Project plan for music generation from MIDI files

# Project goal

Develop a music generation system using MIDI files, which includes creating a specialized dataset and algorithms for music segmentation, pattern recognition, and generating new musical compositions.

# Step 1: Research and analysis of existing resources

1. **Study existing MIDI datasets and tools:**

• **AAM Dataset**: Explore 3000 artificial music tracks with rich annotations for various music information retrieval (MIR) tasks.

• **ComMU Dataset**: Utilize 11,144 MIDI samples created by professional composers with 12 metadata fields for generating diverse music.

• **ArtificialSongGenerator**: Tool for automatically generating MIDI files and audio tracks using algorithmic composition.

• **FMP Notebooks**: Educational materials and code examples for various music information retrieval tasks.

# Step 2: Creating and annotating the dataset

1. **Data preparation:**

• Collect and analyze existing MIDI files from various sources (AAM, ComMU).

• Use ArtificialSongGenerator to create new MIDI files and corresponding audio tracks.

2. **Data annotation:**

• Develop annotation standards for key musical features: onsets, pitches, instruments, keys, tempos, segments, melodies, beats, chords.

• Automatically and manually annotate the collected data using available tools and algorithms.

# Step 3: Developing and testing algorithms

1. **Music segmentation:**

• Develop algorithms for segmenting music into manageable parts (chorus, verse, etc.).

• Test and optimize segmentation algorithms using data from AAM and ComMU.

2. **Pattern recognition:**

• Use machine learning methods (RNN, VAE, transformers) for recognizing musical patterns.

• Evaluate and improve algorithm performance based on obtained data.

3. **Music generation:**

• Develop and test algorithms for generating new musical compositions.

• Use various approaches (algorithmic composition, deep learning) to create high-quality and diverse music.

# Step 4: Optimization and enhancement

1. **Scaling and computational resources:**

• Gradually scale up tests and use external computational resources as needed.

• Use Google Colab?

2. **Automation and verification:**

• Develop methods for automatic evaluation and manual verification to improve data quality.

• Include human-in-the-loop reinforcement learning methods for manual result verification.

# Step 5: Publications and further research

1. **Publications:**

• Prepare scientific papers and publications based on the created dataset and developed algorithms.

2. **Collaborations and media applications:**

• Explore potential applications in media and potential collaborations.

• Develop plans for further research and improvement of the music generation system.

A diagram of a data flow

Description automatically generated with medium confidence

MINDMAP - Music Generation System Project Roadmap.html

A screen shot of a computer

Description automatically generated

MINDMAP - Music Production Software and AI Tools.html