IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-7, Issue-3, March 2019 E-ISSN: 2347-2693

Automatic Music Generation

Lawakesh Patel[1*], Nidhi Singh[2], Rizwan Khan[3]

[1,2,3]Computer Science and Engineering, ABES Institute of Technology, Ghaziabad, India

Corresponding Author: lawakesh23@gmail.com, Contact No..: 8882290934

DOI: https://doi.org/10.26438/ijcse/v7i3.8082 | Available online at: www.ijcseonline.org

Accepted: 08/Mar/2019, Published: 31/Mar/2019

***Abstract*—** In this paper authors describes the automatic music generation system and automatic music evaluation system. The system composes short pieces of music by choosing some factors in music, such as timbre, pitch interval, rhythm, tempo etc. The most important features of the system the music is generated according to the mood and sentiments of person. In the implemented work mode control the pitch interval and density control the rhythm of music. Neural Network Algorithm for automatic evaluation system of music. Music composition is an art, even the task of playing composed music takes considerable effort by humans. Given this level of complexity and abstractness, designing an algorithm to perform both the tasks at once is not obvious and would be a fruitless effort. In this paper authors describe new music composition by using trained music data set to extract useful music pattern and generate the music in the form of chord.In this paper also discussed about method or platform use for automatic music generation.

***Keywords*—** Music Generator,Generative Model,The Restricted Boltzman Machine,MIDI file,Tensorflow

I. INTRODUCTION

The music industry has shifted towards more and more digital distribution through streaming services and online music stores like, Groove shark, Spotifyand, iTunes and Google Play. As a result, automatic music recommendation has become an increasingly relevant problem.Deep Learning has become an essential tool in the field of automated music generation. A various deep learning architectures have been studied for performing the task pop music generation (Fidler,Chu,and Urtasun) and for composing novel melodies music that are similar to classical compositions(Hadjeres, and Nielse,Pachet) Tools now exist for helping artists compose music in order to augment the creative process, and there exist number of commercial music generation systems.The generation tools can be broadly classified into two types: symbolic and raw audio models. Symbolic methods train and operate at the note level—based on melodies and notes from existing training Machine Learning has tool that can generate music by computer.

Machine Learning has enabled to build complex applications with great accuracy. Whether it deal with audio, video and image, using Machine Learning can solve a wide range problem. Tensorflow Framework helps to achieve all of these applications. Google created the Tensorflow framework for creating Deep Learning models. Deep Learning is a class of machine learning models that use multi-layer neural networks.

Suppose you wanted to compose a specific song, a “funny” and “sad” song. The first step would be to get a data set of songs that labeled with emotions. The program would convert the set of speech o to text. For each word, vectors would create by program and train the model. After completing the training, the program would find the associated vectors of given input (“funny” and “sad”) and compare them to the vectors that has been created before. The output would me generated in the form of chord.

Artificial intelligence have also worked on audio samples, but instead of trying to learn and generate from raw audio samples, they work on the frequency domain of the audio. This approach is faster because it allows the network to train and predict a group of samples that make up the frequency domain rather than one sample. Since the frequency domain can still represent all audible audio signals, there are no restrictions on the type of music that it can generate. Authors propose an automated music generation system that enables users to create their own music for their own purpose by controlling the music with high level control parameters. The objective of this project is to learn a structure from a lot of different song examples. This will create some melodic songs. The MIDI format data that will be used for learning. This format is an easy way to have access to a lot of free music data set. The other advantage of the MIDI format is that it is easy to retrieve the needed information.

II. OUR OWN MUSIC GENERATOR

Magenta is a research topic that exploring the role of machine learning in the process of creating art and music. Primarily this involves developing new deep learning and reinforcement learning algorithms for generating, images, drawings, songs, and other materials. Although Magenta is very powerful, we will build our own simple music generator using a Restricted Boltzmann Machine (RBM) to generate short sequences of pop music. Our training data that is in the form of midi format will be around a hundred MIDI files of pop songs (*MIDI* is a format that directly encodes musical notes). Using this format, we won't label them with emotions as described in the introduction, the output will be another new pop melody, just like our training files.

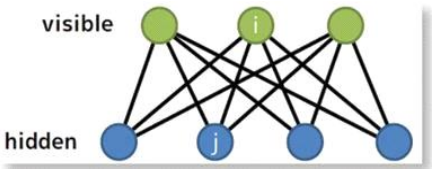

Fig 1 RBM (Restricted Boltzmann Machine)

RBM is a neural network with two layers, that is visible layer and the hidden layer. In which each visible node is connected to each hidden node (and vice versa), but there are no visible-visible or hidden-hidden connections (nodes are simply where calculations take place). This is the *restriction* each node of visible layer takes one chord. Each chord is multiplied by a *weight* and then the output of the nodes at the hidden layer. Most of the neural networks, RBMs are generative models that directly model the probability distribution of data. To sample from an RBM, use an algorithm known as Gibbs Sampling.

We'll use them as follows:

1. Use Panda as our data analytics library
2. For the music files Midi as our helper library
3. TQDM library for printing a progress bar during the process.

III. GENERATIVE MODEL

Generative Models enumerate a probability distribution over a dataset of input vectors. For an unsupervised task, form a model for P(x)P(x), where xx is an input vector. For a supervised task, form a model for P(x|y)P(x|y), where yy is the label for xx. Like discriminative models, most of generative models can be used for classification task. To perform classification with a generative model, we support the fact that if we know P(X|Y)P(X|Y) and P(Y)P(Y), we can use Bayes rule to estimate P(Y|X)P(Y|X).

IV. THE RESTRICTED BOLTZMAN MACHINE

Boltzmann Machines (BMs) are a particular form of log-linear Markov Random Field (MRF), for which the energy function is linear in its free parameters. To make them powerful enough to represent complicated distributions, we consider that some of the variables are never observed. By having more hidden variables, we can increase the modeling capacity of the Boltzmann Machine (BM). Restricted Boltzmann Machines (RBM) further restrict BMs to those without visible-visible and hidden-hidden connections.

The energy function *E (v, h)* define as:

$$E(v, h) = -b'v - c'h - h'Wv$$

Where W represents the weights connecting hidden and visible units and, b, c are the offsets of the visible and hidden layers respectively.

This translates directly to the free energy formula:

$$\mathcal{F}(v) = -b'v - \sum_i \log \sum_{h_i} e^{h_i (c_i + W_i v)}.$$

Because of the RBM's specific structure of, visible and hidden units are conditionally independent given one-another. Using this property, we can write:

$$p(h|v) = \prod_i p(h_i|v)$$

$$p(v|h) = \prod_j p(v_j|h).$$

A. Model Architecture

The RNN-RBM architecture is not tremendously complicated. Each hidden unit of RNN is paired with an RBM. RNN hidden unit *ut* takes input from data vector *vt* as well as from RNN hidden unit *ut*−1.Outputs of hidden unit ut are the parameters of *RBMt*+1, which takes as input data vector *vt*+1. Where t is a time.

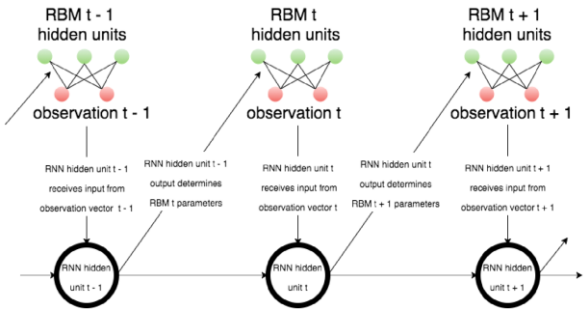

Figure 2. Architecture of RNN-RBM

- *Wuv*, the weight matrix that connects the RNN hidden units to the RBM visible biases.
- *Wvu*, the weight matrix that connects the data vectors to the RNN hidden units.
- *Wuu*, the weight matrix that connects the RNN hidden units to each other.

- *bu*, the bias vector for the RNN hidden unit connections.
- *bv*, the bias vector for the RNN hidden unit to RBM visible bias connections.

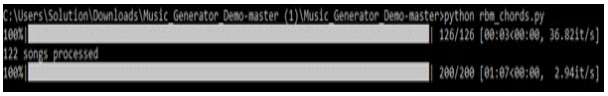

Figure3. Training process

V. CANCLUSION AND FUTURE SCOPE

Automatic music generation have always been an interesting topic for researchers across globe. All the proposed idea in above paper have a crucial advantages of their own as well as limitations as well. The Restricted Boltzmann Machines(RBM) and gibbs algorithm is used in this project has a python and tensorflow implementation To built the auto music lyrics generation, in which machine learning libraries panda, midi and tqdm are used, panda used for data analytics and tqdm is used for print the progress bar during process.

In future I will keep on experimenting with new datasets and newly created Machine Learning models that can generate Indian music lyrics by learning from the inputs fed into it. Focus will be also on integrating a speech synthesiser that will give the algorithm a voice.

REFERENCES

[1] Akshay Jakhotiya, Ketan Kulkarni, Chinmay Inamdar, Bhushan Mahajan, Alka Londhe "Automatic Subtitle Generation for English Language Videos",2018,PP No.2-6
[2] Shubham Jain; A. Pandian; "survey on automatic music generation", 2018, PP NO.2-3
[3] Briot,J.;Hadjeres, G.; and Pachet, F. 2018.DeepLearning Techniques for Music Generation Springer International Publishing.
[4] Chu, H.; Urtasun, R.; and Fidler, S. 2017. Song from PI: A musically plausible network for pop song generation. In Proc.International Conference on Learning Representations (ICLR), Workshop Track.
[5] Blaauw, M., and Bonada, J. 2017. A neural paramet singing synthesizer. In Proc. Interspeech.
[6] Lin, J.-C., Wei, W.-L., and Wang, H.-M. DEMVmatchmaker: emotional temporal course representation and deep similarity matching for automatic music video generation. In ICASSP, 2016.
[7] Eric Malmi, Pyry Takala, Hannu Toivonen, Tapani Raiko, and Aristides Gionis. DopeLearning: A Computational Approach to Rap Lyrics Generation, arXiv :1505.04771v1 [cs.LG], 18 May 2015.
[8] Peter Potash, Alexey Romanov, Anna Rumshisky. GhostWriter: Using an LSTM for Automatic Rap Lyric Generation, proceedings of conference on Empirical Methods in NLP, 2015
[9] Margareta Ackerman and David Loker. Algorithmic Song writing with ALYSIA, arXiv:1612.01058v1 [cs.AI], 4 Dec 2016.
[10] Jukka M. Toivanen and Hannu Toivonen and Alessandro Valitutti. Automatical Composition of Lyrical Songs, in The Fourth International Conference on Computational Creativity, 2013.
[11] Automatic generation of song lyrics on a semantic domain. Journal of Artificial General Intelligence, 6(1):87–110, 2015.
[12] Dekai Wu and Karteek Addanki, Learning to Rap Battle with Bilingual Recursive Neural Networks, proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)
[13] PhilippeHamel,SimonLemieux,YoshuaBengioandDouglasEck,"Temporal Pooling and multiscale learning for automatic annotation and ranking of Music audio",2015,PP.No.2-4.

Authors Profile

Lawakesh Patel is pursing Bachelor of Technology from AKTU, India. His main research work focuses on Machine Learning, IOT and Computational Intelligence based education.

Nidhi Singh is pursing Bachelor of Technology from AKTU, India. Her main research work focuses on Machine Learning, IoT and Computational Intelligence based education.She has one journal paper.

Rizwan Khan Associate Professor Department of Computern Science and Engineering Dr. Rizwan Khan is B.Tech (CSE) M.Tech (CSE) and Ph.D in Computer Engineering. He is a good researcher. He has 24 journal papers and 12 conference papers. He is also author on one book chapter and 3 books. He has 13 years of experience in teaching