Evaluating AI as an assisting tool to create Electronic Dance Music

Niklas Bohm[1], Christian Fischer[1] and Manuel Richardt[1]

Hochschule Fulda
christian.fischer@ai.hs-fulda.de

Abstract. The demands on creatives to complete a jingle or a piece of music even under time pressure are growing. This paper analyzes Google's "Magenta" to identify its possibilities for a more effective production of electronic dance music *(EDM)*, especially in terms of time, without a loss of subjective listening pleasure. For this purpose, the process of EDM music production, which includes artificial intelligence, was analyzed. With a subsequent survey, it was determined whether the music pieces produced in this way differ in their subjective listening pleasure and which of the approaches can be recommended for further production.

Keywords: Google Magenta, Electronic Dance Music, AI, Music Production

1 Introduction

Artificial intelligence is used more and more in everyday life and therefore discussed controversially in media. In this text AI is understood as a digital tool to support humans in their creative tasks. However, when one reads about AI in general, some authors find it obviously still difficult to think of AI in connection with creative tasks [6]. There are reservations of creative processes being mapped by a computer [8]. Music is one such creative construct, cause it is capable of triggering feelings in people and amplifying or even influencing moods. Interestingly, most listeners don't notice how much music composition is already influenced by AI and that nowadays many artists have already introduced AI into their creative production work flow [3]. Still, it is not a question about whether AI will make the human composer obsolete, but how creative people use AI in their process [9]. Therefore, the aim of this work is to find out how the use of AI can affect the efficiency of the music production process and whether it has an influence on how the music is perceived. The main focus was efficiency in time and to figure in which steps within the production, in the case of Electronic Dance Music *(EDM)*, an AI can be particularly helpful.

2 Related Work

There is an area of application for artificial intelligence in almost every field. Among others, also in the creative areas of music and art. Because AI has so many forms, researchers divide it into the subfields of automation, machine learning, neural networks, and deep learning [2]. A common form of neural network is the Convolutional Neural

Network *(CNNs)*. In this paper, and the Magenta project respectively, Recurrent Neural Networks *(RNNs)* and the Long Short-Term Memory Network *(LSTM)*, a type of RNN, are central [11].
While CNNs tend to be hierarchical, RNNs operate more sequentially [1]. Therefore, CNNs are often used for classification tasks such as text recognition, whereas RNNs are more flexible and are therefore increasingly used for language processing and creative tasks such as music generation [11]. The project "Magenta", placed under Open Source by Google is a project which explores machine learning in a creative context. Magenta Studio, basically a collection of music plugins, uses RNN and LSTM networks. To understand how a computer can utilize this to make music, it is important to take a closer look at these RNNs. An RNN is a class of neural networks used to model sequence data. In an RNN, the connections between artificial neurons form a circuit. For many applications such as speech and text, outputs can depend on previous inputs and outputs. The key concept of RNNs is to use sequential information. RNNs are thus given a "memory" in the form of the information and data that have been processed so far [1]. Related to the music context in this paper, this means that the neural network can, e.g. recognize which notes it has generated in previous measures and is able to adjust the next measure to sound coherent. The size of the "memory" determines how many notes, or even bars the neural network can look back. The goal is to have the generated sounds or musical sequences which contain repetition, as repetition is one core feature in music.

3 Environment Setup and AI-Created Music

Magenta is the name of several research projects by Google. The program Magenta Studio, serves as AI support for digital music producers, as a standalone application or when used as a Plugin e.g. with the Digital Audio Workstation Ableton Live. Magenta has five different models, using different neural network types. In this study LSTM networks, a type of RNN, are used. An example of an LSTM network is the DrumsRNN model. Another example is the MelodyRNN model, which has the same structure but is programmed for melody generation [4]. Each model has different configurations that set the way in which it encodes the input data. For example, the DrumsRNN model has a one drum configuration, which stores a drum sequence in a class, and a drum kit configuration, which splits a drum sequence into nine different instruments (kick, clap, hi-hat, etc.) and adjusts the attention length [4]. Magenta provides many pretrained models for download on its site, here the models "DrumsRNN", "MelodyRNN" and "PolyphonyRNN" were used.
To structure the process of digital music production and to make the influence of AI in this process more comprehensible, we divided it into three levels. The higher the level, the less human involvement in the production process.
Level 1 - Inspiration: In this stage, the human producer is supported or inspired by the AI. For example: AI generated melodies are listened to. Being inspired, the human producer composes a new melody which contains no or hardly any parts of the original AI melody.
Level 2 - AI Assistance/Co-production: In AI assistance, one gets support from an AI

in the creative process. Here, bigger parts of AI-generated music can be incorporated into the composition. The AI can be used to help, e.g. by suggesting subsequent notes. Since the AI and human are both working on the same piece, they can both inspire each other.
Level 3 - Automation: In automation, an AI generates new music on its own without input. The producer has no influence on the results. One of many examples of autonomous music generation would be Magenta's Generate function [10]. One other example is AIVA, which stands for Artificial Intelligence Virtual Artist. AIVA has already released self-produced albums in the EDM genre. It was created in early 2016 and has been under constant development ever since [5].

4 Tests Structure and Efficiency Evaluation

First three simple EDM tracks, according to the three stages, each consisting of the three main components bass, chords and melody, were composed. For the later evaluation, some measurement data had to be collected. Thus, the production duration and effort were measured based on mouse movement, mouse clicks and key usage, utilizing the free software Mousotron (Win; version 12.1; Black Sun Software 2017). In order not to confuse the participants of the later survey by different drums and drum patterns in the songs, an identical one was adopted here for all. For the music production itself the Digital Audio Workstation Studio FL (Win; version 20.8; Image-Line Software; 2020) was used. One aim was to verify whether the introduction of AI into the production process is suitable for speeding up and simplifying music production. For this purpose, the individual steps of the production were calculated, and all mouse clicks and movement were measured as the second and third measured values.
The evaluation (collecting data of mouse and keyboard usage) was carried out through the different phases to determine and track the highest increase in efficiency on specific production phases and production steps. In this way, phases for which AI is particularly suitable can be identified (see also Table 1).
Preparation phase: In the preparation phase, human production performed best. Preparation took just under a minute, with both AI automation and AI assistance at around three minutes.
Composition phase: In the composition phase, the tendency of the first phase changes. Here, human production took by far the longest at around nine minutes. With the help of AI assistance and AI automation, up to 85% of the time used could be saved.
Editing phase 1 (without mixing and sound design): In this phase, it is noticeable that the AI assistance has particularly high values. To adjust an AI given melody to one's own desires might take time. The duration of the production phase was therefore about four to five times longer for the AI assistance than for the other two production types.
Editing phase 2 (mixing and sound design): Great inspiration can come from sound design, as melodies can also sound very different depending on the sound. Since a producer usually creates a melody after the sound design, this phase also takes about twice as long as the other production types. With the support of AI, one comes to a time saving of about 65% here. The AI assistance has the lowest effort in this phase, but this could change if the experiment is performed multiple times.

Human production took the longest in terms of total production time. However, this is closely followed by AI assistance. The use of AI automation reduced the duration of production by 40%. For composition only (without processing phase 2), there was even a 42% reduction. AI assistance took about 23% longer than a human production in the pure composition process.

Phase	Song	Duration	Mouse clicks	Mouse distance	Keystrokes
Preparation	Human production	01:12	22	3,51	24
	Automation	03:20	64	18,56	76
	AI-Assistance	02:52	53	15,45	57
Composition	Human production	09:07	203	20,57	239
	Automation	01:43	16	1,41	12
	AI-Assistance	01:23	16	1,56	14
Editing 1	Human production	02:34	41	5,87	12
	Automation	03:02	36	7,19	95
	AI-Assistance	13:01	227	30,28	433
Editing 2	Human production	07:48	141	21,34	78
	Automation	03:47	67	12,17	67
	AI-Assistance	02:34	40	8,41	52

Table 1. Measurements in different production phases

5 Listening Evaluation

The second aim of this study was to find out how listeners like the generated songs and whether they can tell the difference between an AI-generated and a human production. To investigate these questions, 50 participants with different musical backgrounds and individual tastes (in balanced proportions) were asked to do an online questionnaire stating their opinions about the three generated songs.

After general questions (age, gender, knowledge/involvement in music and knowledge of EDM), the participants listened to the songs and were asked to rate the songs based on the following eight evaluation criteria: Creativity, Recognition Value, Arrangement (of a song), Cohesiveness, Variety of Tones, Energy, Danceability, and Emotion. Through a study conducted by the University of California Irvine, these criteria were identified as crucial. The study found that songs are especially popular when they are more upbeat and danceable than other songs [7]. In addition, dynamics and the mood of the listeners play a role [7]. Other criteria such as the vocals or the genre are not considered in this paper it focuses exclusively on instrumental EDM music.

The participants could rate the different criteria from 1 to 5 points. One meaning low and five meaning high. All points from all 50 participants that were given for one song were added to an overall score for that song (see Table 2).

Song 1 - Automation: The first song was generated by the AI alone. The creativity of the song was rated an average of 3 and is therefore exactly in the middle of the scale. The recognition score averaged 2.7 points and the arrangement was rated at 2.76 points

Song	Creativity	Recognition Value	Arrangement of tones	Cohesiveness	Variety of Tones	Energy	Danceability	Emotion
1	3	2.7	2.76	3.22	3.19	3.8	3.4	2.47
2	3.39	3.43	3.8	3.76	3.27	3.43	3.65	3.14
3	3.82	3.58	3.56	3.39	3.37	3.66	3.37	2.98

Table 2. Results of the questionnaire

on the scale. For the AI-generated song, energy got the highest score with an average of 3.8 points on the scale. This was followed by danceability with 3.4 points and cohesiveness with 3.22. Emotion was rated the worst with a mean of 2.62 points. Overall, Song 1 collected 1228 points in all criteria by all participants or an average value of 3.08 points on the scale.

Song 2 - Human production: The second song was produced by humans. The arrangement scored best with an average of 3.8 points on the scale. This was followed by cohesiveness with 3.76 points and danceability with 3.65 points. Emotion was again rated the worst with a mean of 3.14 points. Overall, the song was rated with 1365 points and achieved an average value of 3.49 points on the scale.

Song 3 - AI assistance: The third song was produced in collaboration between a human and an AI. Creativity received the highest score here, with an average of 3.82 points on the scale. This is followed by energy with 3.66 points and recognition with 3.5 points. Emotion was again rated the lowest with a mean score of 2.98 points. Overall, the song was rated with 1368 points and achieved an average value of 3.42 points on the scale.

In comparison, the human-produced song, and the AI-human collaboration both scored about the same. The AI-generated song scored 1228 points, about 10% worse than the other songs. Creativity is rated the highest for the AI-assisted production, and the recognition value is also almost 30% higher. The human production was considered less creative, but the arrangement was easy for the listener to understand, and the song seemed cohesive. In the AI production, the rhythmic criteria such as danceability and energy stood out.

6 Summary and Conclusion

The experiment was designed to test whether the use of AI in digital music production can increase the time efficiency. In the experiment, three songs were produced with the same prerequisites, but each in a different production mode. Since AI can be used to different degrees in music production, three levels were developed: AI inspiration, AI assistance, and AI automation. As a result, it was expected that the producer will have significant savings in time and effort by using AI. However, some of the results differed from the assumptions. A digital music production was divided into four phases. The experiment showed that the use of AI in production increased efficiency especially in the composition phase and the second editing phase. There were no significant differences in the overall production ratio in the preparation phase, but this phase was slightly faster with human production. Overall, the AI automation was convincing with an efficiency increase of 41% on average. AI assistance decreased production efficiency by about

20% on average.
The created songs from the experiment were listened to, rated, and evaluated by 50 people. The result of the survey was that the computer-generated song was rated about 10% less, than the other two songs. It is concluded that although an AI can write songs and melodies on its own, they are perceived not as melodic and creative as productions that involved a human. Since AI assistance was rated 30% better in terms of melodic criteria, it can be stated that the introduction of AI into the production process can exceed these very limits.
Thus, for the further development of AI-assisted programs for music production, it is important to develop AIs that see a song as a whole and when the "memory" (e.g., in LSTM) is large enough to remember a theme or idea in a song and repeat it at the right places. For EDM producers, it should be noted that the introduction of AI as a tool into the right phase of the production process has many positive aspects such as saving time and effort and supporting the creative process of composing melodies.

References

1. Bisharad, D., Laskar, R. H.: Music genre recognition using convolutional recurrent neural network architecture. Expert Systems 36 (4) (2019)
2. Dargan, S., Kumar, M., Ayyagari, M.R., Kumar, G.: A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning. Archives of Computational Methods in Engineering 27 (4), 1071–1092 (2020)
3. Deahl, D.: How AI-generated music is changing how hits are made. The Verge, https://www.theverge.com/2018/8/31/17777008/artificial-intelligence-taryn-southern-amper-music (2018)
4. DuBreuil, A.: Hands-on music generation with Magenta (E-Book). Explore the role of deep learning in music generation and assisted music composition. Birmingham, Packt (2020)
5. Hewahi, N., AlSaigal, S., AlJanahi, S.: Generation of music pieces using machine learning: long short-term memory neural networks approach. Arab Journal of Basic and Applied Sciences 26 (1), 397–413 (2019)
6. Holeman, R.: Why A.I. Can't Replace Human Creativity. Modus Medium, https://modus.medium.com/when-it-comes-to-creativity-is-it-game-over-for-humans-ef12907eb30d (2019)
7. Interiano, M., Kazemi, K., Wang, L., Yang, J., Yu, Z., Komarova, N.L.: Musical trends and predictability of success in contemporary songs in and out of the top charts. Royal Society open science 5 (5) (2018)
8. Müller, V.E.: Risks of Artificial Intelligence. In: Vincent C. Müller (Hg.). Risks of artificial intelligence. Boca Raton, FL, CRC Press, 1–8 (2020)
9. Pennington, A: Artificial Intelligence gets creative. IBC, https://www.ibc.org/how-ai-is-being-used-to-boost-the-creative-process/2937.article (2018)
10. Roberts, A., Engel, J., Mann, Y., Gillick, J., Kayacik, C., Nørly, S., Dinculescu, M., Radebaugh, C., Hawthorne, C., Eck, D.: Magenta Studio: Augmenting Creativity with Deep Learning in Ableton Live (2019)
11. Yin, W., Kann, K., Yu, M., Schütze, H.: Comparative Study of CNN and RNN for Natural Language Processing (2017)
12. Mousotron, https://mousotron.de.softonic.com/
13. Studio FL, https://www.flstudioshop.de/fl-studio-/