Report: audio ai - guitar clustering and melody transcription

Subtask 1: Guitar Audio Transcription

Objective: Convert guitar audio into a sequence of musical notes, identifying the start and end times for each note played.

Approach:

1. Audio Loading: Utilized librosa.load to import audio files, ensuring compatibility with various audio formats and handling exceptions for file access issues.
2. Preprocessing:

– Noise Reduction: Applied using the noisereduce library to minimize background noise and enhance signal clarity.

– Normalization and Pre-emphasis: Boosted the signal's high-frequency components to balance the audio spectrum and improve note distinction.

3. Feature Extraction:

– Harmonic Separation: Separated harmonic elements from percussive sounds to focus on tonal aspects relevant to pitch identification.

– Pitch Tracking: Employed librosa.piptrack, which estimates pitches from the harmonic component, crucial for note identification.

4. Note Event Detection:

– Onset Detection: Detected the start of musical notes using librosa.onset.onset_detect, which identifies significant increases in signal energy.

– Offset Calculation: Estimated note durations by detecting energy drops or the next note's onset.

5. Note Classification:

– Pitch Classification: Converted frequency information into musical note names.

– Duration Handling: Ensured that notes are uniquely identified and temporally bounded without overlaps.

Results: The transcription process outputs a list of tuples, each representing a note with its name and the corresponding start and end times.

```
PS C:\Users\User\PycharmProjects\audio_Proj> python audio_task_1.py "Audio_Guitar/1.wav"
[('B7', 0.035, 0.43), ('D4', 0.43, 0.871), ('A3', 0.871, 1.312), ('C#7', 1.312, 1.544), ('A5', 1.544, 1.985
), ('E5', 1.985, 2.496), ('G#6', 2.496, 2.612), ('C#4', 2.612, 2.856), ('A5', 2.856, 3.506), ('F7', 3.506,
4.319), ('G3', 4.319, 4.981), ('F5', 4.981, 6.316), ('D6', 6.316, 6.548), ('C6', 6.548, 6.908), ('A6', 6.90
8.638, 8.882), ('G6', 8.882, 8.916), ('D#7', 8.916, 9.601), ('F#5', 9.601, 10.043), ('G#5', 10.043, 10.228)
, ('C#7', 10.228, 10.461), ('C6', 10.461, 10.879), ('D#6', 10.879, 11.958), ('A#4', 11.958, 13.015), ('A#6'
, 13.015, 13.491), ('C6', 13.491, 13.723), ('F#5', 13.723, 13.886), ('G#7', 13.886, 14.35), ('D#7', 14.35,
14.791), ('G6', 14.791, 15.43), ('A6', 15.43, 15.859), ('C7', 15.859, 16.742), ('F#5', 16.742, 16.962), ('G
5', 16.962, 17.171), ('G#3', 17.171, 17.392), ('F7', 17.392, 18.286), ('D#7', 18.286, 18.936), ('A#4', 18.9
36, 19.992), ('D#7', 19.992, 20.236), ('D7', 20.236, 20.457), ('C6', 20.457, 20.84), ('E7', 20.84, 20.875),
 ('A6', 20.875, 21.548), ('B7', 21.548, 21.745), ('E6', 21.745, 22.338), ('A6', 22.338, 22.639), ('B7', 22.
639, 22.779), ('D#7', 22.779, 23.475), ('D7', 23.475, 23.963), ('G7', 23.963, 24.184), ('C#7', 24.184, 24.3
81), ('C6', 24.381, 24.822), ('D#6', 24.822, 25.263), ('A#6', 25.263, 25.867), ('D7', 25.867, 26.912), ('C7
 28.247)]
```

Subtask 2: Guitar Audio Clustering

Objective: Cluster audio files into groups based on the specific guitar used for their recording.

Approach:

1. **Feature Extraction**:

– Used **librosa** to extract MFCCs (Mel Frequency Cepstral Coefficients) which capture the timbral aspects of the audio signals.

– Calculated the mean of these coefficients across time to create a feature vector for each audio file.

2. **Preprocessing**:

– Standardized features using **StandardScaler** to nullify scale differences and improve clustering performance.

– Applied PCA (Principal Component Analysis) to reduce dimensionality and focus on the most informative aspects of the data.

3. **Clustering**:

– Implemented **KMeans** clustering to partition the audio files based on their timbral characteristics, assuming two distinct guitars.

– Configured the algorithm with a fixed random state to ensure reproducibility of the results.

Results: The clustering process outputs a list of tuples, where each tuple contains filenames of audio files that are predicted to be played by the same guitar.

```
PS C:\Users\User\PycharmProjects\audio_Proj> python audio_task_2.py "Audio_Guitar"
```

```
('1.wav', '3.wav')
('2.wav', '4.wav')
```