arXiv:2201.08373v1 [cs.SI] 20 Jan 2022

TweetBoost: Influence of Social Media on NFT Valuation

Arnav Kapoor*[1], Dipanwita Guhathakurta *[1], Mehul Mathur *[1] Rupanshu Yadav *[2], Manish Gupta[1], and Ponnurungam Kumaraguru[1]

[1] International Institute of Information Technology, Hyderabad
[2] Indraprastha Institute of Information Technology, Delhi

Abstract. NFT or Non-Fungible Token is a token that certifies a digital asset to be unique. A wide range of assets including, digital art, music, tweets, memes, are being sold as NFTs. NFT-related content has been widely shared on social media sites such as Twitter. We aim to understand the dominant factors that influence NFT asset valuation. Towards this objective, we create a first-of-its-kind dataset linking Twitter and OpenSea (the largest NFT marketplace) to capture social media profiles and linked NFT assets. Our dataset contains 245,159 tweets posted by 17,155 unique users, directly linking 62,997 NFT assets on OpenSea worth 19 Million USD. We have made the dataset public.[3] We analyze the growth of NFTs, characterize the Twitter users promoting NFT assets, and gauge the impact of Twitter features on the virality of an NFT. Further, we investigate the effectiveness of different social media and NFT platform features by experimenting with multiple machine learning and deep learning models to predict an asset's value. Our results show that social media features improve the accuracy by 6% over baseline models that use only NFT platform features. Among social media features, count of user membership lists, number of likes and retweets are important features.

1 Introduction

Blockchain has emerged as a core disruptive technology that has transformed the financial ecosystem. The origin of Blockchain can be traced back to a 2008 whitepaper [10] published under the pseudonym Satoshi Nakamoto, who introduced blockchain in the context of the most popular crypto-currency, Bitcoin. Bitcoin uses blockchain technology to develop the public distributed ledger used to record the transactions on its network. The growing interest in Blockchain technology especially its use in the financial domain from both retail and institutional investors[4] has led to several new products emerging in the crypto-sphere to find the 'next big thing'.

One such emerging Blockchain product that has captured large public attention is Non-Fungible Tokens or NFTs. An NFT is a token that certifies a digital asset to be unique. NFTs use blockchain to store anything that can be converted into digital files, for example, images, music, and videos. Blockchain technology enables the association

---

*∗ denote equal contribution

[3]https://www.dropbox.com/sh/5ubvrrpziwbxxkt/AAAMO737hMxu_4Ykhv8MQC8ka?dl=0

[4]https://www.forbes.com/sites/lawrencewintermeyer/2021/08/12/institutional-money-is-pouring-into-the-crypto-market-and-its-only-going-to-grow/

of proof of ownership to the digital asset. NFTs have grown exponentially in 2021; this phenomenal growth has led to traditional auction houses being receptive to digital art NFTs. On March 11, 2021, *'Everydays: The First 5000 Days'*, an NFT artwork of the prominent digital artist Beeple, sold at Christie's for over $69 million.[5] Jack Dorsey (CEO of Twitter) raised an astounding $2.9 million for charity by auctioning his first tweet as an NFT.[6] All these sales demonstrate the adoption of NFTs by the mainstream community. The highly volatile NFT prices and sudden popularity led many people to create NFTs and in turn, promote and sell them for a profit. The most significant impact of NFTs has been on how they transformed the art world [8,15]. NFTs have allowed artists to sell their art outside the gate-keeping systems and taste-hierarchies [8].

NFT is a token representing digital assets stored on the blockchain with proof of ownership. Smart contracts [2] allow to transfer and retrieve information about an NFT through function calls. Some function calls such as 'transfer' are restricted to the owner of the NFT. Since the underlying blockchain is decentralized, no one can change the state of the contract or the NFT without making these function calls, ensuring high levels of security.

Most NFTs are digital, meaning that consumers do not receive any physical items when they purchase them. In most circumstances, the NFT is only a proof of ownership, not of copyright; i.e., the owner does not have exclusive access to the content of NFTs. For example, 'disaster girl', a popular Internet meme, was sold as an NFT for $495,000 even though the exact image in the NFT is freely available and distributed throughout the Internet.[7] The value of the NFT came from the fact that it was sold by the girl featured in the meme. The same meme sold as an NFT by other people on the marketplace did not receive any traction.[8] Interestingly, most NFTs sold online can be downloaded and shared publicly for free. The value of an NFT is based on the perception of buyers which arises from the recognition of the creator and the overall marketing around the NFT itself. Further, unlike company stocks or crypto-currency exchanges, which are traded at regular intervals, NFTs have sporadic sales. The transaction history of an NFT spans varied durations, and the owner changes in each sale. A single sale cannot account for the overall value of an NFT as the next buyer may be willing to pay much more or less than the previous sale amount depending on their current perception of the NFT. Hence, we use the average of all sales to assign a value to the NFT. The tremendous volatility of crypto-currencies, the lack of any tangible asset and the speculative marketspace makes the asset valuation task for NFTs an extremely challenging one. Unlike traditional assets, NFT asset valuation cannot be modelled directly as a mathematical economic system, but rather as a social phenomenon involving marketing schemes and the recognition and popularity of the NFT.

---

[5] https://www.theverge.com/2021/3/11/22325054/beeple-christies-nft-sale-cost-everydays-69-million

[6] https://www.cnbc.com/2021/03/22/jack-dorsey-sells-his-first-tweet-ever-as-an-nft-for-over-2point9-million.html

[7] https://www.nytimes.com/2021/04/29/arts/disaster-girl-meme-nft.html

[8] https://opensea.io/assets/0x495f947276749ce646f68ac8c248420045cb7b5e/
81178342661852467359012497214474763787141141498310730289985511697868488966145

There are various marketplaces to buy and sell different categories of NFTs, such as OpenSea, Rarible, SuperRare, NiftyGateway and Foundation.[9] OpenSea is the largest and most popular of all such marketplaces, with over 300,000 users with $3.4 Billion volume traded in August 2021 alone. Hence, we chose OpenSea to understand and model the NFT market.

To explore the branding and context around the NFT, we look at social media, particularly Twitter as a vehicle for building public perception and attracting potential buyers. Fig. 1(a) shows a popular tweet and the corresponding NFT asset sold for over $150,000 on OpenSea. Several instances of well-known personalities like Mark Cuban, Jack Dorsey, etc. on Twitter selling high-priced NFTs indicate that social media reach can play a role in influencing asset value. More than 70% of the total traffic from social media on OpenSea is from Twitter.[10]. Thus, we focus on Twitter to understand the influence of social media on the NFT market. We aim to assess if an asset is overvalued or undervalued based on our current asset valuation framework.

In this paper, we aim to answer the following research questions:

- RQ1. What is the relationship between user activity on Twitter and price on OpenSea?
- RQ2. Can we predict NFT value using signals obtained from Twitter and OpenSea; and identify which features have the greatest impact on prediction?

We collect tweets announcing NFTs (Fig. 1a) and follow the linked OpenSea URLs to extract NFT sales information and metadata from OpenSea as show in Fig. 1b. In addition, we crawl NFT images to analyse whether the image in itself exerts an influence on the price. Analysing growth and price trends on OpenSea against popularity metrics from Twitter reveals the possible significance of social media features on NFT value. We motivate average selling price as the metric to assign value to an NFT due to limited sales and highly volatile prices of NFT. We further develop and analyse predictive models using Twitter, OpenSea and NFT image features to assess their impact on asset value. Overall, we make the following contributions in this paper:

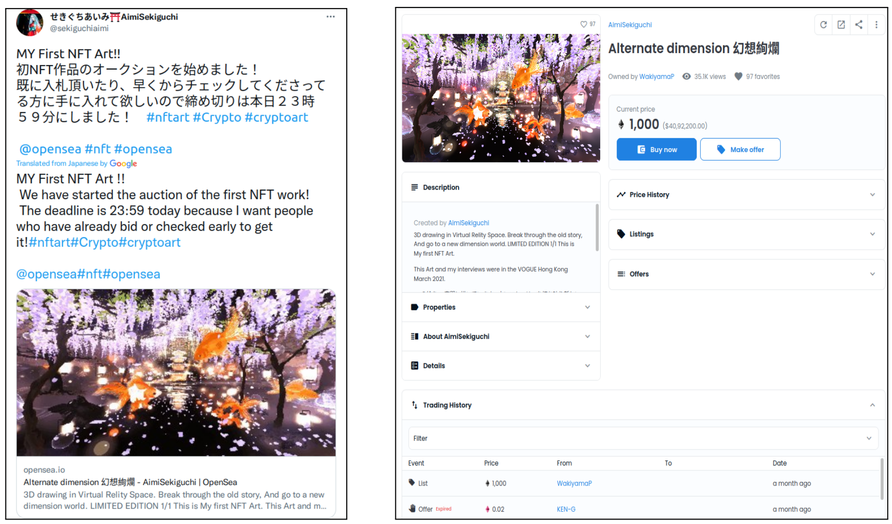
- To the best of our knowledge, we create the first ever data set for NFTs from OpenSea and their corresponding tweets. We have made the dataset public[3] in adherence to the FAIR principles as described in Section 5.1.
- We build models to predict NFT asset value using features from both OpenSea and Twitter, as well as the NFT image itself. Our best model comprising of an ensemble of Twitter and OpenSea features shows an accuracy of 69.5% in a 6-class classification setup.
- We show that social media features like *listed count*, *tweet-likes* and *number of retweets* influence the model output. We also show how branding and metadata (Twitter and OpenSea features) have stronger predictive powers than the NFT product itself. (Image features)

In Section 2, we discuss related work on NFTs and asset valuation in the financial domain. We motivate and present our asset valuation problem setup in Section 3. Sec-

[9] https://opensea.io, https://rarible.com, https://superrare.com, https://niftygateway.com, https://foundation.app

[10] https://www.similarweb.com/website/opensea.io/#social

tion 4 briefly introduces the OpenSea platform and its features, as well as blockchain-specific terms used in our analysis. We discuss our data collection pipeline and analyse the impact of Twitter on the OpenSea marketplace from the temporal dimension and value perspective in Section 5. In Section 6, we discuss multimodal models for NFT asset valuation. Finally, we conclude with limitations and future work in Section 7.



(a) Tweet announcing an NFT (b) Corresponding OpenSea listing page

Fig. 1: One of the most-liked tweets in our dataset from Japanese VR artist Aimi Sekiguchi and its corresponding page on OpenSea. This 14-second NFT video sold for over $150,000 on March 24, 2021.

2 Related Work

Since NFTs are a recent phenomena, there is limited research on NFT related data. Most of the previous work on NFTs is on analysis of relationship between blockchain, crypto-currency and NFTs. Recent work has analyzed the protocols, standards, desired properties, security, and challenges associated with NFTs [5,14]. The sudden rise of NFTs sparked an interest to find a correlation with more traditional crypto-assets. There has been a focus on understanding the correlation between the entire NFT market with Ethereum and Bitcoin [1] and also specific collections like Axie, CryptoPunks, and Decentraland [7]. These sub-markets have millions of dollars traded every day [7], which has led to critical analysis of these individual markets. Other studies focus on fairness in NFT submarkets like CryptoKitties [13].

To the best of our knowledge, there is limited work on NFT asset valuation, which is the focus of this work. Recently, Nadini et al. [9] used simple machine learning algorithms to develop a predictive model using sale history and visual features, but ignore social media features. Social media features have helped predict the prices of assets in traditional markets [4,12] and predict stock prices [4,11]. Hence, in this work, we focus on understanding impact of OpenSea and social media features for this task.

3 The NFT Asset Valuation Problem

NFT markets are highly illiquid in nature, which means sale prices are very volatile and irregular. Similar to physical art collections, the number of transactions per NFT is low as the buyers and sellers are a small niche of collectors. A traditional price prediction setting is not feasible or robust for each NFT as we do not have consistent periodic sales. In our dataset as well, a very small percentage (0.9%) of the total assets had more than 5 sales. Price prediction models used for stock or cryptocurrencies use a large number of historical data points gathered at regular intervals which is not available for an illiquid market like NFT.

We propose an asset valuation task instead of a price prediction task to overcome the challenge of an illiquid market. We define asset value as the average selling price of an asset over all its historic sales. This compensates the large volatility in selling price, and provides a better indicator of asset valuation. Our objective is to provide a qualitative assessment of the NFT's value and identify whether the Twitter reach of the NFT influences the value. We thus divide the NFTs into multiple asset classes (Unprofitable Asset, Very Low Value Asset, Low Value Asset, Low-Medium Value Asset, Medium-High Value Asset, High Value Asset, Very High Value Asset) based on the average sale price. The maximum average price was found to be orders of magnitude more than the minimum, hence the asset classes were binned into logarithmic divisions. Henceforth, we use the term 'asset value' to refer to the average selling price of the NFT.

4 Primer on OpenSea Platform

OpenSea is the first and largest peer-to-peer marketplace for NFTs. It has attracted traders to trade assets, creators to launch their portfolios and developers to build integrated marketplaces for their applications.

The primary product on OpenSea is called the "asset" which is a unique digital item stored as an NFT on the Blockchain. The transactions and ownership of the asset is programmed in a Smart Contract to store the link to the image, music or video in its metadata. Each asset is uniquely identified by its parent contract's *address*, unique *token id* and needs to be listed on OpenSea by the creator to be available for sale. OpenSea permits the following transactions on an asset: (1) listing NFT at a fixed price, (2) listing NFT at first price auction or a Dutch auction, (3) or direct offers from buyers.

Once an asset is sold to a buyer, the buyer can resell it; thus, the asset keeps changing owner and price over time. Assets on OpenSea can be further grouped into *collections*. Collections are a group of homogeneous assets sharing common traits and properties. For example, the OpenSea page for one of the most popular collections, CryptoPunks (Fig. 2) contains similar assets with small variations.

Most transactions on OpenSea are done on the Ethereum blockchain. Since every transaction on Ethereum requires a transaction fee, called gas-fee, accepting offers and buying assets on OpenSea are also associated with a gas-fee in addition to asset price.

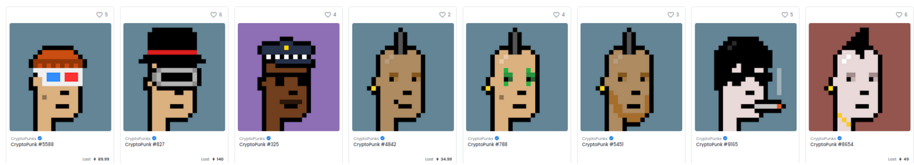


Fig. 2: CryptoPunks collection: A set of 10,000 assets showing similar traits.

5 Data Collection and Analysis

5.1 Data Collection

We collect data from two platforms: Twitter and OpenSea marketplace. We use the unique URL of the OpenSea asset to link these two data sources. Our data collection pipeline is illustrated in Figure 3.

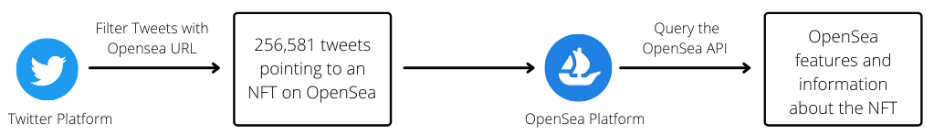


Fig. 3: Pipeline for the data collection process.

Twitter Data First, we curate 245,159 tweets from Jan 1, 2021, to Mar 30, 2021 that contain an opensea.io NFT asset link. A total of 17,155 unique users posted these tweets. Multiple tweets can reference the same OpenSea link. Our dataset contains 62,997 unique OpenSea assets belonging to 16,001 unique collections. We also collect additional information about the tweet (number of likes, number of retweets, tweet timestamp) and the source of the tweet (number of followers, number of followings, bio, account creation date). The additional meta-information allows us to model and understand the users posting about NFTs on Twitter.

OpenSea Data Apart from the social media information extracted from Twitter, the OpenSea platform also provides us with many valuable features about the asset. We keep only those assets created between Jan 1, 2021, to Mar 30, 2021 (same as our tweet collection period). We used the OpenSea API[11] to extract this additional information about the asset and its collection. Besides these features, we also crawl the associated NFT image. If the asset is a video or a GIF, we extract and use the first frame only. Overall, the dataset contains 62,997 images corresponding to unique assets.

FAIR Dataset Principles The gathered data consists of publicly available information about a social network, gathering and examining which would provide significant insights into the platform's characteristics. Our dataset also conforms to the FAIR principles. In particular, the dataset is "findable", as it is shared publicly. This dataset is also

[11] https://docs.opensea.io/

“accessible”, given the format used (CSV) is popular for data transfer and storage. This file format also makes the data “interoperable”, given that most programming languages and softwares have libraries to process CSV files. Finally, the dataset is “reusable”, as the included README file explains the data files in detail. The data was collected through public API endpoints of OpenSea and Twitter, adhering to their privacy policy. The data we collected was stored in a central server with restricted access and firewall protection.

5.2 Twitter and OpenSea Interaction Analysis

We first study the interaction between user activity on the OpenSea and Twitter platform and its influence on the asset value. We perform temporal analysis of our dataset across both the platforms. We also perform a basic correlation analysis of signals like average number of followers with asset price.

Correlation between NFT popularity across platforms We measure the NFT popularity on Twitter by aggregating features of all tweets that mention the asset. We have 245,159 tweets in our dataset, with the majority 89% of them being made in March. We plot the daily number of tweets and the NFT asset creation dates in Fig. 4. We observe a strong correlation between the two timeseries, indicating that users post on Twitter soon after creating an NFT on OpenSea. The Spearman’s correlation coefficient (ρ) for the two timeseries is 0.85 (p-value$<$0.001), showcasing the strong positive correlation. More than half (54.6%) of the tweets are posted less than a day after the NFT creation.

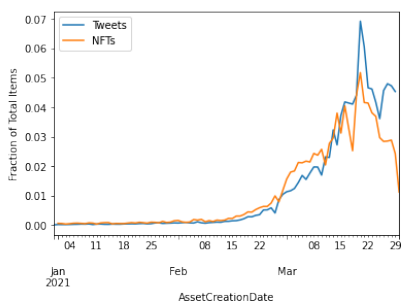


Fig. 4: Fraction of total tweets and NFTs created daily over the three month period from 1st Jan to 30th Mar 2021 versus asset creation dates.

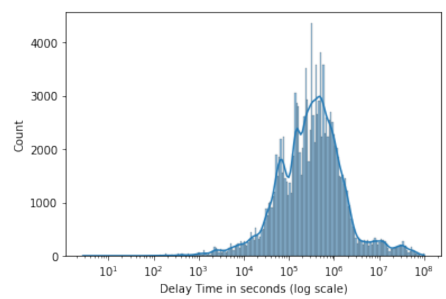


Fig. 5: Histogram of delay in seconds between creation of NFTs and the corresponding tweet. The graph follows an approximate log normal distribution.

Delay between posts on Twitter and OpenSea Next, we analyzed the time delay between the NFT creation date on OpenSea and its mention on Twitter. We plot the histogram of the delay between asset creation on OpenSea and its promotion on Twitter in Fig. 5. The distribution of the time delay approximately follows a log normal distribution. Such distribution has been observed in other studies on online information spread [6] and inter-activity times [3]. The inter-activity time is the duration between two consecutive tasks, like addition of followers on social media or sending of emails [3]. We also analyzed the impact the delay can have on the asset value. However, we found no significant correlation (ρ <0.05) between asset value and time-delay.

Twitter Username Analysis Next, we characterize the users on Twitter that post about these NFTs by analyzing their usernames. Twitter username is a strong indicator of affinity towards a cause or an organization. We computed character n-grams from usernames and manually inspected the most frequent ones. The most frequent relevant 3-gram turned out to be 'nft' which was present in 7.6% of the usernames. The other significant n-grams that came up were 'crypto', 'collect', and 'design'.

Further, we partition users into two buckets: NFT affiliated (having 'nft' in their username) and non NFT-affiliated, and check their account creation dates. Fig. 6 shows that a large proportion of NFT affiliated accounts were created in the first quarter of 2021 indicating that they were specifically created to promote and push NFT related content. Also, over 60% of all NFT-affiliated accounts were created in March 2021, compared to only 18% of non NFT-affiliated accounts. We also found that the mean value of assets promoted by NFT-affiliated usernames was marginally more significant than those by non NFT-affiliated usernames.

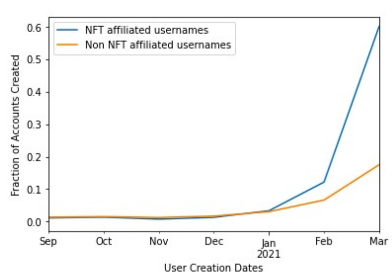


Fig. 6: Creation date of NFT affiliated and non NFT affiliated accounts. We see that a far higher percentage of NFT affiliated accounts were created in 2021 compared to non NFT affiliated accounts.

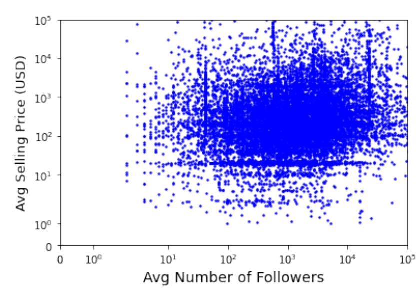


Fig. 7: Scatter plot for average follower count and asset value. We observe a weak positive correlation which suggests that an increase in the number of followers leads to a greater asset value.

Asset Value Analysis To understand the relationship between the asset value and the popularity of the user, we plot the average number of followers (our proxy for user popularity) and the asset value of NFT in Fig 7. Since many users can promote an NFT, we took the average follower count of the users who tweeted about each NFT. Both asset value and number of followers are highly skewed; hence we create the log-log plot. Note that we filtered out the points with asset value and follower count less than one.

The Spearman's coefficient between the follower count and asset value on the log scale is 0.20, which indicates a weak positive correlation. The correlation, albeit weak, leads us to believe that social media features can help in asset value prediction. The following section, discusses how we leverage features across the Twitter and OpenSea platforms for accurate asset value prediction.

6 Models for Asset Valuation

We trained multiple machine learning and deep learning models using a mix of Twitter, OpenSea and Image features to predict the value of an NFT asset. We model the NFT asset valuation problem in two ways - a binary and a multi-class classification problem.
Binary Classification - Our first objective is to gauge if selling an NFT can be a profitable venture. In cases when the NFT remains unsold, or is valued at a nominal amount, users are not able to cover the mandatory gas fees. In our dataset, 78% of assets were unsold or sold for less than \$10. They form the loss bearing NFT class (class 0), while the rest of the assets sold for more than \$10 form the profitable NFT class (class 1).
Multiclass Classification - Next we consider only the profit bearing class and discretize it into further subclasses. The maximum selling NFT was found to be orders of magnitude more expensive (\$552,621) than \$10. We thus propose a multi-class classification, with 5 logarithmically binned classes by powers of 10. For example, Class1 includes assets selling between \$10 and \$100, Class 2 between \$100 and \$1000, and so on. Of the 14,814 total assets with sales >10, distribution per class is as follows: Class1 (4000), Class2 (8391), Class3 (2189), Class4 (211), Class5 (23).

We use 77 Twitter features, 19 OpenSea asset features and 25 OpenSea collection-level features to tackle the asset valuation problem. The salient features from all three aspects are captured in Table 1. We tried both simple machine learning models like SVM, Logistic Regression, and Decision tree-based ensemble models like Random Forests and XGBoost. Across the board, for both tasks, XGBoost Classifier (Gradient Boosted Decision Trees) performed the best. We use the average gain across all splits where the feature was used as the metric for assigning feature importance in case of XGBoost Classifier. All the results in the following sections use the XGBoost classifier and the average gain as the feature importance metric.

Feature	**Type**	**Description**
listed count	Twitter (Account Level)	The number of accounts who have added the account to their lists
has nft in username	Twitter (Account Level)	True if any of the users who tweet about the NFT asset have 'nft' in their screenname or username
n_likes	Twitter (Tweet Level)	The number of likes on the tweet
n_hastags	Twitter (Tweet Level)	The number of hashtags used with the tweet
average price collection	OpenSea (Collection Level)	The average price of the other assets in the same collection
average volume traded	OpenSea (Collection Level)	The average traded volume of other assets in the same collection
is presale	OpenSea (Asset Level)	If the asset is available for pre-sale
verified asset	OpenSea (Asset Level)	If the asset is asset is verified by OpenSea
bid withdrawn	OpenSea (Asset Level)	Binary label if a bid has been revoked for the asset
bid entered	OpenSea (Asset Level)	Binary label if a bid has been placed for the asset

Table 1: Salient features across both Twitter (account level and tweet level) and OpenSea (asset level and collection level).

6.1 Twitter Features

In our dataset, 17,155 users made a total of 245,159 tweets mentioning 62,997 OpenSea NFT assets. Multiple tweets could mention the same assets; we thus aggregate the tweet level properties such as likes, replies, and retweets. Similarly, each NFT could be mentioned by multiple users; hence user level Twitter attributes like followers counts, listed count, favorites were also aggregated across users. We obtained an accuracy of $\sim$84%

for the binary classification and $\sim$67% for the multiclass classification problem using XGBoost. We observe that features like listed count and presence of NFT in Twitter username had the most significant influence on the model prediction.

6.2 OpenSea Features

We retrieve features about the NFT from the OpenSea platform. Since every NFT asset is part of a collection; we have two different sets of features on OpenSea. First is the asset level features, which are properties of a specific NFT. Second is the collection level features; this captures information about other assets in the same collection. The price of other items in the collection is a strong predictor, as prices across the same collection tend to remain in a similar range. In addition, different collections usually contain objects with homogeneous visual features [9].

Collection Level - The model performance using only OpenSea collection features was significantly better than using just Twitter features. XGBoost classifier shows an accuracy of 87% for the binary classification and 76% for the multiclass classification problem. We found that the collection value, i.e., the mean asset value of all the assets in the collection, is the most significant feature. This highlights that the value of existing assets in the same collection is a good indicator to predict value of newer assets.

Since collection level features are aggregation of all the assets in a collection, they are unavailable for newly created collections. Additionally for collections with very few assets the collection level features are not robust enough to help in the prediction. We thus switch and focus on only asset level features from now onwards.

Asset Level - We use asset level features like asset creation date, bid withdrawn and bid entered. Additionally, we engineer and gather several asset features like number of sales, number of bids and number of offers based on events like auctions, sales, bids, transfers, etc. using the events endpoint of the OpenSea API. The classifier provided an accuracy of 63% for the multi-class classification problem.

Table 2 shows that the binary/multiclass accuracy using various feature sets. Since the asset level features include signals like number of sales, using these features for binary classification leads to almost perfect accuracy values. Thus, we report only multiclass classification accuracies when using OpenSea asset features. We find that multiclass classification accuracy is $\sim$13% lower with asset features compared to using collection level features alone. It is interesting to note that even Twitter features alone could predict the asset value better than the inherent asset features. This indicates that social media features highly influence the value of an NFT.

6.3 Twitter and OpenSea Ensemble

Finally, we combine Twitter features along with OpenSea asset features. The ensemble model of Twitter + OpenSea Asset features performs better than each of them individually with a final accuracy of 69.5% for the multiclass classification task, while the OpenSea Asset features alone show an accuracy of 63.44%. There is a significant improvement of over 6% in accuracy compared to using only OpenSea asset features. We computed feature importances and found that both the Twitter and the OpenSea features appear in the top few.

Feature Set Used	Binary Accuracy	Multiclass Accuracy	Binary F1	Multiclass F1
Twitter	83.75	66.65	82.00	63.19
OpenSea (Asset)	-	63.44	-	58.49
OpenSea (Collection)	87.44	76.34	87.22	75.59
Twitter + OpenSea (Asset)	-	69.46	-	67.09
Twitter + OpenSea (Asset) + OpenSea (Collection)	88.17	79.26	88.05	78.64

Table 2: XGBoost classifier accuracy scores for different combination of features sets. Twitter features combined with OpenSea (Asset) increased accuracy by 6% showing the importance of Twitter features.

6.4 Image Features

Besides Twitter and OpenSea asset and collection features, we also attempt to capture the influence of the NFT image in determining the asset value of the NFT on OpenSea. We predict asset value using Neural Network classifiers.

Model Architecture: We experiment with various architectures such as ResNet-50, ResNet-101, AlexNet, DenseNet-121 and ResNeXt-50 for the Convolutional Neural Network-based classifier. For each architecture, we use the Adam optimizer with a learning rate of 0.001. The images are augmented using Random Affine and Random horizontal flip to avoid model overfitting. The image pixel values are normalized and scaled to values between -1 to 1 and are grouped into batches of size 128. We use the cross-entropy loss.

We obtain accuracies of 57.71%, 58.28%, and 58.74% with pretrained ResNet-50, ResNeXt-50, and ResNet-101 models respectively, for the multiclass classification task, which is lower than that of the Twitter model (66.65%) from Section 6.1 as well as OpenSea asset model (63.44%) from Section 6.2. The ResNet-101 model has an accuracy of 79.01% in the binary classification setting. The confusion matrix for the multiclass classification task using the ResNet-101 model is depicted in Table 3. From the confusion matrix we observe that the model learns to classify most assets into the largest class (Class 2). Regardless of the architecture used, the models using image features display accuracies and F1-scores lower than those using Twitter and OpenSea features described in Table 2 under identical training and validation settings. On adding image features to the Twitter-OpenSea ensemble from Section 6.3 we notice no improvement in model performance. The lack of increase in accuracy along with poor results on using the image features independently reinforce our initial hypothesis that the branding and context surrounding an NFT influences the asset value more than the content itself.

		Predicted				
		Class1	**Class2**	**Class3**	**Class4**	**Class5**
Actual	**Class1**	199	769	28	0	0
	Class2	169	1846	86	0	0
	Class3	25	392	131	0	0
	Class4	7	40	6	0	0
	Class5	0	6	0	0	0

Table 3: Confusion matrix for the image-based classifier model. Most assets are classified into the largest class.

7 Conclusion

NFTs have the potential to challenge the collectible and art market worth over $350 Billion. In this paper we track the growth of NFTs and show how social media reach can impact its value. We build models to predict asset value and find that Twitter features listed count, username characteristic have a great impact on the asset valuation. We lay out the first piece of work to characterise and value NFT assets using social media features. Our proposed system can be used to build a profitable trading strategy by identifying overvalued and undervalued assets.

NFTs is a fast evolving industry with numerous challenges appearing daily. For example on 28th October, an NFT asset - CryptoPunk 9988 was sold for over $532 Million USD. [12] However the buyer and seller of the asset was the same person and the goal was to artifically inflate prices and become viral on social media. Our current system would not handle such outlier transactions and schemes. In future we plan to build systems to automatically detect such fraudulent transactions that try to create artificial market signals.

References

1. Ante, L.: The non-fungible token (NFT) market and its relationship with Bitcoin and Ethereum. Tech. Rep. ID 3861106, Social Science Research Network (Jun 2021)
2. Ante, L.: Smart contracts on the blockchain – A bibliometric analysis and review. Telematics and Informatics **57**, 101519 (Mar 2021)
3. Blenn, N., Mieghem, P.: Are human interactivity times lognormal? (07 2016)
4. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. Journal of Computational Science **2**(1), 1–8 (Mar 2011)
5. Chevet, S.: Blockchain Technology and Non-Fungible Tokens: Reshaping Value Chains in Creative Industries. Tech. Rep. ID 3212662, Social Science Research Network (May 2018)
6. Doerr, C., Blenn, N., Mieghem, P.V.: Lognormal Infection Times of Online Information Spread. PLOS ONE **8**(5), e64349 (May 2013)
7. Dowling, M.M.: Is Non-fungible Token Pricing Driven by Cryptocurrencies? Tech. Rep. ID 3815093, Social Science Research Network (Mar 2021)
8. van Haaften-Schick, L., Whitaker, A.: From the Artist's Contract to the Blockchain Ledger: New Forms of Artists' Funding Using NFTs, Fractional Equity, and Resale Royalties. Tech. Rep. ID 3842210, Social Science Research Network (Dec 2020)
9. Nadini, M., Alessandretti, L., Di Giacinto, F., Martino, M., Aiello, L.M., Baronchelli, A.: Mapping the NFT revolution: market trends, trade networks and visual features. arXiv:2106.00647 [physics, q-fin] (Sep 2021)
10. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system. Decentralized Business Review p. 21260 (2008)
11. Nguyen, T.H., Shirai, K., Velcin, J.: Sentiment analysis on social media for stock movement prediction. Expert Systems with Applications **42**(24), 9603–9611 (Dec 2015)
12. Piñeiro-Chousa, J., Vizcaíno-González, M., Pérez-Pico, A.M.: Influence of Social Media over the Stock Market. Psychology & Marketing **34**(1), 101–108 (2017)

[12] https://www.bloomberg.com/news/articles/2021-10-29/here-s-a-532-million-nft-trade-that-wasn-t-what-it-appeared

13. Sako, K., Matsuo, S., Meier, S.: Fairness in ERC token markets: A Case Study of CryptoKitties. arXiv:2102.03721 [cs] (Feb 2021)
14. Wang, Q., Li, R., Wang, Q., Chen, S.: Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges. arXiv:2105.07447 [cs] (May 2021)
15. Whitaker, A.: Art and Blockchain: A Primer, History, and Taxonomy of Blockchain Use Cases in the Arts. Artivate: A Journal of Enterprise in the Arts pp. 21–47 (Oct 2019)