Week 6 Meeting

2357351G - MSci Half Project

What I've done this week

- Read *Learning to Build Natural Audio Production Interfaces (2019).*
- Downloaded and had a play with the *SocialFx* dataset.
- Read and summarised *Word Embeddings for Automatic Equalization in Audio Mixing (2022).*
- Spent some time comparing the SAFE-DB dataset and plugins with the AudioCommons timbre extractor models.
- Watched a talk from the Audio Developer Conference from an *iZotope* engineer around how they apply deep learning to their plugins.
- Read and summarised *Semantic Description of Timbral Transformations in Music Production (2016)*.

Questions

- Would you mind going over what you expect from the Interim Report in terms of breadth of literature review and how defined the research problem should be at the point of submission?
- Also, how much of the implementation of the deep learning model would you expect to be completed by the end of the first semester?

Plan for next week

- Need to consider the inputs/outputs of the model (parameter settings or audio signals) and how this will affect the overall architecture.
- Have a look at whether the SocialFx and SAFE-DB datasets can be combined without compromising the data.
- See if these effects can be applied to general EQ plugins, or whether they will be specific to the plugins used for dataset generation.

Where I am in schedule

- Believe that the SocialFx and SAFE-DB datasets should be enough for the project.
- Have a loose high level idea of what the 'workflow' of the model might look like:
  - User inputs a descriptive word/phrase.
  - Model produces some audio effect (one or multiple plugins? Fixed set of plugins or flexible?)
  - User can then make some alteration to these parameters (in the latent space or the actual low-level parameters? Using NLP or controls?)
  - These user preferences are somehow stored for future inference.