EE336 Project Presentation

Music Generation and Visualization

Mohit Manoj - 200102106

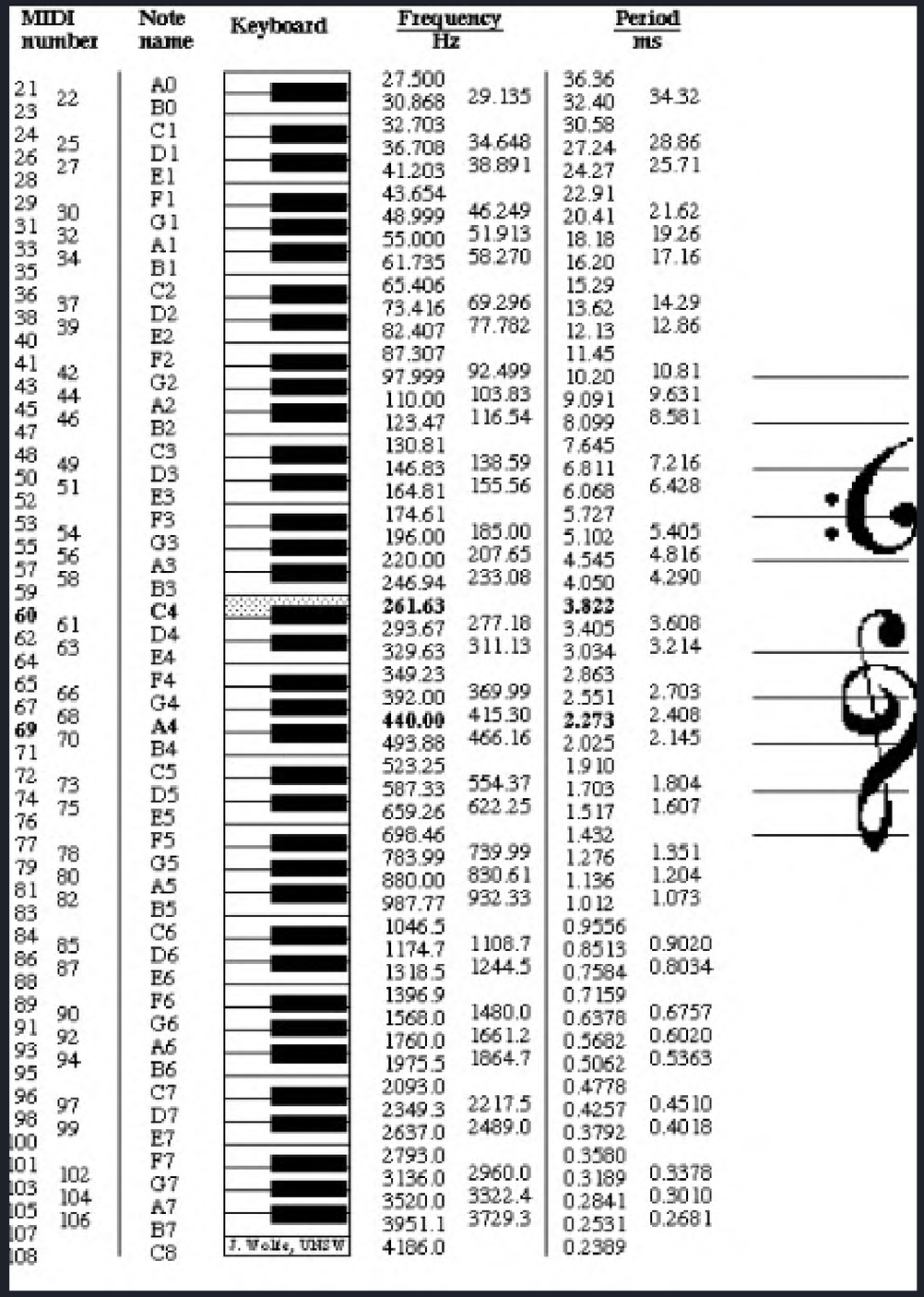
Sanskar Kejriwal - 200102082

Manish Kumar - 200102108

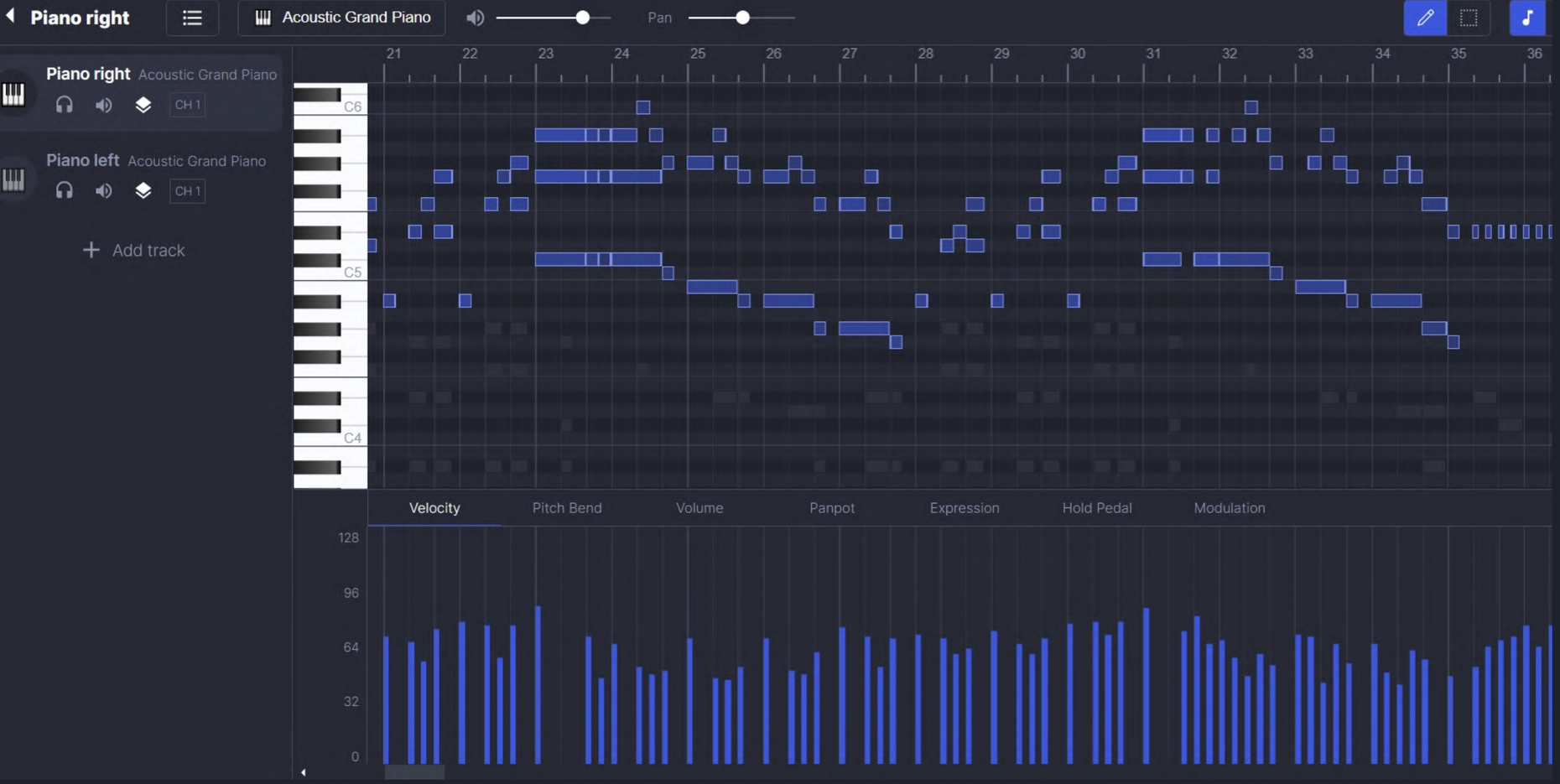
Midi Files

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.



Try Pitch



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.

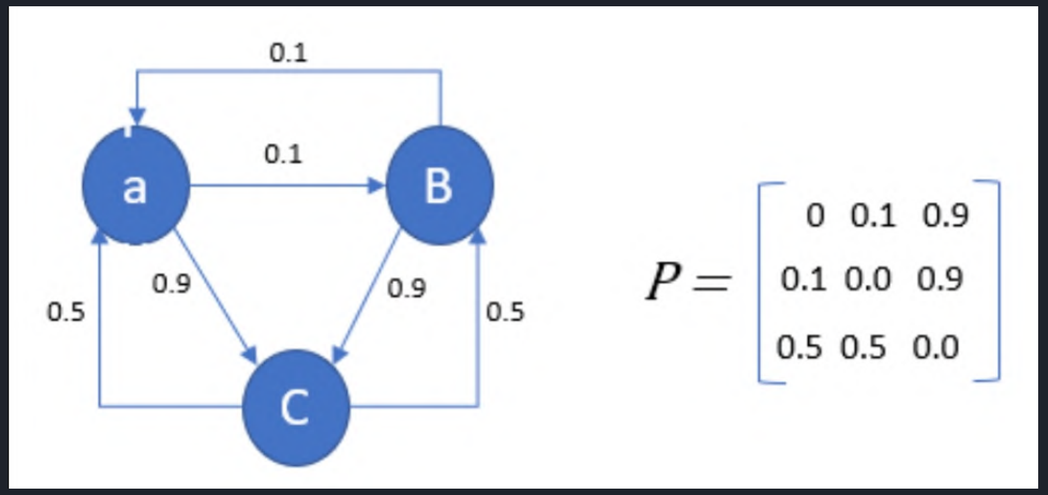
Try Pitch

Markov Chains

VERY BASIC MODEL

Markov Chain

- Discrete-time stochastic process.
- Follows Markovian Property



$$P(Yn|X1, X2, \ldots Xn\text{-}1) = P(Yn \mid Xn\text{-}1)$$

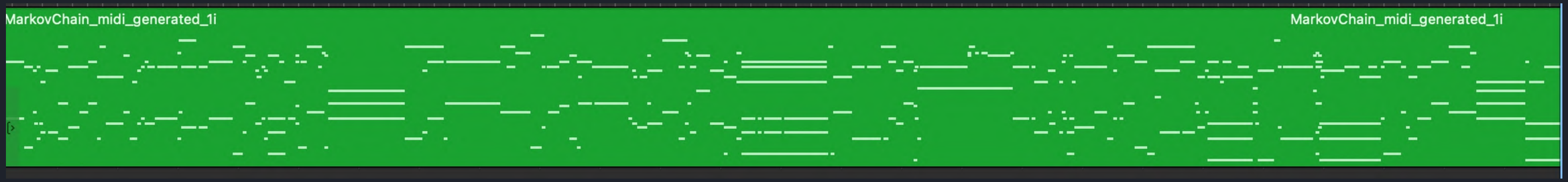
What we did?

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.

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

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