# MA Notes

## Lead Question

How can I use denoising diffusion probabilistic models in deep learning to create music based on DJ sets as training data?

Goal: Create Music in a similar genre like the training data -> Possibly with simple prompts, e.g. BPM or other automatically labeled characteristics.

## Different structures:

* Diffusion
  + Start with random noise and iteratively refine it to generate data. Training is reverse operation.
  + Original Diffusion Paper 2015: <http://arxiv.org/pdf/1503.03585>
  + DDPM: <http://arxiv.org/pdf/2006.11239>
  + Improved DDPM: <http://arxiv.org/pdf/2102.09672>
  + OpenAI Paper Diffusion: <https://arxiv.org/pdf/2105.05233>
  + Apple Paper Diffusion for music gen: <https://arxiv.org/pdf/2311.00613>
  + General Explanations: <https://scholar.harvard.edu/binxuw/classes/machine-learning-scratch/materials/foundation-diffusion-generative-models>, <https://www.superannotate.com/blog/diffusion-models>
  + <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/>
  + Latent Diffusion: <https://arxiv.org/pdf/2112.10752>
  + ETH Paper Diffusion for music gen: <https://arxiv.org/pdf/2301.13267>
  + Hierarchical diffusion <https://arxiv.org/abs/2401.02644> , <https://www.sintra-granulats.ca/aiseo/what-is-reverse-hierarchical-diffusion> . Create to models, m1 low resolution long bits, m2 high resolution
  + Distillation for speedup

## Training

* Unsupervised Learning
* Implicit labels
  + Time segmentation, sliding window, cluster for soft labeling
* Label
  + Possibly label small portion for fine tuning

## Additional Libs

* Spleeter: Possibly split into stems
* Librosa: for labeling e.g. BPM, -> splitting up the dj sets?

## Other Considerations

* Raspberry pi compute module etc. to train the model, API to access model
* Questions: Copyright -> Big training data with a high artist diversity -> should not be a problem because influence of single artist very low