

```
In[ ]:= SetDirectory[NotebookDirectory[]];  
Get["../Howl/HowlMidiTools.wl"]
```

My dataset is a collection of .mid files.

```
datasetPath = "E:\\datasets\\music\\alec_finetune";  
{time, dataset} = Map[HowlMidiImport, HowlFindMidis@datasetPath] // AbsoluteTiming;  
time  
Length@dataset
```

```
Out[ ]:= 637.679
```

```
Out[ ]:= 145
```

We will need to convert the dataset format of

```
{{timeSincePrev, duration, volume, noteInt}, ...}
```

Note, data can fail to load, so we will remove all of the failures first.

```
In[ ]:= datasetNoFailures = Query[Select[! FailureQ[#["Notes"]] &]] [dataset];  
Length @ datasetNoFailures
```

```
Out[ ]:= 145
```

```
Export["dataset_music_finetune_20210718.wxf", datasetNoFailures]
```

```
Out[ ]:= dataset_music_finetune_20200718.wxf
```

We will now extract the encoded data we are training on.

```
In[ ]:= datasetEncoded = (Key["EncodedNotesV1"] /@ datasetNoFailures) /. _Missing -> Nothing;  
Length @ datasetEncoded
```

```
Out[ ]:= 145
```