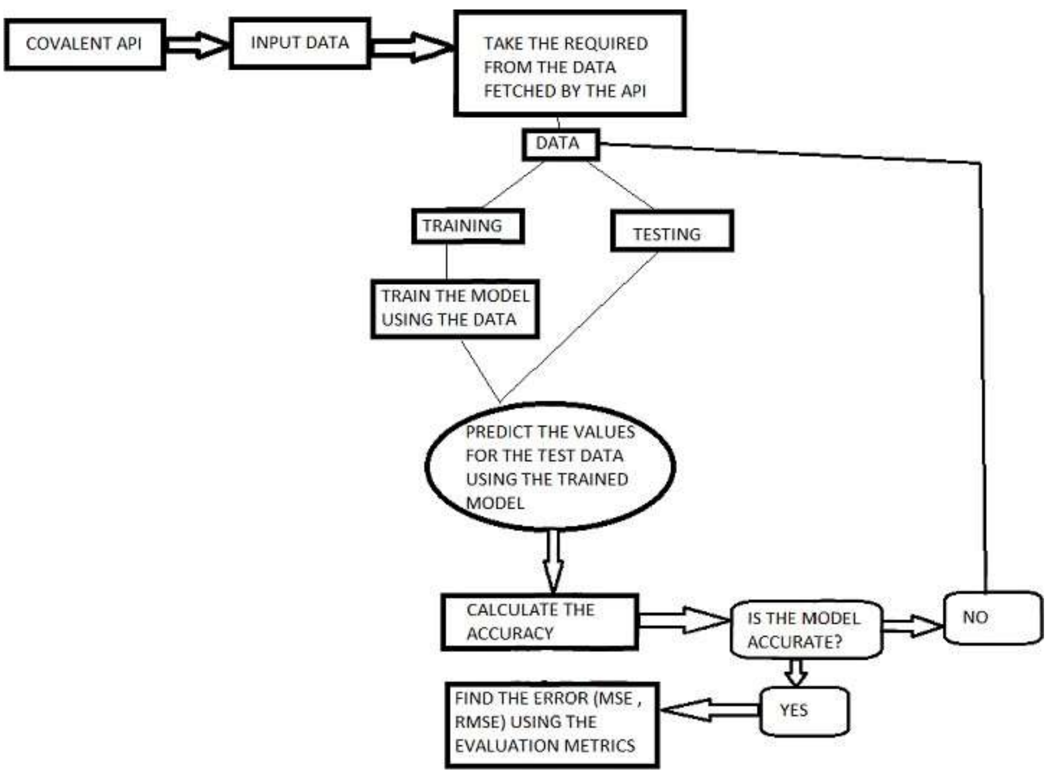
In order to get the data we were interested in gathering related to NFTs we needed to explore more APIs than we initially intended to in our project proposal. The APIs we explored include :

- OpenSea NFT API
- Covalent NFT historical data API
- Etherscan API
- Coingecko API
- Moralis API

We go through a more detailed process of how we converged to using the Covalent API in the Appendix. The Covalent API has become one of the more reputable teams working on indexing blockchain data across multiple chains (i.e., Ethereum, Binance, Polygon, Solana, Ronin, etc.). Covalent is backed by some of the top crypto venture capital firms in the world and have built a strong reputation within the crypto developer community which made our team feel more confident in the quality of data used in this project.

Due to fundamental differences between the crypto market and traditional stock market, additional adjustments were required in order to have data that is properly aligned. In particular, the NFT and crypto markets are open 24 hours a day, 7 days a week, while traditional markets are only open Monday to Friday during trading hours. In order to adapt our NFT data to be usable in conjunction with stock information, we decided to only use the NFT data associated with days that have corresponding stock data. The same adjustment was required for information that has data for every day of the year, such as Google Trends data.

ARCHITECHTURE / FLOW DIAGRAM :



WORKING :

We fetch data using Covalent Api for that we used 4 functions :

fetch_collection_hist(address){used to fetch historical prices of the nft}
fetch_token_id(address) {used to fetch the token id of all the nft in a collection}
fetch_token_tx(address, token_id) {used to identify the features of the nft art}
fetch_token_meta(address, token_id) {used to display nft art and its features}

address is smart contract address of the nft

some publically available smart contract addresses for some of the most popular nft colelcitons are :

Geisha Tea House: ox2ABb22d74Dbc2BoF3C9BAC9f173ef35DdB2C0809t
BAYC: oxBC4CAoEdA7647A8aB7C2061c2E118A18a936f13D
Cryptopunks: oxb47e3cd837dDF8e4c57F05d70Ab865de6e193BBB
Doodles: ox8a90CAb2b38dba80c64b7734e58Ee1dB38B8992e
Azuki: oxED5AF388653567Af2F388E6224dC7C4b3241C544
Deadfellaz: ox2acAb3DEa77832C09420663boE1cB386031bA17B
Gutter Cat Gang: oxEdB61f74Bod09B2558F1eeb79B247c1F363Ae452
Sup Ducks: 0x3Fe1a4c1481c8351E91B64D5c398b159dE07cbc5
Cyber Kongs: ox57a204AA1042f6E66DD7730813f4024114d74f37
Creature World: oxc92cedDtb8dd984A89tb49#376f9A48b999aAFct
Cool Cats: oxlA92f7381B9F03921564a437210bB9396471050C
World of Women: oxe785E82358879F061BC3dcAC6f0444462D4b5330
Alien Frens: oxd23d2D4aA76df5C4A19e1c9b6A83EA83f8c3db18
Lazy Lions : ox8943C7bAC1914C9A7ABa750Bf2B6B09Fd21037Eo

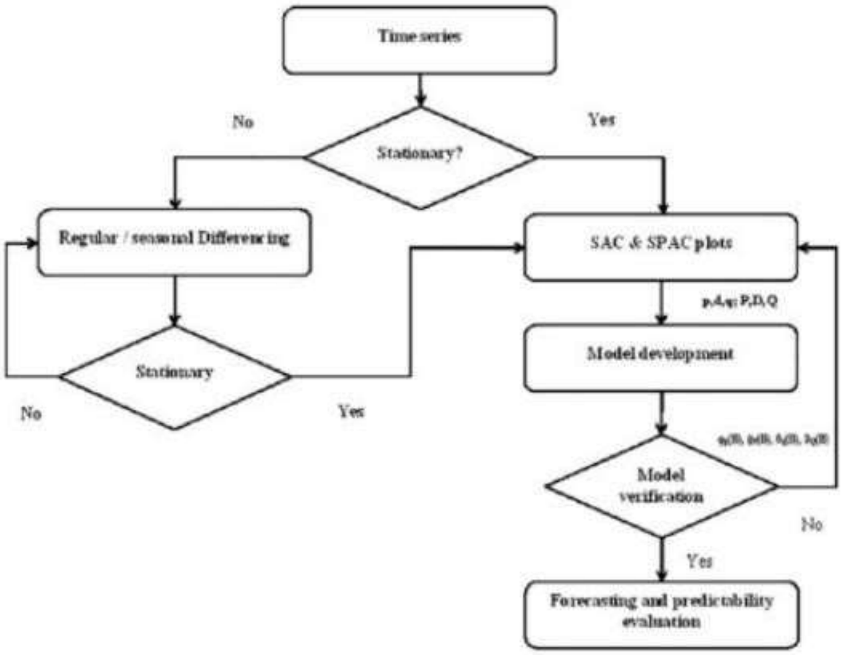
After getting the data we only extract important features such as
'opening_date', 'volume_wei_day', 'volume_quote_day',
'average_volume_wei_day', 'average_volume_quote_day',
'unique_token_ids_sold_count_day', 'floor_price_wei_7d',
'floor_price_quote_7d', 'gas_quote_rate_day', 'quote_currency'
From the taken data And take the them as Inputs features.

The data format is timeseries data.

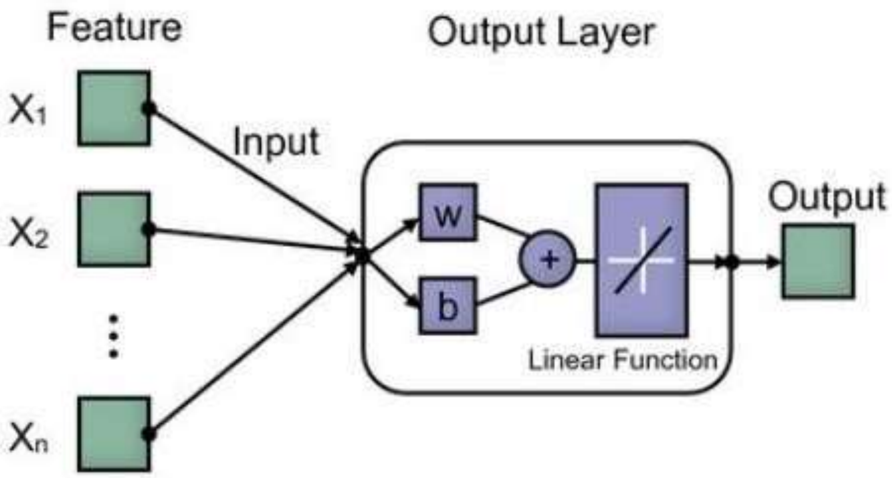
And predict the Sales as the output.

And calculate the accuracy and estimate the error for all the models (linear regression , random forest , XGBoost , ARIMA

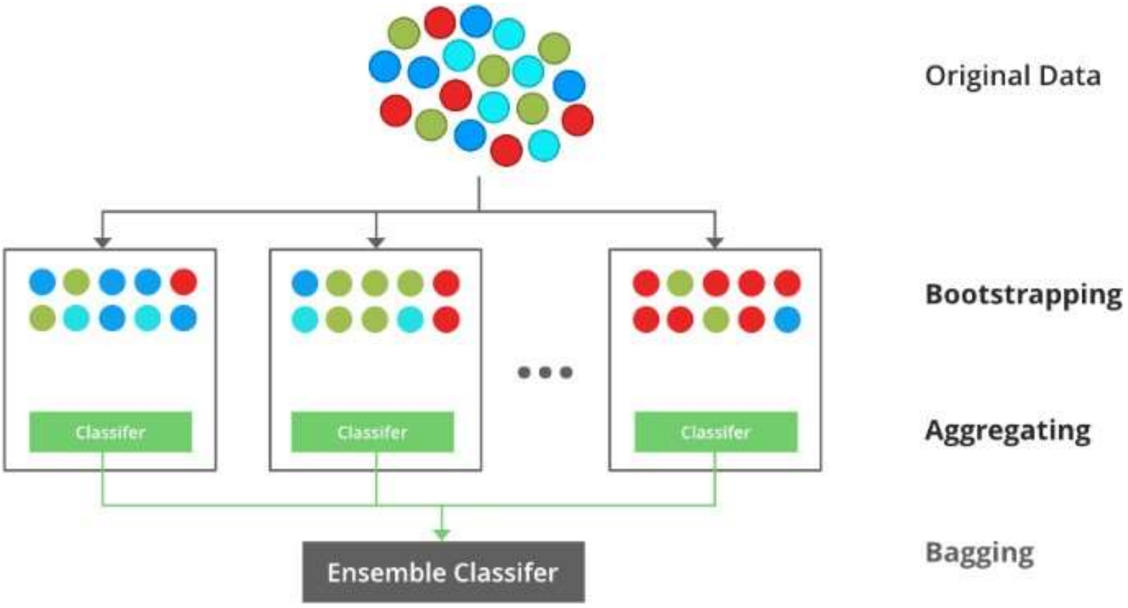
Architechture of ARIMA Model



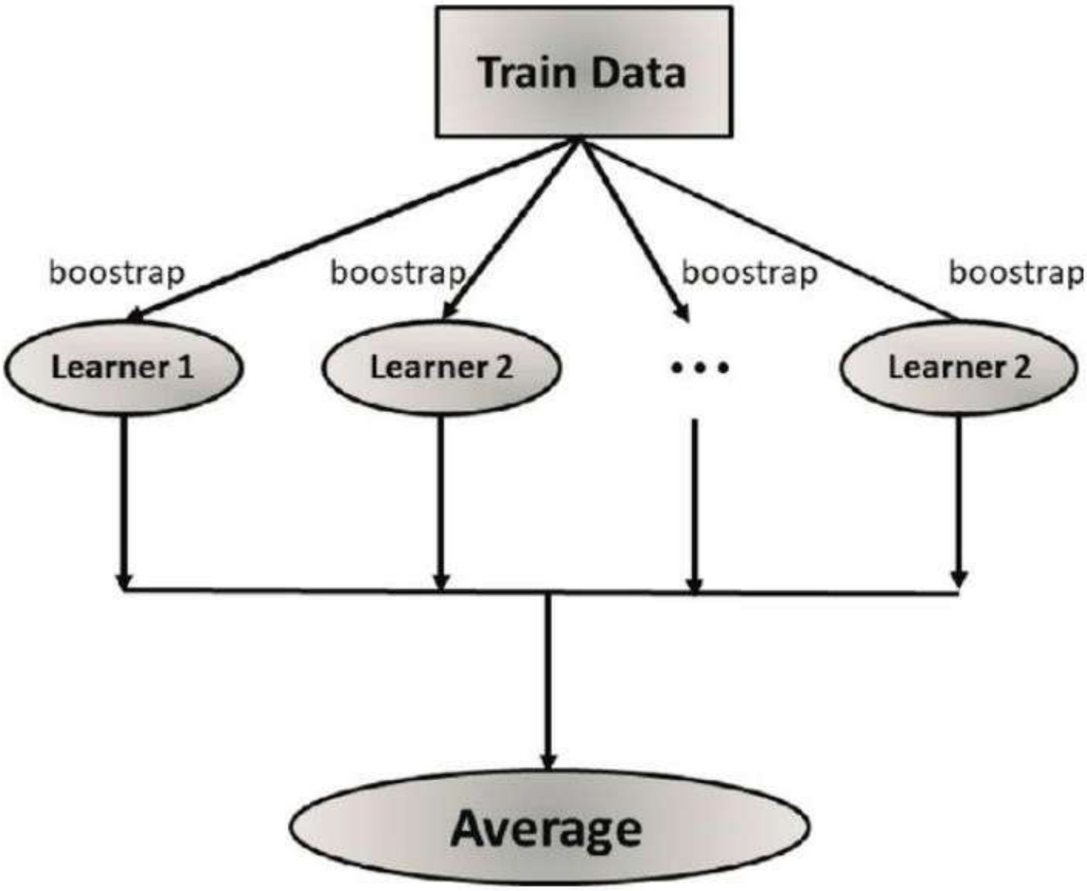
Architechture of linear regression



Architechture of XGBoost Algorithm



Architechture of random forest model :



PERFORMANCE ANALYSIS

IMPORTANCE OF PARAMETERS AND SOFTWARE TESTING DONE ON THE DATASET:

The time-series data used in this analysis are the weekly moving - average prices of NFTs, retrieved from nonfungi-ble.com, displayed on a daily basis in US dollars. We extracted the moving-average prices not only from all available NFTs6(from 2019-06-23 to 2021-12-21), but also from each of the our major projects with different time-scales and categories :

Decentraland (from 2018-03-19 to 2021-12-20),

CryptoPunks(from 2018-05-17 to 2020-12-20),

Ethereum Name Service (from2019-05-04 to 2021-12-20),

ArtBlocks (from 2020-11-27 to 2021-12-20).

X_parameters : 'opening_date', 'volume_wei_day', 'volume_quote_day', 'average_volume_wei_day', 'average_volume_quote_day', 'unique_token_ids_sold_count_day', 'floor_price_wei_7d', 'floor_price_quote_7d', 'gas_quote_rate_day', 'quote_currency'

Y_parameters : ‘Floor_price’

The dependent variable to predict is the floor price of a given NFT collection. Since the past sale price is usually a good indicator of the future price, this can be interpreted as the prediction of a time-dependent event:

Y(c|tn) = f(w,x,Y(c|t0,...,tn-1))
where Y is the floor price of collection c at time tn and x represents the time dependent independent variables; Y from t0 to tn-1 represent the past prices that can be used to model the price now.

We know the prices of a lot NFTs have sky-rocketed since last year and there is surely a trend in the price. Given most of the time-series models require stationarity (no trend) in the data, generally Y needs to be transformed into a difference in order to remove trending in the price:

dY(c|t) = f(w,x,dY(c|t0,...,tn-1))
where dY is the percentage change in price of the collection (or token) c from month t-1 to t

we split the dataset into training and testing dataset and apply partial and auto correlation on them

original dataset :

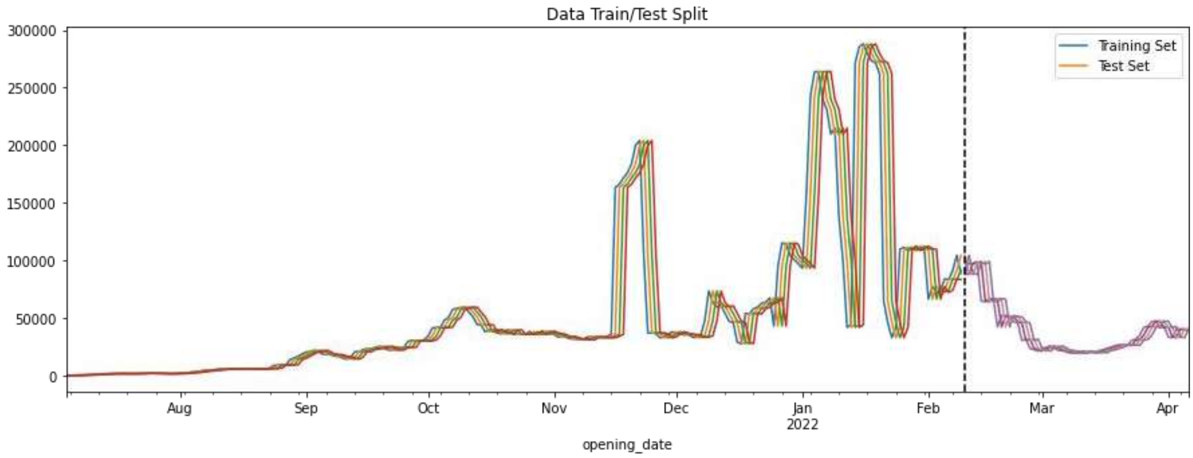
	chain_id	collection_name	collection_address	collection_ticker_symbol	volume_wei_day	volume_quote_day	average_volume_wei_day	average_volume_quote_day
opening_date								
2021-07-01	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	8099144580310900000	17967.137	33193215493077500	73.63581
2021-07-02	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	19961700243270440000	213719.800	74064685823572000	158.89948
2021-07-03	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	36350675977016800000	79322.380	69239382813365400	151.09027
2021-07-04	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	34719873640270500000	792661.060	130650145318806000	238.10495
2021-07-05	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	97001963914198400000	224008.200	166099253277737000	383.57568
...	...	...	...	...	...	...	...	...
2022-04-02	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	12082590000000000000	419521.250	8055000000000000	27968.08200
2022-04-03	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	79062250000000000000	2749465.200	15502401960784300000	52911.08200
2022-04-04	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	14236000000000000000	499238.700	8897500000000000	31202.41800
2022-04-05	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	114024963990118000000	2981498.200	1096703884522290000	38283.63700
2022-04-06	1	Cool Cats	0x1a92f7381b9f03921564a437210bb9396471050c	COOL	11875000000000000000	397669.100	913461538461539000	30589.93200

Training dataset:

opening_date	
2021-07-01	73.63581
2021-07-02	145.43475
2021-07-03	148.71074
2021-07-04	234.94214
2021-07-05	253.67885
	...
2022-02-05	83883.42000
2022-02-06	83401.99000
2022-02-07	91426.36000
2022-02-08	104512.31000
2022-02-09	88761.22000

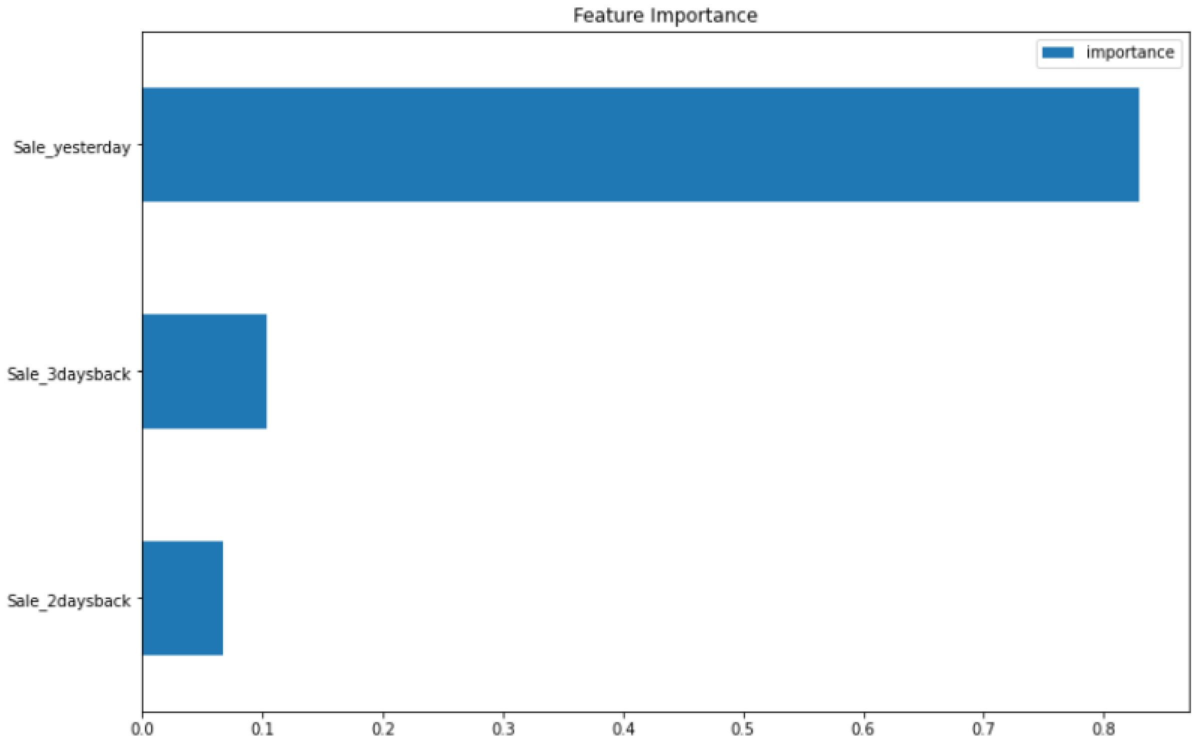
Testing dataset :

opening_date	
2022-02-10	87819.930
2022-02-11	97893.290
2022-02-12	96764.890
2022-02-13	99143.250
2022-02-14	64728.023
2022-02-15	64212.668
2022-02-16	67053.090
2022-02-17	66355.510
2022-02-18	42237.390
2022-02-19	52122.273
2022-02-20	51397.934
2022-02-21	51045.188
2022-02-22	46880.617
2022-02-23	45497.113
2022-02-24	30586.592
2022-02-25	31751.459




```
RandomForestRegressor
RandomForestRegressor(max_features=3, random_state=1)
```

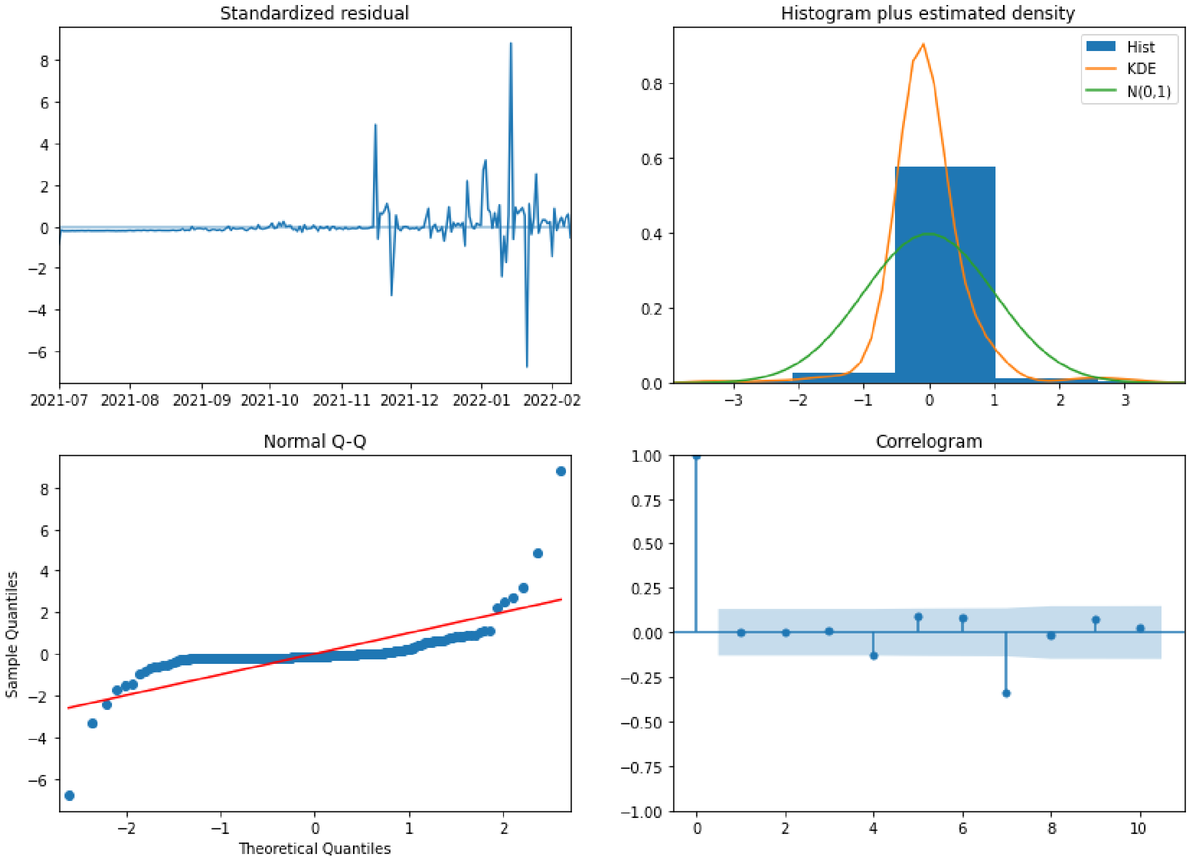
```
XGBRegressor
XGBRegressor(base_score=0.5, booster='gbtree', callbacks=None,
             colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
             early_stopping_rounds=50, enable_categorical=False,
             eval_metric=None, feature_types=None, gamma=0, gpu_id=-1,
             grow_policy='depthwise', importance_type=None,
             interaction_constraints='', learning_rate=0.01, max_bin=256,
             max_cat_threshold=64, max_cat_to_onehot=4, max_delta_step=0,
             max_depth=3, max_leaves=0, min_child_weight=1, missing=nan,
             monotone_constraints='()', n_estimators=1000, n_jobs=0,
             num_parallel_tree=1, objective='reg:linear', predictor='auto', ...)
```

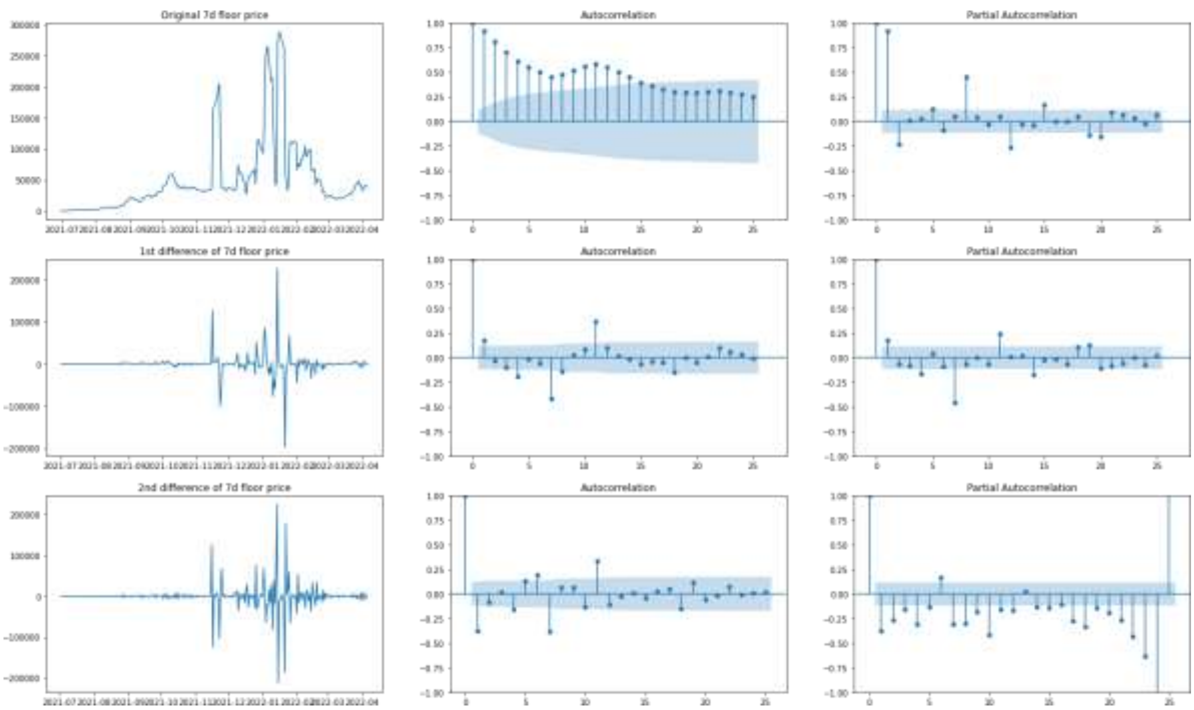


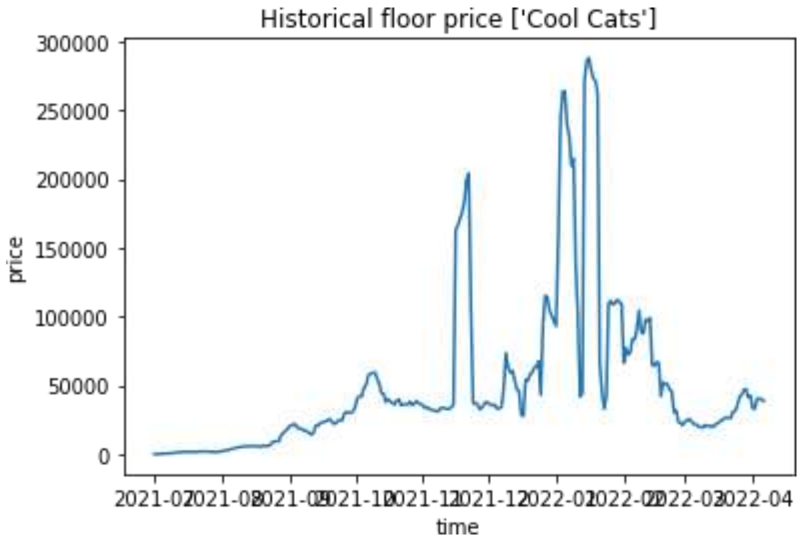
```
ADF Statistic: -2.119959
p-value: 0.236565
Time-series is non-stationary at 5% significance level. Find the order of difference!
ADF Statistic: -5.499733
p-value: 0.000002
Time-series is stationary at 5% significance level.
ADF Statistic: -7.803186
p-value: 0.000000
Time-series is stationary at 5% significance level.
```

```
Performing stepwise search to minimize aic
 ARIMA(1,0,1)(0,0,0)[0]             : AIC=5195.485, Time=0.09 sec
 ARIMA(0,0,0)(0,0,0)[0]             : AIC=5723.986, Time=0.01 sec
 ARIMA(1,0,0)(0,0,0)[0]             : AIC=5203.617, Time=0.02 sec
 ARIMA(0,0,1)(0,0,0)[0]             : AIC=5564.807, Time=0.04 sec
 ARIMA(2,0,1)(0,0,0)[0]             : AIC=5197.409, Time=0.04 sec
 ARIMA(1,0,2)(0,0,0)[0]             : AIC=5197.556, Time=0.05 sec
 ARIMA(0,0,2)(0,0,0)[0]             : AIC=5510.830, Time=0.04 sec
 ARIMA(2,0,0)(0,0,0)[0]             : AIC=5195.573, Time=0.04 sec
 ARIMA(2,0,2)(0,0,0)[0]             : AIC=5199.474, Time=0.16 sec
 ARIMA(1,0,1)(0,0,0)[0] intercept   : AIC=5191.668, Time=0.04 sec
 ARIMA(0,0,1)(0,0,0)[0] intercept   : AIC=5459.745, Time=0.04 sec
 ARIMA(1,0,0)(0,0,0)[0] intercept   : AIC=5201.267, Time=0.02 sec
 ARIMA(2,0,1)(0,0,0)[0] intercept   : AIC=5193.077, Time=0.08 sec
 ARIMA(1,0,2)(0,0,0)[0] intercept   : AIC=5193.387, Time=0.08 sec
 ARIMA(0,0,0)(0,0,0)[0] intercept   : AIC=5611.230, Time=0.01 sec
 ARIMA(0,0,2)(0,0,0)[0] intercept   : AIC=5415.021, Time=0.09 sec
 ARIMA(2,0,0)(0,0,0)[0] intercept   : AIC=5191.084, Time=0.04 sec
 ARIMA(3,0,0)(0,0,0)[0] intercept   : AIC=5193.085, Time=0.07 sec
 ARIMA(3,0,1)(0,0,0)[0] intercept   : AIC=5195.071, Time=0.18 sec

Best model:  ARIMA(2,0,0)(0,0,0)[0] intercept
Total fit time: 1.147 seconds
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  224
Model:               SARIMAX(2, 0, 0)   Log Likelihood               -2591.542
Date:                Sun, 27 Nov 2022   AIC                           5191.084
Time:                        23:19:03   BIC                           5204.730
Sample:                    07-01-2021   HQIC                          5196.592
                         - 02-09-2022
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept   5815.3949   2673.598      2.175      0.030     575.240    1.11e+04
ar.L1          1.1256      0.066     17.058      0.000       0.996       1.255
ar.L2         -0.2311      0.063     -3.641      0.000      -0.355      -0.107
sigma2      6.535e+08      0.445   1.47e+09      0.000    6.53e+08    6.53e+08
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):             13872.45
Prob(Q):                              0.93   Prob(JB):                         0.00
Heteroskedasticity (H):              49.45   Skew:                             2.33
Prob(H) (two-sided):                  0.00   Kurtosis:                        41.27
===================================================================================
```


Standardized residual
Histogram plus estimated density
Hist
KDE
N(0,1)
Normal Q-Q
Sample Quantiles
Theoretical Quantiles
Correlogram
2021-07
2021-08
2021-09
2021-10
2021-11
2021-12
2022-01
2022-02


Original 7d floor price
Autocorrelation
Partial Autocorrelation
1st difference of 7d floor price
Autocorrelation
Partial Autocorrelation
2nd difference of 7d floor price
Autocorrelation
Partial Autocorrelation


Historical floor price ['Cool Cats']
price
300000
250000
200000
150000
100000
50000
0
2021-07
2021-08
2021-09
2021-10
2021-11
2021-12
2022-01
2022-02
2022-03
2022-04
time

The NFT

Token id : “90”

The body is: blue cat skin

The hats is: bucket hat tan

The shirt is: overalls blue

The face is: sunglasses pixel

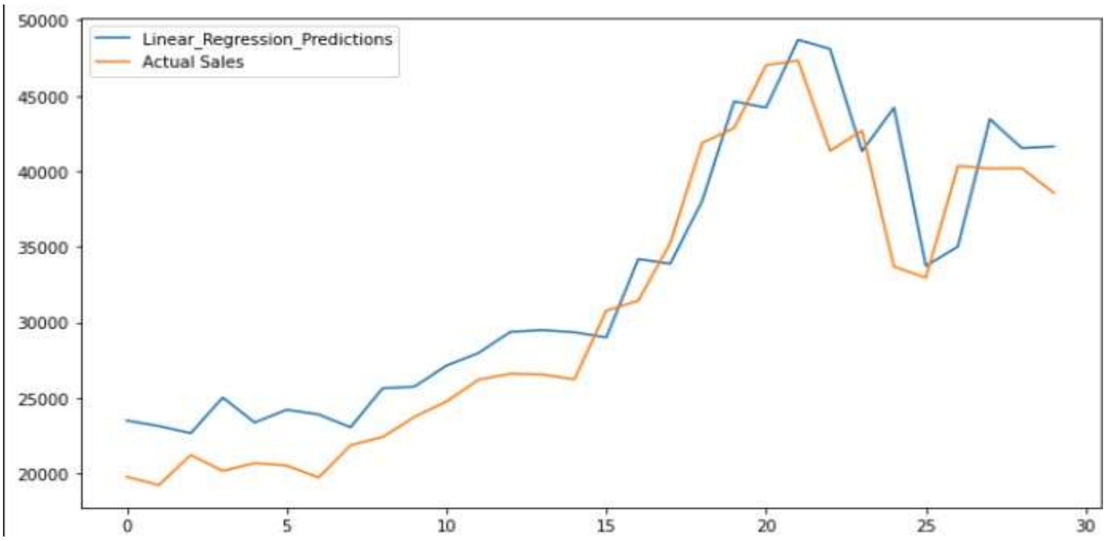
The tier is: cool_2

RESULTS OBTAINED

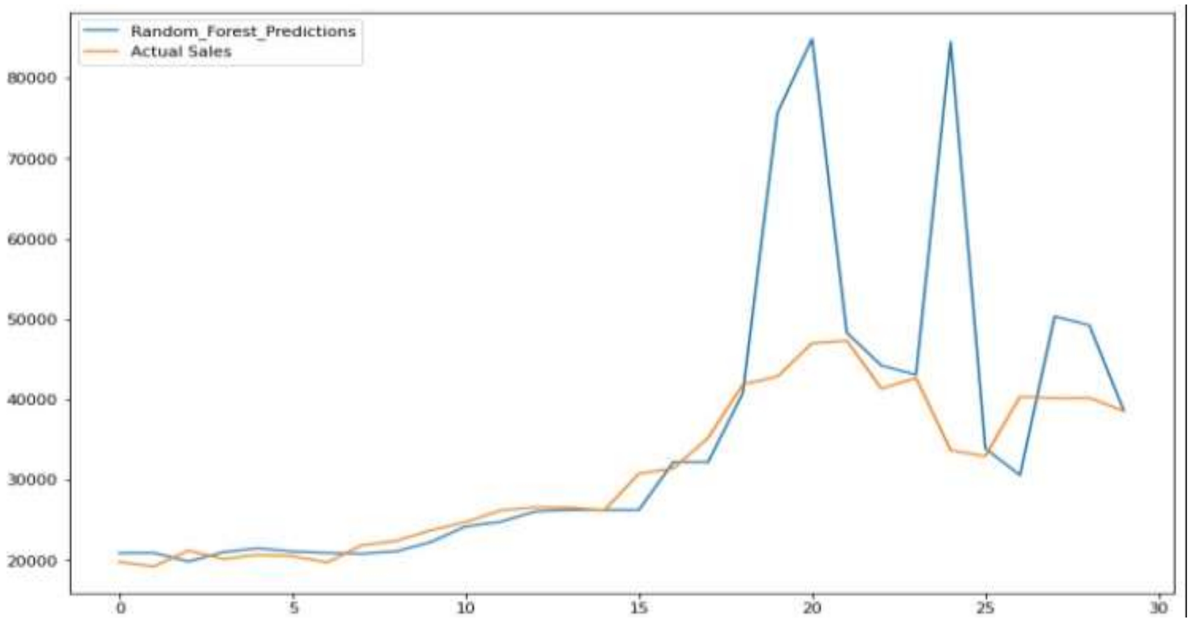
MODEL	ACCURACY	RUNTIME	EXECUTION TIME	MSE	RMSE
RAMDOM FOREST	67.07	0.5s	0.1	13099799.37592	13456.8863
ARIMA	54.11	1.2s	0.1	637977159.46739	25258.20974
LINEAR REGRESSION	71.21	0.7s	0.1	93099799.3759	9619.3644
XGBOOST	78.14	0.3s	0.1	12708487.1030445	3628.52

COMPARITIVE ANALYSIS

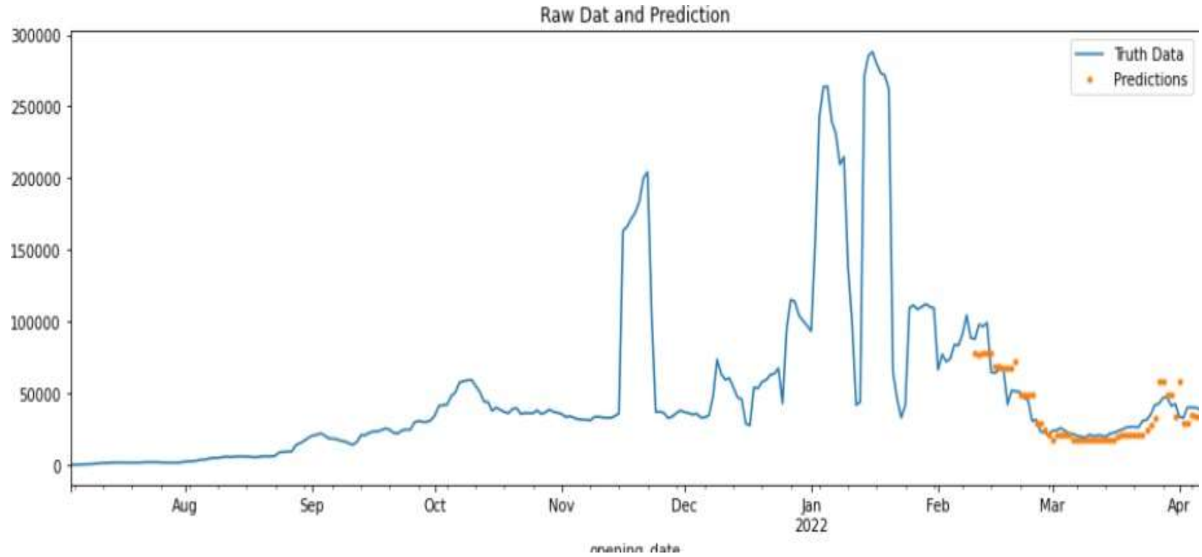
LINEAR REGRESSION :



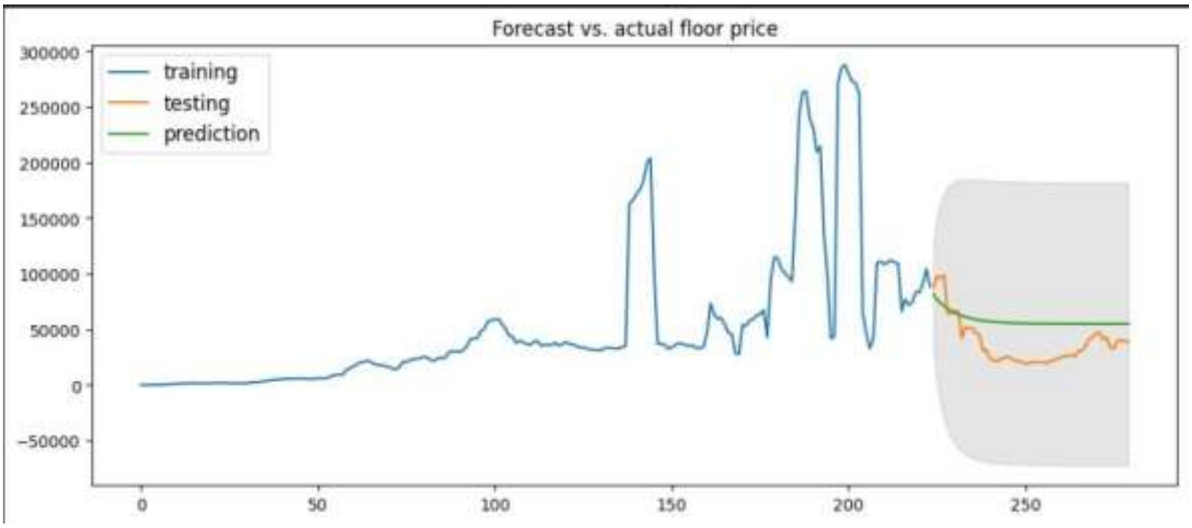
RANDOM FOREST :



XBBOOST :



ARIMA:



CONCLUSION

We have seen how influential NFTs are becoming across various different domains: art, entertainment, sports, academia, and more. In this paper, we have highlighted various methods of appraising an NFT for a specific collection. We showed how to use various different types of data sources to inform our models. We believe that we have aggregated the seminal works published thus far around NFT appraisals in the most sensible way to build a our tool using ARIMA model. Now that we have determined the key predictors that influence an NFT's price, these predictors can be fed into more advanced models to improve the quality of the predictions.

In this project, we focused on the Cool cats Collection, which is a collection of 9,999 programmatically, randomly generated NFTs. Similar to predecessors like CryptoPunks, each Cool Cat NFT is made up of a variety of unique traits.
We believe much more work can be done with language models and NFT appraisals that were out of scope for this project.

Overall, NFTs are a new tool that satisfies some of the needs of creators, users, and collectors of a large class of digital and non-digital objects. As such, they are probably here to stay or, at least, they represent a first step towards new tools to deal with property and provenance of such assets. We anticipate that our Project will help accelerate new research on NFT in a broad array of disciplines, including economics, law, cultural evolution, art history, computational social science, and computer science. The results will also help practitioners make sense of a rapidly evolving landscape and inform the design of more efficient marketplaces as well as the associated regulation.

REFERENCES

[1] "Web3." https://en.wikipedia.org/wiki/Web3 (accessed Feb. 04, 2022).
[2] N. F. T. Labs, "The social value of NFT - NFT Labs," Medium, Jul. 11, 2021. https://nftlabs.medium.com/the-social-value-of-nft-c1d47fa71074 (accessed Feb. 04, 2022).
[3] "Bloomberg." https://www.bloomberg.com/news/articles/2022-01-06/nft-market-surpassed-40-billion-i n-2021-new-estimate-shows (accessed Feb. 04, 2022).
[4] "Non-fungible tokens: can we predict the price they'll sell for?," The Alan Turing Institute. https://www.turing.ac.uk/blog/non-fungible-tokens-can-we-predict-price-theyll-sell (accessed Feb. 04, 2022).
[5] "New World Notes," New World Notes. https://nwn.blogs.com/nwn/2022/01/second-life-nft-ownership-rate.html (accessed Feb. 04, 2022).
[6] E. I. Proposals, "EIP-721: Non-Fungible Token Standard," Ethereum Improvement Proposals. https://eips.ethereum.org/EIPS/eip-721 (accessed Feb. 04, 2022).

[7] E. I. Proposals, “EIP-1155: Multi Token Standard,” Ethereum Improvement Proposals. https://eips.ethereum.org/EIPS/eip-1155 (accessed Feb. 04, 2022).
[8] “yfinance,” PyPI. https://pypi.org/project/yfinance/ (accessed Feb. 04, 2022).
[9] “NFT API Overview,” OpenSea Developer Documentation. https://docs.opensea.io/reference/api-overview (accessed Feb. 04, 2022).
[10] “The Graph,” The Graph. https://thegraph.com (accessed Feb. 04, 2022).
[11] “OpenSea,” Y Combinator. https://www.ycombinator.com/companies/opensea (accessed Feb. 04, 2022).
[12] M. Nadini, “Non fungible tokens dataset,” Aug. 2021, Accessed: Feb. 04, 2022. [Online]. Available: https://osf.io/wsnzr/
[13] M. Nadini, L. Alessandretti, F. Di Giacinto, M. Martino, L. M. Aiello, and A. Baronchelli, “Mapping the NFT revolution: market trends, trade networks, and visual features,” Sci. Rep., vol. 11, no. 1, pp. 1–11, Oct. 2021.
[14] “Website.” https://trends.google.com/trends/explore?q=ethereum&geo=CA
[15] C. Pinto-Gutiérrez, S. Gaitán, D. Jaramillo, and S. Velasquez, “The NFT Hype: What Draws Attention to Non-Fungible Tokens?,” Sci. China Ser. A Math., vol. 10, no. 3, p. 335, Jan. 2022.
[16] A. Kapoor, D. Guhathakurta, M. Mathur, R. Yadav, M. Gupta, and P. Kumaraguru, “TweetBoost: Influence of Social Media on NFT Valuation,” Jan. 20, 2022. Accessed: Feb. 04, 2022. [Online]. Available: http://arxiv.org/abs/2201.08373
[17] T. Gao and Y. Chai, “Improving Stock Closing Price Prediction Using Recurrent Neural Network and Technical Indicators,” Neural Comput., vol. 30, no. 10, pp. 2833–2854, Oct. 2018.
[18] J. Brownlee, “Linear Regression for Machine Learning,” Machine Learning Mastery, Mar. 24, 2016. https://machinelearningmastery.com/linear-regression-for-machine-learning/ (accessed Feb. 04, 2022).
[19] “Website.” https://web.s.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=84af4650-8c48-46d2 -b92b-7e38251be56e%40redis [20]A. C. Rencher and G. Bruce Schaalje, Linear Models in Statistics. John Wiley & Sons, 2008.
[21] “Using Machine Learning to Improve Pricing Forecasts,” www.micro.ai, Sep. 04, 2020. https://www.micro.ai/blog/using-machine-learning-algorithms-to-predict-pricing-trends (accessed Feb. 04, 2022).
[22]K. Kumar and M. T. U. Haider, “Enhanced Prediction of Intra-day Stock Market Using 14 NFT Appraisal Prediction (Early Work) Metaheuristic Optimization on RNN–LSTM Network,” New Generation Computing, vol. 39, no. 1, pp. 231–272, Sep. 2020.
[23]W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, “A CNN-LSTM-Based Model to Forecast Stock Prices,” Complexity, vol. 2020, Nov. 2020, doi: 10.1155/2020/6622927.
[24]A. Tondak, “Recurrent Neural Networks,” Cloud Training Program, Jan. 04, 2021. https://k21academy.com/datascience/machine-learning/recurrent-neural-networks/ (accessed Feb. 04, 2022).
[25]K. Hale, “Goldman Sachs $100K Bitcoin Endorsement Could Buoy ‘Digital Gold,’” Forbes, Jan. 11, 2022.

https://www.forbes.com/sites/korihale/2022/01/11/goldman-sachs-100k-bitcoin-endorse ment-buoys-digital-gold/ (accessed Feb. 04, 2022).
[26]N. Whittemore, “Are NFTs the True Store of Value?,” CoinDesk, Jan. 29, 2022. https://www.coindesk.com/podcasts/the-breakdown-with-nlw/are-nfts-the-true-store-of-v alue/ (accessed Feb. 04, 2022).
[27] L. Schaar and S. Kampakis, “Non-Fungible Tokens as an Alternative Investment: Evidence from CryptoPunks,” The JBBA, p. 31949, Jan. 2022.
[28]D. Fonda, “Crypto and Stocks Look Increasingly Correlated. That Has Raised Risk Fears,” Barrons, Jan. 15, 2022. https://www.barrons.com/articles/crypto-and-stocks-look-increasingly-correlated-thats-ra ised-risk-fears-51642207952 (accessed Feb. 04, 2022).
[29] “Bloomberg.” https://www.bloomberg.com/news/articles/2021-12-03/bitcoin-s-correlation-with-stocksgrows-as-risk-appetite-drops (accessed Feb. 04, 2022).
[30]B. McCullough, “A revealing look at the dot-com bubble of 2000 — and how it shapes our lives today,” ideas.ted.com, Dec. 04, 2018. https://ideas.ted.com/an-eye-opening-look-at-the-dot-com-bubble-of-2000-and-how-it-sh apes-our-lives-today/ (accessed Mar. 10, 2022).
[31] “Gated recurrent unit.” https://en.wikipedia.org/wiki/Gated_recurrent_unit (accessed Mar. 13, 2022).
[32] “Multiseries Featural LSTM for Partial Periodic Time-Series Prediction: A Case Study for Steel Industry.”
https://ieeexplore.ieee.org/abstract/document/8970298?casa_token=WqtKnHvfx4YAAAA
A:BSzpJ6Kgwjz4k19aHI2ZBVlS8kb3JsMyOIkF2STqK_BAk4vJAboZEkk1kJplbQkn0xNb6p Zz (accessed Mar. 13, 2022).
[33] “XGBoost Documentation — xgboost 1.5.2 documentation.”
https://xgboost.readthedocs.io/en/stable/ (accessed Mar. 13, 2022).
[34] “4.2. Permutation feature importance,” scikit-learn. https://scikit-learn/stable/modules/permutation_importance.html (accessed Mar. 13, 2022).
[35]L. Shukla, “Introduction to Hyperparameter Sweeps – A Model Battle Royale To Find The Best Model In 3 Steps,” Oct. 30, 2019. https://wandb.ai (accessed Mar. 13, 2022).
[36]I. M. S. M. Executive and B. P. F. of RTFKT, “The power of hype? how nfts are fuelling exclusivity - green room,” Green Room Design. [Online]. Available: https://greenroomdesign.com/blog/the-power-of-hype-the-nft-virtual-brand-activity-thathas-fuelled-the-appetite-for-exclusivity/. [Accessed: 03-Apr-2022].