Transactions on Artificial Intelligence
https://www.sciltp.com/journals/tai

*Review*

# Recent Advances in Artificial Intelligence for Music Education

Guo Yu [1], Guanyi Zhao [2] and Zhengjie Yang [3,*]

[1] J Major Music Studio, Pymble, NSW 2073, Australia
[2] Department of Computer Science, City University of Hong Kong, Hong Kong SAR, China
[3] Hong Kong Generative AI Research and Development Center, The Hong Kong University of Science and Technology, Hong Kong SAR, China
* Correspondence: zhengjie.yang@sydney.edu.au

**How To Cite**: Yu, G.; Zhao, G.; Yang, Z. Recent Advances in Artificial Intelligence for Music Education. *Transactions on Artificial Intelligence* **2026**, *2*(1), 39–53. https://doi.org/10.53941/tai.2026.100004

Received: 16 January 2026
Revised: 2 February 2026
Accepted: 25 February 2026
Published: 27 February 2026

**Abstract:** In this survey, we present a focused analysis of recent advances in Artificial Intelligence (AI) for music education, structured around four key pedagogical domains: learning and practicing, assessment, creation, and teaching. First, for learning and practicing, we examine AI-driven tools for personalized skill acquisition, including intelligent instruments and adaptive systems that provide real-time feedback. Second, for assessment, we review progress in the automated assessment of musical performance, moving beyond pitch and rhythm to more nuanced expressive qualities. Third, for creation, we investigate innovations in AI-augmented composition, where generative models act as collaborative partners and creative catalysts for students. Finally, for teaching, we explore systems for teacher empowerment, such as automated resource generation and learning analytics dashboards. The analysis highlights the transformative role of Deep Learning (DL) and Generative AI (GAI) in each domain while critically discussing persistent technical limitations and emerging ethical concerns. By synthesizing interdisciplinary developments, this survey aims to chart the current frontier and inform future research at the intersection of AI technology and music pedagogy.

**Keywords:** music education; Artificial Intelligence; human-AI collaboration

## 1. Introduction

Music education, fundamental to cognitive and cultural development, has long been constrained by the inherent limitations of one-to-one instruction, particularly in providing scalable, personalized, and immediate feedback. The development of Artificial Intelligence (AI) technologies offers a paradigm-shifting opportunity to address these persistent challenges [1–3]. This survey presents a focused analysis of the recent advances in AI for music education, a period distinguished by the advancement of Machine Learning (ML) [4], Deep Learning (DL) [5] and the disruptive emergence of Generative AI (GAI) [6,7]. We structure our survey not by technological taxonomy, but by the core pedagogical activities that define music learning, offering a domain-oriented framework to assess the progress and impact of AI in music education, which is both educator-friendly and technician-friendly.

While prior reviews have cataloged applications or focused on specific technical subsystems [8–10], there are still limited synthesized analyses for music education from the perspective of educational workflow, which is friendly to educators who are not familiar with AI technologies. To address this, in this survey, we organize recent research into four key pedagogical domains. First, in *learning and practicing*, we examine AI-driven tools for personalized skill acquisition, such as intelligent instruments [11,12] and adaptive systems that provide real-time feedback [13,14]. Second, in *assessment*, we review progress in the automated evaluation of musical performance, which is increasingly moving beyond basic pitch and rhythm analysis toward more nuanced assessments of expression and technique. Third, in *creation*, we investigate innovations in AI-augmented composition, where generative models are transitioning from novel tools to collaborative partners that can act as creative catalysts for students. Finally, in *teaching*, we explore systems designed for teacher empowerment, including automated resource generation [15,16] and learning analytics dashboards [17,18] that inform pedagogical decision-making.

The primary goal of this survey is to chart the current frontier at the intersection of AI and music pedagogy by synthesizing interdisciplinary developments in recent years. Our analysis critically highlights the transformative role of DL and GAI in each domain while also discussing persistent technical limitations (e.g., data scarcity and model interpretability) and emerging ethical concerns (e.g., algorithmic bias, copyright implications, and the redefined role of the human teacher). Moreover, we aim to provide future research directions, encouraging developments that are not only technologically innovative but also pedagogically sound and ethically considered, so as to ultimately enhance the human experience of musical learning and teaching.

The remainder of this paper is organized as follows. Section 2 provides a concise overview of the key AI technologies that are utilized in recent applications. Section 3 forms the core of the survey, detailing recent advances across four key pedagogical domains: personalized skill acquisition, automated assessment, AI-augmented creation, and teacher empowerment, each followed by a critical discussion of future prospects and challenges. Finally, Section 4 concludes by summarizing the current state of music education, where AI plays an important role in the service of musical and educational growth.

## 2. Overview of AI Technologies in Music Education

The recent breakthroughs in applying AI to music education are fundamentally driven by the utilization of core AI technologies, particularly DL [5], GAI [7], and Multimodal Large Language Models (MLLMs) [19]. These advancements collectively form the foundation for modern intelligent music education systems, enabling them to perceive, understand, evaluate, and even generate musical contents. The representative works of AI technologies in music education are listed in Table 1.

**Table 1.** AI technologies used in music education.

| AI Technologies | Typical Use Cases | Description | Representative References |
|---|---|---|---|
| Classical & feature-based ML | Performance analysis and similarity modeling | Early machine learning approaches relying on handcrafted features for classification, transcription, and similarity estimation. | [20–22] |
| Deep neural networks (CNNs, RNNs) | Audio, visual, and temporal performance analysis | Deep learning models for pitch estimation, posture recognition, and temporal modeling of expressive performance. | [23–25] |
| Transformer-based models | Music understanding and assessment | Large-scale pre-trained models for symbolic and audio music representation learning and performance understanding. | [26–29] |
| Generative models (GAN, VAE, Diffusion) | Music generation and creative support | Generative approaches for composition, accompaniment, and creative exploration in educational contexts. | [30–32] |
| Adaptive & interactive AI (RL, Multimodal, LLMs) | Personalized learning, feedback, and orchestration | Interactive systems combining reinforcement learning, multimodal perception, and language models to support learning and teaching. | [16,33–36] |

### *2.1. Evolution from Rule-Based to Data-Driven Paradigms*

Early technologies in music education rely predominantly on rule-based expert systems and traditional Music Information Retrieval (MIR) methods [20,37]. These systems provide feedback based on predefined logical rules applied to features like pitch and rhythm, but they lack flexibility and struggle with complex dimensions such as musical expression and emotion [38,39]. The shift to a data-driven paradigm, powered by ML and DL has become predominant recently. By training on large-scale datasets of music audio and symbolic data (e.g., MIDI [40] and sheet music), ML models can automatically learn complex patterns and feature representations inherent in music, enabling more precise and nuanced analysis and interaction [21].

### *2.2. Deep Learning in Music Information Retrieval*

DL has significantly expanded and refined the capabilities of MIR, which is crucial for automated assessment and feedback in the following three aspects:

(1) Audio Signal Modeling: The combination of Convolutional Neural Networks (CNNs) [23,41] and Recurrent Neural Networks (RNNs) [42,43] effectively processes the time-frequency characteristics of musical audio, achieving near-human performance in tasks such as instrument recognition, onset detection, and chord recognition.

(2) Music Representation Learning: DL architectures like the Transformer [44] enable deep modeling of musical

sequences (e.g., MIDI event sequences), capturing long-range dependencies between notes, which are vital for tasks like melody generation and style analysis [26,27].

(3) Pre-trained Audio Foundation Models: DL models such as MERT [28] and Jukebox [30] learn general and semantically rich representations of musical audio through self-supervised pre-training on vast amounts of unlabeled audio data. These representations serve as powerful feature extractors for downstream tasks (e.g., emotion recognition, technique assessment), significantly boosting performance in few-shot learning scenarios.

### *2.3. Generative AI and Music Foundation Models*

GAI [7], particularly diffusion models [45–47] and autoregressive language models [44,48] applied to music, represents a revolutionary advance in recent years. These models are able to analyze and create music, opening new avenues for creative pedagogy and music content generation. (1) Music Generation Models: MusicLM [31] and MusicGen [32] can generate high-quality and coherent musical segments from text descriptions (e.g., "a cheerful jazz piano piece") or audio prompts. Their core innovation involves discretizing music into token sequences and leveraging large-scale language modeling techniques for generation; (2) Symbolic Music Generation and Editing: MuseNet [49], MusicBERT [26], and MidiBERT [27] specialize in the generation, in-painting, and variation of symbolic music (MIDI). They have direct applications in music theory instruction, creative inspiration, and generating personalized exercises.

### *2.4. Multimodal Large Language Models in Music*

Modern intelligent music education systems are increasingly moving towards multimodal integration to provide a more immersive and holistic learning experience. (1) Audio-Visual Fusion: By incorporating Computer Vision (CV) [50], systems can use cameras to capture a performer's gestures, fingering, posture, and facial expressions. Pose estimation algorithms can assess technique (e.g., violin bowing and piano hand position), while expression recognition can infer a performer's emotional engagement, adding another dimension to expressivity assessment [51,52]; (2) Cross-Modal Understanding and Generation: MLLMs can correlate music, text, images, and even video [53,54]. For instance, generating imagery based on a musical piece or improvising a melody inspired by a painting provides tools for innovative and cross-art form pedagogy.

## 3. Pedagogical Framework and Taxonomy

Leveraging AI in music education is moving beyond conceptual frameworks into tangible applications that reshape core pedagogical activities. In this section, we structure the pedagogical methodology around the foundational cycle of (1) learning and practice; (2) assessment; (3) creation; and (4) teaching. The proliferation of DL and GAI has been the primary catalyst, enabling music education systems to move from passive analysis to interactive, adaptive, and co-creative partnerships. Our taxonomy is shown in Figure 1. The representative works of the four pedagogical domains are listed in Table 2.

**Table 2.** Core pedagogical domains in music education.

| Pedagogical Domain | Application Focus | Description | Representative References |
|---|---|---|---|
| Learning & practicing | Adaptive and multimodal practice support | AI-driven systems that provide personalized practice guidance through real-time feedback, multimodal sensing, and adaptive scaffolding to support skill acquisition and self-regulated learning. | [11,13,14,55] |
| Assessment | Automated performance assessment and benchmarking | AI-based approaches for evaluating musical performance quality, accuracy, and expressivity, as well as benchmarking and validation frameworks for assessment models. | [29,56–59] |
| Creation | AI-assisted and pedagogically adaptive generation | Generative models that support music composition and creative exploration, with mechanisms for controlling musical structure or difficulty to meet pedagogical and inclusive learning goals. | [30–32,60] |
| Teaching | Analytics, orchestration, and pedagogical discourse | AI tools and conceptual frameworks that support teachers through learning analytics, classroom orchestration, and broader ethical, policy, and pedagogical considerations. | [17,18,36,61] |

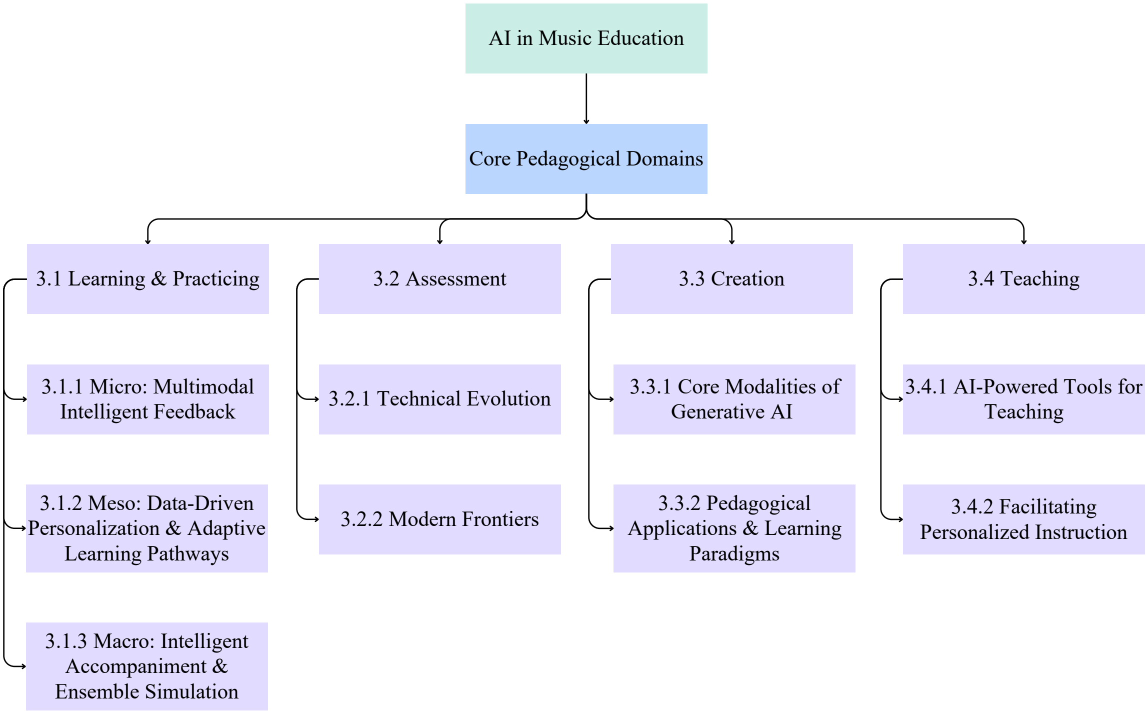


**Figure 1.** Our taxonomy. We structure our survey by four core pedagogical activities (Sections 3.1–3.4) in music education, which is both educator-friendly and technician-friendly.

### *3.1. Learning and Practicing*

The domain of musical skill acquisition has been fundamentally reshaped by AI, moving beyond the paradigm of passive and repetitive practice towards a model of interactive and intelligent partnership. Contemporary systems no longer merely judge a performance after its completion. They now engage in real-time dialog with the learner, offering personalized guidance that adapts to individual needs and progress. This evolution is powered by a confluence of sophisticated audio-visual perception, data-driven learner modeling, and adaptive logic, which effectively scales key functions of an expert tutor to provide strong support in music technical development [62,63]. In the following, we will start analyzing the shift from passive tools to interactive partners. Then, we will analyze key stages of learning from micro to macro scopes. (1) Micro: to provide real-time corrective intervention during practice; (2) Meso: to provide a long-term learning process; (3) Macro: to utilize the learned skills in a real music context.

The traditional model of musical skill acquisition relies heavily on guided and deliberate practice, where expert feedback corrects errors and refines technique. However, this model faces significant challenges: during solitary practice, learners often lack the skills to self-diagnose issues, leading to the unconscious reinforcement of mistakes in pitch, rhythm, and posture [64,65]. Although personalized expert instruction is essential, it is difficult to disseminate due to cost and expert availability, resulting in educational inequities [66]. On the other hand, early technological aids, categorized as Computer-Assisted Instruction (CAI), offer only partial solutions. Tools like score-following MIDI players and pre-recorded lessons are passive and non-adaptive. They could deliver content but could not listen, analyze, or provide responsive feedback [67]. As a result, they address access but fail to replicate the interactive diagnostic loop, which is crucial for learning.

Given the above traditional methods, the current paradigm shift is driven by the utilization of ubiquitous sensors (smartphone microphones/cameras) and advanced ML models. This enables intelligent music education systems to actively perceive and analyze performance. The field has thus evolved from static tools to Intelligent Tutoring Systems (ITS) that function as always-available practice partners, scaling the core analytical functions of a tutor to provide consistent and formative feedback [68,69]. This marks a fundamental transition from one-way instruction to dynamic and interactive pedagogical partnerships.

#### 3.1.1. Micro: Multimodal Intelligent Feedback

Multimodal intelligent feedback establishes the foundation of music pedagogy, focusing on perception and diagnosis. By integrating audio and visual analysis, these systems emulate a teacher's core function of listening and watching, addressing the immediate question of "What just went wrong?" to provide micro-level, real-time, and corrective intervention.

*Audio Analysis:* Modern systems employ sophisticated DL architectures to move beyond basic pitch and rhythm detection. CNNs trained on spectrograms or other time-frequency representations are highly effective for fine-grained feature extraction. This helps identify subtle timbral qualities (e.g., bow noise and pressure in string playing, and articulation clarity in wind instruments), which is crucial for technique improvement [24,70]. What's more, models can assess intonation in non-fixed-pitch contexts (e.g., expressive vibrato and microtonal adjustments) and the evenness of fast passages, the precision of which is unattainable by rule-based systems.

*Visual Analysis:* The visual component provides critical data on the physical execution underlying the sound. Using pose estimation (e.g., OpenPose [71] and BlazePose [72]) and keypoint tracking algorithms, systems can quantitatively assess posture and technique. For string players, this includes measuring bow angle, contact point relative to the bridge, and right-hand flexibility [73]. For pianists, vision systems analyze hand posture, finger curvature, and wrist elevation, which are directly linked to efficiency and injury prevention. In vocal pedagogy, facial action unit analysis can identify excessive jaw or lip tension that impedes resonance. Research on clarinet embouchure [74] demonstrates how specific visual features correlate strongly with expert ratings of technique quality.

*Sensor Fusion and Causal Feedback:* The true pedagogical power emerges from the sensor fusion of audio and visual streams. By correlating a specific physical gesture (e.g., a collapsed wrist) with its immediate acoustic consequence (e.g., a weak and uneven tone), the system can move from describing what is wrong to hypothesizing why. This enables causal feedback (e.g., "Your intonation is flat because your bow is slipping too far from the bridge".) which is fundamentally more actionable for the learner than a simple "You are out of tune". alert. This multimodal correlation is an active research area, which aims to build computational models of the biomechanical-acoustic chain in music performance [25,51,75].

#### 3.1.2. Meso: Data-Driven Personalization and Adaptive Learning Pathways

Data-driven personalization and adaptive learning pathways advance the planning and optimization. It moves from moment-to-moment feedback to the intelligent orchestration of the long-term learning process. By modeling the learner's state and dynamically sequencing practice, it addresses the strategic question of "How to practice most effectively next?", marking an evolution from error correction to personalized coaching. Such a kind of music pedagogy transforms from a largely linear and one-size-fits-all curriculum into a responsive, personalized, and strategic partnership.

*Reinforcement Learning for Optimized Task Sequencing:* The application of Reinforcement Learning (RL) formalizes pedagogical planning as a sequential decision-making problem. Here, the AI tutor (agent) learns a policy for selecting practice tasks (actions), such as specific exercises or repertoire excerpts, based on the learner's current state. The goal is to maximize cumulative long-term progress, using performance outcomes as a reward signal [33,55]. A sight-reading tutor leverages this RL framework to determine not only what to practice but also when and how to introduce concepts [34]. This enables sophisticated strategies (e.g., pre-teaching foundational sub-skills, dynamically adjusting difficulty, and scheduling review based on predictive models of memory decay), thereby constructing a highly optimized and personalized practice schedule.

*Skill Decomposition and Algorithmic Deliberate Practice:* AI enables a granular and data-driven approach to deliberate practice. Through MIR and pattern recognition, complex pieces can be algorithmically decomposed into constituent micro-skills, such as specific arpeggio patterns or polyrhythms. Once a weakness is identified, the system can generate targeted and looped exercises that isolate and vary that specific challenge. For example, attribute-aware practice models can transpose it, alter its tempo, or modify its articulation [76,77]. Furthermore, GAI can create musically coherent variations of these micro-skills, preventing boredom and promoting skill transfer [78]. This automates the design of focused practice regimens, shifting the burden of exercise creation from the learner to the intelligent system.

#### 3.1.3. Macro: Intelligent Accompaniment and Ensemble Simulation

Intelligent accompaniment and ensemble simulation elevate technological affordance to the contextualization and application level. It transcends isolated technical drills by placing skill development in simulated and authentic socio-musical contexts (interacting with an accompanist or a virtual ensemble). This addresses the critical problem of skill transfer, fostering musicianship and collaborative competency within a holistic performance environment.

*Responsive and Adaptive Accompaniment:* AI enables responsive accompaniment systems to go beyond static backing tracks. Employing real-time score-following and generation algorithms, a virtual accompanist listens to the student's performance, analyzes tempo, rhythm, and expression, and dynamically adjusts its playback in real-time [35,79]. This involves musically responding to the performer's expressive fluctuations and errors while maintaining harmonic coherence. Such systems train the core ensemble skill of mutual adaptation. Specifically, they teach students to listen and adjust within a flexible musical framework rather than to follow a rigid tempo [80], thereby building foundational competency for real-world collaborative performance.

*Virtual Ensemble Simulation and Practice:* AI also enables full virtual ensemble simulation. For a student practicing a chamber music part, the system can generate other voices in real-time. Using multi-track music generation techniques, these AI-generated parts are harmonically and rhythmically coherent and responsive to the lead performer's tempo and phrasing [81]. This provides an on-demand rehearsal environment for developing ensemble awareness, intonation blending, and collaborative interpretation, which is historically dependent on gathering multiple players [35]. AI also stimulates the development of systems for practicing with isolated ensemble tracks, further enhancing score-reading and auditory skills.

#### 3.1.4. Discussion

Although existing AI-powered methods stimulate the advancement of music learning and practicing, significant challenges still persist. (1) Expressivity gap: while systems are good at evaluating technical accuracy, they usually fail to assess musical expression, phrasing, and emotional intent (the subjective core of musicality) [82, 83]; (2) Black-box problem: the unexplainable decision-making of complex models undermines pedagogical trust, which leads to research in Explainable AI (XAI) [84,85] aimed at generating feedback that is both accurate and interpretable; (3) Finally, prevailing systems exhibit a pronounced cultural and stylistic bias, being predominantly trained on Western classical and popular music data. Therefore, developing adaptable models that serve diverse global musical traditions is necessary for equitable pedagogical impact.

### *3.2. Assessment*

The assessment of musical performance is a cornerstone of pedagogy, driving learning, providing validation, and certifying achievement. Traditionally, this process has been hindered by subjectivity, inconsistency, and the significant time investment required from expert instructors. By utilizing AI, we can transform assessment from a periodic and summative judgment into a continuous, formative, and multidimensional diagnostic process [22]. By leveraging data-driven models, AI systems aim to provide objective, immediate, and clear feedback that scores and deconstructs performances, offering actionable insights for improvement. This evolution encompasses the automation of core assessment tasks, the expansion of evaluable criteria, and a shift towards supporting the learning process itself, thereby augmenting rather than replacing the teacher's expert judgment [69].

Historically, performance assessment has been an intrinsically subjective endeavor, based on the trained ear and personal experience of the evaluator. While criteria can standardize the process to a degree, factors like rater fatigue, bias, and different pedagogical philosophies result in variability, especially in examinations or competitions [86]. The desire for objectivity and scalability pushes the development of early computational approaches, primarily focusing on MIR tasks (e.g., note-level transcription and score alignment). These systems can compare a performance to a reference score and detect wrong notes or rhythmic deviations. Nevertheless, they treat music as a binary right/wrong sequence, ignoring the nuances of dynamics, articulation, timing, and expression that define musical quality. This gap between measurable accuracy and perceived musicianship frames the core challenge for AI in assessment: moving from judging what is played to evaluating how it is played [82].

#### 3.2.1. Technical Evolution

The progression of automated assessment mirrors the broader shift in AI from hand-crafted features to learned representations. For feature-based approaches, systems rely on extracting predefined acoustic and musical features from audio signals. Such features include timing (tempo, tempo variability, and asynchrony between hands), dynamics (loudness contours), intonation (pitch deviation), and articulation (note onset/offset characteristics). Statistical models or simple machine learning classifiers such as Support Vector Machines (SVM) [56], are then trained on these features, using expert ratings as ground truth, to predict a holistic score [22]. While a step forward, these systems are limited by the comprehensiveness of the engineered features and their inability to capture complex and high-level musical concepts. In recent years, DL revolution has begun. The advent of RNNs and Transformers enables a paradigm shift. Instead of relying on pre-defined features, these DL models learn relevant representations directly from raw or lightly processed data (e.g., spectrograms or MIDI-like note sequences). They are capable of modeling the long-term temporal structure of a performance, understanding context, and capturing subtle interactions between musical elements. This allows for more accurate and nuanced predictions of both holistic scores and multi-dimensional attribute scores (e.g., separate ratings for pitch, rhythm, dynamics, and articulation) from a single performance [57,58].

3.2.2. Modern Frontiers

Current research is pushing automated assessment into the following three sophisticated and interconnected frontiers.

*Multidimensional Attribute Scoring:* Beyond a single performance score (number), state-of-the-art systems provide a diagnostic profile. For instance, a model might output that a piano performance scored 8/10 on note accuracy, 6/10 on rhythmic stability, and 9/10 on dynamic contrast. This is achieved by training models on datasets where expert graders provide separate ratings for different criteria. Such granularity transforms assessment from a judgment into a learning tool, clearly directing the student's focus to specific areas of weakness [29,59].

*Formative Feedback and Error Detection:* The most pedagogically valuable systems move beyond scoring to generate actionable and formative feedback. This involves precise note-level or phrase-level error detection and categorization. For example, a system can identify not just a wrong note, but that a passage is consistently rushed, that dynamics are not sufficiently contrasting between phrases, or that articulation is uneven. Some systems can even suggest corrective exercises and highlight the exact measures requiring remedial practice, closing the loop between assessment and learning [58,59].

*Explainable Assessment (XAI for Music):* Recently, after recognizing the black box problem, researchers have started to focus on explainable AI for assessment. The goal is to not only provide a score but also justify it in musically meaningful terms. This involves visualizing which sections of the spectrogram influenced a poor phrasing score, generating natural language explanations like "The crescendo into measure 12 was too abrupt". or aligning the performance with a graphical representation of tempo and dynamics, highlighting deviations from an expressive ideal. Explainability is key to building student and teacher trust in the system's feedback [22].

3.2.3. Discussion

Despite significant advancements, the field of AI-driven music assessment continues to confront several interconnected challenges. While systems excel at evaluating objective technical parameters, they fundamentally struggle to quantify high-level and culturally-mediated musical qualities, such as expressive depth, stylistic authenticity, and emotional communication. This challenge is compounded by the issue of context and style dependence, because models trained on specific genres (e.g., Western classical piano) often lack the adaptability to fairly evaluate performances from divergent traditions (e.g., jazz improvisation vs. blues). To transition from research prototypes to effective pedagogical tools, we can focus on designing human-AI collaboration frameworks that effectively integrate the granular and data-driven diagnostics of AI with the holistic and experiential judgment of the teacher, ensuring technology augments without disrupting expert pedagogy.

## *3.3. Creation*

The domain of music creation within education has undergone a paradigm shift with the advent of GAI. Moving beyond the role of passive tool or simple accompaniment generator, AI now acts as an active co-creator, inspiration engine, and personalized composition tutor [47,87]. This transformation lowers the threshold for utilizing GAI technologies, enabling students to explore melody, harmony, form, and style of music in an exciting way with an interactive sandbox. GAI in music education does not only focus on content generation. It reframes composition as a dynamic dialog between human intent and computational possibility, fostering creativity through collaboration, iteration, and the immediate auditory realization of ideas [88]. In this section, we explore how generative models augment and accelerate the learning of musical invention and expression.

The integration of computation into music composition has a long history. Early works [89] on algorithmic composition were largely deterministic, requiring expert programming knowledge and yielding outputs that often lacked the fluidity and coherence of human composition. The subsequent development of ML models, such as RNNs [42] and Variational Autoencoders (VAEs) [43,90], enables systems to learn statistical patterns from large corpora of music and generate novel sequences [91]. However, the true revolution arrives with the scale and architecture of modern generative models, such as Transformers and diffusion models, which can produce high-fidelity, structurally coherent, and stylistically diverse music from high-level prompts or even from nothing [31,32]. This technological leap has moved AI from a niche tool for experts to an accessible partner for learners at all levels, fundamentally altering the pedagogy of music creation.

3.3.1. Core Modalities of Generative AI in Creative Pedagogy

Generative AI facilitates music creation through several interconnected modalities, each serving distinct pedagogical functions by leveraging different data representations and interaction paradigms.

*Text-to-Music and Prompt-Based Generation:* In this modality, emerging models like MusicLM [31] and

MusicGen [32] allow students to generate original audio clips from natural language descriptions, effectively translating high-level semantic concepts into sound. The pedagogical power of this approach lies in its capacity to overcome creative inertia. It serves as a dynamic tool for idea generation and brainstorming, instantly providing musical starting points. Furthermore, it enables the exploration of timbre and orchestration, allowing students to experientially understand how different instrumental combinations realize specific emotional or descriptive concepts, without requiring proficiency in each instrument [92]. This process fosters a deeper understanding of musical semantics, because students investigate the intricate relationships between linguistic descriptors and their acoustic correlates, thereby enhancing their overall musical literacy and descriptive vocabulary [93].

*Symbolic Music Generation and Interactive Co-Creation:* Operating within the symbolic domain, such as MIDI or piano roll notation, models like MuseNet [49] and various Transformer-based architectures provide fine-grained control over discrete musical elements [60]. This modality fundamentally supports melodic and harmonic development. Students can input a short motif and direct the AI to generate variations, continuations, or stylistically coherent harmonizations, offering a practical sandbox for studying musical grammar and form. A key pedagogical application is style imitation and analysis. When AI-generated music emulates a specific composer or genre, it allows students to deconstruct and internalize stylistic elements, from characteristic chord progressions to idiomatic rhythmic patterns [89]. Ultimately, this facilitates interactive composition loops, establishing a call-and-response dynamic where students edit the AI's output, which in turn adapts to these edits. This iterative dialog teaches critical compositional decision-making and refines the student's creative intent through action and reaction [94].

*AI as a Pedagogical Content Creator for Composition:* Beyond open-ended creation, GAI is also helpful in generating structured and didactic materials, thereby augmenting the instructor's capacity for personalization. It enables the creation of customized exercises and etudes through specific prompting. For instance, an instructor can request "a 16-bar chord progression for practicing secondary dominants in B major" or "a contrapuntal two-part invention for oboe and bassoon," yielding infinite and tailored practice material that targets precise technical or theoretical skills [16, 95]. Moreover, AI significantly aids in arrangement and re-orchestration, capable of quickly generating alternative versions of a piece. For example, AI can simplify a complex orchestral score for a student ensemble or create a jazz combo adaptation of a classical theme. This demonstrates the practical craft of orchestration and fosters an understanding of how musical material can be adapted across different contexts and performance constraints [96,97].

#### 3.3.2. Pedagogical Applications and Learning Paradigms

The integration of GAI into music education facilitates transformative learning paradigms by fundamentally reshaping students' compositional engagement. Primarily, it significantly lowers technical barriers, enabling learners to bypass traditional instrumental or notational hurdles and immediately participate in high-level compositional thinking. By using textual or symbolic prompts to realize musical ideas, students can focus on developing intentionality and structural planning, while the immediate auditory feedback serves as a motivational scaffold that strengthens the connection between creative intent and sonic outcome [98,99]. Furthermore, AI enables a mode of what-if experiment and rapid sonic prototyping, transforming composition into a dynamic and hypothesis-driven process. This capability allows students to instantly test alternatives and iterate on ideas, thereby significantly accelerating the creative feedback loop from conception to audition and fostering a more exploratory and iterative approach to music-making [31,100]. Additionally, advanced interfaces that expose high-level semantic controls (e.g., for arousal, valence, or style) make abstract compositional parameters tangible and manipulable. Such a method of deconstructing the creative process through parametric exploration helps students articulate intuitive decisions and understand the causal relationships between musical elements, thereby building a more nuanced and actionable theoretical understanding [101,102].

#### 3.3.3. Discussion

While GAI offers significant pedagogical promise, its integration into music education faces several unsolved challenges that constitute critical barriers to its effective and responsible implementation. A foremost concern is the originality and authorship dilemma. Over-reliance on AI tools may reduce composition to mere prompt refinement, potentially restraining the development of the student's authentic voice and core competencies, such as internal hearing, manual notation, and structural thinking [88,92]. This ambiguity directly complicates the assessment of genuine student contributions in collaborative human-AI works. Furthermore, the prevalent lack of explainability and pedagogical transparency in black-box models presents a major instructional hurdle. Students simply receive a musical output without insight into the generative logic, preventing the translation of the result into actionable music-theoretic knowledge. Beyond issues of mechanism and skill acquisition, there exists a profound risk of stylistic bias and cultural homogenization. Since models are predominantly trained on datasets skewed towards Western canonical and popular music, they tend to reproduce these traditions, marginalizing non-Western, folk, and avant-garde expressions and thereby narrowing the educational scope [60]. The imperative future direction is to educate students not only in tool usage but also in critically evaluating outputs, understanding inherent limitations and biases, and developing the ethical dimensions of co-creation. The overall goal is to cultivate discerning and informed musicians capable of navigating a complex technological landscape, transforming a potential passive dependency into a mode of empowered and reflective artistic practice.

### *3.4. Teaching*

While much of the discourse on AI in education focuses on direct student-tool interaction, its most profound impact may lie in augmenting and empowering the human teacher. AI is emerging as a transformative force in teaching, shifting the educator's role from the primary source of information and correction to that of a facilitator, mentor, and learning architect [103]. This transition addresses long-standing pedagogical challenges: the unsustainable workload of individualized lesson planning, the difficulty of obtaining timely and granular insights into each student's progress, and the need to create equitable and personalized learning experiences in classrooms with diverse learners. By automating administrative tasks, providing data-driven insights, and generating personalized content, AI functions as a co-pilot for the teacher, freeing cognitive and temporal resources for the uniquely human aspects of teaching. That is to say, AI helps build relationships, inspire creativity, provide nuanced mentorship, and cultivate a vibrant learning culture [104]. In this section, we examine how AI is not replacing teachers but rather providing them with sophisticated tools to scale their expertise and enhance their pedagogical effectiveness.

The traditional model of the music teacher as a singular authority figure, responsible for all instruction, assessment, and motivation, faces mounting pressures. Educators commonly manage heterogeneous classrooms where students possess varying skill levels, learning paces, and musical interests [105]. Providing truly individualized feedback and lesson plans within this context is a difficult and often impossible challenge, leading to a one-size-fits-all approach. Furthermore, a significant portion of a teacher's time is consumed by administrative and preparatory labor: creating custom exercises, transcribing music, designing assessments, and tracking student progress manually [98]. These tedious tasks detract from the time available for high-quality instruction and student interaction. The advent of data-rich digital learning environments and powerful generative models presents an opportunity to alleviate these burdens. The challenge and opportunity for AI in teaching is to move beyond simple grading aids to become integrated systems that support the complex and holistic practice of pedagogy, thereby enabling teachers to reclaim their role as guides and inspirers.

#### 3.4.1. AI-Powered Tools for Enhanced Teaching Practice

AI augments teaching across several key dimensions, acting as an intelligent assistant in the instructional workflow.

*Learning Analytics and Diagnostic Dashboards:* AI systems aggregate and analyze data from student practice sessions, digital assignments, and assessments. This data is synthesized into visual analytics dashboards that provide teachers with an at-a-glance and evidence-based overview of class and individual student performance [36,61]. These dashboards can visualize (1) longitudinal progress: tracking skill development over time across multiple dimensions (e.g., sight-reading speed, rhythmic accuracy, and repertoire difficulty); (2) problem identification: highlighting common stumbling blocks for the entire class or pinpointing specific and persistent technical issues for individual students (e.g., a student consistently struggles with intonation in the upper register); and (3) engagement metrics: offering insights into practice habits, such as frequency, duration, and time-of-day patterns, which can inform motivational strategies. In this way, it transforms teacher intuition into data-informed decision-making, allowing for proactive and targeted intervention rather than reactive correction.

*AI-Assisted Content and Curriculum Development:* Generative AI models are powerful tools for on-demand pedagogical content creation, dramatically reducing preparation time. For exercise and etude generation, teachers can input parameters (e.g., skill level, key signature, technical focus, and length) to instantly generate a bank of customized technical exercises or sight-reading materials tailored to their students' needs [31]. For arrangement and adaptation, AI can quickly simplify complex scores for beginner ensembles, transpose parts for available instruments, or create practice tracks with a specific part isolated or muted, promoting accessibility and inclusion. For lesson and unit planning, AI can assist in structuring curriculum by suggesting repertoire progressions based on pedagogical goals, generating discussion questions, or creating lesson outlines that incorporate personalized activities for varied skill levels.

*Administrative Automation and Classroom Management Support:* AI can streamline non-instructional tasks that consume teachers' time. For foundational tasks like scale practice, rhythm drills, or basic theory homework, AI can provide initial corrective feedback, allowing the teacher to focus their expert commentary on higher-order musical issues during lesson time. AI-powered tools can also help draft student progress reports, summarize key points from practice data for parent-teacher conferences, or manage routine communications.

#### 3.4.2. Facilitating Personalized and Differentiated Instruction

A primary promise of AI in teaching is to make scalable differentiation a reality. By processing individual learner data, AI systems can help teachers design and manage multiple and parallel learning pathways within a single classroom. Beyond describing what is happening, advanced systems aim to suggest what to do next. A dashboard might recommend specific remedial exercises for a struggling student or propose challenge repertoires for an advanced learner, effectively acting as a teaching assistant for differentiation [106]. Moreover, AI can analyze student profiles and performance data to suggest optimal groupings for peer learning, ensemble work, or collaborative projects, ensuring productive and balanced interactions.

#### 3.4.3. Discussion

Although the integration of AI as a pedagogical partner is promising, it still introduces a matrix of complex ethical, practical, and socio-technical challenges that demand careful navigation. Foremost among these is the ethics of data and analytics, where the pervasive collection of granular student performance data raises critical concerns regarding privacy, ownership, and the potential for surveillance, risking the reduction of learning to a set of optimizable metrics and creating a culture of performance anxiety [107]. This technological intervention simultaneously challenges teacher agency and expertise. An over-reliance on algorithmic recommendations risks the deskilling of educators, undermining their diagnostic intuition and pedagogical autonomy [108]. Furthermore, the distribution of these benefits is profoundly uneven, threatening to exacerbate existing inequalities through a new digital and AI-dependent method, because access to the necessary technology, connectivity, and support depends on resources of institutions, thereby marginalizing underfunded schools [109]. Therefore, empowering educators to critically evaluate AI-generated content, interpret data dashboards, and integrate tools into their practice is a prerequisite for ethical and effective human-AI collaboration in the classroom [110].

## 4. Conclusions

In this paper, we systematically examine the recent advances in AI for music education, structuring the analysis around four core pedagogical domains (learning and practicing, assessment, creation, and teaching). We also highlight the transformative roles of DL, GAI, and MLLMs in reshaping traditional music pedagogies. AI has evolved from passive tools to interactive and collaborative partners throughout the music education life cycle (four pedagogical domains). Developments in assessment models inform adaptive practices and feedback systems, while generative models for creation are increasingly leveraged as learning and teaching aids. In the meantime, learning analytics tools simultaneously shape assessment and instructional decision-making. These cross-domain interactions indicate that advances in one domain often propagate methodological and pedagogical impacts across others, underscoring the importance of considering AI-enabled music education as an interconnected ecosystem rather than a set of isolated applications.

However, persistent challenges remain, including the difficulty of evaluating subjective musical expression, the black-box problem of complex models, cultural and stylistic biases in training data, ethical concerns related to data privacy, and the redefinition of teacher-student roles. Future research should prioritize interdisciplinary integration, reconciling technological innovation with pedagogical soundness and ethical responsibility, such as advancing XAI for transparency, expanding diverse training datasets to foster inclusivity, and designing human-AI collaboration frameworks that augment human expertise. Ultimately, the goal of AI in music education is to enrich the holistic experience of musical learning and teaching, and cultivate creative, discerning, and culturally aware musicians in an increasingly technological landscape.

## Author Contributions

G.Y.: Conceptualization, investigation, writing, and revision; G.Z.: Conceptualization, investigation, writing, and revision; Z.Y.: Conceptualization, investigation, writing, and revision. All authors have read and agreed to the published version of the manuscript.

## Funding

This research received no external funding.

## Data Availability Statement

Not applicable.

## Conflicts of Interest

The authors declare no conflict of interest.

## Use of AI and AI-Assisted Technologies

During the preparation of this work, the authors used ChatGPT to check the grammar. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

## References


1. Achiam, J.; Adler, S.; Agarwal, S.; et al. GPT-4 Technical Report. *arXiv* **2023**, arXiv:2303.08774.
2. Touvron, H.; Martin, L.; Stone, K.; et al. Llama 2: Open Foundation and Fine-Tuned Chat Models. *arXiv* **2023**, arXiv:2307.09288.
3. Liu, A.; Feng, B.; Xue, B.; et al. DeepSeek-V3 Technical Report. *arXiv* **2024**, arXiv:2412.19437.
4. Alpaydin, E. *Introduction to Machine Learning*, 4th ed.; MIT Press: Cambridge, MA, USA, 2020.
5. Goodfellow, I.; Bengio, Y.; Courville, A.; et al. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
6. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; et al. Generative Adversarial Nets. In Proceedings of the 28th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; Volume 27, pp. 2672–2680.
7. Feuerriegel, S.; Hartmann, J.; Janiesch, C.; et al. Generative AI. *Bus. Inf. Syst. Eng.* **2024**, *66*, 111–126.
8. Li, P.P.; Wang, B. Artificial Intelligence in Music Education. *Int. J. Hum.-Comput. Interact.* **2024**, *40*, 4183–4192.
9. Zhang, Y.; Fen, B.W.; Zhang, C.; et al. Transforming Music Education Through Artificial Intelligence: A Systematic Literature Review on Enhancing Music Teaching and Learning. *Int. J. Interact. Mob. Technol.* **2024**, *18*, 76.
10. Zhang, H. Artificial Intelligence in Music Education: A Scoping Review of Practices, Strategies, and Challenges. *J. Adv. Res. Educ.* **2025**, *4*, 69–82.
11. Ruan, S. AI Harmonica: A Smart Electronic Harmonica for Music Learning and Co-Creativity. In Proceedings of the 39th Conference on Neural Information Processing Systems (NeurIPS 2025), San Diego, CA, USA, 2–7 December 2025.
12. Blanchard, L.; Naseck, P.; Egozy, E.; et al. Developing Symbiotic Virtuosity: AI-Augmented Musical Instruments and Their Use in Live Music Performances. In *An MIT Exploration of Generative AI*; MIT Press: Cambridge, MA, USA, 2024.
13. Shu, X.; Shi, L.; Cheng, J.; et al. FretMate: ChatGPT-Powered Adaptive Guitar Learning Assistant. In Proceedings of the 30th International Conference on Intelligent User Interfaces, Cagliari, Sardinia, Italy, 24 –27 March 2025; pp. 715–726.
14. Blanco, A.D.; Tassani, S.; Ramirez, R. Real-Time Sound and Motion Feedback for Violin Bow Technique Learning: A Controlled, Randomized Trial. *Front. Psychol.* **2021**, *12*, 648479.
15. Cheng, D.; Qu, X. Intelligent Generation and Optimization of Resources in Music Teaching Reform Based on Artificial Intelligence and Deep Learning. *Sci. Rep.* **2025**, *15*, 30051.
16. Holster, J. Augmenting Music Education Through AI: Practical Applications of ChatGPT. *Music Educ. J.* **2024**, *110*, 36–42.
17. Possaghi, I.; Vesin, B.; Zhang, F.; et al. Integrating Multi-Modal Learning Analytics Dashboard in K-12 Education: Insights for Enhancing Orchestration and Teacher Decision-Making. *Smart Learn. Environ.* **2025**, *12*, 53.
18. Cabral, L.; Pinto, R.; Goncalves, G. AI-Powered Learning Analytics Dashboards: A Systematic Review of Applications, Techniques, and Research Gaps. *Discov. Educ.* **2025**, *4*, 525.
19. Yin, S.; Fu, C.; Zhao, S.; et al. A Survey on Multimodal Large Language Models. *Natl. Sci. Rev.* **2024**, *11*, nwae403.
20. Downie, J.S. Music Information Retrieval. *Annu. Rev. Inf. Sci. Technol.* **2003**, *37*, 295–340.
21. Benetos, E.; Dixon, S.; Duan, Z.; et al. Automatic Music Transcription: An Overview. *IEEE Signal Process. Mag.* **2018**, *36*, 20–30.
22. Berenzweig, A.; Logan, B.; Ellis, D.P.; et al. A Large-Scale Evaluation of Acoustic and Subjective Music-Similarity Measures. *Comput. Music J.* **2004**, *28*, 63–76.
23. Pons, J.; Serra, X. Musicnn: Pre-Trained Convolutional Neural Networks for Music Audio Tagging. In Proceedings of the 20th International Society for Music Information Retrieval Conference (ISMIR), Delft, The Netherlands, 4–8 November 2019.
24. Kim, J.W.; Salamon, J.; Li, P.; et al. CREPE: A Convolutional Representation for Pitch Estimation. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 161–165.
25. Park, S.H.; Park, Y.H. Audio-Visual Tensor Fusion Network for Piano Player Posture Classification. *Appl. Sci.* **2020**, *10*, 6857.
26. Zeng, M.; Tan, X.; Wang, R.; et al. MusicBERT: Symbolic Music Understanding with Large-Scale Pre-Training. In Proceedings of the Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, Association for Computational Linguistics, Online, 1–6 August 2021; pp. 791–800.

27. Chou, Y.H.; Chen, I.C.; Ching, J.; et al. MidiBERT-Piano: Large-Scale Pre-Training for Symbolic Music Classification Tasks. *J. Creat. Music Syst.* **2024**, *8*. https://doi.org/10.5920/jcms.1064
28. Yizhi, L.; Yuan, R.; Zhang, G.; et al. MERT: Acoustic Music Understanding Model with Large-Scale Self-Supervised Training. In Proceedings of the Twelfth International Conference on Learning Representations, Vienna, Austria, 7–11 May 2024.
29. Zhang, H.; Liang, J.; Dixon, S. From Audio Encoders to Piano Judges: Benchmarking Performance Understanding for Solo Piano. In Proceedings of the 25th International Society for Music Information Retrieval Conference, San Francisco, CA, USA, 10–14 November 2024.
30. Dhariwal, P.; Jun, H.; Payne, C.; et al. Jukebox: A Generative Model for Music. *arXiv* **2020**, arXiv:2005.00341.
31. Agostinelli, A.; Denk, T.I.; Borsos, Z.; et al. MusicLM: Generating Music from Text. *arXiv* **2023**, arXiv:2301.11325.
32. Copet, J.; Kreuk, F.; Gat, I.; et al. Simple and Controllable Music Generation. *Adv. Neural Inf. Process. Syst.* **2023**, *36*, 47704–47720.
33. Li, K. Policy Gradient Reinforcement Learning for Personalized Pathways in Intelligent Music Education Systems. In Proceedings of the 2025 2nd International Conference on Informatics Education and Computer Technology Applications, Kuala Lumpur, Malaysia, 17–19 January 2025; pp. 128–132.
34. Yahya, K.; Fard, P.R. Computational Model of Music Sight Reading: A Reinforcement Learning Approach. *arXiv* **2010**, arXiv:1007.0546.
35. Cancino-Chac'on, C.; Peter, S.; Hu, P.; et al. The ACCompanion: Combining Reactivity, Robustness, and Musical Expressivity in an Automatic Piano Accompanist. In Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23), Macao, China, 19–25 August 2023; pp. 5779–5787.
36. Sedrakyan, G.; Malmberg, J.; Verbert, K.; et al. Linking Learning Behavior Analytics and Learning Science Concepts: Designing a Learning Analytics Dashboard for Feedback to Support Learning Regulation. *Comput. Hum. Behav.* **2020**, *107*, 105512.
37. Schedl, M.; Gomez, E.; Urbano, J.; et al. Music Information Retrieval: Recent Developments and Applications. *Found. Trends Inf. Retr.* **2014**, *8*, 127–261.
38. Morreale, F.; Martinez-Ramirez, M.A.; Masu, R.; et al. Reductive, Exclusionary, Normalising: The Limits of Generative AI Music. *Trans. Int. Soc. Music Inf. Retr.* **2025**, *8*, 300–312.
39. Arcos, J.L. Music and Similarity Based Reasoning. In *Soft Computing in Humanities and Social Sciences*; Springer: Berlin, Heidelberg, 2011; pp. 467–478.
40. Rothstein, J. *MIDI: A Comprehensive Introduction*, 2nd ed.; A-R Editions, Inc.: Middleton, WI, USA, 1995; Volume 7.
41. LeCun, Y.; Bottou, L.; Bengio, Y.; et al. Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE* **2002**, *86*, 2278–2324.
42. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780.
43. Roberts, A.; Engel, J.; Raffel, C.; et al. A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; pp. 4364–4373.
44. Vaswani, A.; Shazeer, N.; Parmar, N.; et al. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017, Volume 30.
45. Sohl-Dickstein, J.; Weiss, E.; Maheswaranathan, N.; et al. Deep Unsupervised Learning Using Nonequilibrium Thermodynamics. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 2256–2265.
46. Ho, J.; Jain, A.; Abbeel, P. Denoising Diffusion Probabilistic Models. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 6840–6851.
47. Chen, J.; Xue, W.; Tan, X.; et al. FastSAG: Towards Fast Non-Autoregressive Singing Accompaniment Generation. In Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, Jeju, South Korea, 3–9 August 2024; pp. 7618–7626.
48. Bengio, Y.; Ducharme, R.; Vincent, P.; et al. A Neural Probabilistic Language Model. *J. Mach. Learn. Res.* **2003**, *3*, 1137–1155.
49. Payne, C. MuseNet. Available online: https://openai.com/blog/musenet (accessed on 12 January 2026).
50. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*.
51. Yoo, S.J.; Shrestha, S.; Muresanu, I.; et al. VioPose: Violin Performance 4D Pose Estimation by Hierarchical Audiovisual Inference. In Proceedings of the 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Tucson, AZ, USA, 26 February–6 March 2025; pp. 1–12.
52. Provenzale, C.; Di Tommaso, F.; Di Stefano, N.; et al. Real-Time Visual Feedback Based on MIMUs Technology Reduces Bowing Errors in Beginner Violin Students. *Sensors* **2024**, *24*, 3961.
53. Mao, Z.; Zhao, M.; Wu, Q.; et al. DeepResonance: Enhancing Multimodal Music Understanding via Music-Centric Multi-Way Instruction Tuning. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, Suzhou, China, 4–9 November 2025; pp. 12926–12948.

54. Ramesh, A.; Pavlov, M.; Goh, G.; et al. Zero-Shot Text-to-Image Generation. In Proceedings of the 38th International Conference on Machine Learning, Online, 18–24 July 2021 ; pp. 8821–8831.
55. Moringen, A.; Rüttgers, S.; Zintgraf, L.; et al. Optimizing Piano Practice with a Utility-Based Scaffold. *arXiv* **2021**, arXiv:2106.12937.
56. Phanichraksaphong, V.; Tsai, W.H. Automatic Evaluation of Piano Performances for STEAM Education. *Appl. Sci.* **2021**, *11*, 11783.
57. Pati, K.A.; Gururani, S.; Lerch, A. Assessment of Student Music Performances Using Deep Neural Networks. *Appl. Sci.* **2018**, *8*, 507.
58. Wang, X.; Fitri bin Haris, M. AI-Driven Violin Performance Evaluation: A Deep Learning Approach for Automated Assessment and Personalized Feedback in Music Education. In Proceedings of the 2025 International Conference on Generative Artificial Intelligence and Digital Media, Dali, China, 14–16 March 2025; pp. 126–132.
59. Zhang, H.; Cheung, V.K.; Nishioka, H.; et al. LLaQo: Towards a Query-Based Coach in Expressive Music Performance Assessment. In Proceedings of the 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hyderabad, India, 6–11 April 2025; pp. 1–5.
60. Jin, C.; Wang, T.; Liu, S.; et al. A Transformer-Based Model for Multi-Track Music Generation. *Int. J. Multimed. Data Eng. Manag.* **2020**, *11*, 36–54.
61. Pan, Z.; Biegley, L.; Taylor, A.; et al. A Systematic Review of Learning Analytics: Incorporated Instructional Interventions on Learning Management Systems. *J. Learn. Anal.* **2024**, *11*, 52–72.
62. Chang, J.; Wang, Z.; Yan, C. MusicARLtrans Net: A Multimodal Agent Interactive Music Education System Driven via Reinforcement Learning. *Front. Neurorobot.* **2024**, *18*, 1479694.
63. Xiao, H. Personalized Learning Path Planning and Effect Verification of Music Education Information System Empowered by Artificial Intelligence. In Proceedings of the 2025 International Conference on Artificial Intelligence and Educational Systems, Beijing, China, 25–27 April 2025; pp. 443–448.
64. Rom, B.A. Scaffolding Autonomy in the Practice Room: A Mixed Methods Study Examining the Impact of Digital Scaffolds on High School String Musicians' Self-Correction and Improvement of Pitch and Rhythmic Accuracy During Independent Music Practicing. Ph.D. Thesis, The University of Nebraska-Lincoln, Lincoln, NE, USA, 2020.
65. Mateos-Moreno, D.; Garc'ıa-Perals, J.; Maxwell, T. Teaching to Practice Productively and Consciously: An Action-Research Study in One-to-One Instrumental Music Teaching. *Music Educ. Res.* **2025**, *27*, 203–214.
66. Butler, A.; Lind, V.L.; McKoy, C.L. Equity and Access in Music Education: Conceptualizing Culture as Barriers to and Supports for Music Learning. *Music Educ. Res.* **2007**, *9*, 241–253.
67. Webster, P.R. Computer-Based Technology and Music Teaching and Learning. In *Critical Essays in Music Education*; Routledge: London, UK, 2017; pp. 321–344.
68. Acquilino, A. Real-Time Computational Feedback Systems for Enriched Music Pedagogy. Ph.D. Thesis, McGill University, Montreal, QC, Canada, 2025.
69. Lu, L. AI-Powered Intelligent Music Education Systems for Real-Time Feedback and Performance Assessment. *Int. J. Inf. Commun. Technol.* **2025**, *26*, 33–47.
70. Chellamani, G.K.; N, A.; C, C.; et al. SpectroFusionNet: A CNN Approach Utilizing Spectrogram Fusion for Electric Guitar Play Recognition. *Sci. Rep.* **2025**, *15*, 16842.
71. Cao, Z.; Hidalgo, G.; Simon, T.; et al. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 172–186.
72. Bazarevsky, V.; Grishchenko, I.; Raveendran, K.; et al. BlazePose: On-Device Real-Time Body Pose Tracking. *arXiv* **2020**, arXiv:2006.10204.
73. Mizuho, Y.; Sugiura, Y. A Comparison of Violin Bowing Pressure and Position among Expert Players and Beginners. In Proceedings of the AsiaHaptics 2024, Selangor, Malaysia, 28–30 October 2024.
74. Almeida, A.; Li, W.; Schubert, E.; et al. Recording and Analysing Physical Control Variables Used in Clarinet Playing: A Musical Instrument Performance Capture and Analysis Toolbox (MIPCAT). *Front. Signal Process.* **2023**, *3*, 1089366.
75. Grosshauser, T.; Hermann, T. Sensor Fusion and Multi-Modal Feedback for Musical Instrument Learning and Teaching. In Proceedings of the Second Vienna Talk Music Acoust, Vienna, Austria, 19–21 September 2010.
76. Matsuoka, Y.; Sako, S. Attribute-Aware Deep Music Transformation for Polyphonic Music. In Proceedings of the Late Breaking Demo in the 22nd International Society for Music Information Retrieval Conference (ISMIR), Online, 7–12 November 2021.
77. Kawai, L.; Esling, P.; Harada, T. Attributes-Aware Deep Music Transformation. In Proceedings of the 21st International Society for Music Information Retrieval Conference (ISMIR), Montreal, QC, Canada, 11–16 October 2020; pp. 670–677.
78. Huo, X. Design of Automatic Generation Model for Teaching Music Based on User Differences and Transfer Learning. *Int. J. Embed. Syst.* **2024**, *17*, 62–72.
79. Nakamura, T.; Nakamura, E.; Sagayama, S. Acoustic Score Following to Musical Performance with Errors and Arbitrary Repeats and Skips for Automatic Accompaniment. In Proceedings of the 10th Sound and Music Computing Conference, Stockholm, Sweden, 30 July–2 August 2013; pp. 299–304.

80. Park, J.; Cancino-Chacon, C.; Chiruthapudi, S.; et al. Matchmaker: An Open-Source Library for Real-Time Piano Score Following and Systematic Evaluation. In Proceedings of the 26th International Society for Music Information Retrieval Conference (ISMIR 2025), Daejeon, Korea, 21–25 September 2025.
81. Wu, Y.; Gardner, J.; Manilow, E.; et al. The Chamber Ensemble Generator: Limitless High-Quality MIR Data via Generative Modeling. *arXiv* **2022**, arXiv:2209.14458.
82. Cancino-Chacon, C.E.; Grachten, M.; Goebl, W.; et al. Computational Models of Expressive Music Performance: A Comprehensive and Critical Review. *Front. Digit. Humanit.* **2018**, *5*, 25.
83. Oore, S.; Simon, I.; Dieleman, S.; et al. This Time with Feeling: Learning Expressive Musical Performance. *Neural Comput. Appl.* **2020**, *32*, 955–967.
84. Ramoneda, P.; Eremenko, V.; D'Hooge, A.; et al. Towards Explainable and Interpretable Musical Difficulty Estimation: A Parameter-Efficient Approach. In Proceedings of the 25th International Society for Music Information Retrieval Conference, San Francisco, CA, USA, 10–14 November 2024.
85. Gohel, P.; Singh, P.; Mohanty, M. Explainable AI: Current Status and Future Directions. *arXiv* **2021**, arXiv:2107.07045.
86. McPherson, G.E.; Thompson, W.F. Assessing Music Performance: Issues and Influences. *Res. Stud. Music Educ.* **1998**, *10*, 12–24.
87. Väkevä, L.; Partti, H. Generative AI as a Collaborator in Music Education: An Action-Network Theoretical Approach to Fostering Musical Creativities. *Action Crit. Theory Music Educ.* **2025**, *24*, 16–52.
88. Briot, J.P.; Hadjeres, G.; Pachet, F.D. *Deep Learning Techniques for Music Generation—A Survey*; Springer: Cham, Switzerland, 2020.
89. Cope, D. *Computer Models of Musical Creativity*; MIT Press: Cambridge, MA, USA, 2005.
90. Kingma, D.P.; Welling, M. Auto-Encoding Variational Bayes. In Proceedings of the 2nd International Conference on Learning Representations (ICLR), Banff, AB, Canada, 14–16 April 2014.
91. Chu, H.; Urtasun, R.; Fidler, S. Song from PI: A Musically Plausible Network for Pop Music Generation. *arXiv* **2016**, arXiv:1611.03477.
92. Manovich, L. *AI Aesthetics*; Strelka Press: Moscow, Russia, 2018.
93. Salganik, R.; Tu, T.; Chen, F.Y.; et al. MusicSem: A Semantically Rich Language-Audio Dataset of Organic Musical Discourse. In Proceedings of the AI for Music Workshop in the 39th Annual Conference on Neural Information Processing Systems, San Diego, CA, USA, 7 December 2025.
94. Fiebrink, R.A.; Caramiaux, B. The Machine Learning Algorithm as Creative Musical Tool. In *The Oxford Handbook of Algorithmic Music*; Oxford University Press: Oxford, UK, 2018.
95. Ramoneda, P.; Suzuki, M.; Maezawa, A.; et al. Difficulty-Aware Score Generation for Piano Sight-Reading. *arXiv* **2025**, arXiv:2509.16913.
96. Ou, L.; Zhao, J.; Wang, Z.; et al. Unifying Symbolic Music Arrangement: Track-Aware Reconstruction and Structured Tokenization. In Proceedings of the Thirty-Ninth Annual Conference on Neural Information Processing Systems, San Diego, CA, USA, 2–7 December 2025.
97. Ramoneda, P.; Parada-Cabaleiro, E.; Jeong, D.; et al. Difficulty-Controlled Simplification of Piano Scores with Synthetic Data for Inclusive Music Education. *arXiv* **2025**, arXiv:2511.16228.
98. Bauer, W.I. *Music Learning Today: Digital Pedagogy for Creating, Performing, and Responding to Music*, 2nd ed.; Oxford University Press: Oxford, UK, 2020.
99. Ruthmann, S.A.; Hebert, D.G. Music Learning and New Media in Virtual and Online Environments. In *The Oxford Handbook of Music Education*; Oxford University Press: Oxford, UK, 2012; Volume 2.
100. Fiebrink, R. Machine Learning Education for Artists, Musicians, and Other Creative Practitioners. *ACM Trans. Comput. Educ.* **2019**, *19*, 1–32.
101. Ilkka, T. *The Impact of Artificial Intelligence on Learning, Teaching, and Education*; European Union: Luxembourg, 2018.
102. Herremans, D.; Chew, E. MorpheuS: Generating Structured Music with Constrained Patterns and Tension. *IEEE Trans. Affect. Comput.* **2017**, *10*, 510–523.
103. Luckin, R.; Holmes, W. *Intelligence Unleashed: An Argument for AI in Education*; UCL Knowledge Lab.: London, UK, 2016.
104. Baker, R.S.; Inventado, P.S. Educational Data Mining and Learning Analytics. In *Learning Analytics: From Research to Practice*; Springer: New York, NY, USA, 2014; pp. 61–75.
105. Hallam, S.; Creech, A.; McQueen, H. Can the Adoption of Informal Approaches to Learning Music in School Music Lessons Promote Musical Progression? *Br. J. Music Educ.* **2017**, *34*, 127–151.
106. Van Leeuwen, A.; Janssen, J. A Systematic Review of Teacher Guidance During Collaborative Learning in Primary and Secondary Education. *Educ. Res. Rev.* **2019**, *27*, 71–89.
107. Zeide, E. The Structural Consequences of Big Data-Driven Education. *Big Data* **2017**, *5*, 164–172.
108. Williamson, B. *Big Data in Education: The Digital Future of Learning, Policy and Practice*; SAGE Publications Ltd.: London, UK, 2017.

109. Reich, J.; Ito, M.; Team, M. From Good Intentions to Real Outcomes. *Digit. Media Learn. Res. Hub.* **2017**. Available online: https://clalliance.org/publications/good-intentions-real-outcomes-equity-designlearning-technologies (accessed on 12 February 2026).
110. Cetindamar, D.; Kitto, K.; Wu, M.; et al. Explicating AI Literacy of Employees at Digital Workplaces. *IEEE Trans. Eng. Manag.* **2022**, *71*, 810–823.