Proceedings of the 3rd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2025.27797

Quantitative Analysis and Prediction of the Popularity of Digital Media Artworks Based on Machine Learning Algorithms

Like Ma*, Xiaoli Song

School of Digital Technology & Innovation Design, Jiangnan University, Wuxi, China
**Corresponding Author. Email: christinemal@163.com*

Abstract. Exploring a quantitative assessment system for the popularity of digital media art has become a central issue linking art creation, technology application and market communication. Aiming at the bottleneck of the current algorithm, this paper proposes a LSTM algorithm based on multi-head attention mechanism optimization. The study began with a correlation analysis. The results showed that the number of interactive elements had the strongest correlation with popularity, and the absolute value of correlation coefficient reached 0.556887. Variables such as creation time, number of colors, and complexity score also have some correlation with popularity. It can be seen that the time invested in creation, the richness of color and the complexity of works will affect the popularity of works to a certain extent. Taking decision tree, random forest, CatBoost, AdaBoost, and XGBoost as comparative experimental objects, our model performed best in terms of accuracy, recall, accuracy, F1, and AUC in various indicators. Its accuracy was 0.855, which was higher than decision tree (0.709), random forest (0.803), CatBoost (0.786), AdaBoost (0.778) and XGBoost (0.744), and the overall classification accuracy was the highest. Recall and accuracy of 0.855 and 0.856, respectively, are also ahead of other models. It performs better at identifying positive samples and predicting the proportion of actual positive samples among positive samples. The F1 value of 0.855 is also higher than that of other models, and the ability to take into account both accuracy and recall is stronger. The AUC reached 0.904, surpassing random forest's 0.875 and CatBoost's 0.876, demonstrating the best ability to differentiate between positive and negative samples. In contrast, the other models are slightly inferior in various metrics, especially the overall performance of the decision tree and XGBoost is significantly lower than ours. This research outcome provides a more effective method for quantitatively assessing the popularity of digital media art. It not only provides a direction for the integration of artistic creation and technology application, but also provides a scientific basis for the formulation of market communication strategy.

Keywords: Machine learning, digital media artworks, popularity quantification

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.27797

1. Introduction

Digital media art, as a product of the fusion of technology and aesthetics, has experienced explosive growth with the development of Internet, artificial intelligence and virtual reality technology [1]. Its form has expanded from early digital images to diverse forms such as generative art, VR interactive works, NFT digital collectibles, and the communication scene has also expanded from professional exhibitions to multi-channel ecosystems including social media, online galleries, metaverse platforms [2]. However, the popularity of such works is intertwined and influenced by many factors such as creative techniques, visual characteristics, transmission paths, etc. The commercial operation of the digital art market also presents an urgent need for objective predictive models [3]. In this context, exploring a quantitative assessment system for the popularity of digital media art has become a key issue connecting artistic creation, technology application and market communication [4].

Machine learning algorithms provide technical support for solving the problem of predicting the popularity of digital media art. Compared with traditional qualitative analysis, machine learning can integrate multi-modal data such as creation duration, visual features, and interactive elements, and mine implicit correlations between variables through classification models to achieve objective prediction of the popularity of works [5]. For example, based on decision tree algorithms, key influencing factors such as artistic style and 16:9 aspect ratio can be identified and generated. Recurrent neural network (RNN) can capture the temporal characteristics during the propagation of works and improve the accuracy of dynamic prediction. This kind of technology not only provides data-driven creative reference for creators, but also helps platforms optimize recommendation mechanisms and enhance user stickiness [6].

Aiming at the bottleneck of existing algorithms, this paper proposes a LSTM algorithm based on multi-head attention mechanism optimization. When the basic LSTM processes sequence data through gated units, it assigns equal weights to all input features, making it difficult to focus on core influencing factors such as artistic style and interactive elements. This optimization scheme breaks through the limitation of feature processing of traditional models, provides a more accurate technical path for predicting the popularity of digital media art, and also provides algorithm support for accurate docking between artistic creation and market demand.

2. Data sources

This dataset is selected from an open-source dataset and contains a total of 389 samples of digital media artworks, covering 17 independent variables and 1 predictor variable. The independent variables are divided into five major categories: basic attributes (creation time, artist experience), technical attributes (AI-assisted creation, file size), visual attributes (resolution, aspect ratio, animation frame rate), content attributes (artistic style, work theme, number of interactive elements), and release attributes (type of creation tool, release platform). The predictor variable is popularity (0 indicates unpopular and 1 indicates popular). This data can be directly used to train machine learning classification models, helping to analyze the laws of digital art dissemination, optimize creative strategies or enhance the effect of platform recommendations. Some datasets are selected for display, as shown in Table 1.

DOI: 10.54254/2755-2721/2025.27797

Table 1. A partial dataset

creation _time	creation _period	artist_ experience	techniques _used	file _size	aspect _ratio	animation _frames	text _elements	interactive _elements	complexity _score	platform	popularity
103	29	3	11	67.15	1.33	212	5	12	1	Gallery Exhibition	0
52	14	7	12	63.79	1.5	2	6	7	7	NFT Marketplace	1
93	8	4	6	22.22	1.5	178	9	21	1	Social Media	1
107	22	3	1	20.8	1.33	163	6	16	6	Art Platform	0

A correlation analysis was performed on each variable to draw a correlation thermal map as shown in Figure 1. From the correlation coefficient, the number of interactive elements has the strongest correlation with heat, and the absolute value of correlation coefficient reaches 0.556887. Variables such as creation time, color quantity and complexity score also have a certain correlation with popularity, indicating that the time invested in creation, color richness and complexity of the work will all affect the popularity of the work to a certain extent.

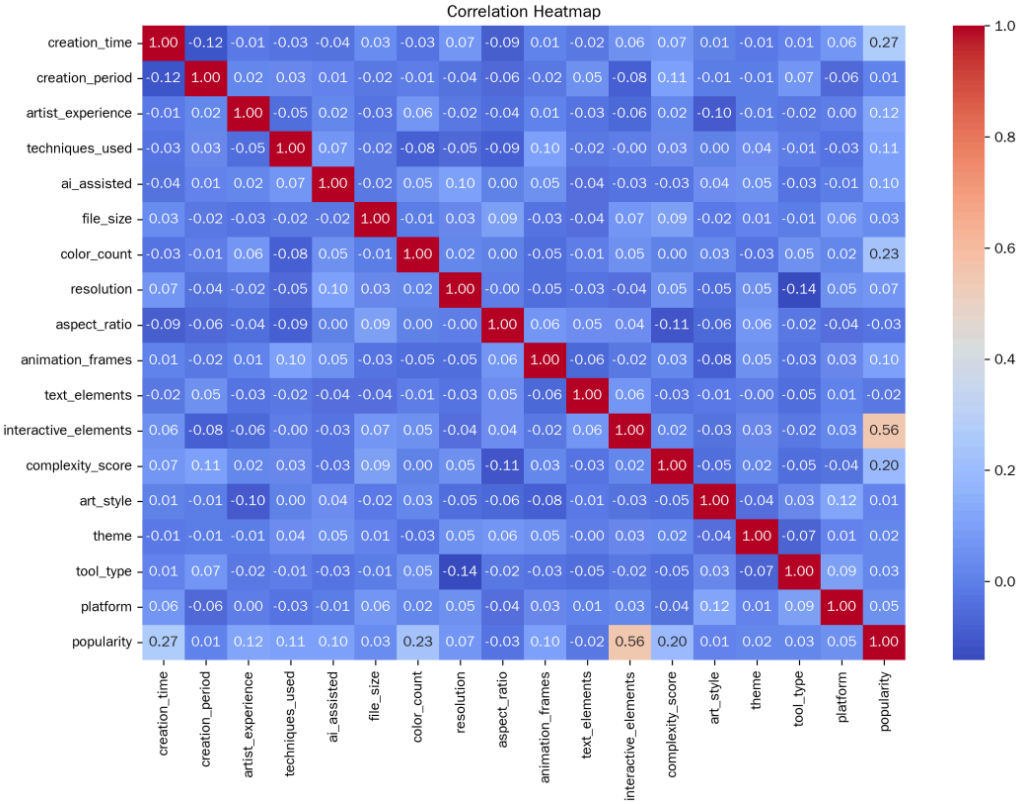

Figure 1. Correlation heat map

DOI: 10.54254/2755-2721/2025.27797

3. Method

3.1. Multi-head attention mechanism

The multi-head attention mechanism is a core component of the Transformer model. The core idea is to enhance the model's ability to model complex sequences by executing multiple attention heads in parallel to capture relevant information in different dimensions in the input data [7]. The principle can be divided into three key steps: firstly, the input query, key and value are mapped to multiple low-dimensional subspaces through linear transformation, so as to avoid the limitation of information interaction in a single high-dimensional space; Secondly, each attention head calculates the attention weight independently-the similarity score is obtained by scaling the dot product, and then it is normalized by the Softmax function to obtain the attention weight of each element to other elements. Finally, the weighted sum of the values and weights is used to generate the output of this header. Finally, the outputs of all attention heads are connected in series, and the multi-dimensional features are fused by a single linear transformation to form the final multi-head attention result. The advantage of this design is that different attention heads can focus on capturing different types of associations, enabling the model to simultaneously learn rich contextual information in the input data, significantly enhancing the understanding and modeling ability of complex sequences [8].

3.2. LSTM

Long short-term memory network (LSTM) is an improved model of recurrent neural network (RNN). Its core lies in solving the problem that the gradient disappears or bursts when traditional RNN deals with long sequences through gating mechanism, so as to effectively capture the long-distance dependence in time series data [9]. Its structure encompasses a cellular state throughout the entire sequence, similar to an information conveyor belt, which can steadily carry critical information. At the same time, the information flow is dynamically regulated through three gating units: the forgetting gate determines which historical information to discard from the cell state through the sigmoid function; The input gate first filters new information through the sigmoid function, and then generates candidate information to be added through the tanh function. The combination of the two updates the unit state. The output gate is based on the current unit state and input, controls the proportion of output information through the sigmoid function, and generates the hidden state at the current moment after being processed by the tanh function [10]. The network structure of LSTM is shown in Figure 2 [11].

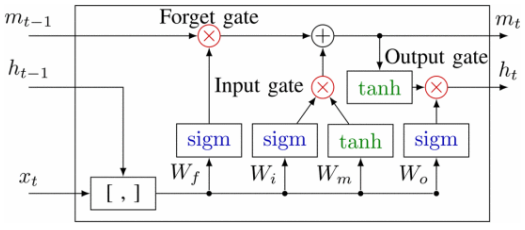

Figure 2. The network structure of LSTM

3.3. Multi-head attention mechanism-LSTM

The core of the multi-head attention mechanism in optimizing LSTM is to compensate for the lack of attention to key information in long sequences by the feature selection ability of the attention mechanism. The principle is as follows: firstly, the gating mechanism of LSTM is used to process

DOI: 10.54254/2755-2721/2025.27797

time series data, capture the long-term and short-term dependencies of the sequences, and generate hidden state sequences containing time series features; Subsequently, these hidden states are taken as inputs to the multi-head attention mechanism. Through multiple parallel attention heads, the attention weights of features for each time step in different subspaces are calculated separately to accurately locate moments or features that are more critical to the task (the number of interactive elements in digital media art data varies with the stage of creation). Finally, the weighted features are fused with the LSTM output, which not only retains the modeling ability of LSTM for temporal dynamics, but also enhances the influence of important information through the attention mechanism and reduces noise interference.

4. Result

In terms of parameter Settings, when dividing the dataset, the proportion of the training set is set to 0.7, the adam gradient descent algorithm is adopted, the maximum number of iterations is 150, the batch size is 128, the initial learning rate is 1e-3, the learning rate scheduling method is piecewise, the descent factor is 0.1, the descent period is 1200, and the dataset is scrambled in each round of iteration.

Decision trees, random forests, CatBoost, AdaBoost and XGBoost were used as comparative experiments, and the experimental results are shown in Table 2. The comparison bar chart of the output experimental results is shown in Figure 3.

Table 2. The experimental results

Model	Accuracy	Recall	Precision	F1	AUC
Decision tree	0.709	0.709	0.71	0.71	0.709
Random Forest	0.803	0.803	0.804	0.803	0.875
CatBoost	0.786	0.786	0.786	0.786	0.876
AdaBoost	0.778	0.778	0.782	0.778	0.75
XGBoost	0.744	0.744	0.744	0.742	0.853
Our model	0.855	0.855	0.856	0.855	0.904

By all metrics, our model performs best in terms of accuracy, recall, precision, F1, and AUC. Its accuracy is 0.855, which is higher than that of Decision Tree 0.709, Random Forest 0.803, CatBoost 0.786, AdaBoost 0.778, and XGBoost 0.744, indicating the highest overall classification accuracy. Recall and accuracy of 0.855 and 0.856, respectively, are also ahead of other models, showing superior performance in identifying positive samples and predicting the proportion of actual positive samples among positive samples. The F1 value of 0.855 is also higher than other models, indicating a stronger ability to balance accuracy and recall. The AUC reached 0.904, surpassing random forest's 0.875 and CatBoost's 0.876, among others, indicating the best ability to distinguish between positive and negative samples. In contrast, other models are slightly inferior in various metrics, especially the overall performance of Decision Tree and XGBoost is significantly lower than ours.

DOI: 10.54254/2755-2721/2025.27797

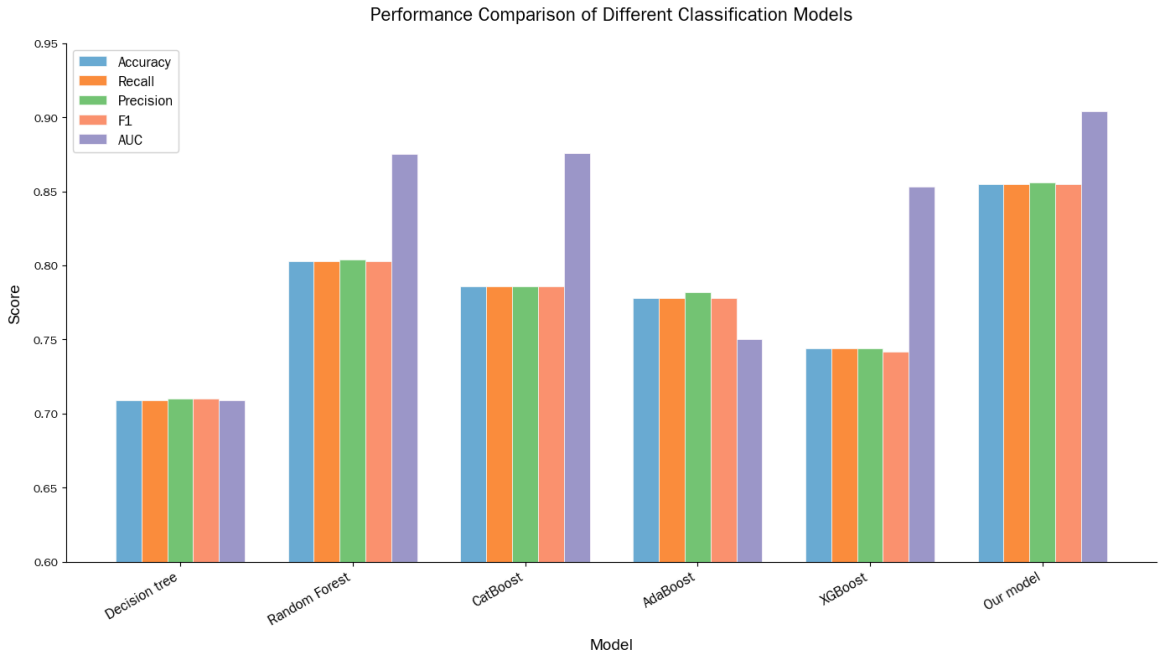

Figure 3. The comparison bar chart of the output experimental results

Output the confusion matrix of the test set of Our model, as shown in Figure 4.

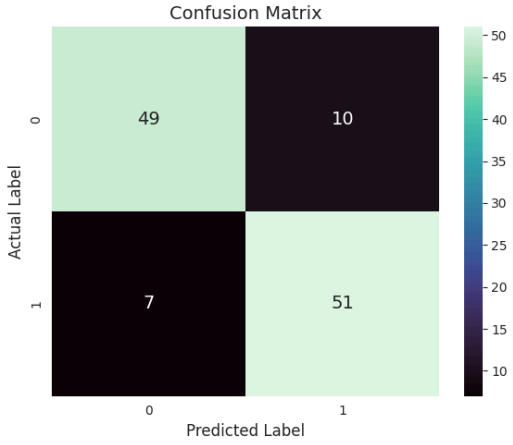

Figure 4. The confusion matrix of the test set of our model

5. Conclusion

Exploring the quantitative assessment system of the popularity of digital media art is an important research direction connecting art creation, technology application and market communication. Aiming at the limitation of existing algorithms, an optimized LSTM algorithm with multi-head attention mechanism is proposed in this paper. The results of correlation analysis show that the correlation between the number of interactive elements and heat is the most significant, and the absolute value of correlation coefficient reaches 0.556887. Variables such as creative length, number of colors, and complexity score also have some correlation with popularity, meaning that creative input, color richness, and the complexity of the work all have an impact on its popularity.

DOI: 10.54254/2755-2721/2025.27797

Through comparison experiments with decision tree, random forest, CatBoost, AdaBoost and XGBoost, it can be seen that the proposed model has advantages in all indexes: the accuracy rate reaches 0.855, which is higher than other models, and the overall classification accuracy is leading; Recall and accuracy were 0.855 and 0.856, respectively, which performed better in positive sample identification and prediction accuracy. The F1 value is 0.855, indicating a stronger ability to balance accuracy and recall. The AUC was 0.904, showing the best performance in distinguishing positive and negative samples. Other models, especially decision trees and XGBoost, lag far behind it. This research result provides more precise technical support for the quantitative evaluation of digital media art. It not only helps creators to optimize creative strategies based on data, but also provides scientific reference for the coordinated development of technology application and market communication, and promotes the standardized and efficient development of digital media art field.

References

[1] Perera, A., & Fernando, P. (2024). Cyberbullying detection system on social media using supervised machine learning. *Procedia Computer Science, 239*, 506-516.
[2] Boddu, R. S. K., et al. (2022). An analysis to understand the role of machine learning, robotics and artificial intelligence in digital marketing. *Materials Today: Proceedings, 56*, 2288-2292.
[3] Kozcuer, C., Mollen, A., & Bießmann, F. (2024). Towards transnational fairness in machine learning: A case study in disaster response systems. *Minds and Machines, 34*(2), 11.
[4] Bhaskaraputra, A., et al. (2022). Systematic literature review on solving personalization problem in digital marketing using machine learning and its impact. In *2022 International Seminar on Application for Technology of Information and Communication (iSemantic)*. IEEE.
[5] Xu, Y., & Nazir, S. (2024). Ranking the art design and applications of artificial intelligence and machine learning. *Journal of Software: Evolution and Process, 36*(2), e2486.
[6] Thylstrup, N. B. (2022). The ethics and politics of data sets in the age of machine learning: Deleting traces and encountering remains. *Media, Culture & Society, 44*(4), 655-671.
[7] Sharma, A., et al. (2022). A conceptual analysis of machine learning towards digital marketing transformation. *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*. IEEE.
[8] Omran, E., Al Tararwah, E., & Al Qundus, J. (2023). A comparative analysis of machine learning algorithms for hate speech detection in social media. *Online Journal of Communication and Media Technologies, 13*(4), e202348.
[9] Lannelongue, L., & Inouye, M. (2024). Pitfalls of machine learning models for protein–protein interaction networks. *Bioinformatics, 40*(2), btae012.
[10] Shaikh, A. A., et al. (2022). The role of machine learning and artificial intelligence for making a digital classroom and its sustainable impact on education during COVID-19. *Materials Today: Proceedings, 56*, 3211–3215.
[11] Altché, F., & de La Fortelle, A. (2017). An LSTM network for highway trajectory prediction. *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE.