International Journal of Information and Communication Technology Education
Volume 20 • Issue 1

Optimization of Piano Performance Teaching Mode Using Network Big Data Analysis Technology

Xiang Wei, Taiyuan University of Technology, China
Shuping Sun, Jining University, China*

ABSTRACT

To effectively avoid subjective bias in manual evaluation. This article proposes a MIDI piano teaching performance evaluation method based on bidirectional LSTM. This method utilizes a three-layer bidirectional LSTM neural network mechanism to make it easier for the model to capture useful information. In addition, the Spark clustering training model is constructed using the deeplearning4j deep learning framework, and the model parameters are adjusted through the UI dependency relationships provided by deeplearning4j to improve work efficiency. The experimental results verified the superiority of the bidirectional LSTM model. The methods provided in this article can improve students' independent practical abilities and reduce the pressure on teachers during the teaching process. These measures can promote the development of music education, improve students' music literacy and learning skills, and make positive contributions to the music education industry.

KEYWORDS

big data (BD), network information, piano performance, piano teaching

INTRODUCTION

Music performance is an art form, and different music performances may require different evaluation methods. For music educators, choosing an effective and reliable evaluation method is a key issue. Research has shown that computer evaluation methods can effectively avoid subjective biases in manual evaluations. Therefore, this article proposes a MIDI piano teaching performance evaluation method based on bidirectional long short-term memory (LSTM) networks to improve the reliability and effectiveness of evaluation. This method utilizes a three-layer bidirectional LSTM neural network mechanism to make it easier for the model to capture useful information. In addition, the Spark clustering training model is constructed using the deeplearning4j deep learning framework, and the model parameters are adjusted through the UI dependencies provided by deeplearning4j to improve work efficiency. The experimental results verified the superiority of the bidirectional LSTM model. The methods provided in this article can enhance students' independent practical abilities and reduce the pressure on teachers during the teaching process. These measures can promote the development of

DOI: 10.4018/IJICTE.341266 *Corresponding Author

music education, improve students' music literacy and learning skills, and make positive contributions to the music education industry.

LITERATURE REVIEW

Music education is formed naturally with the development of human society, culture, and physiology. This is one of the important signs of human evolution. Piano teaching is one of the important basic courses for a music education major, and it is also a required course (Li, 2020). Every step of music education professional development is accompanied by the development of piano curriculum. After the Opium War, under the influence of western music culture, with the establishment of new schools, China's music education began to improve (Ranjan, 2014). Some women's schools in Shanghai set up special piano courses, and music teachers gradually learned to spread music cultural knowledge in the form of piano performance and to carry out music teaching activities in the classroom (Sun, 2021). Later, music associations, such as the Peking University Music Research Association, emerged and gradually became the earliest music education institutions in China (Wang, 2022). Subsequently, Xiao Youmei officially established the China Conservatory of Music, marking the official establishment of China's modern professional music colleges and opening a new chapter in Chinese piano teaching (Xia, 2020). With the continuous progress and development of society, there is an increasing demand for compound talents (Yan, 2022).

Some major music colleges are still cultivating piano postgraduates in the direction of music education, and gradually develop master training centers (Shuo & Xiao, 2019). They have also cultivated many excellent piano professionals for the country (Wang et al., 2020).

Since the 1990s, information technology (IT) has been constantly updated and iterated. From the information highway to the information explosion, big data (BD), cloud computing, Internet of things, and Internet +, IT has gradually become a necessity in people's daily life (Yu & Yang, 2022). Using IT to improve people's work and lifestyle and trying to apply internet thinking to economic, educational, and social construction are important issues that have attracted the attention of the state and society in recent years (Li & Robin, 2019). The so-called network thinking in piano teaching is the thinking mode and teaching method of "network + piano teaching" (Huang, 2023). Specifically, in terms of educational technology, network thinking should make comprehensive use of the advantages of IT and the Internet and carry out technological innovation in combination with the characteristics of piano teaching and the Internet. In terms of communication, network should make use of the media advantages brought by the Internet to combine traditional communication theories with emerging media to form a new way of knowledge communication (Livingstone, 2015). This reflects the benefits that the Internet can provide and how teachers can improve teaching methods and content under the background of modern education (Yao, 2023). Education under internet thinking does not mean subverting the traditional education model (Jing, 2023). At the level of educational theory, education under internet thinking is the integration of the Internet and education. When it is implemented in the teaching process, it is the reflection of teachers and learners on their own roles. On this basis, it reconstructs the meaning of those roles to improve the teaching process (Sha & Zhang, 2023).

The performance evaluation function of piano education software available on the market is based mainly on rules (Fang & Liang, 2023). This method cannot evaluate the quality of music from the perspective of musical expressiveness and continuity (Xie, 2023). With the continuous enhancement of computing power and the improvement of deep learning, more and more researchers use high-performance distributed clustering and deep learning methods to solve classification problems (Yu et al., 2023). The parallelization method is used to improve classification efficiency (Luo et al., 2023). The emergence of BD makes efficient parallelization possible (Gao & Li, 2023). Google laid the foundation for BD technology after successively publishing three papers on MapReduce and Bigtable (Zhong, 2023). Hadoop initially implemented MapReduce, and later became a multi-faceted Hadoop ecosystem (Zhong, 2023). Spark is a member of the Hadoop ecosystem and a computing framework

(Zhao et al., 2023). Its computing speed is 100 times faster than MapReduce in Hadoop, but it still needs to rely on the distributed file storage system HDFS (Hadoop distributed file system) and the resource management system Yan in Hadoop. Spark has become the first choice for a BD parallel processing engine since its development (Cao, 2023). Deeplearning4j is another member of Hadoop ecosystem. It is a deep learning framework based on data parallelization theory and is compatible with Spark. Based on a Spark and Deeplearning4j framework, a bidirectional long-short-term-memory (LSTM) MIDI piano performance evaluation scheme is proposed. The scheme monitors the model training process in real time through the user interface to facilitate real-time adjustment of parameters. It has certain practical value for improving the performance evaluation efficiency of piano lessons (Zhang, 2023).

RELATED MATERIALS AND METHODS

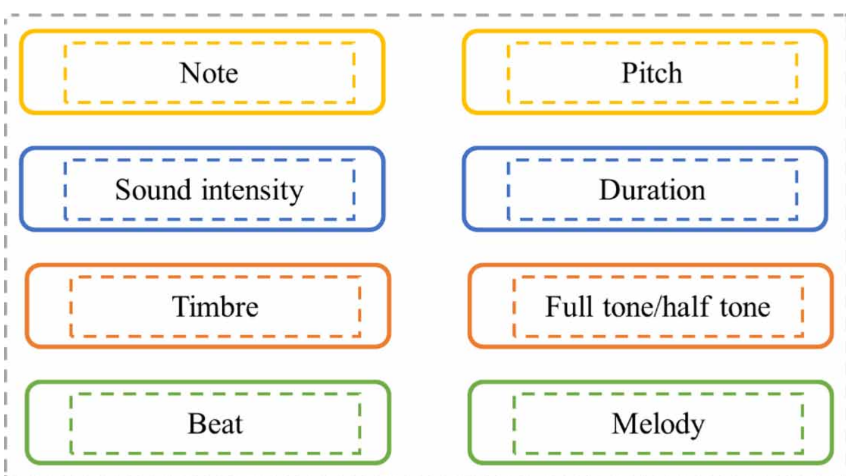
Basis of Music Theory

Sound comes from the vibration of an object. Regular and resonant sounds are called music. In daily life, voice can not only allow communication between people, but also fosters the spiritual development of human civilization. Music represents a special form of sound, and its origins can be traced back to the origin of human beings. The development of human beings must be accompanied by the development of music. Music has evolved into a substantial professional field that has developed a theoretical approach to music, known as music theory.

As illustrated in Figure 1, music theory mainly includes 8 parts:

- Note: a special symbol for recording different long and short notes. It is the smallest unit of music.
- Pitch: the loudness of a sound. The pitch depends on the vibration value of the sound part in a specific time. The higher the value, the louder the pitch.
- Sound intensity: the intensity of sound. Tone intensity is determined by the amplitude of the sound producing part. The greater the amplitude, the greater the intensity.
- Duration: the duration of sound and a measure of sound maintenance ability. It depends on the duration of vibration of the sound part.
- Timbre: the basic attribute of an object. Different substances have different timbres in the same vibrational state. It is mainly reflected in that different timbres give people different feelings.

Figure 1. Basic knowledge structure of music theory



- Full tone/half tone: in the twelve melodies, the relationship between two adjacent tones is called half tone, and the relationship between two half tones is called full tone. Specifically, the relationship between the two adjacent keys on the 88-key piano keyboard is half tone, and the relationship between the two keys separated by the key distance is full tone.
- Beat: the beat can be understood by splitting the segment and the beat. In the process of music production, music is usually divided into isochronous basic units, called beats. The time value of the beat is the time value of the note. The strong beat and the weak beat are combined according to specific rules to form a beat. The beat is a measure of rhythm.
- Melody: from a physical point of view, music melody is a sequence of pitch and intensity in the time dimension, and it is the focus of emotional expression. Melody is divided into main melody and accompaniment melody. The main melody accompanies the whole music process and is the main framework of music expression. The accompaniment melody is a means to strengthen the musical expression.

Music Signal Storage Format

With the development of the computer, the format of music storage is also changing. Music storage formats can be divided into two categories. The first category is analog storage, which includes mainly recording and tape. The second category is digital storage, mainly including MP3, wma, WAV, and MIDI. The disadvantages of analog storage are as follows:

- The storage carrier is not easy to save for a long time, resulting in music distortion.
- The storage operators have many specifications and are incompatible with each other.

Digital storage can overcome the above shortcomings, so this format has become the mainstream of music storage. The dataset format is MIDI format here. MIDI format has the advantage of small storage capacity and can completely express the internal structure of music. It is easier to analyze music under the interpretation of computers.

According to standard MIDI1.0, a standard MIDI file SMF is composed of multiple file blocks, which are composed of flag characters (4 bytes), data string length (4 bytes), and data (6 bytes). For each MIDI file, SMF only contains one header chunk file and one or more track chunks.

SMF=<header chunk>+<track chunk>+(<track chunk>...) (1)

Structure of Header Chunk

The header chunk structure is composed of three parts: mark character, data string length, and header chunk data (Figure 2).

00 00 represents the single-track format, 00 01 represents the multi-track synchronous format, 00 10 represents the multi-track asynchronous format, and nn nn determines the number of tracks; only the single track is involved here (00 01). Dd represents the basic format of time. When the highest bit is 0, it indicates the scale. The last three bits represent the number of divisions in the rounding. When the highest bit is 1, it indicates SMPTE format timing, and the last three bits indicate SMPTE frames per second. Most MIDI files currently use the former format.

Structure of Track Chunk

The track chunk structure is composed of three parts: mark character, data string length, and track chunk data (Figure 3).

Figure 2. Structure diagram of head chunk

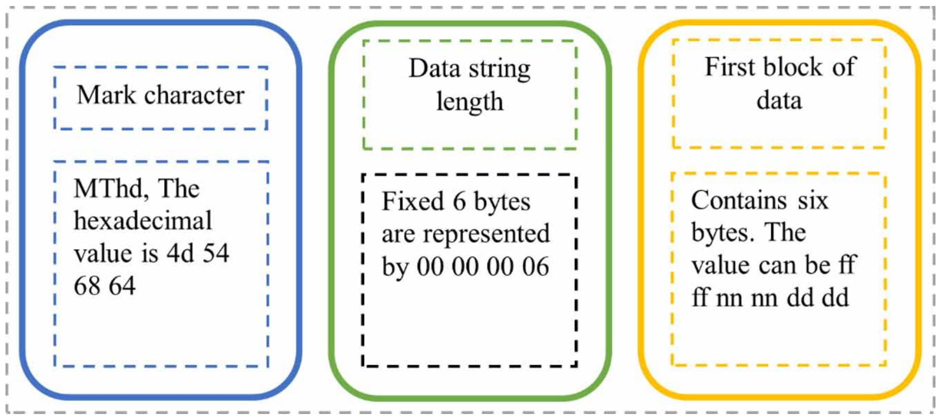
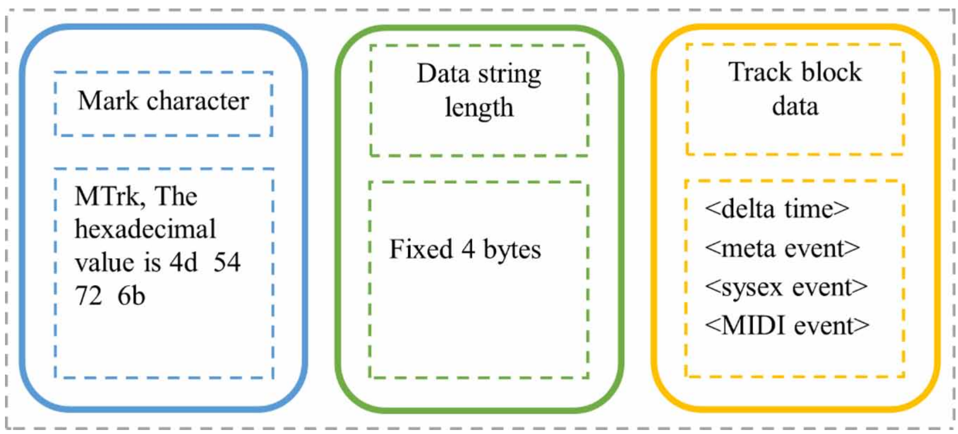


Figure 3. Structure diagram of track chunk



In the track chunk data, <delta time> indicates the time interval between this event and the previous event. dd dd in the header chunk file determines the time format, generally indicating ticks timing. Because the time duration is not fixed, it is represented by dynamic bytes. For example, 380ticks is represented as 128*2+124, binary is 10000010 01111100, and hexadecimal is 827C.

<meta event> indicates the channel, lyrics, labels, and other information of MIDI files. Although the event does not contain the description information of music features, it plays an important role in the preliminary screening of music. The basic format of <meta event> is FF XX YY ZZ. FF is the fixed symbol, XX is the description of function type, YY is the byte length of data, and ZZ is the description of data bytes. It should be noted that when the last byte after FF is 03, this <meta event> event indicates the music name, and this information is usually only contained in the first track chunk. <sysex event> is mostly used for real-time equipment controlling, which is a fixed format here.

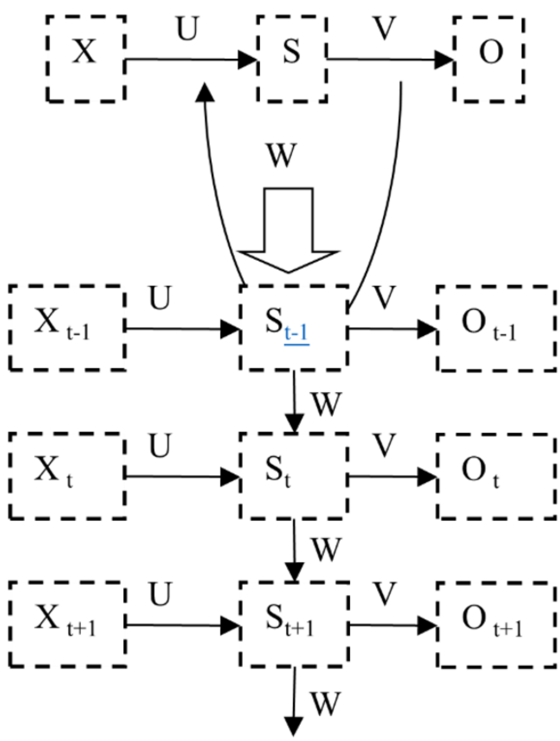
A MIDI event usually consists of a status byte and several data bytes. The last four bits of the status byte represent the channel information, and the first four bits represent the event type. Here it mainly masters the bill opening event and bill closing event, that is, the first four digits 1001 represent the bill opening event, and the first four digits 1000 represent the bill closing event. The next two characters indicate the number of notes and the strength of the keys. There are usually two ways to describe notes. The first way consists of on events and off events. The second consists of two public activities. When the second on event is pressed with zero force, the status byte can be omitted. In fact, since the second method is 1 byte less than the first method, it is usually used to write <MIDI event> files.

Recurrent Neural Network (RNN)

A recurrent neural network (RNN) is an important type of artificial neural network, which is particularly suitable for modeling and analyzing serialized data (such as time series, text, etc.). Unlike other neural networks, RNN has a feedback loop that can use previous state information to influence the current output when processing sequence data, thus achieving more complex models. The basic structure of RNN includes input layer, hidden layer, and output layer. At each time step, input data is fed into the input layer and then entered into the hidden layer through a weight matrix for calculation. The hidden layer integrates the current input data with the previous state information, generates a new state information, and transmits it to the next time. Finally, the state information of the last hidden layer will be fed into the output layer for calculation, resulting in the model's prediction results. The advantage of RNN is that it can handle sequential data and has no limit on the length of input data. This makes RNN widely applicable in fields such as natural language processing, speech recognition, and machine translation. However, RNN also has some drawbacks, such as the tendency to encounter problems such as vanishing gradients and exploding gradients, making it difficult to train. Meanwhile, RNN can only utilize the state information from the previous moment, and cannot utilize earlier historical information, leading to long-term dependency issues. To address these issues, improved models based on RNN have emerged, such as long short-term memory networks (LSTM) and gated recurrent units (GRU).

The diagram of the RNN is shown in Figure 4.

Figure 4. Recurrent neural network



Compared with the traditional neural network, the RNN structure has one more feedback input, which contains matrix *W* and the output information of the previous time. When the time is *t*, the neural node receives not only the input *x* at t*,* but also the weight and *s*t-1 at t-*1*.

The expression of output value o_t at *t* is as follows:

$$o_t=g(Vs_t) \tag{2}$$

where *g* indicates the activation function, *V* indicates the weight, and s_t represents the sum of weights at *t*.

The expression of sum of weights s_t at *t* is as below:

$$s_t=f(Ux_t+Ws_{t-1}) \tag{3}$$

where *f* represents the activation function, *U* represents the weight, *W* means the state transition weight matrix from the previous time to the next time, x_t means input at *t*, and s_{t-1} means the sum of weights at *t-1*. Infinite substitution of Equation (3) into Equation (2) obtains o_t expansion Equation (4); this shows that RNN has strong memory function for sequence information.

$$o_t=Vf(Ux_t+Wf(Ux_{t-1}+Wf(Ux_{t-2}+Wf(Ux_{t-3}+\cdots)))) \tag{4}$$

Forward Propagation Algorithm (FPA)

The FPA is an algorithm that realizes the function of data along the forward propagation. The RNN forward propagation algorithm can handle serialized data, such as time series, text, and so on. It can use historical information to predict future states or outputs. The parameters in the RNN model are shared at each time step, which means that the model has a relatively small number of parameters, which can better handle long sequence data and reduce the computational resources required for training the model. The RNN model has a feedback loop structure, which can transmit information through the hidden state of time steps, thereby capturing the time dependency relationship in sequence data. This enables the RNN model to demonstrate good performance in dealing with long-term dependency issues. This section introduces the process of forward propagation of RNN and introduces the calculation of each parameter of FPA according to the model introduced in the previous section.

At *t*, the hidden state s_t is as follows:

$$s_t=\tilde{A}(Ux_t+Ws_{t-1}+b) \tag{5}$$

where A means activation function, generally tanh; *b* means bias.

At *t*, the output o_t is as below:

$$o_t=Vs_t+c \tag{6}$$

where *c* means bias.

At *t*, the predicted output $\bar{y}_t$ is as follows:

$$\overline{y_t}=\tilde{A}(o_t) \tag{7}$$

Back Propagation Algorithm

The RNN backpropagation algorithm can handle serialized data, such as time series, text, and so on. It can use historical information to predict future states or outputs. Unlike traditional feedforward neural networks, the RNN model has a feedback loop structure, making its gradient calculation more complex. The RNN backpropagation algorithm adopts gradient backpropagation technology, which updates network parameters by backpropagating errors from the output layer to the hidden layer. The long-term dependency problem in RNN models is a common challenge. The RNN backpropagation algorithm can backpropagate errors from the output layer to the hidden layer through the BPTT algorithm and calculate the gradient of each time step based on the backpropagation error, thereby solving the long-term dependency problem. Based on RNN FPA, the RNN back propagation algorithm can be obtained. The process of the RNN back propagation algorithm is to calculate the gradient of each parameter of the model (i.e., U, W, V, b, c) through the transfer property of gradient descent error.

It is assumed that the loss function is the cross-entropy loss function L, the output activation function is the softmax function, and the activation function in the hidden layer is the tanh function.

The total loss function is calculated as:

$$L=\sum_{t=1}^{T} L_t \tag{8}$$

The gradients of V and c are as below:

$$\frac{\partial L}{\partial c}=\sum_{t=1}^{T}\frac{\partial L_t}{\partial L}=\sum_{t=1}^{T}\overline{y_t}-y_t \tag{9}$$

$$\frac{\partial L}{\partial V}=\sum_{t=1}^{T}\frac{\partial L_t}{\partial V}=\sum_{t=1}^{T}(\overline{y_t}-y_t)(s_t)^T \tag{10}$$

The gradient δ of the hidden state at t is defined as follows:

$$\delta_t=\frac{\partial L}{\partial s_t} \tag{11}$$

The recurrence of δ_t from δ_{t+1} is performed as:

$$\delta_t=(\frac{\partial o_t}{\partial s_t})^T\frac{\partial L}{\partial o_t}+(\frac{\partial o_{t+1}}{\partial s_t})^T\frac{\partial T}{\partial o_{t+1}}=V^T(\overline{y}_t-y_t)+W^T diag(1-(s_{t+1})^2)\delta_{t+1} \tag{12}$$

Because it is at the end of the sequence, the expression of δ_T is as follows:

$$\delta_T=(\frac{\partial o_T}{\partial s_T})^T\frac{\partial L}{\partial o_T}=V^T(\overline{y_t}-y_t) \tag{13}$$

The gradient expression of W is given in Equation (14):

$$\frac{\partial L}{\partial W}=\sum_{t=1}^{T}diag(1-(s_t)^2)\delta_t(h_{t-1})^T \tag{14}$$

The gradient expression of U is given in Equation (15):

$$\frac{\partial L}{\partial U}=\sum_{t=1}^{T}diag(1-(s_t)^2)\delta_t(x_t)^T \tag{15}$$

The gradient expression of b is given in Equation (16):

$$\frac{\partial L}{\partial b}=\sum_{t=1}^{T}diag(1-(s_t)^2)\delta_t \tag{16}$$

From the above derivation, the gradients of W, U, and b are related to \mathrm{s}\uMathrm{t} at the same time. Because \mathrm{s}\uuMathrm{t} is a TANH function with a value between 0 and 1, when the value of t is too large, that is, when the length of the sequence data processed by RNN is too long, the gradients of W, U, and b are infinitely close to zero, resulting in the disappearance of gradients. A gradient vanishing problem exists not only in RNN network, but also in other depth network models. At present, there are several commonly used methods to solve the gradient disappearance problem, such as using the ReLu activation function to replace the original activation function. In the field of Natural Language Processing (NLP), LSTM or LSTM deformation is usually used to replace the traditional RNN.

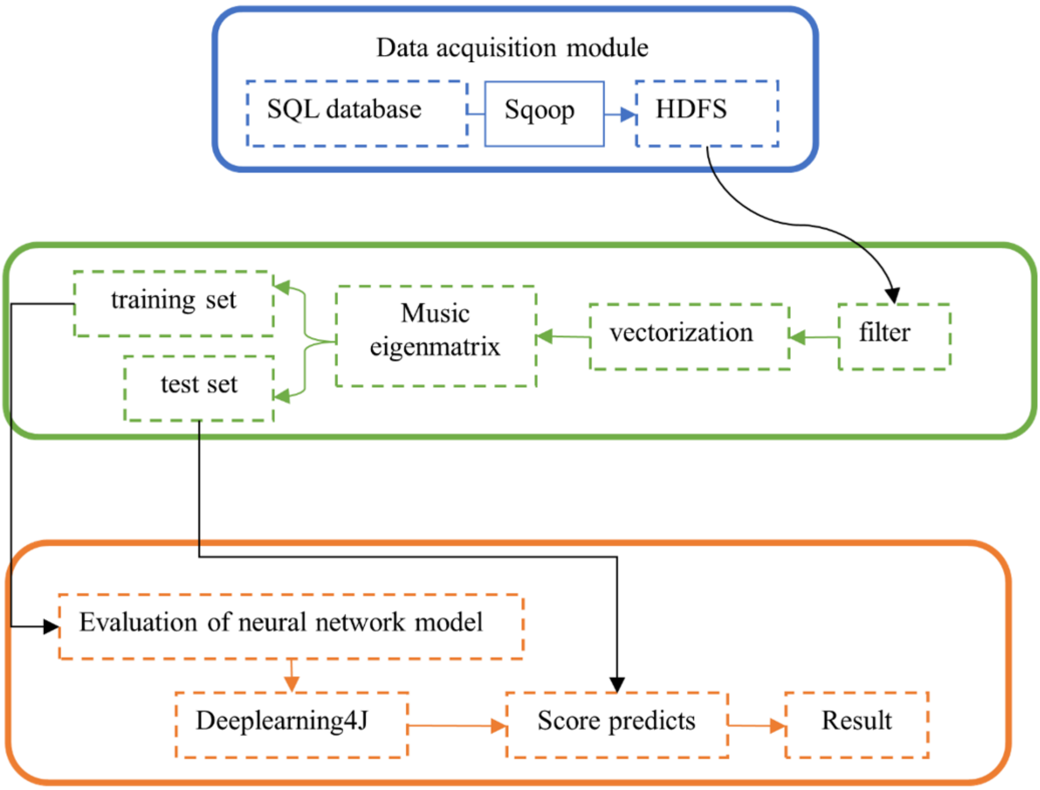
Framework Design of Performance Evaluation Model in MIDI Piano Teaching

The MIDI piano performance evaluation neural network model mentioned in this method is implemented on a Spark deeplearning4j framework. The bidirectional LSTM neural network model with attention mechanism can achieve efficient and accurate MIDI Piano performance evaluation. The model framework is composed of data acquisition, data pre-processing, and music evaluation and classification modules (Figure 5).

In the data acquisition module, the sqoop tool is used to migrate data to the distributed data storage system HDFS. In the data pre-processing module, it is necessary to filter the original data that is not suitable for training, convert the original data into the input matrix form suitable for neural network model training, and divide it into training, verification, and test sets. In the music evaluation and classification module, it is necessary to establish Spark Yan clustering, establish a neural network model on the distributed framework, send the pre-processing data into the model training, and adjust the model parameters in real time through the UI interface provided by deeplearning4j to obtain the model parameters with good evaluation effect.

After constructing the corresponding scoring system based on the above BD analysis and algorithm, the grade evaluation method is applied to assist piano teaching and realize the optimization of piano teaching mode. Two music teachers were selected as the auxiliary evaluation, and five music majors were selected as the evaluation subjects. Each subject received e-mail before participating in the test, including signing the informed consent form and a detailed description of the research purpose and problems.

Figure 5. Framework of performance evaluation model in MIDI piano teaching



RESULTS AND ANALYSIS

Analysis of Model Results

Data Acquisition Implementation

The model evaluates MIDI piano music, and the acquisition of MIDI piano data is the premise of training the model. The initial MIDI piano data is in the MySQL database. The piano music works are shown in Table 1. The SoutID field uses (1,5) intervals to represent the five grades of MIDI piano music evaluation and classification, and FileAdrr stores the relative path address of MIDI files.

Table 1. Piano music works

Name	Type	Other
PainoMusicID	Decimal (10)	VarChar (200)
PainoMusicName	VarChar(50)	N/A
StorageTime	Datetime	N/A
Size	Float	N/A
SoutID	Decimal (5)	N/A
FileAdrr	VarChar (200)	N/A

The number of partitions for HDFS is set to three, and the block of 128 MB_SUCCESS is an index file. Because MIDI files are small files and the size of a MIDI music song ranges from hundreds of KB to 2MB, HDFS is not suitable for storing these small files directly. MIDI piano music files are packaged as ZIP files and migrated to HDFS.

Music Analysis Implementation

The music analysis module is the core module, which realizes the function of the MIDI music score. Using the bidirectional LSTM network model with attention mechanism combined with the advantages of this model can not only capture the bidirectional timing information of MIDI music, but also label the important information with greater weight. Since each subnet model is independent of the others, this section only introduces the deployment and training process of the single subnet model based on the deeplearning4j deep learning framework in Spark.

The music analysis module can use the static internal class generator in the neural net configuration class to declare the super parameters of the model, such as the learning rate and gradient descent method. The layer method is used to define the neural network parameters of each layer and the P value of dropout. The parameter object is obtained by the generator, and the model parameter object is created by the multi-layer network class. By adding UI server dependencies and binding listeners, it can directly view the classifier detection graph in the training UI visual interface (Figure 6).

The corresponding layer is clicked to view the specific parameters on the right side of the figure. For example, the output layer is RnnOutput (Lay3), the number of input nodes is 88, the number of output nodes is 5, the activation function is softmax function, and the tuner is Adam. In the training process of the neural network model, the change process of loss function can be checked in real time. When divergence or violent fluctuation occurs, the training process can be stopped and the parameters can be adjusted. In addition, it can check the memory usage of the JVM system through the TrainingUI (Figure 6C).

Selection of the Number of Hidden Layers and Nodes

As illustrated in Figure 7, the pitch matrix is an 88-dimensional matrix, so the number of nodes is 88. Firstly, the single-layer bidirectional LSTM model is studied. The results show that when the number of nodes is 88, 176, 352, 704, 1,408, and 2,816, the L value also changes. The bidirectional LSTM model of double-layer hidden layer is further explored, and it is determined that the number of nodes in the first layer is 352 and the number of nodes in the second layer is 88, 176, 352, 704, 1,408, and 2,816 respectively. L value decreases. According to the bidirectional LSTM model of three hidden layers, it is determined that the number of nodes in the first layer is 352, the number of nodes in the second layer is 176, and the number of nodes in the third layer is 88, 176, 352, 704, 1,408, and 2,816. The change of L value is more obvious. Finally, the bidirectional LSTM model with four hidden layers is studied. When the number of nodes in the first layer is 352, the number of nodes in the second layer is 176, the number of nodes in the third layer is 88, and the number of nodes in the fourth layer is 88, 176, 352, 704, 1,408, and 2,816, respectively. The change of L value is similar to that in the third layer. To sum up, it can be determined that the reasonable structure of the music evaluation network is as follows: the number of hidden layers is 3 and the number of nodes in each layer is 352, 176, and 88.

Analysis of Music Classification Results

When binary classification problems are studied, precision, recall, and K value are commonly used parameters. The true value is a positive class; if it is detected as a positive class, it is called TP class and if it is detected as a negative class, it is called FN class. If it is detected as a positive class, it is called an FP class; and if it is detected as a negative class, it is called a TN class. The expressions of precision, recall, and K value are as follows:

Figure 6. Relevant logic change diagram of training model

Note. A is the detail diagram of the UI interface, B is the change monitoring diagram of loss function, and C is the memory utilization of the JVM system.

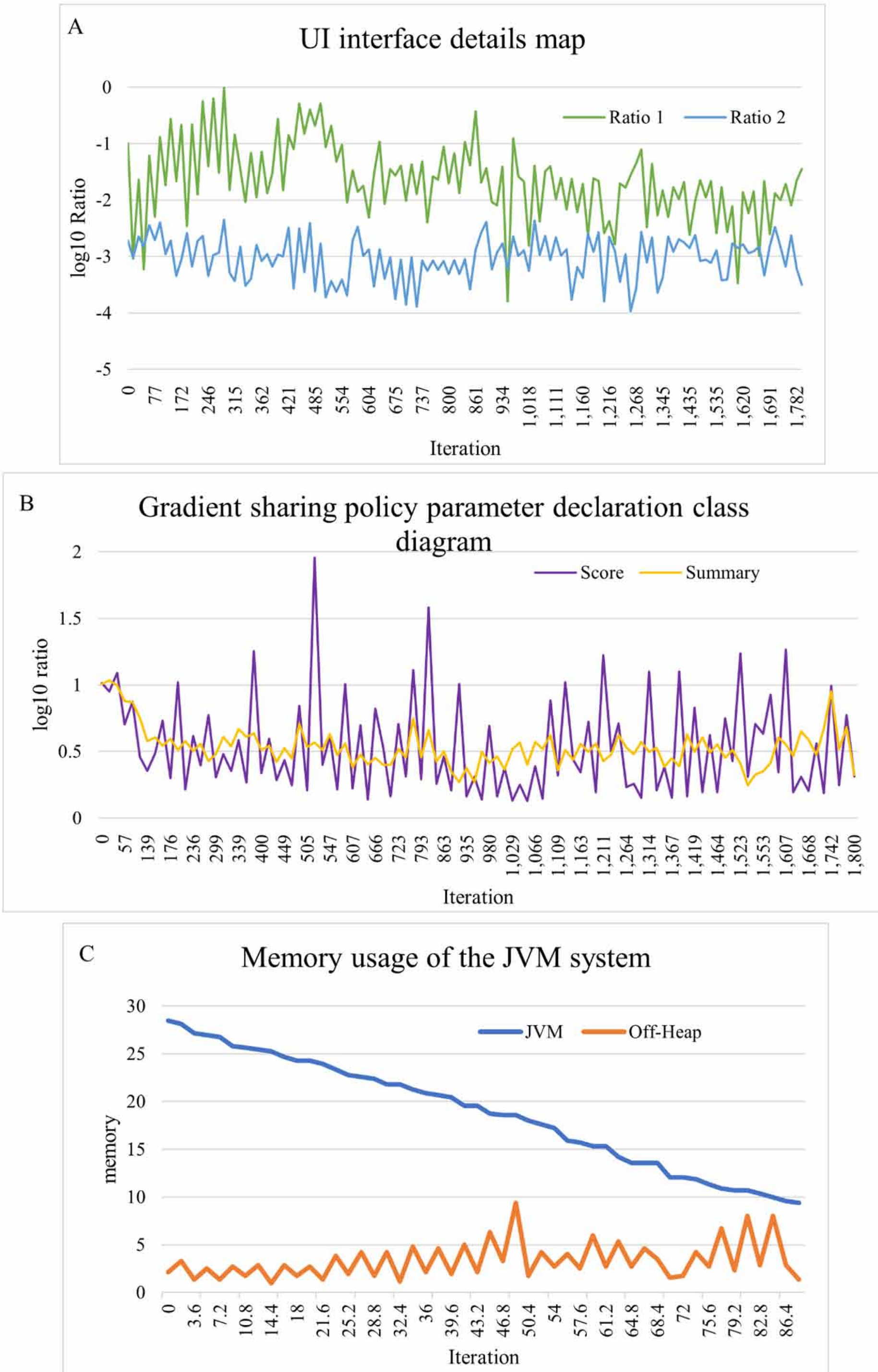
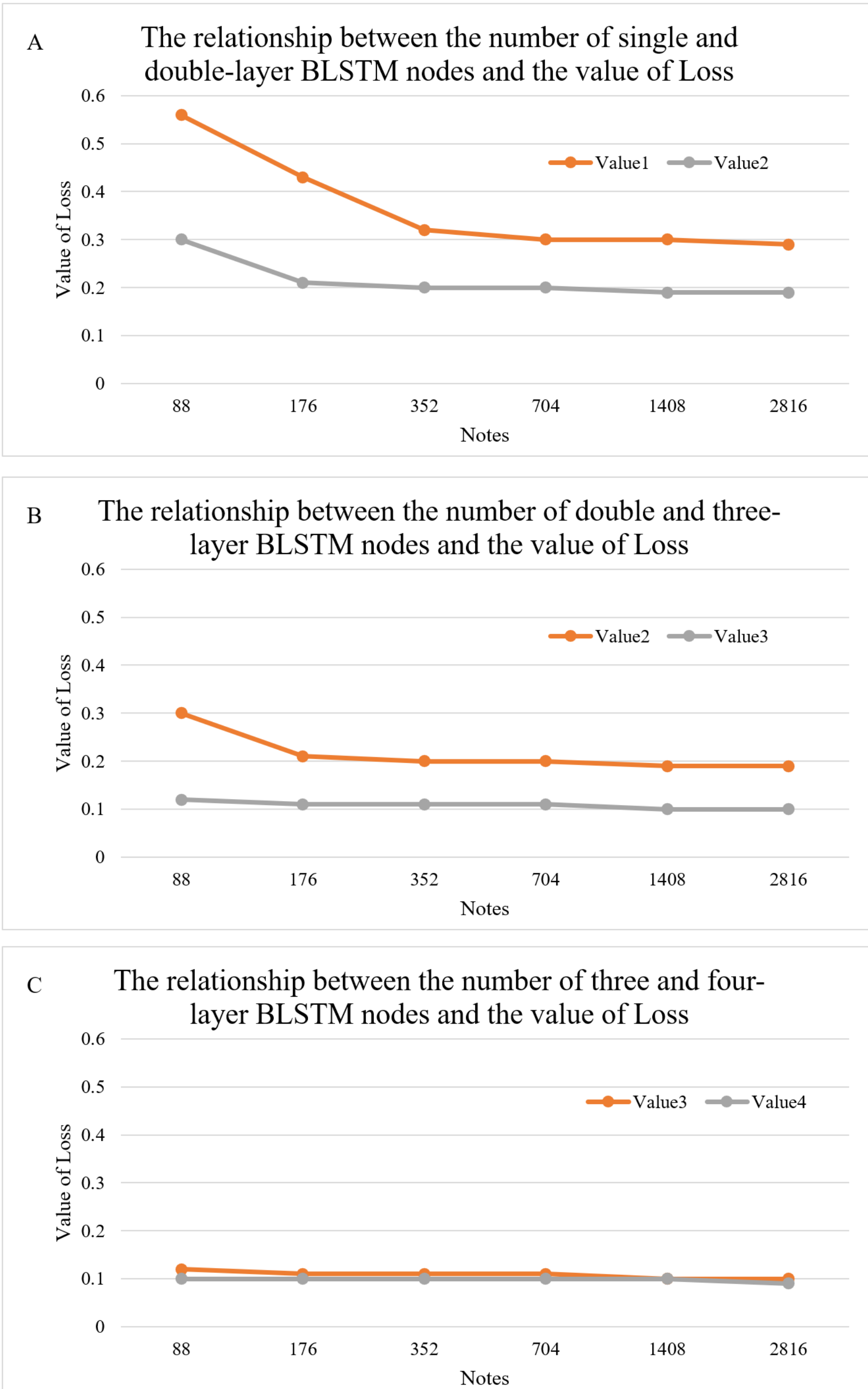


Figure 7. Relationship between LSTM node number and L value

Note. A represents single layer and double layer, B represents double layer and three layer, and C represents three layer and four layer.



$$\text{precision}=\frac{\text{TP}}{\left(\text{TP+FP}\right)} \tag{17}$$

$$\text{recall}=\frac{\text{TP}}{\left(\text{TP+FN}\right)} \tag{18}$$

$$F_1=\frac{2^*\text{precision}^*\text{recall}}{\text{precision+recall}} \tag{19}$$

When multivariate classification is studied, it can calculate the parameters of the category. The five evaluation grades are divided into *excellent*, *good*, *medium*, *not good*, and *poor*. When calculating the category *excellent*, excellent is regarded as a positive category, and the *good*, *medium*, *not good* and *poor* are regarded as negative categories. For category calculation, the multivariable classification problem can be transformed into multiple binary classification problems.

The bidirectional LSTM models of single hidden layer and three hidden layers are tested through the test set, and the tables are designed. The horizontal coordinate is the actual value, and the vertical coordinate is the test value. The obtained data is shown in Figure 8.

Figure 8A indicates TP=456, FP=153, FN=144, and TN=2,247 for the category *excellent*. According to Equation (17), it can be calculated that the accuracy value of the *excellent* category is 0.75, the recall value is 0.76, and the F1 value is 0.75. Similarly, it can calculate three parameters for the *good*, *medium*, *not good*, and *poor* categories, respectively. The same parameter values can also be obtained from Figure 8B, and the parameters can be sorted out as shown in Figure 9.

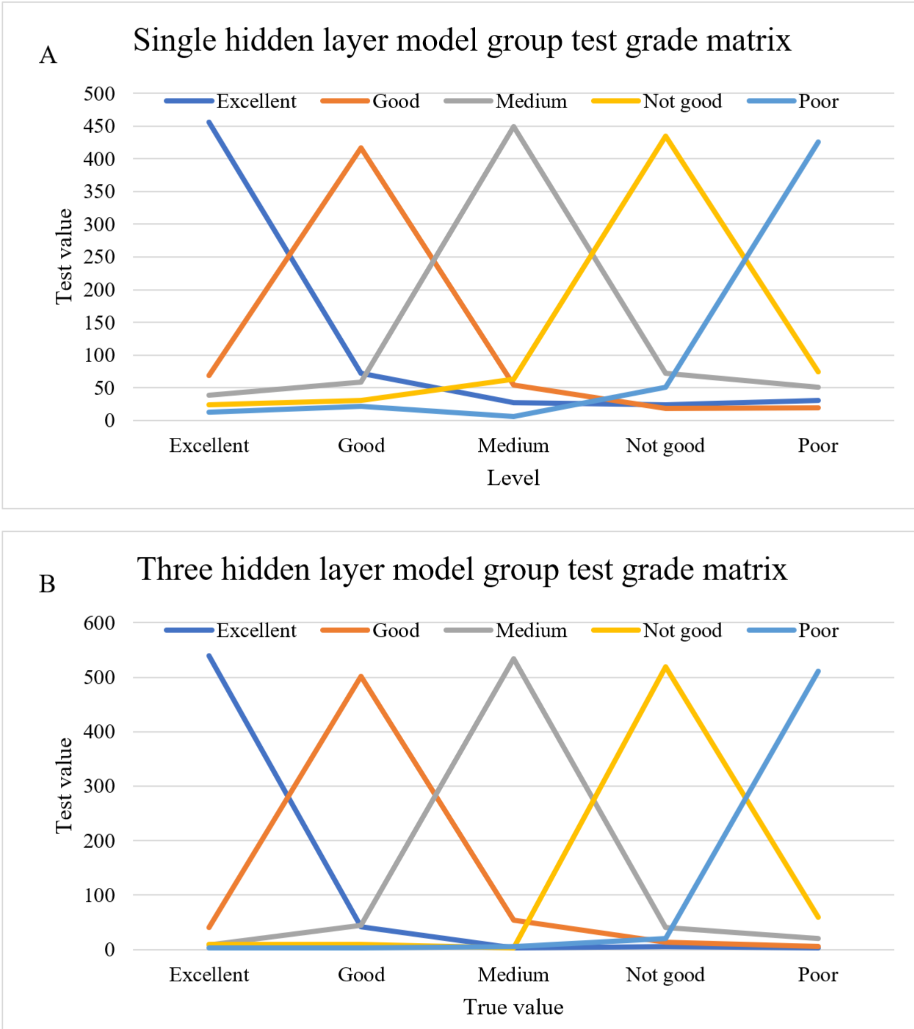
Figure 9 suggests that in the five categories of *excellent*, *good*, *medium*, *not good*, and *poor* the precision, recall rate, and F1 value of the bidirectional LSTM model of three hidden layers are higher than those of the bidirectional LSTM model of a single hidden layer.

Analysis of Practical Applications

One of the goals of music education is to cultivate students' performance abilities. However, traditional music performance evaluation methods are easily influenced by subjective factors, resulting in unstable and inaccurate evaluation results. In order to solve this problem, some music performance evaluation methods based on computer technology and deep learning models have emerged in recent years. These methods can objectively and accurately evaluate students' performance ability and provide better guidance and feedback for teachers. This article introduces a MIDI piano teaching performance evaluation method based on bidirectional LSTM and explores its practical application prospects in the field of music education.

- Music teaching evaluation: This method can be applied in the music teaching process to help teachers evaluate students' piano performance. By using a bidirectional LSTM model, teachers can obtain more objective and accurate evaluation results, thereby better guiding students' learning and improvement.
- Music grading evaluation: Music grading is an important part of music education, which evaluates students' music skills and performance abilities. The method proposed in this article can serve as an auxiliary tool for objective evaluation of grading performance, providing more accurate scores and feedback, helping candidates understand their shortcomings and improve.
- Music competition judges: In a music competition, judges need to evaluate and score the performance of the contestants. The use of bidirectional LSTM based evaluation methods can provide more objective and consistent evaluation criteria, avoid the influence of subjective bias, and ensure the fairness and reliability of competition evaluation.

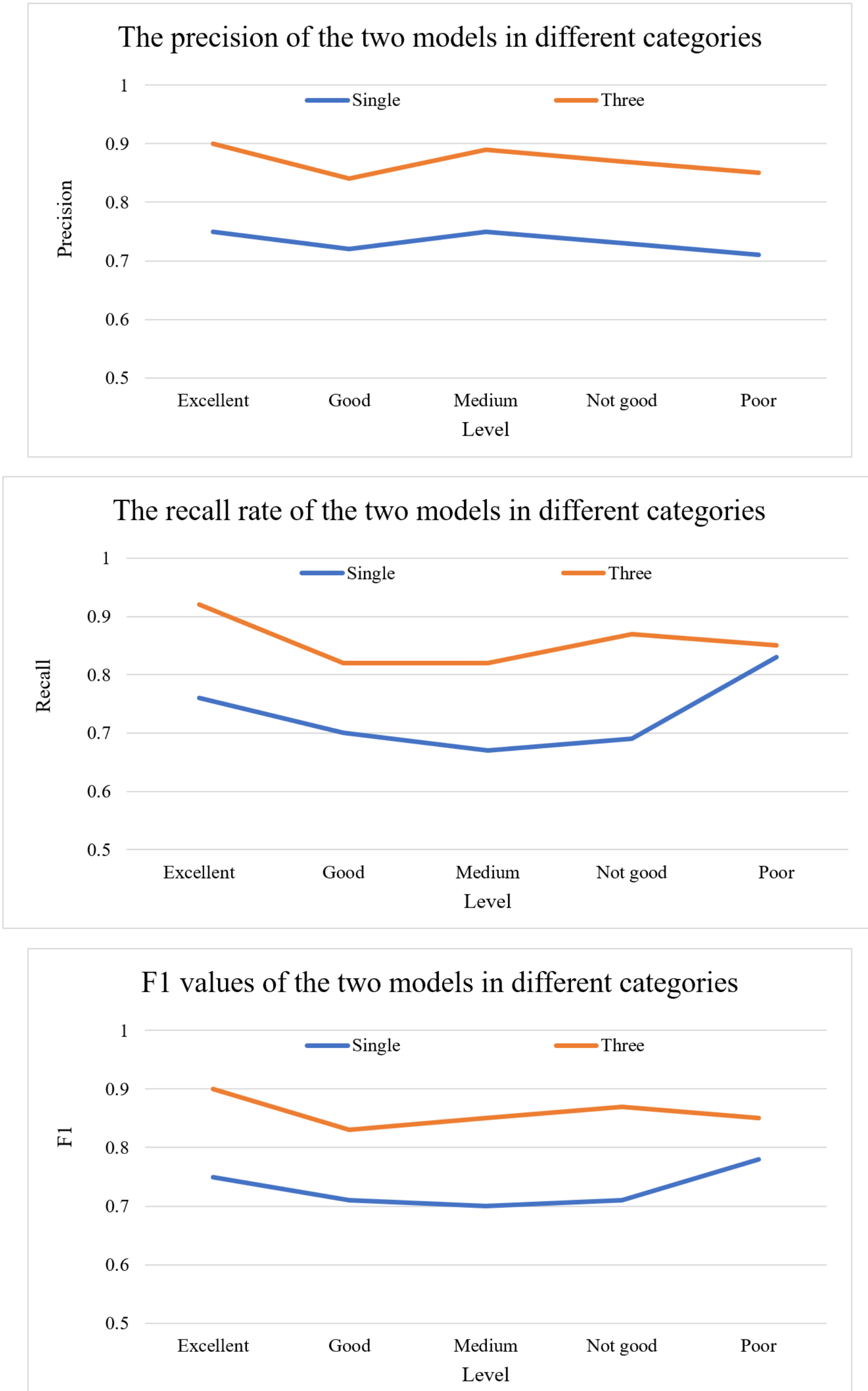
Figure 8. Test grade of hidden layer model group



- Music creation assistance: This method can be combined with deep learning models to assist in music creation. By analyzing a large amount of music data, models can learn the patterns and styles of music, helping music creators generate new music segments or provide creative inspiration.
- Music education research: The evaluation method based on bidirectional LSTM can provide a new tool and method for music education research. Researchers can use this method to evaluate and compare music performances of different teaching strategies and student groups, further exploring effective music education methods and teaching effectiveness.

In summary, the MIDI piano teaching performance evaluation method based on bidirectional LSTM proposed in this article has broad practical application prospects, which can provide more accurate and objective evaluation and guidance in the field of music education, promote the development of music education and improve students' learning ability. Although the MIDI piano

Figure 9. Comparison of precision, recall rate, and F1 value of bidirectional LSTM model between three-layer hidden layer and single-layer hidden layer



teaching performance evaluation method based on bidirectional LSTM has broad practical application prospects, there are also some challenges and limitations:

- Data collection and annotation: This method requires a large amount of piano performance data as the training set, and these data also need to be accurately annotated. The process of data collection and annotation may require a significant amount of time and human resources. To address this issue, it is possible to consider utilizing existing open-source datasets and reducing the workload of data annotation through automation technologies such as MIDI file conversion.
- Model training and tuning: Evaluation models based on bidirectional LSTM require a complex training and tuning process, including determining appropriate network structures, parameter settings, and training algorithms. This requires professional knowledge and technical skills, and it may take a long time to achieve ideal model performance. To improve efficiency, pre-trained models or transfer learning methods can be used to reduce training time and resource consumption.
- Diversity and personalization: Music is an art form that is rich in personality and creativity, and different people may have different performance styles and expressions. The evaluation method based on bidirectional LSTM may not fully capture individual differences and diversity, leading to bias and inaccuracy in the evaluation results. Therefore, it is possible to consider combining other machine learning methods, such as attention mechanisms or generative adversarial networks, to better handle personalization and diversity.
- Difficulty in evaluating non-technical factors: Music performance not only involves technical abilities, but also includes non-technical factors such as emotional expression, artistic feelings, and creativity. The evaluation method based on bidirectional LSTM may have limitations in evaluating these non-technical factors, making it difficult to provide comprehensive evaluation and guidance. Therefore, a combination of artificial intelligence and human experts can be considered to provide more comprehensive evaluation and guidance.
- Real time and interactivity: In actual music teaching environments, students need real-time feedback and guidance. The evaluation method based on bidirectional LSTM may require a certain amount of computation time to generate evaluation results, which limits its application in real-time and interactivity. To solve this problem, the inference speed of the model can be optimized, and parallel computing or distributed systems can be used to accelerate the evaluation process.

In summary, the MIDI piano teaching performance evaluation method based on bidirectional LSTM faces some challenges in the application process. However, with the continuous development and improvement of technology, these problems may gradually be solved, providing better evaluation tools and methods for music education.

CONCLUSION

Different music performances require different evaluation methods. It is crucial to provide guidance for music educators in selecting an effective and reliable evaluation method. Research has shown that computer evaluation methods can effectively avoid subjective biases in manual evaluations. Therefore, BD evaluation is the most suitable evaluation method for classroom teaching. BD evaluation improves the reliability and effectiveness of evaluation through a series of strict evaluation criteria, avoiding the influence of subjective assumptions. In order to compensate for the shortcomings of rule-based evaluation methods that cannot consider music continuity and expressiveness, this paper proposes a MIDI piano teaching performance evaluation method based on bidirectional LSTM. This method utilizes a three-layer bidirectional LSTM neural network mechanism to make it easier for the model to capture useful information. In addition, the Spark clustering training model is constructed using the deeplearning4j deep learning framework, and the model parameters are adjusted through the UI dependencies provided by deeplearning4j to improve work efficiency. The experimental results verified the superiority of the bidirectional LSTM model. The methods provided in this article can enhance students' independent practical abilities and reduce the pressure on teachers during the teaching process. These measures can promote the development of music education, improve students' music literacy and learning skills, and make positive contributions to the music education industry. This article only used specific datasets for experimentation and validation, without considering music data from other sources. Therefore, the applicability and generalization ability of evaluation methods require more empirical research to verify. In the future, deep learning models can be considered to study how to automatically generate or assist in creating music works, providing more resources and tools for music education, and helping students cultivate creativity and expression abilities.

DATA AVAILABILITY

The figures and tables used to support the findings of this study are included in the article.

CONFLICTS OF INTEREST

The author declares that there are no conflicts of interest.

FUNDING STATEMENT

This work was not supported by any funds.

ACKNOWLEDGEMENTS

The author would like to show sincere thanks to those who have contributed to this research

REFERENCES

Cao, H. (2023). The analysis of edge computing combined with cloud computing in strategy optimization of music educational resource scheduling. *International Journal of System Assurance Engineering and Management*, *14*(1), 165–175. doi:10.1007/s13198-021-01452-w

FangB.LiangX. (2023). Design and application of cloud computing recommendation based on genetic algorithm in piano online course video system. Research Square [Preprint]. 10.21203/rs.3.rs-2709888/v1

Feng, Y. (2023). Design and research of music teaching system based on virtual reality system in the context of education informatization. *PLoS One*, *18*(10), e0285331. doi:10.1371/journal.pone.0285331 PMID:37796970

Gao, H., & Li, F. (2023). The application of virtual reality technology in the teaching of clarinet music art under the mobile wireless network learning environment. *Entertainment Computing*, *49*, 100619. doi:10.1016/j.entcom.2023.100619

Huang, Y. (2023). Research on the intelligent teaching mode of college piano classroom based on big data technology. *Applied Mathematics and Nonlinear Sciences*, *9*(1), 1–18. doi:10.2478/amns.2023.2.00889

Jing, Z. (2023). Construction and application of piano to intelligent teaching system based on multi-source data fusion. *Journal of Circuits, Systems, and Computers*, *32*(4), 2350071. doi:10.1142/S0218126623500718

Li, H. (2020). Piano automatic computer composition by deep learning and blockchain technology. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 188951–188958. doi:10.1109/ACCESS.2020.3031155

Li, X., & Robin, H. (2019). Construction and analysis of hidden Markov model for piano notes recognition algorithm. *Journal of Intelligent & Fuzzy Systems*, *37*(3), 3293–3302. doi:10.3233/JIFS-179131

Livingstone, S. (2015). Critical reflections on the benefits of ICT in education. *Oxford Review of Education*, *38*(1), 9–24. doi:10.1080/03054985.2011.577938

Luo, G., Ning, Y., & Zong, L. (2023). The expression of sound art in piano performance. *Journal of Education and Educational Technologies*, *3*(11), 17–22. doi:10.53469/jeet.2023.03(11).04

Ranjan, R. (2014). Modeling and simulation in performance optimization of big data processing frameworks. *IEEE Cloud Computing*, *1*(4), 14–19. doi:10.1109/MCC.2014.84

Sha, X., & Zhang, Y. (2023). Simulation of music education multimedia mode in network environment based on QNSO algorithm and dynamic optimization. *Soft Computing*, *2023*, 1–11. doi:10.1007/s00500-023-08920-4

Shuo, C., & Xiao, C. (2019). The construction of internet+ piano intelligent network teaching system model. *Journal of Intelligent & Fuzzy Systems*, *37*(5), 5819–5827. doi:10.3233/JIFS-179163

Sun, S. (2021). Evaluation of potential correlation of piano teaching using edge-enabled data and machine learning. *Mobile Information Systems*, *2021*, 6616284. doi:10.1155/2021/6616284

Wang, S., Li, J., Cao, T., Wang, H., Tu, P., & Li, Y. (2020). Dance emotion recognition based on Laban motion analysis using convolutional neural network and long short-term memory. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 124928–124938. doi:10.1109/ACCESS.2020.3007956

Wang, T. (2022). Neural network-based dynamic segmentation and weighted integrated matching of cross-media piano performance audio recognition and retrieval algorithm. *Computational Intelligence and Neuroscience*, *2022*, 9323646. doi:10.1155/2022/9323646 PMID:35602641

Xia, Y. (2020). Resource scheduling for piano teaching system of internet of things based on mobile edge computing. *Computer Communications*, *158*, 73–84. doi:10.1016/j.comcom.2020.04.056

Xie, Y. (2023). Use of big data technology for network classroom teaching quality management. *International Journal of Emerging Technologies in Learning*, *18*(14), 256–270. doi:10.3991/ijet.v18i14.41917

Yan, H. (2022). Design of online music education system based on artificial intelligence and multiuser detection algorithm. *Computational Intelligence and Neuroscience*, *2022*, 9083436. Advance online publication. doi:10.1155/2022/9083436 PMID:35371210

Yao, Y. (2023). Online and offline hybrid teaching mode of piano education under the background of big data and Internet of Things. *Journal of Computational Methods in Sciences and Engineering*.

Yu, X., Ma, N., Zheng, L., Wang, L., & Wang, K. (2023). Developments and applications of artificial intelligence in music education. *Technologies*, *11*(2), 42. doi:10.3390/technologies11020042

Yu, Z., & Yang, B. (2020, October). Research on piano performance strength evaluation system based on gesture recognition. In *2020 13th International Conference on Intelligent Computation Technology and Automation (ICICTA)* (pp. 198–205). IEEE. doi:10.1109/ICICTA51737.2020.00049

Zhang, N. (2023). Video transmission and teaching simulation of piano network course based on mobile service terminal development. *Soft Computing*, *2023*. doi:10.1007/s00500-023-08895-2

Zhao, Y., Gao, W., & Ku, S. (2023). Optimization of the game improvement and data analysis model for the early childhood education major via deep learning. *Scientific Reports*, *13*(1), 20273. Advance online publication. doi:10.1038/s41598-023-46060-9 PMID:37985677

Zhong, M. (2023, August). A framework for piano online education based on multi-modal AI technology. In *2023 4th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI)* (pp. 89–93). IEEE. doi:10.1109/ICHCI58871.2023.10278015

Xiang Wei was born in Linfen, Shanxi, P.R. China, in 1989. He received the bachelor's degree from Xi'an Conservatory of Music, Shaanxi P.R. China. He received the master's and D.M.A degree from North Dakota State University, U.S. Now, he works at College of Art, Taiyuan University of Technology. His research interest includes collaborative piano and coaching, piano pedagogy, music theory.