**Evaluation of Machine Learning Models for Music Generation**

Music is a complex combination of various things likes Harmonies and Melodies even for humans who have advanced cognitive abilities. For automatic music generation we are not at the liberty to assume that machines would possess similar kinds of abilities simply because machines are not there yet. This makes the process of understanding music even more difficult for machines.

Classical machine learning techniques such as Support Vector Machines or K-Nearest Neighbors are what we might term as single layered or shallow architectures that do not provide a great amount of flexibility when it comes transforming the raw data set of complex notes, chords, velocities, offsets, and other parameters into a feature space that can be useful for training models on single layered architecture (something SVM can do using Kernel trick). Having said that, the only upside to using these models would very quick times that data would be trained in (if we can work up a feature space). That brings the other important thing to keep in mind with Classical machine learning techniques i.e., feature engineering. There would be a lot more expertise required in music theory to generate a workable feature set if we were to use classical machine learning theory which again boils down to the fact that the shallow architecture doesn’t allow us to capture contexts and regularities in data.

That’s where something similar to how a human perceives music or in general Audio signal helps us. Since our ability to comprehend these signals are made up of complex processing mechanisms, it only makes sense for the machine to have a similar kind of architecture that can break down complex contexts and regularities from data.

Deep Learning architectures are composed of many layers of non-linear steps for processing raw data. All the steps share data amongst themselves, and this allows the developed model to be generative in nature and helps in making effective use of large amounts of data for extracting meaningful patterns that our model can use to make future predictions or just cluster unlabelled data in general (something that’s not possible to do in classical machine learning approaches)

The most basic of deep learning algorithms such as Feed forward network with back-propagation is one of the approaches that can be followed to do all the things listed above those classical approaches cannot do. The downside with this approach though is the loss function/objective function which in most practical cases in a non-convex function and finding whose global minimum is something that this approach cannot do efficiently.