International Core Journal of Engineering Volume 9 Issue 10, 2023
ISSN: 2414-1895 DOI: 10.6919/ICJE.202310_9(10).0021

# Study of Generation based on Recurrent Neural Networks

Lifan Sun
The Hill School Pottstown, PA, USA
*Lsun24@thehill.org

## Abstract

**This article mainly uses the Magenta framework to explore the two models of Melody RNN and Performance RNN, They are all recurrent neural networks based on LSTM (Long short-term memory).. The experiment uses a data set containing 120 songs. I set the Melody RNN model is trained for approximately 20,000 steps, while the Performance RNN model is trained for approximately 40,000 steps. The results show that the accuracy of Melody RNN is 80%, while that of Performance RNN is 90%.**

## Keywords

**Deep Learning; Generation; Melody RNN; Performance RNN.**

## 1. Introduction

Since Ancient Greece, the therapeutic uses of music have been recognized. A legend reported by Boethius (480-524 CE) stated that Pythagoras, one of the first notable people to recognize the importance of music in healing and treatment, upon hearing of a youth who had been jilted by his lover and was preparing to set her house on fire, determined that the youth was under the influence of a certain musical mode (scale). By suggesting that he changed his tune and employ a melody based on an alternate scale, Pythagoras was able to restore the youth to a state of calmness. [1].

Other philosophers of the time such as Socrates and Plato also had similar beliefs in that music could affect ones personality and emotions. Plato extended his belief to the necessities of music censorship as it had a direct correlation to a person's ethos. [2] Many Greek philosophers opined that listening to music in certain states could affect the listener in either a positive or negative way. On the other hand, listening to music that does not benefit the body or soul, or that was not conventional would lead to the loss of societal standards and the corruption of youth.

In the Renaissance, musical properties such as harmony were viewed to be therapeutic and later in the 17th and 18th centuries, the curative properties of music were elucidated through vibrations. During the 18th century, in conjunction with the rise in notability in psychiatry and the classification of diseases, stimulation of senses become prominent, which music was believed to achieve effectively. [3] But only at the turn of the 19th century, were scientific studies conducted on the therapeutic purposes of music, resulting in undeniable proof of the benefits of music.

This belief that music has therapeutic uses is still employed today in similar fields the Ancient Greeks believed it could be purposed. In 2013, the American Psychology Association conducted a study in which 42 children ages 3-11 were inserted with an IV. Researchers discovered that patients listening to music reported less pain and demonstrated less distress. [4] Other studies have also found positive and even curative purposes in music from simple music therapy to encourage positive physiological behaviors [5] to vibration therapy for muscle weakness or Parkinson's disease. [6].

One such issue in which therapeutic music can aid in is sleep. Sleep in a biological process which provides rest and restoration for the mind and body. Without enough quality sleep, physical and mental health, thinking, and daily functioning can be negatively affected. [7] Sleep disorders refer to

ISSN: 2414-1895 DOI: 10.6919/ICJE.202310_9(10).0021

the conditions that impact a person's sleep quality, which in turn, affects their function while awake. One of the most common forms of sleep disorders is insomnia, defined by an inadequate quality of sleep that interferes with daytime functioning. Approximately 30% of American adults have short-term insomnia, and 10% of American adults have chronic insomnia. [8] Causes of insomnia include, but are not limited to, stress, travel or work schedule, and poor sleep habits.

Most attempts to remedy poor sleep such as the use of pharmaceutical sleep aids have negative side effects which must be considered. Long term uses of pharmaceutical sleep aids lead to many negative side effects including nausea, dizziness, amnesia, and even increased morality. [9] Due to dependency, cost, and harmful side effects of sleep aids, an alternate solution should be prioritized.

With growing affirmation on musical therapy, it should be considered as a possible solution to poor sleep. As mentioned above, listening to music has the effect of reducing anxiety, as well as regulate ones mood. Current findings indicate that music played at around 60bpm can cause the brain to synchronize with the beat causing alpha brainwaves, which is required to be present in a relaxed state. Devoting 45 minutes in a relaxed position and listening to calming music can induce sleep. [10] The accessibility of music is what allows it to be a feasible alternative to medications.

A study conducted in 2008 showed that music statistically improved sleep quality. [11] Using a group of 94 students with sleep complaints, participants were asked to listen to classical music for 45 minutes at bedtime for 3 weeks, listen to an audiobook or do neither. Participants who listened to music showed a significant improvement in sleep quality compared to those who listened to an audiobook or the control group. Music also positively affected quality of sleep for people with sleep disorders. In a study involving 557 patients, it was shown that music can assist in improving sleep quality of patients with acute and chronic sleep disorders. [12].

The ability to compose music suitable for sleep aid is difficult, but through deep learning, the process can be simplified drastically. Deep learning is a part of a broader family of Machine learning, in which a machine attempts to simulate the behavior of a human brain, allowing it to "learn" from data. It is incredibly useful for certain scenarios in our day to day lives and will only continue to improve as time progresses. Artificial Intelligence has been commonly used for pattern recognition and prediction, but in recent years, it has also become useful for mimicking human creativity, bringing up an interesting philosophical question. Is it possible for a machine to produce original music?

In this case of creating melodies, instead of continually receiving input from a programmer, deep learning allows the machine to listen to many examples of songs, learn the patterns, and use the patterns it learned to generate a "new" song, similar to what musicians do. Magenta, the framework used in this project, is an open-source research project exploring the role of machine learning as a tool in the creative process. Through creating a dataset of MIDI files, Magenta is capable of analyzing and creating its own composition.

This project is by no means the first of its kind, as artificial intelligence to write music has been around for years, and software similar to Magenta such as Flow Machines or AIVA have been used to create extraordinary musical works. On April 2nd, 2021, RollingStone magazine published an article on how an AI software wrote a "New" Nirvana song. [13] Utilizing Magenta, the creators were able to select riffs and sections of songs to generate new melodies which could be fused into a song. This is exactly the purpose of Deep learning in music: to create a composition based on a certain theme. A similar experiment, DeepJ, is capable of composing music in a specific style through a mixture of composer styles, which yields practical benefits for aiding composers in their compositions. [14] As is shown in Fig.1.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202310_9(10).0021


**Fig. 1** Google's Magenta, an open-source research project exploring the role of machine learning.

Deep learning has become more prevalent in the past few years and the number of studies directed towards music have increased as well. Most studies in recent years focus on the technicalities of deep music generation, specifically, which methods of generation have the best results. One such example is Music SketchNet, inspired by the original automatic image completion system SketchNet. Music SketchNet is a neural framework which allows users to specify partial musical ideas guiding automatic music generation, allowing it to complete incomplete monophonic pieces. [15] Many other projects have done similar studies on the effects of LSTM network applications to music and composition. Through training the networks using tens of thousands of music transcriptions, it is possible to generate new transcriptions. [16].

Other frameworks allow multi-track compositions. MuseGan is a Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment. Under the framework of generative adversarial networks, MuseGan is capable of producing a rock composition of five tracks. [17].

As an avid musician, I believe that music is a prominent part of everyday life. Sleep, although a completely different area in our lives, is arguably even more important to the well-being of humans. The ultimate goal is to utilize music to help people sleep better. The general idea of this project is simple: Utilize deep learning to simplify and automate the process of generating new and unique music which can be used to help people sleep. Utilizing Magenta, I was able to achieve the goal of generating unique music through using its models and training them using selective pieces.

## 2. METHODS

The framework I used to create music is Magenta, a program started by Google researchers and engineers that employ deep learning to generate compositions. On Magenta's official website, it states that the magenta research project is "an exploration in building smart tools and interfaces that allow artists and musicians to extend (not replace!) their processes using these models." [18] I took the liberty to take it a step forwards and replace the need of a composer, allowing the software to interpret data on its own and create compositions without interference.

The two models I used within Magenta to create the composition were Melody Rnn and Performance Rnn, each with their individual specialties and produced different results.

Melody Rnn has four configurations: Basic, Mono, Lookback, and Attention. In Melody Rnn, I chose to utilize the basic configuration as the other three resulted in bugs. The basic configuration returns a monophonic melody with no key or mode, resulting in a rudimentary melody. This configuration acts as a baseline for melody generation with an LSTM model. For its training, all transposed examples and outputs are in the MIDI range 48 to 84. [19].

On the other hand, Performance Rnn returns a more complete piece involving different chords, complimented with a variety of note lengths and overall rhythm. Yet, it still lacked a key and a pleasing melody. The most significant application of Performance Rnn is its ability to generate music with human-like performance and phrasing. At a fundamental level, MIDI consists of precisely-timed events, each of which specifies the pitch of the note. These events are then imported into a standard synthesizer to create the 'sound' of the piano. Performance Rnn uses a set of events including time-

ISSN: 2414-1895 DOI: 10.6919/ICJE.202310_9(10).0021

shifts and velocity changes to encapsulate a more dynamic expression and feel, similar to how a musician would perform it. [20].

## 3. Experiment

The first stage of generating a new composition using Magenta was creating a dataset. The dataset I created was composed of approximately 120 songs which were mostly classical pieces with a few other piano compositions. I specifically focused on using slow piano compositions because the output is a piano composition as well. I ensured that all compositions were approximately the same genre and style, which conveyed the feeling of calmness and relaxation, and had similar characteristics of being slow and soft, similar to songs one would listen to when sleeping.

One detail I realized was that the similarity between the compositions in the dataset and the generated music was not as apparent as I imagined it to be, most likely due to the relatively small size of the dataset.

The training process for both the Melody Rnn model and the Performance Rnn model were both a success. The Melody Rnn model trained for approximately 20,000 steps while the Performance Rnn model trained for approximately double the steps at around 40,000. Both ended with a relatively high accuracy with the Melody Rnn having about an 80% accuracy and the Performance Rnn having about a 90%.

In the completed score, the difference between the two models is displayed. Melody Rnn outputs a much simpler melody with few notes and expressive quantities. All pieces begin with middle C, are in the key C major and do not have accidentals or key changes.

Performance Rnn is much more different. In Fig.2, the first measure of the score generated by Performance Rnn shows its expressiveness. The notes are difficult to read on paper due to the odd rhythm, but that is the purpose of the model. It generates pieces based on how they are meant to be performed to be more expressive. Everything including the notes, rhythm, chords, and tempo, all contribute to the complexity of the piece.

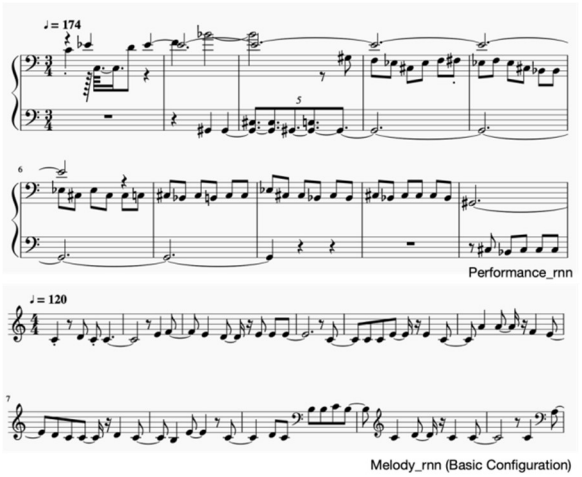


**Fig. 2** Final Score of Performance Rnn and Melody Rnn

## 4. Conclusion

The overall object of this project was achieved through utilizing Magenta to generate music based off a data set, Melody RNN is the simplest and most basic model in Magenta's automatic composition model, Performance RNN can generate composite music based on temporal and dynamic factors, They are both classic sequence generative models.

But the music generated in the experimrnt did not resemble sleep music. Some parts were clashing, loud, and did not give a soothing feel. One solution to this problem would be to use a more appropriate

ISSN: 2414-1895 DOI: 10.6919/ICJE.202310_9(10).0021

and larger data set. Although most pieces in the data set were approximately the same genre and style, parts of them changed styles and augmented the training process. The data set used in this project had under 150 songs, while other studies have up to thousands. By expanding the data set, I would most likely have obtained a better result. Secondly, apart from Performance Rnn and the Basic configuration in Melody Rnn, there are many models which could be used, each with their own benefits. I believed Performance Rnn would have the best result, but it would be worth the try with the other models.

The ability to generate music using Melody Rnn or Performance Rnn faster than a composer could write music is the main benefit. Finely-tuned, the application of this product could achieve its goal of helping people sleep. The unique music produced by it would be the main factor why people would listen to machine generated music rather than music they are familiar with.

## References

[1] Richards, Melanie. “Pythagoras and Music.” Pythagoras and Music, 2009, https://0b90d72963ef1d10103a-57da1f3fe93ccf3a9828e6ce67c3d52c.ssl.cf5.rackcdn.com/07_richards.pdf.

[2] Schoen, Mary B. “Plato and Aristotle on the Ends of Music - Érudit.” Plato and Aristotle on the Ends of Music, 1978, https://www.erudit.org/fr/revues/ltp/1978-v34-n3-ltp3386/705684ar.pdf.

[3] Dobrzyńska, Ewelina, et al. “Music Therapy - History, Definitions and Application - Researchgate.” Music Therapy - History, Definitions, and Application, Mar. 2006, https://www.researchgate.net/publication/282647747_Music_Therapy_-_History_definitions_and_application.

[4] Novotney, Amy. “Music as Medicine.” Monitor on Psychology, American Psychological Association, Nov. 2013, https://www.apa.org/monitor/2013/11/music.

[5] “Music Therapy: What Is It and How Does It Work?: Musicians Institute.” Musicians Institute Hollywood, 21 May 2021, https://www.mi.edu/in-the-know/music-therapy-work/.

[6] CG;, Kapur SS;Stebbins GT;Goetz. “Vibration Therapy for Parkinson's Disease: Charcot's Studies Revisited.” Journal of Parkinson's Disease, U.S. National Library of Medicine, 2012, https://pubmed.ncbi.nlm.nih.gov/23939405/.

[7] Medline Plus. “Sleep Disorders.” Sleep Disorders, U.S. National Library of Medicine, 3 Jan. 2020, https://medlineplus.gov/sleepdisorders.html.

[8] Harvard Heath. Insomnia: Restoring Restful Sleep, 1 Feb. 2009, https://www.health.harvard.edu/staying-healthy/insomnia-restoring-restful-sleep.

[9] Trahan, Tabitha, et al. “The Music That Helps People Sleep and the Reasons They Believe It Works: A Mixed Methods Analysis of Online Survey Reports.” PloS One, Public Library of Science, 14 Nov. 2018, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6235300/#pone.0206531.ref015.

[10] Counseling. “Releasing Stress through the Power of Music: Counseling Services.” University of Nevada, Reno, https://www.unr.edu/counseling/virtual-relaxation-room/releasing-stress-through-the-power-of-music.

[11] Harmat, László, et al. “Harmat - 2008 - Wiley Online Library.” Music Improves Sleep Quality in Students, 18 Apr. 2008, https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2648.2008.04602.x.

[12] Wang, Chun-Fang, et al. “Music Therapy Improves Sleep Quality in Acute and Chronic Sleep Disorders: A Meta-Analysis of 10 Randomized Studies.” International Journal of Nursing Studies, Pergamon, 9 Apr. 2013, https://www.sciencedirect.com/science/article/abs/pii/S0020748913000965.

[13] Grow, Kory. “In Computero: Hear How AI Software Wrote a 'New' Nirvana Song.” Rolling Stone, Rolling Stone, 5 Apr. 2021, https://www.rollingstone.com/music/music-features/nirvana-kurt-cobain-ai-song-1146444/.

[14] Mao, Huanru Henry, et al. “DeepJ: Style-Specific Music Generation.” ArXiv.org, 3 Jan. 2018, https://arxiv.org/abs/1801.00887.

[15] Chen, Ke, et al. “Music Sketchnet: Controllable Music Generation via Factorized Representations of Pitch and Rhythm.” ArXiv.org, 4 Aug. 2020, https://arxiv.org/abs/2008.01291.

[16] Sturm, Bob L., et al. “Music Transcription Modelling and Composition Using Deep Learning.”ArXiv.org, 29 Apr. 2016, https://arxiv.org/abs/1604.08723.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202310_9(10).0021

[17] Dong, Hao-Wen, et al. “Musegan: Multi-Track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment.” Proceedings of the AAAI Conference on Artificial Intelligence, https://ojs.aaai.org/index.php/AAAI/article/view/11312.

[18] Magenta. “Magenta/Magenta: Magenta: Music and Art Generation with Machine Intelligence.” GitHub, https://github.com/magenta/magenta.

[19] Magenta. “Magenta/Magenta/Models/melody_rnn at Main Magenta/Magenta.” GitHub, https://github.com/magenta/magenta/tree/main/magenta/models/melody_rnn.

[20] Simon, Ian. “Performance RNN: Generating Music with Expressive Timing and Dynamics.” Magenta, 29 June 2017, https://magenta.tensorflow.org/performance-rnn.