**Discussion Points for the Report**

**1. Summary of Findings**

* **Objective Accomplished:** Successfully developed and trained a Transformer-based model for music generation in the style of Bach’s Cello Suites.
* **Dataset Preparation:** The dataset was preprocessed to extract musical notes and durations from the provided MIDI files. These were tokenized to create a structured input for the Transformer model.
* **Model Training:**
  + The Transformer model effectively learned relationships between notes and durations, enabling sequential generation.
  + Metrics such as loss and accuracy indicate a steady improvement during training.
* **Music Generation:**
  + The model generates sequences that adhere to tonal harmony and melodic structure reminiscent of Bach’s compositions.
  + Outputs include stylistically consistent patterns, although occasional anomalies (e.g., unusual note transitions) highlight areas for improvement.

**2. Reflection on the Model’s Ability to Generate Music in the Style of Bach**

* **Strengths:**
  + The model captures key stylistic elements of Bach’s Cello Suites, such as harmonic progressions and phrasing.
  + It mimics Bach’s counterpoint to a reasonable extent, demonstrating the potential of Transformer architectures for musical imitation.
  + The generated music exhibits coherence in terms of rhythm and note sequencing.
* **Limitations:**
  + Long-term structure (e.g., overarching musical themes) is sometimes inconsistent, reflecting challenges in modeling global dependencies.
  + The model occasionally introduces redundant or dissonant notes that deviate from Bach’s strict harmonic style.

**Extra Credit Research Problems**

**3. Quantitative Metrics**

* Evaluating the training and performance of a music generation model can include:
  + **Perplexity:** Measures the uncertainty of the model’s predictions. Lower perplexity indicates better learning of musical patterns.
  + **Pitch Class Entropy:** Assesses the diversity of pitch classes in the generated music. Too low or too high values may indicate issues.
  + **Rhythmic Consistency:** Measures how well the generated durations align with common rhythmic patterns in the training data.
  + **Chord Progression Similarity:** Compares the harmonic sequences in generated pieces with those in Bach’s compositions.

**4. Musical Quality**

* **Subjective Evaluation:**
  + Listening tests can be conducted where musicians or listeners rate the resemblance to Bach’s style.
  + Metrics include perceived complexity, harmony, and emotional impact.
* **Objective Analysis:**
  + Use statistical comparisons with Bach’s compositions for note distributions, interval transitions, and rhythmic patterns.
  + Evaluate adherence to common compositional rules (e.g., voice leading, counterpoint).

**5. Beyond Bach**

* **Adapting to Other Composers:**
  + A similar pipeline can be applied to other datasets, such as:
    - Beethoven’s string quartets for intricate harmonic progressions.
    - Mozart’s piano sonatas for classical elegance.
    - Jazz improvisations for more complex rhythms and harmonies.
  + This would require:
    - Preprocessing MIDI files for the chosen composer.
    - Adapting tokenization to account for stylistic differences (e.g., syncopation in Jazz).
* **Example Use Case:**
  + Train a model on Beethoven’s works to generate piano sonatas that feature his dramatic and dynamic range.