PROJECT REPORT

Approach and Trade-offs

The project adopts a two-pronged approach: melody generation using predefined musical templates and genre classification leveraging machine learning. The melody generation relies on the music21 library, defining musical scales, tempo, and rhythmic patterns for different genres (Afrobeat, Fuji, Jazz). The genre classification uses the GTZAN dataset from Kaggle, which contains labeled audio clips across multiple genres.

Trade-offs include:

- **Rule-based Melody Generation:** While efficient, it lacks the flexibility of deep generative models, limiting the variety of compositions.
- **Predefined Genre Features:** Relying on fixed musical scales and patterns reduces generalization to sub-genres or experimental music.
- **GTZAN Dataset Dependence:** While widely used, the GTZAN dataset has been criticized for possible dataset contamination and limited diversity.

Challenges and Solutions

1. **Melody Generation Complexity**
    - Challenge: Generating diverse and realistic melodies with simple rule-based methods.
    - Solution: Introduced syncopation and rhythmic variations to make melodies more dynamic.
2. **Dataset Storage and Access**
    - Challenge: Limited local storage for large datasets.
    - Solution: Integrated Kaggle API to directly fetch data, avoiding manual downloads.
3. **Feature Extraction for Classification**
    - Challenge: Extracting meaningful audio features for classification.
    - Solution: Leveraged feature extraction techniques such as MFCCs (Mel-frequency cepstral coefficients) to improve genre detection accuracy.
4. **Model Performance and Overfitting**
    - Challenge: Overfitting due to the relatively small dataset.
    - Solution: Implemented data augmentation and dropout layers to improve model generalization.

Potential Improvements

- **Deep Generative Models:** Implementing RNNs, Transformers, or VAEs for more sophisticated melody generation.

- **Hybrid Classification Model:** Combining spectrogram-based CNNs with text-based NLP models for genre tagging.
- **Larger and More Diverse Datasets:** Incorporating datasets beyond GTZAN, such as FMA (Free Music Archive), to enhance model robustness.
- **Model Deployment:** Hosting the classifier as a web app for real-time music genre classification and generation.