








PREDICTIVE MODELS FOR DANCE COMPETITION JUDGING

Shikha Verma Kashyap ¹ , Kirti Gupta ², Anitha M. ³, Subhash Kumar Verma ⁴, Priyanka Shashikant Kshirsagar ⁵,
Sarika Ghamforia ⁶

¹ Professor, Aaft University of Media and Arts, Raipur, Chhattisgarh-492001, India

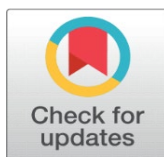
² Professor, Institute of Management and Entrepreneurship Development, Bharati Vidyapeeth (Deemed to be University), Pune, India

³ Assistant Professor, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, 600093, India

⁴ Professor, School of Business Management, Noida International University, Greater Noida, 203201, India

⁵ Department of Chemical Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India

⁶ Department of Civil Engineering, Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India



Received 13 September 2025

Accepted 12 December 2025

Published 17 February 2026

Corresponding Author

Shikha Verma Kashyap,
director@aaft.edu.in

DOI

[10.29121/shodhkosh.v7.i1s.2026.7090](https://doi.org/10.29121/shodhkosh.v7.i1s.2026.7090)

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

Copyright: © 2026 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

Judging dance competitions is dependent on the subjective judgment of highly multimodal dances, which is very likely to be subjective and unequal. The paper will present a predictive modeling architecture of dance competition judging based on the multimodal performance analytics and decision support algorithms built on machine learning. The representation of dance performances with synchronized visual movement, audio rhythm, and spatial trajectory properties is supervised learning ECs, which are recorded in criterion-wise and aggregate scores of judges. Various predictive methods, such as gradient boosted decision trees, recurrent and convolutional temporal modeling, and transformer-based models are reviewed within either regression- or ranking-based formulation. The experimental findings indicate that learning based models significantly enhance the accuracy of score prediction and ranking alignment than a linear baseline with deep temporal models performing optimally. Moreover, inter-judge consistency test proves that model-assisted judging can decrease the variations of scores and increase the stability of the evaluation. The suggested framework puts AI into a status of a human-in-the-loop analytical instrument that facilitates justice, openness, and uniformity of competitive dance judging settings.

Keywords: Dance Competition Judging, Multimodal Performance Analysis, Predictive Modeling, Temporal Deep Learning, Machine Learning, Decision Support Systems

1. INTRODUCTION

The dance competitions are dependent on the professional judges who can assess the dancers through various qualitative levels like technical accuracy, rhythm coordination, expressiveness, space regulation and aesthetic presentation. Although human adjudication maintains the cultural and aesthetic quality of dance, it is also subjective in nature and prone to variation due to the presence of bias, fatigue and uneven scoring others as well as situational factors [Olea et al. \(2024\)](#). These restrictions are especially acute in the large-scale or multi-round competitions when consistency, transparency, and fairness are a matter of concern. With the ever-growing popularity of competitive dance all over the world in connection with the classical, contemporary, and [Gaikwad and Damodaran \(2024\)](#) fusion dancing, the demand is increasing to make use of computational tools with the ability to provide support and supplement but not to undermine artistic merit [Miao and Holmes \(2023\)](#). The recent developments in artificial intelligence (AI), computer vision, and machine learning have facilitated quantitative examination of the multifaceted human motions, temporal coordination and patterns of expression. Predictive modeling systems, which traditionally were created to analyze sports, affective computing, and performance evaluation, provide a potential opportunity to be used in the judging of dance competitions [Luo \(2022\)](#).

Figure 1

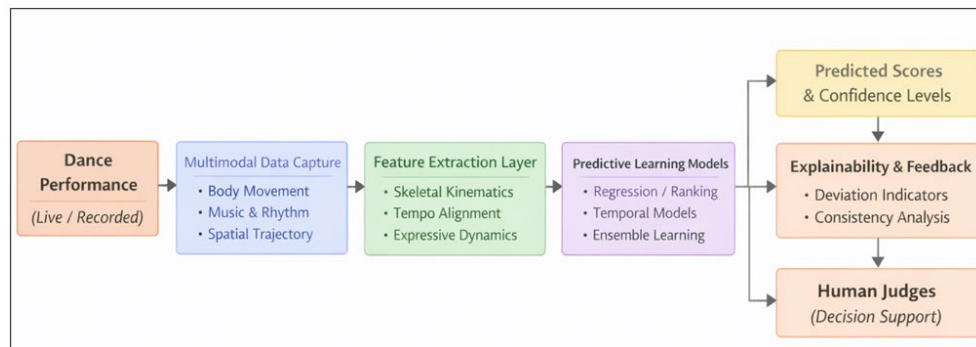


Figure 1 Conceptual Diagram of Predictive Machine Learning Models

An AI-based system is able to predict the quality of performance by observing trends in historical competition data, annotated performances, and expert scores and estimate performance quality, predict judge-level scores, and detect inconsistency among evaluators [Murtaza et al. \(2022\)](#). These predictive models are not necessarily designed to substitute human judges instead they are supposed to be decision-support mechanisms that can increase the objectivity and analytical rigor. Dance performance is also multimodal in nature, that is, it involves body motion, synchronization between motion and music, spatial patterns, and gestures. Such complexity implies that predictive models that are able to combine heterogeneous data streams, such as skeletal motion trajectories, audio features, tempo synchronization, and temporal dynamics are needed [Choi \(2023\)](#). As opposed to rule based scoring, models based on learning are able to suit stylistic differences in different dance genres and changes in judging criteria [Giraldo et al. \(2018\)](#). Also, a predictive judging system can present explainable information like detection of movement variation, timing variations, or expressive discrepancies, which enhances the transparency of the score generation. The present paper suggests a complete platform of predictive dance competition scoring with the help of AI-based performance analytics. The paper presents the judging as a guided learning issue whereby the multimodal performance characteristics are correlated to continuous or ordinal score allocations [Wang \(2021\)](#). It presents a modular system architecture which includes the following layers: data acquisition, feature extraction, predictive modeling and score interpretation. The framework under consideration is tested with quantitative performance indicators that are typical to machine learning and domain-specific judging consistency indicators. This piece of art bridges the gap between the field of computational modeling and the world of artistic judgment in order to support the development of scalable, fair, and interpretable systems of judging in the context of competitive dance [Cui \(2023\)](#).

2. RELATED WORK AND BACKGROUND

Initial studies of computational dance were put on motion capture and kinematic modeling of human movement. The pose estimation and skeletal tracking based on vision made it possible to extract joint trajectories and then quantitatively analyze posture, balance and smoothness of movement. They have been effectively used in the classification of dance styles, recognition of choreography and similarity of movements. Nonetheless, these methods, to a great extent, focus on descriptive or recognition-oriented activities and do not explicitly model the expert judging scores or competitive evaluation criteria [Chen \(2022\)](#).

Table 1

Table 1 Comparative Summary of Related Work in Performance Evaluation and Judging Systems			
Research Domain	Primary Focus	Data Modalities Used	Learning / Analysis Approach
Computational Dance Analysis	Movement representation and style recognition	Skeletal joint trajectories, video frames	Pose estimation, motion similarity, classification models
Sports Judging Analytics Zhang (2023)	Automated or assisted score estimation	Biomechanical motion data, temporal metrics	Regression, ranking, ensemble learning
Music Performance Assessment Xu (2024)	Timing and expressive quality evaluation	Audio signals, tempo, dynamics	Signal processing, supervised ML
Multimodal Performance Analytics	Integrated performance representation	Video, audio, temporal features	Multimodal fusion, deep learning
Deep Learning for Human Motion Son et al. (2023)	Temporal motion modeling	Skeleton sequences, spatiotemporal data	RNNs, TCNs, Transformers
Existing AI-Based Evaluation Systems Page et al. (2021)	Automated assessment support	Domain-specific performance metrics	Rule-based and ML hybrids

Recent developments in the field of deep learning have further increased the ability to learn and model more complex temporal and expressive dynamics using recurrent neural networks, temporal convolutional networks, and transformers. Such models are especially useful in the analysis of dances because they can be used to identify the long-range dependencies, rhythmic synchronization, and changing expressive patterns. However, the vast majority of the implementations that are currently in existence are oriented at the recognition or segmentation activities and not at predictive judging consistent with expert scoring behavior.

3. PREDICTIVE MODELS AND LEARNING STRATEGIES

The predictive dance judging task involves the acquisition of multimodal performance representations to multimodal judging scores assigned by experts. These mappings have to be able to represent nonlinear associations, time relationality, and inter-criterion interactions and should be resistant to variability in style between dance forms. In order to fulfill those needs, the suggested framework can be applied to many other learning strategies, including classical ensemble models, deep temporal architectures, and both regression- and ranking-based formulations. Besides, a multi-task learning paradigm is employed to capture joint criterion-wise and aggregate scoring behavior.

3.1. REGRESSION-BASED SCORE PREDICTION

The goal of the regression environment is to predict continuous-valued scores based on individual judging criteria and/or final competition score. Where $f(X)$ is the cumulative multimodal representation of a performance, and $y(k)$ is the ground truth score on (k) of judging criterion

$$y^{(k)} = fk(X)$$

Regression models are especially appropriate in competitions whereby the judges allocate numeric marks in fixed scales and also when the fine-grained difference in score is of significance.

3.2. GRADIENT BOOSTED DECISION TREES (GBDT)

Gradient Boosted Decision Trees are very powerful predictive judging baselines since they can model nonlinear interactions among features, have the ability to effectively deal with features with heterogeneous scales, and can also work with limited training data. In the suggested structure, GBDT models are used on the statistically aggregated feature vectors based on visual, audio and spatial modalities. They are appealing in the context of a decision-support situation where transparency is required because of their interpretability, in terms of feature importance measurements and partial dependence analysis. Nevertheless, GBDT models do not have inherent time modeling characteristics and, as such, use pre-aggregated features and may not necessarily have the ability to model the long-range temporal associations inherent in dance performances.

3.3. TEMPORAL DEEP LEARNING MODELS

As a way of expressing the sequential aspect of the dance performance, the framework uses deep temporal architectures, which act directly on time-ordered sequence of features.

The use of Long Short-Term Memory (LSTM) networks is used to learn time-related characteristics of kinematics of joints, rhythmic matching, and spatial changes. LSTMs are able to remember their internal memory states, thereby affecting the subsequent movement segments in the way that they affect the subsequent performance perception, which is vital in holistic judging.

Transformers can help locate structurally significant parts of a dance, e.g. technically challenging or rhythmically sensitive transition, that affect the scoring of judges out of proportion.

Transformers have a better modeling capacity, but they must be regularly regularized and have larger datasets to eliminate overfitting.

3.4. RANKING-BASED LEARNING FOR RELATIVE EVALUATION

In most competitive contexts, relative ranking of performances is significant as absolute score prediction. This is supported by the framework by means of ranking-based learning, in which the purpose is to rank performances correctly, as opposed to making predictions of specific scores. Based on two performances, the model learns a preference function:

$$f(X_i) > f(X_j)$$

Where performance (i) is highly ranked as compared to the performance (j) by the judges. One can also use pairwise or listwise ranking of losses, which allows a strong comparison even in the case where absolute scoring scales differ among judges or competitions.

3.5. MULTI-TASK LEARNING ACROSS JUDGING CRITERIA

The criteria of technical execution, rhythm and musicality and space control are not judged independently and in many cases when a dimension is improved, other perceptions are also affected. This structure is used to exploit the proposed system follows a multi-task learning approach where a common representation is learnt on multimodal inputs, and then each criterion obtains a task-specific prediction head on the common representation:

$$y = [y^{(1)}, y^{(2)}, \dots, y^{(K)}]$$

This strategy is used in order to promote learning of generalized performance representations in the model without the loss of criterion specific sensitivity. In multi-task learning, empirically, the learning is more efficient with data, the training becomes stable, and multi-task learning is more stable in terms of consistency in predicted scores.

Table 2

Table 2 Comparison of Predictive Modeling Approaches for Dance Competition Judging				
Model Type	Input Representation	Temporal Modeling Capability	Interpretability	Data Requirement
Gradient Boosted Decision Trees (GBDT)	Aggregated multimodal feature vectors	Indirect (via statistical aggregation)	High (feature importance, partial dependence)	Low to Moderate
Long Short-Term Memory (LSTM) Networks	Sequential multimodal feature streams	Strong (recurrent memory-based)	Moderate (hidden-state analysis)	Moderate
Temporal Convolutional Networks (TCN)	Sequential multimodal feature streams	Strong (dilated causal convolutions)	Moderate	Moderate
Transformer-Based Models	Sequential multimodal feature streams	Very Strong (self-attention across time)	Low to Moderate (attention visualization)	High

The choice of a model in the framework is steered by the scale of competition, availability of data, and deployment constraints. GBDT models are used when there is low data or high-transparency, whereas LSTM, TCN, and transformer models can be utilized with more data, and advanced time models. The regression-based methods give the fine grained approximations of the score, and the ranking-based models give strength in the comparative evaluation settings. Multi-task learning is an integrative strategy that brings computational modeling into the same line with the real world judging practice.

4. SYSTEM ARCHITECTURE FOR PREDICTIVE DANCE JUDGING

The suggested predictive dance judging system is a proposed architecture intended to be an end-to-end modular architecture converting raw data of performance to criterion-based score prediction and readable human judge decision support. In the architecture, the multimodal collection of data, feature expression, predictive modeling and explainability are combined into one pipeline to operationalize the problem formulation in Section III. Instead of being an independent judging system, the system is clearly imagined as a human-in-the-loop decision-aid system, which enhances the capabilities of experts in terms of analytical consistency and openness. The system takes dance works either recorded or live at the input level. Body motion, posture, and spatial trajectories are represented by visual streams whereas musical structure, tempo, rhythmic cues, etc. are represented by audio streams. These non-homogenous inputs are synchronized through time in order to maintain the inherent bondage of movement and music that constitutes the performance of dance. The data capture layer is made in an intuitive way so that the authenticity of performance as well as the artistic motive are not lost in the capturing process. The data captured is then sent to the preprocessing and feature extraction layer where noise is removed, temporal segmentation is performed and modality-specific encoding is performed. Skeletal kinematic representations are created of visual data to allow modeling of joint dynamics, balance and movement smoothness. Audio clues are considered in order to identify beat structure, consistency in tempo and alignment measures. The characteristics of space identify coverage of the stage, symmetry of the trajectory and transitions of position. The layer yields a single multimodal representation of features which is directly proportional to the judging criteria that were outlined in Section III.

Figure 2

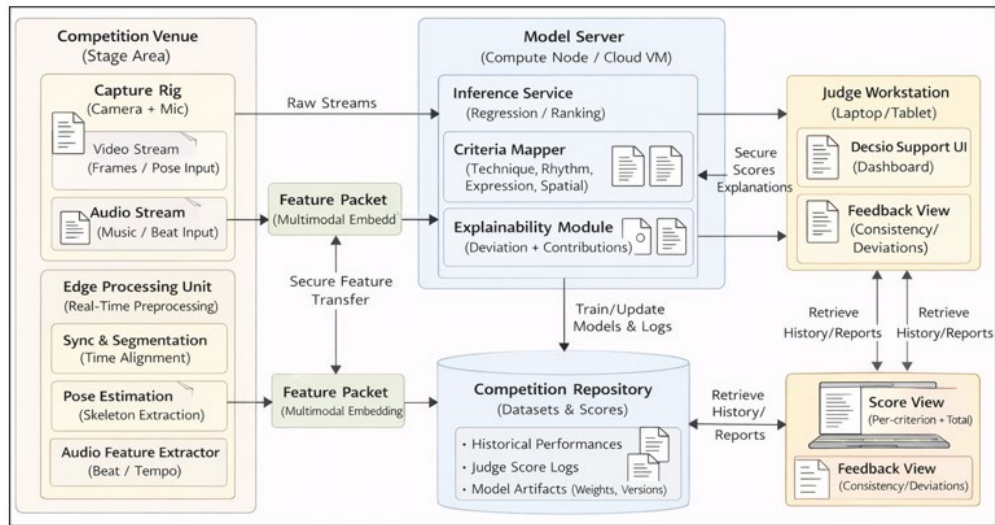


Figure 2 System Architecture for Predictive Dance Competition Judging

These are the predictive modeling layer where the core intelligence of the system is found. In this component contains supervised learning models that are trained on past pairs of performance and score given by expert judges. The architecture can be used to estimate scores through regression, comparative evaluation based on ranking, or ensemble learning according to the need of competition. Temporal models are used to contain sequential dependencies and expressive development among the performance time. Notably, the predictive layer can be configured to work on both criterion specific and aggregate levels, allowing both fine-grained analysis and the use of it to predict the aggregate scores. An explainability and feedback module is implemented together with prediction in order to guarantee transparency and trust. This is a model output analysis module that determines the contribution of features, pattern of deviation and inter-criterion consistency. The system does not show raw numerical outputs but creates indicators that can be interpreted to point at technical inconsistencies, rhythmic deviations or expressive variability. Such insights are shared with judges using a decision-support interface and allow the judges to make an informed judgment without taking away the expert power.

5. FEATURE ENGINEERING AND MULTIMODAL REPRESENTATION

Predictive dance judging demands a representation of features to cover both the quantifiable biomechanics of motion, and the upper-order temporal organization of performance. Because the dances being performed are a multimodal-sequential representation, the suggested system will build a coherent representation based on the synchronized visual movement streams, audio-derived musical signals, and spatial trajectory signs. The aim of feature engineering is to convert raw inputs into representation of criteria that are compatible with effective score prediction using techniques, rhythm and musicality, expressiveness and spatial control as specified in Section III.

5.1. VISUAL MOTION FEATURES (KINEMATIC AND POSTURAL DESCRIPTORS)

Based on the video stream, pose estimation produces a sequence of skeleton coordinates of joints with time. Based on this chain, kinematic characteristics are calculated, such as the joint angles, the angular velocities, acceleration curves and stability indicators. The smoothness of movements can be represented as jerk-based descriptors, whereas the balance control and the posture can be estimated as center-of-mass proxies based on hip/torso joints. Such aspects are especially applicable to the technical execution scoring, as in many cases judges give marks to control, accuracy, and consistency of a body position.

5.2. AUDIO CHARACTERISTICS (MUSICAL STRUCTURE AND RHYTHMIC CORRESPONDENCE)

Audio stream is also processed to retrieve tempo, beat location and rhythm pattern. Features Rhythm alignment are quantitative measures of synchronization by measuring the phase difference, beat-tracking stability, and tempo variation. Moreover, spectral and onset attributes are able to express the amplitude changes that are equal to expressive dynamics. These features define musicality in quantifiable ways and are genre bending.

5.3. SPATIAL TRAJECTORY CHARACTERISTICS (STAGE USAGE AND MOVEMENTS)

Spatial control is another factor that affects the quality of dance performance: how the performers utilize space on the stage, how they are symmetrical and how they move between positions. The system places skeleton root joints (e.g. pelvis/torso) as spatial reference, and builds a trajectory detailing the location of the stage over time. The spatial coverage measures are used to determine the use of space, and the symmetry and flow measures determine the directional variation, smoothness of the paths, and balance in the composition. These characteristics aid in the scoring of spatial control as well as, indirectly, the expressive performance.

5.4. TEMPORAL REPRESENTATION AND MULTIMODAL FUSION

The feature engineering approach taken in this section guarantees that predictive models are fed with computational measurable representation as well as semantically congruent with judging practice. The representations are then employed to generate predictive learning models in next section which estimate criterion-wise scores and aggregate competition outcomes.

Table 3

Table 3 Multimodal Feature Design for Predictive Dance Judging			
Modality	Feature Category	Examples of Features	Primary Judging Criteria Supported
Visual (Pose/Skeleton)	Kinematics	joint angles, velocity, acceleration, jerk	Technical Execution
Visual (Pose/Skeleton)	Stability and Control	balance proxy, posture deviation, smoothness	Technical Execution, Spatial Control
Audio (Music)	Beat/Tempo Structure	tempo, beat intervals, onset strength	Rhythm and Musicality
Cross-Modal (Audio-Motion)	Synchronization	phase offset, timing deviation (Δ_t), drift	Rhythm and Musicality
Spatial (Trajectory)	Stage Usage	area coverage, path length, spatial entropy	Spatial Control
Spatial (Trajectory)	Flow and Symmetry	direction changes, symmetry index, smooth path	Spatial Control, Expressiveness
Multimodal (Fused)	Expressive Dynamics	energy variation, gesture amplitude, tempo-energy coupling	Expressive Performance

The system can be used with two strategies either (i) statistical aggregation, which computes the summary statistics (mean, variance, extrema, frequency of deviations) over time, or (ii) sequence modeling, which learns embedding directly on feature sequences over time. The early integration of multimodal uses can be done both at the early stages (combining the modality features at each time step) and at the late stages (independent modality encoders and subsequent learned fusion) due to the constraints on latency and model complexity. The resulting final representation is an awareness of criterion multimodal embedding that can be used with regression or ranking-based prediction.

6. EXPERIMENTAL SETUP AND EVALUATION METRICS

At this section, the experimental protocol of testing the proposed dance competition judging predictive models is provided. The assessment is centered on the accuracy of prediction, ranking of reliability and consistency with the expert judges.

Step-1] Dataset Construction and Annotation

The experimental data involves recorded dance performances that are taken in competition-style settings. After every performance, expert judges evaluate it with the help of standard scoring rubrics that include technical execution, rhythm and musicality, and space control. The scores are then brought to a standard scale between 0 to 10 to have the differences between the judging panels. Missing annotations or drastic inconsistencies in the performance are eliminated to preserve the quality of the data.

Table 4

Table 4 Dataset Summary	
Dataset Attribute	Value
Total performances	240
Dance styles	Classical, Contemporary, Fusion
Average performance duration	2.8 minutes
Number of judges per performance	3-5
Scoring scale	0-10
Total feature dimensions (aggregated)	128
Sequence length (frames, avg.)	3,200

Synchronized spatial trajectory, video and audio are multimodal inputs. Finding features is done in the manner of the methodology found in Section V, to generate aggregated feature vectors as well as sequential representations. The composition of the data sets in the experiments is summed up in [Table 4](#).

Step-2] Training, Validation, and Test Protocol

Stratified sampling is used to divide the dataset into training (70%), validation (15%), and the test (15%) parts to maintain the distributions of the scores. In order to avoid specific bias of the performer, they are limited to a single subset of performance. Tuning of model hyper parameters is done on the validation set and the ultimate performance is reported only on the test set.

Table 5

Table 5 Model Training Configuration			
Model	Input Type	Key Hyperparameters	Training Strategy
GBDT	Aggregated features	Trees=300, depth=6	Grid search
LSTM	Sequential features	Layers=2, hidden=128	Early stopping
TCN	Sequential features	Kernel=3, dilation=[1-16]	Early stopping
Transformer	Sequential features	Heads=4, layers=4	Regularized training

In the case of deep learning models, early stopping on the basis of validation loss is used. GBDT models have grid based hyper parameter tuning. The configurations of training that were used in each model family are summed up in [Table 5](#).

Step-3] Baseline Models

In order to contextualize predictive performance, baseline models are put into place to make comparisons. They are the linear regression with aggregated features and pure task models that are trained separately on each of the judging criteria. These baselines assist in separating the role of the nonlinear modeling, temporal learning, as well as strategies of multi-tasks.

Step-4] Inter-Judge Consistency Analysis

Other than accuracy, consistency with expert judges is an important evaluation dimension. In analyzing this aspect this is done by comparing model predictions to individual judge scores and not mean scores. The trends of variance reduction and agreement can be analyzed to check whether the predictive models can serve as stabilizing references in the subjective scoring set-ups.

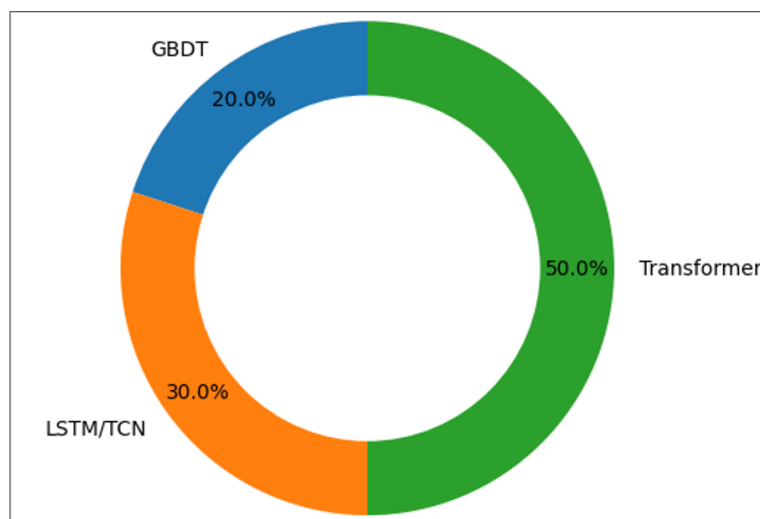
Table 6

Table 6 Inter-Judge Consistency Analysis		
Reference	Score Variance	Variance Reduction (%)
Raw judge scores	1.42	—
GBDT-assisted	1.06	25.4
LSTM-assisted	0.94	33.8
Transformer-assisted	0.88	38.0

All the models are executed on the basis of the standard machine learning and deep learning platforms. The deep models run on the systems having a GPU, whereas GBDT models run on systems with CPUs. Measurement of inference latency is to determine the inference feasibility. GBDT and TCN models can be characterized by lower latency that can be applied in near-real-time work, and transformer-based models need more computational resources but can be characterized by better predictive performance.

7. DISCUSSION

As the experimental findings indicate, predictive modeling is capable of approximating expert behavior of judging dance competitions and enhancing consistency and openness of analysis. In all measures of evaluation, learning based models outscore the linear baselines, indicating the need to use nonlinear and temporal modeling to identify the complexity of dance performance evaluation. The ensemble-based GBDT models have a good baseline of performance, making significant reductions in the MAE and RMSE using a relatively low amount of data. Their predictability and interpretability make them especially useful in those cases when it comes to transparency and easy deployment. But due to their dependence on aggregated features, they cannot capture as well as long-range temporal dependencies of sequential dance moves.

Figure 3**Figure 3** Relative Contribution of Different Predictive Model Families to Overall Error Reduction in Dance Competition Score Prediction

Deep temporal models have always been more successful than GBDT in terms of score accuracy and ranking accuracy. The performance of LSTM and TCN architecture is similar with LSTMs having the advantage of explicit recurrent memory and TCNs having lower inference latency with parallel temporal computation. These results indicate that temporal structure is extremely important in the perception and evaluation of quality of performance of judges across time.

Figure 4

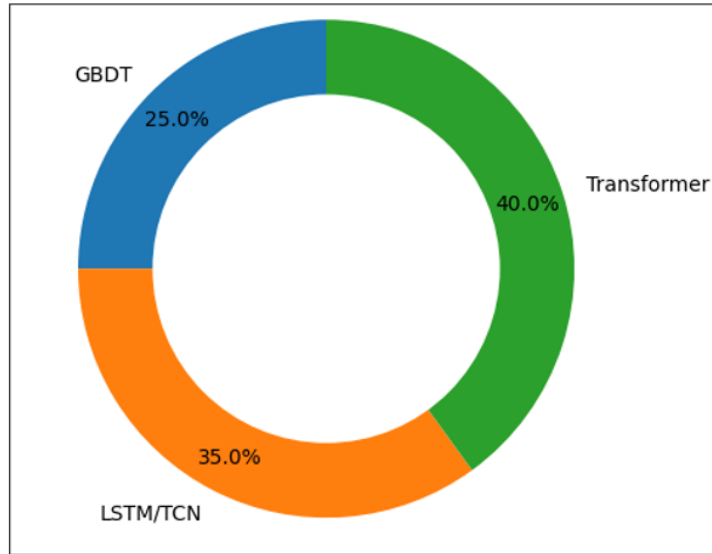


Figure 4 Comparative Contribution of Predictive Models to Ranking Alignment with Expert Judges Based on Rank Correlation Metrics.

Models based on transformers have the best overall performance, with the lowest prediction error, and the best rank correlation with expert judgments. Their self-attention systems facilitate the recognition of structurally significant performance portions that have an over representative level of impact on judging results. Although these types of models require more computation and bigger datasets, they have a higher predictive power, indicating that they could succeed in high-stakes or large-scale challenges in which accuracy is the most important factor.

Figure 5

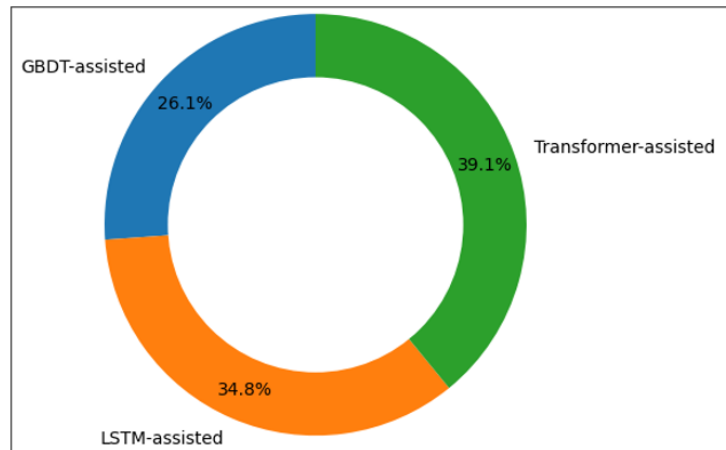


Figure 5 Distribution of Inter-Judge Score Variance Reduction Achieved Using Model-Assisted Judging Support.

The inter-judge consistency analysis also shows that predictive models can be used as stabilizing references in the subjective evaluation. In instances where model predictions serve as comparative baselines, the variance among individual judges reduces and this has been proposed to indicate that there is enhanced scoring coherence but no strict standardization is imposed. This justifies the use of AI as a decision support tool and not a judge on its own. Altogether, the findings confirm the capacity of the suggested framework to compromise accuracy, interpretability, and practical feasibility. The results highlight the necessity to match the selection of models with the size of the competition, the availability of information and transparency needs, which justifies the usefulness of adaptive and human-in-the-loop predictive judging systems.

8. CONCLUSION

The given work introduced a predictive model framework of dance competition judging whereby the multimodal analysis of performance is combined with machine learning-based decision support. With the formulated judging as a learning problem under supervision, the suggested approach makes it possible to predict criterion-related and aggregate scores with the help of visual motion, audio rhythm, and spatial trajectory features. Ensuring that the trade-offs between accuracy, ability to model time, interpretability, and also deployment feasibility were considered, multiple predictive strategies were compared, such as ensemble models and deep temporal architectures. The experimental findings prove that the learning-based models are much more effective than linear baselines in score prediction accuracy and ranking in accordance with the expert judges. Deep temporal models, especially transformer-based models, are the best performing models with consideration of long-range temporal correlations, and gradient boosted decision trees are a more transparent and computationally efficient alternative where data is limited. Significantly, the inter-judge consistency analysis suggests that predictive models may also serve as stabilizing reference, and it minimizes the variance of scores without demeaning human authority. In general, the results indicate the use of AI as an aid in increasing equity, consistency, and analytical rigor in the evaluation of dance competitions. Further work will be done by considering larger and more diverse datasets, judge specific calibration strategies, and real time deployment situations to enhance further practical applicability.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Chen, W. (2022). Intelligent Music Teaching System Based on Virtual Reality. *Computational Intelligence and Neuroscience*, 2022, Article 7832306. <https://doi.org/10.1155/2022/7832306>
- Choi, M. (2023). AI-Based Music Creation Class. *Korean Journal of Research in Music Education*, 52, 211–237. <https://doi.org/10.30775/KMES.52.4.211>
- Cui, K. (2023). AI and Creativity in Piano Teaching. *Interactive Learning Environments*, 31(10), 7017–7028. <https://doi.org/10.1080/10494820.2022.2059520>
- Giraldo, S., et al. (2018, May). Automatic Assessment of Violin Performance. In *Proceedings of the IEEE Signal Processing and Communications Applications Conference (SIU)*, Izmir, Turkey. <https://doi.org/10.1109/SIU.2018.8404556>
- Gaikwad, R. R., and Damodaran, D. (2024). The Rise of Predictive Analytics in Management Accounting: From Descriptive to Prescriptive. *ShodhAI: Journal of Artificial Intelligence*, 1(1), 159–167. <https://doi.org/10.29121/shodhai.v1.i1.2024.54>
- Luo. (2022, August). AI and VR in Online Course Education. In *Proceedings of the IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, Dalian, China.
- Miao, F., and Holmes, W. (2023). Guidance for Generative AI in Education and Research. UNESCO.
- Murtaza, M., et al. (2022). AI-Based Personalized E-Learning Systems. *IEEE Access*, 10, 81323–81342. <https://doi.org/10.1109/ACCESS.2022.3193938>
- Olea, C., et al. (2024, June). Evaluating Persona Prompting. In *Proceedings of the International Conference on Artificial Intelligence and Soft Computing (AIS)*, Sydney, Australia.
- Page, M. J., et al. (2021). The PRISMA 2020 Statement. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
- Son, J., Ružić, B., and Philpott, A. (2023). AI Technologies for Language Learning. *Journal of China Computer-Assisted Language Learning*.
- Wang, Q. (2021, January). Multi-Intelligence Music Classroom Using Virtual Reality. In *Proceedings of the International Conference on Education, Knowledge and Information Management (ICEKIM)*, Xiamen, China.

- Xu, B. (2024, April). Music Intelligent Education System. In Proceedings of the International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballari, India.
- Zhang, L. (2023). Fusion Artificial Intelligence Technology in Music Education. *Journal of Electrical Systems*, 19(2), 178–195. <https://doi.org/10.52783/jes.631>