FIRST RESEARCH PLAN FORM

Name:
Nikit Uppala, Sai Akarsh

The tentative title of my project is *A problem/statement / question / phenomenon?*

Musical Tempo and Key Estimation using CNN with Directional Filters

I want to find out / explore *Main purpose / research question / issues*

Explore how the different semantics of spectrograms' time and frequency axes can be exploited for musical tempo and key estimation using CNN. By addressing both tasks with the same network architectures ranging from shallow, domain-specific approaches to deep variants with directional filters, we show that axis-aligned architectures perform similarly well as common VGG-style networks developed for computer vision, while being less vulnerable to confounding factors and requiring fewer model parameters.

Because it is important to understand *Context and significance of the study*

Tempo and key estimation can be seen as tasks from two different ends of a spectrum of common MIR tasks, which are addressed by systems relying more or less on temporal or spectral signal properties. Systems for other tasks like general purpose tagging or genre recognition are found more towards the center of the spectrum as they usually require both spectral and temporal information.

This means that my research design will be *Qualitative? Quantitative? Case Study? etc.*

The research design of the project will be to estimate the key and tempo of a given musical piece. Hence it is quantitative in nature.

In order to find out what is already known about my topic, I have to read at least about

Key estimation in music
Tempo estimation in music
CNN using rectangular filters
VGG-style architecture

And define the following concepts

Key: The key of a piece is the group of pitches that forms the basis of a music composition that can be classified as 12 minor and major tones.
Tempo: Speed or pace of given piece of music that is measured in BPM (beats per minute)

I plan to collect my data by means of *Research instrument?*

Using datasets that are already available online.

This way of data collection is suitable for my purpose, because *Rationale: advantages*

The data is already available and hence does not require us to manually record it.

But I also have to remember that this way of data collection has the following drawbacks / limitations

Not all the data points are suitable because we are using a big data set. We need to make sure that we are using pieces that are relevant.

My research instrument should be ready by: 09-03-2022

My data collection should be ready by: 09-03-2022