THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Music Generation with Deep Neural Networks

for the course

DATS 6303 ‘Deep Learning’
spring 2024

Team 7
Aman Jaglan
Aneri Patel
Smit Pancholi
Tushar Sharma

Introduction

Employing computational music and time series prediction approach for extrapolating monophonic melodies using deep neural networks.

Data: ESAC folksong database consisting of 5000+ melodies

Deep neural networks under consideration:

- Gated Recurrent unit (GRU)
- Variational Auto-encoder
- LSTM with Multi-Head Attention

Tools and softwares:

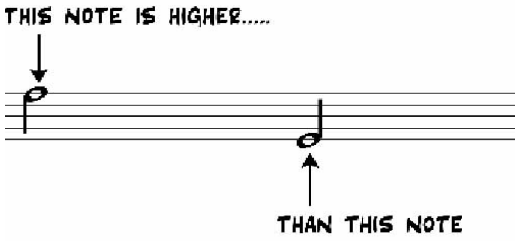
- Music21 (computational musicology)
- PyTorch (training models)
- MuseScore (rendering music scores)
- Streamlit (web application)

Source: ESAC Folk Song Dataset

Music theory fundamentals

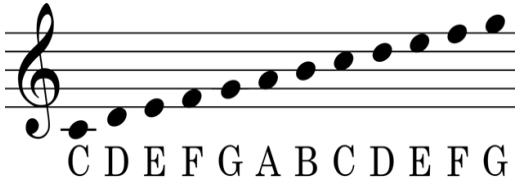
Pitch

Position of a note in the frequency scale.



Key

The tonic of a melody/song on which it seems to resolve.



Scale/Mode

Collection of notes in specific intervals (frequency difference).

Tempo

Speed of the beat (measured in ‘beats per minute’).

Time signature

The amount and type of notes each measure/bar contains. E.g. – 3/4, 4/4, 9/8

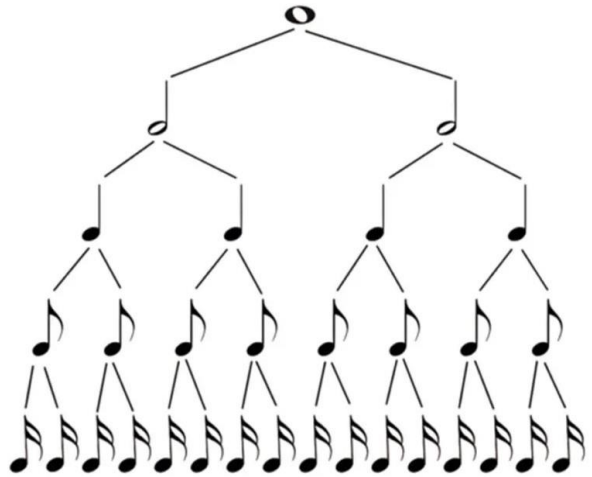
Whole note (1 measure)

Half note (1/2 measure)

Quarter note (1/4th measure)

8th note (1/8th measure)

16th note (1/16th measure)

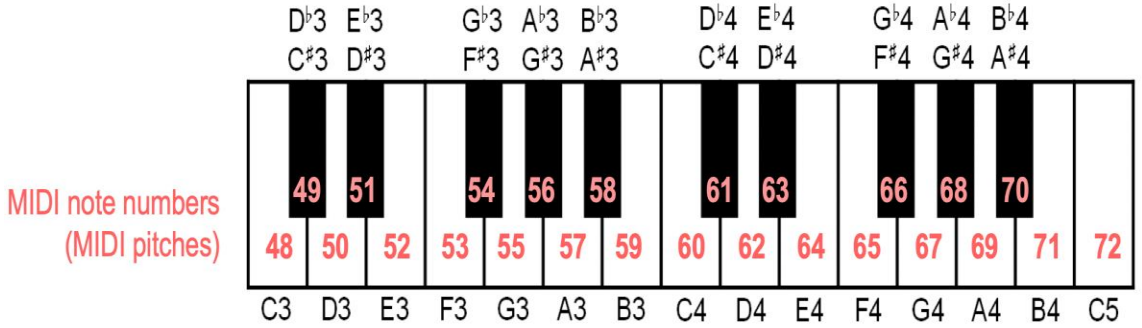


The division of beats in western music.

Symbolic music representation

Musical Instrument Digital Interface (MIDI):

- Protocol to allow electronic music instruments and computers to communicate with each other.
- MIDI note numbers range: **0** to **127**
  Eg - here, middle C4 is the number 60 (not a fixed convention).

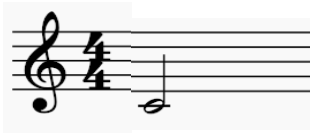


Symbols

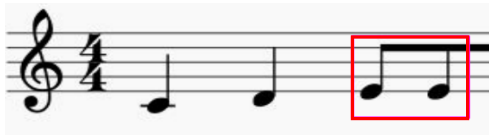
- Each pulse/time-step indicates a **16th note** (quarter of a measure)
- **'_'** indicates ***sustain***
- **'r'** indicates ***rest***

Translation to time series representation

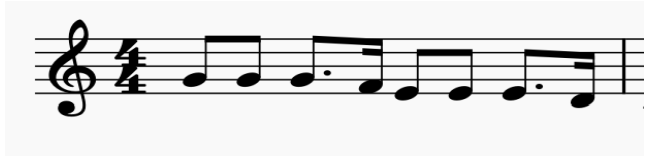
[60 , _ , _ , _]



[60 , _ , _ , _ , 62 , _ , _ , _ 64 , _ 64 , _]



[67 , _ , 67 , _ , 67 , _ , _ , 65 , 64 , 64 , 64]



Sequence encoding

Score

- Symbolic sequences are mapped to integers for input to models
- Sliding window approach: input: [10, 11, 12, …] target: [13]
  input: [11, 12, 13, …] target: [14]
  … …

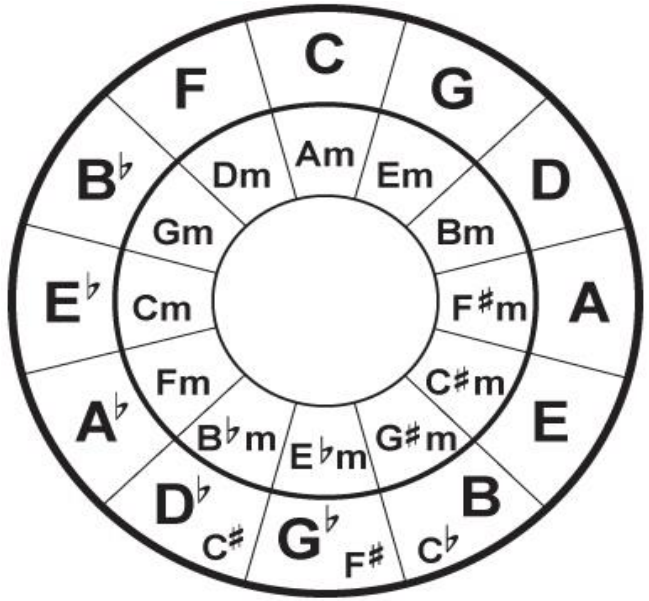
Transposition

Changing the key of a melody but maintaining the relative intervals between the notes.

Keys retained**:** **12** major and **12** minor

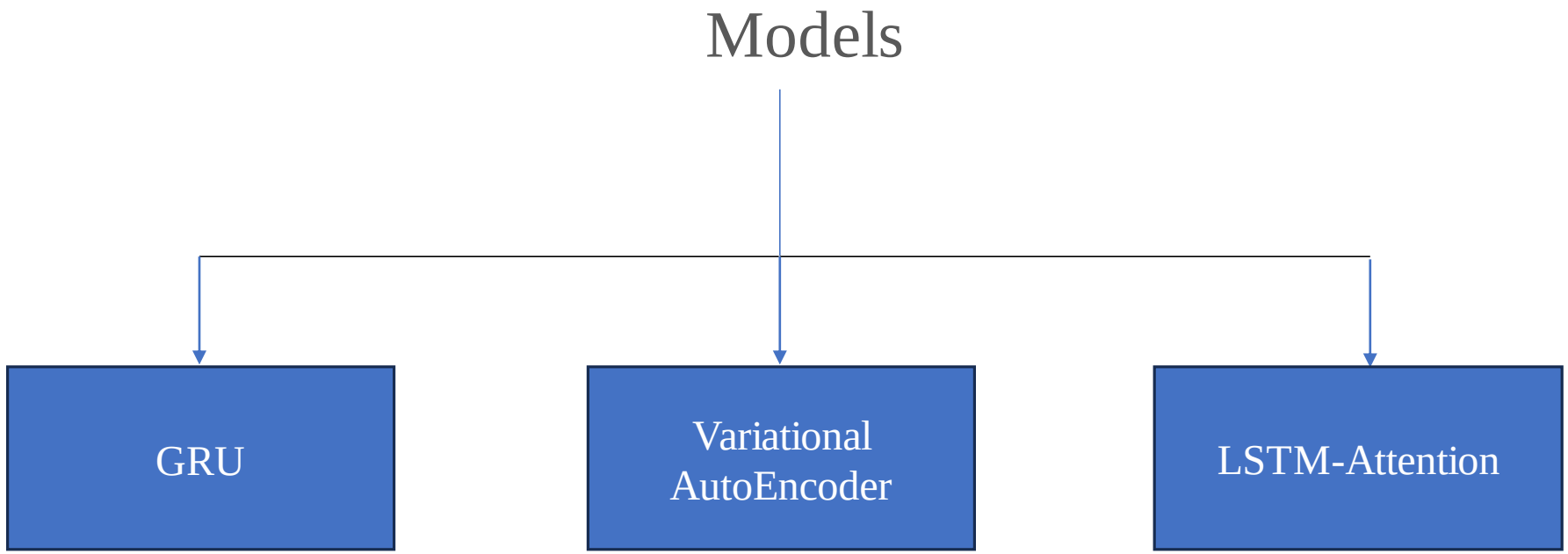
Each melody is transposed to all the **24** keys (equivalent of data augmentation!)

Final count of data samples (melodies) : 40,800

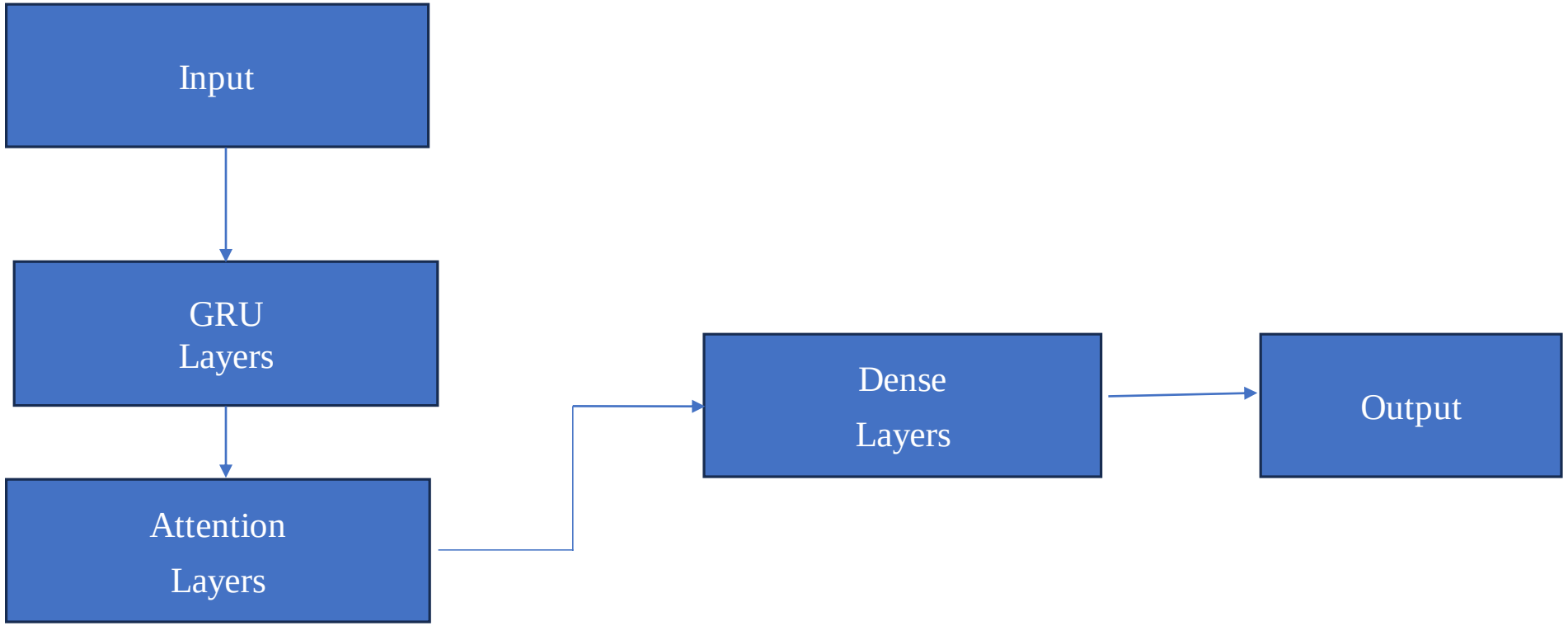


The 24 keys represented as the circle-of-fifths

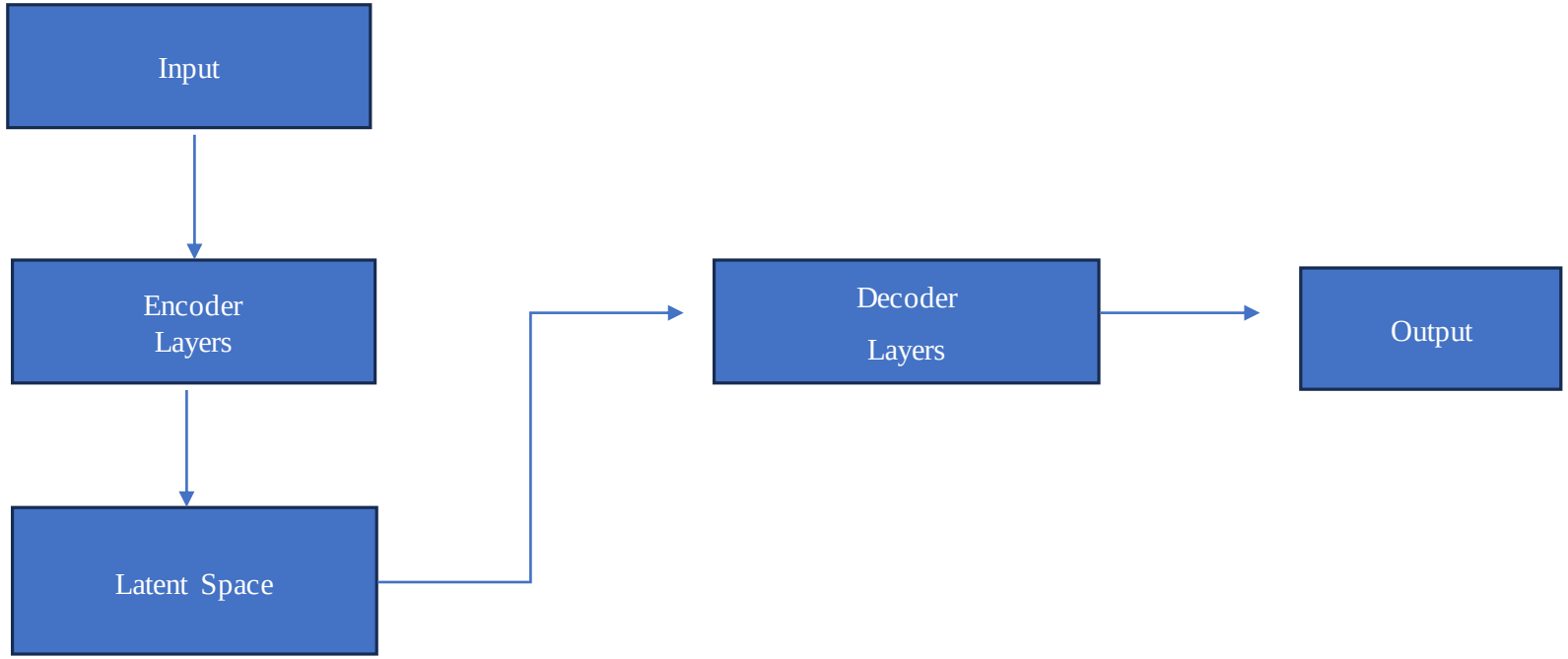
Model Selection



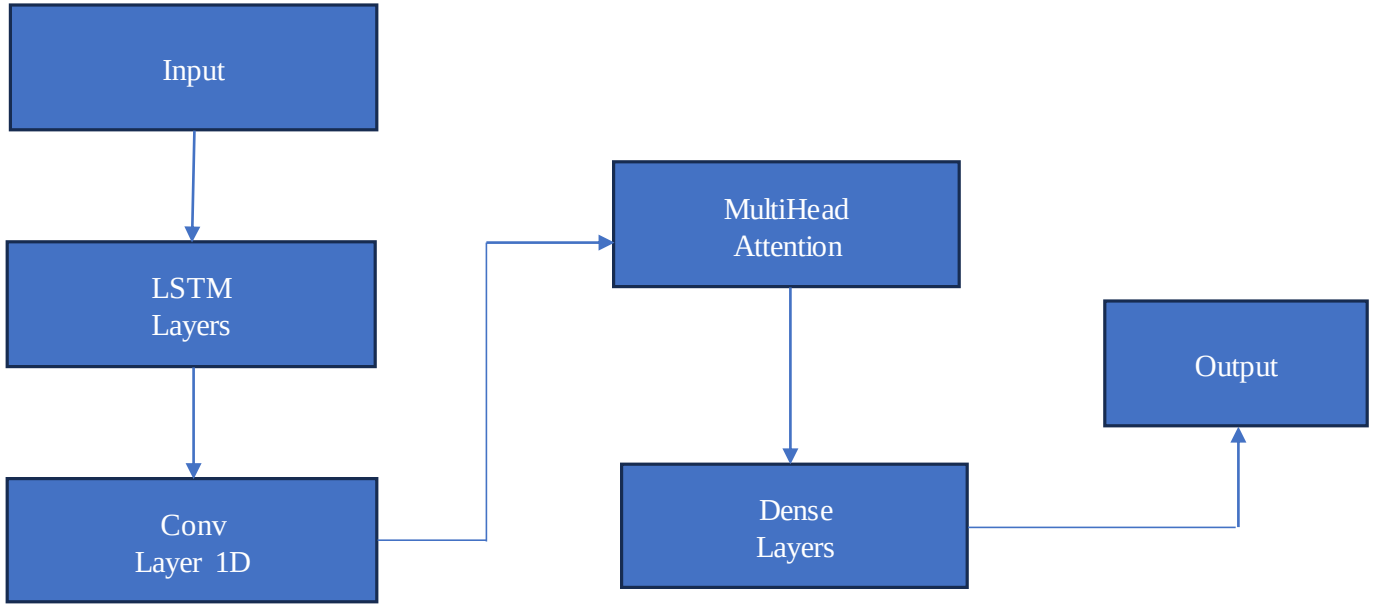
GRU



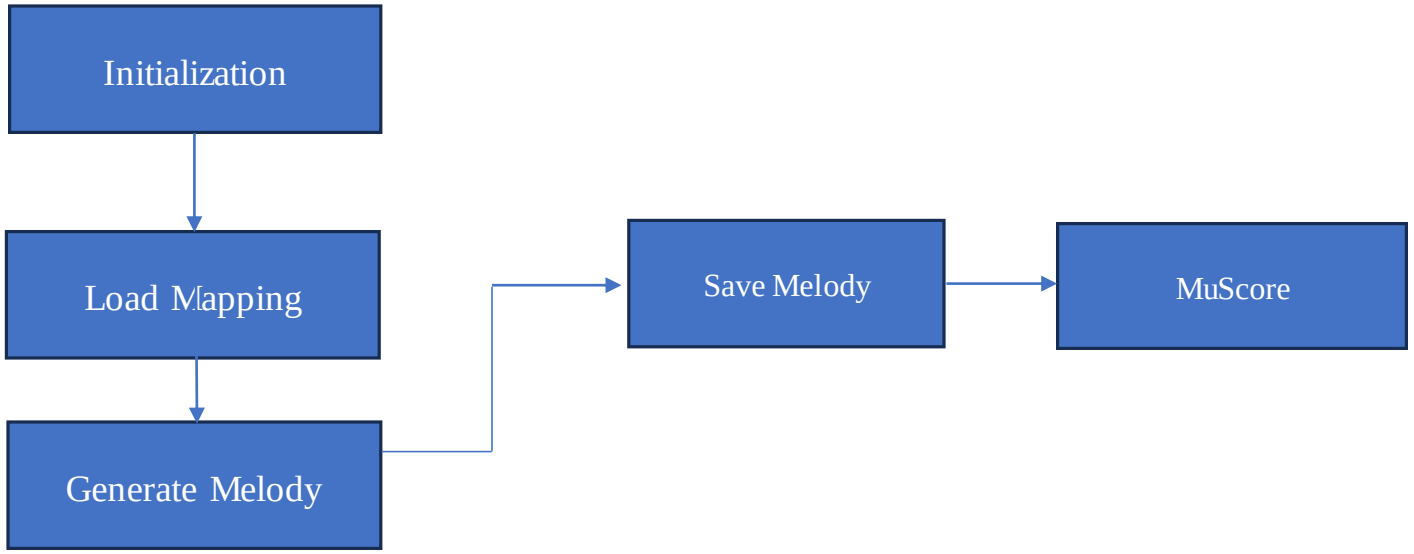
Variational-Auto Encoder



LSTM- With Multi head attention

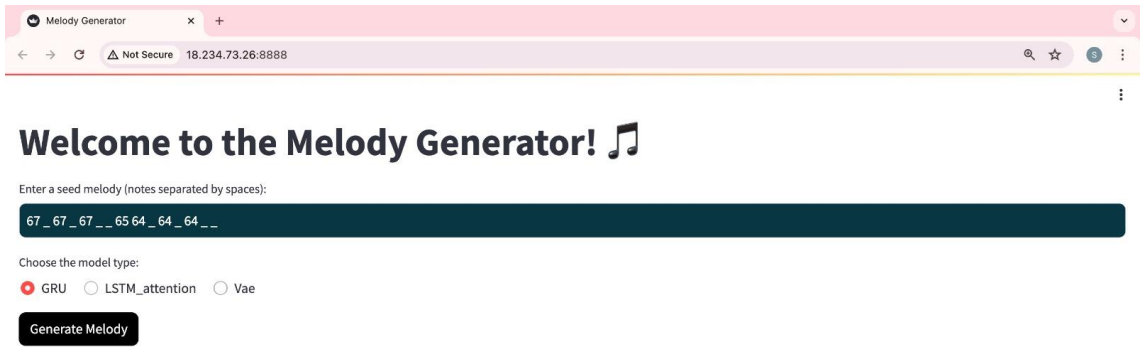


Melody Generator



Streamlit


Streamlit


Melody Generator
Not Secure
18.234.73.26:8888
Welcome to the Melody Generator! 🎵
Enter a seed melody (notes separated by spaces):
67 _ 67 _ 67 _ _ 65 64 _ 64 _ 64 _ _
Choose the model type:
GRU
LSTM_attention
Vae
Generate Melody

Thank You!