Mid-Evaluation Report for Music Generation Model Project

1. Summary of Progress

So far, I've focused on setting up essential components and initial tests to create a music generation model. Here's what I've accomplished:

• Data Collection and Preprocessing: I sourced MIDI files and other symbolic data for melody generation. I also started with vocal datasets that have emotion and lyric labels. Managing data on Kaggle's limited storage has been challenging, so I've focused on selectively sampling data without sacrificing quality.

• Model Training and Initial Output: I configured a custom GPT-2 model to test its ability to generate melodies using Kaggle's T4 GPU and bf16 precision for efficiency. The model produces simple musical patterns, but the output still lacks complexity and coherence. Improving this is a key focus for the next phase.

2. Self-Assessment on Mid-Evaluation Goals

Goal	Status	Comments
Basic Transformer model for melody	Achieved	Set up and tested a GPT-2 model on symbolic melody data
Prepare datasets for training	Achieved	Cleaned and processed MIDI and vocal datasets
Generate initial music output	In Progress	Model produces simple patterns, but needs more musical quality
Preliminary Literature Review	Partial	Conducted basic review; need deeper insights for the next phase

3. Unanticipated Challenges

Some unexpected issues have come up that I'll need to work through:

• Output Simplicity: The model is generating music, but it's very basic. Adjusting model configurations to increase the quality and complexity has proven tricky and will need more experimentation.

• Data Constraints on Kaggle: Handling large datasets like MAESTRO has been difficult with Kaggle's limited storage. I've had to make trade-offs, which could impact the model's learning.

• Preprocessing Audio for Consistency: Preparing MIDI and vocal data for alignment took

more time than expected, especially for synchronizing melody and lyrics. This will need further attention to get the singing AI results I'm aiming for.

4. Plan of Action for the Second Phase (Starting December 1)

To meet the final deadline on December 12, my plan for the second phase includes:

• Improving Model Complexity and Output Quality: I'll continue adjusting model parameters to create more sophisticated music. This involves testing different model depths and sequence lengths to generate richer and longer pieces.

• Expanded Literature Review: To deepen my understanding of how to improve music generation, I'll conduct a more thorough review of research on melody generation and transformer models, especially Music Transformer and Jukebox. I expect this will provide ideas for model tweaks that enhance musical complexity.

• Enhanced Data Management on Kaggle: I'll look into ways to handle larger datasets more effectively on Kaggle, potentially through dynamic loading or optimized sampling. This will help the model learn from a greater variety of data without hitting storage limits.

• Synchronizing Melody and Vocals: I'll experiment with methods for better aligning melody and lyrics, gradually refining this through weekly tests. I plan to produce and review samples weekly, making adjustments based on quality.

This approach gives me a structured path to refine the model's output and improve its musicality by the final deadline on December 12.

Mid-Evaluation Report for Music Generation Model Project

Course of Action for the Next Phase (starting December 1)

1. Enhancing Model Complexity and Quality: Experiment with tuning model parameters, including depth and sequence length, to achieve more nuanced musical output.

2. Deepening Literature Review: Extend the literature review, especially on music generation and transformer-based models. This will help refine the approach for creating a cohesive and rich musical composition.

3. Synchronizing Melody and Vocals: Test alignment techniques for melody and lyrics and review the weekly samples for quality adjustments, aiming to meet the final December 12 deadline.