

4-bar phrase + tempo
input: midi, tempo

Melody Continuation

n-bar phrase + structure
output: midi + "A4A4B4A4"

Baseline: Music SketchNet

Original dataset: Irish Folk Song' *monophonic and continuous*

Targeted dataset: Pop 90s *polyphonic and discontinuous*

Data Preprocessing

multi-track $\xrightarrow{\text{extract}}$ melody $\xrightarrow{\text{split}}$ phrases (sections) $\xrightarrow{\text{filter}}$ 4-bar / 8-bar phrases

Training

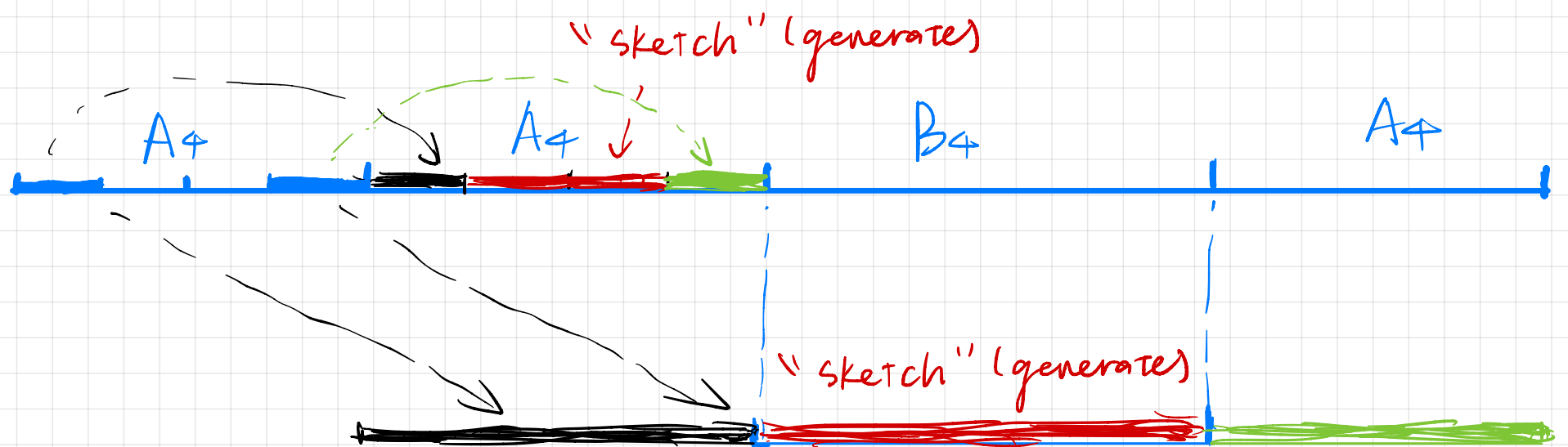
Dataset $\xleftarrow{\text{tempo}=80 \text{ bpm}}$

dataset $\xrightarrow{\text{DataLoader}}$ train/test/val (4-bar phrases)

numpy array shape: each phrase \rightarrow 24 frames

- Note
- Melody $\xrightarrow{\text{tokenize}}$
- Rhythm
- melody length (int)

Structure



Past

generate

future