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

Like Ma[1*], Xiaoli Song[1]

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

Abstract. Exploring a quantitative evaluation system for the popularity of digital media art has become a core issue that connects artistic creation, technological application and market communication. In view of the bottlenecks existing in the current algorithms, this paper proposes an LSTM algorithm optimized based on the multi-head attention mechanism. The study first conducted a correlation analysis. The results showed that the number of interactive elements had the strongest correlation with popularity, with an absolute value of the correlation coefficient reaching 0.556887. Variables such as creation time, the number of colors, and complexity scores also have a certain correlation with popularity. It can be seen that the time invested in creation, the richness of colors, and the complexity of the work will, to a certain extent, affect the popularity of the work. Taking decision tree, random forest, CatBoost, AdaBoost and XGBoost as the comparative experimental objects, in terms of various indicators, Our model performed the best in Accuracy, Recall, Precision, F1 and AUC. Its Accuracy is 0.855, which is higher than that of decision trees (0.709), random forests (0.803), CatBoost (0.786), AdaBoost (0.778), and XGBoost (0.744), with the highest overall classification accuracy. The Recall and Precision were 0.855 and 0.856 respectively, also leading other models. It performed better in identifying positive samples and the proportion of actually positive samples among those predicted to be positive. The F1 value of 0.855 is also higher than that of other models, and it has a stronger ability to balance accuracy and recall. The AUC reached 0.904, surpassing the 0.875 of random forest and the 0.876 of CatBoost, demonstrating the best ability to distinguish between positive and negative samples. In comparison, other models are slightly inferior in various indicators, especially the overall performance of the decision tree and XGBoost, which is more significantly lower than that of the Our model. This research achievement provides a more efficient method for the quantitative assessment of the popularity of digital media art. It not only offers direction for the integration of artistic creation and technological application but also provides a scientific basis for the formulation of market communication strategies.

Keywords: Machine learning, digital media artworks, popularity quantification.

1. Introduction

Digital media art, as a product of the integration of technology and aesthetics, has witnessed explosive growth along with the development of Internet, artificial intelligence and virtual reality technologies [1]. Its forms have expanded from early digital images to diverse forms such as generative art, VR interactive works, and NFT digital collectibles, and its dissemination scenarios have also extended from professional exhibitions to multi-channel ecosystems including social media, online galleries, and metaverse platforms [2]. However, the popularity of such works is interwoven and influenced by multiple factors such as creative techniques, visual features, and dissemination paths. The commercial operation of the digital art market also poses an urgent demand for objective predictive models [3]. Against this backdrop, exploring a quantitative evaluation system for the popularity of digital media art has become a key issue connecting artistic creation, technological 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 multimodal data such as creation duration, visual features, and interactive elements, and through classification models, mine the implicit correlations among variables to achieve objective predictions of the popularity of works [5]. For instance, based on the decision tree algorithm, key influencing factors such as artistic style and 16:9 aspect ratio can be identified and generated. Recurrent neural networks (RNN) can capture the temporal features during the dissemination of works and improve the accuracy of dynamic prediction. This type of technology not only provides creators with data-driven creative references, but also assists platforms in optimizing recommendation mechanisms and enhancing user stickiness [6].

In view of the bottlenecks of existing algorithms, this paper proposes an LSTM algorithm optimized by the multi-head attention mechanism. When the basic LSTM processes sequence data through the gating unit, 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 feature processing limitations of traditional models, providing a more accurate technical path for predicting the popularity of digital media art, and also offering algorithmic support for the precise connection 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.

Table 1. A partial dataset

creation _time	creation _period	artist_ex perience	techniqu es_used	file_ size	aspect _ratio	animation _frames	text_ele ments	interactive_ elements	complexit y_score	platform	popul arity
103	29	3	11	67.1 5	1.33	212	5	12	1	Gallery Exhibition	0
52	14	7	12	63.7 9	1.5	2	6	7	7	NFT Marketpla ce	1
93	8	4	6	22.2 2	1.5	178	9	21	1	Social Media	1
107	22	3	1	20.8	1.33	163	6	16	6	Art Platform	0

Correlation analysis was conducted on each variable, and a correlation heat map was drawn, as shown in Figure 1. From the perspective of the correlation coefficient, the number of interactive elements has the strongest correlation with popularity, with the absolute value of the correlation coefficient reaching 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 the complexity of the work will to some extent affect the popularity of the work.
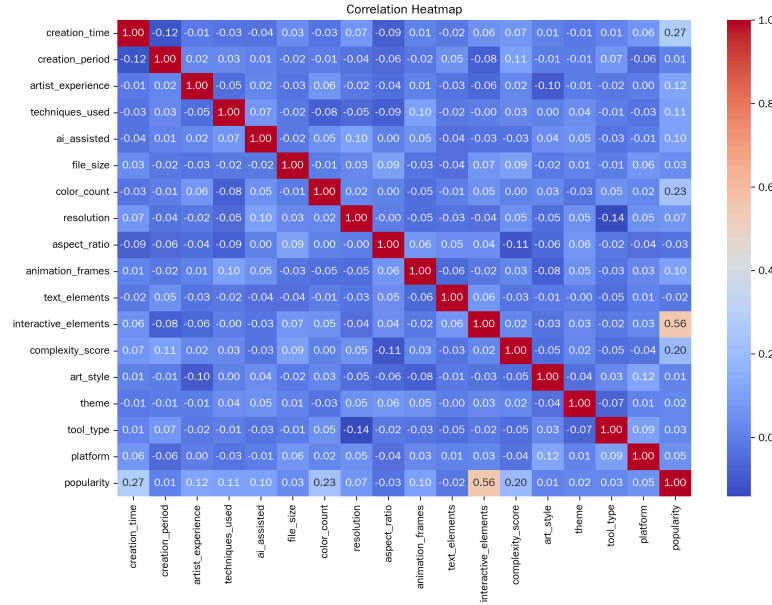

Figure 1. Correlation heat map

3. Method

3.1. Multi-head attention mechanism

The multi-head attention mechanism is a core component of the Transformer model. Its core idea is to capture the correlation information of different dimensions in the input data by executing multiple attention heads in parallel, thereby enhancing the model's modeling ability for complex sequences [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 to avoid the limitations of information interaction in a single high-dimensional space; Secondly, each attention head independently calculates the attention weight - by scaling the dot product to obtain the similarity score, and then normalizing it through the Softmax function to obtain the attention weight of each element to other elements. Finally, the weighted sum of the values with the weights is used to generate the output of this head. Finally, the outputs of all attention heads are concatenated, and multi-dimensional features are fused through a single linear transformation to form the final multi-head attention result. The advantage of this design lies in that different attention heads can focus on capturing different types of associations, enabling the model to simultaneously learn the rich contextual information in the input data and significantly enhance 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 of vanishing or bursting gradients when traditional RNN processes long sequences through gating mechanisms, thereby effectively capturing long-distance dependencies in time series data [9]. Its structure contains a cellular state that runs through the entire sequence, similar to an information conveyor belt, which can stably convey key information. At the same time, the flow of information 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 cell state. The output gate, based on the current cell state and input, controls the proportion of output information through the sigmoid function and generates the hidden state at the current moment after processing by the tanh function [10]. The network structure of LSTM is shown in Figure 2 [11].
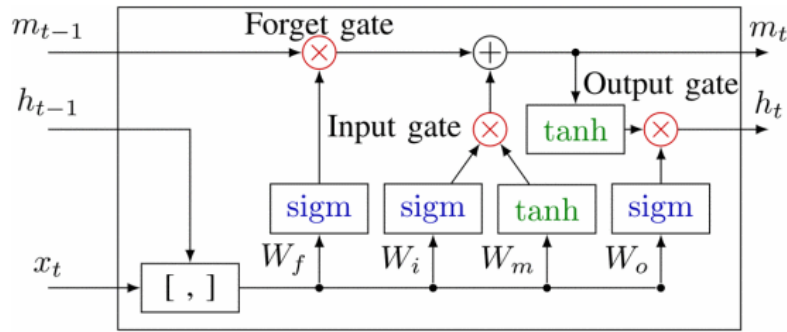

Figure 2. The network structure of LSTM

3.3. Multi-head attention mechanism-LSTM

The core of the multi-head attention mechanism in optimizing LSTM is to make up for the deficiency of LSTM in paying insufficient attention to key information in long sequences through the feature selection ability of the attention mechanism. The principle is as follows: Firstly, the gating mechanism of LSTM is utilized to process time series data, capture the long-term and short-term dependencies of the sequence, and generate a hidden state sequence containing time series features; Subsequently, these hidden states are taken as the input of the multi-head attention

mechanism. Through multiple parallel attention heads, the attention weights of the features at each time step in different subspaces are calculated respectively to accurately locate the moments or features that are more critical to the task (the number of interactive elements in digital media art data changes with the creation stage). Finally, the weighted features are fused with the LSTM output, which not only retains the LSTM's modeling ability 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

Judging from various indicators, Our model performs the best in 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. The Recall and Precision were 0.855 and 0.856 respectively, also leading other models, demonstrating superior performance in identifying positive samples and the proportion of actually positive samples among those predicted to be positive. The F1 value of 0.855 is also higher than that of 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, etc., indicating the optimal ability to distinguish between positive and negative samples. In contrast, other models are slightly inferior in various indicators, especially the overall performance of the decision tree and XGBoost is more significantly lower than that of the Our model.
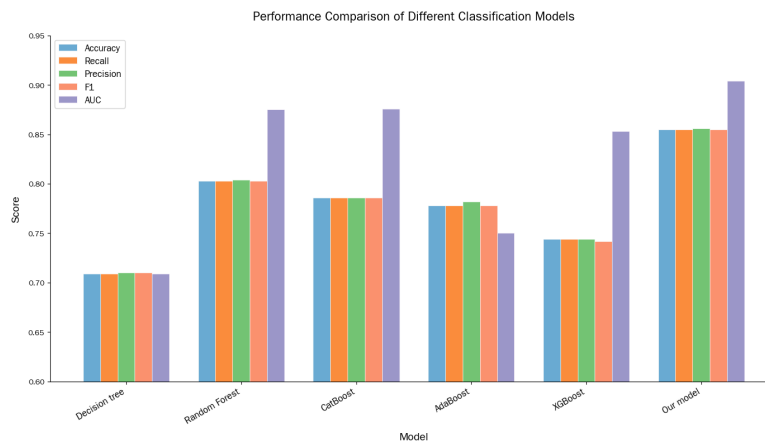
Figure 3. The comparison bar chart of the output experimental results

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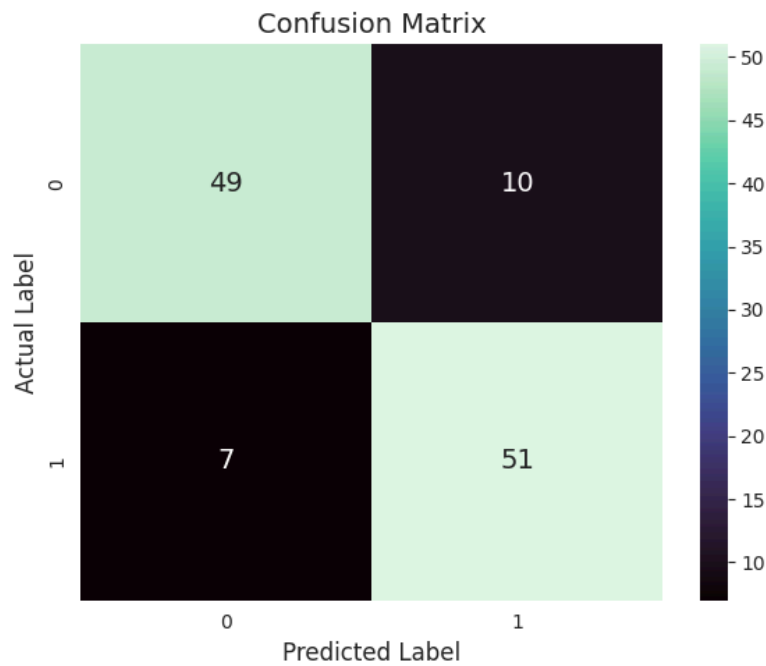

Figure 4. The confusion matrix of the test set of Our model

5. Conclusion

Exploring a quantitative evaluation system for the popularity of digital media art is an important research direction that connects artistic creation, technical application and market communication. In view of the limitations of the existing algorithms, this paper proposes an optimized LSTM algorithm integrating the multi-head attention mechanism. The results of the correlation analysis show that the correlation between the number of interactive elements and popularity is the most significant, and the absolute value of the correlation coefficient reaches 0.556887. Variables such as the duration of creation, the number of colors, and the complexity score also have a certain correlation with popularity, meaning that creative investment, color richness, and the complexity of the work will all have an impact on its popularity.

Through comparative experiments with decision trees, random forests, CatBoost, AdaBoost and XGBoost, it can be known that the proposed model has advantages in all indicators: the Accuracy reaches 0.855, which is higher than that of other models, and the overall classification accuracy is leading; The Recall and Precision were 0.855 and 0.856 respectively, and it performed better in positive sample recognition and prediction accuracy. The F1 value is 0.855, indicating a stronger ability to balance precision and recall. The AUC is 0.904, which shows the best performance in distinguishing positive and negative samples. Other models, especially decision trees and XGBoost, lag significantly behind it. This research achievement provides more precise technical support for the quantitative assessment of digital media art. It not only helps creators optimize their creative strategies based on data but also offers a scientific reference for the coordinated development of technology application and market communication, promoting the standardized and efficient development of the digital media art field.

References

[1] Perera, Andrea, and Pumudu Fernando. "Cyberbullying detection system on social media using supervised machine learning." Procedia Computer Science 239 (2024): 506-516.

[2] Boddu, Raja Sarath Kumar, et al. "An analysis to understand the role of machine learning, robotics and artificial intelligence in digital marketing." Materials Today: Proceedings 56 (2022): 2288-2292.

[3] Kozcuer, Cem, Anne Mollen, and Felix Bießmann. "Towards transnational fairness in machine learning: a case study in disaster response systems." Minds and Machines 34.2 (2024): 11.

[4] Bhaskaraputra, Aryo, et al. "Systematic literature review on solving personalization problem in digital marketing using machine learning and its impact." 2022 International Seminar on Application for Technology of Information and Communication (iSemantic). IEEE, 2022.

[5] Xu, Yandong, and Shah Nazir. "Ranking the art design and applications of artificial intelligence and machine learning." Journal of Software: Evolution and Process 36.2 (2024): e2486.

[6] Thylstrup, Nanna Bonde. "The ethics and politics of data sets in the age of machine learning: Deleting traces and encountering remains." Media, Culture & Society 44.4 (2022): 655-671.

[7] Sharma, Alpana, et al. "A conceptual analysis of machine learning towards digital marketing transformation." 2022 5th international conference on contemporary computing and informatics (IC3I). IEEE, 2022.

[8] Omran, Esraa, Estabraq Al Tararwah, and Jamal Al Qundus. "A comparative analysis of machine learning algorithms for hate speech detection in social media." Online Journal of Communication and Media Technologies 13.4 (2023): e202348.

[9] Lannelongue, Loïc, and Michael Inouye. "Pitfalls of machine learning models for protein–protein interaction networks." Bioinformatics 40.2 (2024): btae012.

[10] Shaikh, Asmat Ara, et al. "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 (2022): 3211-3215.

[11] Altché, Florent, and Arnaud de La Fortelle. "An LSTM network for highway trajectory prediction." 2017 IEEE 20th international conference on intelligent transportation systems (ITSC). IEEE, 2017.