https://doi.org/10.31449/inf.v48i5.5474 Informatica **48** (2024) 145-156 **145**

Multi-genre Digital Music Based on Artificial Intelligence Automation Assisted Composition System

Anna Liu
Shanghai Dance School; Shanghai, 200336, China
E-mail: kebi001155@163.com

Keywords: artificial Intelligence, multi-genre, digital music; Automation; Assisted Composition

Received: November 27, 2023

In the present paper, we conduct an in-depth study on multi-styled digital music and design an artificial intelligence-based multi-styled digital theme automation-assisted composition system. In the process of system development, B/S software design architecture is selected, J2EE ecosystem-related technology is used, and the system framework is built following the principle of sub-module development. A web hosting platform is provided for the whole Internet publishing system to realise the online and real-time functions. In addition, to improve the robustness, scalability, security, and ultimately, the system's quality of service, the functional design analysis, software architecture system, development technology, and optimisation strategy are carefully studied and discussed. After learning and analysing the file format of MIDI digital music, we chose the feature extraction method for each instrument to preserve the instrument's characteristics. We used its parameter model for each device during the composition to ensure the instrument's characteristics. After optimisation of the model, composition experiments were conducted, and the compositional effect was measured; the average ratio of adjacent notes with intervals within one octave was 83.57% for the composed pieces. The composition system eliminates using only a single algorithmic composition technique method in the system platform. Instead, it adopts the direction of a hybrid system that integrates multiple processes, an inevitable trend. The composition system provides flexible human-computer interaction at all music composition levels to improve the system's usefulness and effectiveness.

Povzetek: Razvit je AI sistem za avtomatizirano komponiranje glasbe z več žanri, ki uporablja B/S arhitekturo, J2EE tehnologijo in omogoča spletno objavo.

1 Introduction

Artificial Intelligence (AI) can be considered a technical science utilised to investigate, simulate, and extend human intelligence; it belongs to the computer science area. AI, nano-science, and genetic engineering are the world's top forward technologies in the 21st century, and the quick development of AI technology has impacted all areas of our lives. Since the 20th century, the use of computers has brought revolutionary breakthroughs in the development of music, and technology has become more closely linked to music [1]. The rise of electronic instruments, electronic music, and computer composing has given music a broader scope for development. With the help of specialised musical equipment such as computers, musicians' compositions can be directly translated into actual scores or even specific sounds [2]. This has freed musicians from many simple repetitive tasks and, at the same time, has dramatically expanded their creative abilities and skills. Thus, musicology must change the traditional thinking of composition, composing mode, and composing technology so that with the help of science and technology, the technicality and artistry of music works can strengthen and give people an audio experience beyond imagination [3]. The establishment of a music knowledge project will effectively analyse and summarize huge data in the "music information management system" systematically, establish a systematic understanding of the regularity of the creation process of different genres, composers, and styles, and make the information management system intelligent so that the music information accumulated for many years can be explored to the greatest extent [4]. Firstly, our ancestors have some empirical knowledge of music composition, which is insufficient to become a music composition theory. Secondly, as the accumulation of experience increases, the expression of knowledge and the exploration of laws become very important.

With the increasing demand for music information and access and utilization, as well as the rapid development of the music industry, the issues related to music information have been actively explored Its exploration substantially covers the conception, features, and functions of music information, music information demand and behavior, music information reclamation, library music information resource construction and service, digital music intellectual property rights, and the music assiduity [5].still, not important research has been assigned to the introductory proposition of digital music information and music information accession behavior, the exploration content isn't comprehensive, the exploration depth isn't sufficient, and there's a lack of in-depth discussion on the theoretical issues similar as the provocation of uses' digital music information accession behavior, the factors impacting the object of accession behavior, the medium of selecting the way of accession behavior, etc. There is no comprehensive and systematic analysis of users' digital music information accession behavior and relative analysis of different types of users [6]. Because of this, this article, grounded on the current study’s results, explores the fundamental theories on music information and music information geste and conducts a detailed investigation about the motivations of users’ digital music information acquiring behavior, acquiring behavior objects, mechanisms of developing behavior ways, the commonalities of users' digital music information acquiring behavior and the diversities of different various of users' digital music information developing behavior, to enhance music information [7]. It's anticipated to be very valuable and used in enhancing music information proposition and stoner information behavior proposition, directing users to effectively achieve digital music information and the healthy development of digital music assiduity.

With the broad scope of digital media art, the classification of digital media art is not distinguished based on the technology used. The primary basis for classifying digital media art is the main field of artistic experimentation or related categories in which digital media technology is engaged [8]. Digital art refers to art forms that are digitally expressed in computers, often created, and stored in "Binary." The research areas covered in this paper include the fields of composition, digital media art, and computing. Composition is the technical term used in this study to express a creator's musical ideas using a specialized, theoretical system of basic music theory, harmonics, polyphony, orchestration, and composition structure [9]. While discussing algorithmic composition, it is necessary to consider the system design's practicality and think about it simultaneously with the concept and development of computer music. From a macro perspective, algorithmic composition is the process by which a creator uses an algorithm to analyze and program musical fragments, elements, or potential laws embedded within them and drive a computer to generate a musical composition [10]. Computers have come to play an essential role in various aspects of the field of music composition. In contrast, algorithmic composition has been retained as a technical term to define the process of composing compositions in which algorithmic programs are the primary means of creation.

2 Related Works

The use of computers to assist in the creation of music and the use of algorithms to create music has a long history since the extensive use of computers and digital systems. The core concept is derived from the deconstruction and reorganization of digital information. Its development is an essential theoretical support for this study, a significant research source, and a basis [11]. On the one hand, more advanced synthesizers have been introduced; the establishment of MIDI protocols in the early 1980s has extensively enhanced the development of computer music and provided the most critical technical support for studying the structure of music and creating music in more dimensions [12]. The product of computer music is the expansion and derivation of traditional music, thus entering a new era of artistic creation. Developing mathematical and computer disciplines has provided significant technical support for artistic creation [13]. Regarding the scope of the definition of artificial intelligence, many scholars have explored it from different perspectives. It is hypothetically argued that machine intelligence can achieve subjective human thinking through machine computing and human thought. It can be briefly summarized that AI is the process of simulation of human thinking and behavior by machines. The music program A circle canon, such as Frère Jacques, is programmed using rule-based artificial intelligence [14]. The rules for generating melodies and harmonies are based on the rules for combining notes and chords. He uses his exploratory concepts and their role in society as a possible way to shed light on this issue. From all these early experimental studies, much valuable experience has been provided for research.

AI technology helps us to learn, compose and analyze music more efficiently. Customs-Carnicer J argues that combining AI technology and musical instruments enables the combination of software and hardware to make complex playing easy to learn and the training process more enjoyable. In contrast, the rise of online education forms makes music education more convenient [15]. The AI can also learn by itself and can be more efficient in performing tasks that require the "mechanization" of the human brain. Simultaneously, AI can learn independently, identify related characteristics in rich and diverse music, and store large amounts of data. AI technology can facilitate music classification, retrieval, and library construction. Kwak J believes it is revolutionary to promote the tagging management classification of ethnic music, to integrate and process data, and can be applied to the tagging classification management of ethnic music, combined with an information retrieval system to build an ethnic music database [16]. AI technology is beneficial to the dissemination of music and promotes the commercial development of music. De Beukelaer C believes AI's self-learning features will be highly convenient for music identification, dissemination types, and genres [17].

There are three main types of computer music systems based on artificial intelligence technology: composition, improvisation, and performance systems. It is a very challenging task to give computers the expressive power to play with the characteristics of human musicians. Previous approaches have severe limitations based on fixed or empirical musical rules. A closer approach to the human observation of the imitation process is to use implicit interpretations of music extracted directly from recordings of human performers. Brown A. E. et al.

proposed a neural network model based on variable self-coding, MIDI-VAE, in parallel with their first attempt to apply generative adversarial networks to symbolic music domain transfer [18]. Daniel R et al. used an autoregressive model and Gibb's sampling to transform the style of an arbitrary piece with harmonic structure into two different types, Bach choir and jazz [19]. Also, based on symbolic music, Guichardaz R et al. developed the first fully supervised algorithm based on synthesized data; this codec model can convert suggestive musical accompaniment between many different styles, and the accumulation of compositional levels is mainly based on experience [20]. The remarkable difference from natural science is that this experience is not studied through noteworthy controlled experimental observation but directly through the practical activity of music composition by gradually auditioning and correcting it while imitating and interpreting it according to traditional composition principles and compositional techniques, thus forming a musical work progressively with a specific style and particular aesthetic significance [21]. Musicology analysis can help us explore and discover specific characteristics of human creative thinking, especially for the guidance of imaginative thinking and innovative, inspirational thinking. Therefore, the object of our research is to investigate the structural characteristics of music and the cognitive thinking model of musicology and to explore the principles of the higher thinking model of human cognition and the possible methods of its implementation, which is of great significance for both music composition and cognitive science [22].

3 Artificial Intelligence for digital music modeling of multiple song styles construction

The theoretical model of digital music information acquisition behavior is a model that reflects the interrelationship and interaction among the elements of digital music information acquisition behavior. Digital music information acquisition behavior includes three key factors: digital music information acquisition behavior motivation, digital music information acquisition behavior object, and digital music information acquisition behavior mode. From the psychological perspective, the demand induces the engine, and the motivation governs the behavior [23]. The request for music information is also inseparable from the digital music information acquisition behavior model. Users' music information needs are influenced by their factors and environmental factors. Users' music information needs will trigger users' music information acquisition behavior motivation and influence users' digital music information acquisition behavior. Users' digital music information acquisition behavior determines what kind of digital music information acquisition method and what digital music information users acquire. Digital music information acquisition behavior objects and digital music information acquisition behavior interact. Digital music information acquisition behavior object is obtained through digital music information acquisition behavior. Different digital music acquisition behavior objects may have to be obtained through other digital music information acquisition behavior ways so that digital music information acquisition behavior objects will influence the choice of digital music information acquisition behavior way. Digital music information acquisition behavior objects satisfy users' music information needs. The digital music information acquisition behavior model is illustrated in Figure 1.

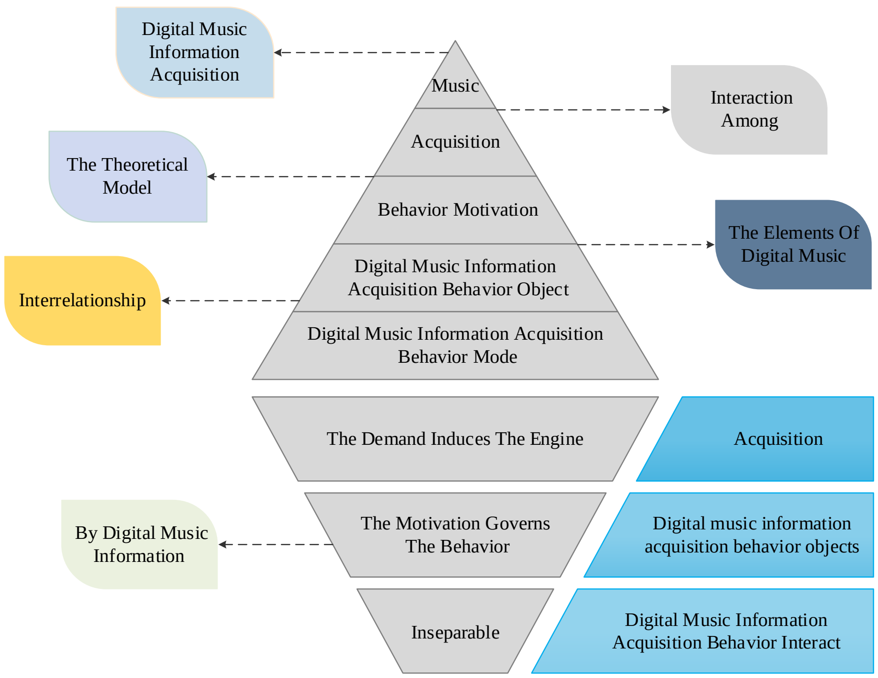

Figure 1: Theoretical model of digital music information acquisition behavior

Artificial intelligence can assist musicians in generating ideas, be applied to music creation, and directly produce music works. Although music is an artistic creation that reflects the creator's spirit, it is logical and calculable. Music composition techniques also reflect a rich and rigorous mathematical logic, such as melodic progression, fundamental transformation, harmonic pitch arrangement, instrument timbre matching, etc., all of which can be defined as a single or combined set of algorithms. Artificial intelligence music composition is to capture the mathematical logic implied behind the music through artificial intelligence technology and the use of extensive data analysis, a series of music data units, and intelligent algorithms in the computer software so that the software forms a machine learning, supervised learning, deep learning artificial intelligence model and neural network, according to the user's individual needs, select the related material to complete the automated composition.

Deep learning is a branch of machine learning, which is a branch of artificial intelligence. The concept of deep learning is derived from traditional neural networks but is not the same as conventional neural networks. However, deep learning algorithms usually include the word "neural network", for example, recurrent neural networks, convolutional neural networks, etc. Deep learning is an upgrade to the traditional neural network, a semi-empirical and semi-theoretical modelling approach in which human mathematical knowledge and overall architectures are built using computer algorithms. It then combines a large amount of training data and the computer's large-scale computing power to continuously adjust internal parameters to achieve problem goals. Compared to traditional machine learning, relying on manual feature extraction, manual feature extraction is simple and effective for specific tasks but not general and highly subjective. In contrast, deep learning is better at pattern recognition for data that cannot form symbols, such as image and waveform data. It relies mainly on automatic machine extraction when performing feature extraction, thus avoiding the influence of the subjectivity of manual recognition. Therefore, deep learning is less interpretable, but from the analysis of results, deep learning has better results than traditional machine learning.

The closer the actual value is to the predicted value, or the closer the proper distribution is to the expected distribution, the smaller the loss value and the better the model's performance. Conversely, the more significant the difference between the two, the larger the loss value and the worse the model version. Cross-entropy comes from information theory and measures the similarity of probability distributions. If there are two probability distributions$p(x)$ and$q(x)$ about the sample set, where $p(x)$ is the proper and non-true distribution, the cross-entropy H of p and q is defined in information theory as Equation (1).

$$H_{(p,q)} = \int \frac{p(x)}{log\, q\,(x) - 1} \times [p(x) - q(x)] \tag{1}$$

The application of cross-entropy in neural networks, where$p(x)$ the actual value$q(x)$ is the predicted value, and the function value H (p, q) measures the similarity between the predicted and actual values. The smaller the cross-entropy, the smaller the error between the predicted and actual values, and the goal of the neural network is to minimize the cross-entropy. For regression problems, the Mean Absolute Error (MAE) of Equation (2) and the Mean Square Error (MSE) of Equation (3) are often used as loss functions. Wherey_i denotes the actual value andy_{i-1} represents the predicted value.

$$MAE = \sum_{i=1} \frac{y_i + y_{i-1}}{n - y_i} \tag{2}$$

$$MSE = \int_{i=1} (y_i + y_{i-1}) \times (y_i - y_{i-1}) \tag{3}$$

The loss function solves the problem of complex metrics between the predicted and actual extents of the model, and the backpropagation algorithm solves another challenge, namely, how to make this metric effect drive the weights of the network to be constantly updated so that the loss function is minimised and the network is continuously trained [24]. The automatic learning of neural networks includes forward and backward propagation. In forward propagation, the pre-processed data is first input to the input layer. It then enters the hidden layer, where the result is passed to the output layer through a series of neuron operations, and the predicted outcome is output in the output layer. If the expected value is not equal to the actual value, then this error is calculated using the loss function. After the error is obtained, the backpropagation algorithm starts to operate. According to the estimated error of the output layer, the backpropagation is performed in some form layer by layer in the intermediate layers, and the model parameters are updated by gradient descent for each layer. Through several iterations, the error in the calculated value of the loss function is continuously reduced, and the neural network converges to an optimal state.

The activation function is introduced to make the neural network model linear. Without the activation function, the layers of the neural network would only be linearly multiplied or summed so that the output would only be a linear combination of the inputs, no matter how many layers there are. The activation function is a joint two-number activation function, which is calculated by the formula:

$$tanh(x - 1) = \sum \frac{sinh(x - 1)}{cosh(x - 1)} \times \frac{e^{2x} - 1}{e^{2x} + 1} \times \sqrt{e^x - e^{-x}} \tag{4}$$

The ReLU activation function has been more used in recent years, characterized by the fact that the gradient does not disappear when the input is non-negative and is calculated as Equation (5).

$$f(x+1) = \sum max \frac{x+1}{\sqrt{x-1}} \quad (5)$$

Since the commonly used Sigmoid function only supports the solution of binary classification problems. Still, the actual output results are usually more than two categories; the analogous extension is wanted to get a general method that can solve multi-classification problems. Softmax is thus generated, which is a kind of chemical function with a discrete probability distribution that will output the results derived from multi-classification issues in the form of probabilities, Softmax function equation (6).

$$s_i = \sum_{j=1} (e^j - n) \times e^j \quad (6)$$

Different predicted values are transformed into probabilities mapped in the (0,1) interval by the Softmax function and then normalized so that the likelihood of different values sums to 1. This way, the size of the probability directly reflects the possible size of the corresponding value as the prediction result, and the value having the highest chance is chosen as the prediction result. Softmax uses an exponential function so that the original large values are more significant and the small initial values are more minor, improving learning efficiency; secondly, the Softmax function is continuously derivable, and there are no inflexion points in its function image.

4 Artificial intelligence-based digital music automation-assisted composition system design for multiple song styles

The system uses J2EE ecosystem technology for architecture design, following the MVC design pattern and applying each technology's functional characteristics to the view, controller, and model layers. (1) Server layer: The J2EE platform-based system must run on the Web application server; Tomcat alone can release the system to reflect the idea of separation of movement; the system in Tomcat added Apache, which can further enhance the system scalability robustness. (2) View layer: The view layer for the user is the browser displays static web pages; the system in the view layer technology mainly uses Html to build the page framework, CSS is responsible for page style, Java Script, and JQuery to achieve effects and foreground validation, and JSP is responsible for displaying the controller layer dynamic data. (3) Controller layer: The traditional J2EE platform uses Servlet as the system controller layer, but in the actual project development, simply using Servlet will bring a lot of inconveniences, such as in accepting parameters, data validation, and returning JSON data. For this reason, Spring MVC is used as the controller layer technology implementation to improve development efficiency. (4) Model layer: The model layer is the core business layer of the system and is responsible for data calculation and storage. The technology used is also based on these two features Java writes the logic code classes of the business layer, Spring is responsible for the instantiation of business classes, and Mybatis uses XML to achieve access and storage with the database. And Memcached builds a memory cache of commonly used data to access such data quickly. Figure 2 indicates the system technical architecture design diagram.

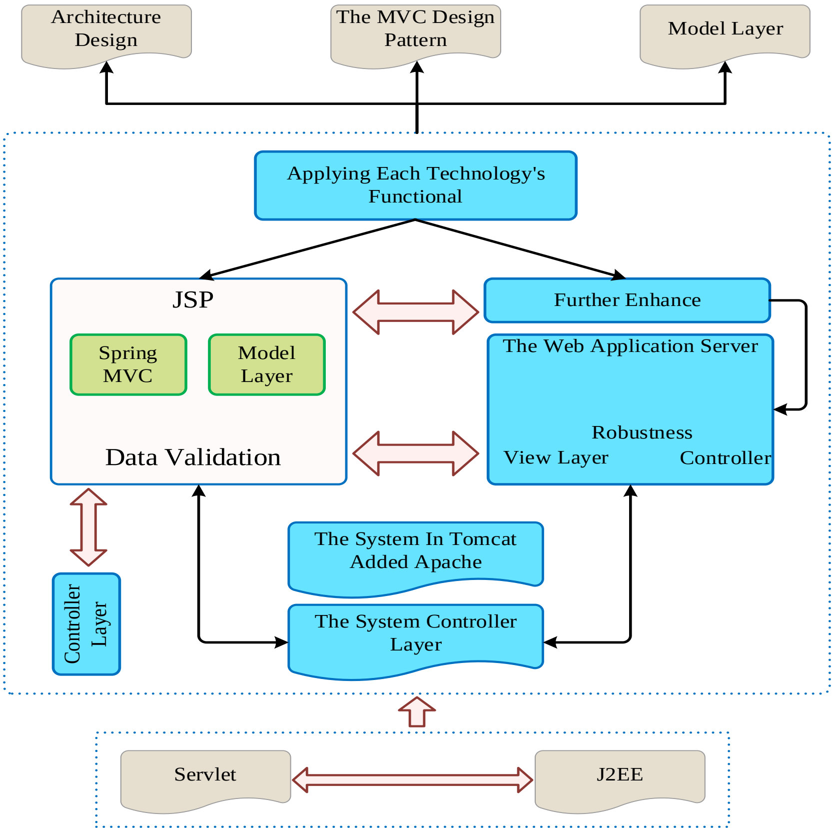

Figure 2: System technical architecture design diagram.

According to the results of user requirements analysis, the website users can be divided into three major categories of visitors interested in composing, composing users using the system and administrators of the management system, and then the analyzing the operational behavior of the three types of users respectively, and can be divided into two types of roles: front-end functional users and backend system administrators, where the front-end operational users include visitors and composing users [25]. Based on the big difference in the operation behavior of these two types of roles, the system is split into two separate websites: the front-end composing output website visited by front-end users and the back-end management website called by administrators.

(1) reduce the complexity of development; the system split is based on the role of the distinction, which makes each system focus on different user groups, and the corresponding functional modules of each system are significantly reduced. Secondly, the system design focuses on the front-end system; the front-page layout must be aesthetically pleasing, and the functional form must be reasonable to increase the workload when developing. On the contrary, the website owner uses the background system, and the page requirements are simple if the function meets the demand. Therefore, the front and backend are designed separately, and the development strategy is adapted to this, which reduces the extra workload brought by excessive design.

(2) Enhance the system's quality of service; after the system's release, two factors need to be redeployed to the system, and the system will not be accessible during deployment. If the system is not split, the entire system must be shut down regardless of the functional module in which such a situation occurs, which is a very unfriendly experience for the user. On the contrary, after the system is split, because the front and backend run independently under different servers, they do not interfere with each other, which reduces the frequency of the entire system being inaccessible.

The system deployment design mainly studies the operation environment and access policy after the system is released, which can be designed separately from three aspects: software, hardware, and network media. The system deployment configuration design diagram has been provided in Figure 3.

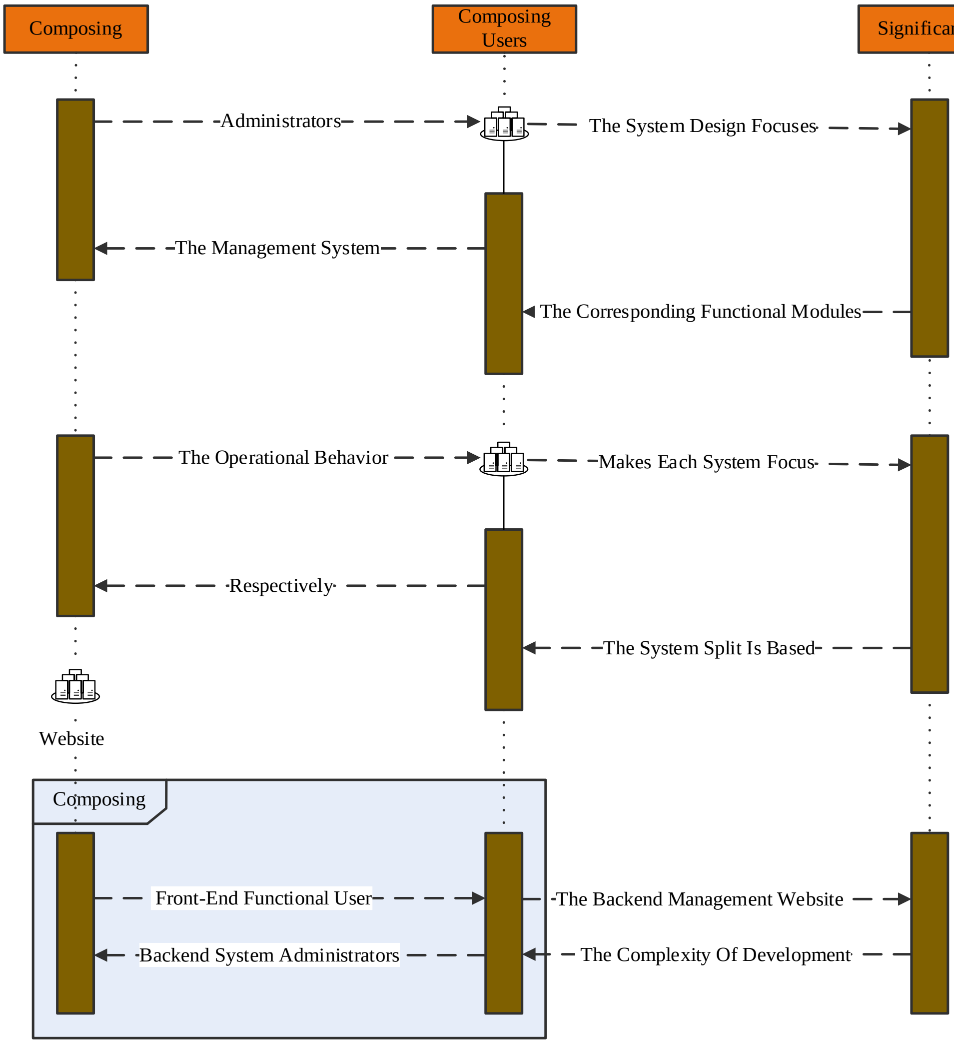

Figure 3: System deployment architecture design diagram.

(1) Software deployment environment: Through the analysis of the technical architecture, the required software can be divided into three major categories, which are database and cache server, reverse proxy server, and application server.

(2) Hard deployment environment: The system adopts B/S architecture, and the system needs at least one host with a public IP address to publish the design. To create a public and open cloud computing service platform, through the virtualization of hardware and software resources, the primary resources are turned into a "pool" that can be freely dispatched, thus realizing the rationing of resources on demand.

(3) Network medium: The most significant advantage of B/S architecture is that no additional software must be installed, and the device only needs to support a browser to access the operating system so that the system can be accessed through the traditional Internet and the mobile Internet. However, since the system is a music output system, there will be many upload and download operations, so the bandwidth of the server should be adjusted according to the website's usage.

The dataset used for the automatic composition experiments is the Enya MIDI music set, a dataset of digital music files for multi-instrument ensembles. Most of the musical works in the dataset are performed by multiple instruments, with different bars of the songs played solo or by various agencies in an ensemble, sometimes melodic, passionate, quiet, lively, and crisp. In addition, most of the songs in the dataset choose 4/4 time as the main rhythm, making integrating the instrumental characteristics of different pieces better when composing ensembles after learning them [26]. Therefore, it is advantageous to use it as a dataset: the variety of instruments used in the dataset makes it possible to learn and compose ensembles with multiple devices; most of the songs in the dataset are in 4/4 time, which is conducive to the integration of instruments when ordering ensembles; the dataset is in MIDI format, which is convenient for interpretation and feature extraction.

5 Analysis of results

5.1 Digital music model analysis of multiple song styles with artificial intelligence

In a situation where computer science and artificial intelligence have not yet solved how common sense and meta-knowledge can help solve problems that domain knowledge cannot, domain knowledge still must be expressed concretely as the derivation of a specific rule and state. However, understanding such as musical composition is subjective from the beginning to the end, full of emotions and aesthetics. The work must be varied to obtain exciting results while maintaining consistency. The model was adjusted for each parameter the model through several comparative experiments. Experiments on instrument identification of music composed by automatic composition were started, using the mentioned MIDI results of design, 20 audio tracks of 75S in each of four genres: piano, guitar, bass, and strings, transcribing each MIDI file to WAV format through MUSeScore3, and then cutting the yellow frequencies, cutting each the 75S audio was cut into audio collections with a length of 10S offset of 5S. Finally, 280 audio tracks were obtained for every category, of which 220 tracks were training sets while 60 tracks were test sets, and the resultant loss rate of training is illustrated in Figure 4.

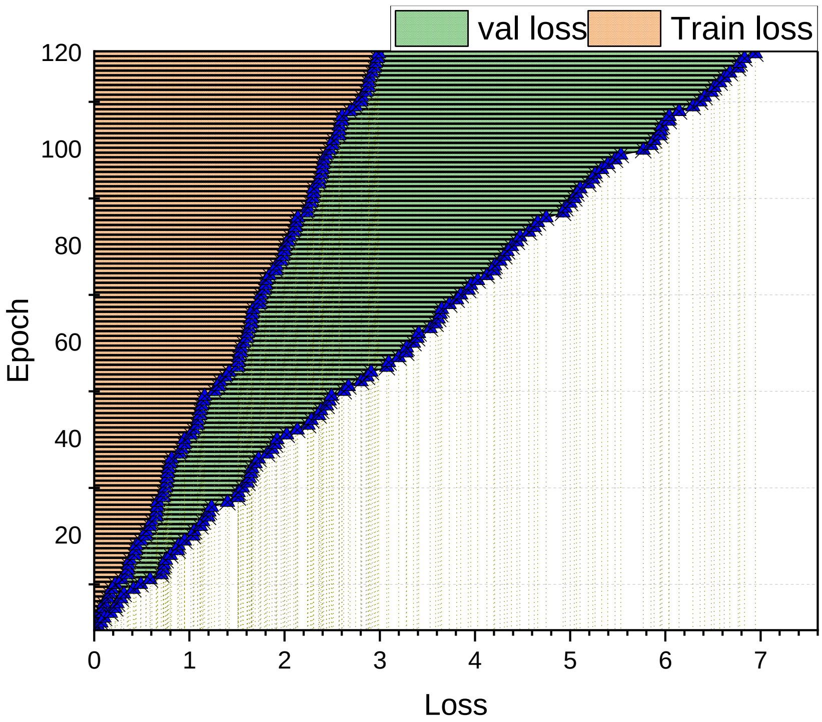

Figure 4: Resultant loss rate of training

The core requirement of the system is to automate the composing process by simply setting some data. Still, it is also necessary to implement a series of functions related to the automatic composing process, such as composing data set definition, composing result visualization and analysis, effective management, a preview of the composing data set, and recording audio and writing results [27]. The system allows users to record, play and edit audio, remove noise from audio to obtain the data sets needed for automatic composition training, compose music automatically by audio or compose instrumental music by MIDI, and visualize the information in the audio by spectral analysis of the results of the compositions and by note-recognition of the melodies and their conversion into short scores. The system can also be used to visualize the information in the audio by performing spectral analysis on the results of the tracks, note recognition of the melody, and its conversion into a simple score. When the user starts the system and clicks on record first, the system turns on the microphone and begins recording and timing. The system will generate temporary audio for the user to listen to, evaluate the recording effect, and choose whether to keep it.

Audio analysis is a function designed to support the user to perform a visual analysis of the music composed or the audio note the user is interested in and wants to know more about. This section mainly provides the user with a visualization of audio sentiment analysis, melody analysis, instrument type analysis, time domain spectrum, etc. The subject matter involved is instrument-related compositions, so for the time being, the development of this section has only completed the visualization of instrument type, time domain diagram, and time-frequency diagram information. An audio file is selected and loaded, and when OK is clicked, the file can be analyzed accordingly; next to the training function is a training of the instrument recognition model, thus refreshing the training result parameters of the model. After the audio analysis, the results of the instrument recognition type of the audio are displayed, as well as the audio's time domain and frequency map information. The audio's time domain and frequency maps are drawn using the Librosa library. The audio's time domain and frequency maps facilitate the study of waveforms, which describe mathematical functions or physical signals versus time. The audio signal's time domain waveforms can express the movement's changes over time. Time-frequency diagrams reflect the relationship between time and audio frequency variation. The experimental setup sends 10,000 concurrent requests to the FISCO BCOS federated chain nodes to invoke the transaction log contract and observes the interpretation of the number of transaction log contract transactions (TPS) that can be processed per second for a different number of nodes. By increasing the number of nodes, the TPS remains in the range of 280 to 290. The TPS values for various numbers of nodes are shown in Figure 5.

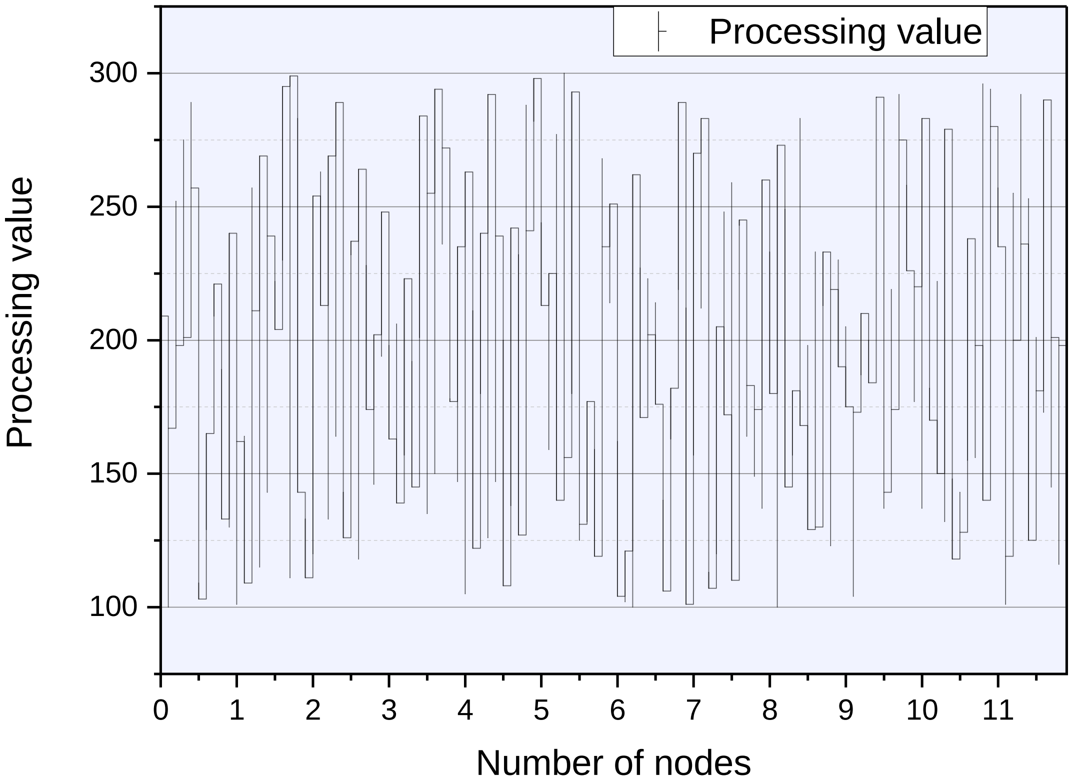

Figure 5: Processing value per second with different numbers of nodes.

5.2 Artificial intelligence of multi-track style digital music automation assisted composition system implementation

The system has been studied and implemented for MIDI digital music composition, so only the MIDI-related sections of the system have been developed and implemented so far, which are recording, management of recorded music, MIDI digital music composition, identification of instrument categories of composing tracks, and visual analysis of the corresponding characteristic frequency and spectrogram of audio [28]. The system records audio data and generates recording objects, records the related description information in the database, and puts the recording data in the static resource folder for storage; when managing the recording files, the first step is to obtain the audio information in the database and then process it accordingly; when composing model training with customized datasets, the dataset files are loaded to record the model training results of each instrument in the static When ordering, select the dataset, load the corresponding model training result and description file for action song, and then save the composing result in the static resource folder; when training the instrument recognition model, first load the WAV dataset for model training, and then save the model result in the static resource folder after training; when analyzing the universal frequency, load the instrument recognition model for When studying the expected frequency, the instrument recognition model is loaded to identify the instrument type, and after that, various spectral maps of the audio files are calculated and exported and saved in the static resource folder. The system test line diagram is shown in Figure 6.

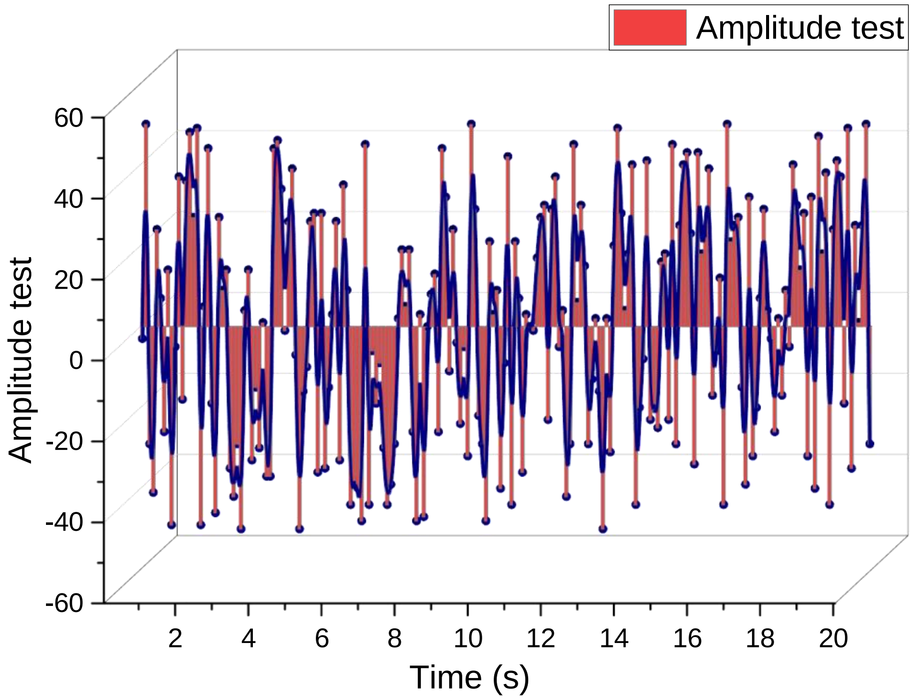

Figure 6: System test line graph

Through the above operation, three different versions of the network model are successfully built, and then we need to set some reasonable hyperparameters to train and evaluate these networks. The whole training process is shown as follows: firstly, two large loops are set up; the first is used to update the number of iterations of the network, and the second is used to update the batch block data obtained each time. Then, the batch block data is adjusted to the specific format required by the network; then, the data blocks are considered inputs for the set network model, and the output results are achieved by weight calculation. An error value is calculated based on the output result and the accurate result, and the backpropagation algorithm is used to calculate the corresponding deviation value to update the weight value of the network; finally, the above operations are repeated until the network achieves the accuracy requirement or reaches the set number of iterations. The following table shows the hyperparameters used in the first version of the network and the default values of these hyperparameters. The above operations resulted in many LSTM-based music generation network models, which were then validated by using loss curves to see the convergence of the various models and by the next evaluation of the effect of the produced MIDI files. The entire model evaluation process is shown as follows: first, the input note data is preprocessed by the prepare sequences output (notes, pitch names, n_ vocab) function, where notes denote the input notes, pitch names represent the different note names, and n_ vocab refers to the number of notes. Then the network is built and initialized with the create_ network_ _add_ weights () function; then, the notes are predicted with the generating_ notes () function; finally, the output notes are reconverted to MIDI files and saved with the create_ midi () function. The experimental data on learning rate curves for diverse models are shown in Figure 7.

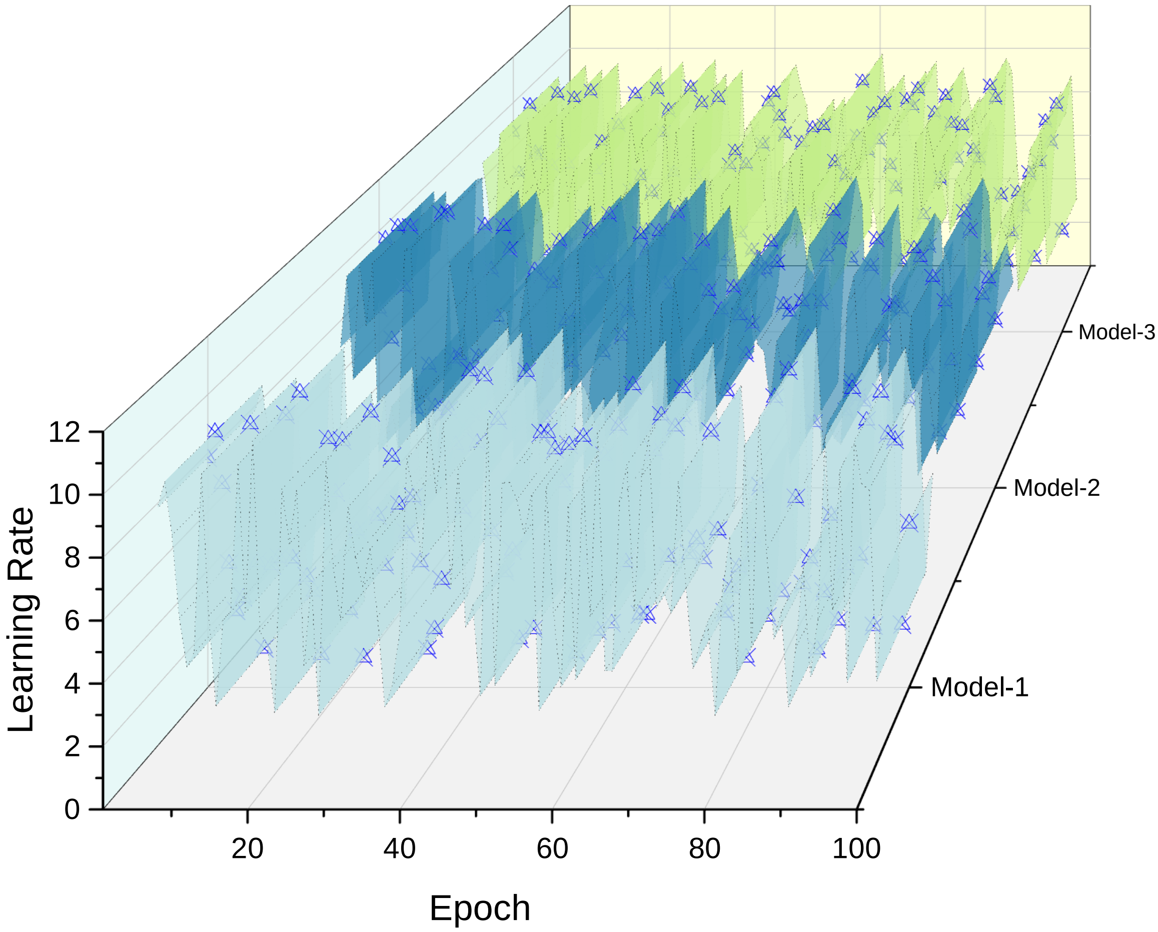

Figure 7: Experimental data for learning rate curves for various models.

According to the structure of musical language analysis, the mode of musical knowledge processing, and the characteristics of musical thinking variation, the basic ideas and methods of possibility construction space theory are used to explore the parts of the space of trait features of melodic motives, the opening of musical variation features and the length of musical style features, as well as the basic patterns of evolutionary reasoning among these three spaces. The specific application of variational operators of possibility construction space theory in computer composition is proposed. To study how to describe the characteristics of musical motives more completely, i.e., to find the space of traits of melodic motives, we first analyze the process of musical composition by studying the reasons for two musical fragments, from which we find out how musical works describe the characteristics of motives, how motives develop and change into variants, and how some variants of explanations are selected according to composition theory and experience to form a particular style of compositions with certain The design is a work with a specific aesthetic meaning, which expresses the thoughts and feelings of a particular group of people. Due to the diversity and ambiguity of musical composition knowledge, the above method causes a significant redundancy of data in a computerized composition system, directly affecting the system operation's speed and convergence; 500 simultaneous access requests were made in 2 seconds. For the music audio file upload interface, the mean response time was 208ms, and the maximum response time was 221 seconds, of which the system within 210ms responded to 99% of the user requests; for the start comparison interface, the average response time was 229ms, and the maximum response time was 258 seconds, of which 99% of the user requests were for the start matching interface, the average response time is 229ms, the maximum response time is 258 seconds, and 99% of the user requests get the system response within 232ms. According to the results, all requests get a conventional response, the exception rate is 0.00%, and there is no error. Hence, the results obtained by the backend interface performance test are located within an acceptable area, and the result is overtaken. The response time of the interface under the simulated high concurrency is also within an affordable range, and the system keep the normal flow. Therefore, the performance test result for the music comparison and analysis system is passed. The test results of the digital music automation-assisted composition system are shown in Figure 8.

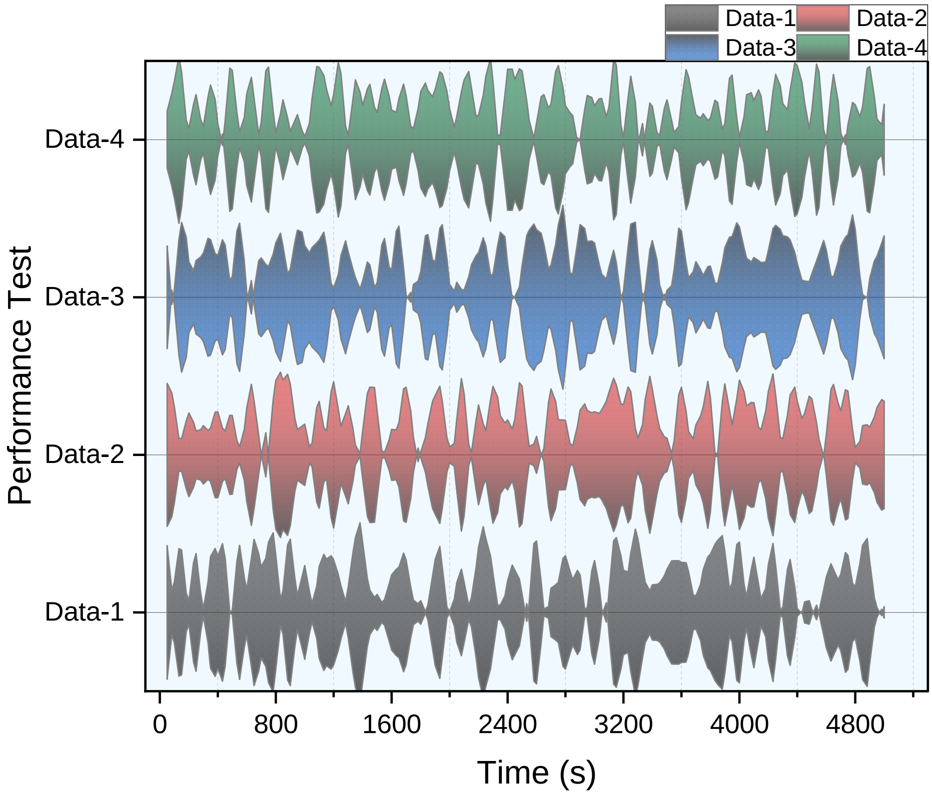

Figure 8: Test results of digital music automation-assisted composition system

6 Conclusion

This paper is based on the artificial intelligence field of music automation-assisted composition by comparing the music signal of computer language with music language, starting from the formal aspect of music and dismantling the melody recognition, rhythm recognition, harmony recognition, song structure recognition, music style recognition, and music emotion recognition claimed by the current artificial intelligence according to the dimensions of the essential elements of music, such as pitch, length, intensity, and quality of sound. This paper is based on the original C/S architecture. In this paper, based on the inconvenience of the original C/S architecture system, user interaction functions are extracted and developed using B/S architecture, and the automatic composition output system is conceived and designed. For the effectiveness of automatic composition, we can verify the diversity of composition by viewing the score of composition results, and by analysing the pitch sequence of composition results, the percentage of adjacent pitches of each instrument with intervals of not more than one octave is 83.57% illustrates the continuity of the compositional results. By analysing within bars, it was verified that the repertoire we composed corresponded to the characteristics of each instrument. The study of the representation and reasoning of the rhythm, pitch, intensity, and accompaniment of the music, the use of modern music technology for a more comprehensive musical composition, and the construction of the overall framework of the music. The music is also refined semi-supervised, thus simplifying the work of arranging the music. The experiments show that we combine artificial intelligence with automatic computer compositions that respond to this property of musical uncertainty and generate a piece that conforms to the rules of musical theory. Defining and extracting more independent evidence based on AI and music libraries makes musical reasoning more flexible and richer, improves the musical accompaniment knowledge base, and increases the efficiency of automated-assisted composition.

Competing of interests

The authors declare no competing of interests.

Authorship Contribution Statement

Anna Liu: Writing-Original draft preparation, Conceptualization, Supervision, Project administration.

Data Availability

On Request

Declarations

Not applicable

References

[1] Z. Li *et al.*, “Text compression-aided transformer encoding,” *IEEE Trans Pattern Anal Mach Intell*, vol. 44, no. 7, pp. 3840–3857, 2021.

[2] L. M. Meier and V. R. Manzerolle, “Rising tides? Data capture, platform accumulation, and new monopolies

in the digital music economy," *New Media Soc*, vol. 21, no. 3, pp. 543–561, 2019.

[3] M. Eriksson, "The editorial playlist as container technology: on Spotify and the logistical role of digital music packages," *J Cult Econ*, vol. 13, no. 4, pp. 415–427, 2020.

[4] S. Murphy, "Music marketing in the digital music industries–An autoethnographic exploration of opportunities and challenges for independent musicians," *International Journal of Music Business Research*, vol. 9, no. 1, pp. 7–40, 2020.

[5] D. Nakano, "Digital music, online outlets and their business models," *Brazilian Journal of Operations & Production Management*, vol. 16, no. 4, pp. 581–591, 2019.

[6] M. Peng and S. Y. Bae, "A Study on the Chinese Consumers' Purchase Intention of Digital Music Using an Extended Model of Goal-directed Behavior (EMGB): Focused on Mobile Digital Music," *The Journal of the Korea Contents Association*, vol. 20, no. 9, pp. 332–343, 2020.

[7] K. Negus, "From creator to data: the post-record music industry and the digital conglomerates," *Media Cult Soc*, vol. 41, no. 3, pp. 367–384, 2019.

[8] K. C. H. Kim, "The impact of blockchain technology on the music industry," *International journal of advanced smart convergence*, vol. 8, no. 1, pp. 196–203, 2019.

[9] I. B. Gorbunova, "Music computer technologies in the perspective of digital humanities, arts, and researches," *Opción: Revista de Ciencias Humanas y Sociales*, no. 24, pp. 360–375, 2019.

[10] A. Danielsen and Y. Kjus, "The mediated festival: Live music as trigger of streaming and social media engagement," *Convergence*, vol. 25, no. 4, pp. 714–734, 2019.

[11] T. Roshandel Arbatani, A. Omidi, and E. Norouzi, "Music Industry in the Age of Modern Technologies: Presenting Innovative Strategies for Digital Music Distribution in Iran," *Journal of Culture-Communication Studies*, vol. 21, no. 52, pp. 277–308, 2020.

[12] H. Supiarza and I. Sarbeni, "Teaching and learning music in digital era: creating keroncong music for gen z students through interpreting poetry," *Harmonia: Journal of Arts Research and Education*, vol. 21, no. 1, pp. 123–139, 2021.

[13] T. Shen *et al.*, "Peia: Personality and emotion integrated attentive model for music recommendation on social media platforms," in *Proceedings of the AAAI conference on artificial intelligence*, 2020, pp. 206–213.

[14] Q. Zhang and K. Negus, "East Asian pop music idol production and the emergence of data fandom in China," *International Journal of Cultural Studies*, vol. 23, no. 4, pp. 493–511, 2020.

[15] D. Calderón-Garrido, J. Gustems-Carnicer, and X. Carrera, "Digital technologies in music subjects on primary teacher training degrees in Spain: Teachers' habits and profiles," *International Journal of Music Education*, vol. 38, no. 4, pp. 613–624, 2020.

[16] J. Kwak, K. Anderson, and K. O'Connell Valuch, "Findings from a prospective randomized controlled trial of an individualized music listening program for persons with dementia," *Journal of applied gerontology*, vol. 39, no. 6, pp. 567–575, 2020.

[17] C. De Beukelaer and A. J. Eisenberg, "Mobilising African music: how mobile telecommunications and technology firms are transforming African music sectors," *Journal of African Cultural Studies*, vol. 32, no. 2, pp. 195–211, 2020.

[18] A. E. Brown, K. Donne, P. Fallon, and R. Sharpley, "From headliners to hangovers: Digital media communication in the British rock music festival experience," *Tour Stud*, vol. 20, no. 1, pp. 75–95, 2020.

[19] R. Daniel, "Digital disruption in the music industry: The case of the compact disc," *Creative Industries Journal*, vol. 12, no. 2, pp. 159–166, 2019.

[20] R. Guichardaz, L. Bach, and J. Penin, "Music industry intermediation in the digital era and the resilience of the Majors' oligopoly: the role of transactional capability," *Ind Innov*, vol. 26, no. 7, pp. 843–869, 2019.

[21] Q. Zhang and K. Negus, "Stages, platforms, streams: The economies and industries of live music after digitalization," *Popular Music and Society*, vol. 44, no. 5, pp. 539–557, 2021.

[22] M. Miller, D. Fürst, H. Hauptmann, D. A. Keim, and M. El-Assady, "Augmenting digital sheet music through visual analytics," in *Computer Graphics Forum*, Wiley Online Library, 2022, pp. 301–316.

[23] D. Kaimann, I. Tanneberg, and J. Cox, "'I will survive': Online streaming and the chart survival of music tracks," *Managerial and Decision Economics*, vol. 42, no. 1, pp. 3–20, 2021.

[24] O. Sesigür, "How to approach collecting music on streaming services," *Interactions: Studies in Communication & Culture*, vol. 11, no. 1, pp. 65–74, 2020.

[25] C. Peters, "Acquiring PDF scores for the music library: a progress report," *Music Reference Services Quarterly*, vol. 22, no. 3, pp. 131–144, 2019.

[26] J. Ilan, "Digital street culture decoded: Why criminalizing drill music is street illiterate and counterproductive," *Br J Criminol*, vol. 60, no. 4, pp. 994–1013, 2020.

[27] B. A. Morgan, "Revenue, access, and engagement via the in-house curated Spotify playlist in Australia," *Popular Communication*, vol. 18, no. 1, pp. 32–47, 2020.

[28] G. B. V. de Melo, A. F. Machado, and L. R. de Carvalho, "Music consumption in Brazil: an analysis of streaming reproductions," *PragMATIZES-Revista Latino-Americana de Estudos em Cultura*, vol. 10, no. 19, pp. 141–169, 2020.