Learning AI with Music: A Sound-based Dissemination Activity for High School Students

Guido Vallarino[1,*], Giorgio Delzanno[1,*], Giovanna Guerrini[1], Matteo Moro[1] and Lorenzo Luciano Morelato[1,†]

[1] *Università degli Studi di Genova, Italy*

Abstract

We present a dissemination activity developed as an interactive laboratory for high school students aimed at introducing the basic principles of Artificial Intelligence and Machine Learning through music as an entry point. The activity is designed to be engaging, using the familiarity of the students with musical instruments to explain the concept of machine learning from data. Participants are invited to record sounds of real instruments, which are then used to train a classifier in real time, showing how a machine can learn to recognize different types of input data. Then we briefly introduce current music creation tools based on Machine Learning, such as splitters that can separate individual instruments from a mixed track. A second part of the activity focuses on Generative AI, introducing the core ideas through links between using text and signals and a musical Turing test: students are asked to distinguish between songs played by human musicians and others generated by AI models. We report the results collected from the hundreds of students who have already participated, along with the educational objectives of the activity, which include stimulating curiosity and critical thinking about the field and the possible future impacts.

Keywords

Artificial Intelligence, Machine Learning, Generative AI, Creativity, Music Production, Education, Critical Thinking

1. Introduction

Background The adoption of tools based on Artificial Intelligence (AI) has become a concrete reality within the music industry, and it is a field undergoing constant transformation. Beyond the well-known recommendation systems used by streaming platforms, which significantly influence how music is consumed by listeners, recent advancements in Machine Learning (ML), have introduced new support systems for musicians and producers and brought significant ethical and economic implications in the music production industry, see [1, 2]. AI systems such as OpenAI's MuseNet, Google's Magenta, and Amper Music can analyze vast datasets of existing music to learn complex patterns and styles, resulting in the ability to generate music that can mimic human creativity. These tools are capable to assist in various tasks related to music production that were previously far more labor-intensive and time-consuming. They have unlocked opportunities that were previously unimaginable, due to the limitations of earlier technologies that were either inaccurate or ineffective. A notable example is the development of stem splitters, see e.g. Spleeter based on TensorFlow [3], which can isolate individual instruments or vocals from a mixed audio track, which is something that was extremely difficult to achieve using traditional methods. Another important innovation is the evolution of auto-tuners or tools that correct tempo defects used in recording studios: today, these tools are significantly faster and more accurate at recognizing melodies and rhythms, allowing producers to correct imperfections in a singer's or musician's performance with greater efficiency than ever before.

2nd Workshop on Education for Artificial Intelligence (edu4AI 2025, https://edu4ai.di.unito.it/), Co-located with ECAI 2025, the 28th European Conference on Artificial Intelligence which will take place on October 26, 2025 in Bologna, Italy

*Corresponding author.

†These authors contributed equally.

guido.vallarino@edu.unige.it (. G. Vallarino); giorgio.delzanno@unige.it (. G. Delzanno); giovanna.guerrini@unige.it (. G. Guerrini); matteo.moro@unige.it (. M. Moro); lorenzo.luciano.morelato@edu.unige.it (. L. L. Morelato)

0000-0001-7030-1050 (. G. Delzanno); 0000-0001-9125-9867 (. G. Guerrini); 0000-0002-8644-0175 (. M. Moro)

Perhaps with an even greater impact on large-scale music production, Generative AI enables the creation of new compositions with minimal input. Similarly to GPT models that break down text into tokens, music sheets or wave files can be broken down into discrete units that AI algorithms can process [4, 5, 6]. The AI algorithms can then predicts the next likely token in a sequence, similar to how it predicts the next word in a sentence for text generation. Furthermore, tools based on Generative AI such as Suno, ElevenMusic , Soundraw, and many others, now enables the creation of original music directly from text prompts making music composition accessible to everyone. AI-based tools also gave rise to major controversies on the intellectual property rights of generated compositions, see e.g. [2] and the 2023 lawsuit against Anthropic by different music publishers [7].

Motivations and Goals Given the rich and complex scenario mentioned above, creating music seems a highly relevant and timely way to disseminate AI methods and tools. The main objective of our project is to utilize these innovative tools alongside targeted educational resources. The idea is to offer students a general introduction to artificial intelligence using music, a topic that naturally sparks their interest. This will stimulate curiosity and encourage critical thinking, ideally equipping participants with the foundational skills needed to analyze and forecast the potential impacts of the current AI revolution and, in the meantime, reason on ethical and legal issues. For this purpose, we use the concrete and engaging example of the artistic field of music production. In the rest of the paper we present the format of our dissemination lab and present preliminary results obtained via questionnaires proposed to participants of different editions proposed as an orientation activity for the Computer Science Degree courses of our University. The lab proposal has also been accepted at the 2025 edition of the Pisa's Internet Festival and at the Genoa's Science Festival that will take place in between October and November 2025.

2. Activity overview

The general structure and plan of the lab is summarized in Table 1. In the rest of the section we will describe each part in detail.

First part: AI basics explanation The activity is designed to be presented as a live show. Visitors are welcomed while a small group of two or three musicians performs a song (for instance, using two guitars and an electric bass) while following a backing track (that might include drums and a lead vocal).

Only at the end of the performance, once all visitors have taken their seats, will the surprise be revealed: the song they just heard was actually "composed" using generative AI. From there, the session will continue with an exploration of the principles behind AI-based and generative AI tools, as well as practical demonstrations showing how these technologies can be useful to musicians in their creative and production processes.

To begin, visitors are invited to reflect on how they themselves learned to recognize different things, focusing specifically on the example of musical instruments. Most likely, they first heard the sound of an instrument and, only afterwards, someone told them its name. After hearing just a few examples, they were able to identify that instrument among others.

At this point, the facilitator (the person guiding and explaining the activity) plays sounds from well-known and easily recognizable instruments, asking the audience to guess their names (for example, an acoustic guitar or a drum kit). After a series of such examples, for which the audience can easily provide the correct answer, the facilitator suddenly plays the sound of a less familiar instrument, such as an oboe or a bassoon. At this moment, the facilitator (or, if a classical musician is in the audience, it might step in) reveals the instrument's name. The audience is then given a couple more listening examples of this "new" instrument, while the facilitator points out that they are now more likely to recognize it, having just learned to distinguish it from others.

[0] https://suno.com/home
[0] https://elevenlabs.io/
[0] https://soundraw.io/

Activity	Description	Duration	Type of Interaction
Welcome	Backgroud music	2 min	Live show
Basic Concepts	Supervised machine learning explanation with a comparison with human experience	10 min	Explanation with some Q&A
Training a Model	Training a sound classifier using Google's Teachable Machine	10 min	Participants train a model playing musical instruments
Inference: hearing accuracy and precision	Participants test the classifier and discuss outcomes of the experiment (scores, errors, etc)	10 min	Participants play the instruments and test the model accuracy
	Example tool: music STEM splitter based on ML	3 min	Live show
Generative AI	Introduction to the basic concepts of Generative AI: difference from non-generative AI and its output features	10 min	Explanation with some Q&A
	Turing test: audience tries to recognize recorded song from others generated with AI	5 min	Live quiz with Wooclap or similar tools
	Discussion of the results and the impact of AI and Gen-AI in our lives	5 min	Q&A

Table 1
Detailed plan of the lab activities.

This exercise is designed to mirror the process of supervised learning, the same way humans acquire knowledge through examples. It then serves as a bridge to introduce a simplified, accessible explanation of how computers use supervised machine learning to identify patterns and make predictions: there is a problem to solve (for example: given a sound, which instrument is producing it) and the machine attempts to provide an answer based on the features it has learned from previous examples of this same problem. During the training phase, it is “fed” with annotated data samples (in our case, recordings of different instruments labeled with their correct names) so that it can learn the patterns that distinguish each instrument's sound.

Next, a hands-on example of training a simple model is provided following a more human-centered approach for a more empirical interactive explanation of ML [8, 9] by inviting volunteers from the audience. At their choice, they can either play one of the instruments previously introduced at the start of the activity, or try out a selection of other instruments to spark curiosity, such as small percussion instruments, unique ethnic instruments (e.g. Tibetan singing bowls or maracas), or even DIY instruments built using Makey Makey [10]. The sounds they produce are recorded and used to train a basic algorithm via Google's Teachable Machine [11], chosen for its intuitive graphical interface, which makes every step of both the training and inference process easy to understand. During the brief technical pause required for training, the facilitator explains what the neural network is doing: namely, identifying which features are most likely to correspond to one instrument or another. Once the inference phase begins, the system demonstrates how, for each input sample, it estimates the probability that the sound belongs to one of the trained instrument categories. The audience is also shown what happens when the model encounters an instrument it has never “seen” before, causing it to make incorrect or uncertain predictions. This provides a perfect opportunity to explain the challenges of creating large and comprehensive datasets, even for relatively simple classification problems, and how this issue becomes even more significant for complex, real-world scenarios, where training can require substantial computational resources.

Second part: tool based on AI and Generative AI In the second half of the activity, after providing a general explanation of what all supervised machine learning systems have in common (particularly in the context of classification) we focus on showcasing some useful and visually engaging applications.

We begin by presenting a STEM track separator: given the audio file of a pop or light music track, it splits the song into separate components: drums, bass, vocals, and the rest of the arrangement. The audience learns that these tools work by analyzing the song's spectrogram and applying an algorithm trained on a dataset containing pre-separated tracks. This essentially allows the AI to recognize the typical frequency ranges associated with each instrument [12]. A famous song is first played (e.g. Hit the Road Jack), and then the algorithm is used to separate it into its individual stems. At this point, to keep the session entertaining and performance-oriented, the theoretical explanation is paused: the musicians who performed at the start of the event take the stage again, this time playing live along with the AI-separated tracks. For example, if the musicians play bass and guitar, the drums and vocals come from the original track, isolated by the algorithm, while the live instruments are performed in real time.

The theoretical segment wraps up with a brief introduction to Generative AI. Using the same explanatory framework applied earlier to define AI, we clarify that generative AI learns patterns from its training data and then produces new examples that follow those patterns. Continuing with the musical analogy, if the AI is trained on a dataset of blues songs, it can generate entirely new blues tracks that stylistically match the originals.

Finally, we return to a possible definition of AI [13], describing it, paraphrasing Alan Turing thoughts in [14], describing it, as an artificial system capable of exhibiting intelligent behavior (that is, behavior comparable to human actions). At this point, the audience is introduced to the Turing Test, which evaluates a system's ability to appear human.

An interactive version of this test is then carried out with the participants. They are asked to listen to four short instrumental blues tracks (A,B,C,D): one (B) is an original recording performed by us, while the other three were generated by AI. Using a platform such as Wooclap, the audience votes on which track they believe was the genuine human performance.

Prologo As discussed in the previous section, the Turing test usually results in being lost by the majority of the audience. At this point, to conclude the activity, the audience is asked: What does this all mean? The discussion then reveals that AI has already reached capabilities comparable to the human level in certain tasks. This means that our focus should not be on what might happen in the future but rather on what is already happening now. These tools are designed to serve humans, and it is our responsibility to understand them so that we can decide how best to use them.

In fact, during our trials with professional musicians (about ten attempts in total) they were consistently able to identify which track was played by humans. Almost all of them pointed out that what is missing from AI-generated music is the subtle "human touch": interpretation, nuance, and even imperfections. This leads to our final invitation to the audience: use these tools to better understand what your own "added touch" is, the unique quality you can bring to your work.

3. Educational goals and preliminary results

The main educational goals of the activities are to provide participants with a foundational understanding of AI including basic definitions. Participants are also guided to develop an intuitive grasp of core Machine Learning concepts such as models, training, inference, and supervised learning. The hands-on experiment using Google's Teachable Machine introduces the principles of feature extraction, labeling, and model evaluation through real-time testing. The use of pre-trained models for music STEM separation serves to illustrate practical applications of deep learning in music processing. Finally, the exploration of generative models for the creation of musical content fosters critical reflection on creativity, randomness, and pattern learning in AI, encouraging participants to evaluate both the potential and the limitations of these technologies. In general, the lab promotes curiosity-driven learning through experimentation and provides an opportunity to engage with real-world AI tools that integrate concepts from multiple disciplines.

The activity was carried out during several editions of the laboratory, involving high school classes from various areas of the Liguria and Tuscany regions. As suggested in [15] throughout the sessions, students were asked to share their thoughts and emotions related to recent news about Artificial Intelligence . Their responses were anonymously collected through the Wooclap platform and visualized as a word cloud. A preliminary analysis revealed that the most frequent keywords were *curiosity* (which consistently emerged as the dominant term in all editions), *interest*, but also *concern*, *fear*, and *anxiety*. These results show the importance of outreach and educational initiatives such as this one, which can help students better understand and critically approach the ongoing digital revolution. The other interesting statistics derived from the Wooclap questionnaires concern the results of the interactive Turing test. In all editions, the majority of participants were unable to correctly identify the track performed by humans among those generated by AI. In the first 3 editions the average number of responses for the correct answer (song B) was between (35%), song C being the most voted answer (40%). The percentage of wrong answers (in particular, for songs C and D) increased in the other editions of the lab. In some editions we also introduce an additional test using three trap songs generated and one minor artist's song. The generated songs turned out to be indistinguishable from the real one (around 5% of votes only), thus showing the possible impact of generative AI in music production.

4. Conclusion

Generative algorithms - aside from ongoing debates over copyright on the music they are trained on - can be useful for artists. However, the true artistic value lies in what a human wants to communicate to another human: the primary goal of these algorithms is essentially to "pass" the Turing test [16]. However, the cognitive process that leads to something produced by generative AI has nothing in common with the human reasoning behind a creative work. Their effectiveness is evaluated solely on the outcome — that is, how convincingly human-like the result appears — rather than on the process that produces it. This is why it is important to develop a solid understanding of these systems: to be aware of their limitations and, consequently, to identify the contexts in which they can be considered reliable tools, and those where they cannot.

In future editions of the activity we are considering to introduce an activity entirely dedicated to AI-based signal processing using tools such as ML-machine[1] and Vitta Science AI[2]. Furthermore, we plan to introduce a more explanatory activity on generative AI using Google's Music VAE [17] for a better understanding of the functioning of generative models. Moreover, we planned to develop an initial and final questionnaire to gather information on the activity's outcome, evaluating both the content learned and the audience's emotional response. This will allow us to perform a meaningful analysis after several editions of the activity, once we have a larger statistical sample.

Acknowledgment This research has received funding from the EU program NextGenerationEU and the Ministry of University and Research, National Recovery and Resilience Plan, Mission 4, Component 2, Investment 1.5, project RAISE (ECS00000035).

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT, DeepL in order to: Paraphrase and reword, Text Translation. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

[1]https://ml-machine.org/
[2]https://it.vittascience.com/ai

References

[1] S. Gera, The impact of Artificial Intelligence on music production: Creative potential, ethical dilemmas, and the future of the industry, 2025. https://nhsjs.com/2025/.

[2] G. Chou, N. Lang, The sound shift: How Generative AI is redefining the music industry's business model, 2024. https://www.artefact.com/blog/.

[3] R. Hennequin, A. Khlif, F. Voituret, M. Moussallam, Spleeter: a fast and efficient music source separation tool with pre-trained models, Journal of Open Source Software 5 (2020) 2154. URL: https://doi.org/10.21105/joss.02154. doi:10.21105/joss.02154, deezer Research.

[4] A. Agostinelli, T. I. Denk, Z. Borsos, J. Engel, M. Verzetti, A. Caillon, Q. Huang, A. Jansen, A. Roberts, M. Tagliasacchi, M. Sharifi, N. Zeghidour, C. Frank, Musiclm: Generating music from text, 2023. URL: https://arxiv.org/abs/2301.11325. arXiv:2301.11325.

[5] C. A. Huang, A. Vaswani, J. Uszkoreit, I. Simon, C. Hawthorne, N. Shazeer, A. M. Dai, M. D. Hoffman, M. Dinculescu, D. Eck, Music transformer: Generating music with long-term structure, in: 7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019, OpenReview.net, 2019. URL: https://openreview.net/forum?id=rJe4ShAcF7.

[6] J. Copet, F. Kreuk, I. Gat, T. Remez, D. Kant, G. Synnaeve, Y. Adi, A. Défossez, Simple and controllable music generation, in: A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, S. Levine (Eds.), Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023, 2023. URL: http://papers.nips.cc/paper_files/paper/2023/hash/94b472a1842cd7c56dcb125fb2765fbd-Abstract-Conference.html.

[7] W. Tan, Defining authorship for the copyright of AI-Generated music, 2024. https://www.artefact.com/blog/the-sound-shift-how-generative-ai-is-redefining-the-music-industrys-business-model/.

[8] J. A. Fails, D. R. Olsen Jr, Interactive machine learning, in: Proceedings of the 8th international conference on Intelligent user interfaces, 2003, pp. 39–45.

[9] M. Gillies, R. Fiebrink, A. Tanaka, J. Garcia, F. Bevilacqua, A. Heloir, F. Nunnari, W. Mackay, S. Amershi, B. Lee, et al., Human-centred machine learning, in: Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems, 2016, pp. 3558–3565.

[10] B. M. Collective, D. Shaw, Makey makey: improvising tangible and nature-based user interfaces, in: Proceedings of the sixth international conference on tangible, embedded and embodied interaction, 2012, pp. 367–370.

[11] M. Carney, B. Webster, I. Alvarado, K. Phillips, N. Howell, J. Griffith, J. Jongejan, A. Pitaru, A. Chen, Teachable machine: Approachable web-based tool for exploring machine learning classification, in: Extended abstracts of the 2020 CHI conference on human factors in computing systems, 2020, pp. 1–8.

[12] Z. Rafii, A. Liutkus, F.-R. Stöter, S. I. Mimilakis, D. FitzGerald, B. Pardo, An overview of lead and accompaniment separation in music, IEEE/ACM Transactions on Audio, Speech, and Language Processing 26 (2018) 1307–1335.

[13] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed., Prentice Hall Press, USA, 2009.

[14] A. M. Turing, Computing machinery and intelligence, in: Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer, Springer, 2007, pp. 23–65.

[15] L. Cesaro, G. Dodero, Generazione digitale, ma non consapevole: giovani e IA fra percezioni e pratiche, in: Atti del Convegno Italiano sulla Didattica dell'Informatica (ITADINFO 2025), 2025.

[16] L. Floridi, The ethics of artificial intelligence: Principles, challenges, and opportunities (2023).

[17] A. Roberts, J. Engel, C. Raffel, C. Hawthorne, D. Eck, A hierarchical latent vector model for learning long-term structure in music, in: International conference on machine learning, PMLR, 2018, pp. 4364–4373.