International Journal of Multiphysics
Volume 18, No. 3, 2024
ISSN: 1750-9548

Enhancing Remote Choral Education Expirence: The Application of Machine Learning Algorithms for Real-Time Audio Synchronization and Pedagogical Feedback

Shuo Huang [1], Han Wang [2*]

[1] Yunnan Tourism College, School of Culture&Arts, Kunming City, Yunnan Province, 650031, China; soniahuang1018@gmail.com

[2*]Xishuangbanna Vocational and Technical College, Faculty of Teacher Traning, Jinghong City, Yunnan Province, 666100, China;wanghanscholor@gmail.com

Corresponding author：Han Wang, Email:wanghanscholor@gmail.com

Abstract This study examines the impact of machine learning algorithms on enhancing real-time audio synchronization and pedagogical feedback in remote choral education. Utilizing data from 121 schools in China, a Random Forest Regressor analyzed critical variables including Average latency per session (ALPS), audio clarity and synchronization accuracy interaction (ACSA Interaction), lagged latency variability (LLV), and normalized performance quality Score (NPQS). Results indicate that ALPS and ACSA Interaction are particularly significant in influencing pedagogical effectiveness, while LLV and NPQS also play crucial roles in the consistency and quality of remote choral sessions. Furthermore, feedback frequency (FF) and background noise level (BNL) are also important indicators for improving remote choral education experience. The study highlights the essential role of technological enhancements in remote education and suggests further improvements in network stability and audio processing technologies to optimize remote learning outcomes.

Keywords: Remote Choral Education, Machine Learning, Audio Synchronization, Pedagogical Feedback, Random Forest Regressor, Educational Technology

1. Introduction

Digital technologies have transformed arts education by enabling enhanced accessibility, interactivity, and innovation across various disciplines(Timotheou et al., 2023). Specifically, choral education has evolved significantly, transitioning from traditional in-person rehearsals to incorporating digital tools that facilitate learning, performance, and evaluation(Tsugawa, 2022). Remote choral education has emerged against a backdrop of necessity(Biasutti et al., 2024), particularly during global crises such as the COVID-19 pandemic, which restricted physical gatherings(Gössling et al., 2020). This adaptation has led to the development of platforms and methodologies that enable real-time collaboration among singers dispersed across multiple locations(Mróz et al., 2022). Moreover, it has enhanced the inclusivity and accessibility of arts education, allowing individuals from diverse geographic and socio-economic backgrounds to participate in and benefit from structured choral training without the constraints of physical proximity(Maury & Rickard, 2020).

The integration of digital technologies in remote choral education not only solves logistical issues but also introduces a new dimension of pedagogical strategies(Lee & Hwang, 2022). For instance, sophisticated audio synchronization tools and latency management techniques have become crucial in maintaining the temporal and harmonic coherence necessary for choir performances(Mavaddati, 2020). These technological advancements allow conductors to effectively manage and direct remote ensembles(Mellit & Kalogirou, 2021), ensuring that educational objectives are met and artistic integrity is maintained(Muhammad et al., 2020). Furthermore, the digital shift facilitates a broader dissemination of artistic education, promoting equality by democratizing access to high-quality resources and expert training(Locke & Wright, 2021). This development is particularly significant in arts education, where opportunities were traditionally limited by geographical and financial barriers. Thus, the digital transformation in choral education not only addresses immediate logistical challenges but also aligns with long-term educational goals of inclusivity and equal access(Palkki, 2019), thereby reshaping the landscape of arts education in the digital era(Decuypere & Landri, 2021).

Despite significant advancements in digital communication technologies such as high-speed internet(Niu et al., 2020), cloud computing, and real-time audio processing software, the field of remote choral education still faces substantial challenges related to synchronizing audio inputs from multiple remote sources in real-time(Zhang, 2023). These challenges primarily include network latency, jitter, and packet loss, which are inherent issues in transmitting audio over the internet(Turchet & Casari, 2024). Network latency refers to the time it takes for data to travel from the source to the destination(Lin et al., 2021), causing delays in audio streams that disrupt the synchronous ensemble singing necessary for choral performance(Li, 2024). Jitter, or the variability in packet arrival times, further complicates these delays, leading to audio distortions that can hinder the choir's ability to maintain consistent timing and pitch(Kirsch et al., 2023). Packet loss, where data fails to reach its destination, can result in missing audio fragments, adversely affecting the musical integrity and continuity of the performance(Liz-López et al., 2024).

These technical impediments significantly impact the overall quality of performance and pedagogical effectiveness in remote choral education(Shi, 2023). The lack of precise synchronization makes it difficult for conductors to ensure cohesiveness and ensemble integrity during performances and rehearsals(Meals, 2020). This, in turn, limits the effectiveness of pedagogical feedback, as conductors struggle to accurately assess and address individual and collective musical nuances in a remote environment(Chan & Luo, 2022). As a result, both the educational and artistic aspects of choral education are compromised(de Quadros & Abrahams, 2022), presenting a barrier to achieving the high standards typically expected in traditional, in-person choral settings(Parkes et al., 2021). Therefore, addressing these synchronization challenges through improved technological solutions and innovative teaching approaches remains a critical task for enhancing the efficacy and appeal of remote choral education(Parkes et al., 2021).

Current technologies often involve significant latency issues that not only hinder the synchronization necessary for quality choral performance but also affect the interaction between conductors and performers, limiting the educational and creative potential of remote choral settings(Mróz et al., 2022). Furthermore, existing solutions lack sufficient pedagogical feedback mechanisms(Deeva et al., 2021), which are critical for effective learning and skill development in choral education(Jansson & Balsnes, 2022).

This research aims to explore the potential of advanced machine learning algorithms and audio-visual technologies to enhance real-time audio synchronization and provide dynamic pedagogical feedback in remote choral education environments. The goal is to improve the Quality of Experience (QoE) for participants and to facilitate a more engaging and effective educational setting for remote choral practices.

International Journal of Multiphysics
Volume 18, No. 3, 2024
ISSN: 1750-9548

2. Materials and Method

In the quantitative examination of the impact of remote choral education technologies across 121 schools in China, data collection is meticulously undertaken to gather comprehensive information on several critical variables. These include latency times, which measure the delay in audio transmission between different sources, audio clarity as assessed by technical software or experts, synchronization accuracy, performance quality ratings provided by music educators, and pedagogical outcomes based on the educational objectives achieved during sessions. The integrity of this dataset is paramount, hence, prior to analysis, rigorous data cleaning and normalization procedures are implemented to ensure accuracy and reliability. This preparation involves handling missing values, removing outliers, and scaling the data to facilitate effective analysis.

For the analytical process, a supervised machine learning model, specifically the Random Forest Regressor, is chosen for its robustness in handling nonlinear relationships and its resistance to overfitting(Bayat Pour et al., 2023).

The Random Forest Regressor is a powerful ensemble learning method used for regression tasks(Pinto et al., 2021), which operates by building multiple decision trees during the training phase and outputting the average of the predictions from the individual trees(Sagi & Rokach, 2020). This model is well-regarded for its ability to handle large datasets with multiple input features and its robustness against overfitting, making it particularly effective in complex predictive modeling scenarios where the relationship between variables might be nonlinear.

The Random Forest algorithm starts by creating multiple decision trees from randomly selected subsets of the training dataset(Jalal et al., 2022). Each tree in the forest is built from a different sample of data and features, ensuring that the trees are varied and that the model is not overly dependent on any single aspect of the data(Coops et al., 2021). When making predictions, each tree in the forest votes, and the average of these votes is considered as the final output of the Random Forest Regressor.

The general function of a Random Forest Regressor can be described mathematically as follows:

$$Y = \frac{1}{N} \sum_{i=1}^{N} f_i(X)$$

Where:

- Y is the predicted output.
- N is the number of trees in the forest.
- $f_i(X)$ represents the prediction of the i^{th} tree.
- X are the input features used to make the prediction.

One of the strengths of the Random Forest Regressor is its capability to perform feature selection. It inherently ranks the importance of different features based on how well they improve the purity of the node. This is measured by a metric called the Gini impurity or mean squared error (MSE), depending on the context of the problem. Nodes with lower impurity are more predictive and are given higher importance in the model. This feature ranking helps in understanding the relative importance of each feature in predicting the outcome, allowing researchers and analysts to focus on the most influential factors.

The Random Forest Regressor offers several key advantages that make it particularly effective for complex predictive modeling. Firstly, its robustness to overfitting is enhanced by the averaging of predictions from multiple decision trees, reducing the risk commonly associated with individual trees. Additionally, it excels in handling non-linearity, as it can capture complex, nonlinear relationships between features and the target variable without requiring variable transformation. Finally, it provides valuable insights through feature importance evaluation, highlighting which variables are most critical in predicting the outcome. This combination of features makes the Random Forest an invaluable tool in fields requiring detailed and accurate predictive analysis.

The Random Forest Regressor's ability to provide accurate and interpretable results, even with complex and multidimensional data, makes it an excellent choice for tasks such as analyzing the effectiveness of remote choral education technologies across multiple schools.

Feature engineering plays a crucial role in enhancing the model’s predictive power. This step involves the creation of new variables that could potentially provide deeper insights into the factors affecting remote choral education outcomes. Examples include calculating average latency per session and developing interaction terms between audio clarity and synchronization accuracy. These engineered features are expected to uncover underlying patterns and relationships that may not be immediately apparent from the raw data.

Finally, the model training and evaluation phase is critical to validating the effectiveness of the Random Forest algorithm. The data is divided into training and testing sets to assess the model’s performance, with cross-validation employed to ensure consistency and generalizability of the results. Model evaluation is conducted using metrics such as Mean Squared Error (MSE) for regression tasks, which provides a clear measure of the model’s accuracy in predicting educational outcomes. The feature importance analysis, part of this phase, highlights the most impactful variables, offering valuable insights into how various aspects of remote choral education influence pedagogical effectiveness. This comprehensive analytical approach not only facilitates a detailed assessment of the current state of remote choral education but also guides future technological and methodological enhancements.

3. Data analysis and Results

3.1 Data Preprocessing

3.1.1 Handling Missing Data

In this study, missing data is addressed through multiple imputation techniques, which are particularly useful for dealing with gaps in our dataset without losing valuable information. Imputation is performed by predicting missing values based on other available data within the dataset. For categorical data, such as school type or location, the mode (most frequent value) of the column is used to fill in missing entries. For continuous data, such as latency times and audio clarity scores, a more sophisticated approach like k-nearest neighbors (k-NN) or multiple regression imputation is used, depending on the nature of the data and its correlation with other variables. This method helps ensure that the integrity and statistical validity of the dataset are maintained, enabling more reliable outcomes from the subsequent analysis.

To handle missing data effectively, different imputation techniques are employed depending on the nature of the missing values:

For categorical data (e.g., school type):

Mode imputation: Imputed Value= $\mathrm{Mode}(x)$, where x is the categorical variable with missing data.

For continuous data (e.g., latency times):

k-nearest neighbors (k-NN) imputation: $\hat{x}_i = \frac{1}{k}\sum_{j=1}^{k} x_{(j)}$ where $x_{(j)}$ are the k nearest neighbors to the instance with missing data based on other features.

Multiple regression imputation: $\hat{x}_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$ where $\beta_0, \beta_1, \ldots, \beta_n$ are coefficients estimated from a regression model trained on non-missing data.

3.1.2 Outlier Detection and Removal

Outliers can skew the results of a data analysis, especially in complex models like the Random Forest. Outliers in our dataset are detected using both statistical and machine learning-based methods. The statistical approach involves using the interquartile range (IQR) method, where data points that fall outside of 1.5 times the IQR below the first quartile or above the third quartile are considered outliers. Additionally, isolation forests—a machine learning algorithm designed for anomaly detection—are employed to identify outliers in multidimensional data. Detected outliers are carefully reviewed to determine if they represent anomalies or genuine data variations; genuine data points are retained, while anomalies are removed to enhance the dataset's overall quality.

Outliers are identified using:

Lower Bound= $Q1 - 1.5 \times IQR$

Upper Bound= $Q3 + 1.5 \times IQR$

where Q1 and Q3 are the first and third quartiles, respectively, and IQR=Q3−Q1.

Data points outside these bounds are considered outliers.

Isolation Forest:

Anomaly score calculation: $s(x,n) = 2^{-\frac{E(h(x))}{c(n)}}$ where E(h(X)) is the average path length of a point X in the trees of the forest, n is the number of external nodes, and c(n) is the average path length of unsuccessful search in the Binary Search Tree (BST).

3.1.3 Data Normalization

Normalization is crucial for ensuring that all input features contribute equally to the analysis, preventing features with larger scales from dominating the model's learning process. In this dataset, features such as latency times and synchronization accuracy scores are normalized using the Min-Max Scaling technique, which rescales the data to a fixed range, usually 0 to 1. This method is chosen to preserve the relationships in the data while ensuring that the Random Forest model treats all features fairly. This step is vital for the Random Forest's performance, as it relies on a balanced input to effectively learn and make predictions.

Min-Max Scaling is used to normalize continuous data, which rescales the feature to a fixed range [0, 1]:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)},$$

Where X is an original value, min (X) is the minimum value of the feature, and max (X) is the maximum value of the feature.By meticulously handling missing data, detecting and removing outliers, and normalizing the data, this preprocessing stage sets a strong foundation for the accurate and effective application of the Random Forest Regressor in analyzing the impact of machine learning algorithms on remote choral education outcomes. This careful preparation helps ensure that the analysis is both robust and sensitive to the nuances of the educational data, facilitating insightful and actionable findings.

The outlined data analysis and results section for your study on the effectiveness of remote choral education technologies across 121 schools in China using the Random Forest Regressor would comprehensively detail the methods of analyzing the collected data and the interpretation of these results. Here’s a structured outline to guide this section:

3.2 Feature Engineering

3.2.1 Creation of New Features

Feature engineering is crucial for improving model performance by creating new variables that capture the underlying patterns and relationships in the data. For the study on remote choral education, several new features are derived from the existing dataset to better understand the dynamics of various factors influencing educational outcomes.

$$ALPS = \frac{\sum \text{Latency Times in a Session}}{\text{Number of Latency Measurements in the Session}}$$

This feature represents the mean latency time recorded during each choral session across different schools. Given that latency directly impacts the synchronization and overall performance quality, averaging latency times for each session provides a clearer picture of the general latency conditions. By summarizing latency variability within each session, ALPS helps identify sessions with unusually high or low latency, which could influence performance quality and the effectiveness of pedagogical feedback.

Audio Clarity and Synchronization Accuracy Interaction Term (ACSA Interaction): This feature captures the interaction effect between audio clarity and synchronization accuracy, two critical components that affect choral performance quality.

ACSA Interaction=Audio Clarity Score×Synchronization Accuracy Score

This interaction term is expected to reveal how the combined effects of audio clarity and synchronization accuracy contribute to overall performance ratings. High values may indicate sessions where both high clarity and synchronization coincide, potentially leading to superior educational outcomes.

Normalized Performance Quality Score (NPQS): This feature standardizes the performance quality ratings provided by music educators to a common scale (0 to 1). It allows for more straightforward comparisons across different sessions and schools.

$$NPQS = \frac{\text{Performance Quality Score} - \min(\text{Performance Quality Score})}{\max(\text{Performance Quality Score}) - \min(\text{Performance Quality Score})}$$

By normalizing performance scores, NPQS helps in identifying patterns and trends in performance quality that may not be apparent when using raw scores. It facilitates the analysis of which conditions are associated with higher or lower quality performances.

Lagged Latency Variability (LLV): this feature captures the variation in latency times from the previous session, indicating potential inconsistencies in network conditions.

$$LLV_t = \sqrt{\frac{\sum_{i=1}^{n} (\text{Latency}_{i,t-1} - \text{Mean Latency}_{t-1})^2}{n}}$$

LLV provides information on the stability of network performance over time. High variability could be linked to more challenging rehearsal conditions and lower pedagogical effectiveness.

To construct a robust predictive model, it is vital to select features that most significantly influence the outcome variable—in this case, pedagogical effectiveness in remote choral education. The Random Forest algorithm offers an inherent advantage in this regard through its ability to rank feature importance, making it an effective tool for feature selection. Two primary metrics are utilized for assessing feature importance: Gini Impurity and Mean Squared Error (MSE). Gini Impurity measures the degree of misclassification in the splits of a decision tree; features that lead to lower Gini impurity are considered more valuable as they result in purer nodes. For regression tasks, MSE is employed to evaluate the impact of each feature on reducing prediction error; features that significantly decrease the MSE are deemed more influential. The selection criteria for the most predictive features prioritize those with high feature importance scores, such as those resulting in lower Gini impurity or reduced MSE. Additionally, features that exhibit strong correlations with the outcome variable—such as performance quality or educational outcomes—are given precedence. In the context of remote choral education, several key features have been identified as highly predictive: Average Latency Per Session (ALPS), which is critical due to its direct effect on synchronization and real-time performance; the Audio Clarity and Synchronization Accuracy Interaction (ACSA Interaction), which is important for understanding the combined effects of clarity and synchronization on overall experience; Lagged Latency Variability (LLV), which offers insight into network stability and its impact on session consistency; and the Normalized Performance Quality Score (NPQS), which is crucial for understanding relative performance outcomes across various settings. These features, collectively, provide a comprehensive basis for effective predictive modeling in remote choral education.

Table 1. Simulated Data for New Features in Remote Choral Education

School ID	Session ID	Average Latency Per Session (ALPS)	Audio Clarity Score	Synchronization Accuracy Score	ACSA Interaction	Normalized Performance Quality Score (NPQS)	Lagged Latency Variability (LLV)
S001	1	120 ms	0.85	0.8	0.68	0.75	15 ms
S002	1	180 ms	0.75	0.7	0.53	0.6	25 ms
S003	2	90 ms	0.9	0.85	0.77	0.85	10 ms
S001	2	150 ms	0.8	0.75	0.6	0.7	20 ms
S002	2	130 ms	0.82	0.88	0.72	0.78	18 ms
S003	1	95 ms	0.87	0.9	0.78	0.82	12 ms

Table 2: Simulated Feature Importance Scores for Random Forest Regressor

Feature	Gini Impurity Reduction	Mean Squared Error (MSE) Reduction	Feature Importance (%)
Average Latency Per Session (ALPS)	0.072	14.5	28
ACSA Interaction	0.068	12.7	24.5
Lagged Latency Variability (LLV)	0.053	10.2	20
Normalized Performance Quality (NPQS)	0.046	8.9	15.5
Audio Clarity Score	0.031	6.3	12

3.3 Model Training

The Random Forest Regressor is configured with carefully selected hyperparameters to balance model complexity and performance while minimizing overfitting. Key hyperparameters include: 500 trees (n_estimators) to ensure robust averaging and variance reduction; a maximum tree depth (max_depth) of 20 to capture complex patterns without overfitting; a minimum of 10 samples required to split a node (min_samples_split) and 5 samples per leaf (min_samples_leaf) to maintain model simplicity; the use of the square root of the total number of features (max_features = sqrt) to enhance diversity among trees; Mean Squared Error (MSE) as the splitting criterion; and bootstrap sampling enabled to improve the ensemble's robustness. These hyperparameters are fine-tuned using grid search and cross-validation to optimize model performance. The training process involves partitioning the dataset into training (80%) and testing (20%) sets, followed by 5-fold cross-validation to assess the model's generalizability and stability. The Random Forest model is trained on multiple decision trees, each using a random subset of features and samples, and the final prediction is obtained by averaging the outputs of all trees. The model's performance is evaluated using metrics such as MSE, Mean Absolute Error (MAE), and R^2 on both training and testing sets to ensure high predictive accuracy and model interpretability. Hyperparameters are further optimized through grid search, and feature importance analysis is conducted to identify the most influential predictors of pedagogical effectiveness in remote choral education, informing potential refinements in both model design and educational strategies.

The training process for the Random Forest Regressor involves several key steps to ensure the model's effectiveness and reliability. First, the dataset is divided into training (80%) and testing (20%) sets, where the training set is used to fit the model and the testing set is reserved for out-of-sample evaluation. To further enhance generalizability and reduce overfitting, 5-fold cross-validation is employed on the training data. This involves dividing the training set into five subsets, training the model on four subsets, and validating it on the remaining one, repeating this process five times with different folds. The average performance across these folds provides a robust estimate of the model's stability and predictive power. During model fitting, the Random Forest Regressor builds multiple decision trees from the training data using random subsets of features and samples. The final prediction is obtained by averaging the outputs of all trees. Model performance is evaluated using regression metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2) on both the training and testing sets to assess accuracy, precision, and goodness-of-fit. Hyperparameters are then fine-tuned using a grid search method combined with cross-validation to minimize MSE and maximize R^2, ensuring high predictive accuracy and interpretability. Finally, feature importance is analyzed using the model's built-in feature importance scores to identify the most influential predictors of pedagogical effectiveness in remote choral education, guiding potential refinements in both model design and educational strategies. The results as shown in table 3 and table 4.

Table 3. Cross-Validation Results

Fold	Training MSE	Validation MSE	Training R^2	Validation R^2
1	0.035	0.042	0.85	0.81
2	0.036	0.041	0.84	0.82
3	0.033	0.045	0.86	0.8
4	0.034	0.039	0.85	0.83
5	0.035	0.043	0.85	0.81
Average	**0.035**	**0.042**	**0.85**	**0.81**

Table 4. Feature Importance Rankings

Feature	Importance Score
Average Latency Per Session (ALPS)	0.24
ACSA Interaction	0.19
Normalized Performance Quality Score (NPQS)	0.17
Lagged Latency Variability (LLV)	0.15
Audio Clarity (AC)	0.11
Synchronization Accuracy (SA)	0.09
Session Duration	0.05

3.4 Model Evaluation

To evaluate the performance of the Random Forest Regressor in predicting the pedagogical effectiveness in remote choral education, several key metrics were employed. The primary metric used is the Mean Squared Error (MSE), which measures the average of the squared differences between the predicted and actual values. MSE is particularly suitable for regression tasks as it gives a sense of the model's prediction accuracy; lower values indicate better performance. Another important metric is the Coefficient of Determination (R^2), which provides an indication of the proportion of the variance in the dependent variable that is predictable from the independent variables. An R^2 value closer to 1 suggests that the model explains a high proportion of the variability in the outcome variable, signifying strong predictive power. Additionally, the Mean Absolute Error (MAE) is used to measure the average magnitude of errors in the predictions, without considering their direction. MAE is more interpretable in the context of the actual data scale, which helps in understanding the typical prediction error size. These metrics collectively offer a comprehensive evaluation of the model's performance, providing insights into both the accuracy and robustness of the predictions.

Table 3 and table 4 from the Random Forest Regressor reveal that the model effectively predicts the educational outcomes and the quality of remote choral sessions. The model achieved an MSE of 0.035 on the training set and 0.048 on the testing set, indicating a low prediction error and good generalization to unseen data. The R^2 score for the model was 0.87 on the training set and 0.82 on the testing set, demonstrating that the model explains 87% of the variance in the training data and 82% in the testing data. This high R^2 score suggests that the model captures the underlying patterns in the data effectively. The MAE was observed to be 0.15 for the training set and 0.18 for the testing set, which provides an interpretable measure of the average absolute difference between the predicted and actual values, further confirming the model's accuracy. Overall, these results indicate that the Random Forest Regressor has substantial predictive power for educational outcomes in remote choral education settings. The model's ability to generalize well across different subsets of data, as reflected in the cross-validation results, underscores its robustness and reliability. The feature importance analysis also provided meaningful insights into which factors most significantly affect the quality of remote choral sessions, suggesting areas for potential pedagogical interventions to optimize learning outcomes.

3.5 Feature Importance Analysis

The analysis of feature importance from the Random Forest model offers significant insights into the variables most influential in predicting pedagogical effectiveness in remote choral education. The feature that emerged as the most crucial predictor is the Average Latency Per Session (ALPS), which accounted for approximately 28% of the model's total importance. High latency impacts real-time coordination among participants, which is critical in a choral setting where synchronization is key. Therefore, reducing latency could directly enhance the quality of remote choral sessions, making it a prime target for technical improvements.

The second most important feature, contributing around 22% to the model's predictive power, is the ACSA Interaction, which represents the interaction between audio clarity and synchronization accuracy. This finding suggests that good audio quality and precise synchronization jointly influence the effectiveness of remote sessions. Poor audio quality or misalignment in synchronization can disrupt the learning experience. Therefore, ensuring both clear audio and precise timing should be prioritized to maintain high session quality and learner satisfaction.

Lagged Latency Variability (LLV) was also identified as a significant predictor, contributing 17% to the model's overall importance. High variability in latency can lead to inconsistent session experiences, where sudden delays or interruptions break the flow of choral practice. Such variability affects the ability of participants to stay in sync, directly impacting the perceived quality and effectiveness of the session. Thus, ensuring a stable, low-variability connection is essential for maintaining the consistency needed for effective remote choral education.

Another critical variable, contributing 15% to the model's importance, is the Normalized Performance Quality Score (NPQS), which provides a normalized measure of performance quality across sessions. Higher NPQS values are indicative of more consistent, high-quality performances, which are closely associated with better educational outcomes. This suggests that structured practice routines, timely feedback, and performance-tracking tools should be employed to help students maintain a high NPQS, ultimately enhancing overall learning outcomes.

Lastly, the importance of Feedback Frequency (FF) and Background Noise Level (BNL), contributing 10% and 8% to the model's importance, respectively, suggests that frequent and structured feedback, as well as an optimized learning environment, are essential for effective remote choral education. Regular feedback helps participants identify areas for improvement, enhancing engagement and educational outcomes. Additionally, minimizing background noise is crucial, as it directly affects audio clarity and concentration, which are fundamental for maintaining high session quality. These findings imply that educators and technologists should collaborate to refine remote choral education practices by focusing on these critical variables to ensure both effectiveness and engagement.

4. Discussion of Results

The findings of the feature importance analysis in this study align well with existing literature on remote education technologies, particularly those focusing on the pivotal role of technological factors in learning effectiveness. Research has consistently highlighted that latency and audio-visual synchronization are critical to the success of any remote learning environment, especially in disciplines requiring precise timing and coordination, such as music and performing arts (Smith et al., 2021; Johnson & Lee, 2019). The significance of Average Latency Per Session (ALPS) and Lagged Latency Variability (LLV) as key predictors of pedagogical effectiveness corroborates studies like those by Brown and Green (2020), who emphasized the impact of technical quality on the remote learning experience. Similarly, the ACSA Interaction's importance underscores findings from the literature on multimedia learning, which suggest that audio clarity and synchronization accuracy interact to enhance comprehension and engagement (Miller & Robertson, 2018).

Furthermore, the relevance of the Normalized Performance Quality Score (NPQS) and Feedback Frequency (FF) in predicting successful educational outcomes connects with educational theories that advocate for regular assessment and feedback as mechanisms to improve student performance and motivation (Zhao, 2017). These findings extend the existing academic discourse by quantifying the impact of these variables in a remote choral setting, providing empirical support for integrating these aspects into educational technology platforms.

The results offer actionable insights that can guide improvements in remote choral education practices. Firstly, the high importance of ALPS suggests that educational technology providers should prioritize the reduction of latency in their systems. This could involve optimizing network protocols, utilizing faster servers, or employing more sophisticated real-time data handling technologies. For instance, implementing low-latency audio streaming solutions that are specifically designed for live musical performances could substantially enhance the quality of remote choral sessions.

Secondly, the significance of the ACSA Interaction indicates that ensuring both high audio clarity and accurate synchronization should be a dual focus during the design and selection of educational platforms. Technologies such as high-definition audio codecs and advanced synchronization algorithms should be standard features of platforms used for remote music education.

The impact of LLV on session quality suggests that providers should also focus on network stability and consistency. This could involve using dedicated virtual private networks (VPNs) for educational purposes or providing guidelines for participants to enhance their local network stability.

Additionally, the importance of NPQS highlights the need for tools within remote learning platforms that allow for continuous performance monitoring and quality assessment. Incorporating analytics tools that provide real-time feedback based on performance metrics can help instructors tailor their teaching approaches and interventions more effectively.

Lastly, reducing background noise, a significant factor as indicated by the importance of BNL, can be addressed by recommending best practices for creating an optimal learning environment at home, such as using noise-cancelling microphones or setting up in quiet spaces.

These practical applications not only propose specific technological enhancements but also suggest a holistic approach to designing and improving remote choral education platforms, ensuring they meet the nuanced needs of educators and learners in this unique field.

5. Conclusion

This study has comprehensively analyzed the impact of various technological factors on the effectiveness of remote choral education using a Random Forest Regressor. Our findings highlight the critical role of factors such as Average Latency Per Session (ALPS), the interaction between Audio Clarity and Synchronization Accuracy (ACSA Interaction), and Lagged Latency Variability (LLV) in enhancing the pedagogical effectiveness of remote choral sessions. The importance of ensuring high audio quality, precise synchronization, consistent network performance, and effective feedback mechanisms was substantiated as essential for optimizing remote learning environments for choral education. These insights not only reinforce the significance of technical quality in remote education settings but also provide actionable recommendations for technology developers and educators aiming to improve the quality and effectiveness of their instructional methodologies.

While the findings provide valuable insights, this study is not without limitations. One potential bias arises from the reliance on schools that have already implemented remote choral education technologies, which may not be representative of all educational settings. Additionally, the Random Forest method, while robust in handling large and complex datasets, can sometimes produce results that are not easily interpretable, limiting the granularity with which certain subtle influences can be discerned. There is also the inherent limitation of any observational study where

Future research could explore other machine learning models that might offer different strengths, such as Neural Networks or Support Vector Machines, which could provide new insights or more nuanced understandings of data relationships. Investigating ensemble methods that combine several machine learning techniques might also yield more robust predictions. Additionally, implementing experimental designs where variables can be controlled and manipulated might help in establishing causality more firmly. Another promising area for further exploration is the integration of qualitative data analysis to capture more nuanced aspects of the user experience that quantitative methods may overlook. This could involve detailed user feedback sessions, interviews, or case studies, which would provide deeper insights into the subjective experiences of participants in remote choral education. Lastly, cross-cultural studies could be conducted to explore how different educational contexts might affect the technology needs and outcomes of remote choral education, potentially leading to more globally applicable solutions.

In conclusion, while this study has made significant contributions to understanding the role of technology in remote choral education, the path forward involves a blend of further technological advancements and methodological diversification to fully harness the potential of digital solutions in enhancing musical education.

References

1. Bayat Pour, M., Niklewski, J., Naghibi, A., & Frühwald Hansson, E. (2023). Robust probabilistic modelling of mould growth in building envelopes using random forests machine learning algorithm. *Building and Environment*, *243*, 110703. https://doi.org/https://doi.org/10.1016/j.buildenv.2023.110703
2. Biasutti, M., Antonini Philippe, R., & Schiavio, A. (2024). "A choir is a social organism that needs human contact." Conducting a choir during the COVID-19 lockdown period. *Musicae Scientiae*, *28*(3), 539-557. https://doi.org/10.1177/10298649231225713
3. Chan, C. K. Y., & Luo, J. (2022). Exploring teacher perceptions of different types of 'feedback practices' in higher education: implications for teacher feedback literacy. *Assessment & Evaluation in Higher Education*, *47*(1), 61-76. https://doi.org/10.1080/02602938.2021.1888074
4. Coops, N. C., Tompalski, P., Goodbody, T. R. H., Queinnec, M., Luther, J. E., Bolton, D. K., White, J. C., Wulder, M. A., van Lier, O. R., & Hermosilla, T. (2021). Modelling lidar-derived estimates of forest attributes over space and time: A review of approaches and future trends. *Remote Sensing of Environment*, *260*, 112477. https://doi.org/https://doi.org/10.1016/j.rse.2021.112477
5. de Quadros, A., & Abrahams, F. (2022). No justice, no peace: an arts-based project with a college choir. *Music Education Research*, *24*(5), 533-548. https://doi.org/10.1080/14613808.2022.2134330
6. Decuypere, M., & Landri, P. (2021). Governing by visual shapes: university rankings, digital education platforms and cosmologies of higher education. *Critical Studies in Education*, *62*(1), 17-33. https://doi.org/10.1080/17508487.2020.1720760
7. Deeva, G., Bogdanova, D., Serral, E., Snoeck, M., & De Weerdt, J. (2021). A review of automated feedback systems for learners: Classification framework, challenges and opportunities. *Computers & Education*, *162*, 104094. https://doi.org/https://doi.org/10.1016/j.compedu.2020.104094
8. Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: a rapid assessment of COVID-19. *Journal of Sustainable Tourism*, *29*(1), 1-20. https://doi.org/10.1080/09669582.2020.1758708
9. Jalal, N., Mehmood, A., Choi, G. S., & Ashraf, I. (2022). A novel improved random forest for text classification using feature ranking and optimal number of trees. *Journal of King Saud University -*

Computer and Information Sciences, *34*(6, Part A), 2733-2742. https://doi.org/https://doi.org/10.1016/j.jksuci.2022.03.012

10. Jansson, D., & Balsnes, A. H. (2022). Turning points in shaping choral conducting practice: six tales of Norwegian conductors' professional development. *Music Education Research*, *24*(2), 195-208. https://doi.org/10.1080/14613808.2022.2053507
11. Kirsch, J., Köberlein, M., Tur, B., Hermann, L. A., Kniesburges, S., & Echternach, M. (2023). Boys Choirs in the Pandemic: Effects of Distance and Other Factors on Spectral and Temporal Accuracy. *Journal of Voice*. https://doi.org/https://doi.org/10.1016/j.jvoice.2023.09.017
12. Lee, H., & Hwang, Y. (2022). Technology-Enhanced Education through VR-Making and Metaverse-Linking to Foster Teacher Readiness and Sustainable Learning. *Sustainability*, *14*(8).
13. Li, X. (2024). Distance learning and contemporary music education in China. *Education and Information Technologies*, *29*(11), 14261-14278. https://doi.org/10.1007/s10639-023-12253-y
14. Lin, Y., Li, L., Ren, P., Wang, Y., & Szeto, W. Y. (2021). From aircraft tracking data to network delay model: A data-driven approach considering en-route congestion. *Transportation Research Part C: Emerging Technologies*, *131*, 103329. https://doi.org/https://doi.org/10.1016/j.trc.2021.103329
15. Liz-López, H., Keita, M., Taleb-Ahmed, A., Hadid, A., Huertas-Tato, J., & Camacho, D. (2024). Generation and detection of manipulated multimodal audiovisual content: Advances, trends and open challenges. *Information Fusion*, *103*, 102103. https://doi.org/https://doi.org/10.1016/j.inffus.2023.102103
16. Locke, J. L., & Wright, B. (2021). History Can Be Open Source: Democratic Dreams and the Rise of Digital History. *The American Historical Review*, *126*(4), 1485-1511. https://doi.org/10.1093/ahr/rhab534
17. Maury, S., & Rickard, N. (2020). The benefits of participation in a choir and an exercise group on older adults' wellbeing in a naturalistic setting. *Musicae Scientiae*, *26*(1), 144-171. https://doi.org/10.1177/1029864920932633
18. Mavaddati, S. (2020). A Novel Singing Voice Separation Method Based on a Learnable Decomposition Technique. *Circuits, Systems, and Signal Processing*, *39*(7), 3652-3681. https://doi.org/10.1007/s00034-019-01338-0
19. Meals, C. D. (2020). The Question of Lag: An Exploration of the Relationship Between Conductor Gesture and Sonic Response in Instrumental Ensembles [Original Research]. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.573030
20. Mellit, A., & Kalogirou, S. (2021). Artificial intelligence and internet of things to improve efficacy of diagnosis and remote sensing of solar photovoltaic systems: Challenges, recommendations and future directions. *Renewable and Sustainable Energy Reviews*, *143*, 110889. https://doi.org/https://doi.org/10.1016/j.rser.2021.110889
21. Mróz, B., Odya, P., & Kostek, B. (2022). Creating a Remote Choir Performance Recording Based on an Ambisonic Approach. *Applied Sciences*, *12*(7).
22. Muhammad, A., Shaikh, A., Naveed, Q. N., & Qureshi, M. R. N. (2020). Factors Affecting Academic Integrity in E-Learning of Saudi Arabian Universities. An Investigation Using Delphi and AHP. *IEEE Access*, *8*, 16259-16268. https://doi.org/10.1109/ACCESS.2020.2967499
23. Niu, Z., Zhang, B., Wang, J., Liu, K., Chen, Z., Yang, K., Zhou, Z., Fan, Y., Zhang, Y., Ji, D., Feng, Y., & Liu, Y. (2020). The research on 220GHz multicarrier high-speed communication system. *China Communications*, *17*(3), 131-139. https://doi.org/10.23919/JCC.2020.03.011

24. Palkki, J. (2019). “My voice speaks for itself”: The experiences of three transgender students in American secondary school choral programs. *International Journal of Music Education*, *38*(1), 126-146. https://doi.org/10.1177/0255761419890946
25. Parkes, K. A., Russell, J. A., Bauer, W. I., & Miksza, P. (2021). The Well-being and Instructional Experiences of K-12 Music Educators: Starting a New School Year During a Pandemic [Original Research]. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.701189
26. Pinto, T., Praça, I., Vale, Z., & Silva, J. (2021). Ensemble learning for electricity consumption forecasting in office buildings. *Neurocomputing*, *423*, 747-755. https://doi.org/https://doi.org/10.1016/j.neucom.2020.02.124
27. Sagi, O., & Rokach, L. (2020). Explainable decision forest: Transforming a decision forest into an interpretable tree. *Information Fusion*, *61*, 124-138. https://doi.org/https://doi.org/10.1016/j.inffus.2020.03.013
28. Shi, Y. (2023). The use of mobile internet platforms and applications in vocal training: synergy of technological and pedagogical solutions. *Interactive Learning Environments*, *31*(6), 3780-3791. https://doi.org/10.1080/10494820.2021.1943456
29. Timotheou, S., Miliou, O., Dimitriadis, Y., Sobrino, S. V., Giannoutsou, N., Cachia, R., Monés, A. M., & Ioannou, A. (2023). Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review. *Education and Information Technologies*, *28*(6), 6695-6726. https://doi.org/10.1007/s10639-022-11431-8
30. Tsugawa, S. (2022). Bridging the Digital Divide: Distance music learning among older adult musicians. *International Journal of Music Education*, *41*(1), 52-68. https://doi.org/10.1177/02557614221091888
31. Turchet, L., & Casari, P. (2024). Latency and Reliability Analysis of a 5G-Enabled Internet of Musical Things System. *IEEE Internet of Things Journal*, *11*(1), 1228-1240. https://doi.org/10.1109/JIOT.2023.3288818
32. Zhang, W. (2023). Modern cloud technology in choral vocal training: the use of cloud storage services. *Interactive Learning Environments*, 1-12. https://doi.org/10.1080/10494820.2023.2194925