Machine Learning Project Course Proposal

Jayakrishna Sahit (16110104), Sai Praneeth (16110145), Rendla Aditya (16110137)

Abstract

With electronic music and instruments being a growing trend, it is often easy to be able to convert anything we play on these instruments into MIDI files and be able to retrieve score cards for them, however for acoustic instruments we must rely purely on the audio we capture using a microphone. Given that most of the research has been done on a piano, in this project we choose to present our research on data from an acoustic guitar. We wish to approach this by doing a comparative analysis on piano data and guitar data, and using these results propose a method to be able to retrieve musical notes from music of an acoustic guitar. We then evaluate our method, by comparing them with methods which have been proposed exclusively on piano data.

Introduction

What is the motivation? Or, why is this an important problem to solve? How would the world change if this problem is solved?

Music identification from acoustic instruments is an added advantage for people interested in acoustic instruments, because they now have an easier way of learning songs using such instruments. On a broader note, understanding such instruments, can give us better insight on understanding sound from non-electrical devices, such as the human voice box, etc

How do you envision this to be a machine learning problem?

For a person who is a music enthusiast and has learned basics of music, music in the form of notes is very important. Since he/she cannot store notes of all the music he/she wants to learn, there should be someway to automate this. So, the main problem is to detect a bit of music which note it belongs to and we can achieve this by machine learning methods. Training the

model on different notes at different instances(samples) and then predicting the note when a music bit is given.

What kind of an ML problem is this?

It is a supervised ML problem, with ground truth labels as musical notes of training data

Is it a regression or classification problem?

It is a classification problem, where the task is to classify each time slice of input audio as one note of the guitar

Related Work

What is the state-of-the art? Are baseline implementations available? Do you have a table of results from a previous paper?

With the goal of becoming the ultimate platform for notation based music creation Doremir's unique technology enables users to easier than ever before capture and digitize their musical creativity – just like "Google Translate", but for music. The technology behind this product is based on 20 years of research in music cognition.

Website: https://scorecloud.com/

Previous work on this topic had been done by Karey Shi, Michael Bereket and titled their paper as "*An AI Approach to Automatic Natural Music Transcription*". Here is the results table of the above mentioned paper

LR Schedule	Accuracy	F1-score	Recall	Precision
Initial LR = 0.1, halving every 5 epochs	98.33%	52.85%	50.10%	59.74%
Initial LR = 0.05, halving every 10 epochs	**98.72%**	**55.07%**	**52.54%**	**60.51%**
Manual step decay	98.72%	54.45%	51.91%	59.49%

Table 1: Validation set results with step-decay learning rate schedules after 40 training epochs.

Datasets

How many are available? If not, how will you go about it?

There are various MIDI files available for guitar. But still work has to be done to extract the notes from these MIDI files. Here is one link where we can find classical MIDI guitar archives.

However if the necessary arrives, we can make our own datasets. Collecting our own data can be done, by having an audio recording of one of the group members playing the guitar, which can then be manually labelled for musical notes.

Experimental Plan

What experimental setting will you choose? How to find best params for your algos and baselines?

We are plan to use deep learning models where we split the data into a normal 70:30 for train and test. We would also like to try out our model for music generated by us. We would run the following in comparison to algorithms which have been studied in a piano setting.

Plan is to try out various architectures which will mostly be based on previous papers and work, for the models and use near exhaustive methods to find various hyperparameters.

What metric will you choose? And why? Does this relate to the motivation?

One of the base metric is accuracy. We also plan to use recall and f1-score to evaluate the model. Since when predicting a binary vector(of size 88) to indicate the presence of note ‘i’(out of 88 notes), one could give a naive predict of zero vector as an output which does very well when evaluated using accuracy. So this brings the necessity of using recall and also f1-score to evaluate the trained model.

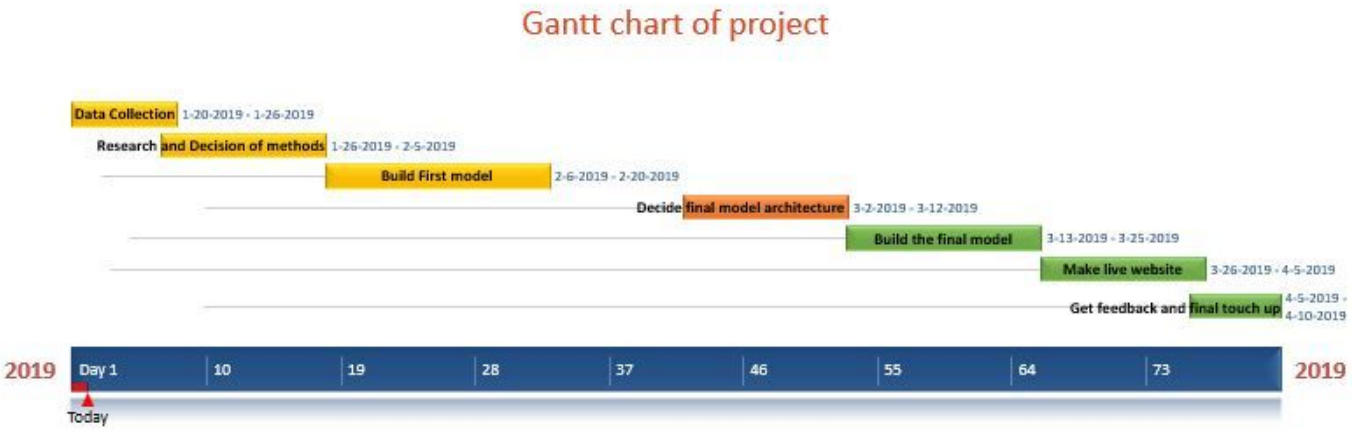
What level of systems effort will you require? Please paint a scenario for the final demo that you will present. Will it be a live website? Or, an Android app?

We plan to build live website, where user inputs an mp3/wav audio file or simply plays the

guitar live to get the guitar notes.

Project Management

Project Gantt Chart			DURATION (days)
START DATE	END DATE	DESCRIPTION	
1-20-19	1-26-19	Data collection	7
1-26-19	2-5-19	Research and decision of methods	11
2-6-19	2-20-19	Build first model	15
3-2-19	3-12-19	Decide final model architecture	11
3-13-19	3-25-19	Build the final model	13
3-26-19	4-5-19	Make live website	11
4-5-19	4-10-19	Get feedback and give final touch up	6
			74



Computation resources needed

If we are to work with a small dataset (most likely will not suffice), our local machines are to be enough, however otherwise we would need Google cloud credits(TPUs)

When will you consider project a success?

When we are able to built a model that performs better on the given data when compared with the model in above mentioned paper. When it meets our expectations by running successfully live without any flaws

What is the biggest risk that will lead to project failure and how will you address it?

Decision of final model architecture will decide the fate of the project. We are spending 1/3rd of the project time to research and decide the final architecture for model. Hence, by spending sufficient time towards the above decision we hope to reduce the risk in architecture decision as much as possible

How do you wish to divide the work amongst yourselves?

Working on the architecture will be done by all members of the group. Division of work would come during the final coding up of the model.