Using Generative Model to Produce Instrumental Music

Samiksha Pandit[1] Shachita Jain[2] Siddharth Varangaonkar[3] Prof Kavita Namdev[4]
[4]Senior Assistant Professor
[1,2,3,4]AITR Indore, India

Abstract— This Research lets you play with new sounds created with machine learning. It's built using NextGenMusic, a research project that trained a neural network on over 300,000 instrument sounds. This application is a Music Streaming Web App and it's major feature is Song Mixing that is aimed towards the general public. The application allows users to obtains new set of notes from exiting instrumental notes. To improve user experience, the application also helps to adjust the meter of notes. The application also provides functionality to users to create personal profiles.

Keywords: Deep Learning, Generating Music, VAE (Variation Auto Encoder)

I. INTRODUCTION

NextGenMusic lets you select and combine two sound sources. Use the slider to change the interpolation between the two sound sources. Technology has always played a role in creating new types of sounds that inspire musicians—from the sounds of distortion to the electronic sounds of synths. Today, advances in machine learning and neural networks have opened up new possibilities for sound generation.

It's a machine learning algorithm that uses a deep neural network to learn the characteristics of sounds, and then create a completely new sound based on these characteristics. Rather than combining or blending the sounds, NextGenMusic synthesizes an entirely new sound using the acoustic qualities of the original sounds—so you could get a sound that's part flute and part sitar all at once.

II. LITERATURE SURVEY

Generative models in vision have seen rapid progress due to algorithmic improvements and the availability of high-quality image datasets. In this paper, we offer contributions in both these areas to enable similar progress in audio modeling. First, we detail a powerful new WaveNet-style auto encoder model that conditions an autoregressive decoder on temporal codes learned from the raw audio waveform. Second, we introduce NSynth, a large-scale and high-quality dataset of musical notes that is an order of magnitude larger than comparable public datasets. Using it, we demonstrate improved qualitative and quantitative performance of the WaveNet auto encoder over a well-tuned spectral auto encoder baseline. Finally, we show that the model learns a manifold of embeddings that allows for morphing between instruments, meaningfully interpolating in timbre to create new types of sounds that are realistic and expressive. Audio synthesis is important for a large range of applications including text-to-speech (TTS) systems and music generation. Audio generation algorithms, know as vocoders in TTS and synthesizers in music, respond to higher-level control signals to create fine-grained audio waveforms. Synthesizers have a long history of being hand-designed instruments, accepting control signals such as pitch', 'velocity', and filter parameters to shape the tone, timbre, and dynamics of a sound (Pinch et al., 2009).

III. PROPOSED WORK

VAEs learn latent space of musical sequences, providing different modes of interactive musical creation, including:
1) Random sampling from the prior distribution,
2) Interpolation between existing sequences,
3) Manipulation of existing sequences via attribute vectors or a latent constraint model.
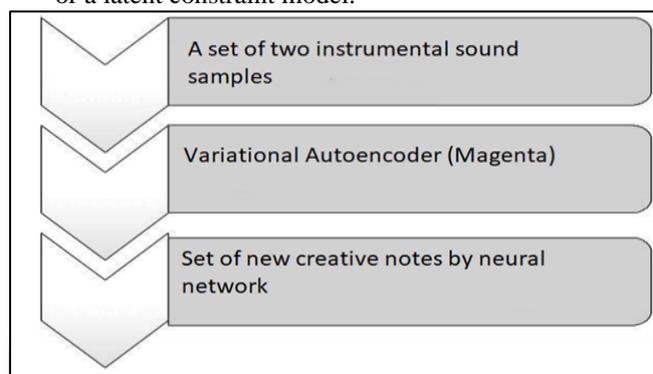

Fig. 1: process flow

Magenta is a research project 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 songs, images, drawings, and other materials. But it's also an exploration in building smart tools and interfaces that allow artists and musicians to extend (not replace!) their processes using these models.
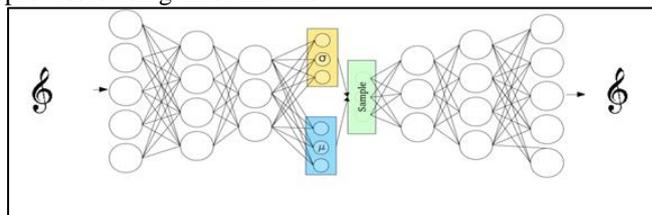
We found that users stress state is firmly identified with that of his/her friends in online life, and we utilize an enormous scale dataset from true social stages to methodically think about the relationship of users' stress states and social associations. We initially characterize a lot of stress related printed, and social properties from different viewpoints, and afterward propose a novel trained model joined with the Naïve Bayes Classifier to use social media Tweets and social association data for stress detection. By further breaking down the social association information, we likewise find a few captivating marvels, the substance of social communications, e.g., textual comment content, while disregarding the intrinsic auxiliary data like how users are associated.
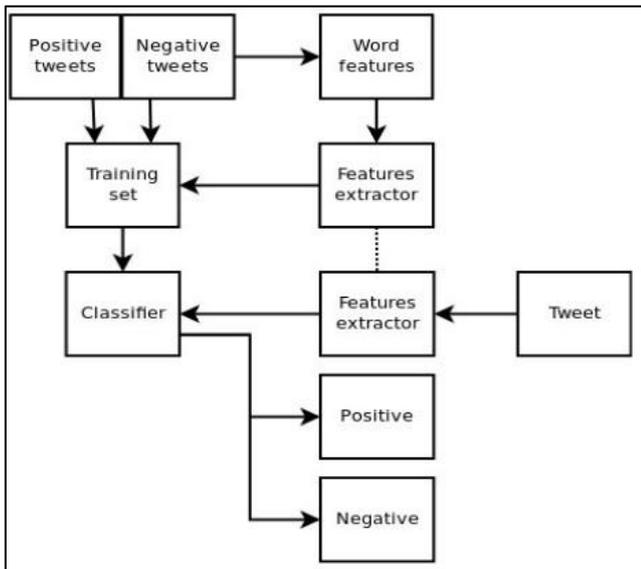
Fig. 2: System Architecture

The main motto of this system is to reduce the manual efforts made till now in the process of Stress detection through long and lengthy surveys. Thus, this project is operationally feasible and very – efficient, the operational feasibility is highly achieved through this system.

IV. PROBLEM STATEMENT AND METHODOLOGY

This application is a Music Streaming Web App and it's major feature is Song Mixing that is aimed towards the general public. The application allows users to obtaina new set of notes from exiting instrumental notes. To improve user experience, the application also helps to adjust the meter of notes. The application also provides functionality to users to create personal profiles.

Deep learning can be used to solve the problem of creativity in music. Neural networks in a way can be used as generative models. Generative models are neural networks which generate an output based on the input with its own set of learning. GANs, VAEs and autoencoders are widely used for such tasks .We can use an Variational autoencoder in our project.
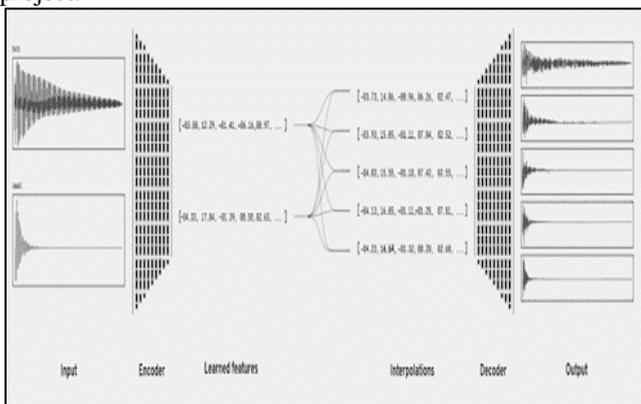

Fig. 3: encoder and deoder matrix

In contrast to the more standard uses of neural networks as regressors or classifiers, Variational Autoencoders (VAEs) are powerful generative models, now having applications as diverse as from generating fake human faces, to producing purely synthetic music.

When using generative models, you could simply want to generate a random, new output, that looks similar to the training data, and you can certainly do that too with VAEs. But more often, you'd like to alter, or explore variations on data you already have, and not just in a random way either, but in a desired, specific direction. This is where VAEs work better than any other method currently available. Variational Autoencoders (VAEs) have one fundamentally unique property that separates them from vanilla autoencoders, and it is this property that makes them so useful for generative modeling: their latent spaces are, by design, continuous, allowing easy random sampling and interpolation. Hence by combining two different instrumental sounds we can feed this data to our VAEs accordingly so that it will generate a combined different music based on that.

V. S

A set of experiments carried out on the MagentaJs used for generating music ,exploring the role of machinelearning. The end result is thus a creative system which helps us in generating new notes and advances with the concept of bringing enhancements in the musical industry and breaking that monotonous generation of same kind of music. This system would combine the notes of two different instruments to generate a new musical note of some amazingly new instrument.

VI. CONCLUSION

The research will help users to generate music based on pre existing instrumental notes. With limited scales and limited choice of notes, nowadays users cannot create more creative music hence to help them our web app can give them a new variety of music.

REFERENCES

[1] Madakam, Somayya, R. Ramaswamy, and Siddharth Tripathi. "Internet of Things (IoT): A literature review." Journal of Computer and Communications 3.05 (2015): 164.

[2] Alpaydin, Ethem. Introduction to machine learning. MIT press, 2009.

[3] Xiao, Laisheng, and Zhengxia Wang. "Internet of things: A new application for intelligent traffic monitoring system." Journal of networks 6.6 (2011): 887.

[4] Zhao, Dongbin, Yujie Dai, and Zhen Zhang. "Computational intelligence in urban traffic signal control: A survey." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 42.4 (2012): 485-494.

[5] Zhou, Binbin, Jiannong Cao, Xiaoqin Zeng, and Hejun Wu. "Adaptive traffic light control in wireless sensor network-based intelligent transportation system." In Vehicular technology conference fall (VTC 2010-Fall), 2010 IEEE 72nd, pp. 1-5. IEEE, 2010.

[6] Arel, Itamar, Cong Liu, T. Urbanik, and A. G. Kohls. "Reinforcement learning-based multi-agent system for network traffic signal control." IET Intelligent Transport Systems 4, no. 2 (2010): 128-135.

[7] Gartner, Nathan. (1983).