ASSESSING THE ROLE OF ARTIFICIAL INTELLIGENCE IN MUSIC CREATION: COLLABORATIVE TOOLS, GENERATIVE MODELS, AND AUTHORSHIP

Anugwa Udokaku Lovelyn PhD
Department of Music, Faculty of Humanities, Ignatius Ajuru University of Education, Rumuolumeni, Port Harcourt, Nigeria.
amaudokaku@gmail.com

ABSTRACT
The integration of Artificial Intelligence (AI) in music creation has revolutionized the industry, transforming composition, production, and performance. This study explores the evolution of AI-driven music tools, highlighting the progression from early algorithmic compositions to contemporary generative models. AI has significantly enhanced music composition through advanced systems like AIVA, MuseNet, and MusicVAE, enabling the generation of melodies, harmonies, and full compositions that emulate human creativity. AI-powered production tools, such as LANDR and iZotope's Neutron, streamline mastering and mixing, democratizing access to professional-grade audio production. Additionally, AI plays a growing role in live performances, with robotic musicians and real-time interactive visuals redefining audience engagement. The study also examines state-of-the-art generative AI models, including Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), Autoregressive Models, and Diffusion Models, which enhance creative possibilities in music generation. Despite these advancements, challenges remain in authorship attribution, ethical considerations, and the evolving role of human musicians. The paper further explores AI's role in music classification, recommendation systems, and plagiarism detection. Looking ahead, AI-driven music research presents numerous opportunities, including improving AI-human collaboration, enhancing creativity through multimodal learning, and addressing ethical dilemmas related to ownership and originality. As AI continues to reshape the music landscape, understanding its capabilities, limitations, and potential future developments remains essential for both artists and technologists. This study provides a comprehensive overview of AI's impact on music, setting the stage for future research.
Keywords: *Artificial Intelligence in Music, Generative AI, Music Composition, AI Music Production, AI Live Performance, Generative Models, Music Recommendation, and AI Music Classification.*

1. Introduction

Music has been a fundamental aspect of human culture for centuries, serving as a universal language that transcends geographical and linguistic barriers. It plays a crucial role in communication, self-expression, and emotional articulation, whether through composition, performance, or listening (Barton, 2018; Mycka & Mańdziuk, 2025). With the advent and rapid advancement of artificial intelligence (AI), music has become one of the many fields where AI technologies are being explored, tested, and implemented. The integration of AI into music has evolved from initial challenges, such as representing musical data and generating simple sounds, to more complex applications, including full-scale composition, music classification, recommendation systems, and interactive creative tools (Kaliakatsos-Papakostas, Floros & Vrahatis, 2020; Ndou, Ajoodha & Jadhav, 2021). These developments have led to significant transformations in how music is created, consumed, and understood.

AI in music creation can broadly be categorized into two primary domains: Generative AI and Assistive AI. Generative AI refers to the capability of AI to create new musical content, often trained on vast datasets and employing machine learning techniques such as neural networks to generate compositions that closely resemble human-made music. This domain has seen remarkable advancements in AI-generated compositions, leading to debates on the authenticity, originality, and artistic merit of AI-generated music (Atanacković, 2024). On the other hand, Assistive AI enhances human creativity by providing tools that support music composition, arrangement, and production, optimizing workflow efficiency and expanding creative possibilities. While these AI-driven advancements have introduced new opportunities, they have also sparked concerns regarding the authenticity of AI-generated compositions, the potential displacement of human musicians, and ethical considerations surrounding authorship and copyright.

The proliferation of AI-powered music creation tools has led to polarized perspectives regarding its role in artistic expression. Some scholars and practitioners view AI as a revolutionary force, enhancing creative possibilities, democratizing music production, and providing tools that enable both amateur and professional musicians to explore new musical frontiers (Hong, 2018; Roberts, Engel, Mann et al., 2019; Moruzzi, 2020; Newman, Morris & Lee, 2023). Conversely, others argue that AI-generated music lacks the depth, intention, and emotional nuance inherent in human-created works, raising questions about the implications of AI's involvement in creative processes and the potential erosion of traditional artistic values (Mazzone & Elgammal, 2019; Zulic, 2019; Hong, Peng & Williams, 2021). The debate extends beyond mere creative considerations to include issues of intellectual property, ownership, and the impact on the professional music industry.

A key challenge in assessing AI's role in music creation is the absence of objective evaluation tools for measuring the artistic and technical quality of AI-generated compositions. While AI has demonstrated the ability to generate coherent and stylistically consistent pieces, determining their aesthetic and emotional impact remains a subjective endeavor (Yang & Lerch, 2020). Additionally, generating long, expressive compositions that convincingly mimic human creativity, even under expert scrutiny, remains an ongoing challenge. As AI continues to evolve, understanding its influence on the creative process, musician-AI collaboration, and the broader music industry is essential for fostering a balanced integration of AI technologies.

This study aims to explore the implications of AI in music composition by addressing three key research questions: (1) How do musicians and composers perceive the use of AI in their creative processes? (2) What role does the creative context play in shaping AI's involvement in music composition? (3) How can AI tools be designed to better support music creators while preserving artistic integrity? By examining these aspects, this study contributes to the growing discourse on AI's role in music creation, building upon existing research in Music Information Retrieval (MIR) and human-AI collaboration in the creative industries (Pasquier, Eigenfeldt, Bown & Dubnov, 2017; Knees, Schedl & Got, 2020; Cella, 2020; Gioti, 2020).

To address these research questions, this study employs qualitative and case study methodologies, examining diverse creative contexts in which AI are utilized for music composition. Through this approach, we aim to develop a comprehensive framework that illustrates the fluid and dynamic roles AI can play in the creative process. The findings of this study will provide valuable insights into how AI tools can be designed and implemented to enhance, rather than replace, human creativity, fostering a collaborative synergy between musicians and AI. Hence, systematically assessing AI's capabilities, limitations, and ethical implications in music composition, this research seeks to inform future developments in AI-assisted music creation. Ultimately, understanding AI's role in music will help pave the way for more effective, ethical, and artistically meaningful integrations of AI in creative practices.

The Evolution of AI in Music Creation

The integration of AI into music creation has been a progressive journey, evolving from rudimentary algorithmic experiments to sophisticated compositions that challenge traditional notions of creativity. This evolution reflects both technological advancements and the dynamic relation between human artistry and machine capabilities. In the early experiments and algorithmic foundations, the inception of AI in music dates back to the 1950s, marked by Alan Turing's pioneering efforts in generating computer-based musical tones. In 1957, Lejaren Hiller and Leonard Isaacson utilized the ILLIAC I computer to compose the

"Illiac Suite for String Quartet," demonstrating that computers could follow musical rules to create structured compositions (Mishra, 2021). This period also saw composers like Iannis Xenakis applying stochastic processes and mathematical models to introduce elements of randomness into music, thereby expanding the horizons of algorithmic composition.

The introduction of the Musical Instrument Digital Interface (MIDI) in 1983 revolutionized the music industry by standardizing communication between electronic instruments and computers. This innovation facilitated precise control over musical parameters, leading to the development of Digital Audio Workstations (DAWs) and sophisticated algorithmic composition tools. Concurrently, John Chowning's discovery of digital frequency modulation (FM) synthesis significantly influenced the sound of 1980s pop and R&B music, exemplified by the Yamaha DX7 synthesizer (The Guardian, 2024). The late 20th and early 21st centuries witnessed significant strides in AI-driven music composition. David Cope's Experiments in Musical Intelligence (EMI), developed in 1981, enabled the emulation of various composers' styles, allowing for the creation of new compositions that mirrored classical works (Mishra, 2021). In 2010, the computer system Iamus composed "Opus One," the first fragment of professional contemporary classical music entirely generated by a computer in its own style. These developments underscored AI's potential to not only analyze but also create music autonomously (Diaz-Jerez, 2011; Ball, 2012).

However, the 2010s and early 2020s saw the emergence of AI platforms that democratized music creation. AIVA (Artificial Intelligence Virtual Artist), created in 2016, specializes in composing classical and symphonic music. By analyzing a vast repertoire of classical works, AIVA can detect patterns and compose original pieces, earning recognition from the French music rights organization SACEM (Youvan, 2024). Similarly, platforms like Suno AI and Udio have enabled users to generate music based on text prompts, blending vocals and instrumentation to produce realistic songs (Ward, 2023; Robinson, 2024). Recent advancements have led to the development of AI tools that enhance music production. Nvidia's Fugatto, unveiled in late 2024, can generate unprecedented sounds, such as a trumpet that meows, by interpreting textual descriptions (Roth, 2024). This capability allows for the creation of intricate sound compositions and the manipulation of vocal attributes, pushing the boundaries of traditional music production.

With the rise of machine learning and deep learning technologies, AI-driven music creation entered a new phase of sophistication. The introduction of neural networks and generative adversarial networks (GANs) enabled AI to analyze and generate music with a higher degree of complexity and stylistic coherence (Goodfellow et al., 2014). Projects like OpenAI's MuseNet and Google's Magenta have demonstrated AI's ability to generate multi-instrument compositions, blending different styles seamlessly (Brown, Mann, Ryder, Subbiah, Kaplan et al., 2020). Music Information Retrieval (MIR) has played a critical role in advancing AI-based music analysis and generation. By leveraging large datasets, AI models can recognize patterns, predict musical structures, and generate compositions that adhere to traditional and contemporary music styles (Knees, Schedl & Goto, 2020). Despite these advancements, challenges remain, including the need for better interpretability, ethical considerations, and the refinement of AI-generated music to meet human artistic expectations (Rohrmeier, 2022). The impact of AI on music creation continues to evolve, with ongoing research focused on human-AI collaboration, improving AI's adaptability to creative contexts, and addressing concerns about authorship and originality (Pasquier et al., 2017; Newman et al., 2023).

Applications of Artificial Intelligence in Music Creation

AI in Music Composition

The application of artificial intelligence (AI) in music composition has revolutionized the way melodies and harmonies are created, significantly reducing the time required for composing original pieces. Traditionally, musical composition demanded years of training and a deep understanding of music theory. However, AI now assists composers by generating melodies, harmonies, and even entire compositions, facilitating the creative process (Briot, Hadjeres, & Pachet, 2020). AI-driven tools such as AIVA and Amper Music serve as prime examples of how AI is utilized to generate original music. These platforms analyze extensive datasets of existing music, learning patterns, chord progressions, and stylistic elements from various genres to create compositions that emulate human creativity (Youvan, 2024). AIVA, for instance, has been officially recognized as a composer by music rights organizations, with its compositions featured in video games and film soundtracks. Similarly, Amper Music allows musicians to compose original tracks by selecting parameters such as mood, tempo, and instrumentation, enabling them to generate custom-tailored compositions in a matter of minutes (Pasquier et al., 2017). Despite AI's ability to generate music, critics argue that AI-composed pieces may lack the emotional depth and nuance of human-created works. However, AI serves as a valuable collaborative tool, offering musicians fresh perspectives and innovative ideas. By working alongside human composers, AI opens up new possibilities for musical exploration, pushing the boundaries of traditional composition methods (Kaliakatsos-Papakostas, Floros, & Vrahatis, 2020).

Composition's Structure

Compositions are made of a melodic part, played by different instruments whose frequency range may or may not be similar, and an accompaniment or harmonic part that gives the piece a deep and structured feel. Moreover, music is based on two dimensions, time dimension, represented by the note's duration or rhythm and harmony dimension related to the note values or pitch. (Hernandez-Olivan & Beltran, 2022). Based on the ideas of Walton in 2005, the basic music principles or elements are identified:

Harmony

It is the superposition of notes that form chords which compose a chord progression. The note-level could be considered as the lowest level in harmony, followed by the chord-level, while the highest-level can be considered as the progression-level which usually belongs to a certain key.

Music Form or Structure

It is the high-level structure that of the composition and it is related with the time dimension. The smallest part of a music piece is the motif which is developed in a music phrase and the combination of music phrases form a section. Sections in music are ordered depending on the music style such as intro-verse-chorus verse-outro for some pop songs (also represented as ABCBA) or exposition development- recapitulation or ABA for Sonatas. The concatenation of sections which can be in different scales and modes gives us the entire composition.

Melody and Texture

Texture in music terms refers to the melodic, rhythmic and harmonic contents that must be combined in a composition in order to form the music piece. Music can be monophonic or polyphonic depending on the notes that are played at the same time step, homophonic or heterophonic depending on the melody, if it has or not accompaniment.

Instrumentation and Orchestration

These are music techniques that take into account the number of instruments or tracks in a music piece. Whereas instrumentation is related to the combination of musical instruments which compose a music piece, orchestration refers to the assignment of melodies and accompaniment to the different instruments that compose a determined music piece. In recording or software-based music representation, Instruments are organized as tracks, each of those containing the collection of notes played on a single instrument (a piece of music played by more than one instrument is called multi-track). Each track can contain one note (monophonic tracks) or multiple notes that sound simultaneously (polyphonic tracks). (Candusso, 2024).

AI in Music Production

AI has also made significant contributions to the music production process, from recording to mastering. Traditionally, production required skilled sound engineers and producers to fine-tune tracks manually, ensuring professional sound quality. AI-driven tools now streamline these tasks, making high-quality music production more accessible to independent artists and producers (Sturm et al., 2019). Platforms such as LANDR, an AI-powered mastering service, allow musicians to upload tracks and receive professionally mastered versions within minutes. The AI analyzes the track, adjusts levels, and applies effects to optimize sound quality, democratizing the mastering process and making it more affordable (Lee, Hitt, Terada & Lee, 2022). Additionally, AI-based mixing tools such as iZotope's Neutron analyze individual elements of a mix and provide suggestions for adjusting equalization, compression, and spatial effects, enabling producers to achieve a balanced sound more efficiently (Hong, 2018). Furthermore, AI-driven virtual instruments and plugins have expanded creative possibilities for musicians. AI synthesizers, such as Google's Magenta Studio, allow for the generation of entirely new and unique sounds, fostering innovation in experimental music production. By automating repetitive tasks, AI allows producers and musicians to focus on the artistic aspects of production, ultimately enhancing creativity and efficiency (Manning, 2020).

AI in Live Performances

Live music performances have traditionally relied on human musicians' improvisational skills and technical abilities to create engaging experiences. AI is now enhancing live performances by adding interactivity, real-time composition, and visual synchronization, expanding the boundaries of musical expression (Yellowbrick, 2020). AI-driven technology is used to generate dynamic visual displays that synchronize with live music. For example, artist Holly Herndon integrates AI-generated vocals and visuals into her performances, creating a multi-sensory concert experience. Additionally, AI-powered robotic musicians such as Shimon, developed at Georgia Tech, improvise in real-time alongside human musicians by analyzing musical patterns and generating complementary melodies and rhythms (Gao et al., 2024). Beyond stage performances, AI is also revolutionizing virtual concerts and audience engagement. Platforms such as Wave XR use AI and augmented reality (AR) to create immersive virtual concerts, allowing artists to reach global audiences without the constraints of physical venues. These technological advancements redefine the concert-going experience, merging music with digital interactivity (Mazzone & Elgammal, 2019).

Generative AI Tools for Music Composition

As noted earlier, one of the most prominent themes in AI for music is composition. At the core of generative music tools is a range of algorithms and methodologies designed to mimic and enhance the process of creativity. Generative AI tools apply machine learning models, like recurrent neural networks (RNN), variational auto encoders (VAE), and generative adversarial networks (GAN) (Atanacković, 2024). They encompass a range of different platforms with varying approaches to algorithmic creativity and musical exploration. Generative AI tools can be categorized in line with the musical elements in which they specialize, like melody generation, harmonic progression, rhythmic patterns, and sound synthesis. There has been a recent surge in software tools for this purpose, with a range of companies working in this area, making it difficult to discuss them all in detail.

Music VAE and MuseNet

One such tool is Music VAE, a product from researchers in Google Magenta. Music VAE applies variational autoencoders (VAE) to learn underlying structures in musical data and generate compositions from learned patterns and style. Improvements in Music VAE enable users to generate and explore a vast range of musical ideas, ranging from traditional melodies to experimental works. Similarly, OpenAI's MuseNet applies deep neural networks to generate original music in a range of genres and styles (Atanacković, 2024). By training on a vast dataset of musical scores, MuseNet captures subtleties in different musical traditions and genres and generates compositions that capture coherence and creativity. This online tool can replicate jazz improvisations and classical works by supplying its algorithm with compositions and performances from specific composers and artists. However, it is principally designed to create simple melodies and short motifs.

Soundraw and AIVA

Impressive software is Soundraw, a program with which users can create royalty-free music according to their own tastes and intentions. Employing machine learning to process user data and musical variables, Soundraw creates original pieces, providing a seamless solution for content creators, multimedia producers, and filmmakers who require proper soundtracks for their productions. Comparable to this class, AIVA is one of the oldest AI-based music generation software. This software allows users to compose new songs in over 250 different genres in seconds. It is meant for professional and amateur musicians. Additionally, AIVA enables users to define their own style models from uploaded audio or MIDI files to influence the algorithm. Compositions can then be saved in a variety of file formats. However, instrumental music is still its prime focus (Atanacković, 2024).

Suno AI and MelodyRNN

Likewise, Suno AI not only creates music from text-based input but also generates accompanying lyrics, as shown on its front page. Another helpful melodic generator is Magenta's MelodyRNN, which is meant to generate melodic sequences. Comparable to the above software, it employs deep learning algorithms to process melodic contours from familiar compositions to create new melodies according to established stylistic patterns but with additional original variations and permutations (Atanacković, 2024).

Harmonic and Rhythmic Generators

Regarding harmonic progression, which is crucial in determining a piece's tonal and emotive character, harmonic generation software like ChordChord and the ToneGym Chord Progression Generator are effective in producing harmonic sequences accompanying melodic lines. The tools use machine-learning algorithms to process harmonic structures, chord progressions, and voice-leading rules from different genres and styles of music to allow them to generate rich and beautiful progressions. Rhythmic-wise, GroovePizza is programmed to generate rhythmic patterns and drum loops that can serve as a starting point for a piece of music. Processing rhythmic motifs, syncopated grooves, and polyrhythmic interactions in different musical traditions and cultures, the tool generates varied and interesting rhythmic patterns.

Sound Synthesis and Transformation Tools

Additionally, generative AI can synthesize and transform sounds and timbres that contribute to a composition's sonic palette. Tools such as Google's Nsynth and OpenAI's SynthGPT-based GPT-3 break down and analyze the spectral features, harmonic overtones, and transient behavior in various acoustic and synthetic instruments and sound sources and can hence generate new sounds with a balance between expressiveness and realism. One of the more advanced ideas taken from ChatGPT and generative art Discord servers (such as Midjourney) and utilized in SynthGPT is text-to-prompt. This feature enables users to inform SynthGPT what sound they want in their production and composition and have a fully generated synth patch. SynthGPT, however, has not yet been updated to allow creatives to modify some of its settings. Another intriguing AI-based sound synthesis tool is Magenta's Tone Transfer. Any audio file can be uploaded and processed and then synthesized into acoustic instruments such as saxophone, violin, flute, cello etc.

Tone Transfer is also available as a VST (Virtual Studio Technologies) plugin, which music professionals can use in their DAW (Digital Audio Work-station) of choice. One of the more advanced technologies with the same text-to-prompt idea is used in the Synplant plugin, a VST instrument highly used in modern electronic music. In its new version (2.0), not only is it possible to describe the sound and get the full patch, but also producers and composers are allowed to shape the sound using more advanced parameters such as filtering, shaping the length of the sound using ADSR envelope (Attack, Decay, Sustain and Release) and more. Even though SynthGPT and Tone Transfer (along with Synplant) could be thought of as generative AI tools, due to the fact of their text prompt or sound analytic possibilities, they can also potentially be used as assistive tools for composers and producers (Atanacković, 2024).

AI Models in Music Creation

The application of AI in music creation has advanced significantly due to the development of deep learning models that can analyze, generate, and manipulate complex patterns. These advancements have led to the emergence of various generative AI (GenAI) models, which play a crucial role in enhancing composition, production, and performance. Some of the most influential models in AI-driven music creation include Generative Adversarial Networks (GANs), Large Language Models (LLMs), Variational Auto encoders (VAEs), Autoregressive Models, and Diffusion Models. These models serve as collaborative tools that assist musicians in creating, modifying, and enhancing musical compositions.

Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs), introduced by Good fellow et al. (2014), consist of two neural networks—a generator and a discriminator—that operate in a competitive learning process. The generator creates outputs (such as melodies, harmonies, or synthesized instrument sounds), while the discriminator evaluates their authenticity by distinguishing AI-generated compositions from real human-made music. Over time, the generator improves its ability to produce realistic outputs, leading to highly convincing synthetic music. GANs have been widely adopted in music generation for their ability to learn musical patterns and create novel compositions. Researchers have used GAN-based models like MuseGAN (Dong et al., 2018) to generate polyphonic music by analyzing large datasets of existing musical works. These models enable AI to compose multi-instrumental pieces, making them valuable tools for musicians seeking inspiration or automated accompaniment. Additionally, GANs are instrumental in style transfer, allowing musicians to generate music that mimics the stylistic elements of specific composers or genres.

Large Language Models (LLMs)

Large Language Models (LLMs), including OpenAI's GPT-4 and Google's Bard, have revolutionized text-based AI applications, including lyric generation, songwriting, and music-related storytelling. These models are trained on extensive datasets, allowing them to generate lyrics that align with specific themes, emotions, and styles. The transformer architecture (Vaswani et al., 2017), which underpins modern LLMs, enables these models to understand linguistic nuances and long-range dependencies in text, making them capable of producing coherent and contextually relevant lyrics. Beyond lyric writing, LLMs are increasingly integrated into AI-powered music generation tools. By processing textual prompts, these models assist in composing melodies or structuring musical pieces. For example, OpenAI's Jukebox (Brown et al., 2020) is an AI model that combines LLM capabilities with music synthesis, generating complete songs with lyrics, harmonies, and instrumentals. Moreover, multimodal large language models (MLLMs) are expanding beyond text-based input, enabling AI systems to interact with audio and symbolic music data, enhancing their role in collaborative music creation.

Variational Auto encoders (VAEs)

Variational Auto encoders (VAEs), introduced by Kingma and Welling (2013), are generative models designed to encode and reconstruct data. They consist of an encoder that compresses input data into a latent representation and a decoder that reconstructs it into meaningful output. In music creation, VAEs have been widely applied in sound synthesis, instrument modeling, and music generation. VAEs play a crucial role in AI-driven composition by learning latent features of musical structures. For instance, models like MusicVAE (Agostinelli, Denk, Borsos et al., 2023) use VAEs to interpolate between different musical styles, generating smooth transitions between melodies and harmonies. This technique allows musicians to explore creative variations of a composition and experiment with novel musical ideas. Additionally, VAEs contribute to timbre synthesis, enabling the creation of unique instrumental sounds that can be integrated into digital audio workstations (DAWs) for music production.

Autoregressive Models

Autoregressive models are probabilistic frameworks that predict the next element in a sequence based on previous elements. These models are widely used in natural language processing (NLP) and have been adapted for music composition, where they generate melodies and harmonies by sequentially predicting musical notes. In AI-driven music generation, autoregressive models analyze patterns in musical sequences and generate compositions note by note. Notable examples include OpenAI's MuseNet and Google's Transformer-based Music Transformer (Huang et al., 2018), which leverage autoregressive techniques to produce stylistically diverse music. These models have demonstrated success in tasks such as melody continuation, where AI extends a given musical phrase while maintaining coherence with the original theme. Furthermore, autoregressive approaches are used in AI-powered improvisation tools, allowing musicians to collaborate with AI in real-time performances.

Diffusion Models

Diffusion models, originally inspired by physical diffusion processes, have gained attention for their ability to generate high-quality images and audio. These models work by adding random noise to data and then learning to reverse the process to reconstruct meaningful output. In music creation, diffusion models have been employed in audio synthesis, sound design, and text-to-music generation. Recent advancements in diffusion-based AI models, such as Stable Audio (Harmonai, 2023) and OpenAI's Riffusion, demonstrate the potential of these techniques in generating complex musical textures and ambient soundscapes. By training on large datasets of musical recordings, diffusion models can generate high-fidelity instrumentals and experimental sounds. They are particularly useful in electronic music production, where AI-generated sounds can be manipulated and integrated into compositions.

Challenges and Ethical Considerations

Music Generation

The problem of music generation consists in creating music in any of its forms i.e. sound, notes or other representation. Music generation specifically relates to creativity—a very human ability. For this reason, the problem has attracted a lot of attention since the onset of AI application to music (Kaliakatsos-Papakostas et al., 2020; Mycka & Mańdziuk, 2025). Two aspects of the music generation should be distinguished. One is generation of new musical compositions (e.g. generation of sheet music), and the other one is generation of expressive performance (referring to already existing works). Both aspects emulate human creativity process.

Music Classification

The second key problem related to AI in music is music classification. This task can take different forms, but in general can be stated as determining a particular feature of a given piece (genre, composer, performer, etc.). Another, more specific form, is piece recognition, which usually refers to identifying the author and name of the piece, sometimes also the performer. An important aspect of this problem is dealing with distortions that can occur when music is recorded. Systems for classification/recognition should work properly despite these inefficiencies. Attempts to solve the problem have a long history, and include, for instance, creation of a set of simple rules for musical performance analysis (Mycka & Mańdziuk, 2025). As an example, genres classification problem is presented in Fig. 1.

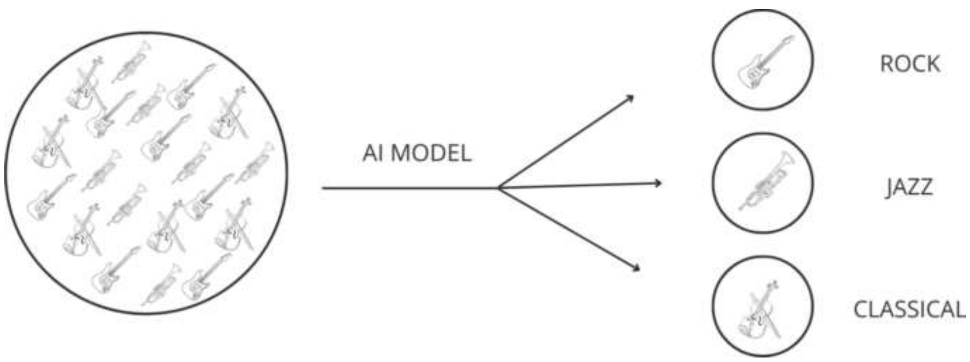

Fig. 1: Genres classification problem (Mycka & Mańdziuk, 2025).

Music Recommendation

Creating recommendation systems is the third major music-related topic (Fig. 2). This problem has recently gained much popularity, partly due to the growing interest in various music-related streaming services. The basic task of recommendation is to suggest music (usually songs) to the user in accordance with his/her preferences. Thus, the problem is to determine the preferences of the user, both general and moment-specific, and then to select songs from a pool of music resources that best match these preferences. The essence of the problem is that the recommended songs should not only be selected from the songs that the user has already listened to, but especially from the songs that are available in the pool and with which the user is not familiar yet (Mycka & Mańdziuk, 2025).

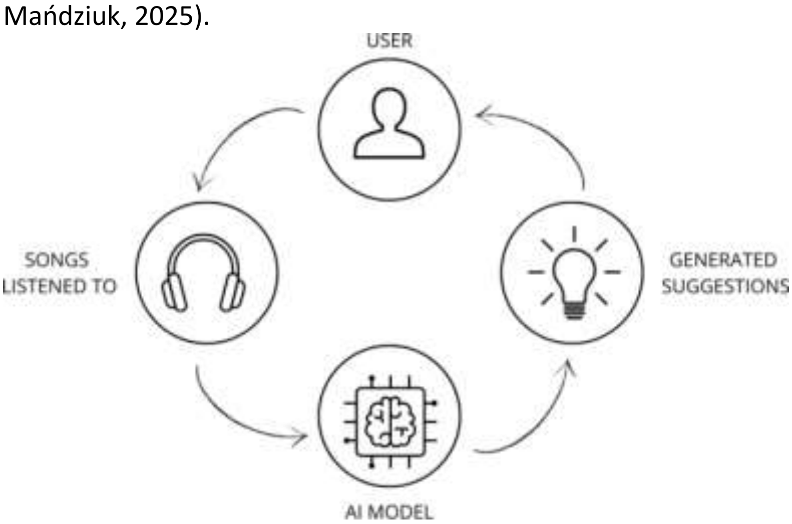

Fig. 2: Basic flow of the recommendation process (Mycka & Mańdziuk, 2025).

Other Problems

Apart from the three main above-mentioned trends, i.e. generation, classification and recommendation, there are several other topics that do not fit under any of these categories. One of the less popular, though interesting, and research avenues involves both human vision and hearing (Mycka & Mańdziuk, 2025). According to Raphael in 2006, an endeavour is made to synchronize music (hearing) with its notation (vision) automatically. The system relies solely on pitch-based representation, and the solution model considers both discrete (notation position) and continuous (song tempo tracking) variables. The task is to find the correct position in a musical score given an audio frame. The estimation of a score position is carried out using tree search algorithm. But Li, Tao, Maybank and Yuan in 2008 noted that another pairing of visual and auditory senses is presented, where the goal is to associate images with a given song. Separate sets of features are extracted from both song (MFCC) and image (colour and texture) and input together into the system (Mycka & Mańdziuk, 2025).

The detection of music plagiarism is an example of another less popular music-related topic. The system proposed by De Prisco, Malandrino, Zaccagnino and Zaccagnino (2017) uses fuzzy vector-based similarity to identify plagiarism. The detection process starts with a transformation of the suspected piece into a vector representation, then the search for similar pieces is performed, and finally, the fuzzy similarity score is calculated. The system has demonstrated high detection accuracy when tested on a collection of musical pieces containing confirmed instances of plagiarism. An interesting initial task that can serve as a foundation for further analysis or data processing is automatic feature detection. An example of such is proposed in Das and Kolya (2020), where extraction is performed using unsupervised training of a feedforward network with Hebbian learning.

Future Directions and Open Research Questions

The role of artificial intelligence in music creation is expanding rapidly, and as technology continues to evolve, new opportunities and challenges arise. Several key areas for future research and development include:

Enhancing AI Creativity and Originality

One of the most pressing questions in AI-generated music is whether AI can truly be "creative" in the human sense. Current AI models primarily analyze existing datasets and generate compositions based on learned patterns, but the ability to create genuinely novel and emotionally resonant music remains a challenge. Future research could explore new architectures, such as hybrid AI models that combine deep learning with cognitive science approaches, to enhance AI's capacity for creativity.

Improving Human-AI Collaboration

AI is increasingly being used as a collaborative tool rather than a replacement for human musicians. Future developments could focus on designing interactive AI systems that adapt to individual artists' styles, allowing for real-time co-creation. AI tools that can respond dynamically to human input, suggest musical variations, and refine compositions based on user feedback will enhance creative partnerships between humans and machines.

Ethical and Legal Considerations in AI-Generated Music

As AI-generated music becomes more prevalent, issues surrounding copyright, ownership, and intellectual property rights will need clearer resolution. Who owns the rights to music created by AI? Should AI-generated works be treated as derivative or original compositions? Research in this area will be crucial for shaping legal frameworks and ensuring fair compensation for musicians, producers, and developers.

AI in Personalized Music Recommendation

Music streaming services rely heavily on AI-powered recommendation algorithms. However, these systems often reinforce existing listening habits rather than encouraging musical exploration. Future advancements could focus on more sophisticated personalization techniques that balance familiarity with diversity, introducing listeners to new genres and artists while respecting their preferences.

AI and Emotional Understanding in Music

Music is deeply connected to human emotions, yet AI-generated compositions often lack the depth of feeling that characterizes human-made music. Developing AI systems that can analyze and respond to emotional cues, both in composition and performance, remains an open challenge. Emotionally aware AI could lead to more meaningful music for therapeutic, entertainment, and creative applications.

AI in Adaptive and Interactive Music Systems

Another exciting avenue for future research is the application of AI in adaptive and interactive music, particularly in gaming, virtual reality, and augmented reality experiences. AI-generated soundtracks that adjust dynamically to user actions, emotions, or environmental cues could revolutionize immersive audio experiences.

Expanding AI's Role in Music Education

AI tools have the potential to democratize music education by providing personalized learning experiences, real-time feedback, and composition assistance. Future research could explore AI-driven tutoring systems that cater to different learning styles, helping students understand music theory, composition, and production more effectively.

2. Conclusion

Artificial intelligence has significantly transformed music creation, enabling new forms of composition, production, and performance. From early algorithmic experiments to sophisticated deep learning models, AI has expanded the possibilities for musicians, producers, and audiences alike. Tools such as AIVA, MuseNet, and Soundraw showcase AI's ability to generate compositions, while advancements in music production and live performance demonstrate AI's role in enhancing human creativity. Despite its progress, AI-generated music raises important questions about originality, authorship, and ethical considerations. As AI continues to develop, future research must address the challenges of creativity, collaboration, and emotional depth to ensure that AI remains a tool that enhances, rather than replaces, human artistry. By focusing on the synergy between technology and musicianship, AI has the potential to redefine the way music is created, experienced, and appreciated in the digital age.

References

Agostinelli, A., T. Denk, Z. Borsos, J. Engel, M. Verzetti, A. Caillon, Q. Huang, A. Jansen, A. Roberts, M. Tagliasacchi, M. Sharifi, N. Zeghidour & C. Frank (2023). MusicLM: Generating music from text. Available at: https://arxiv.org/abs/2301.11325

Atanacković, D. (2024). Artificial Intelligence: Duality in Applications of Generative AI and Assistive AI in Music. INSAM Journal of Contemporary Music, Art and Technology, 12, 12-31. https://doi.org/10.51191/issn.2637-1898.2024.7.12.12

Ball, P. (2012). Iamus, classical music's computer composer, live from Malaga. The Guardian.

Barton, G. (2018). Music learning and teaching in culturally and socially diverse contexts: implications for classroom practice. Springer

Briot, J-P., Hadjeres, G., & Pachet, F. (2020). Deep Learning Techniques for Music Generation. In: Computational Synthesis and Creative Systems. Springer Cham. https://doi.org/10.1007/978-3-319-70163-9

Brown, T.B., B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D.M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever & D. Amodei (2020). Language models are few-shot learners. Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2020). Curran Associates Inc., Red Hook, NY, USA, Article 159, 1877–1901. doi: 10.5555/3495724.3495883.

Candusso, S. (2024). Exploring the impact of generative AI on the music composition market: A study on public perception, behavior, and industry implications. Master's Thsis in Engineering and Management, Politecnico di Torino.

Cella, C.-E. (2020). Music information retrieval and contemporary classical music: A successful failure. Transactions of the International Society for Music Information Retrieval, vol. 3, 126–136.

Das, S., & Kolya, A.K. (2020). Detecting generic music features with single layer feedforward network using unsupervised Hebbian computation. Int J Distrib Artif Intell (IJDAI) 12(2):1–20

De Prisco, R., Malandrino, D., Zaccagnino, G., & Zaccagnino, R. (2017). Fuzzy vectorial-based similarity detection of music plagiarism. In: 2017 IEEE international conference on fuzzy systems (FUZZ-IEEE), pp 1–6. https://doi.org/10.1109/FUZZ-IEEE.2017.8015655

Diaz-Jerez, G. (2011). Composing with Melomics: delving into the computational world for musical inspiration. Leonardo Music Journal, 21, 13–14. doi:10.1162/LMJ_a_00053. S2CID 57569752.

Dong, H.-W., Hsiao, W.-Y., Yang, L.-C., & Yang, Y.-H. (2018). MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation. In: Proceedings of AAAI Conference on Artificial Intelligence.

Gao, X., Rogel, A., Sankaranarayanan, R., Dowling, B., & Weinberg, G. (2024). Music, body, and machine: gesture-based synchronization in human-robot musical interaction. Frontiers in robotics and AI, 11, 1461615. https://doi.org/10.3389/frobt.2024.1461615

Gioti, A.-M. (2020). From artificial to extended intelligence in music composition," Organised Sound, vol. 25, no. 1, p. 25–32, 2020.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., &

Bengio, Y. (2014). Generative Adversarial Nets. In: Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014), pp. 2672–2680.

Harmonai. (2023). Stable Audio: AI-powered Music and Sound Generation. Retrieved from https://harmonai.org

Hernandez-Olivan, C., & Beltran, J. R. (2022). Music Composition with Deep Learning: A Review. Advances in Speech and Music Technology, 25–50. http://arxiv.org/abs/2108.12290

Hong, J. (2018). Bias in perception of art produced by artificial intelligence. In: International Conference on Human-Computer Interaction: Interaction in Context, 2018.

Hong, J., Peng, Q., & Williams, D. (2021). Are you ready for artificial mozart and skrillex? an experiment testing expectancy violation theory and ai music. New Media & Society, 23, 1920–1935.

Huang, C.-Z. A., Vaswani, A., Uszkoreit, J., Simon, I., Hawthorne, C., Dai, A. M., Hoffman, M. D., & Eck, D. (2018). Music Transformer: Generating Music with Long-term Structure. In: Proceedings of the International Society for Music Information Retrieval (ISMIR).

Kaliakatsos-Papakostas, M., Floros, A., & Vrahatis, M.N. (2020). Artificial intelligence methods for music generation: a review and future perspectives. In: Nature-inspired computation and swarm intelligence, pp 217–245

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. arXiv:1312.6114. Available at: https://arxiv.org/abs/1312.6114

Knees, P., Schedl, M., & Goto, M. (2020). Intelligent user interfaces for music discovery. Transactions of the International Society for Music Information Retrieval, 3, 165–179, 2020.

Lee, K., Hitt, G., Terada, E., & Lee, J. (2022). Ethics of singing voice synthesis: Perceptions of users and developers," in Proc. of the 23rd Int. Society for Music Information Retrieval Conf., Bengaluru, India, 2022, pp. 733–740

Manning, C. (2020). Artificial Intelligence Definitions. Stanford University.

Mazzone, M., & Elgammal, A. (2019). Art, creativity, and the potential of artificial intelligence. Arts, 8, 26.

Mishra, M. (2021). Living, Singing AI: An evolving, intelligent, scalable, bespoke composition system. Master's thesis in Media Arts and Sciences, School of Architecture and Planning, Massachusetts Institute of Technology.

Moruzzi, C. (2020). Should human artists fear AI? A report on the perception of creative ai. In xCoAx 2020: Proceedings of the Eighth Conference on Computation, Communication, Aesthetics & X, M. Verdicchio, M. Carvalhais, L. Ribas, and A. Rangel, Eds. Porto: Universidade do Porto, 2020, pp. 170–185.

Mycka, J., & Mańdziuk, J. (2025). Artificial intelligence in music: recent trends and challenges. Neural Comput & Applic., 37, 801–839. https://doi.org/10.1007/s00521-024-10555-x

Ndou, N., Ajoodha, R., & Jadhav, A. (2021). Music genre classification: a review of deep-learning and traditional machine-learning approaches. In: 2021 IEEE international IOT, electronics and mechatronics conference (IEMTRONICS), pp 1–6. https://doi.org/10.1109/IEMTRONICS52119.2021.9422487

Newman, M., Morris, L., & Lee, J.H. (2023). Human-AI Music Creation: Understanding the Perceptions and Experiences of Music Creators for Ethical and Productive Collaboration. In: Proc. of the 24th Int. Society for Music Information Retrieval Conf., Milan, Italy, 2023.

Pasquier, P., Eigenfeldt, A., Bown, O., & Dubnov, S. (2017). An introduction to musical metacreation.

Comput. Entertain., 14, 2017. https://doi.org/10.1145/2930672

Roberts, A., Engel, J., Mann, Y., Gillick, J., Kayacik, C., Nørly, S., Dinculescu, M., Radebaugh, C., Hawthorne, C., & Eck, D. (2020). Magenta Studio: Augmenting Creativity with Deep Learning in Ableton Live. In Proceedings of the 6th International Workshop on Musical Metacreation. Charlotte, United States: MUME, Jun. 2019, p. 7. https://doi.org/10.5281/zenodo.4285266

Robinson, K. (2024). Major Labels Sue AI Firms Suno and Udio for Alleged Copyright Infringement. Billboard.

Roth, E. (2024). Nvidia claims a new AI audio generator can make sounds never heard before. The Verge.

Sturm, B. L., Ben-Tal, O., Monaghan, Ú., et al. (2019). The Musical Turing Test: AI and the Future of Music. Computers in Entertainment (CIE), 14(3), 1-14.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems 30. arXiv: 1706.03762v7.

Ward, A. (2023). How to Use Microsoft Copilot's New Suno AI Music Creation Tool. Tech.co.

Yang, L-C., & Lerch, A. (2020). On the evaluation of generative models in music. Neural Comput Appl., 32:4773–4784

Yellowbrick. (2023). Understanding the Music Industry's Business Structure.

Youvan, D.C. (2024). The Future of Music: Leveraging Advanced AI and Computational Speed to Revolutionize Music Creation, Production, and Consumption. https://www.researchgate.net/publication/381698513

Zulic, H. (2019). How AI can change/improve/influence music composition, performance and education: Three case studies. INSAM Journal of Contemporary Music, 1, 100–114.