SMAI Project Handout Document

Team: Team 1

Project ID: 11

Difference in Playing Styles of Different Composers

Fig. 1

Cellular Automata Output

Fig. 2 Here y-axis denote the notes number and x-axis denote time. White Ones are on and Black ones are off

MLP

```
=======================================================================
Layer (type:depth-idx)              Output Shape          Param #
=======================================================================
MLP                                 [128]                 --
├─Linear: 1-1                       [1000]                12,801,000
├─ReLU: 1-2                         [1000]                --
├─Linear: 1-3                       [128]                 128,128
=======================================================================
Total params: 12,929,128
Trainable params: 12,929,128
Non-trainable params: 0
Total mult-adds (G): 12.82
=======================================================================
```
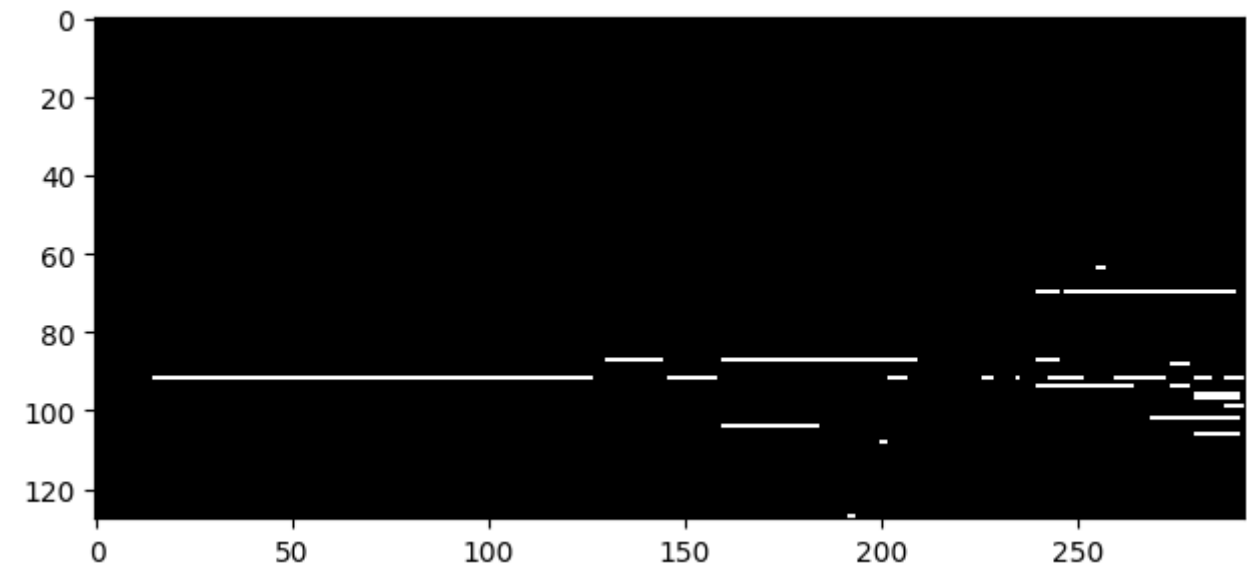
Fig. 3 MLP Architecture

Fig. 4 Output of MLP. Here y-axis denote the key pressed and the x-axis denote the time. White Ones are presses and Black ones are not pressed

LSTM

```
=======================================================================
Layer (type:depth-idx)              Output Shape          Param #
=======================================================================
RNN                                 [128, 2]              --
├─Linear: 1-1                       [1, 1, 512]           66,048
├─BatchNorm1d: 1-2                  [1, 512, 1]           1,024
├─LSTM: 1-3                         [1, 512]              4,202,496
├─BatchNorm1d: 1-4                  [1, 512, 1]           (recursive)
├─Linear: 1-5                       [1, 1, 128]           65,664
=======================================================================
Total params: 4,335,232
Trainable params: 4,335,232
Non-trainable params: 0
Total mult-adds (G): 2.15
=======================================================================
```
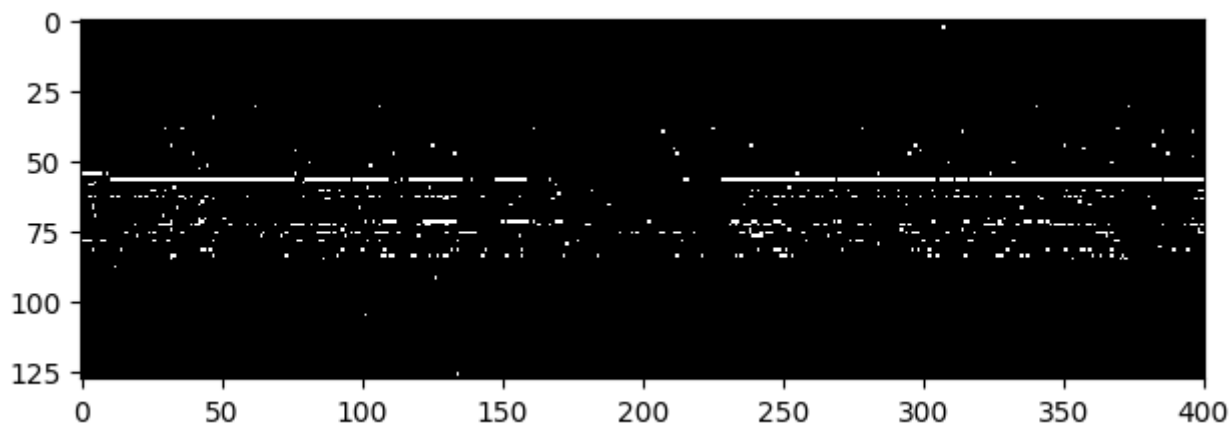
Fig. 5 LSTM Architecture

Fig. 6 Output of LSTM. Here y-axis denote the key pressed and the x-axis denote the time. White Ones are presses and Black ones are not pressed

Encoder Decoder

```
========================================================================
Layer (type:depth-idx)              Output Shape            Param #
========================================================================
RNN                                 [128, 2]                --
├─encoder: 1-1                      [2, 1, 512]             --
│   └─Linear: 2-1                   [1, 1, 512]             66,048
│   └─GRU: 2-2                       [1, 512]               3,151,872
├─decoder: 1-2                      [1, 1, 128]             --
│   └─GRU: 2-3                       [1, 1, 512]             3,151,872
│   └─Linear: 2-4                   [1, 1, 128]             65,664
========================================================================
Total params: 6,435,456
Trainable params: 6,435,456
Non-trainable params: 0
Total mult-adds (G): 1.62
========================================================================
```

Fig. 7 Encoder Decoder Architecture

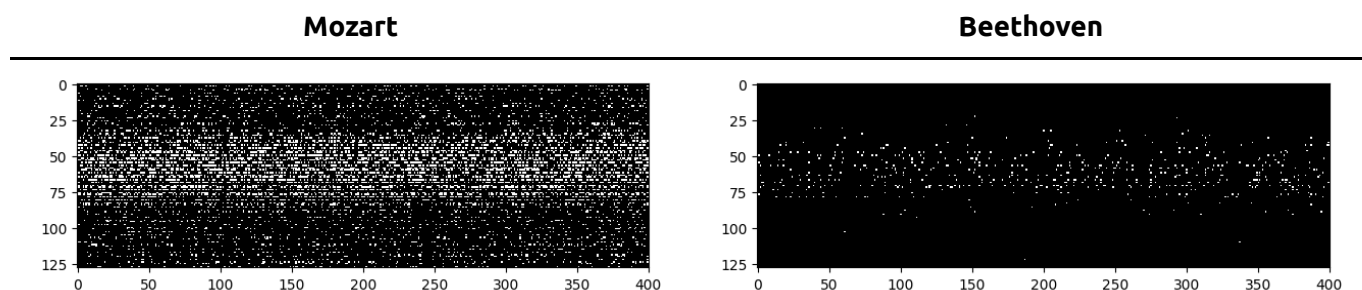**Mozart**   **Beethoven**

Fig. 8 Output of Encoder Decoder. Here y-axis denote the key pressed and the x-axis denote the time. White Ones are presses and Black ones are not pressed

Classifier

Filename	KNN	RFC	MLP	DT	Stacking
ceullular.mid	Pop	Pop	Jazz	Pop	Pop
MLP_bad.mid	Pop	Pop	Pop	Pop	Pop
LSTM_1.mid	Classical	Classical	Classical	Classical	Classical
LSTM_2.mid	Classical	Classical	Classical	Classical	Classical
lstm1_train.mid	Classical	Classical	Classical	Classical	Classical
ED_Beethoven_good.mid	Classical	Classical	Classical	Classical	Classical
ED_Bach_trash.mid	Classical	Pop	Classical	Pop	Jazz

Fig. 9 Outputs of the Classifier on the various shown midi files