AI-Composed Music for User Preference Using Reinforcement Learning

Dominik Myśliwiec
University of Twente
Enschede, The Netherlands

ABSTRACT

Artificial intelligence has become a powerful tool in automatic music generation and music recommendation systems thanks to the field's rapid development. Nonetheless, few solutions have been proposed for easily generating music based on a non-expert user's preference. The results of an extensive literature review show that the main issue to combat was that of subjectivity in music, and gave an indication that a reinforcement learning algorithm combined with a deep learning model for user preference could be an appropriate solution. Based on these findings a new automatic music composition system that relies on a reinforcement learning algorithm and models preferences based on the user's ratings of transformer-generated pieces, improving tailoring to the user over iterations of the algorithm. The system was evaluated through human interaction and shows promising results with a 44.7% improvement to the mean user rating of generated pieces in 6 iterations of the algorithm. The rating prediction model also achieves high accuracy with a difference between the predicted and received ratings of just 0.81 on a 10-point scale. Because the system relies on no prior knowledge, with such effectiveness it could give access to music compositions for user preference to a wider non-expert audience.

CCS CONCEPTS

• **Applied computing → Sound and music computing**; • **Computing methodologies → Machine learning**.

KEYWORDS

music generation, user preference modelling, music transformer, music preference

ACM Reference Format:
Dominik Myśliwiec. 2023. AI-Composed Music for User Preference Using Reinforcement Learning. In *Proceedings of 39th Twente Student Conference on IT (TScIT 39).* ACM, New York, NY, USA, 9 pages.

1 INTRODUCTION

Artificial intelligence is taking the world by storm. Among others, it also found many uses in the music world, whether it is transforming songs to sound like they are covered by a different artist, recommending music to listeners of popular streaming services, or predicting the future popularity of newly-released songs. Systems capable of generating new musical compositions are some of the other many examples of applied use of AI in music [2, 7, 14]. Such systems could not only provide entertainment, or be used to educate about music theory and music creation, but they can also provide high-in-demand music in the world of marketing. Regardless, automatic music-composing solutions are still nowhere near as popular as large language models such as ChatGPT-4 [22] or image generators such as DALL-E 2 [6].

One of the potential reasons for the lack of popularity of automatic music composers is their capability to consider the user's preference. Currently, popular systems require inputting a list of the user's favourite songs or a very detailed description of the expected composition. The first approach requires significant effort and may be ineffective when gathering a big enough dataset is not possible. The second approach, on the other hand, is problematic for non-expert users who are unable to express their preference in terms of providing enough detail. It is therefore important to investigate other options of automatic compositions for user preference in these situations.

Reinforcement learning is a powerful method of navigating unknown environments to reach a specified objective. The presented problem can be described as an environment consisting of all possible compositions, where the objective is to maximize how fitting a composition is to the user's preference. Often, the advantage of applying reinforcement learning is that, if done correctly, it requires minimum effort from the user and can maximize its objective without prior knowledge. This paper will therefore focus on answering the following research question:

RQ How can artificial intelligence compose music for user preference using reinforcement learning?

In this paper, a thorough and systematic literature review is conducted to identify the main problems of music composition and possible solutions to them. Based on the result of the review, a design for a new system is proposed that could serve as a potential solution to the problem of music preference extraction and preference-based music generation without prior preference knowledge. The system will then be implemented and evaluated through human interaction.

2 RELATED WORK

Multiple methods will be used in this research. The first step in answering the research question is conducting a systematic literature review. The review must be extensive and cover state-of-the-art

TScIT 39, July 7, 2023, Enschede, The Netherlands

solutions, drawbacks and advantages of existing systems, and potential future improvements. The literature review aims to bring the research as close as possible to answering the posed research question. Based on the findings of the literature review, a new system will be proposed as a way of generating music to user preference without prior knowledge.

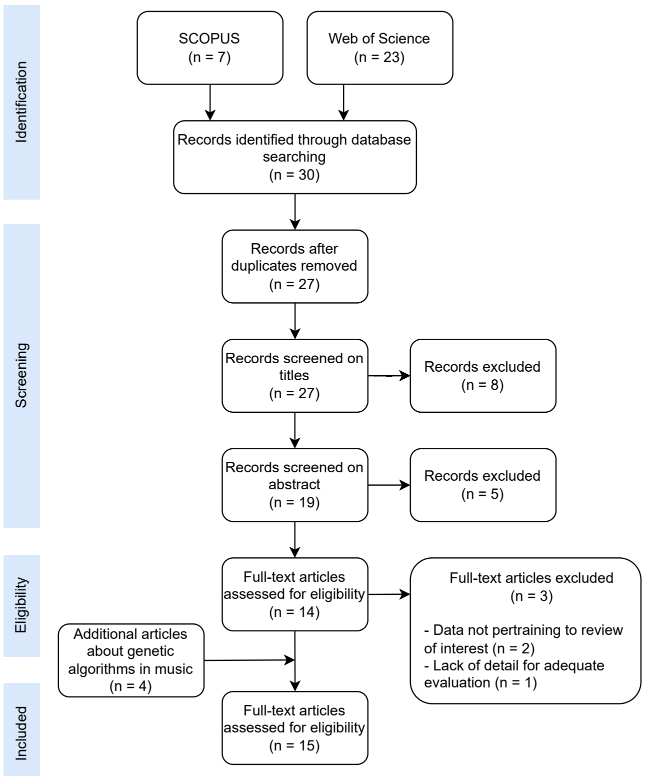

Figure 1: Flow diagram of the conducted search according to the PRISMA 2020 statement guidelines.

Method

To ensure a systematic approach that allows for transparent reporting, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [23] was used. The PRISMA statement allows reproducibility and specifies how the search and selection of the research process can be represented using a PRISMA flow diagram as seen in Figure 1. The goal of this process is to find answers to the earlier stated research question in existing scientific literature. To do so, a search has been conducted in two scientific research databases: Scopus and Web of Science. Similar search strings have been used to produce a more significant number of results, which were then further filtered in the process of selection based on the usefulness of the paper in answering the research question. Figure 1 contains more information on the number of search results and the number of available results meeting the criteria.

Search queries

Scopus — The main query used in Scopus was "preference AND "music generation" AND ("artificial intelligence" OR "machine learning" OR "neural network")". Additionally, another query was used to specifically research the use of genetic algorithms in music " ("music generation" AND "genetic algorithm") ". This query produced 34 results, 30 of which were deemed irrelevant.

Web of Science — The query used in Web of Science was "ALL = (preference) AND ALL = (music generation) AND (ALL = (artificial intelligence) OR ALL = (machine learning) OR ALL = (neural network))"

Several found papers listed in Table 2 center around the idea of music recommendation rather than generation, but the process is also based on user preference extracted through their listening or streaming data [28, 27, 4, 25], which is relevant to the scope of this paper.

A full overview of all found articles is presented in Table 2, while a more specific overview of what aspect of music generation is relevant for each paper, as well as what implications the findings of the paper might have for the design of the system proposed in this paper can be found in Table 1.

2.1 Subjectivity

To understand the problem of music composition and why it is not more popular, apart from analyzing the findings of the literature the main difficulties faced were also identified. The following three main difficulties of music composition can be recognized in the papers:

Format limitations Symbolic representations, such as musical scores or Musical Instrument Digital Interface (MIDI) are a simplified way of describing music. Many aspects of sound cannot be encoded within these formats. Although it might seem that a simple solution would be to use waveforms or raw audio data, research shows that this approach is significantly more difficult due to the much higher complexity of the medium [8].

Short and long term patterns Music is unique in consisting of strong patterns over the entire length of a track (so-called song structure) and patterns between notes following each other to create a melody. Awareness of both kinds of patterns is troublesome to model [13].

Subjectivity Many of the automatic music composition or recommendation systems are based on the principles of artificial intelligence, which requires a clearly defined objective function. Music is highly subjective, and therefore it is difficult to define what universally "good" and "bad" mean in the context of creating a new piece. To address this issue, one option might be to attempt to tailor to each specific user separately. This would require a way to evaluate composed pieces based on user preferences, which in turn adds another dimension to the problem.

Table 1 indicates that although the format of music representation and its complex structure pose difficulty in music composition, the subjectivity of music preference is clearly the most relevant of

Table 1: Overview of what area of music generation complexity is relevant in the paper's presented problem and what its results and limitations (as seen in Table 2) mean for the design of the system proposed in this paper. A bold "X" under an area of the problem indicates it was the most major of the faced difficulties, while an "x" means it was relevant but to a lesser extent.

Reference	Area of problem			Implication for system design
	Format	Structure	Subjectivity	
[5]		**X**	x	If a song is preferred by a user, another one with similar structure and rhythmic qualities will also be liked by them — the system should utilize that to generate music more similar in these features to the liked pieces
[17]			**X**	A deep neural network can be used to successfully model user preference for automatic music generation
[20]		**X**	**X**	An interactive iteration-based algorithm can be applied to extract user preference based on their feedback Transformers can successfully compose music given user preference
[16]	x	**X**	**X**	An iterative reinforcement learning algorithm based on user feedback has a high potential in music generation
[27]	x		**X**	Low-level features of a piece may be used to estimate whether the user will prefer it
[28]	x		**X**	Reinforcement learning may be used to train a model that predicts user rating given a song's musical features
[25]			**X**	Most recent user interaction should be used to estimate short-term music preference
[21]		**X**	x	Subjectivity should be part of the objective function, otherwise, results of music generation vary a lot between users
[31]		x	**X**	An interactive algorithm with similarities to the genetic algorithm can improve generating music for user preference over generations
[4]			**X**	Explicit feedback in the form of user rating can be used as a parameter for learning user preference in a reinforcement learning solution
[26]	**X**		**X**	The developed algorithm should be based on iterative human interaction to optimize for user preference
[19]		**X**	**X**	A real genetic algorithm requires high parameter adjustability and a clear heuristic — without them, a different iterative and interactive algorithm might be more appropriate
[29]			**X**	The system should be studied through many iterations to investigate signs of early convergence, although the study suggests it might continuously improve
[32]			**X**	Asking for user feedback iteratively from non-expert users is a successful method of understanding their preference in the context of music
[18]			**X**	Music preference is non-linear, therefore, similarly to a genetic algorithm, a deep learning approach is appropriate in composing music for user preference

the three. It is the major problem that almost all the found papers attempt to resolve. This is not unexpected, as music tastes differ greatly between people and it is difficult, if not impossible, to define objectively "good" music.

2.2 Modelling user preference

The main differences between the explored systems are the features they model and what kind of models they use. One approach is to use a Bayesian regression model to estimate user preference [28]. Such a model can successfully be trained through reinforcement learning. Due to advancements in the field, newer research suggests different models. The most recent paper focused highly on the emotionality of music [27]. It relies on a deep neural network used to estimate the emotion of a music track using the valence-arousal model [3] given certain low-level features of the track as input. The system for music recommendations is based on existing preference information from the user, as well as collaborative filtering, which takes into account such aspects as the user's social media presence and major events. It can be argued that this system is based on older research from 2010, in which a system was proposed that estimates music emotion and user preference based on a window of their recently listened to songs [4]. This paper already showed success in using user feedback to understand their preference.

What can be deduced from these articles, is that there are many ways to model user preference that can be successful. The method of modelling is highly dependent on the purpose of the model — a system should model preference for the same features of music that it can extract to be relevant and useful. This is logical, since the preference for the genre of music, for example, would have little to no value in a system that recommends music of a certain emotion. Additionally, it shows that even low-level features of a music piece can be useful in reinforcement learning-based training.

2.3 Generating music for user preference

There is also significant research on using user preference in generating music and sound, with vastly differing approaches. One such paper suggests using an interactive genetic algorithm to control the direction of audio generation for sound effects using popular preference [26]. While showing improvement thanks to iterative human interaction, the system still faced challenges, one of which was using the popular preference, rather than tailoring to a specific user. This is troublesome since preference has a high variance between different users.

Two other deep-learning-based music composition tools have been studied [20, 21]. The papers show that the subjectivity of music quality is a crucial aspect to take into account when composing. These systems use an iterative process, in which, using the user's feedback, a song or chord progression can be generated step by step. Although the approach is interesting and shows improvement over time, it does not provide the framework for the instantaneous generation of entire songs or tracks but rather provides aid to the creator of a piece. Another invention in this domain is an accompaniment generator, the aim of which is to create a duet with an already existing music track, or improvisation [16]. This system is another example of a promising implementation of an iterative algorithm based on human interaction.

The most relevant papers for answering the posed research question, however, aim to generate entire songs taking user preference into account [5, 17]. One of these approaches even proposes an iterative learning process in which a deep neural network models user preference and is fine-tuned using user feedback [17]. This system, however, still relies on a large dataset of the most popular hit songs and struggles with a bias toward the exploitation of assumptions as opposed to the exploration of new songs. Finally, Yamaguchi and Fukumoto [31] propose an interactive genetic algorithm (IGA) that consists of iterations of melodies being generated and evaluated by users to understand their preferences, based on which song recommendations are made. Though the recommendations did not become better over generations, which might potentially be related to the correlation between the simple model based on melodies as opposed to complex songs, there was a significant improvement over the many iterations of the algorithm when it comes to the generated melodies. At the end of their experiment, the system was able to generate melodies much more preferred by the user than those generated in the beginning.

The reason Yamaguchi and Fukumoto's [31] research is so relevant is that not only does it showcase a simple way to connect a preference model to music generation, but it also proves that a system could potentially work without any initial input from the user or knowledge about their listening habits or song ratings. To a non-expert user, the process of simply listening to multiple generated melodies and rating them or selecting the favourite is very approachable and easy to navigate. This provides a great foundation for developing a new generation system and answering the research question.

2.4 Optimization Algorithms in Music Generation

Because the amount of pleasure derived from a piece of music is very subjective, usually optimization for an automatic music generator/composer is required. Many proposed systems rely on different forms of genetic algorithms. One possible use case of this type of algorithm is utilizing researched heuristics of general music enjoyability (for example types of intervals between every two next notes) [19]. Similarly, Wiafe et al. [29] use certain features of a music piece (chord progression, vocal smoothness, etc.) to optimize a composition system and find that the enjoyability of the music continuously increased with generations with their approach. Another possible approach is to use an interactive genetic algorithm instead, which replaces fitness functions with a human evaluator, allowing for tuning specifically to the user's opinion. Zhu et al. [32] use this approach to successfully generate music based on the emotionality of the piece desired by a non-expert user. In case the objective of music generation cannot be easily modelled, a powerful method such as deep learning may be used allowing for higher abstraction and complexity. A combination of a genetic algorithm and deep learning can be successful in predicting a user's satisfaction with a piece regarding its pleasantness, which may be otherwise difficult to model numerically [18].

Although the interactive genetic algorithm seems like a very promising method, it requires the possibility of crossover and mutation of a population of generated pieces, which might be more difficult when using a pre-trained transformer, a promising method of generating music. In that case, the deep learning element of a system like the one used by Majidi and Toroghi [18] may be used in a reinforcement learning approach. The loop of reinforcement learning can be adjusted to include some elements of used genetic algorithms, such as rating a certain population in each loop to improve the population of the next group. Such an algorithm has the potential to quickly learn the user's preference by taking inspiration from interactive genetic algorithms.

2.5 Literature Review Conclusions

The literature review provides a good basis for the design of a system for music generation for user preference. The following conclusions may be drawn based on the findings of the papers:

(1) One of the hardest obstacles in automatic music composition is the subjective nature of the field. Since music tastes differ greatly, tailoring to the preference of each specific user separately might be required.
(2) Research shows that many methods for user preference modelling may be successful, but one should choose a method depending on the possibilities and requirements specific to the used system. That being said, for the purpose of modelling something as abstract as user preference, a complex and abstract approach such as deep learning may be the most appropriate.
(3) It is possible to successfully improve melody generation to the user's preference using user feedback without any initial input based on listening habits or song ratings. One tested method to do so is to use an interactive genetic algorithm.
(4) Taking inspiration from an interactive genetic algorithm, a reinforcement learning approach using user feedback may be used to extract the user's preference and optimize the composition of music.

With the knowledge provided by the review of existing literature, it is possible to design the system. The combination of the findings from all analyzed papers as found in Table 1 can be used to design a system based on an iterative and interactive algorithm capable of learning the preference of any non-expert user and generating music based on it. The system is expected to utilize a model that will show improvements in the accuracy of user rating predictions, as well as continuously improving music composition to the user's preference, which can be measured by the means of user rating. Whether the system meets the expectations will be tested through an experiment based on human interaction with the system.

3 SYSTEM DESIGN

In this step of the study, a system is designed using standard design science steps. First, the problem is identified. Next, the objectives of the system are defined to establish the expectations of the prototype. Then, a high-level prototype is developed and evaluated. Sadly, although usually the system design process is iterative, due to the time constraints of the study, only one cycle can be completed. Nevertheless, the results of the evaluation are still valuable and can be a driver for future work. The results of the literature review will be crucial in establishing what problems must be addressed in the design of the system, as well as in identifying potential solutions to the problems. To evaluate the system after its development, a human evaluation will be conducted based on interaction with the system in the intended way and the collection of basic data that will allow us to judge the effectiveness of the system. The form of the human evaluation will be a carefully planned experiment.

3.1 Problem identification

One of the primary problems in music generation is subjectivity. This issue is even more transparent when it is difficult for a user to input their preference, whether it is a set of existing songs or a text prompt. An example of such a situation might be generating music for a video game by a person unfamiliar with existing pieces. Therefore the system needs to satisfy the following requirements:

(1) The system must generate music
(2) The system must adapt to the preference of the user
(3) The system should not require any expertise or knowledge from the user

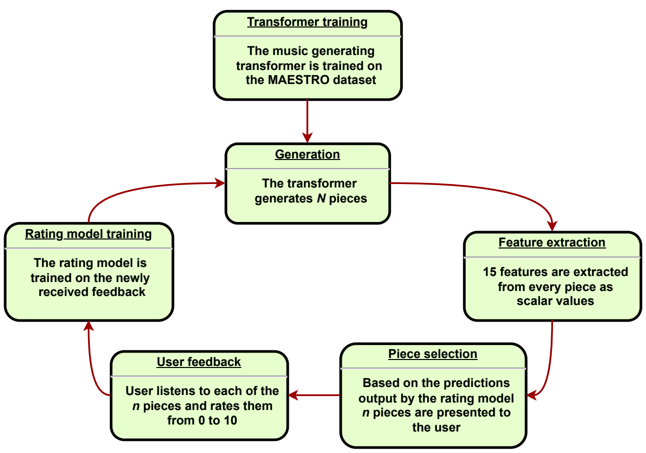

Figure 2: Flow diagram depicting the functioning of the proposed system

3.2 Solution proposal

The majority of state-of-the-art music generators rely on artificial intelligence models. Thanks to rapid advancements in the field, these models are versatile, reliable and flexible. To produce results that reflect the user's expectations, preference must be extracted in some way. One way to do this would be to ask the user for a rating of a generated composition. Then, through many ratings, a model may be created which reflects the user's likes and dislikes. Instead of a time-consuming process of asking the user for many ratings at once, a reinforcement learning algorithm may be used to continuously improve on generating music to the user's preference while receiving new feedback. Drawing inspiration from interactive genetic algorithms, the reinforcement learning loop consists of generating a population of N pieces, of which n (such that $n < N$), that are predicted to receive the highest rating from the user, are chosen to be presented to the user. The user then rates the presented pieces. The rating agent (utilizing a deep neural network) will then be trained to more accurately predict the preference ratings of the user in the future based on the feedback. These loops may be repeated indefinitely until a desired musical composition is reached. The depiction of the steps essential to the functioning of the proposed solution can be seen in Figure 2.

3.3 Prototype

A prototype was developed for the purpose of the experiment. This process consisted of collecting an appropriate example database, training a model, and preparing the experiment environment.

3.3.1 *Training data.*

A large amount of training data is required to achieve an optimal music-generating transformer. Additionally, since the chosen format for the generator model was MIDI, the musical pieces should all be available in the given format. Because the time for executing the experiment is limited, it is also more optimal to train a model generating the performance of a single instrument rather than an orchestrated piece. An available database fulfilling all the criteria is the MAESTRO dataset [12], consisting of many hours of virtuosic piano performances of classical pieces, all available in MIDI format. More specifically, version V2.0.0 of the dataset was used.

3.3.2 *Generator model.*

A trained model of a music-generating transformer is the most crucial element. As a significant breakthrough in the field of image generation, transformers are neural network models consisting of a decoder and an encoder [15]. They are capable of learning from the context of sequential data and were originally meant for use in language translation. Because text generation (done by the decoder) is part of this process, these models are capable of creation based on the learned patterns and have since found application in other domains than just text. Transformers may also be adapted to be used in music generation. Similarly to image generation, they provide very promising and convincing results, and are a well-researched, and therefore very popular, method for music generation [9, 10, 24, 30, 14]. One of the most successful implementations of music-generating transformers is the "Music Transformer" [14]. A reproduction of the Music Transformer was used in the experiment, which was largely possible thanks to an open-source repository available with the foundations of the system implemented [11].

3.3.3 *Preference model.*

The used model for user preference prediction was a fully connected deep neural network (DNN) with an input dimension of 15 scalar values and an output of a single value between 0 and 10 — the predicted rating of a musical piece. The DNN consisted of 3 hidden layers, each consisting of 128 neurons and a ReLU activation [1]. The 15 input values used as the input of the model were features extracted from generated pieces — the number of notes, tempo, number of beats, as well as the mean, median and standard deviation of each of the following: note lengths, note pitches, note velocities, and the number of notes per second.

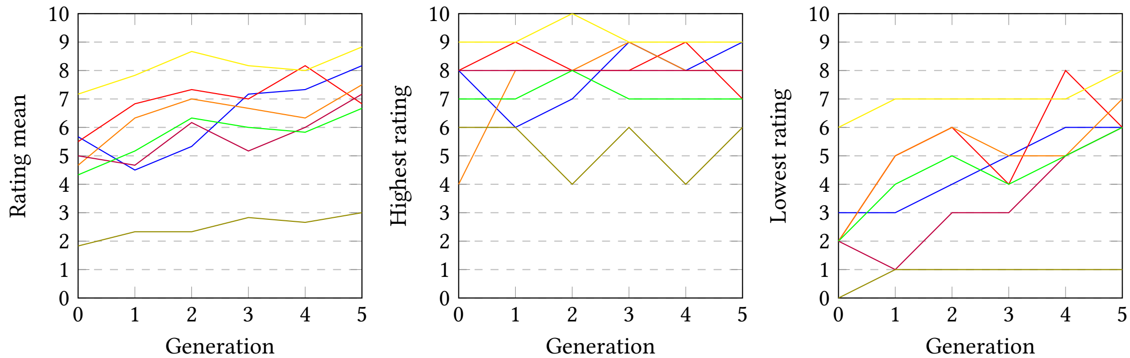

Figure 3: Plot representations of experiment results. Presented information (from left to right): rating mean, highest rating, and lowest rating per generation for each experiment. Each colored line represents the evaluation of a single user, and the colors are consistent between the three graphs as well as Figure 4.

4 EXPERIMENT

Due to the subjective nature of music enjoyment, this element of the research is very valuable. Once the system was developed, an experiment was conducted with the goal of evaluating the system in practice. A significant portion of the experiment consisted of the users interacting with the prototype system in order to fine-tune the model to their preference and evaluate it. This process was observed and measured for statistical analysis.

4.1 Experiment design

In the experiment, 7 participants interacted with the system as a potential user would with the goal of generating a music composition to their preference. Each participant was asked to complete 6 cycles of the algorithm, where a cycle begins with music pieces being generated and ends with the rating model being trained on the user feedback. In each cycle of the algorithm, $N = 30$ music compositions were generated, out of which the $n = 6$ best were presented to the participant based on the rating model's predictions. The participant was then asked to listen through each composition and rate it on a 10-point Likert scale, where a 0 rating is described as "not to my liking at all" and 10 as "the piece could not be more to my liking". Because the prototype transformer is only able to generate piano compositions imitating classical music performances, the participants were made aware of this beforehand.

For all participants the same music-generating transformer was used, with the same parameters and the same training, however, each experiment began with the rating model initialized to always predict 5 as the rating. The rating model was only trained based on the feedback of one particular participant. This is because the purpose of the algorithm is not to fine-tune the music transformer but rather to test the effectiveness of interacting with the system as a user would in order to generate a particular piece to their liking.

For each participant, the following information is recorded about their interaction with the system:

- The rating of each of the 6 compositions per generation (reinforcement learning cycle)
- The predicted rating of each of the pieces by the rating model before training in this cycle

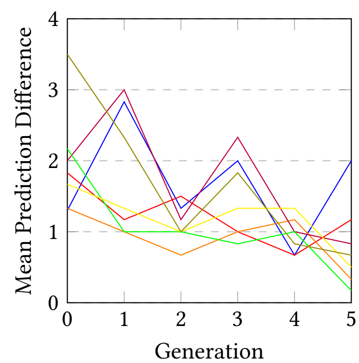

Figure 4: Plot showing the mean difference between ratings predicted by the rating model and ratings given by participants over generations of the algorithm.

- The differences between the expected and the real rating for each composition
- The rated compositions are also saved, though this has no relevance to the results of the experiment

This information will be used to assess how effective the system is in generating music to the participant's preference, but also measure how accurately the rating model is able to predict the ratings received from the user.

4.2 Results

For each participant, the plotted rating mean, as well as the highest and lowest rating per generation can be seen in Figure 3. Additionally, for the purpose of assessing the rating model, the difference between received user ratings and expected ratings may be seen in Figure 4.

For every participant, the mean rating increased significantly from generation 0 to generation 5 with a mean increase of the rating mean of 44.7%. Interestingly, although one of the participants rated all songs much lower than the other participants, the mean rating showed a steady increase for them as well. Furthermore, the mean rating difference between the last and first generation was the highest for this participant at 64%. On the other hand, the participant who found the generated pieces most to their preference from the beginning and consistently rated pieces highly showed the lowest mean rating increase of 23%.

The rating model also shows satisfactory accuracy, as seen in Figure 4. On average, after just 6 generations of the algorithm, the predicted rating was just 0.81 away from the rating given by the participant. Furthermore, for 6 out of 7 participants, the mean difference between prediction and received rating at generation 5 (the last of the experiment) is in the $[0.17, 1.17]$ range.

5 CONCLUSIONS

5.1 Literature

In this paper, an attempt was made of answering the research question: "How can artificial intelligence compose music for user preference using reinforcement learning?". The literature review conducted to understand the problem and existing state-of-the-art solutions provided the following findings:

- Subjectivity is the main issue of composing music for user preference
- Deep learning can be successfully applied to model user preference
- Melody generation to user preference can be improved over time using an interactive algorithm
- Reinforcement learning can be used to successfully extract user preference for music

5.2 Solution

Applying the main findings and other more specific suggestions from the literature review, a new system was designed. The proposed solution utilizes two machine learning models — a transformer for generating music and a deep neural network to predict what rating a generated piece might receive from the user. The rating model is trained through a reinforcement learning loop inspired by a genetic algorithm, where in every loop a population of pieces is generated by the transformer, and a part of the population selected by the rating model is rated by the user based on their preference. These ratings are then used as additional data to further train the rating model. A human evaluation in the form of an experiment conducted with 7 participants interacting with the system prototype showed promising results.

5.3 Evaluation

Although the result shows that the mean rating improves on average by 44.7% over 6 generations for each participant, this does not mean that the highest-rated single composition was achieved at a later generation of the algorithm (see Figure 3). This means that in each generation a larger portion of the rated population fits the user's preference. It is safe to say that the system is successful in improving the rate at which high quantities of preferred pieces are composed, but not necessarily how quickly a single highly preferred piece can be created. It is also clear, that very low ratings appear more rarely with each generation. These results meet the expectations set based on the literature review. Consistently with other papers, we find that not only the melody but also classical piece composition to user preference can improve over iterations of an interactive algorithm.

The plot of the rating means over generations in Figure 3 additionally confirms the overarching problem of music subjectivity. Although the rating mean of 6 out of 7 users consistently remains between 5 and 9 over the experiment, the seventh user rated the pieces on average in the $[2, 3]$ range, indicating that the type of music the transformer is capable of composing was not to their taste. Even though the ratings improved over generations, subjectivity remained highly relevant throughout the experiment as a reason for the lower satisfaction of this user.

Equally as important are the results of the rating model, which, after just 6 generations of the algorithm (where each generation consists of rating 6 pieces), was able to predict ratings on average 0.81 away from the rating given by the participant on a 10-point scale. This proves that a neural network can model music preference, but also that an interactive reinforcement learning algorithm can successfully extract preference with no prior knowledge. These findings coincide with the conclusions of the literature review which show promising capabilities of deep learning in modelling user preference, and most importantly, the potential of reinforcement learning in automatic music composition.

5.4 Limitations and Future Work

The main limitations that can be identified in this work are the level of the prototype and the extensiveness of the evaluation, both of which stem from the lack of time and advanced hardware available for this research. Before utilizing a similar model at a larger scale, it would be strongly advised to train the music-generating transformer on a larger dataset and tune the training hyperparameters. The system should also be evaluated more significantly, preferably by more users and through a longer experiment generation-wise. Highly specialistic hardware would make this more achievable by making the music-generation step of the experiment much faster.

5.5 Implications

In this paper, we were able to answer the posed research question by designing a successful automatic music composition system that uses reinforcement learning to compose for user preference. Thanks to the system consisting of two separate agents it is possible to apply the rating agent to other music-generating systems, enabling the generation of music for user preference across many other genres. Since music is used extensively in the modern world, there is high demand for new compositions. This means that a system such as the one proposed in this paper, capable of quickly composing highly tailored pieces has many potential applications in entertainment and marketing. In addition, the system requires no prior knowledge, reducing the barrier to entry for non-expert users which in turn makes it accessible to a wider audience. It would be reasonable to expect more similar solutions to gain popularity in the near future.

ACKNOWLEDGMENTS

Special thanks to my supervisor Ton Spil, who was supportive in my selection of a topic of interest, and guided me in improving my paper based on his experience and valuable criticism.

REFERENCES

[1] Abien Fred Agarap. 2018. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*.

[2] Andrea Agostinelli et al. 2023. Musiclm: generating music from text. (2023). arXiv: 2301.11325 [cs.SD].

[3] Saikat Basu, Nabakumar Jana, Arnab Bag, Mahadevappa M, Jayanta Mukherjee, Somesh Kumar, and Rajlakshmi Guha. 2015. Emotion recognition based on physiological signals using valence-arousal model. In *2015 Third International Conference on Image Information Processing (ICIIP)*, 50–55. doi: 10.1109/ICIIP.2015.7414739.

[4] Chung-Yi Chi, Richard Tzong-Han Tsai, Jeng-You Lai, and Jane Yung-Jen Hsu. 2010. A reinforcement learning approach to emotion-based automatic playlist generation. In *INTERNATIONAL CONFERENCE ON TECHNOLOGIES AND APPLICATIONS OF ARTIFICIAL INTELLIGENCE (TAAI 2010)* (Conference on Technologies and Applications of Artificial Intelligence). 15th Annual International Conference on Technologies and Applications of Artificial Intelligence (TAAI), Hsinchu, TAIWAN, NOV 18-20, 2010. IEEE Comp Soc; Taiwanese Assoc AI, 60–65. isbn: 978-0-7695-4253-9. doi: 10.1109/TAAI.2010.21.

[5] S. Dai, X. Ma, Y. Wang, and R.B. Dannenberg. 2023. Personalised popular music generation using imitation and structure. *Journal of New Music Research*. cited By 0. doi: 10.1080/09298215.2023.2166848.

[6] [n. d.] Dall·e 2. (). https://openai.com/product/dall-e-2.

[7] Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, and Ilya Sutskever. 2020. Jukebox: a generative model for music. (2020). arXiv: 2005.00341 [eess.AS].

[8] Sander Dieleman, Aaron van den Oord, and Karen Simonyan. 2018. The challenge of realistic music generation: modelling raw audio at scale. *Advances in Neural Information Processing Systems*, 31.

[9] Chris Donahue, Huanru Henry Mao, Yiting Ethan Li, Garrison W Cottrell, and Julian McAuley. 2019. Lakhnes: improving multi-instrumental music generation with cross-domain pre-training. *arXiv preprint arXiv:1907.04868*.

[10] Chuang Gan, Deng Huang, Peihao Chen, Joshua B Tenenbaum, and Antonio Torralba. 2020. Foley music: learning to generate music from videos. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XI 16*. Springer, 758–775.

[11] [SW] Damon Gwinn, MusicTransformer-Pytorch version 1.0, 2020. url: https://github.com/gwinndr/MusicTransformer-Pytorch.

[12] Curtis Hawthorne, Andriy Stasyuk, Adam Roberts, Ian Simon, Cheng-Zhi Anna Huang, Sander Dieleman, Erich Elsen, Jesse Engel, and Douglas Eck. 2019. Enabling factorized piano music modeling and generation with the MAESTRO dataset. In *International Conference on Learning Representations*. https://openreview.net/forum?id=r1lYRjC9F7.

[13] Dorien Herremans, Ching-Hua Chuan, and Elaine Chew. 2017. A functional taxonomy of music generation systems. *ACM Computing Surveys (CSUR)*, 50, 5, 1–30.

[14] Cheng-Zhi Anna Huang et al. 2018. Music transformer. *arXiv preprint arXiv:1809.04281*.

[15] Max Jaderberg, Karen Simonyan, Andrew Zisserman, et al. 2015. Spatial transformer networks. *Advances in neural information processing systems*, 28.

[16] Nan Jiang, Sheng Jin, Zhiyao Duan, and Changshui Zhang. 2020. Rl-duet: online music accompaniment generation using deep reinforcement learning. In *THIRTY-FOURTH AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE, THE THIRTY-SECOND INNOVATIVE APPLICATIONS OF ARTIFICIAL INTELLIGENCE CONFERENCE AND THE TENTH AAAI SYMPOSIUM ON EDUCATIONAL ADVANCES IN ARTIFICIAL INTELLIGENCE* (AAAI Conference on Artificial Intelligence). Vol. 34. 34th AAAI Conference on Artificial Intelligence / 32nd Innovative Applications of Artificial Intelligence Conference / 10th AAAI Symposium on Educational Advances in Artificial Intelligence, New York, NY, FEB 07-12, 2020. Assoc Advancement Artificial Intelligence, 710–718. isbn: 978-1-57735-835-0.

[17] X. Ma, Y. Wang, and Y. Wang. 2022. Content based user preference modeling in music generation. In cited By 0, 2473–2482. doi: 10.1145/3503161.3548169.

[18] Maryam Majidi and Rahil Mahdian Toroghi. 2023. A combination of multi-objective genetic algorithm and deep learning for music harmony generation. *Multimedia Tools and Applications*, 82, 2, 2419–2435. Cited by: 1; All Open Access, Green Open Access. doi: 10.1007/s11042-022-13329-6.

[19] Dragan Matić. 2010. A genetic algorithm for composing music. *Yugoslav Journal of Operations Research*, 20, 1, 157–177. Cited by: 41; All Open Access, Gold Open Access, Green Open Access. doi: 10.2298/YJOR1001157M.

[20] F. Mo, X. Ji, H. Qian, and Y. Xu. 2022. A user-customized automatic music composition system. In cited By 0, 640–645. doi: 10.1109/ICRA46639.2022.9812396.

[21] Maria Navarro-Caceres, Marcelo Caetano, Gilberto Bernardes, and Leandro Nunes de Castro. 2019. Chordais: an assistive system for the generation of chord progressions with an artificial immune system. *SWARM AND EVOLUTIONARY COMPUTATION*, 50, (Nov. 2019). doi: 10.1016/j.swevo.2019.05.012.

[22] OpenAI. 2023. Gpt-4 technical report. (2023). arXiv: 2303.08774 [cs.CL].

[23] Matthew J Page et al. 2021. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372. eprint: https://www.bmj.com/content/372/bmj.n71.full.pdf. doi: 10.1136/bmj.n71.

[24] Serkan Sulun, Matthew EP Davies, and Paula Viana. 2022. Symbolic music generation conditioned on continuous-valued emotions. *IEEE Access*, 10, 44617–44626.

[25] Juan Sun. 2022. Variational fuzzy neural network algorithm for music intelligence marketing strategy optimization. *COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE*, 2022, (Jan. 2022). doi: 10.1155/2022/9051058.

[26] Maho Taniguchi, Kense Todo, Shoya Yasuda, and Masayuki Yamamura. 2021. A system for generating audio influenced by audience evaluation using interactive genetic algorithm. In *ADVANCES IN ARTIFICIAL INTELLIGENCE* (Advances in Intelligent Systems and Computing). K Yada, D Katagami, Y Takama, T Ito, A Abe, E SatoShimokawara, J Mori, N Matsumura, and H Kashima, (Eds.) Vol. 1357. 34th Annual Conference of the Japanese-Society-for-Artificial-Intelligence (JSAI), ELECTR NETWORK, JUN 09-12, 2020. Japanese Soc Artificial Intelligence, 208–215. isbn: 978-3-030-73113-7; 978-3-030-73112-0. doi: 10.1007/978-3-030-73113-7_20.

[27] Shu Wang, Chonghuan Xu, Austin Shijun Ding, and Zhongyun Tang. 2021. A novel emotion-aware hybrid music recommendation method using deep neural network. *ELECTRONICS*, 10, 15, (Aug. 2021). doi: 10.3390/electronics10151769.

[28] Xinxi Wang, Yi Wang, David Hsu, and Ye Wang. 2014. Exploration in interactive personalized music recommendation: a reinforcement learning approach. *ACM TRANSACTIONS ON MULTIMEDIA COMPUTING COMMUNICATIONS AND APPLICATIONS*, 11, 1, (Aug. 2014). doi: 10.1145/2623372.

[29] Abigail Wiafe, Charles Nutrokpor, Ebenezer Owusu, Ferdinand Apietu Kastriku, and Isaac Wiafe. 2022. Using genetic algorithms for music composition: implications of early termination on aesthetic quality. *International Journal of Information Technology (Singapore)*, 14, 4, 1875–1881. Cited by: 0. doi: 10.1007/s41870-022-00897-x.

[30] Xianchao Wu, Chengyuan Wang, and Qinying Lei. 2020. Transformer-xl based music generation with multiple sequences of time-valued notes. *arXiv preprint arXiv:2007.07244*.

[31] Genki Yamaguchi and Makoto Fukumoto. 2019. A music recommendation system based on melody creation by interactive ga. In *2019 20TH IEEE/ACIS INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERING, ARTIFICIAL INTELLIGENCE, NETWORKING AND PARALLEL/DISTRIBUTED COMPUTING (SNPD)*. M Nakamura, H Hirata, T Ito, T Otsuka, and S Okuhara, (Eds.) 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019), Toyama, JAMAICA, JUL 08-11, 2019. Inst Elect & Elect Engineers; Int Assoc Comp & Informat Sci; IEEE Comp Soc, 286–290. isbn: 978-1-7281-1651-8.

[32] Hua Zhu, Shangfei Wang, and Zhen Wang. 2008. Emotional music generation using interactive genetic algorithm. In vol. 1. Cited by: 14, 345–348. doi: 10.1109/CSSE.2008.1203.

A EXAMPLES OF GENERATED AND RATED MUSIC PIECES

An example of generated music pieces rated by the participants of the experiment may be found here: https://drive.google.com/drive/folders/1fV4R560uMPC-hF0N4xUa41x-D-Z2Lf0s?usp=sharing

B LITERATURE REVIEW OVERVIEW

Table 2: Overview of articles found through literature search

Author(s) & year	Title	Methods	Scope	Limitations
Dai et al. (2023) [5]	Personalized Popular Music Generation Using Imitation and Structure	Objective and subjective evaluation	Stylistic Imitations of seed songs preferred by users	Limited to imitating existing songs, no deep learning involved
Ma et al. (2022) [17]	Content-based User Preference Modeling in Music Generation	Objective and subjective evaluation	Modeling user music preference and integrating it in automatic music generation	Limited to users' listening histories and ratings and user preference can be applied easily because a rule- and statistics-based model is used rather than a deep learning model
Mo et al. (2022) [20]	A User-customized Automatic Music Composition System	Experiment and evaluation	Music composition segment by segment to user preference	Limited to generating short segments, composing a piece takes a long time
Jiang et al. (2020) [16]	RL-Duet: Online Music Accompaniment Generation Using Deep Reinforcement Learning	Experiment and evaluation	Reinforcement learning-based accompaniment model for real-time interaction	Limited to accompaniment of existing man-made music, not capable of real-time improvisation
Wang et al. (2021) [27]	A Novel Emotion-Aware Hybrid Music Recommendation Method Using Deep Neural Network	Experiment on model	Modelling music emotion representation and emotion preference prediction	Requires information about listening habits of the user and current events, no direct interaction with user
Wang et al. (2014) [28]	Exploration in Interactive Personalized Music Recommendation: A Reinforcement Learning Approach	Experiments, user study	Interactive music recommendation, balancing exploitation and exploration	Limited to Bayesian model, recommendation requires user rating of existing songs
Sun (2022) [25]	Variational Fuzzy Neural Network Algorithm for Music Intelligence Marketing Strategy Optimization	Performance analysis of developed model	Preference modelling, music feature extraction, music recommendation for preference	Limited to the recommendation and based on listening habits, trouble with outliers
Navarro-Caceres et al. (2019) [21]	ChordAIS: An assistive system for the generation of chord progressions with an artificial immune system	Performance analysis	Interactive iterative chord progression suggestion generation	Scope is only chord progressions, requires chord sequence as input
Yamaguchi and Fukumoto (2019) [31]	A Music Recommendation System based on Melody Creation by Interactive GA	Experiment	Using a melody-creating genetic algorithm to recommend music for user preference	Limited to generating melodies, unsuccessful in recommendation based on preference
Chi et al. (2010) [4]	A Reinforcement Learning Approach to Emotion-based Automatic Playlist Generation	User study on model	Playlist generation based on emotion with human evaluation	Limited to playlist generation of existing songs, limited to user preference of song emotions
Taniguchi et al. (2021) [26]	A System for Generating Audio Influenced by Audience Evaluation Using Interactive Genetic Algorithm	Quantitive and Human Evaluation	Generating single short sounds using an interactive genetic algorithm	Generating sound effects based on short audio rather than composition, certain types of sounds less effective, not individual
Matić (2010) [19]	A genetic algorithm for composing music	System prototyping	Non-interactive genetic algorithm for music composition	Limited to researched heuristic for genetic algorithm fitness, no human interaction
Wiafe et al. (2022) [29]	Using genetic algorithms for music composition: implications of early termination on aesthetic quality	Objective and subjective evaluation	Investigating the impact of genetic algorithm generation number on the quality of generation	Limited to generating monophonic melodies, lack of human interaction and long training time
Zhu et al. (2008) [32]	Emotional music generation using interactive genetic algorithm	System prototyping and subjective test	Interactive genetic algorithm for music composition based on non-expert user feedback on the emotionality of music	Limited to only two emotions of a song, ineffective convergence algorithm
Majidi and Toroghi (2023) [18]	A combination of multi-objective genetic algorithm and deep learning for music harmony generation	System prototyping and subjective evaluation	Interactive genetic algorithm and deep learning with a non-linear objective in the context of music	Requires expert users, training optimization of generation model rather than generating for specific user