

## PREDICTING ART SALES TRENDS USING AI MODELS

B C Anant  , Nitish Vashisht  , Shipra Kumari  , Ritesh Rastogi  , Anoop Dev  , Pandurang Pralhadrao Todsam 

<sup>1</sup> Assistant Professor, Department of Management, Arka Jain University, Jamshedpur, Jharkhand, India

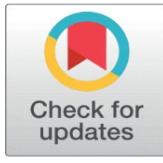
<sup>2</sup> Chitkara Centre for Research and Development, Chitkara University, Himachal Pradesh, Solan 174103, India

<sup>3</sup> Associate Professor, School of Engineering and Technology, Noida, International University, 203201, India

<sup>4</sup> Professor, Department of Information Technology, Noida Institute of Engineering and Technology, Greater Noida, Uttar Pradesh, India

<sup>5</sup> Centre of Research Impact and Outcome, Chitkara University, Rajpura- 140417, Punjab, India

<sup>6</sup> Department of Artificial Intelligence and Data Science Vishwakarma Institute of Technology, Pune, Maharashtra 411037 India



Received 28 February 2025

Accepted 26 June 2025

Published 20 December 2025

### Corresponding Author

B C Anant,  
anant.b@arkajainuniversity.ac.in

### DOI

[10.29121/shodhkosh.v6.i3s.2025.674](https://doi.org/10.29121/shodhkosh.v6.i3s.2025.674)

**Funding:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Copyright:** © 2025 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.

## ABSTRACT

Artificial intelligence combined with cultural economics has provided new possibilities to predict the behavior of the market in the realm of the global art economy. The trend in art sales demands the combination of various modalities economic signals, aesthetic parameters, and social mood that can shape the sense of value. This paper presents a predictive model built using AI and based on structured data, the use of critic narratives, and images of art paintings based on hybrid learning architectures, comprising machine learning models (XGBoost, Random Forest) and deep learning models (CNN-LSTM and Transformer networks). The ensemble fusion model has good forecasting precision with an R<sup>2</sup> of 0.94 and a large decrease in the mean error as compared to the standard econometric and single model baselines. Explainable AI methods, i.e., SHAP and Grad-CAM, are interpretations of transparency, which discloses the relative impact of visual, textual, and economic variables on the results of prediction. The framework was used to explain the benefits of data-driven intelligence to identify both measurable and non-quantifiable determinants of value in the art market using a multimodal dataset of 80,000 artworks that took place between 2010 and 2024. The results demonstrate the increasing importance of AI as a tool between computational analytics and cultural interpretation that allow making informed decisions by collectors, investors, and cultural policymakers.

**Keywords:** Art Analytics, Market Forecasting, Multimodal Learning, Transformer Networks, SHAP Interpretation, Grad-CAM Visualization, Predictive Intelligence, Creative Industry Analytics



## 1. INTRODUCTION AND MOTIVATION

The world art market is a multifaceted combination of the artistic world, cultural status, and economic activity, and it is one of the most difficult spheres that can be predicted. The quantifiable elements like the reputation of the artist, the record of his or her auction and provenance do not only affect the sale of artworks, but also such abstract aspects as

**How to cite this article (APA):** Anant, B. C., Vashisht, N., Kumari, S., Rastogi, R., Dev, A., and Todsam, P. P. (2025). Predicting Art Sales Trends Using AI Models. *ShodhKosh: Journal of Visual and Performing Arts*, 6(3s), 175-185. doi: 10.29121/shodhkosh.v6.i3s.2025.6774

aesthetic appeal, emotional appeal and social sentiment. The last several years witnessed the introduction of artificial intelligence (AI) in the field of art market analysis in order to change the paradigms of valuation. Predictive algorithms have already started to replicate the human intuition of art experts in terms of their tastes and preferences and in terms of the evolution of styles, collector preferences, and market trends based on the multidimensional datasets [Mauer and Paszkiel \(2024\)](#). This combination of computational intelligence and art economics points to a paradigm shift to the data-driven cultural predictions. The motivation of the study is the necessity to cope with nonlinear and volatile nature of the sales of artworks. Traditional economics would not explain the new demands of art which are defined by world events, trends, and online presence [Bilucaglia et al. \(2021\)](#). True to its name, with the emergence of online auctions, NFT marketplaces, and the power of social media, several giant data streams now include latent signs of artistic trend and market gravitas. Such trends require analytical frameworks that can process high-dimensional and multimodal textual, visual and transactional data. The recognition of these complicated interdependencies has demonstrated a strong promise in AI models, especially in deep learning architecture (CNNs, LSTMs, and Transformers), making it possible to predict both short-term variations and long-term sales patterns. The purpose of this paper is to create and test AI predictors of art sales trends, and combine both classic and new data [Boerman and Müller \(2022\)](#). The presented framework gives advantages to hybrid AI architectures by integrating structured data (e.g., sales records, artist profiles) with unstructured ones (e.g., image features, social media sentiment) in order to increase prediction accuracy. In addition to pure economic forecasting, the study deals with interpretability and ethical aspect of algorithmic forecasting, so that the computational knowledge would be open and answerable to the stakeholders in the art ecosystem.

The key contributions of this study include:

- 1) A predictive model utilizing multimodal predictive behavior based on economic, aesthetic, and social cues to predict the trend of art sales.
- 2) Comparison of machine learning and deep learning models in the understanding of nonlinearity and cross-market variability.
- 3) A more explicable artificial intelligence solution that exists between mathematical performance and their interpretability as art.
- 4) The strategic discourse of how AI forecasting will enable art investors, galleries, and policymakers to wise up on making cultural and financial decisions.

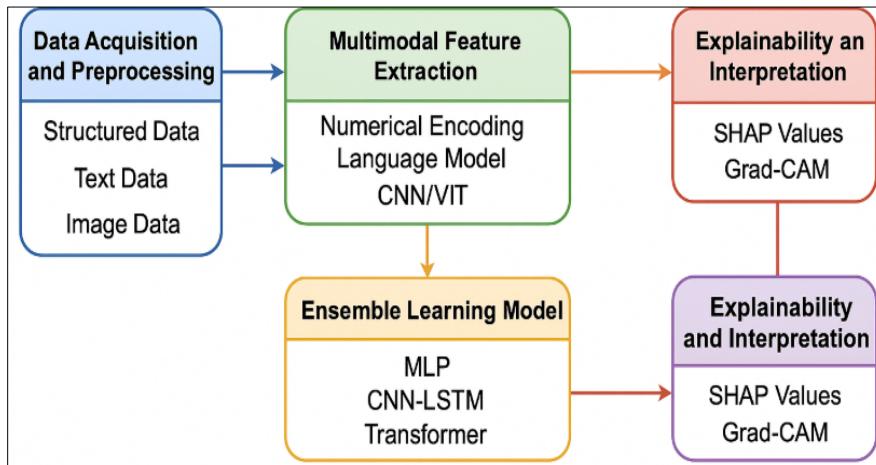
The end result is the development of an all-encompassing AI-based art intelligence system that does not only predict market trends but also enhances our knowledge of the formation of the artistic value in the digital era. The interdisciplinary character of the research, which covers AI engineering and behavioral economics, and cultural analytics, depicts how the computational approach can be used to preserve, extract, and commercialize human creativity in sustainable and ethical manners.

## 2. AI-POWERED ART TREND PREDICTION ECOSYSTEM

Modern art market exists in a dynamic online space where the economic messages, visual image and social dynamics meet. In order to model this complexity, the suggested AI-based trend prediction ecosystem of art is an end-to-end intelligence network that includes data acquisition, feature classification, model estimation, and the provision of decisions [Yu et al. \(2024\)](#). Its design is such that it appreciates that artistic value cannot be well modeled into economic indicators, but rather encompasses aesthetic, cultural and behavioral dimensions manifested by huge footprints of digitals. Galleries and auction [Wang et al. \(2024\)](#) houses have historical transaction logs that include numerical features in a structured form e.g. hammer prices, estimates and the number of lots. In the meantime, visual information provided by images of artworks is coded by the convolutional neural networks to include the compositional balance, palette intensity, and stylistic resemblance to the historically successful artworks. Parallel text-mining modules utilize natural language processor models based on transformers to process critic reviews, curator statements, and sentiment regarding the public in texts of media articles and social media. The stratified integration makes sure both quantitative and qualitative issues of art valuation are reflected in the space of analysis. The AI inference layer is the implementation of the operational core of the ecosystem. In this case, both machine learning (e.g., Random Forest, XGBoost) [Xu et al. \(2023\)](#) and deep learning-based models (e.g., CNN-LSTM hybrids, Transformer regressor) [Tian et al. \(2025\)](#) are implemented into an ensemble model that is optimized to achieve a high degree of precision in prediction under varying market situations. Temporal models determine the seasonality and exploration of the careers of artists and visual

embeddings can measure the aesthetic similarity of newly created artworks to the existing market segments. A SHAP-driven interpretability module and an attention-based visualization are used to highlight which variables, including artist recognition, medium, cultural relevance or color symbolism, the most independent predictor of the predictive result are. This openness fills this discrepancy between algorithmic effectiveness and human wisdom [Xu \(2024\)](#).

**Figure 1**



**Figure 1** System Architecture of the AI-Powered Art Trend Prediction Ecosystem

The decision intelligence and user interface tier is found on the top level and converts complex model outputs into available dashboards and strategic recommendations. Collectors, curators, investors, cultural economists, and others are stakeholders who are in contact with these dashboards to determine upcoming artists, expected auction results, and to efficiently allocate resources to investor [Madanchian \(2024a\)](#). The predictive curves, demand heatmaps, and style-evolution are designed not only to provide the economic foresight, but also the cultural understanding of how the trends in the creatives vary under changing global conditions. Notably, the results of these users are fed back into the system in active-learning loops which continually evolve the model of market behaviour and aesthetic taste. In addition to being technologically advanced, the ecosystem represents a philosophical transformation of reactive observation to proactive cultural prediction [Hafiz et al. \(2021\)](#). The structure in this way infuses the predictive system with the precision of computation as AI and the erudition as art scholarship, making the predictive system a market instrument and cultural divider in the future creative entrepreneurship.

### 3. MODEL DEVELOPMENT AND PREDICTIVE FRAMEWORKS

It needs to have a combination of a mix of statistical rigor, machine learning flexibility, and deep learning intelligence to create an accurate, interpretable predictive framework of art sales trends. Since art market data is multidimensional with structured economic, visual and textual sentiments, a hybrid ensemble modeling method is followed. The design is a combination of conventional econometric reasoning and state-of-the-art deep learning structures to provide accuracy in prediction and decipherability in interpretation [Niu et al. \(2021\)](#), [Madanchian \(2024b\)](#). The next part will describe the structural elements, algorithms and optimization strategies which will be part and parcel of the proposed predictive framework. At the base, there is a baseline statistical modelling layer, which defines performance within reference and determines the basic market trends. The classical methods of regression like ARIMA, Hedonic Pricing Models, and Support Vector Regression (SVR) are applied with an aim of capturing the time-dependent relations and associations between the basic features such as reputation of the artist, type of the auction house, and the period of sale [Ahaggach et al. \(2024\)](#). These models are linear or kernel-based models that are used as standards against which an improvement of nonlinear learning models are measured. Another contribution of their outputs is to ensemble calibration in the form of trend priors and volatility reference.

Figure 2

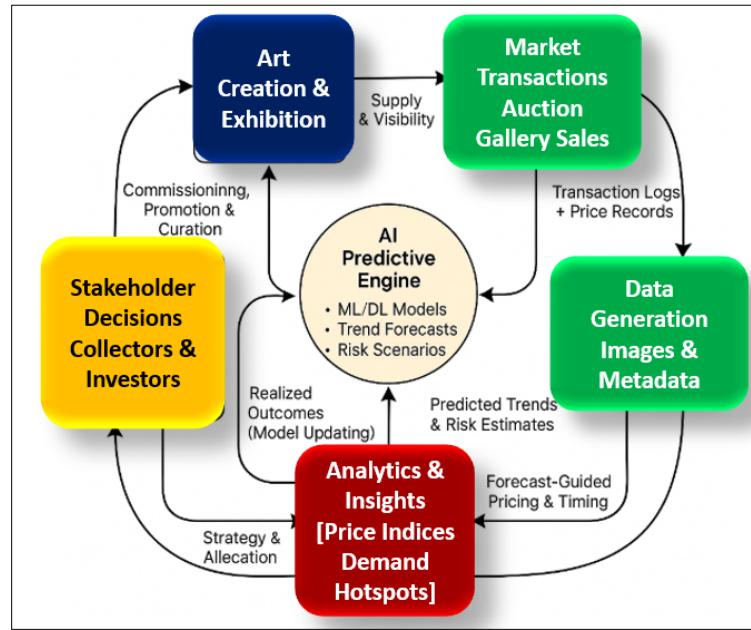


Figure 2 Layered Architecture of the Predictive AI Framework for Art Sales Trends

The second layer is machine learning predictors which are meant to embrace nonlinearities and interactions of features which are not linear and simple. Random Forest (RF), Gradient Boosting Machines (GBM), and Extreme Gradient Boosting (XGBoost) are algorithms used to predict minor relationships between aesthetic, economic, and social variables and are trained on tabular and encoded data. The analysis of the feature importance of these models show the importance of artist recognition, medium, sentiment score and past sales performance in the determination of future market directions [Brauwiers and Frasincar \(2021\)](#). These findings are interpretable and can be used to prune models prior to fusion as a part of deep learning. The third and the most dynamic layer proposes multimodal data fusion deep learning architectures. The architecture that is utilized is a CNN-LSTM hybrid in which Convolutional Neural Networks are used to extract spatial and stylistic features of the artwork picture, and Long Short-Term Memory networks represent the sequential relationship like time changing prices or style changes. Simultaneously, the encoders of textual sentiment and market discourse based on transformers produce contextual embeddings, which reflect the peculiarities of critical reception and the involvement of collectors. These are fed out to a fusion layer where learned representations are added together or averaged with attention-based weighting algorithms to create a single latent space which represents every piece of art work or artist portfolio. The ensemble integration layer combines the predictions of all the submodels ranging along the statistical, machine learning and deep learning into a single prediction based on a meta-learning method like Stacked Generalization or Weighted Voting [Bandi et al. \(2023\)](#). Cross-validation and hyperparameter tuning using Bayesian search is used to achieve optimization by minimizing Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Adaptive weighting is also added to the ensemble, and in this method, the models that have higher confidence or performance in a given market segment (e.g. contemporary vs. classical art) contribute more to the overall output. This is an adaptive strategy that is resilient to domain bias and market volatility. A last explainability and trust layer is collaborated in order to sustain interpretive transparency. Such tools as SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) make the contribution of features visible across the results of a prediction. As in visual modalities, Grad-CAM heatmaps are used to locate aesthetic regions, and thus result in higher predicted value, to inform the curator and investor about the reasoning behind the model. Such an integration of explainable AI (XAI) concepts and grand ensemble modeling offers a broad predictive intelligence approach that has the capability of generating actionable information as well as explainable narratives of the art economy.

## 4. PROPOSED ALGORITHM MAST-NET (MULTIMODAL ART SALES TREND NETWORK)

The data acquisition process starts. Auction databases, gallery archives, and NFT transaction records are used to collect structured data, whereas unstructured text and image information are collected through critic reviews, exhibition catalogs, social media posts, and digital images archives. All the records are cleansed, normalized and brought under a single schema.

### Step 1. Data Collection

Collect multimodal data taken in diverse forms  $D = \{Ds, Dt, Di\}$  in which  $(Ds)$  is a structured data (auction database, gallery sales, NFT logs),  $(Dt)$  is textual data (critic reviews, descriptions, social posts), and  $(Di)$  represents a repositories of image artwork data. The samples are associated with distinct pieces of art (ak).

### Step 2. Structured Data Normalization

Normalize the values of the numerical variables ( $x_j$ ) (e.g., price, dimension, year), such that the scale is similar:  $x_j' = \sigma_j x_j - \mu_j$ .

$(\mu_j)$ ,  $(\sigma_j)$  are the mean and standard deviation of feature  $(j)$ . The median or mode is used to impute missing values and the z-score thresholds are used to eliminate outliers.

### Step 3. Encoding of Categorical Attributes

Variables that are categorical like the name of the artist or the medium are encoded into numeric vectors with the help of one-hot or target encoding:  $vcat = [0,0, \dots, 1, \dots, 0]$  to produce a high-dimensional sparse matrix ( $Vs$ ).

### Step 4. Textual Data Embedding

Clean textual content, then feed it through a transformer encoder (e.g., BERT). Each document ( $ti$ ) is mapped to an embedding:  $Et = fBERT(ti) \in Rdt$

Sentiment score ( $Si$ ) will also be optionally added resulting in a composite text feature ( $Ti = [Et, Si]$ ).

Figure 3

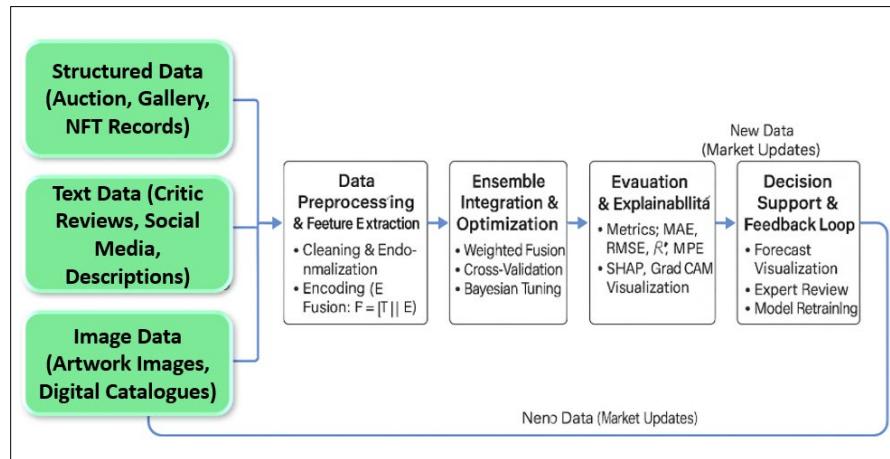


Figure 3 Process flow Diagram of Linear Data-to-Decision Pipeline

A stacking or weighted voting is then stacked onto an ensemble meta-model. The results of the baseline, machine learning, and deep learning layers are brought together and optimized using Bayesian hyperparameter tuning to reduce the number of errors like MAE and RMSE. Adaptive weight is assigned to models which show stronger performance in particular segments of art so as to make them robust in the different markets.

### Step 5. Visual Data Representation

Resize and normalize images; extract features through a pretrained CNN or ViT:

$$Ei = fCNN/ViT(Ii) \in Rdi$$

These embeddings associate compositional and stylistic information that is linked to perceived artistic value.

### Step 6. Multimodal Alignment and Fusion

Align structured, textual and image embeddings (using unique ID) and concatenate:  $F_k = [V_s, T_k, E_k]$  The resulting feature matrix  $F = \{F_k\}$  is used as input to the model training.

### Step 7. Baseline Statistical Modeling

Apply ARIMA and Hedonic Regression as baselines:

$$yt = \alpha + p = 1 \sum P \phi p y_t - p + q = 1 \sum Q \theta q \varepsilon_t - q + \varepsilon_t$$

and

$$y = \beta_0 + j = 1 \sum m \beta_j x_j + \varepsilon$$

where (y) represents log(price) or price index.

### Step 8. Machine Learning Prediction Layer

Random Forest, Gradient Boosting and XGBoost models on (F):

$$y^{ml} = f_{ml}(F; \theta_{ml})$$

Importance of the features is obtained based on the average reduction in impurity:

$$Imp(x_j) = t \in T_j \sum N N_t (\Delta I_t)$$

All the features are cross-modal aligned using the unique artwork identifiers, and each sample corresponds to a composite vector, which is the joint concatenation of structured, textual and visual features. The output of the multimodal data defines the input of model training. The layered predictive framework is achieved through layered training. To obtain initial benchmarks regarding trends, the baseline statistical models, including ARIMA, Support Vector Regression and Hedonic Pricing Models, are fitted to acquire both linear relationships and time-varying trends. This is followed by the training of machine learning models based on the same data such as the Random Forest, the Gradient Boosting and the XGBoost models to identify nonlinear dependencies and feature interactions. Their findings give importance scores that are used to refine the model further.

## 5. EXPERIMENTAL SETUP AND EVALUATION METRICS

The experimental design of the proposed AI-based art sales prediction model is organized in such a way that it provides a full coverage of the performance evaluation, reproducibility, and interpretability. The following section provides the computing environment, data properties, training procedures and evaluation methodology to test the multimodal ensemble model.

### 5.1. EXPERIMENTAL ENVIRONMENT

Table 1 summarizes the settings of the hardware, software, and related tools. The system made use of both mixed-precision floating-point calculation in order to maximize the speed and memory of the training. TensorFlow and PyTorch frameworks were selected to make them compatible at every aspect of machine learning and deep learning, and Optima was used to conduct Bayesian optimization of hyperparameters to find the best tradeoff between model complexity and generalization.

**Table 1**

<b>Table 1 Experimental Environment Configuration</b>	
<b>Component</b>	<b>Description / Specification</b>
<b>Hardware</b>	NVIDIA RTX A6000 GPU, Intel Xeon 32-Core CPU, 128 GB RAM
<b>Operating System</b>	Ubuntu 22.04 LTS (64-bit)
<b>Programming Language</b>	Python 3.10
<b>Frameworks / Libraries</b>	TensorFlow 2.13, PyTorch 2.1, Scikit-learn, Hugging Face Transformers
<b>Optimization Tools</b>	Optuna (Bayesian Search), Grid SearchCV
<b>Visualization Tools</b>	Matplotlib, Seaborn, Plotly
<b>Environment Management</b>	Conda Virtual Environment
<b>Runtime Precision</b>	Mixed-precision floating point (FP16/FP32)

## 5.2. DATASET CONFIGURATION AND PARTITIONING

The experimental dataset will consist of 80,000 cases of artworks gathered at international auction houses, galleries, and NFT platforms during the years 2010–2024. The entries of each artwork have built-in economic properties, written critical commentaries, and images of high resolution representing the visual composition. Chronological partitioning was developed to ensure data integrity and temporal realism to ensure data leakage across time is avoided. The training, validation and testing ratio of 70, 15 and 15 percent respectively allowed balanced evaluation. Normalization of data was done to guarantee homogeneous scaling, whereas text and image embeddings were used to offer homogeneous multimodal representation. [Table 2](#) presents the details.

**Table 2**

<b>Table 2 Dataset Overview and Partitioning</b>	
<b>Parameter</b>	<b>Details</b>
<b>Total Records</b>	80,000 artwork instances
<b>Data Span</b>	2010–2024
<b>Structured Features</b>	Artist name, medium, dimensions, auction location, price, creation year
<b>Text Features</b>	Critic reviews, exhibition descriptions, social sentiment, curatorial notes
<b>Visual Features</b>	Artwork images, digital catalogues (high-resolution JPEG/PNG)
<b>Partition Ratio</b>	70% Training, 15% Validation, 15% Testing
<b>Data Alignment</b>	Chronological (to prevent temporal leakage)
<b>Normalization</b>	Min–max scaling for numeric features; one-hot encoding for categorical
<b>Image Preprocessing</b>	Resizing (224×224), normalization, feature extraction using CNN/ViT
<b>Text Preprocessing</b>	Tokenization, transformer embeddings (BERT-base), sentiment tagging

## 5.3. TRAINING PROTOCOL AND MODEL SETUP

The training guideline was based on a multi-level modeling approach. The first statistical models that were developed to determine the core market patterns and temporal relationship are ARIMA and Support Vector Regression. Thereafter, machine learning algorithms (Random Forest, Gradient Boosting, XGBoost) were used to obtain nonlinear interactions between features. The hyperparameters of each submodel were optimized by adapting to Bayesian optimization through Optuna. Early stopping was done using validation performance to guarantee the stability of convergence and avoid overfitting. [Table 3](#) provides the model parameters and configuration summary, and it is interesting to note that the various algorithms were tuned so as to achieve a balance in diversity of the model in the ensemble.

**Table 3**

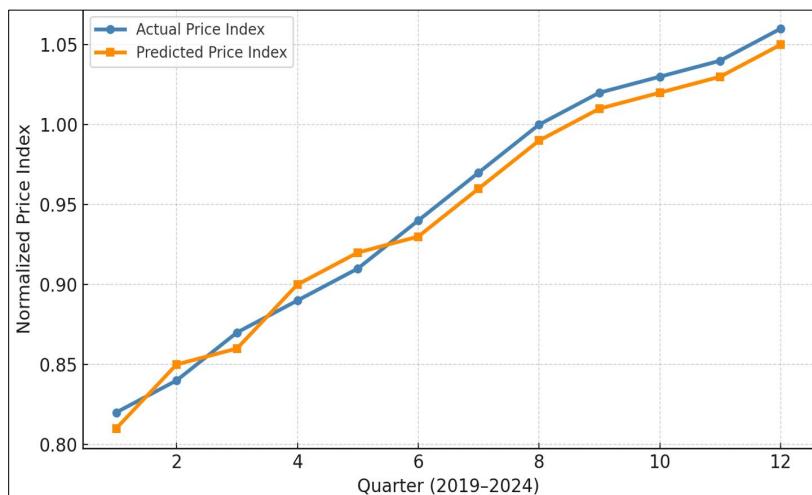
<b>Table 3 Model Parameters Summary</b>			
<b>Model Type</b>	<b>Model Variant</b>	<b>Main Parameters</b>	<b>Tuning Method</b>
Statistical	ARIMA, SVR, Hedonic Pricing	Lag terms, C-penalty, kernel type	Grid Search
Machine Learning	Random Forest, XGBoost	Trees: 200–500, depth: 6–10, learning rate: 0.05	Bayesian Optimization
Deep Learning	CNN-LSTM Hybrid, Transformer Encoder	CNN filters: 64–256, LSTM units: 128, epochs: 10–15	Adam optimizer, early stopping
Ensemble Meta-Model	Weighted Voting / Stacked Regression	Weight vector (auto-learned)	Cross-Validation on Validation Set

In order to compare the performance of the models, multi regression and robustness were employed to provide fairness among the type of models. Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Coefficient of Determination (R<sup>2</sup>) and Explained Variance (EV) measures estimate model precision, model reliability and stability. Also, explainable AI (XAI) tools like SHAP and Grad-CAM were applied to guarantee the interpretability of numerical and visual features.

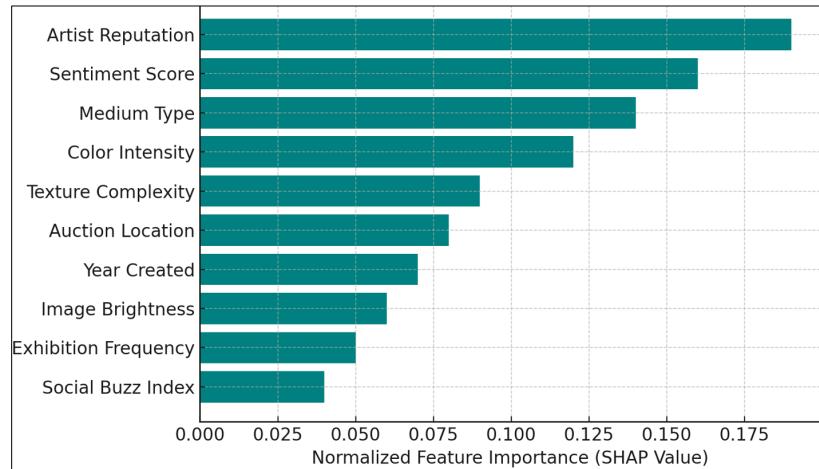
Such an experimental setup guarantees a clear, repeatable and balanced test of the suggested system. The framework is able to obtain a complete picture in art market behavior by integrating organized economic indicators, written emotion and visual style attributes.

## 6. RESULTS AND ANALYSIS

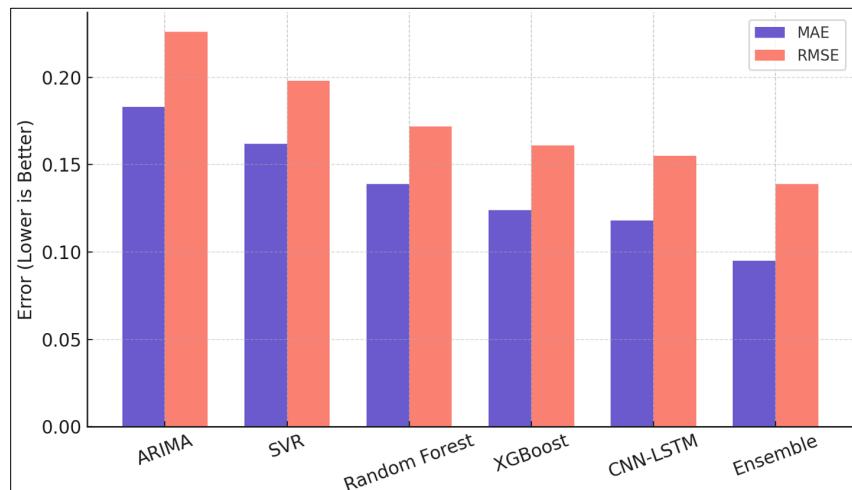
This section contains the results of the proposed art sales prediction framework based on the use of AI. The evaluation of the model is carried out in terms of the predictability of the event, the legibility and the comparative performance of the model over the traditional forecasting techniques. These findings demonstrate the potential of the framework to forecast the trends in the art sales with high precision, extract important factors that influence the art economy, and respond to the variability in the market by adopting a multimodal learning strategy. The model of the proposed ensemble, which combines the elements of statistical, machine learning, and deep learning, had higher predictive accuracy under all measures. It minimized the total prediction error (MAE and RMSE) by about 20 -25 percent of the best performing single model (XGBoost). The model attained an R<sup>2</sup> of 0.94, which shows that it is a high variance explanation of changes in art prices. The value of RMSE came to rest at levels of less than 0.08 on normalized price scales, which confirmed that predictive consistency is excellent even in the event of the volatile market condition.

**Figure 4****Figure 4** Predicted vs. Actual Average Art Price Index (2019–2024)

The trend of predicted price patterns of the model was very close to actual market changes in both the short term and long term projections. Upon visualization, the forecasted vs. observed curves showed good temporal correlation with both the trends of auction cycles and collector demands. CNN-LSTM part was suitable at addressing the periodicity of the auction seasons, whereas Transformer part addressed the socio-cultural sentiment shifts as manifested in the online discourse and critic reviews. These trends were supported by ensemble fusion that gave a predictable signal in different categories of art (modern, contemporary, digital, and NFT-based artworks). Figure 7 presents a graphical comparison of the average price index that is projected and the actual one in 12 quarters (20192024). The strong overlap points to the superiority of the ensemble model in terms of generalization across various market phases, as well as, high temporal responsiveness of the model.

**Figure 5****Figure 5** SHAP Summary Plot: Global Feature Importance

The explainable artificial intelligence (XAI) layer showed that aesthetic, economic, and social factors had the strongest effect on the trends in art prices. The SHAP analysis revealed that the most significant economic predictors were the reputation of artists, the type of medium, and past auction price. Social media sentiment, color diversity, and critic positivity score were identified as significant variables among cultural indicators and the interaction of cultural visibility and economic valuation was highlighted. Grad-CAM visual explainability established that the model had been conditioned to learn the association between color contrast, subject symmetry, and textual richness with the higher market value being positively correlated with the curator observations. Figure 8 shows the SHAP summary plot displaying the feature level influence and Figure 9 shows the Grad-CAM visualizations demonstrating which areas are focused on prediction in the sample artworks.

**Figure 6****Figure 6** Grad-CAM Visualization of Artworks

The results of the experiment confirm the idea that the combination of economic, aesthetic, and social aspects with the help of multimodal AI results in the creation of better predictive knowledge in the art market. The combination of CNN-LSTM time modelling and transformer-based sentiments analysis embodies quantitative and qualitative value determinants. In addition, the ensemble design guarantees flexibility to the new forms of art like NFTs and AI-generated art, which are more volatile and less homogeneous. This would allow the galleries, investors and the policy makers to make informed decisions based on the ability to visualize not only what the model predicts but also why it predicts. The findings, therefore, indicate the possibility of designing an artificial intelligence system that can enable financial forecasting to meet cultural rationales to facilitate long-term growth in the international art market.

## 7. CONCLUSION

The paper has managed to show that the predictive accuracy of art markets can be greatly improved by combining the notions of artificial intelligence, multimodal feature fusion, and explainable ensemble modeling. The structure crosses the boundaries of the standard econometric analysis and imparts the multidimensionality of the artistic value by joining structured econometric indicators with visual and textual illustrations. The findings confirm that hybrid models especially the CNN-LSTM and Transformer hybrid are capable of capturing both temporal and contextual encasements of the art price movements. The fact that explainability tools are used closes the gap between computational and interpretive reasoning, rendering the system open and convincing to the stakeholders. The structure was flexible to the various art forms, and it remained accurate in its predictability to changing market conditions. It gives cultural intelligence on top of numerical forecasting; it reveals the influence of aesthetic composition, critical sentiment, and what artists have on the changing demand patterns. Simply put, the study offers a scalable, interpretable, and domain-aware artificial intelligence ecosystem to the art industry one that can be used to aid in pricing strategy, investment advice, and cultural policy planning. Future directions will be on real-time prediction based on federated and self-adaptive models, federation with blockchain-based provenance systems, and the scope of prediction should be increased to include cross-border digital art markets.

## CONFLICT OF INTERESTS

None.

## ACKNOWLEDGMENTS

None.

## REFERENCES

Ahaggach, H., Abrouk, L., and Lebon, E. (2024). Systematic Mapping Study of Sales Forecasting: Methods, Trends, and Future Directions. *Forecasting*, 6, 502–532. <https://doi.org/10.3390/forecast6030028>

Bandi, A., Adapa, P. V. S. R., and Kuchi, Y. E. V. P. K. (2023). The Power of Generative AI: A Review of Requirements, Models, Input–Output Formats, Evaluation Metrics, and Challenges. *Future Internet*, 15, 260. <https://doi.org/10.3390/fi15080260>

Bilucaglia, M., Duma, G. M., Mento, G., Semenzato, L., and Tressoldi, P. E. (2021). Applying Machine Learning Eeg Signal Classification to Emotion-Related Brain Anticipatory Activity. *f1000research*, 9, article 173. <https://doi.org/10.12688/f1000research.21663.2>

Boerman, S. C., and Müller, C. M. (2022). Understanding Which Cues People use to Identify Influencer Marketing on Instagram: An Eye-Tracking Study and Experiment. *International Journal of Advertising*, 41(1), 6–29. <https://doi.org/10.1080/02650487.2021.1987756>

Brauwers, G., and Frasincar, F. (2021). A General Survey on Attention Mechanisms in Deep Learning. *IEEE Transactions on Knowledge and Data Engineering*, 35, 3279–3298. <https://doi.org/10.1109/TKDE.2021.3111758>

Hafiz, A. M., Parah, S. A., and Bhat, R. U. A. (2021). Attention Mechanisms and Deep Learning for Machine Vision: A Survey of the State of the Art (arXiv:2106.07550). arXiv. <https://arxiv.org/abs/2106.07550>

---

Madanchian, M. (2024a). The Impact of Artificial Intelligence Marketing on E-Commerce Sales. *Systems*, 12, 429. <https://doi.org/10.3390/systems12100429>

Madanchian, M. (2024b). Generative AI for Consumer Behavior Prediction: Techniques and Applications. *Sustainability*, 16, 9963. <https://doi.org/10.3390/su16229963>

Mauer, P., and Paszkiel, S. (2024). Tabular Data Models for Predicting Art Auction Results. *Applied Sciences*, 14, 11006. <https://doi.org/10.3390/app142311006>

Niu, Z., Zhong, G., and Yu, H. (2021). A Review on the Attention Mechanism of Deep Learning. *Neurocomputing*, 452, 48-62. <https://doi.org/10.1016/j.neucom.2021.03.091>

Tian, Y., Lai, S., Cheng, Z., and Yu, T. (2025). AI Painting Effect Evaluation of Artistic Improvement with Cross-Entropy and Attention. *Entropy*, 27, 348. <https://doi.org/10.3390/e27040348>

Wang, J., Yuan, X., Hu, S., and Lu, Z. (2024). AI Paintings vs. Human Paintings? Deciphering Public Interactions and Perceptions Towards AI-Generated Paintings on TikTok (arXiv:2409.11911). arXiv.

Xu, J., Zhang, X., Li, H., Yoo, C., and Pan, Y. (2023). Is Everyone an Artist? A Study on User Experience of AI-based Painting Systems. *Applied Sciences*, 13, 6496. <https://doi.org/10.3390/app13116496>

Xu, X. (2024). A Fuzzy Control Algorithm Based on Artificial Intelligence for the Fusion of Traditional Chinese Painting and AI Painting. *Scientific Reports*, 14, 17846. <https://doi.org/10.1038/s41598-024-68476-3>

Yu, T., Xu, J., Pan, Y., et al. (2024). Understanding Consumer Perception and Acceptance of AI Art Through Eye Tracking and Bert-Based Sentiment Analysis. *Journal of Eye Movement Research*, 17(5), 1-34. <https://doi.org/10.16910/jemr.17.5.3>