

DEEP MUSICIAN

AUTOMATIC MUSIC GENERATION USING DEEPLARNING

Fabian Gabelberger

January 18, 2023

Music Generation

- Create a model that can automatically generate music
- Train the model on a large dataset of existing music
- Use the model to generate new, unheard music of arbitrary length

Preprocessing - get the data in shape

- Use symbolic representation of music
 - Midi Files
 - **Piano roll**, that captures the notes played at each time step (2D-Array)

Architecture

- Sequence aware **encoder-decoder** architecture
 - Each consist of a two-layered **GRU**

Sequence 2 Sequence

- During **training** the model sees sequences and tries to predict the next note
 - After training the model is started with an empty step
- This triggers the model to incrementally **generate** a sequence of notes

RESULTS

- The results measured in terms of the loss function and accuracy are very promising.
 - Yet the generated music is not very good and most of the time only an empty sequence was returned.
- The model is not able to generate a coherent melody, since the training data is very **imbalanced**: Most of the time the notes are not played at all.

Preprocessing

- Endless possibilities to represent the underlying data
- There is not a single best representation: each have there own advantages and disadvantages
- Beside the piano roll representation, there are also symbolic representations that represent the music as a sequence of words.

Evaluation

- Although the loss function and accuracy are a good indicator for the performance of the model, they are not a good measure for the quality of the generated music.
 - Instead of a classical BCE-loss, I used a loss function that was initially developed for image recognition and object detection: **Focal Loss**.
- The main idea behind focal loss is to down-weight the contribution of easy examples in the training data and focus more on the hard examples