International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 6477-6489

https://doi.org/10.70517/ijhsa463557

Multidimensional Computational Analysis of Structured Data and Assessment of Audience Emotional Response for Classical Dance Theater Works

Miao Jin[1,*]
[1] Dance Academy of Nanjing University of the Arts, Nanjing, Jiangsu, 210000, China
Corresponding authors: (e-mail: nuajinmiao1995@163.com).

Abstract Transforming dance drama works into structured data through multi-dimensional computational analysis and combining it with audience emotional response assessment can evaluate dance drama works from a quantitative perspective, provide data support for the creation, performance and evaluation of dance dramas, and promote the innovative development of the dance drama art and the enhancement of audience experience. This study constructs a multi-dimensional computational analysis and audience emotional response assessment system for structured data of classical dance drama works. The study adopts a multidimensional structured data modeling framework to perform computational analysis of dance drama works from entity dimension, attribute dimension and history dimension, and extracts the core features of dance drama works by combining the multidimensional information fusion mechanism and the entity level encoder; at the same time, it designs an audience emotion assessment model based on ATAAE-BERT-BiLSTM to classify and identify the audience emotional response of dance drama works. The experimental results show that the proposed model outperforms the best-performing PIVOT in the baseline model by 2.88 and 0.67 in BLEU and NIST metrics, respectively; in the multidimensional correlation validation, the mean values of the structured data of the storyline, characterization, dance language, music performance, and stage scene reach more than 3000; in the emotion assessment model, the complete segmentation based on ATAAE-BERT-BiLSTM BERT-BiLSTM model achieved 83.25% accuracy and 83.88% F1 value in 4-segment segmentation mode, which were 5.14% and 6.72% higher than the BERT-LSTM model, respectively. The study provides a data-based method for the scientific evaluation of dance drama works and effectively supports the creation and dissemination of dance drama art.

Index Terms Structured data, Multidimensional computational analysis, Audience sentiment assessment, ATAAE-BERT-BiLSTM

I. Introduction

Dance drama, as a type of stage art, brings together artistic elements such as physical narrative, music rendering, visual display, etc., such as “Only This Green”, “Swan Lake”, “Confucius”, “The White-haired Girl” and other classic works [1]-[3]. In these classic dance drama works, both reflecting historical stories, showing ethnic customs, and exploring the philosophy of life, these works have left a deep impression on the audience both in terms of choreography, choreographic techniques, and performances of the actors. Through the body language of the performers and the image of the art work, the artistic emotions of the creators are conveyed, and the audience is presented with a concrete and vivid artistic image, which originates from life and is higher than the emotions in life [4]-[7].

Traditional artistic analysis research is limited to shallow analysis, and it is difficult to cope with the quantitative analysis of complex dance drama works. With the development of digital technology, data-driven analysis of artworks and their multimodal analysis have become an important research direction in artistic computing [8]. The research progress of dynamic capture (facial expression and body movement), three-dimensional modeling, emotional computing and other technologies provide technical support for collecting data such as emotional semantics, visual symbols, audience emotional feedback, and constructing structured data of dance dramas to analyze the creation law of the work, and the assessment of the audience's emotional response [9]-[12]. In the traditional method of audience emotional response assessment through the form of questionnaires and interviews often presents shortcomings such as subjective goodwill bias, response delay, memory bias or forgetfulness, which makes the assessment results have a large error. And through facial emotion recognition, machine learning and other technologies, the construction of facial expression, electroencephalogram, physiological signal changes, limb following, eye changes and other multi-dimensional data of the emotional response framework, can be all-round analysis of audience emotional response [13]-[15].

Dance drama, as a comprehensive art form, contains multi-dimensional contents such as storyline, character image, dance language, music performance and stage scene, etc. A systematic and structured analysis of it is of great significance to the development of dance drama art. Traditional dance drama evaluation mostly relies on experts' experience and subjective judgment, which makes it difficult to form a standardized and quantifiable evaluation system, and lacks systematic research on audience's emotional response. With the development of computer technology and artificial intelligence, it has become possible to conduct multi-dimensional quantitative analysis of dance drama works through structured data. This study focuses on the structured data modeling and audience emotional response assessment of classic dance drama works, and constructs a complete analysis framework. First, a multidimensional structured data modeling framework is established to analyze the dance drama works from the entity dimension, attribute dimension, and history dimension in an all-round way. The entity dimension models the relationship between structured data of the same entity through the self-attention mechanism, while the attribute dimension compares the relationship between data of the same attribute, and the historical dimension considers the temporal relationship of structured data. Through the multidimensional information fusion mechanism, the overall representation of the data is obtained by adaptively fusing the data representations according to the importance of the different dimensional representations. In the entity-level encoder part, the average pooling operation and content selection gating mechanism are used to compute the importance of the representation, and the hierarchical attention mechanism is used to enhance the ability of the decoder to select important data. Meanwhile, in order to assess the audience's emotional response to the dance theater piece, the study constructed the ATAAE-BERT-BiLSTM model, which combines the BERT word embedding with the BiLSTM sentence representation layer to capture the key emotional features through the attention mechanism. The model is trained with cross-entropy loss function and Adam optimization algorithm to effectively improve the accuracy of sentiment analysis in dance drama. The experimental part firstly verifies the effectiveness of the structured data model by comparing with several baseline models, and conducts sample learning comparison and multivariate phrase performance evaluation to deeply analyze the generalization ability of the model in different domains. The reasonableness of the model design is further verified by loss function change records and comparisons. In the audience sentiment assessment experiments, the performance of different pre-trained models and network structures are compared, and the advantages of the proposed ATAAE-BERT-BiLSTM model in classifying the audience sentiment of dance theater are verified. Combining the art of dance drama and data science, this study innovatively proposes a multi-dimensional computational analysis and audience emotional response assessment method for structured data of classical dance drama works, which provides a new scientific tool for the creation, evaluation and dissemination of dance drama art.

II. Multi-dimensional computation and emotional assessment of classic dance theater works

II. A. Multidimensional Structured Data Modeling Framework

II. A. 1) Structured data encoder

(1) Entity dimension

Since the structured data of the same entity is disordered, the relationship between the structured data of the same entity is modeled through the self-attention mechanism. The $u_{i,j}^{ent}$ is used to represent the structured data of the i th entity and the j th attribute in the entity dimension. In this chapter, the representation $c_{i,j}^{ent}$ of the interrelationships between data that imply the same entity is obtained through the self-attention mechanism. The specific computational procedure is shown below:

$$c_{i,j}^{ent} = \sum_{j',j'\neq j} \alpha_{i,j,j'}^{ent} u_{i,j'} \tag{1}$$

where $\alpha_{i,j,j'}^{emt} \propto \exp(u_{i,j}^{T} W_o u_{i,j'})$ is the attentional weight after a normalization operation between the data for the same entity and takes a value ranging from 0 to 1, W_o is a trainable parameter.

Then, the original representation $u_{i,j}$ of the structured data and the representation $c_{i,j}^{ent}$ that implies the interrelationships among the data of the same entity are combined to obtain the final representation $u_{i,j}^{ent}$ of the structured data in the entity dimension, which is computed as shown below:

$$u_{i,j}^{ent} = \tanh(W_f (u_{i,j} \oplus c_{i,j}^{ent})) \tag{2}$$

where W_f is a trainable parameter and $\oplus$ represents vector splicing.

(2) Attribute dimension

In this paper, we propose to model the structural information of attribute dimensions of structured data, using a mechanism similar to self-attention to compare the relationship between data with the same attribute. The $u_{i,j}^{attr}$ is used to represent the structured data of the i th entity and the j th attribute on the attribute dimension. The representation $c_{i,j}^{attr}$ that implies the interrelationships between the data of the same attribute is first computed on the attribute dimension as shown below:

$$c_{i,j}^{attr} = \sum_{i',i'\neq i} \alpha_{j,i,i'}^{attr} u_{i',j} \tag{3}$$

where $\alpha_{j,i,i'}^{attr}$ is the attention weight after a normalization operation on the data from different entities i' but the same attribute j, with a value ranging between 0 and 1. The final representation of the attribute dimension $u_{i,j}^{aut}$ is computed as shown below:

$$u_{i,j}^{attr} = \tanh(W_{fc}(u_{i,j} \oplus c_{i,j}^{attr})) \tag{4}$$

where W_{fc} is a trainable parameter and $\oplus$ represents vector splicing.

(3) Historical dimension

Using the constructed timeline from which structured data with the same entities and attributes are identified, all data on the timeline at time k earlier than the current structured data $u_{i,j}$ can be considered as historical information. Due to the large amount of historical data, this chapter sets a time window and models the data within the time window as historical information $hist(u_{i,j})$. This chapter models structured information in the historical data dimension through a self-attention mechanism. However, unlike the entity and attribute dimensions where there is no order between structured data, historical information on the historical data dimension is ordered. Therefore, a trainable position vector $emb_{pos}(k')$ is introduced to model this temporal ordering relationship and added to the structured data representation to obtain a new structured data representation $rp_{k'}$. This representation represents data that has the same entities and attributes as the structured data $u_{i,j}$, but occurs at the moment of time k' (earlier than time k within the time window). Use $u_{i,j}^{time}$ to represent the historical representation of the structured data for the i th entity and the j th attribute, and then model the relationship between the data in the time window in terms of the historical data dimensions through the self-attention mechanism. The specific computation is shown below:

$$\alpha_{k,k'}^{time} \propto \exp(MLP(rp_k, rp_{k'})) \tag{5}$$

$$c_{i,j}^{time} = \sum_{k'<k} \alpha_{k,k'}^{time} rp_{k'} \tag{6}$$

$$u_{i,j}^{time} = \tanh(W_{ft}(u_{i,j} \oplus c_{i,j}^{time})) \tag{7}$$

The attentional weight $\alpha_{k,k'}^{time}$ is normalized between the data in the time window and takes values between 0 and 1. W_{ft} is a trainable parameter and $\oplus$ represents vector splicing.

II. A. 2) Multidimensional information fusion mechanisms

After obtaining the representation of structured data in each of the three dimensions, it is necessary to determine which dimension of the structured data is more reflective of the information embedded behind the structured data. If a piece of structured data stands out more compared to the structured data of other entities with the same attribute, then this piece of structured data is more important to represent on the attribute dimension. In addition, if certain types of structured data for an entity occur frequently, then the representation on the entity dimension captures the relationship between the different attribute data for the entity.

Therefore, a multidimensional information fusion mechanism is proposed to adaptively fuse data representations from three dimensions according to the importance of different dimensional representations of structured data, and finally obtain a holistic representation of the data. Specifically, in this paper, the three dimensional representations $u_{i,j}^{ent}, u_{i,j}^{attr}$ and $u_{i,j}^{time}$ are vectorially spliced, and then a multilayer perceptron is employed to obtain an initial overall representation of the data, $u_{i,j}^{gen}$, which is considered as the baseline representation of the structured data for that article. Then, the data representation of each dimension is compared with this baseline representation to determine which dimensions are more important in terms of information, and to calculate the weights of the representations of each dimension. As an example, the specific formula for calculating the weight of the attribute dimension representation is shown below:

$$\delta_{i,j}^{attr} \propto \exp(MLP(u_{i,j}^{attr}, u_{i,j}^{gen})) \tag{8}$$

The weights of entity and historical data dimension representations can be obtained in a similar way. Finally, the computed weights of the different dimensional representations are used to fuse the multiple dimensional representations to obtain the overall representation of the final structured data $\tilde{u}_{i,j}$, as shown below:

$$\tilde{u}_{i,j} = \delta_{i,j}^{ent} u_{i,j}^{ent} + \delta_{i,j}^{attr} u_{i,j}^{attr} + \delta_{i,j}^{time} u_{i,j}^{time} \tag{9}$$

II. A. 3) Entity-level encoders

Each row of data in structured data represents various aspects of an entity's performance, and this chapter combines the data of the same entity through the operation of average pooling to obtain an overall representation that can reflect the overall performance of the entity, which is calculated as shown below:

$$ent_i = MeanPooling(\tilde{u}_{i,1}, \tilde{u}_{i,2}, ..., \tilde{u}_{i,A}) \tag{10}$$

where A denotes the number of attributes. Then the importance g_i of each entity representation is calculated using the content selection gating mechanism, and using the calculated importance weights, the entity representation ent_i is filtered to obtain a new entity representation $e\tilde{n}t_i = g_i \odot ent_i$. The representation of each piece of data is obtained through the multidimensional information fusion mechanism, and the representation of entities is obtained through the entity level encoder. In order to improve the decoder's ability to select important data, the hierarchical attention mechanism is used to help the model locate key information more accurately. The decoder selects the important entities first and then the important data when generating the classical dance theater works. The entity-level attention weights are calculated as shown below:

$$\beta_{t,i} \propto \exp(score(d_t, ent_i)) \tag{11}$$

where d_t represents the hidden state of the decoder at step t and ent_i represents the representation of each entity. The entity-level attentional weights $\beta_{t,i}$ are normalized between all entity representations and take values between 0 and 1. The data level weights are then computed as shown below:

$$\gamma_{t,i,j} \propto \exp(score(d_t, \tilde{u}_{i,j})) \tag{12}$$

The $\gamma_{t,i,j}$ is normalized between data of the same entity, taking values between 0 and 1. Finally this chapter fuses the entity-level attention weights $\beta_{t,i}$ and the data-level attention weights $\gamma_{t,i,j}$ to obtain the new attention weights $\tilde{\alpha}_{t,i,j}$ computed based on the hierarchical attention mechanism, which is shown below:

$$\tilde{\alpha}_{t,i,j} = \beta_{t,i}\gamma_{t,i,j} \tag{13}$$

In the training stage, given a batch of structured data $\{S\}_G$ and a reference to a classic dance theater work $\{Y\}_G$, this chapter uses the negative log-likelihood function as the loss function of the model, and the specific computational process is shown below:

$$L = -\frac{1}{G}\sum_{g=1}^{G}\sum_{t=1}^{T_g}\log P(y_{t,g} \mid y_{<t,g}, S_g) \tag{14}$$

Training the model by minimizing L is equivalent to maximizing the probability that the model generates a reference classical dance theater production based on the input structured data, G is the number of training samples in each batch of data, and T_g denotes the length of the first g reference classical dance theater production.

II. B.ATAAE-BERT-BiLSTM model based on audience emotion

II. B. 1) Word Embedding Layer

First of all, it is necessary to convert classical dance theater works containing n words evaluating classical dance theater works into word vectors $\{w_1, w_2, \cdots\cdots, w_n\}$ by means of the pre-trained BERT word vector model.

II. B. 2) BiLSTM sentence representation layer

BiLSTM it is better able to capture bidirectional semantic dependencies and breaks the bottleneck of the limited range of contextual information accessed by traditional recurrent neural networks to understand sentence meanings more deeply [16]. In the BiLSTM network model, two gates of LSTM with opposite directions exist at each moment, where $\vec{h}_t$ denotes the forward output of LSTM at moment t, which is jointly determined by the current input w_t and the forward output $\vec{h}_{t-1}$ of the previous moment, and $\overleftarrow{h}_t$ denotes the reverse output of LSTM at moment t, which is jointly determined by the current input w_t and the previous moment $\overleftarrow{h}_{t-1}$ [17]. h_t denotes the final output of the BiLSTM at moment t, which is jointly determined by moments $\vec{h}_t$ and $\overleftarrow{h}_t$ of moment t and their respective weight matrices W_t and V_t and bias b_t. The final output h_t of the BiLSTM at moment t is computed as shown in Eq. (15):

$$\begin{aligned}\vec{h}_t &= LSTM(w_t, \vec{h}_{t-1}) \\ \overleftarrow{h}_t &= LSTM(w_t, \overleftarrow{h}_{t-1}) \\ h_t &= W_t \vec{h}_t + V_t \overleftarrow{h}_t + b_t\end{aligned} \tag{15}$$

II. B. 3) Attention layer

First, the Attention weight α of the Attention mechanism needs to be computed, which is calculated by Eqs. (16) and (17):

$$g_{a_i} = \tanh\left[(a_i^T * H) * W_{a_i}\right] \tag{16}$$

$$\alpha = softmax(g_{a_i} * W_{a_i}^o) \tag{17}$$

where d is denoted as the spatial dimension trained by the word vector model, $a_i \in \mathbb{R}^d$, $H \in \mathbb{R}^{d \times n}$, $g_{ai} \in \mathbb{R}^{1 \times n}$, $\alpha \in \mathbb{R}^{1 \times n}$, W_{ai} and W_{ai}^o are the training parameters of the model, $W_{ai} \in \mathbb{R}^{n \times n}$, $W_{ai}^o \in \mathbb{R}^{n \times n}$. Then the Attention weighted summation of the Attention weights is obtained by Eq. (18) for the course evaluation of the classical dance theater work representation r:

$$r = H * \alpha^T \tag{18}$$

where $r \in \mathbb{R}^d$. The final sentence representation after the Attention layer is transformed into:

$$h^* = \tanh(W_p r + W_x h_N) \tag{19}$$

In Eq. (19), $h^* \in \mathbb{R}^d$, W_p and W_x are the parameters trained in the model.

II. B. 4) Output layer

The final model needs to map can h^* into the six classifications by softmax fully connected layer to get the probability distribution y of these six emotional tendency classifications, which is calculated as:

$$y = softmax(W_s h^* + b_s) \tag{20}$$

In Eq. (20), W_s and b_s are denoted as the parameters to be trained in the softmax function.

II. C. Model training

II. C. 1) Loss function

During experimentation, we need a way to be able to measure the error for subsequent adjustment of the model and evaluation of the experimental results, and when we use a mathematical formula to measure the algorithmic error, the method we use is called the loss function. Usually, when the experimental error is low, the value of the loss function output is also low, which indicates that the algorithm is highly accurate. Similarly: when the value of the loss function is high, it proves that the model is less accurate and the results are unsatisfactory. However, it is worth mentioning that overfitting may occur when the loss value is too low, which is due to the difference in data between the training set and the test set. So it is essential to keep the loss value of the model within a reasonable range.

Therefore, choosing a loss function with a high degree of fit to the task to improve the accuracy of the final experimental results is also a common measure in research. The loss function used in the experiments of this thesis is the cross-entropy error between the real emotion polarity value and the model predicted value. The cross-entropy error loss function overcomes the problem that the variance cost function is too slow to update the weights, which is affected by the error, and when the error is large, the weights are updated quickly. Similarly, when the error is smaller, the weights are updated slower. The cross-entropy loss function is also a class of commonly used methods in the task of sentiment analysis of classic dance theater works, and its specific formula is as follows:

$$L = -\frac{1}{N}\sum_x \left[y \ln a + (1-y)\ln(1-a)\right] \tag{21}$$

where L is the loss value, x is the sample, and N is the total number of samples. This loss function is chosen because it has the following advantages: firstly, during the construction process, the dataset usually does not change during the training process. Secondly, this loss function achieves local minimum convergence and higher accuracy during each iteration of model training.

II. C. 2) Optimization algorithms

In the experimental process, when we use a more complex model, or the data set is large in size, etc., it will lead to the convergence speed of the algorithm is too low, and the time of a single experiment is too long, which will lead to a decline in the efficiency of the experiment. In order to accelerate the convergence speed of the model, researchers will use a variety of optimization algorithms, which can play a role in accelerating the role of deep neural network training iterations, the process of the algorithm of the initial parameter is constantly optimized, which in turn makes the model convergence speed faster. Currently, the commonly used optimization algorithms in the industry include Adam, SGD, Adagrad, etc. In this paper, the model selected is the Adam optimization algorithm, which is able to iteratively update the parameter weights in the neural network, and dynamically adjust the adaptive learning rate through the continuous training data, unlike traditional algorithms in which the learning rate is fixed and unchanging.

III. Multidimensional computational analysis and sentiment assessment of structured data

III. A. Multidimensional computational analysis of structured data

III. A. 1) Structured Data Model Validation Analysis

(1) Experimental data and preprocessing

Experimental validation and analysis were performed using the dance drama work dataset A. The training set (80%), test set (10%) and validation set (10%) were divided. It contains structured data with its corresponding descriptive classical dance drama works, and the key information of descriptive classical dance drama works is taken from the structured data, and the amount of data in the structured data is usually larger than that of the generated descriptive classical dance drama works, which indicates that the model needs to understand and filter the structured data in the training process, rather than directly perform contextual vocabulary prediction and derivation.

(2) Experimental Setting

In order to avoid errors on floating-point comparisons, in this paper, after several adjustments, the ε in the original loss function $Loss1$ is set to 10-6 and the copy loss function $Loss_2$ The hyperparameter λ is set to 0.68. The optimizer uses Adam and sets the learning rate to 0.0003 and the mini-batch to 40. The size of the word embedding and the field are both set to 768, and 5 is used as the size of the positional embedding. For an attribute with multiple words, we find the arithmetic mean of the embeddings of all the words it contains as the embedding size of the attribute.

(3) Evaluation Criteria

In addition, the change of BLEU (Bilingual Evaluation of Understudies) value with the same “number of training instances” is utilized to judge whether the model is able to perform well with fewer training samples. In order to evaluate the effectiveness of increasing the penalty term of the loss function and the comparison of different penalties in the generative adjustment strategy proposed in this paper, we conduct experiments and observations on the loss value and BLEU under multiple rounds of iterations, and the BLEU calculation formula is as follows:

$$BLEU = \exp\left(\sum_{i=1}^{N} w_n \log P_n\right) BP \tag{22}$$

where BP is the number of penalties, w_n is the weight, and P_n is the precision.

NIST, on the other hand, increases the weight of those key words that occur infrequently by accumulating the amount of information in the n-grams and then dividing it by the number of n-gram segments in the entire model output of the classic dance theater work.

(4) Baseline model

(a) Seq2Seq-based RNN: A single RNN neural network model constructed in the Seq2Seq framework, which belongs to the baseline model for solving the problem of generating structured data to classical dance theater works using sequence-to-sequence framework.

(b) Transformer: a model consisting of only the attention mechanism, built on top of the Seq2Seq framework, often used in major NLP tasks.

(c) PretrainedMT: Skorokhodov et al. compressed the amount of training data and constructed a sequence-to-sequence framework by combining pre-trained models through a semi-supervised approach, thus realizing a deep model with low-resource training.

(d) SemiMT: a semi-supervised approach to corpus reconstruction using a self-encoder, and explored the effects of OOV word ratio and the size of the number of samples on model performance.

(e) PIVOT: a two-stage model based on BiLSTM and Transformer to implement the structured data to classical dance theater works generation task view about small sample learning.

(5) Result Analysis

Comparing the model of this paper with the baseline model, the results are shown in Figure 1. For the evaluation of the model generated data and the standard reference data, the present model improves the BLEU and NIST metrics by 2.88 and 0.67 over the PIVOT model which is the best performer in the baseline model. This study shows that the method of the present paper has a better performance on the structured data of the dance drama.

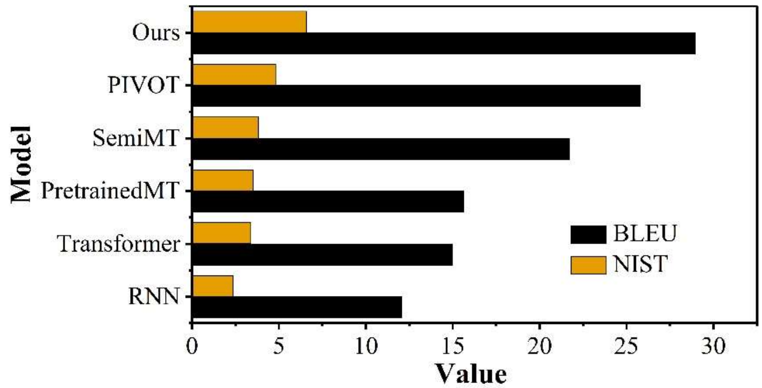


Figure 1: The model in this paper is compared with the baseline model

In order to verify the generalization of the model on different domains, this paper on three different domains were trained with 100, 200, 300, 400, 500 instance samples to observe the above results, sample learning and comparison as shown in Fig. 2, it can be found that, in the case of less than 500 training samples, in the case of the three domains of music, books, as well as dance and drama dataset, are able to show a BLEU>25 at a stable level, verifying that the model is not only applicable to specific domains, but is able to have generalization ability after learning from fewer samples in multiple domains. Further exploration reveals that the BLEU values of the Books and Dance Theater domains are, on average, more than 8.5 higher than that of the Character domain, which is due to the fact that the tabular attributes of these two domains are easier to understand than those of the Character domain, and thus the tabular information can be learned to be more fully realized when generating less-sample learning.

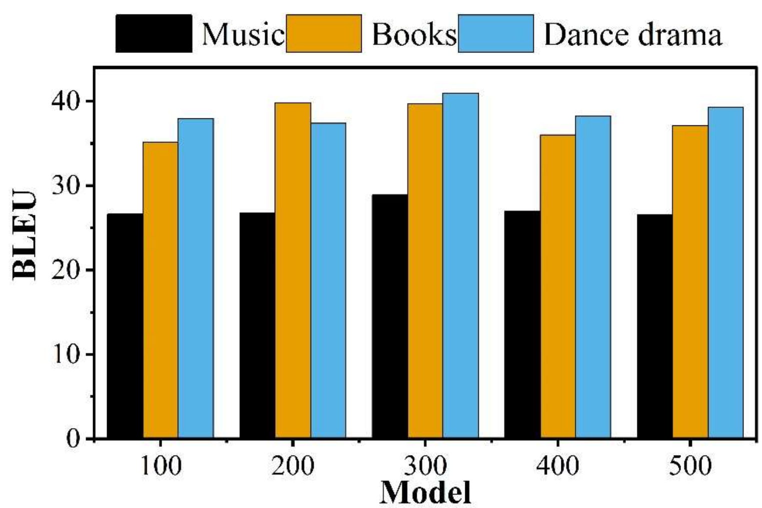
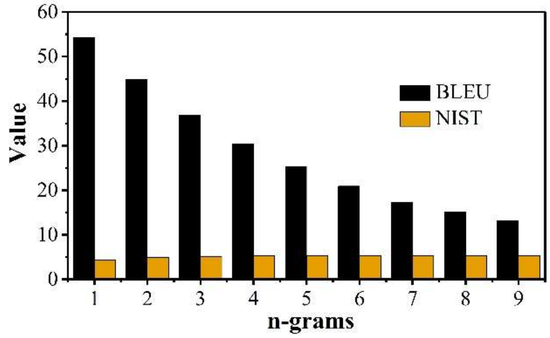
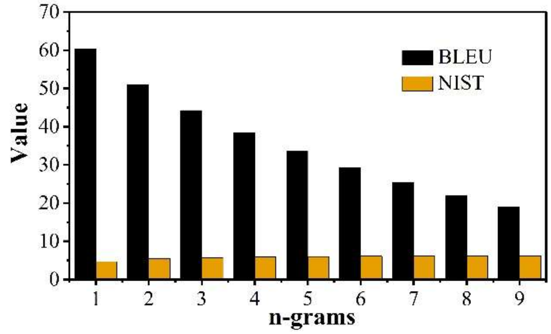


Figure 2: Sample learning and comparison

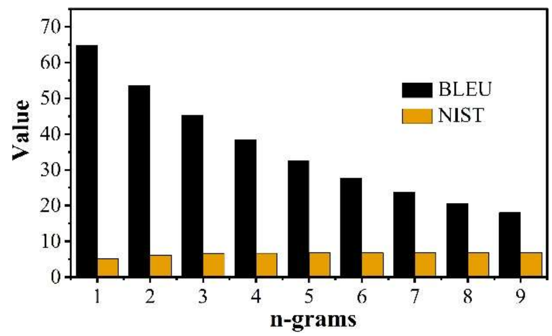
With only 500 instance samples in the dataset of three different domains (including the music domain, the book domain, and the dance theater domain) retained, this paper obtains the BLEU and NIST metrics of our proposed model from 1-grams to 9-grams, and then evaluates the performance effect of the model in the multivariate phrase, and the performance evaluation is shown in Fig. 3, in which (a) to (c) are respectively music domain, book domain, and dance theater domain. From the table, it can be seen that in n-grams (n-grams), when n is larger the BLEU is smaller, and n shows a negative correlation trend with BLEU. And when n is larger then NIST is larger, n shows a positive correlation trend with BLEU. Usually, n takes the middle number between 1 and 9 is more stable. Combined with the referenced related literature, the experiment here takes BLEU-4 (n=4) and NIST-5 (n=5) for observation respectively. We can clearly find that the BLEU-4 for all three domains is greater than 30, which is already much greater than the performance of the baseline model, while the NIST is greater than 5, which is also basically able to reach a close level with the baseline model trained on the original dataset.



(a)The field of music



(b)The field of books



(c)Dance drama field

Figure 3: Performance evaluation

In order to further validate the effectiveness of the structured data model in this paper, this paper records and compares the changes for the loss function. The role of the loss function in the model consists of two stages: in the forward propagation stage, it is calculated layer by layer from end to end, and after the last layer, the loss function is obtained by comparing it with the objective function, and the error update value is calculated. In the back-propagation phase, it is calculated layer by layer from end to end, and after the first layer, all the weights are updated.

In this paper, an auxiliary loss function is added in addition to the original loss, i.e., the copy loss function $Loss2$, and the loss variation of the model under total loss (i.e., joint loss) and copy loss is obtained, as shown in Fig. 4. It can be found that when the model makes an error in selecting the predicted word between the pre-trained model decoder and the decoder, $Loss2$ as a penalty term penalizes the training of the model accordingly, thus reducing this error bias.

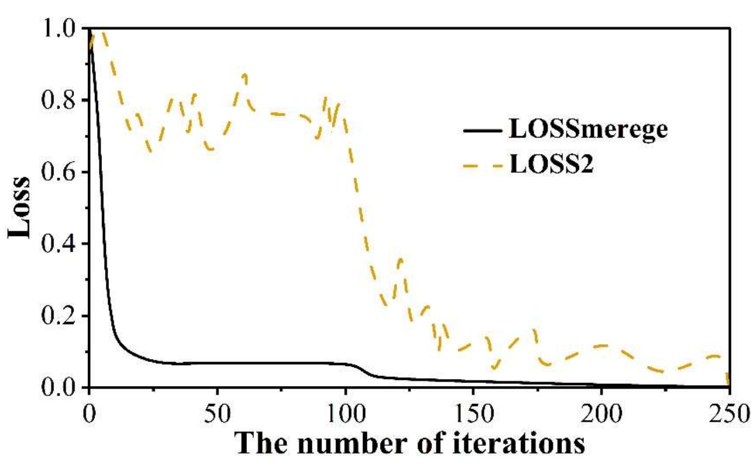


Figure 4: The variation of joint loss and copy loss

In this paper, the auxiliary loss is used to train the model in two cases, copy loss and coverage loss, and the corresponding BLEU values are obtained, and the performance comparison of copy loss or coverage loss is shown in Fig. 5. It can be seen that if the coverage loss is used as an auxiliary loss $Loss2$, the required distribution of attention for the model can be obtained, and thus the part of the model that has been fed the same content multiple times can be penalized to some extent during the training process to improve the performance of the model. The auxiliary loss function is the copy loss.

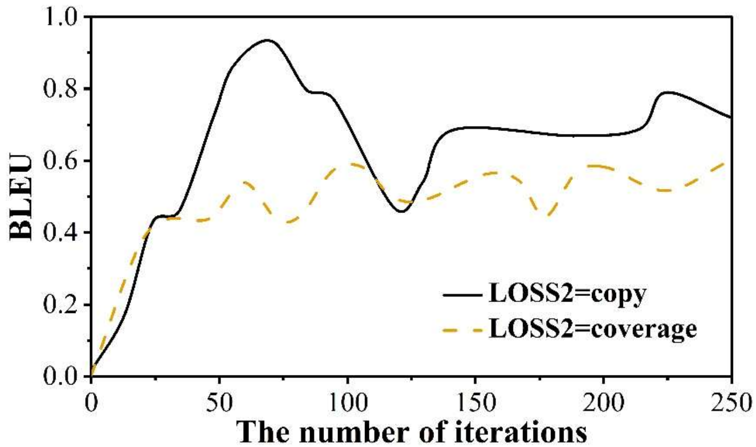


Figure 5: Performance comparison of copy loss or coverage loss

III. A. 2) Multidimensional descriptive statistical analysis

The structured data model was identified above, and in this subsection, we will take the data generated by the structured data model as the baseline to calculate the structured data features of the classical dance drama works in terms of multiple dimensions (storyline, characterization, dance language, music performance, and stage scenes). Before the multi-dimensional correlation calculation and analysis of the structured data of classical dance drama works, we first do a simple descriptive statistics on the samples, so as to have a general understanding of the overall data characteristics of the data. Descriptive statistics is the foundation of the statistical analysis process, which mainly analyzes the centralized trend of the data, the degree of discrete and the correlation between variables. The descriptive statistics analysis of the structured data is shown in Figure 6. It can be seen that the mean value of the structured data for storyline, character image, dance language, music performance, and stage scene is above 3000, indicating that the data generated by the structured data model can satisfy the subsequent research on audience assessment analysis.

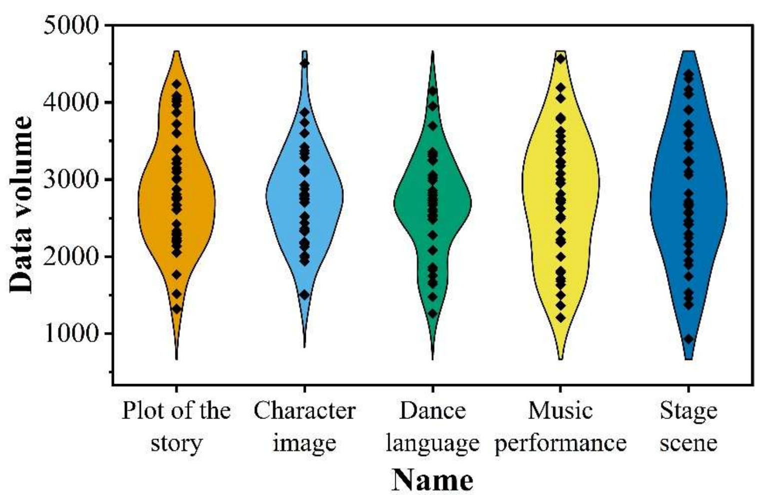


Figure 6: Descriptive statistics of structured data

The above figure gives the descriptive statistics results of the structured data of dance drama works generated by the model, and then the correlation verification calculation will be carried out on the dimensional variables of the structured data of dance drama works, and the results of the correlation verification calculation are shown in Figure 7. It can be seen that there is a strong positive correlation between the dimensional variables, which is in line with the structured data characteristics of dance drama works.

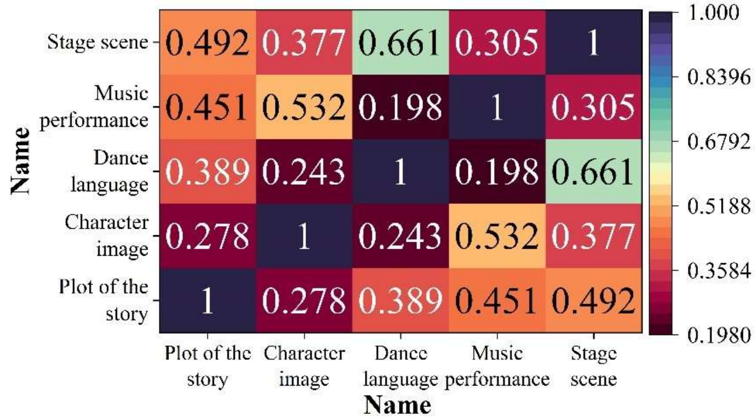


Figure 7: Verify the calculation results of the correlation

III. B. Sentiment Assessment Analysis of Structured Data

III. B. 1) Data sets

The experimental data is generated by the structured data model, a total of 16000 structured data volume of classical dance drama works are generated, and the above described statistical calculation analysis gives the mean value of structured data in five dimensions. According to a certain ratio, the above 16000 structured data volume of classical dance and theater works are divided into test set, validation set and training set.

III. B. 2) Parameters and evaluation indicators

After the experimental platform is built, the experimental parameters are set as shown in Table 1. In this paper, in order to verify the effectiveness of the model, the test dataset is used to evaluate the model. For the results obtained from different models, three evaluation indexes are used to evaluate the model effectiveness, namely, accuracy A, recall R, F1 value, and time T is used to measure the complexity and computation time of the model. The accuracy rate is the ratio of all positive and negative examples that are predicted correctly to the total number of structured dance theater works. Recall is the proportion of correctly predicted positive examples to the proportion of positive examples that are actually positive.The F1 value is the reconciled average of recall and precision. Time T is the time required to train an epoch. The evaluation metrics are calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{23}$$

$$\mathrm{Re}call = \frac{TP}{TP + FN} \tag{24}$$

$$F1 = \frac{2 * P * R}{P + R} \tag{25}$$

Table 1: Experimental parameter Settings

Parameter name	Parameter value
Epoch	100
Batch size	4
Learning rate	4E-5
Embedding size	500
Hidden size	759
Hidden activate	Relu
Dropout	0.8
Optimizer	Adam

In this experiment, positive category data are positive audience emotion categories and negative category data are negative audience emotion categories. Where, TP and TN are the number of correct predictions for positive and negative category samples, respectively. FP and FN are the number of incorrect predictions for positive and negative category samples, respectively.

III. B. 3) Model validation analysis

In order to verify the effectiveness of the model in this paper, three different pre-trained models, Word2vec, GloVe, and BERT, are firstly compared, and the experimental results of pre-trained model comparison are shown in Fig. 8. From the experimental results in Fig. 8, it is concluded that BERT has more powerful characterization ability. Therefore, BERT is chosen as the pre-training model.

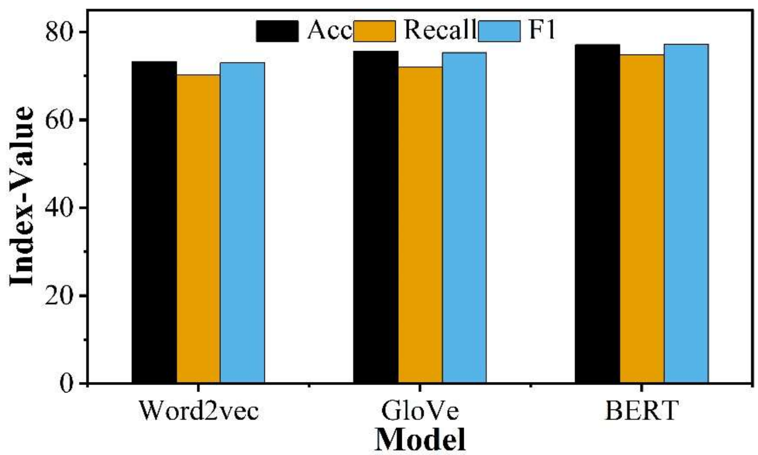


Figure 8: Compare the experimental results of the pre-trained model

The direct truncation method is used as the basic processing method for sentiment analysis, and the structured data is segmented into 2, 4 & 8 segments using BERT-LSTM, BERT-BiLSTM, ATAAE-BERT-BiLSTM, and ATAAE-BERT-BiLSTM based on the segmentation of complete sentences, respectively, for comparison. The experimental results are shown in Table 2. The accuracy, recall, and F1 values of different models for the audience sentiment of classic dance theater works are given in Table 2. Among them, the experimental results of the BERT-BiLSTM model are better than those of the BERT-LSTM model, which proves that BiLSTM has a stronger ability to learn the features of classic dance drama works compared with LSTM, which is conducive to extracting more information. After adding the ATAAE mechanism, the accuracy and F1 value are improved by 1.48% and 1.2%, respectively, which proves that the ATAAE mechanism helps to capture important features of classical dance theater works. The accuracy and F1 value of the BERT-BiLSTM-ATAEE model based on complete segmentation with final segmentation of 4 segments are improved compared to the BERT-BiLSTM-ATAEE model with direct truncation method, which proves that this method helps to improve the accuracy of sentiment classification. Through the analysis, it is obtained that the method of splitting into 2 segments is more difficult to calculate than the direct truncation method, but there is the same problem of information loss, and it is more ineffective than the splitting into 4 segments. The method of splitting into 8 segments maintains the original number of batch, but the memory is not enough. Reducing the number of batch and increasing the number of segments for this purpose will make it easier to lose the connection between segments, and the probability of problems will also increase. Although its slightly better than splitting into 2 segments, it is less effective than splitting into 4 segments and the training time is substantially increased. Training time is also an important factor to consider when there is little difference between F1 and accuracy.

Table 2: Compare the experimental results of the model

Model	A	R	F1	Time/min
BERT-LSTM	78.11	76.23	77.16	4.06
BERT-BiLSTM	80.51	80.24	80.37	5.08
ATAAE-BERT-BiLSTM	81.33	81.72	81.52	5.35
ATAAE-BERT-BiLSTM (2)	82.21	82.88	82.54	8.14
ATAAE-BERT-BiLSTM(4)	83.25	84.51	83.88	12.26
ATAAE-BERT-BiLSTM(8)	82.65	83.52	83.08	19.36

III. B. 4) Analysis of model applications

It is known that the audience's emotional response to dance drama works are happy, sadness, grief and anger, fear of the four categories, the use of this paper's model to assess and analyze the audience's emotional reflection of the 10 classic dance drama works, the results of the model application analysis are shown in Table 3, Table 1, 0 expresses the output of the model. The model evaluates the audience's emotional reaction of classic dance drama work A as fear, and the others are the same, and will not be repeated.

Table 3: Model Application Analysis

Classic dance drama works	Happiness	Sadness	Grief and indignation	Fear
B	1	0	0	1
C	0	0	1	0
D	1	0	0	0
E	0	0	1	0
F	0	1	0	0
G	0	0	0	1
H	1	0	0	0
I	0	1	0	0
J	0	0	1	0
K	0	0	1	0

IV. Conclusion

In this paper, we constructed a multi-dimensional computational analysis and audience emotional response assessment system for structured data of classic dance drama works, and achieved a series of valuable research results. In terms of structured data modeling, the proposed model outperforms the best-performing baseline model, PIVOT, by 2.88 and 0.67 in BLEU and NIST metrics, respectively, which proves the model's high efficiency on structured data of dance dramas. In the sample less learning experiments, the model demonstrates strong generalization ability with less than 500 training samples by showing a stable level of BLEU>25 in all three domains: music, books as well as dance drama. The BLEU-4 for all three domains reaches above 30, and the NIST is above 5, which far exceeds the performance of the baseline model. In the multidimensional descriptive statistics analysis, the mean values of structured data for storyline, character image, dance language, music performance, and stage scene all reach above 3000, and there is a strong positive correlation between the dimensional variables. In terms of audience emotion assessment, the proposed ATAAE-BERT-BiLSTM model achieved an accuracy of 83.25% and an F1 value of 83.88% under the 4-segment segmentation approach, which were 5.14% and 6.72% higher than the BERT-LSTM model, respectively. The experiments proved that the BiLSTM structure has a stronger feature learning ability compared to LSTM, the ATAAE mechanism helps to capture important dance drama features, and the method based on complete sentence segmentation improves the accuracy of sentiment categorization more than the direct truncation method.

This study provides a data-based method for the scientific evaluation of dance drama works, and the constructed multi-dimensional computational analysis and emotion assessment system can effectively support the creation and dissemination of dance drama art, and promote the digital transformation and innovative development of dance drama art.

References

[1] Cayul, M. (2024). Understanding dance and its non-human agency in Ekman's "A Swan Lake". Danza e ricerca. Laboratorio di studi, scritture, visioni, 139-158.

[2] Guo, S., & Zulkafli, N. (2025). Artistic Value and Confucius Culture in Kong Dexin's Dance Drama 'Confucius'. Malaysian Journal of Social Sciences and Humanities (MJSSH), 10(2), e003160-e003160.

[3] Huang, H. (2023). Tracing the Evolution of Self-identity in Performances of The White-haired Girl Since 1945. New England Theatre Journal, 34, 27-52.

[4] Zhanga, M., & Youb, C. (2024). Dance Drama Creation-A Case Study of the Children's. In Proceedings of the 2024 5th International Conference on Mental Health, Education and Human Development (MHEHD 2024) (Vol. 857, p. 427). Springer Nature.

[5] Bao, Q. (2024). The Application and Development of Modern Technology in Dance Art: A Case Study of the Dance Drama Only This Green. Journal of Art, Culture and Philosophical Studies, 1(3).

[6] Pan, C., & Alizadeh, F. (2024). Costume and Body Language of Ethnic Dance Drama under Semiotics: An Analytical Study. Journal of Ecohumanism, 3(7), 741-754.

[7] Huan, G. (2024). Dance drama" Never Gone Radio Wave": the challenge and significance of genre dimension, emotional expression and artistic expression. Advances in Education, Humanities and Social Science Research, 11(1), 80-80.

[8] Song, Y. (2021). Modelling and analysis of aesthetic characteristics using digital technology in artworks. International Journal of Modelling, Identification and Control, 38(3-4), 320-331.

[9] Jin, M. (2025). Analysis of Structured Data Processing and Aesthetic Experience in Classical Dance Theatre Works. J. COMBIN. MATH. COMBIN. COMPUT, 127, 469-485.

[10] Sun, K. (2022). Research on dance motion capture technology for visualization requirements. Scientific programming, 2022(1), 2062791.

[11] Zhong, Y., Fu, X., Liang, Z., & Ning, H. (2024). Application of Three-Dimensional Vision Technology in Dance. Applied System Innovation, 7(5), 80.

[12] Gratch, J. (2021). The field of Affective Computing: An interdisciplinary perspective. Jinko chino, 31(1).

[13] Gloor, P. A., Araño, K. A., & Guerrazzi, E. (2019, October). Measuring audience and actor emotions at a theater play through automatic emotion recognition from face, speech, and body sensors. In Collaborative innovation networks conference of Digital Transformation of Collaboration (pp. 33-50). Cham: Springer International Publishing.

[14] Wigham, P., & Challis, B. (2019, November). Real-time measurement and analysis of audience response. In International Conference on ArtsIT, Interactivity and Game Creation (pp. 38-48). Cham: Springer International Publishing.

[15] Kanipriya, M., Krishnaveni, R., & Krishnamurthy, M. (2020). Recognizing audience feedback through facial expression using convolutional neural networks. International Journal of Engineering Research and Technology, 13(12), 4230-4235.

[16] Peishun Ye,Jiyan Lin,Yaming Kang,Tolga Kaya,Kubra Yildirim,Abdul Hafeez Baig... & Turker Tuncer. (2025). MobileTransNeXt: Integrating CNN, transformer, and BiLSTM for image classification. Alexandria Engineering Journal,123,460-470.

[17] MwataVelu Tat'y,AvinaCervantes Juan Gabriel,CruzDuarte Jorge Mario,RostroGonzalez Horacio & RuizPinales Jose. (2021). Imaginary Finger Movements Decoding Using Empirical Mode Decomposition and a Stacked BiLSTM Architecture. Mathematics,9(24),3297-3297.