SWE HW 5

Sakka Mohamad-Mario 1241EB

1. Dataset

I used a public dataset of lo-fi hip hop MIDI files from Kaggle. The dataset contains multiple piano-style MIDI loops that fit the lo-fi hip hop sound. I downloaded it directly in my environment and pointed my code to the folder that contains the .mid files.

Dataset link: https://www.kaggle.com/datasets/zakarii/lofi-hip-hop-midi/data

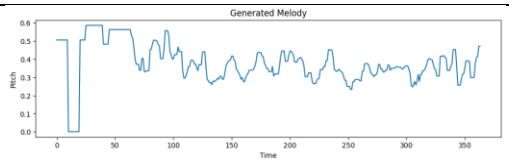
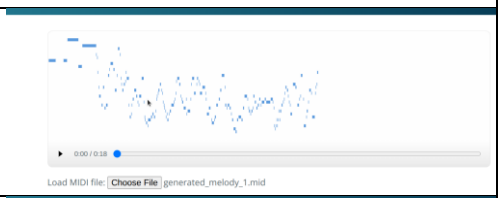
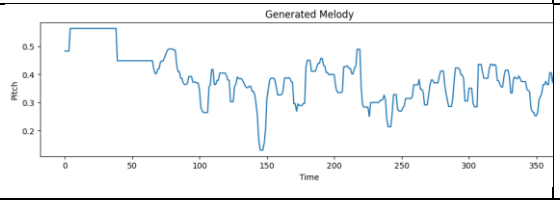
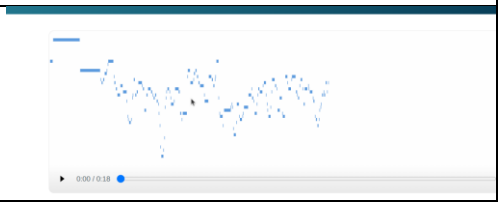
2. Generating New Melodies

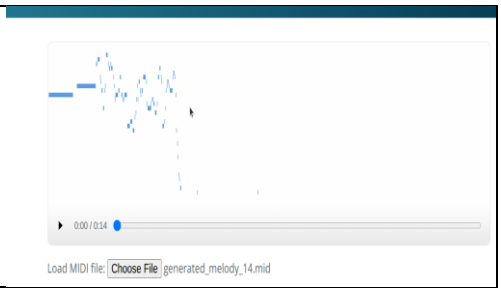
After training the model, I used some of the original songs as seeds to generate new melodies. For each one, I take the last window_size steps, feed them into the model to predict the next value, add that prediction to the sequence, and repeat this many times to build a complete generated melody.

To save each generated melody, I convert the normalized values back into MIDI pitch values and write them to a new MIDI file using pretty_midi. I reuse the first instrument track from each MIDI file instead of choosing specific instruments, because I couldn't easily set or manage different instrument sounds for each track. I generated several songs this way, but in my report I only show the results for songs 1, 2, and 14.

3. Results and Screenshots

In Google Colab, I plotted each generated melody as a line chart to show how the pitch changes over time. I also saved several generated MIDI files using different seed songs. To listen to the results and inspect them, I used this online MIDI player: https://cifkao.github.io/html-midi-player/. Below are some screenshots from my Colab notebook that show a few of the generated melodies.

Melody	Google Collab Screenshot	MIDI Player
1		
2		

3		

4. Results Interpretation

When I look at the generated melody plot, I can see that the model is not just producing random noise. The pitch values stay in a certain range and move smoothly most of the time, which suggests that the LSTM learned some typical patterns from the lo-fi MIDI dataset. For example, in the beginning of the plot the melody stays around a stable level, then it moves into a more active region with small up-and-down changes, and later it gradually drops closer to zero, which I can interpret as the phrase “calming down” or ending.

When I listen to the generated MIDI files using the online MIDI player, I notice that some parts sound like simple lo-fi piano phrases: there is repetition, small variations, and the pitches stay in a comfortable mid-range. At the same time, there are also some weaknesses. Sometimes the model holds the same note for too long, or there are long sections that are almost silent with only a few notes. This makes sense, because the model only sees one note at a time (no chords, rhythm, or velocity changes), and it tries to predict the next pitch based only on the previous 64 steps.

Overall, I would say the results show that the model has learned a basic sense of melodic continuity and pitch range typical for the dataset, but it does not yet capture full musical structure like chord progressions, rhythm, or more complex lo-fi patterns. The generated melodies are a good first step, but they would need more advanced modeling (for example, multiple tracks, rhythm, and dynamics) to sound closer to real lo-fi hip hop tracks.