MAIS 202 Deliverable 2

Problem statement:

We aim to develop a convolutional neural network classification model that can recognize a key press on a certain keyboard given the audio of said key press.

Data preprocessing:

We will use a short time Fourier transform (STFT) to separate individual keystroke from audio sampling.

Our objective is to generate a spectrogram for every isolated key (L/R if stereo permits) and be able to match each of these spectrograms with a key.

Some hyperparameter that have been identified as of now are the number of samples per segment (npergs) which range from 0 – 2048 and affects the resolution of the Fourier spectrogram. A smaller value works best with fast typing.

The cutoff threshold for filtering noise for the fourrier transform is another hyperparameter that can be calibrated and tailored for a specific keyboard.

Machine learning model:

We will stick with the convolutional neural network model proposed in deliverable 1. We have not actually implemented our model yet, but we will likely use a library such as PyTorch for its neural network functionality, as well as NumPy for numerical manipulation. As we have not implemented this yet, the exact architecture is not confirmed.

We have written part of the training for our model, which will use a combination of key recording with audio recording to build a dataset.