

Tweet Based Reach Aware Temporal Attention Network for NFT Valuation

Ramit Sawhney*
Georgia Tech

Megh Thakkar*
BITS, Pilani

Ritesh Soun*
SVC, DU

Atula Neerkaje
MIT, Manipal

Vasu Sharma
Amazon Science

Dipanwita Guhathakurta
IIIT, Hyderabad

Sudheer Chava
Georgia Tech

rsawhney31@gatech.edu, sudheer.chava@scheller.gatech.edu

Abstract

Non-Fungible Tokens (NFTs) are a relatively unexplored class of assets. Designing strategies to forecast NFT trends is an intricate task due to its extremely volatile nature. The market is largely driven by public sentiment and "hype", which in turn has a high correlation with conversations taking place on social media platforms like Twitter. Prior work done for modelling stock market data does not take into account the extent of impact certain highly influential tweets and their authors can have on the market. Building on these limitations and the nature of the NFT market, we propose a novel reach-aware temporal learning approach to make predictions for forecasting future trends in the NFT market. We perform experiments on a new dataset consisting of over 1.3 million tweets and 180 thousand NFT transactions spanning over 15 NFT collections curated by us. Our model (TA-NFT) outperforms other state-of-the-art methods by an average of 36%. Through extensive quantitative and ablative analysis, we demonstrate the ability of our approach as a practical method for predicting NFT trends.

1 Introduction

Non Fungible Tokens (NFTs) are digital assets that represent objects like art, collectibles, and in-game items¹. Public attention towards NFTs exploded in 2021 when their market experienced record sales (NonFungible, 2021), but little is known about the overall structure and evolution of its market. The NFT space is characterized by extreme growth along with highly skewed and uncertain returns that typify speculative markets (White et al., 2022). Little to no work has been done to forecast future trends in the NFT market, and unlike other, more stable assets, investing in NFTs is associated with extremely high amounts of risk (Mazur, 2021a; Nardini et al., 2021) as they are highly volatile (Kong

*Equal contribution.

¹<https://ethereum.org/en/nft/>

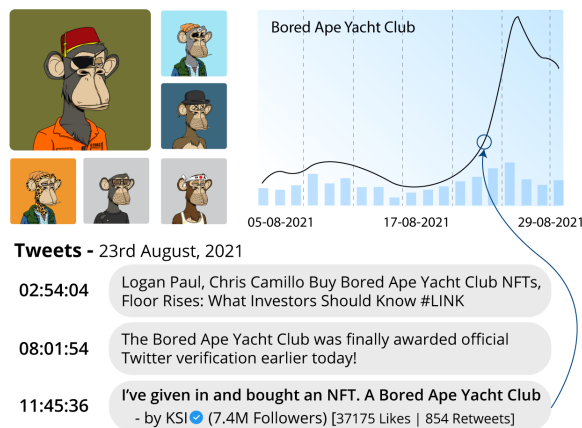


Figure 1: We visualize a sample of tweets related to Bored Ape Yacht Club NFTs. We also plot the daily average price of the same NFT collection to observe the impact of tweets by influential users.

and Lin, 2021). Additionally, social media has emerged as a space for NFT holders and creators to shape community opinion and drive public sentiment about NFT projects (van Slooten, 2022). Therefore, conventional forecasting approaches and contemporary ML models which utilize only numerical historic NFT data fail to capture sufficient information.

Behavioral finance theories (Chu et al., 2019) suggest that people are more likely to make decisions based on overconfidence bias (Slovic and Fischhoff, 1977; Gervais, 2001) and herd behavior (Bikas et al., 2013; Bikhchandani and Sharma, 2000) when faced with uncertainty. The abundance of tweets about various NFT collections help in creating "hype" around them which drives their sales, reinforcing herd behavior. Studies have shown that NFTs valued by experts are more successful (Franceschet, 2020), and that the structure of the the NFT co-ownership network is highly centralized, and small-world-like (Barabasi, 2021; Barrat and Weigt, 1999).

As shown in Figure 1, the daily average price of an NFT collection, namely Bored Ape Yacht

Club, reacts immediately to a highly influential individual tweeting positively about it and spikes up. However, numerous challenges arise while analyzing such texts. For instance, there are inherent dynamic timing irregularities (Sawhney et al., 2021d) when influencers or "alpha" users make such tweets and as communities react to them. Simultaneously capturing temporal granularities along with popularity information (Savaş, 2021) is crucial, as more widely the content is shared over time, the greater the user's impact becomes (Anger and Kittl, 2011).

Therefore, to develop a robust method for predicting NFT trends, we curate a dataset (§3.1), and formulate a new time and popularity aware financial modelling approach, where the influence and reach of individual tweets is captured effectively translating its effects in their market value.

Our contributions can be summarized as:

- We curate a dataset consisting of over 1.3 million tweets and 180 thousand NFT transactions spanning over 15 NFT collections for two downstream tasks, namely daily average price prediction and price movement classification (§3).
- We plan to make this data publicly available and hope that it could further the research in this field. To the best of our knowledge, this will be the first publicly available, large scale dataset on NFTs based on social media "hype" and sentiment.
- We propose a novel tweet based reach-aware temporal attention network to predict NFT trends (§5), and analyze the impact of social media on NFT price prediction.
- Through quantitative (§6.1), ablative (§6.2) and exploratory (§6.3, §6.4) experiments, we build the case for our approach as a practical method for modelling NFT market data.

2 Related Work

Non-Fungible Tokens NFTs are digital assets with relatively recent origins (Nadini et al., 2021). NFT pricing involve more complex valuations in comparison to traditional assets such as equity (Kong and Lin, 2021), and are associated with higher returns along with high volatility (Mazur, 2021a). Existing research on NFTs focus mostly on technical aspects such as components, protocols, standards, & desired properties (Wang et al., 2021) and new blockchain-based protocols to trace physical goods (Westerkamp et al., 2018) and the implications that

NFTs have on the art world (Whitaker, 2019; van Haften-Schick and Whitaker, 2021). Furthermore, little to no work has been done to forecast future trends in the NFT market.

NLP in Finance Traditional financial forecasting techniques have been applied in areas such as stock markets (Ariyo et al., 2014; Rundo et al., 2019), currency exchange markets (Kamruzzaman and Sarker, 2003), and energy economics (Bento et al., 2018). Conventional financial models previously relied on numerical features (Nikou et al., 2019) and technical indicators (Shynkevich et al., 2017). These include discrete (Ariyo et al., 2014; Bollerslev, 1986), continuous (Jacquier et al., 2002; Andersen, 2007), and neural approaches (Luo et al., 2018; Kim et al., 2019). Efforts have since shifted towards utilizing textual data such as social media posts (Xu and Cohen, 2018), news reports (Li et al., 2020; Schumaker and Chen, 2009), web searches (Zhong and Raghieb, 2019; Liu et al., 2012), etc.. These studies confine their analyses to stock markets. Recently, Sawhney et al. (2022) explored cryptocurrency bubble prediction based on user behavior on social media. However, there is a gap in leveraging social media and NLP to analyse and forecast future trends in the NFT market.

Time-Aware Modelling Temporal data is omnipresent in several real-world applications, including healthcare (Baytas et al., 2017a), recommender systems (Rabiu et al., 2020), and finance (Selvin et al., 2017). As a result, sequential neural models such as LSTMs (Hochreiter and Schmidhuber, 1997) have gained popularity due to their ability to capture sequential context dependency (Hu et al., 2018). Time-aware modelling of time series data has shown improvements over conventional sequential neural models on various tasks such as patient subtyping (Baytas et al., 2017a), suicide ideation detection (Sawhney et al., 2020), and disease progression (Gao et al., 2020). Recently, time-aware modelling has been adapted in the realm of financial NLP, such as stock recommendation (Ying et al., 2020), price prediction (Sawhney et al., 2021a), and ranking (Sawhney et al., 2021d). However, these approaches do not take into account the engagement and popularity of social media posts. Hence, such methods do not scale to NFTs, which are more closely correlated with user sentiment and social media "hype" in comparison to traditional asset classes (Bouraga, 2021; Franceschet, 2020). With this work, we seek to explore a promising

research avenue i.e the intersection of NFTs and financial NLP, along with time and hype aware neural modelling.

3 Dataset and Tasks

3.1 Dataset

We utilise two sources, Twitter and Etherscan² (Ethereum Blockchain access point) to collect qualitative and quantitative data respectively for 15 NFT collections. We shortlist NFT collections which are launched before January 1st 2022, and appear among the top 40 collections by all-time sales volume on Opensea³, the most popular marketplace for NFTs. Using the data described below, we construct two datasets for the tasks described in the subsequent section.

3.1.1 Qualitative Data - Tweets

We collect qualitative data by extracting tweets related to shortlisted NFT collections from Twitter. We search for tweets consisting of the official Twitter handle of the collection, the Twitter handles of its creators, as well as a curated list of most frequently used hashtags and search terms related to each collection. Tweets matching any of the above search criteria are extracted. In addition to the tweet text and engagement information (number of likes, retweets, etc.), we also associate each tweet with information about the user who posted it, such as user bio, followers and friends count etc.

We have a total of 1,354,427 tweets corresponding to 15 NFT collections posted in the one-year period between January 1 2021 to January 31 2022. The median number of tweets over the collections is 65,158 with a maximum of 363,506 corresponding to the NFT collection Cool Cats NFT.

3.1.2 Quantitative Data - Transactions

We gather quantitative data, that is NFT transactions between January 1 2021 to January 31 2022 for shortlisted collections from Etherscan which is an Ethereum blockchain explorer. We filter out confirmed NFT sales and extract all relevant data for each transaction comprising of the seller and buyer address, transaction timestamp, amount and meta-data of the NFT sold/purchased.

We have a total to 188,535 transactions over the one year time span for 15 NFT collections. A

²<https://etherscan.io/>

³https://opensea.io/rankings?sortBy=total_volume

detailed breakdown of the dataset is given in Appendix B.

3.2 Tasks

We aim to predict future NFT trends based on historic tweets about an NFT collection.

Daily Average Price Prediction We regress the future daily average price of an NFT collection n given as, $\theta = \frac{\sum_{k=1}^{k=t_d} s_{kd}}{t_d}$, where s_{kd} is the transaction value of the k^{th} NFT sale on day d and t_d is the number of sales on that day. Given L historic tweets for a collection, we aim to predict the average price of the NFT collection on the next day. It is evaluated using mean squared error loss.

Price Movement Classification We formulate movement prediction as a binary classification task. For an NFT collection n , label $y_k = 1$ if $s_k > s_{k-1}$, $y_k = 0$ otherwise. Thus y_k refers to the price movement of the NFT collection since s_{k-1}^{th} transaction. We evaluate the model performance on this task using macro F1 score.

4 Experimental Setup

Preprocessing Following (Nguyen et al., 2020), we use NLTK to preprocess tweets by converting mentions (@) and URLs to special tokens @USER and HTTPURL. We treat emoticons by converting them to strings using emoji Python package.

Training Setup We perform all our experiments on a Tesla T4 GPU. We use Optuna (Akiba et al., 2019) to find optimal hyperparameter values based on the validation MSE/Macro F1 scores by performing 25 search trials. We explored the lookback window length $L \in [2, 40]$ and the hidden state dimensions $\in [64, 768]$. We use 10%, 10% and 80% of the samples for testing, validation and training respectively for both tasks. We use learning rate $\in [1e^{-5}, 1e^{-2}]$ and train the models using Adam as our optimizer for 2,150 seconds and 10,845 seconds for daily average price prediction and price movement classification tasks, respectively.

Evaluation Metrics We evaluate methods using Mean Squared Error (MSE) loss for daily average price prediction task and Macro F1-score (M.F1) for price movement classification task.

4.1 Baseline Models

- **Prophet** A decomposable time-series model utilising interpretable model components (Taylor and Letham, 2017)
- **ARIMA** A moving average based autoregressive model that uses past prices as input (Adebiyi et al., 2014).
- **MLP** A simple Multi-Layer Perceptron that uses averaged BERT embeddings of tweet sequences as input.
- **LSTM** Utilizes an LSTM (Hochreiter et al., 1997), which is capable of learning long term dependencies, to encode textual streams.
- **FastText + CNN** A CNN based architecture (Kim, 2014) with a convolution layer on top of FastText (Joulin et al., 2016) embeddings.
- **FAST** A time-aware LSTM capable of modelling temporally irregular text stream data (Sawhney et al., 2021d).

5 Methodology

5.1 Features

Text Embeddings We use Bidirectional Encoder Representations from BERTweet (Nguyen et al., 2020) to encode each preprocessed tweet p_k to features $\mathbf{m}_k = \text{BERTweet}(p_k) \in \mathbb{R}^d$ where $d = 768$, obtained by taking the [CLS] token output from the final layer.

User Feature Vector We use the Twitter user metadata for each tweet p_k , to construct a user feature vector $\mathbf{u}_k \in \mathbb{R}^d$ where $d = 5$, normalised column-wise. This vector contains essential information about the author of the tweet like the number of followers, whether the author is verified or not, their status count, their favourites count, and friends count. This helps the model learn not only from the contextualized BERT representations but also find potential correlations between user meta data and the tweet’s influence on NFT valuation.

5.2 Model Components

In this section we present the architecture of our framework, **TA-NFT: Time and Reach Aware Network for NFT Price Prediction**, designed to forecast NFT prices based on social media trends by explicitly modelling the temporal irregularities and engagement of tweets.

Reach Aware Temporal Network Fine-grained timing irregularities play a crucial role in modelling online text stream data. For instance, the time interval between two tweets about an NFT collection can vary widely, from a few minutes to several days. Therefore, its influence on the value of the NFT collection may drastically vary overtime. There is a decay or increase in the influence of the tweet in relation to other tweets about the collection. Furthermore, every tweet does not have the same reach. The reach/engagement of two consecutive tweets about the same collection may vary by thousands of likes and retweets. In addition to this, the sentiment polarity between tweets may also vary drastically.

Thus, in order to capture these reach, polarity and time dependent complexities, we modify Time-aware LSTM (Baytas et al., 2017b) into reach-aware temporal network (RTN(\cdot)). Intuitively, the greater the time elapsed between tweets, the lesser the impact, and the greater the reach, the higher the impact in the direction of sentiment polarity. Thus, for a given day and time k , RTN applies a decaying function over Δk , the elapsed time between two tweets $[p_k, p_{k-1}]$. It also applies a function over the number of likes l , retweets r and polarity s of a tweet, transforming the reach, polarity and time differences into weights:

$$\begin{aligned} \mathbf{C}_{k-1}^s &= \tanh(\mathbf{W}^d \mathbf{C}_{k-1} + \mathbf{b}^d) \\ \hat{\mathbf{C}}_{k-1}^s &= \mathbf{C}_{k-1}^s * g(\Delta k) * q(l, r, s) && \text{(Discounted short-term memory)} \\ \mathbf{C}_{k-1}^T &= \mathbf{C}_{k-1} - \mathbf{C}_{k-1}^s && \text{(Long term memory)} \\ \mathbf{C}_{k-1}^* &= \mathbf{C}_{k-1}^T + \hat{\mathbf{C}}_{k-1}^s && \text{(Adjusted previous memory)} \end{aligned}$$

where \mathbf{C}_{k-1}^s is the previous cell memory, $\mathbf{W}^d; \mathbf{b}^d$ are the network parameters, $g(\cdot)$ is a heuristic decaying function. Following (Baytas et al., 2017b) we set $g(\cdot)$ as,

$$g(\Delta k) = 1/\Delta k$$

and $q(\cdot)$ as,

$$q(l, r, s) = \begin{cases} s * (l + r) & \text{if } s \neq 0 \\ \zeta * (l + r) & \text{if } s = 0 \end{cases}$$

where $\zeta \approx 0$.

Using the adjusted previous memory \mathbf{C}_{k-1}^* , we define the current hidden state and current memory states for RTN as:

$$\begin{aligned} \tilde{\mathbf{c}}_k &= \tanh(\mathbf{W}^c \mathbf{h}_{k-1} + \mathbf{U}^c \mathbf{m}_k + \mathbf{b}^c) \\ \mathbf{C}_k &= \mathbf{i}_k * \tilde{\mathbf{c}}_k + \mathbf{f}_k * \mathbf{C}_{k-1}^* && \text{(Current memory)} \\ \mathbf{h}_k &= \mathbf{o}_k * \tanh(\mathbf{C}_k) && \text{(Current hidden state)} \end{aligned}$$

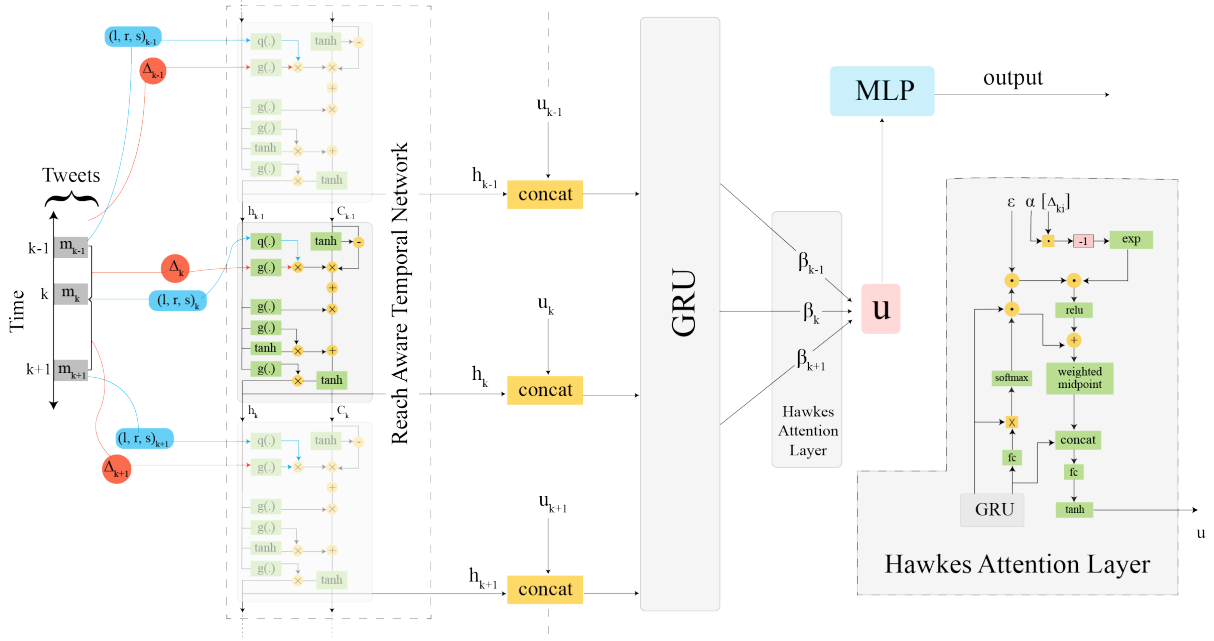


Figure 2: An overview of TA-NFT - Reach Aware Temporal Network. TA-NFT feeds tweet embeddings to a reach-aware temporal network (RTN). User-features are concatenated to the output of RTN and fed to a GRU, followed by a Hawkes Attention layer. Finally, the aggregated representation is passed to an MLP for prediction.

where $\mathbf{W}^c; \mathbf{U}^c; \mathbf{b}^c$ are the learnable parameters, $\mathbf{i}_k; \mathbf{f}_k; \mathbf{o}_k$ are input, forget and output gates. Finally, given tweets $[p_1, \dots, p_T]$ over a lookback length L , we define the update rule of RTN as,

$$\mathbf{h}_k = \text{RTN}(\mathbf{m}_k, \Delta k, \mathbf{h}_{k-1}); \quad k \in [1, T] \quad (1)$$

where, \mathbf{h}_k represents the hidden states of RTN.

The hidden states obtained from RTN are then updated by concatenating the user feature vectors \mathbf{u}_k to it,

$$\mathbf{h}_k = \mathbf{h}_k \oplus \mathbf{u}_k \quad (2)$$

to obtain feature vectors $\in \mathbb{R}^d$ where $d=773$.

Hawkes Attention Layer Existing work show that all historical sequence features are not equally informative and have a *varied influence* over the predictions (Sawhney et al., 2021c). We use a temporal attention mechanism (Luong et al., 2015) to emphasize sequence features likely to have substantial influence. This mechanism learns attention weights β_k for each hidden state $\mathbf{h}_k \in \bar{\mathbf{h}} = [\mathbf{h}_1, \dots, \mathbf{h}_T]$ as,

$$\beta_k = \text{Softmax}_k((\mathbf{h}_k)^T (\mathbf{W}^a \bar{\mathbf{h}})) \quad (3)$$

where, \mathbf{W} denotes learnable weights.

Next, we enhance the temporal attention using the Hawkes process (Mei and Eisner, 2017) with a

Hawkes attention mechanism. The Hawkes process is a temporal point process that models a sequence of arrival of features over time. Each feature item “excites” the process in the sense that the chance of a subsequent arrival is increased for some time. Studies (Zuo et al., 2020; Sawhney et al., 2021b) show that the Hawkes process can be used to model sequences from social media and discourses. The Hawkes attention mechanism learns an excitation parameter ϵ corresponding to excitation induced by tweet p_k and a decay parameter α to learn the decay rate of this induced excitement. Formally, we use a weighted average to aggregate hidden states $\bar{\mathbf{h}}$ via Hawkes process as,

$$\mathbf{u} = \text{TA-NFT}(\{p_k, t_k\}_{k=1}^T) = \sum_k \frac{\beta_k \mathbf{q}_k}{\sum_\tau \beta_\tau \mathbf{q}_\tau} \mathbf{q}_k \quad (4)$$

$$\mathbf{q}_k = \beta_k * \mathbf{h}_k + \epsilon * (\text{ReLU}(\mathbf{h}_k)) * e^{-\alpha \Delta k} \quad (5)$$

6 Results

6.1 Performance Comparison

Table 1 shows a comparison of TA-NFT against baselines spanning commonly used approaches for asset price prediction tasks. We observe that our model outperforms most baselines by an average of 36%. ARIMA (Adebiyi et al., 2014) and Facebook Prophet (Taylor and Letham, 2017), being time-series models using only historical price

Model	Price Pred. MSE ↓	Mov. Pred. M.F1 ↑
Prophet (Taylor and Letham, 2017)	0.4084	0.2576
ARIMA (Adebiyi et al., 2014)	0.1510	0.3278
MLP	0.1363	0.3621
LSTM (Hochreiter et al., 1997)	0.1287	0.3914
FastText + CNN (Kim, 2014)	0.1630	0.3076
FAST (Sawhney et al., 2021d)	0.1253	0.4032
TA-NFT (Ours)	0.0914*	0.4618*

Table 1: Performance comparison with baselines. * indicates improvement over SOTA is significant ($p < 0.01$) under Wilcoxon’s signed rank test.

data, are unable to capture sufficient information. FastText+CNN (Kim, 2014) applies Convolutional Neural Networks on text embeddings from tweets, and FAST (Sawhney et al., 2021d) is a time-aware model using both text and historical features. We postulate that our model’s superior performance over them is due to, 1) time-aware Hawkes attention mechanism, 2) incorporation of tweets’ reach, polarity and timing based irregularities, and 3) accounting for author influence on the impact of individual tweets. TA-NFT outperforms other time-aware networks due to the Hawkes attention mechanism, tweet meta data and user information which serve as proxies for the popularity of the NFT on Twitter. These observations reveal that a combination of these features contribute towards NFT valuation, and by capturing all these features, our model is practically applicable for NFT average price prediction and price movement classification.

6.2 Ablation Study

We account for the importance of various components of TA-NFT in Table 2. First, we observe that replacing the standard LSTM (Hochreiter et al., 1997) with Time-aware LSTM (Baytas et al., 2017b) leads to significant performance improvement. This validates that incorporating the time irregularities helps in modelling the NFT market. Further improvement is noted on modifying it into Reach-aware T-LSTM which accounts for the reach of individual tweets. Enriching the temporal network with Hawkes process leads to performance boosts. This is possibly due to the ability of the Hawkes attention layer to capture excitations caused by influential tweets. Finally enriching the tweet embeddings with user feature vector in combination with reach-aware temporal network and Hawkes attention layer leads to best results, indicat-

Reach Weights	User Feature Vector	Model	Price Pred. MSE ↓	Mov. Pred. M.F1 ↑
×	×	LSTM	0.1287	0.3914
×	×	T-LSTM	0.1248	0.4325
✓	×	T-LSTM	0.1196	0.4372
×	×	T-LSTM + Hawkes	0.1031	0.4561
✓	×	T-LSTM + Hawkes	0.1026	0.4601
✓	✓	T-LSTM + Hawkes (Ours)	0.0914*	0.4618*

Table 2: Ablation study over TA-NFT (mean of 10 runs). *,† indicate improvements are significant ($p < 0.01$) under Wilcoxon’s signed rank test.

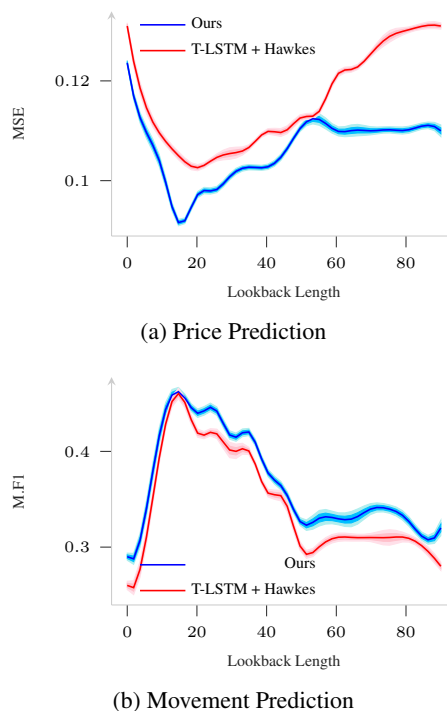


Figure 3: Impact of lookback length L on TA-NFT’s performance with error bounds. Results are averaged over 10 independent runs.

ing that capturing the author’s influence is significantly advantageous to understand the full extent of a tweet’s impact on the NFT market.

6.3 Impact of Lookback Length

We study the impact of varying the lookback length L , referring to the number of historical tweets used as input for each data point, on our model’s performance for average price prediction task. We observe that with no historical context, both models perform the worst. As we increase the lookback length L , the model performance improves up to an optimal point, indicating that the naturally decaying impact of past tweets on NFT valuation is being captured by the model. As we further increase L beyond the optimal value, we observe a gradual drop in performance. This is possibly due to the noise introduced by older tweets, which are

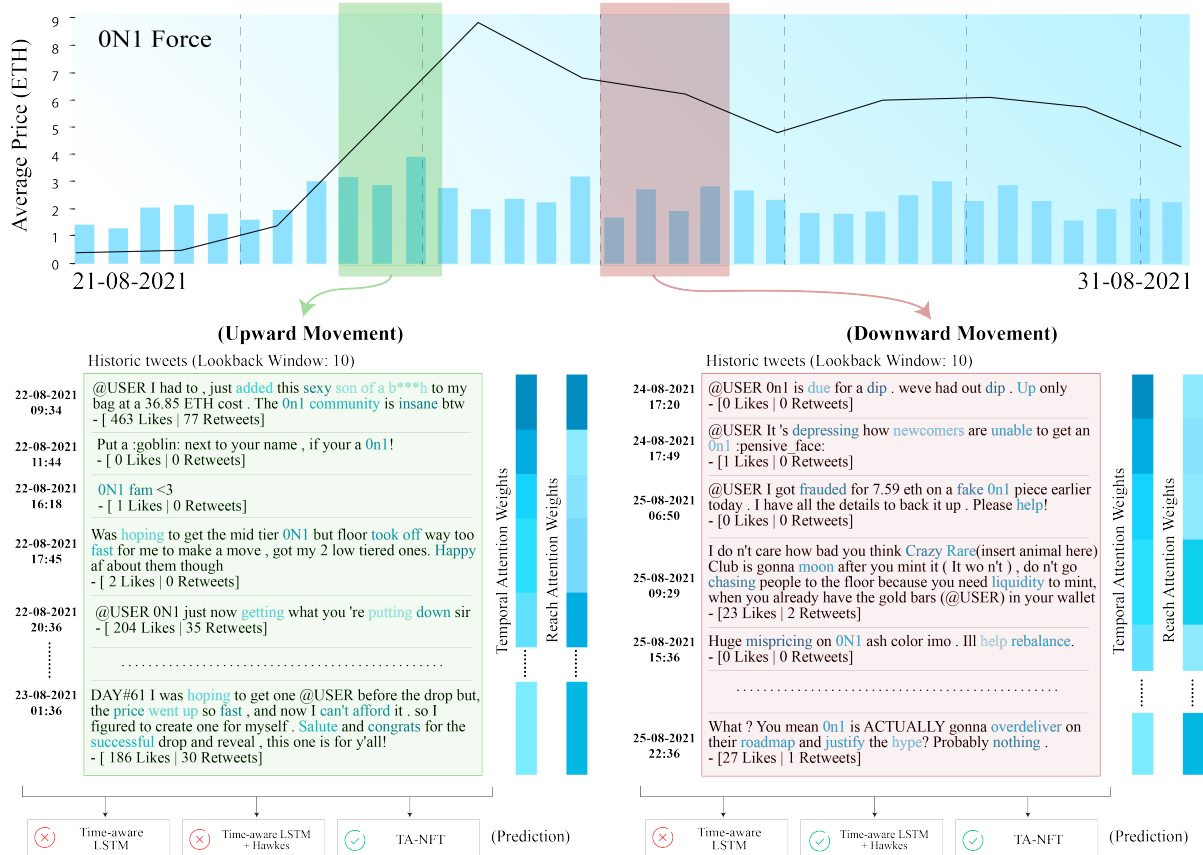


Figure 4: Qualitative analysis of Tweets about ON1 Force NFTs and performance of TA-NFT on price movement prediction task with temporal, reach and token level attention visualised.

Model	Avg. Price. Pred. MSE ↓	Movement Pred. M.F1 ↑
LSTM	0.1943	0.3414
T-LSTM	0.1781	0.3536
T-LSTM + Hawkes	0.1702	0.3819
TA-NFT (Ours)	0.1627*	0.4117*

Table 3: Performance comparisons in a zero shot setting. * indicates improvement over SOTA is significant ($p < 0.01$) under Wilcoxon’s signed rank test.

relatively insignificant to model the temporal state of the community around the NFT collection. The short term dependence of NFT valuation on tweets indicates the fast-moving and volatile nature of the NFT space.

6.4 Zero-shot Transfer Analysis

We compare the performance of our model in a zero-shot setting in Table 3, where we train the models on a set of collections and test them on a set of previously unseen collections. Our model outperforms other text-based and temporal models. This shows that it is able to effectively generalize

Model	Visual Features Used	Avg. Price Pred. MSE	Mov. Pred. M.F1
TA-NFT	None	0.0914	0.4618
TA-NFT	All	0.1879	0.3291
TA-NFT	Reduced using PCA	0.1989	0.3382
TA-NFT	Selected using Boruta	0.1829	0.3432

Table 4: Impact of visual features on the performance of TA-NFT. Results are averaged over 10 independent runs.

better for unseen collections. Further, it indicates that NFT collections share some inherent characteristics and have overlapping latent representations that can be learnt using online text streams.

6.5 Impact of Visual Features

We perform a study to account for the impact of the contents of NFTs, i.e., images towards its valuation. We compare the performance of our modelling approach with and without visual features in Table 4. We pretrain the Barlow Twins model (Zbontar et al., 2021) on all NFT images, minimizing the redundancy between the embeddings of two identi-

cal networks in order to produce information rich representations for the images. We take the output of the last fully connected layer of the model as the vector representation $v_i \in \mathbb{R}^d$ where $d = 1000$ for each image. Further, we concatenate these visual features with text features and carry out training and evaluation as usual. We also explore different approaches to reduce/select feature dimensions, namely Principal Component Analysis and Boruta (Kursa et al., 2010). We observe that utilising visual features does not lead to any improvements, but rather degrades the model performance. This observation suggests that visual features do not provide any useful information for NFT valuation and induce noise to the data. We hypothesize that this could be possibly due to inter-collection and intra-collection content similarities spawned by the market responsiveness to the success of a collection (Nadini et al., 2021).

6.6 Qualitative Analysis

We conduct a qualitative study in an attempt to interpret the predictions of TA-NFT by taking examples of tweets about 0N1 Force NFT collection as shown in Figure 4 for two cases.

Following a series of positive tweets with significant reach, we observe an upward movement in the price of 0N1 Force NFT collection. Similarly, a downward movement appears to be caused by a series of relatively negative tweets with lower reach. This suggests NFTs follow hype-driven pricing where more wide-reaching social media traffic and positive sentiment leads to an upward trend and vice-versa. Our modelling approach (TA-NFT) is able to contextualize the impact of social media hype by accounting for the reach of individual tweets as well as the influence of its authors in addition to the timing irregularities. Thus, it is able to correctly classify the price movement in both cases as opposed to strictly time-aware modelling techniques. Unlike traditional assets like stocks and gold, the intensity and polarity of public sentiment on social media platform drives price fluctuations (Semenova and Winkler, 2021) which is in turn affected by influential individuals.

7 Conclusion

Building on the rising popularity and hype-driven dynamics of NFT markets, we curate a dataset for forecasting NFT trends through two downstream tasks consisting of daily average price prediction

and price movement classification. We introduced TA-NFT, a time and reach-aware neural network for modelling temporal granularities and engagement dynamics of NFT discourse on social media. Through extensive experiments, we show that TA-NFT empirically outperforms other SOTA models by an average of 36%, and present TA-NFT as a practical modelling approach and a strong benchmark for forecasting NFT trends. We hope the proposed dataset can enable more academic progress in the field of financial NLP.

Ethical Considerations

While the predictive power of models like TA-NFT relies on data, we work within the purview of acceptable privacy practices to avoid coercion and intrusive treatment. We utilize publicly available data in a purely observational and non-intrusive manner. Although informed consent of each user was not sought as it may be deemed coercive, we follow all ethical regulations set by our data sources. Since financial markets are transparent (Bloomfield and O'Hara, 1999) and heavily regulated (Edwards, 1996), we discuss the ethical considerations and potential risks pertaining to our work.

Potential risks: Our contributions are meant as an exploratory research in the financial domain and no part of the work should be treated as financial advice. All financial investments decisions are subject to market risk (Mazur, 2021b; Antonakakis et al., 2019; Campbell, 1996) and should be made after extensive testing. Practitioners should check for various biases (demographic, modelling, randomness) before attempting to use the provided code/data/methods for real-world purposes.

Intended use of data artifacts: Our dataset will be made available to use for research purposes. The intended use of financial datasets is to enable investors to take informed financial decisions (Cooper et al., 2016), research and development to foster progress of AI methods and financial modelling for public good (Velooso et al., 2021).

We additionally follow Cooper et al. (2016) and focus on the following ethical considerations for automated trading systems:

Blocking Price Discovery Trading systems should not block price discovery, nor interfere with the ability of other market participants to add to their own information (Angel and McCabe, 2013). Examples of such scenarios include Quote Stuffing (Egginton et al., 2016) and Wash Trading (von

Wachter et al., 2022). TA-NFT does not block price discovery in any manner.

Circumventing Price Discovery A trading system should not hide information, such as by participating in dark pools or placing hidden orders (Zhu, 2014). While we evaluate our approach only on public data, it is possible for TA-NFT, just as any other automated trading system, to be exploited to hinder market fairness (Sako et al., 2021). We follow broad ethical guidelines to design TA-NFT and encourage readers to follow both regulatory and ethical considerations pertaining to the market.

Limitations

While our dataset has been curated using data for the entire year of 2021, the NFT market is fast paced, new and ever-changing, which may lead to the need of adapting newer approaches. Apart from this, there are 1000s of NFT collections, and we conduct our analysis on only 15 of them, which might leave out a lot of independent NFT collections and related trends. We also acknowledge the presence of demographic bias in our study as the tweet data is limited to English, and thus our approach may not directly generalize to non-English settings. Additionally, there is a vast scope for future work accounting for the influence of buyer/seller network, correlation between the NFT and Cryptocurrency market, other sources of qualitative data like news, blogs, Reddit etc. and NFT metadata attributes/value proposition.

References

- Ayodele Adebiyi, Aderemi Adewumi, and Charles Ayo. 2014. Stock price prediction using the arima model.
- Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. 2019. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- Leif BG Andersen. 2007. Efficient simulation of the heston stochastic volatility model. Available at SSRN 946405.
- James J Angel and Douglas McCabe. 2013. Fairness in financial markets: The case of high frequency trading. *Journal of Business Ethics*, 112(4):585–595.
- Isabel Anger and Christian Kittl. 2011. Measuring influence on twitter. In *Proceedings of the 11th international conference on knowledge management and knowledge technologies*, pages 1–4.
- Nikolaos Antonakakis, Ioannis Chatziantoniou, and David Gabauer. 2019. Cryptocurrency market contagion: Market uncertainty, market complexity, and dynamic portfolios. *Journal of International Financial Markets, Institutions and Money*, 61:37–51.
- Adebiyi A Ariyo, Adewumi O Adewumi, and Charles K Ayo. 2014. Stock price prediction using the arima model. In *2014 UKSim-AMSS 16th international conference on computer modelling and simulation*, pages 106–112. IEEE.
- Albert-Laszlo Barabasi. 2021. The art market often works in secret. here’s a look inside. <https://www.nytimes.com/2021/05/07/opinion/nft-art-market.html>.
- Alain Barrat and Martin Weigt. 1999. On the properties of small-world network models. *The European Physical Journal B - Condensed Matter and Complex Systems*, 13:547–560.
- Inci M Baytas, Cao Xiao, Xi Zhang, Fei Wang, Anil K Jain, and Jiayu Zhou. 2017a. Patient subtyping via time-aware lstm networks. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 65–74.
- Inci M. Baytas, Cao Xiao, Xi Zhang, Fei Wang, Anil K. Jain, and Jiayu Zhou. 2017b. Patient subtyping via time-aware lstm networks. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’17*, page 65–74, New York, NY, USA. Association for Computing Machinery.
- PMR Bento, JAN Pombo, MRA Calado, and SJPS Mariano. 2018. A bat optimized neural network and wavelet transform approach for short-term price forecasting. *Applied energy*, 210:88–97.
- Egidijus Bikas, Daiva Jurevičienė, Petras Dubinskas, and Lina Novickytė. 2013. Behavioural finance: The emergence and development trends. *Procedia-social and behavioral sciences*, 82:870–876.
- Sushil Bikhchandani and Sunil Sharma. 2000. Herd behavior in financial markets. *IMF Staff papers*, 47(3):279–310.
- Robert Bloomfield and Maureen O’Hara. 1999. Market transparency: who wins and who loses? *The Review of Financial Studies*, 12(1):5–35.
- Tim Bollerslev. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.
- Sarah Bouraga. 2021. On the popularity of non-fungible tokens: Preliminary results. In *2021 3rd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS)*, pages 49–50. IEEE.
- John Y Campbell. 1996. Understanding risk and return. *Journal of Political economy*, 104(2):298–345.

- Jeffrey Chu, Yuanyuan Zhang, and Stephen Chan. 2019. The adaptive market hypothesis in the high frequency cryptocurrency market. *International Review of Financial Analysis*, 64:221–231.
- Ricky Cooper, Michael Davis, and Ben Van Vliet. 2016. The mysterious ethics of high-frequency trading. *Business Ethics Quarterly*, 26(1):1–22.
- Franklin R Edwards. 1996. *The new finance: regulation and financial stability*. American Enterprise Institute.
- Jared F Egginton, Bonnie F Van Ness, and Robert A Van Ness. 2016. Quote stuffing. *Financial Management*, 45(3):583–608.
- Massimo Franceschet. 2020. [Art for space](#). *J. Comput. Cult. Herit.*, 13(3).
- Junyi Gao, Cao Xiao, Yasha Wang, Wen Tang, Lucas M Glass, and Jimeng Sun. 2020. Stagenet: Stage-aware neural networks for health risk prediction. In *Proceedings of The Web Conference 2020*, pages 530–540.
- S Gervais. 2001. T., olean., 2001. learning to be overconfident. *Review of Financial Studies*, 14(1):1.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Sepp Hochreiter et al. 1997. [Long short-term memory](#). *Neural computation*, 9:1735–80.
- Ziniu Hu, Weiqing Liu, Jiang Bian, Xuanzhe Liu, and Tie-Yan Liu. 2018. Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. In *Proceedings of the eleventh ACM international conference on web search and data mining*, pages 261–269.
- Eric Jacquier, Nicholas G Polson, and Peter E Rossi. 2002. Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, 20(1):69–87.
- Armand Joulin, Edouard Grave, Piotr Bojanowski, Matthijs Douze, Herve Jégou, and Tomas Mikolov. 2016. Fasttext.zip: Compressing text classification models. *arXiv preprint arXiv:1612.03651*.
- Joarder Kamruzzaman and Ruhul A Sarker. 2003. Forecasting of currency exchange rates using ann: A case study. In *International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003*, volume 1, pages 793–797. IEEE.
- Raehyun Kim, Chan Ho So, Minbyul Jeong, Sanghoon Lee, Jinkyu Kim, and Jaewoo Kang. 2019. Hats: A hierarchical graph attention network for stock movement prediction. *arXiv preprint arXiv:1908.07999*.
- Yoon Kim. 2014. [Convolutional neural networks for sentence classification](#).
- De-Rong Kong and Tse-Chun Lin. 2021. Alternative investments in the fintech era: The risk and return of non-fungible token (nft). Available at SSRN 3914085.
- Miron Bartosz Kursa, Aleksander Jankowski, and Witold R. Rudnicki. 2010. Boruta - a system for feature selection. *Fundam. Informaticae*, 101:271–285.
- Xiaodong Li, Pangjing Wu, and Wenpeng Wang. 2020. Incorporating stock prices and news sentiments for stock market prediction: A case of hong kong. *Information Processing & Management*, 57(5):102212.
- Ying Liu, Benfu Lv, Geng Peng, and Qingyu Yuan. 2012. A preprocessing method of internet search data for prediction improvement: application to chinese stock market. In *Proceedings of the Data Mining and Intelligent Knowledge Management Workshop*, pages 1–7.
- Rui Luo, Weinan Zhang, Xiaojun Xu, and Jun Wang. 2018. A neural stochastic volatility model. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.
- Thang Luong, Hieu Pham, and Christopher D. Manning. 2015. [Effective approaches to attention-based neural machine translation](#). In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1412–1421, Lisbon, Portugal. Association for Computational Linguistics.
- Mieszko Mazur. 2021a. Non-fungible tokens (nft). the analysis of risk and return.
- Mieszko Mazur. 2021b. Non-fungible tokens (nft). the analysis of risk and return. Available at SSRN 3953535.
- Hongyuan Mei and Jason Eisner. 2017. [The neural hawkes process: A neurally self-modulating multi-variate point process](#).
- Matthieu Nadini, Laura Alessandretti, Flavio Di Giacinto, Mauro Martino, Luca Maria Aiello, and Andrea Baronchelli. 2021. [Mapping the NFT revolution: market trends, trade networks, and visual features](#). *Scientific Reports*, 11(1).
- Dat Quoc Nguyen, Thanh Vu, and Anh Tuan Nguyen. 2020. BERTweet: A pre-trained language model for English Tweets. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 9–14.
- Mahla Nikou, Gholamreza Mansourfar, and Jamshid Bagherzadeh. 2019. Stock price prediction using deep learning algorithm and its comparison with machine learning algorithms. *Intelligent Systems in Accounting, Finance and Management*, 26(4):164–174.
- NonFungible. 2021. Yearly nft market report. <https://nonfungible.com/reports/2021/en/yearly-nft-market-report>.

- Idris Rabi'u, Naomie Salim, Aminu Da'u, and Akram Osman. 2020. Recommender system based on temporal models: a systematic review. *Applied Sciences*, 10(7):2204.
- Francesco Rundo, Francesca Trenta, Agatino Luigi di Stallo, and Sebastiano Battiato. 2019. Machine learning for quantitative finance applications: A survey. *Applied Sciences*, 9(24):5574.
- Kentaro Sako, Shin'ichiro Matsuo, and Sachin Meier. 2021. Fairness in ERC token markets: A case study of cryptokitties. In *International Conference on Financial Cryptography and Data Security*, pages 595–610. Springer.
- Serkan Savaş. 2021. Analysis of the social media impact on the popularity of crypto-currencies. In *2021 6th International Conference on Computer Science and Engineering (UBMK)*, pages 67–72. IEEE.
- Ramit Sawhney, Shivam Agarwal, Vivek Mittal, Paolo Rosso, Vikram Nanda, and Sudheer Chava. 2022. Cryptocurrency bubble detection: A new stock market dataset, financial task & hyperbolic models. *arXiv preprint arXiv:2206.06320*.
- Ramit Sawhney, Shivam Agarwal, Megh Thakkar, Arnav Wadhwa, and Rajiv Ratn Shah. 2021a. Hyperbolic online time stream modeling. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1682–1686.
- Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa, Tyler Derr, and Rajiv Ratn Shah. 2021b. [Stock selection via spatiotemporal hypergraph attention network: A learning to rank approach](#). *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1):497–504.
- Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa, and Rajiv Shah. 2021c. [Tec: A time evolving contextual graph model for speaker state analysis in political debates](#). In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, pages 3552–3558. International Joint Conferences on Artificial Intelligence Organization. Main Track.
- Ramit Sawhney, Harshit Joshi, Saumya Gandhi, and Rajiv Shah. 2020. A time-aware transformer based model for suicide ideation detection on social media. In *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, pages 7685–7697.
- Ramit Sawhney, Arnav Wadhwa, Shivam Agarwal, and Rajiv Ratn Shah. 2021d. [FAST: Financial news and tweet based time aware network for stock trading](#). In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2164–2175. Online. Association for Computational Linguistics.
- Robert P Schumaker and Hsinchun Chen. 2009. Textual analysis of stock market prediction using breaking financial news: The azfin text system. *ACM Transactions on Information Systems (TOIS)*, 27(2):1–19.
- Sreelekshmy Selvin, R Vinayakumar, EA Gopalakrishnan, Vijay Krishna Menon, and KP Soman. 2017. Stock price prediction using lstm, rnn and cnn-sliding window model. In *2017 international conference on advances in computing, communications and informatics (icacci)*, pages 1643–1647. IEEE.
- Valentina Semenova and Julian Winkler. 2021. Social contagion and asset prices: Reddit's self-organised bull runs.
- Yauheniya Shynkevich, T Martin McGinnity, Sonya A Coleman, Ammar Belatreche, and Yuhua Li. 2017. Forecasting price movements using technical indicators: Investigating the impact of varying input window length. *Neurocomputing*, 264:71–88.
- Paul Slovic and Baruch Fischhoff. 1977. On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4).
- Sean J Taylor and Benjamin Letham. 2017. [Forecasting at scale](#).
- Lauren van Haaften-Schick and Amy Whitaker. 2021. From the artist's contract to the blockchain ledger: New forms of artists' funding using equity and resale royalties. *Social Science Research Network*.
- Jelmer van Slooten. 2022. Predictive value of tweet sentiment on the bored ape yacht club's trading volume and floor price.
- Manuela Veloso, Tucker Balch, Daniel Borrajo, Prashant Reddy, and Sameena Shah. 2021. Artificial intelligence research in finance: discussion and examples. *Oxford Review of Economic Policy*, 37(3):564–584.
- Victor von Wachter, Johannes Rude Jensen, Ferdinand Regner, and Omri Ross. 2022. Nft wash trading: Quantifying suspicious behaviour in nft markets. *arXiv preprint arXiv:2202.03866*.
- Qin Wang, Rujia Li, Qi Wang, and Shiping Chen. 2021. Non-fungible token (nft): Overview, evaluation, opportunities and challenges. *ArXiv*, abs/2105.07447.
- Martin Westerkamp, Friedhelm Victor, and Axel Küpper. 2018. Blockchain-based supply chain traceability: Token recipes model manufacturing processes. *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 1595–1602.
- Amy Whitaker. 2019. Art and blockchain: A primer, history, and taxonomy of blockchain use cases in the arts. *Artivate*, 8:21 – 46.

Joshua T White, Sean Wilkoff, and Serhat Yildiz. 2022. The role of the media in speculative markets: Evidence from non-fungible tokens (nfts). *Available at SSRN 4074154*.

Yumo Xu and Shay B. Cohen. 2018. [Stock movement prediction from tweets and historical prices](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1970–1979, Melbourne, Australia. Association for Computational Linguistics.

Xiaoting Ying, Cong Xu, Jianliang Gao, Jianxin Wang, and Zhao Li. 2020. Time-aware graph relational attention network for stock recommendation. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 2281–2284.

Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, and Stéphane Deny. 2021. Barlow twins: Self-supervised learning via redundancy reduction. *arXiv preprint arXiv:2103.03230*.

Xu Zhong and Michael Raghieb. 2019. Revisiting the use of web search data for stock market movements. *Scientific reports*, 9(1):1–8.

Haoxiang Zhu. 2014. Do dark pools harm price discovery? *The Review of Financial Studies*, 27(3):747–789.

Simiao Zuo, Haoming Jiang, Zichong Li, Tuo Zhao, and Hongyuan Zha. 2020. Transformer hawkes process. In *International Conference on Machine Learning*, pages 11692–11702. PMLR.

A Experimental Setup

Parameter	Value
Optimizer	Adam
Learning Rate	2e-4
Batch Size	64
$\beta_1, \beta_2, \epsilon$	0.9, 0.999, 1e-6
# Epochs	20
Evaluation Metric	MSE/Macro F1
Base Model	BERTweet
Classifier (over architecture)	Linear layer
Number of Parameters	4,817,035
Hardware	Nvidia Tesla T4

Table 5: Model and training setup for TA-NFT.

Collection	# of NFTs	# of tweets	# of transactions
ON1 Force	7,777	11,153	8,473
Bored Ape Yacht Club	10,000	28,651	19,472
Cool Cats NFT	9,933	363,506	16,890
CrypToadz by GREMPLIN	7,025	134,339	9,408
CyberKongz	5,000	298,710	4,357
DeadFellaz	9,999	65,158	14,489
FLUF World	10,000	67,379	10,059
Hashmasks	16,370	92,903	16,642
Loot	7,779	393	6,642
Mutant Ape Yacht Club	17,961	5,154	14,819
Meebits	20,000	108,237	13,221
Pudgy Penguins	8,888	2,017	15,997
SupDucks	10,001	169,909	12,965
VOX Collectibles	8,888	2,190	11,787
World of Women	10,000	4,728	13,314
Total	159,621	1,354,427	188,535

Table 6: NFT-collection wise data distribution.

Task	# of data points
Daily Average Price Prediction	2,679
Price Movement Classification	188,535

Table 7: Task-wise data distribution.

B Dataset Details

A detailed collection-wise breakdown of the collected data is given in Table 6. In addition to this, Table 7 gives task-wise distribution (number of data points) for the tasks defined above.