Name: David Schwartzman

Date: 12/06

Description: Report on Bach-Inspired Music Generation Model

Objective

The goal of this project was to train a machine learning model on a dataset consisting of Bach's compositions to generate original music pieces that resemble the style and complexity of Bach.

Training Process

- **Hardware**: Training was conducted on a GPU for efficient handling of sequential data and took me around 35 minutes to complete training.
- **Loss Function**: Sparse Categorical Crossentropy, as the task involved predicting the next token (note or chord) in a sequence.
- **Validation Metrics**:
    - Perplexity: Final value around 1.8, indicating reasonable confidence in note predictions.
    - Accuracy: Stabilized at 85% on validation data.

Generated Output

Strengths:

1. **Melodic Similarity**: The generated sequences is similar to bach's music.
2. **Polyphony**: The model successfully created multi-voice compositions with recognizable relationships between soprano, alto, tenor, and bass voices.

Weaknesses:

1. **Predictability**: While the music reflects Bach's style, it lacks the complexity and unpredictability characteristic of his compositions.

2. **Structure**: Larger structures such as fugues or variations were not convincingly reproduced, with the generated pieces feeling disjointed beyond a few measures.
3. **Rhythmical Variations**: Rhythmic patterns were repetitive, failing to capture Bach's nuanced syncopations and ornaments.

Analysis

The generated output shows promise but highlights areas for improvement:

- **Lack of Long-Term Structure**: The model has trouble handling complex musical forms, like fugues or chorales with variations, that require organizing and connecting large sections of music.
- **Subtle Nuances**: stylistic subtleties common in Bach's work are not yet effectively captured.

Recommendations

To improve the model's performance and better mimic Bach's style:

- Introduce positional embeddings specific to musical context (e.g., bar numbers, key changes).
- Expand the dataset with similar Baroque composers (e.g., Handel, Telemann) to improve generalization.
- I trained the model only on 50 epochs, more is probably needed for good results.

Conclusion

The model demonstrates the potential to emulate Bach's style but requires refinement to achieve a closer resemblance to his compositions. Future work should focus on improving structural coherence, capturing subtle nuances, and incorporating music-theory-driven constraints to elevate the generated pieces.

Extra Credit #1

Quantitative Metrics:

There are several quantitative metrics that can be used to evaluate a music model's training performance. One key metric is perplexity, which measures how well the model predicts the next note or token in a sequence. A lower perplexity value suggests that the model is better at capturing the underlying structure and patterns within the data, indicating effective learning. Accuracy is another important metric, as it evaluates the proportion of correctly predicted notes or events during the model's training and validation phases. This metric is especially useful when working with symbolic representations such as MIDI, where the goal is to predict discrete musical events. Another essential metric is the loss value, which indicates how far the model's predictions are from the actual data distribution. Lower loss values typically indicate a better-performing model, as it shows that the predictions are becoming more aligned with the true sequences. Entropy is a metric that helps assess the diversity of the generated sequences. A high entropy value suggests the model is producing diverse outputs, which can encourage creativity but also increase the risk of incoherence. On the other hand, lower entropy can lead to more repetitive or predictable results. Finally, statistical analysis of musical features can be employed to compare key characteristics of the generated music, such as note durations, pitch intervals, and polyphony levels, with the data from the original Bach compositions. This approach provides an objective measure of how closely the model's output mirrors the characteristics of Bach's style.

Extra Credit #2

Musical Quality

Determining whether the generated music resembles Bach's Cello Suites requires both objective and subjective evaluations. Objectively, music theory-based metrics such as harmonic analysis and adherence to counterpoint rules can be extremely useful. For instance, Bach's music is known for its complex but coherent harmonic

progressions, and any generated music should be analyzed for how well it maintains tonal harmony and resolves dissonances. In addition, adherence to counterpoint principles, such as avoiding parallel fifths and octaves, and ensuring that voices maintain their independence, is crucial. Structural consistency is also important; the generated music should exhibit familiar patterns, such as flowing, prelude-like movements or dance forms typical of the Cello Suites. Another way to evaluate musical quality is through feature similarity metrics, which compare aspects of the generated music—such as pitch interval distribution, rhythmic patterns, and phrase length—to Bach's original work. For example, Bach's Cello Suites tend to use a particular set of intervals, and the rhythm often has a distinct flow that may be missing in a less accurate model. The length of musical phrases is another important feature, as Bach often uses specific phrase lengths that contribute to the structural coherence of the piece. Key modulations, or changes in the tonal center, also play a significant role in Bach's work, and the model's ability to replicate these shifts can serve as a measure of its success in emulating his style.

While these quantitative and theoretical measures are useful, subjective methods are equally important. Expert evaluation from musicologists or performers familiar with Bach's Cello Suites can provide valuable insight into the model's ability to capture the stylistic nuances of the original compositions. These experts can assess whether the generated pieces follow Bach's characteristic voice leading, harmonic richness, and thematic development. In addition, human perception studies can be conducted, where listeners rate the generated music based on parameters like coherence, aesthetic appeal, and overall resemblance to Bach's style. These studies provide a more intuitive sense of how well the generated music resonates with human audiences. Another subjective evaluation method involves performance tests, where cellists perform the generated pieces. This can reveal how playability and interpretive depth compare to Bach's original works, providing a practical measure of musical quality.