

EE336 Project Presentation

Music Generation and Visualization

Mohit Manoj - 200102106

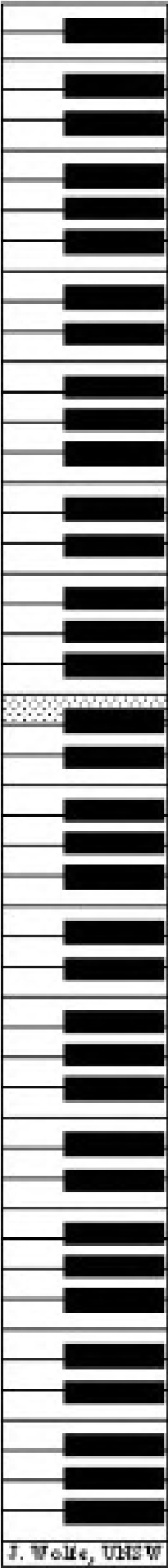
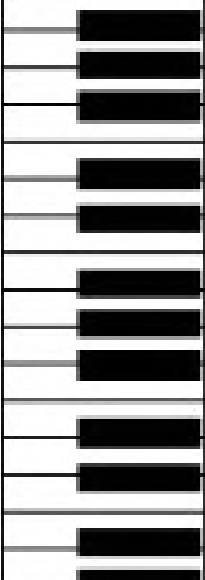
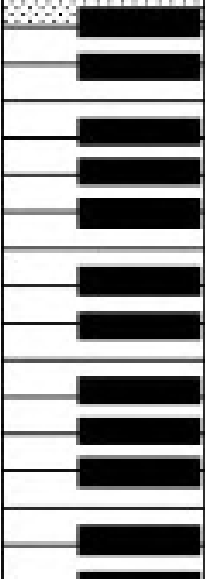
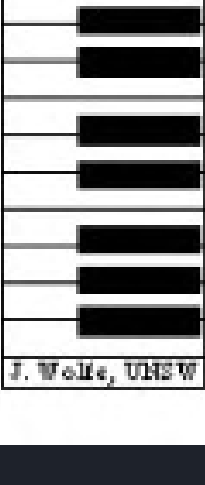
Sanskar Kejriwal - 200102082

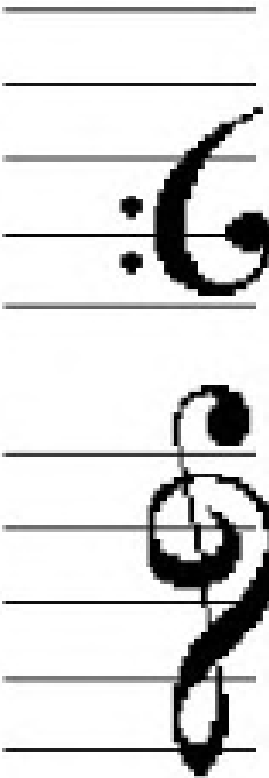
Manish Kumar - 200102108

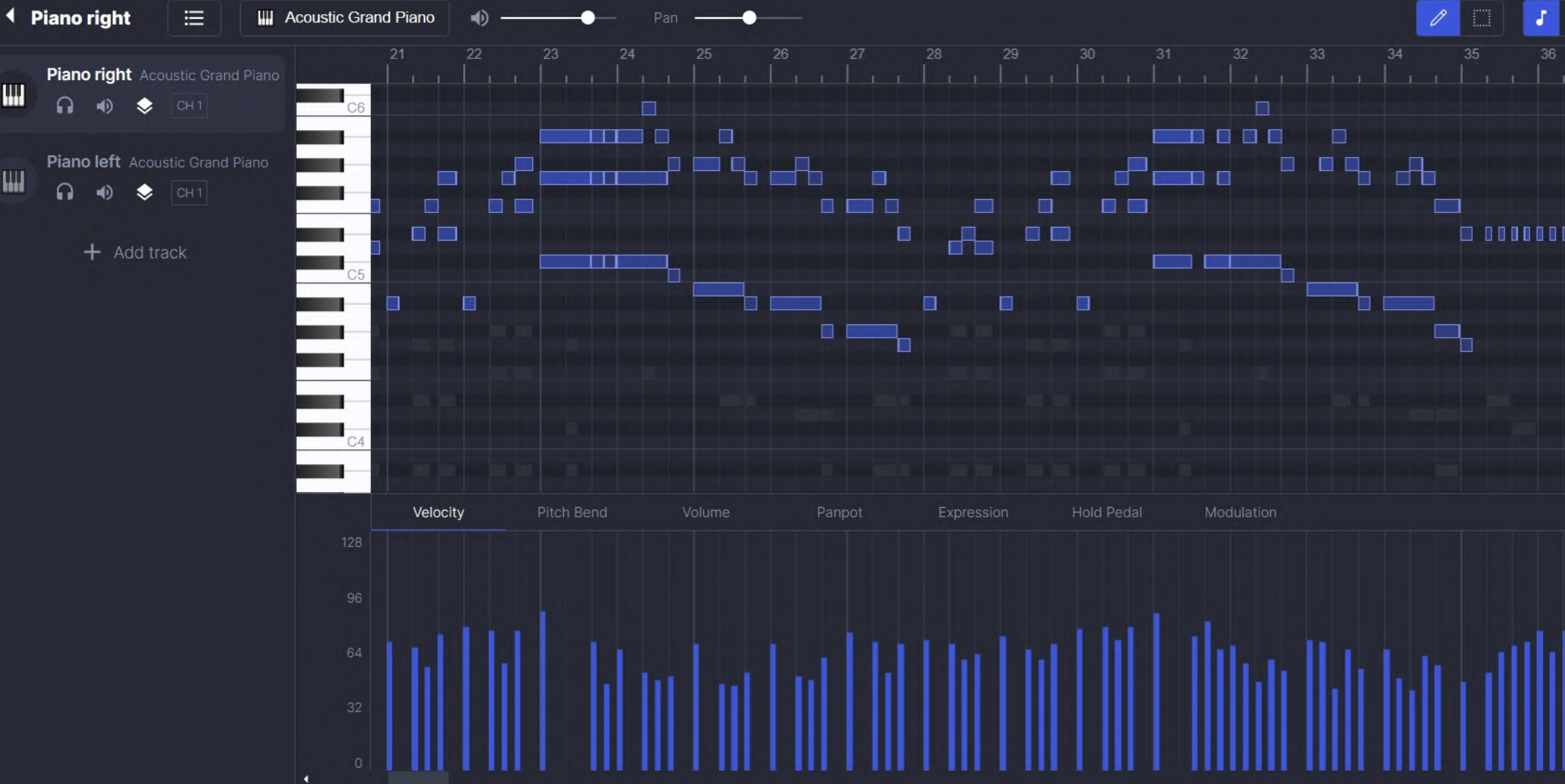
Midi Files

WHAT ARE THESE ?

A file with the .MID or .MIDI file extension is a Musical Instrument Digital Interface file. Unlike regular audio files like MP3s or WAVs, these don't contain actual audio data and are therefore much smaller in size. They instead explain what notes are played, when they're played, and how long or loud each note should be.

MIDI number		Note name	Keyboard	Frequency Hz		Period ms	
21	22	A0		27.500		36.36	
23		B0		30.868	29.135	32.40	34.32
24	25	C1		32.703		30.58	
26		D1		36.708	34.648	27.24	28.86
28	27	E1		41.203	38.891	24.27	25.71
29		F1		43.654		22.91	
31	30	G1		48.999	46.249	20.41	21.62
33		A1		55.000	51.913	18.18	19.26
35	34	B1		61.735	58.270	16.20	17.16
36		C2		65.406		15.29	
38	37	D2		73.416	69.296	13.62	14.29
40		E2		82.407	77.782	12.13	12.86
41	42	F2		87.307		11.45	
43		G2		97.999	92.499	10.20	10.81
45	44	A2		110.00	103.83	9.091	9.631
47		B2		123.47	116.54	8.099	8.581
48	49	C3		130.81		7.645	
50		D3		146.83	138.59	6.811	7.216
52	51	E3		164.81	155.56	6.068	6.428
53		F3		174.61		5.727	
55	54	G3		196.00	185.00	5.102	5.405
57		A3		220.00	207.65	4.545	4.816
59	58	B3		246.94	233.08	4.050	4.290
60		C4		261.63		3.822	
62	61	D4		293.67	277.18	3.405	3.608
64		E4		329.63	311.13	3.034	3.214
65	66	F4		349.23		2.863	
67		G4		392.00	369.99	2.551	2.703
69	68	A4		440.00	415.30	2.273	2.408
71		B4		493.88	466.16	2.025	2.145
72	73	C5		523.25		1.910	
74		D5		587.33	554.37	1.703	1.804
76	75	E5		659.26	622.25	1.517	1.607
77		F5		698.46		1.432	
79	78	G5		783.99	739.99	1.276	1.351
81		A5		880.00	830.61	1.136	1.204
83	82	B5		987.77	932.33	1.012	1.073
84		C6		1046.5		0.9556	
86	85	D6		1174.7	1108.7	0.8513	0.9020
88		E6		1318.5	1244.5	0.7584	0.8034
89	90	F6		1396.9		0.7159	
91		G6		1568.0	1480.0	0.6378	0.6757
93	92	A6		1760.0	1661.2	0.5682	0.6020
95		B6		1975.5	1864.7	0.5062	0.5363
96	97	C7		2093.0		0.4778	
98		D7		2349.3	2217.5	0.4257	0.4510
100	99	E7		2637.0	2489.0	0.3792	0.4018
101		F7		2793.0		0.3580	
103	102	G7		3136.0	2960.0	0.3189	0.3378
105		A7		3520.0	3322.4	0.2841	0.3010
107	106	B7		3951.1	3729.3	0.2531	0.2681
108		C8		4186.0		0.2389	





Midi Files

MIDI messages are made up of 8-bit *words* that are transmitted serially at a rate of 31.25 kbps. The first bit of each word identifies whether the word is a status byte or a data byte, and is followed by seven bits of information. A start bit and a stop bit are added to each byte for framing purposes, so a MIDI byte requires ten bits for transmission.

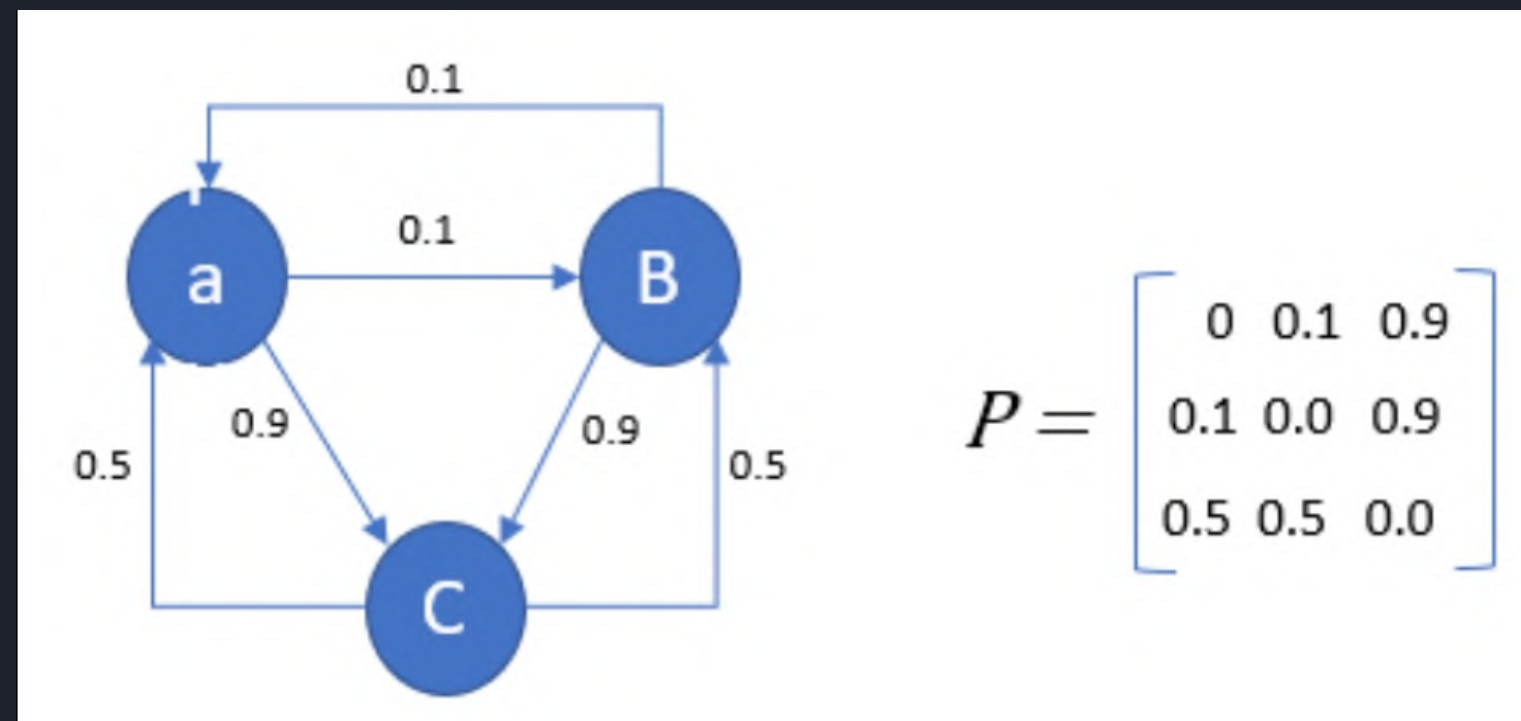


Markov Chains

VERY BASIC MODEL

Markov Chain

- Discrete-time stochastic process.
- Follows Markovian Property



$$P(Y_n | X_1, X_2, \dots, X_{n-1}) = P(Y_n | X_{n-1})$$

What we did?

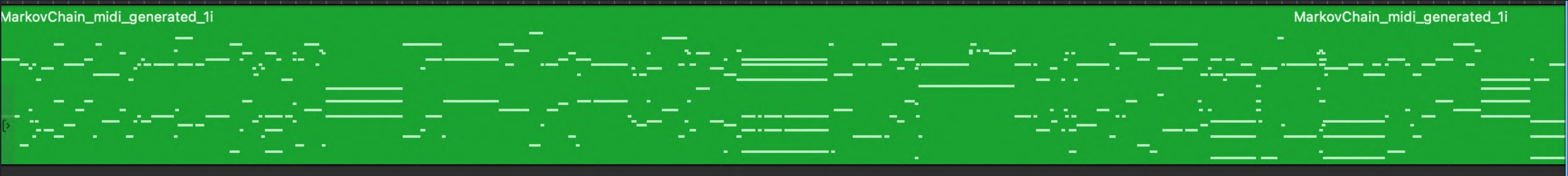
STEPS :

- Representing Midi sequence as vectors of size $88 * N$.
- Elements of vectors contain notes of velocity in range 0-128.
- Represent all unique vectors as states.
- Construct a Transition Matrix using the frequency of notes transition.
- Doing random walk to generate new notes



Results

Without Memory



With Memory





LSTMs

LONG SHORT TERM MEMORY

LSTMs(Neural Network)

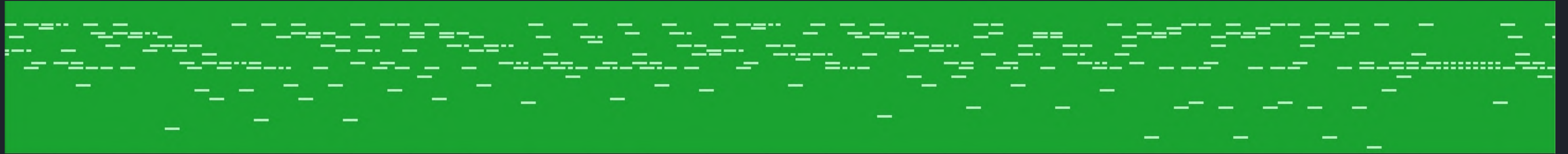
WHAT IS LSTM?

LSTM stands for **long short-term memory networks**, used in the field of Deep Learning. It is a variety of recurrent neural networks (RNNs) that are capable of learning long-term dependencies, especially in sequence prediction problems.

STEPS INVOLVED:

- Extract Notes from Midi file.
- Prepare input and output sequence for training and mapping the sequence to integers.
- Training the neural network
- Taking a random sequence and predict next notes from it by unmapping using same dictionary.
- Convert the notes predicted into Midi file.

Results





GAN

GENERATIVE ADVERSERIAL NETWORK

GANs

WHAT ARE GANS?

GANs are a class of Generative Models, i.e, they create new models that resemble the training data. It is composed of two neural networks - a generator and a discriminator which compete against each other. The generator creates fake samples and the discriminator decides whether these samples are fake or real.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

DCGANS

Deep Convolutional GANs uses convolutional and transpose convolutional layers in the discriminator and generator respectively. DCGANs use images to train and generate new examples that resemble the training data.

Steps followed

PREPROCESSING

- Load Dataset containing MIDI files.
- Preprocess all MIDI files into numpy arrays.
- Convert all MIDI files into images of shape (106, 106)

BUILDING AND TRAINING THE MODEL

- Using tensorflow, build models of the Generator and Discriminator.
- Train on the model.
- Generate new images from the trained model.
- Convert the generated images into MIDI files.

Results (plots)





Visualization

BONUS :)



Thanks