



# Musical Instrument Classification Data Book

By: Rohan Sikand



This is my data book for my project regarding musical instrument audio classification. View the Github repository [here](#) for more details and the code.

## Table of Contents

[Table of Contents](#)

[Data](#)

[Audio Files](#)

[Preprocessing](#)

[Waveforms](#)

[Acoustic Guitar](#)

[Bass Drum](#)

[Cello](#)

[Clarinet](#)

[Double Bass](#)

[Flute](#)

[Hi-hat](#)

[Saxophone](#)

[Snare Drum](#)

[Violin/Fiddle](#)

[Mel Spectrograms](#)

[Acoustic Guitar](#)

[Bass Drum](#)

[Cello](#)

[Clarinet](#)

Double Bass  
Flute  
Hi-hat  
Saxophone  
Snare Drum  
Violin/Fiddle

---

# Data

## Audio Files

The raw data files used in this experiment was obtained from [here](#). In total, there are ten classes:

```
['Acoustic_guitar', 'Bass_drum', 'Cello', 'Clarinet', 'Double_bass', 'Flute', 'Hi_hat', 'Saxophone', 'Snare_drum', 'Violin_or_fiddle']
```

Each class has 30 audio files. From here, a 90:10 ratio was used to split the data between training and testing. Thus, each class was split into 27 training audio files and 3 test audio files.

## Preprocessing

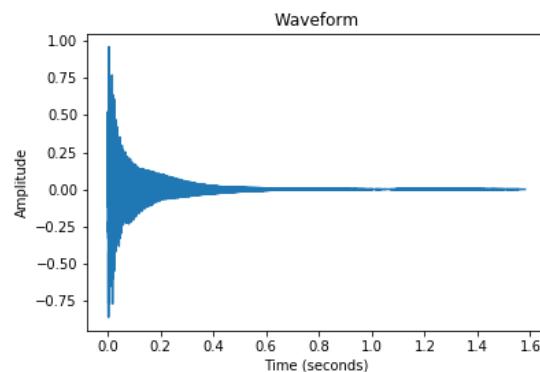
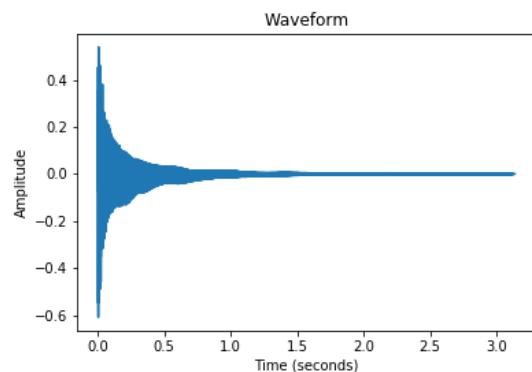
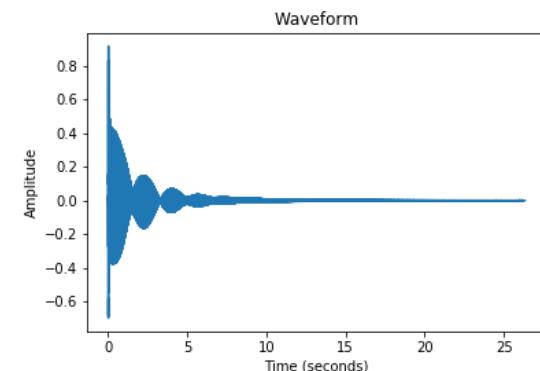
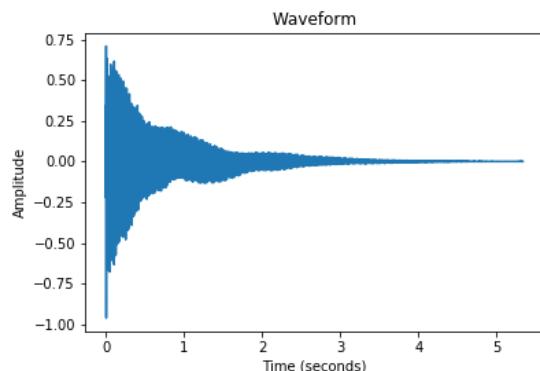
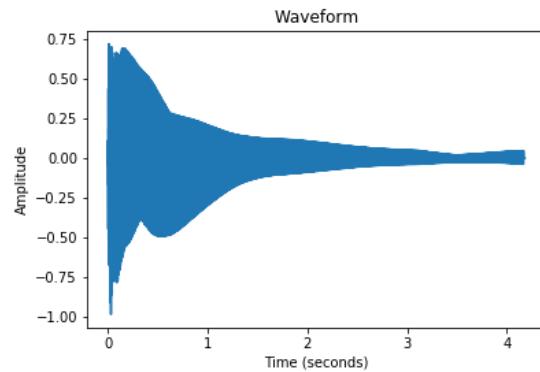
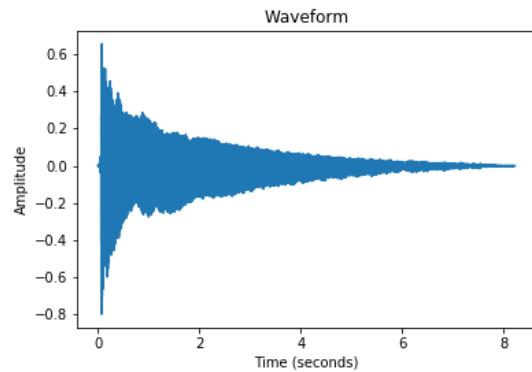
The overall preprocessing pipeline was to take the raw audio files and convert them into Mel Spectrograms which were then subsequently fed into the machine learning algorithms. The following shows the preprocessing pipeline:

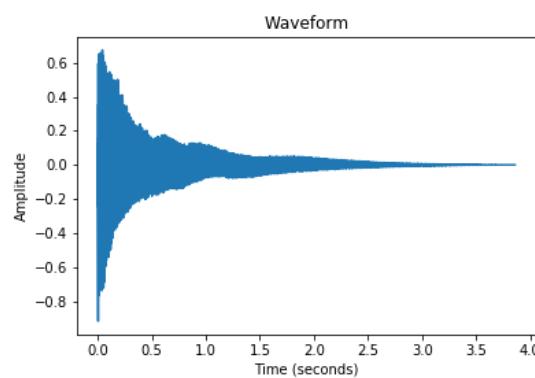
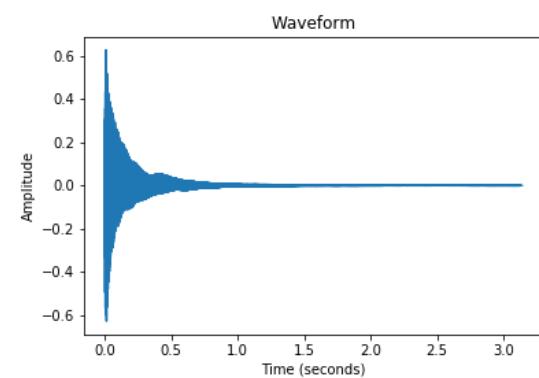
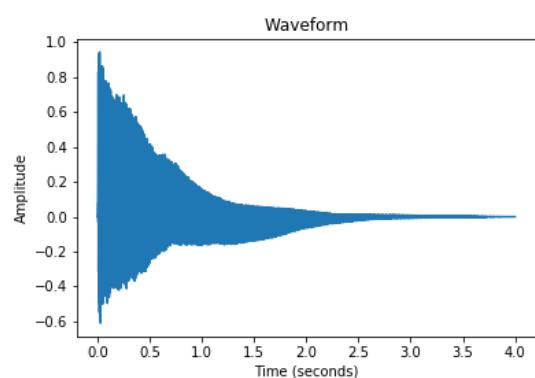
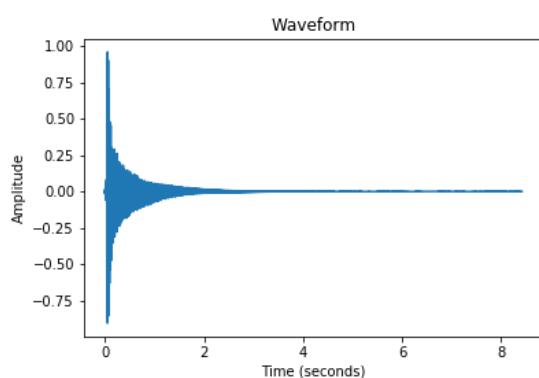
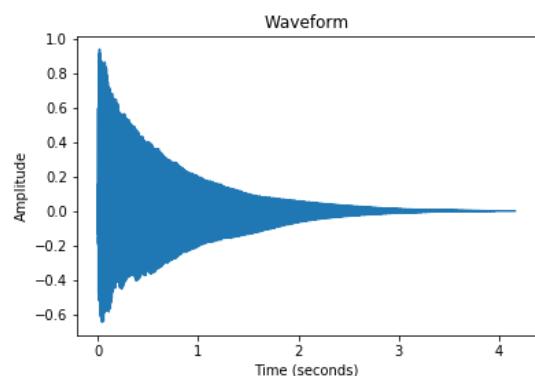
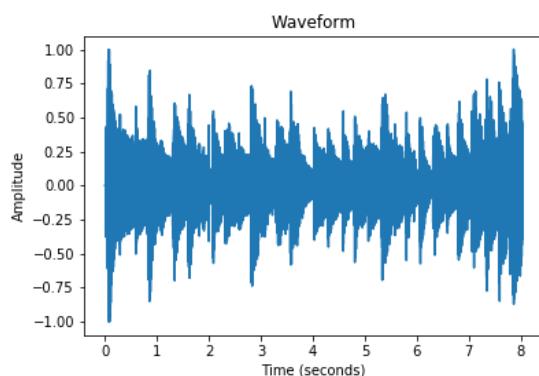
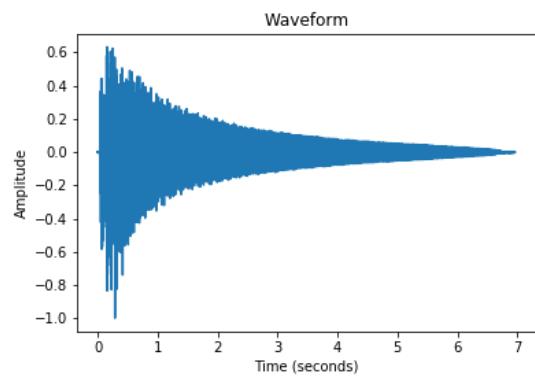
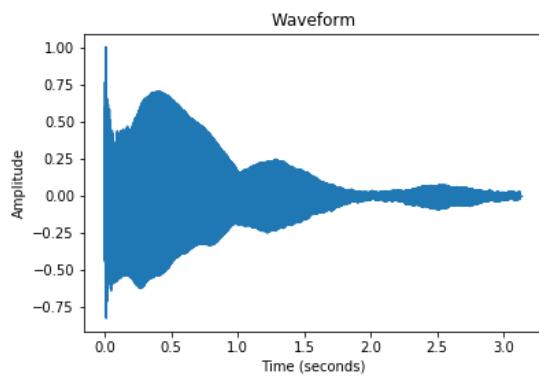
1. Load in raw audio file (`.wav`) format.
2. Sample the signal such that the resultant is discrete data.
3. Convert to Mel Spectrogram—resultant is an image.
4. Save resultant image to respective directory.
5. Repeat for all images.

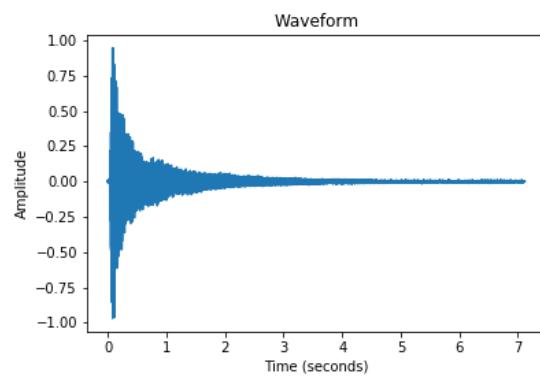
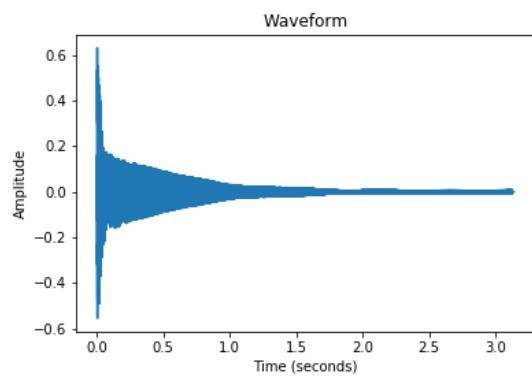
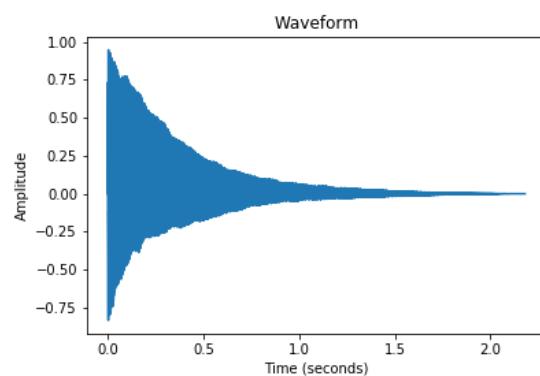
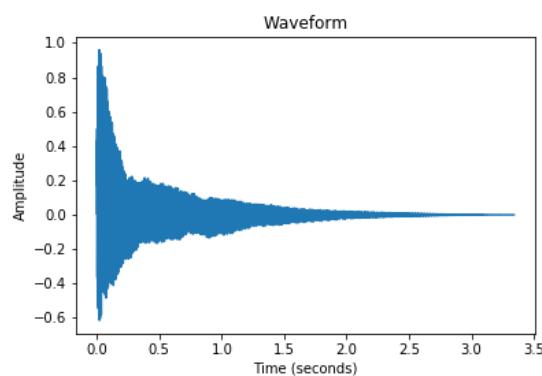
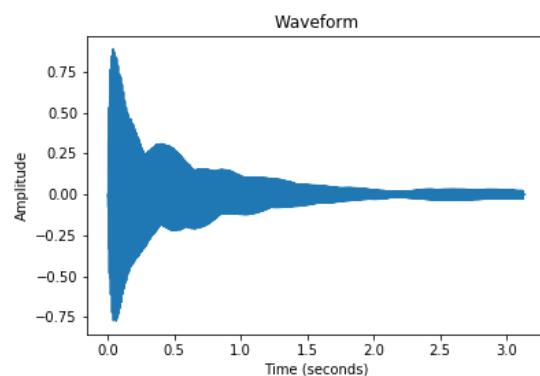
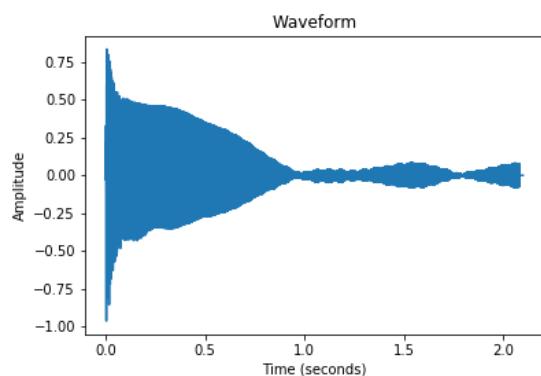
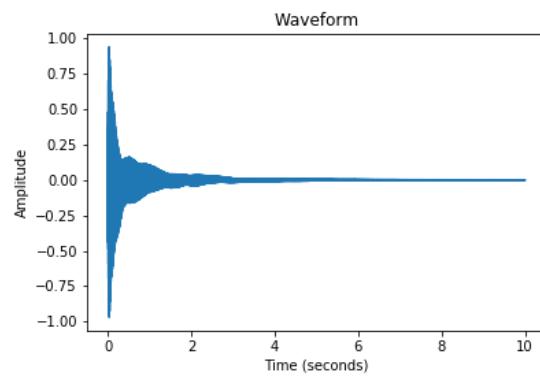
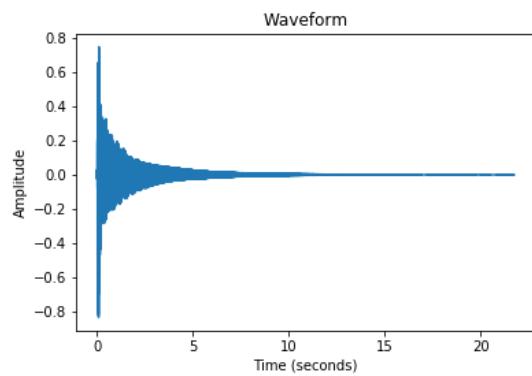
## Waveforms

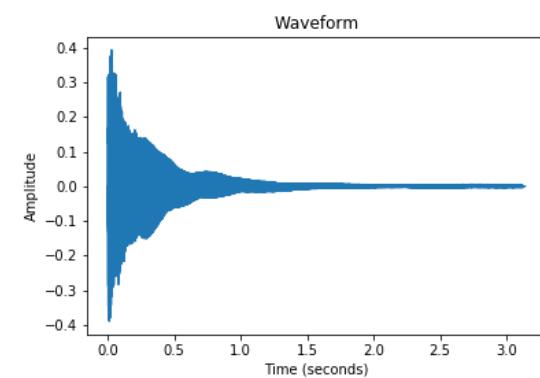
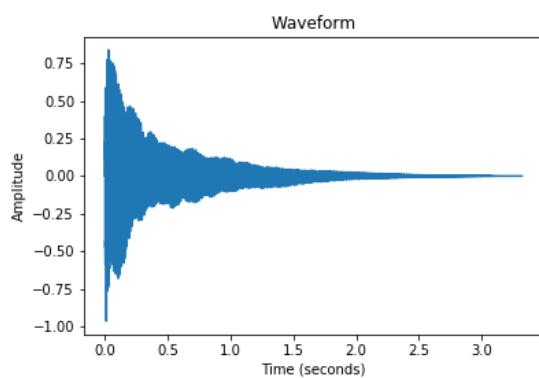
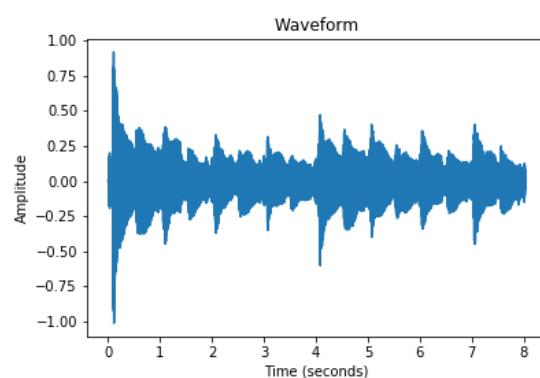
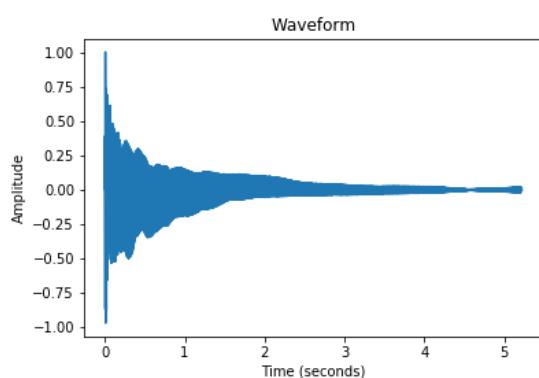
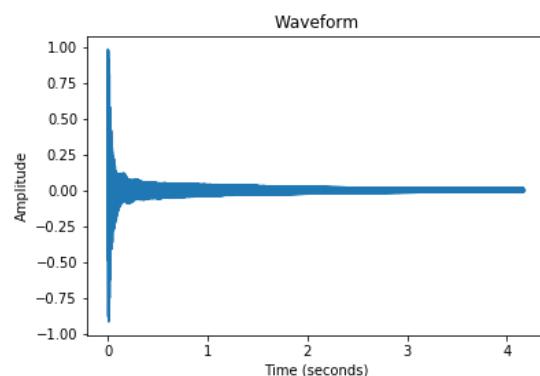
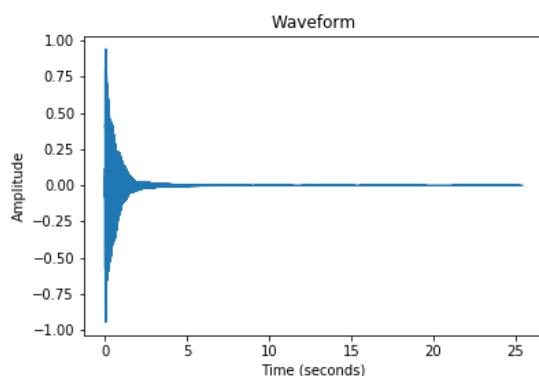
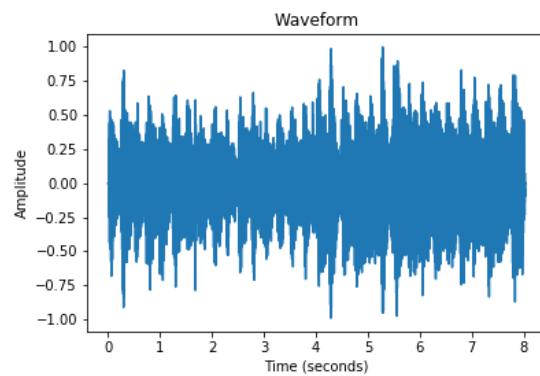
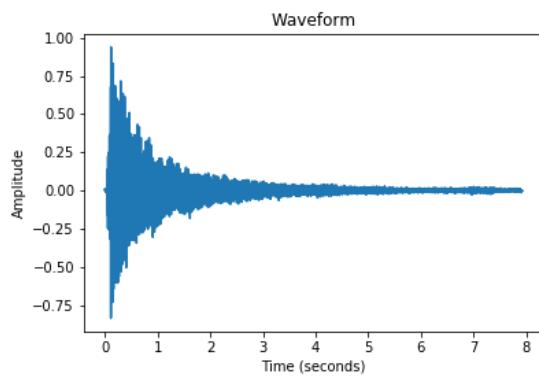
The following are plotted waveforms for each audio signal in each class. Out of the 30 audio samples per class, the first 15 are shown on the left side and the remaining 15 are shown on the right side.

## Acoustic Guitar

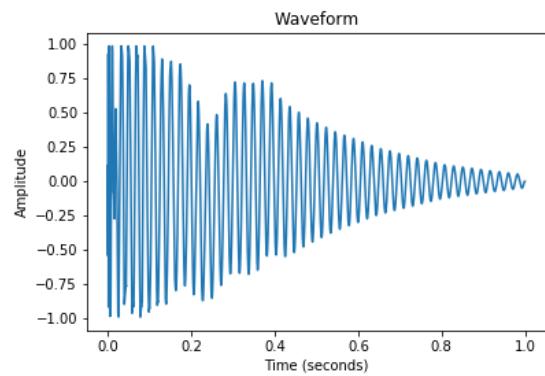
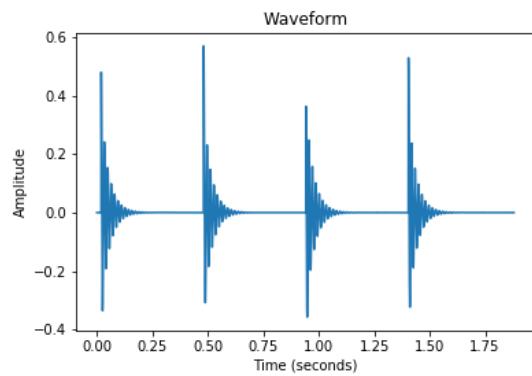
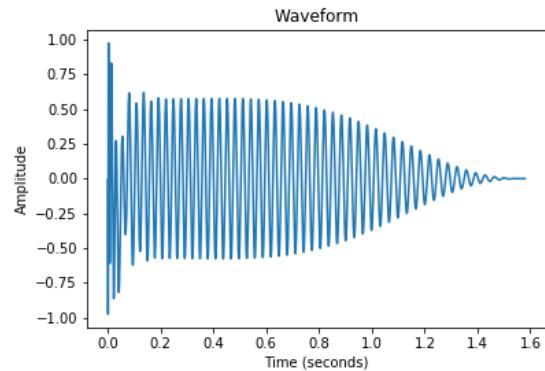
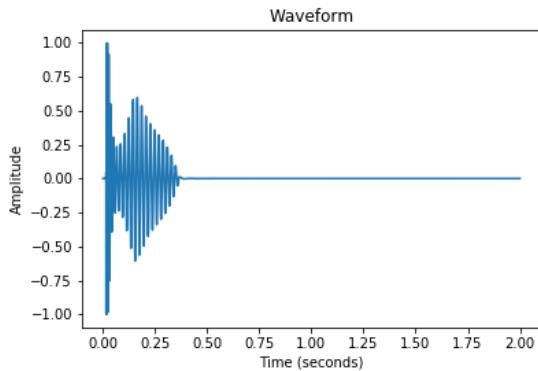
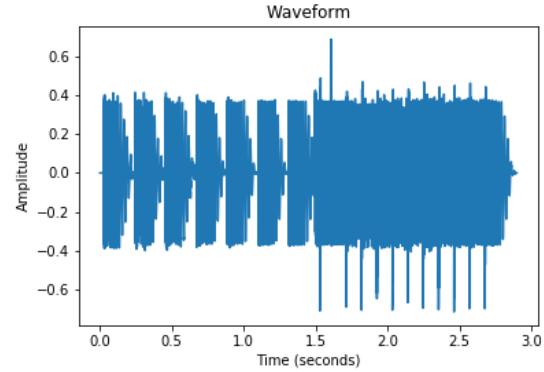
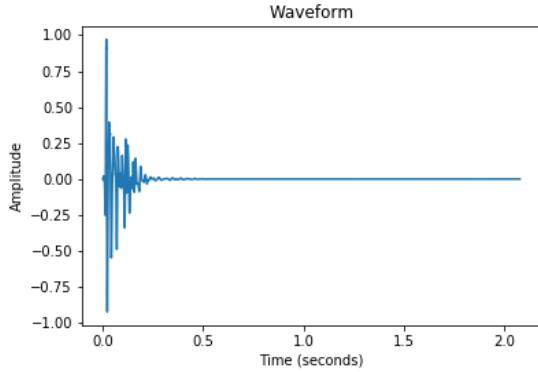


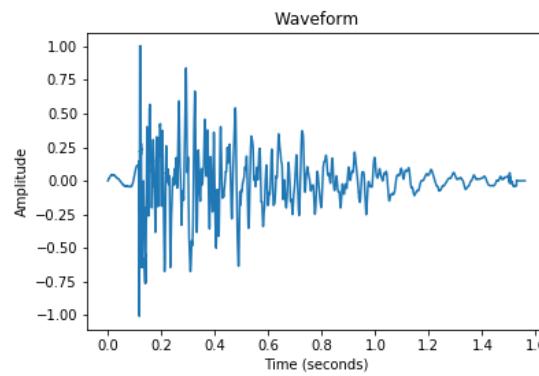
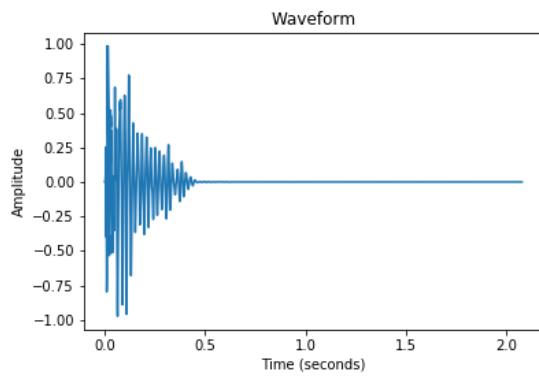
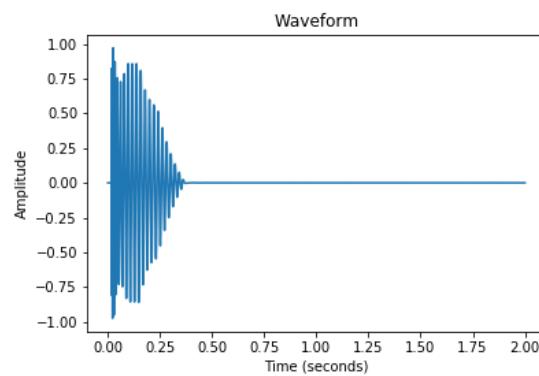
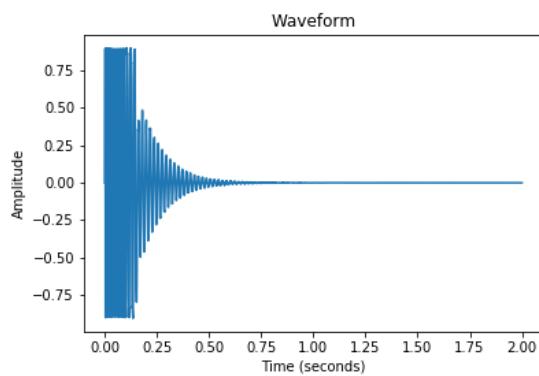
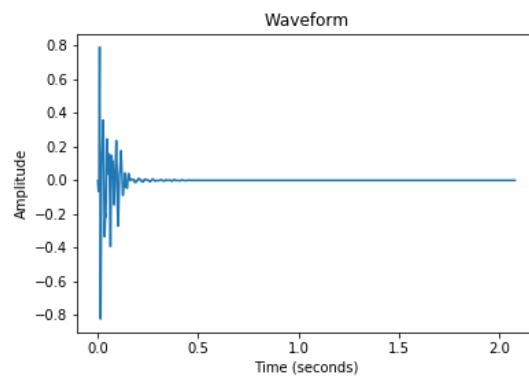
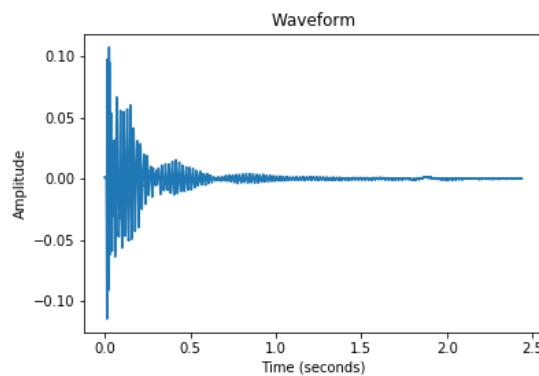
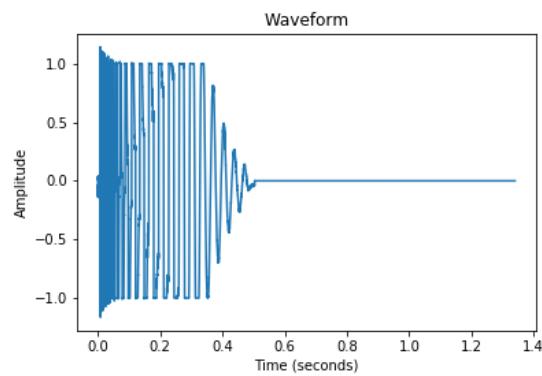
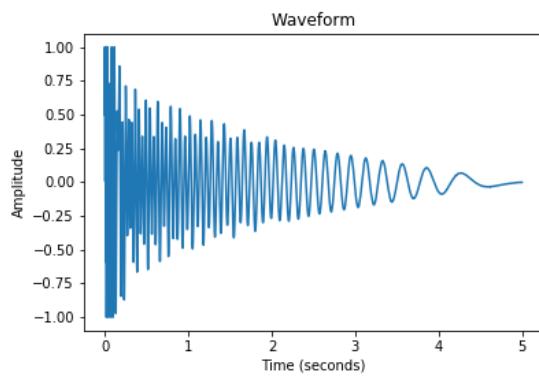


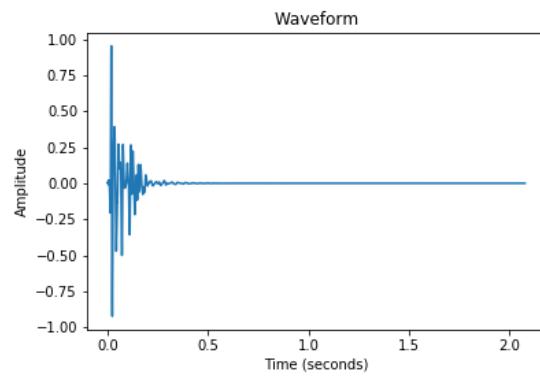
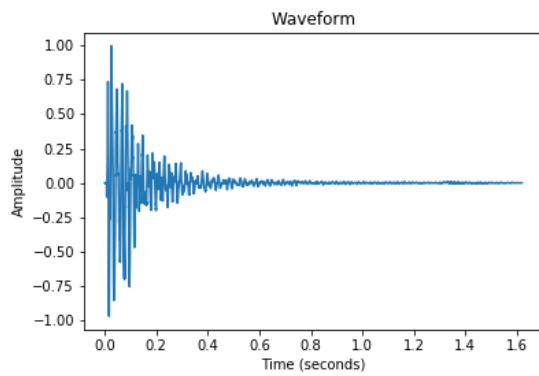
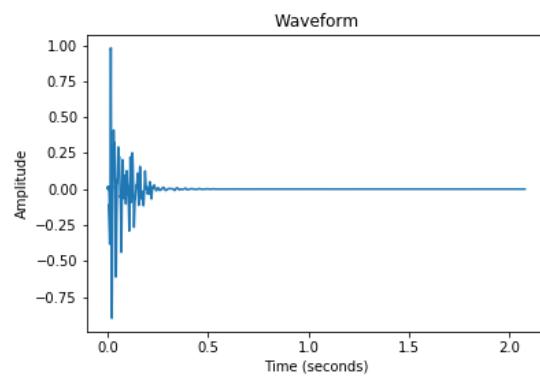
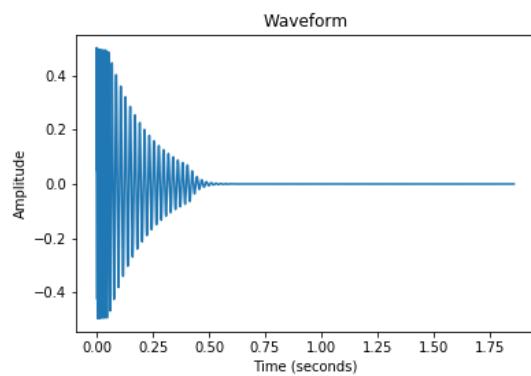
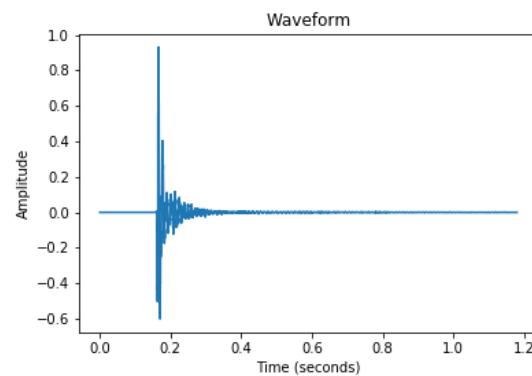
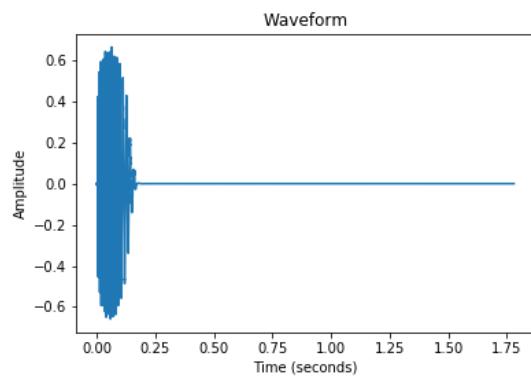
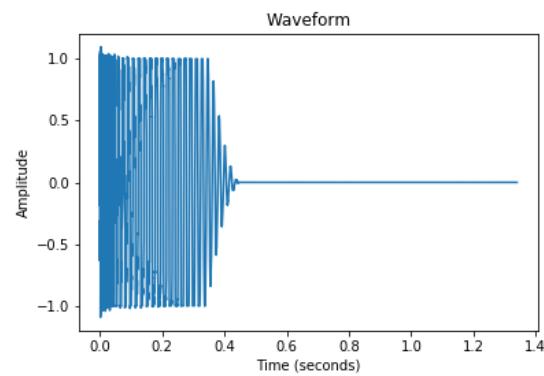
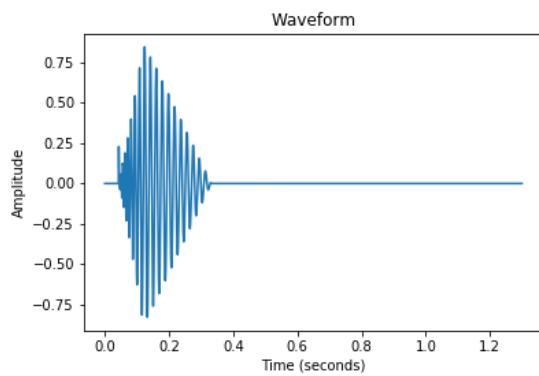


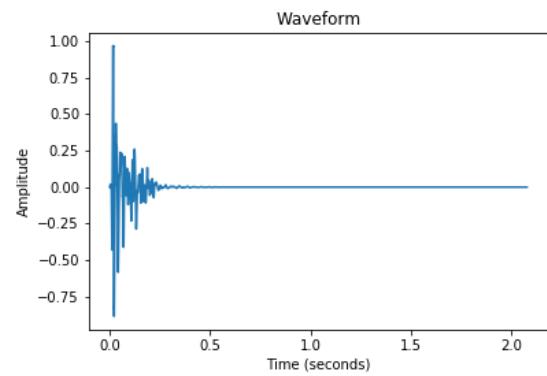
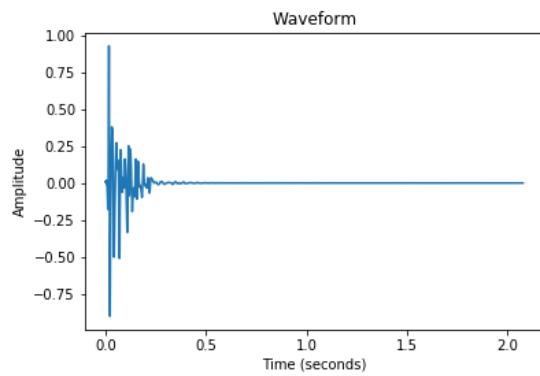
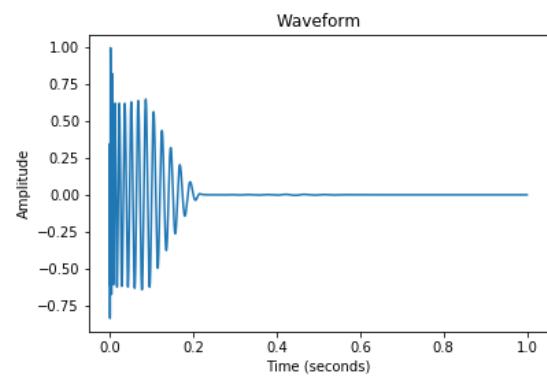
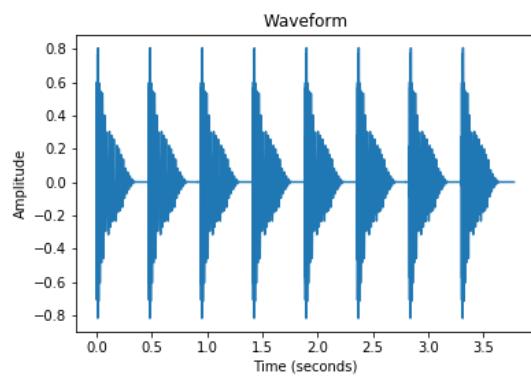
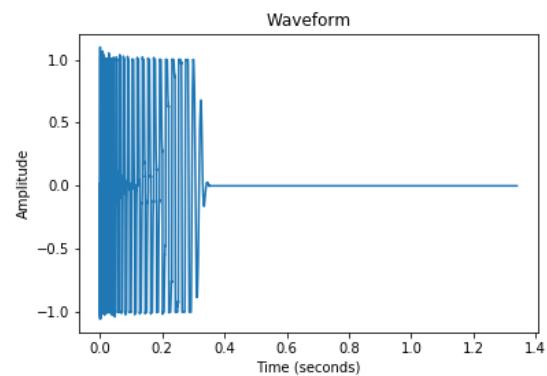
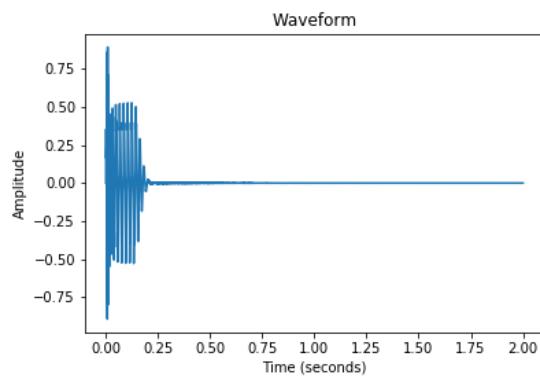
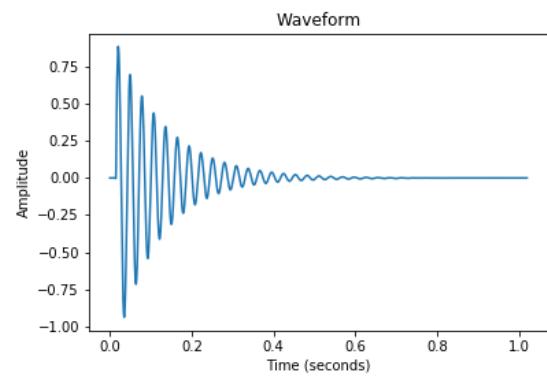
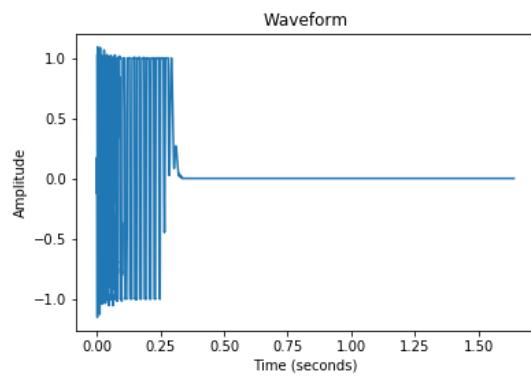


## Bass Drum

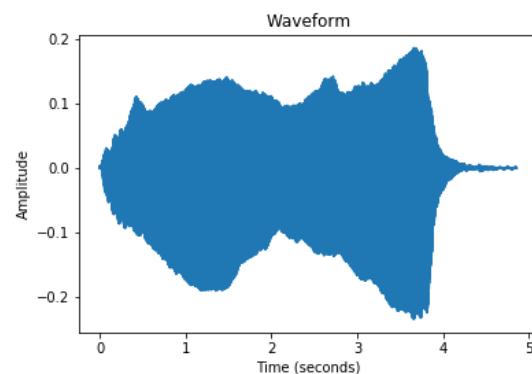
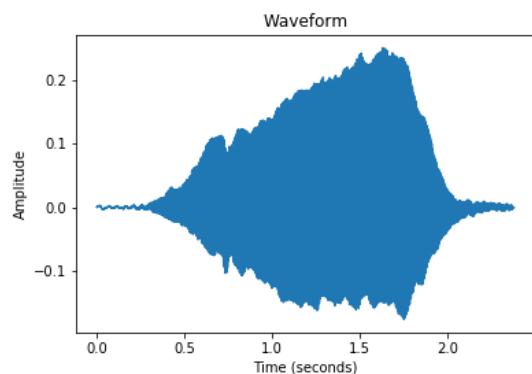
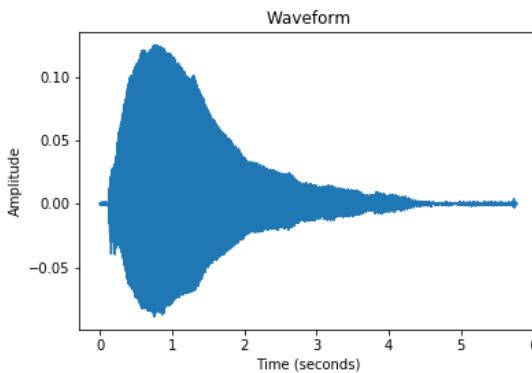
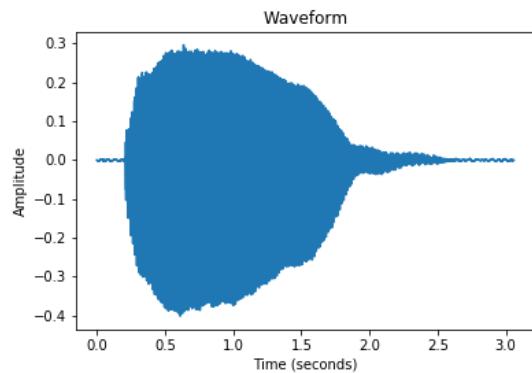
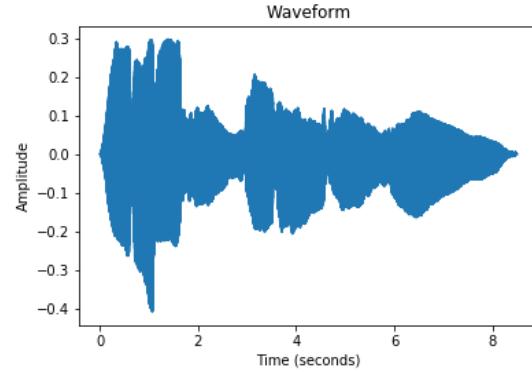
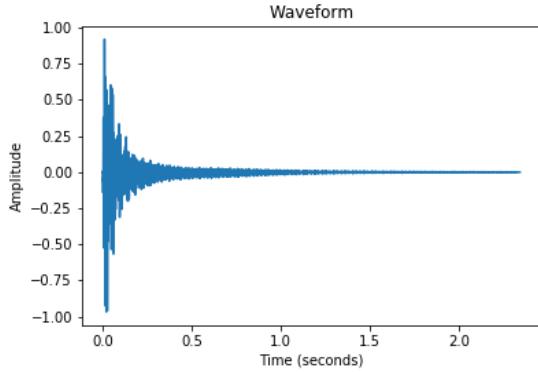


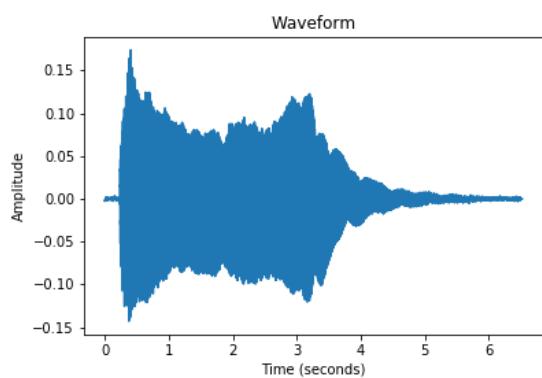
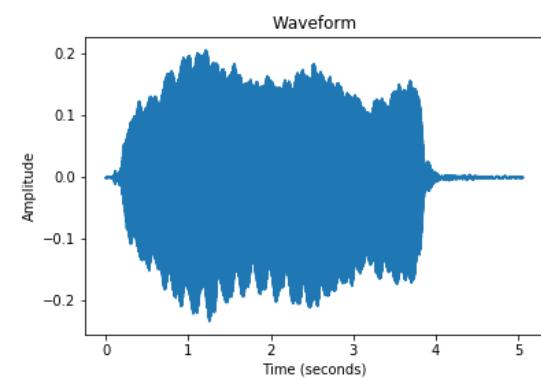
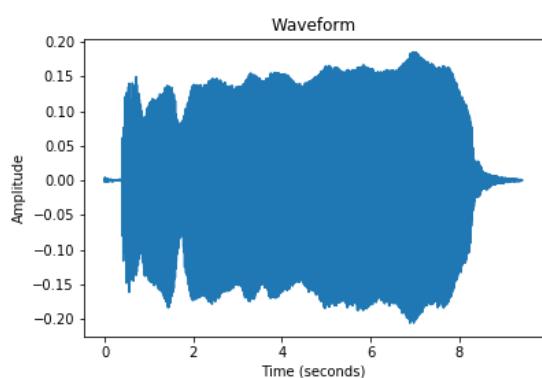
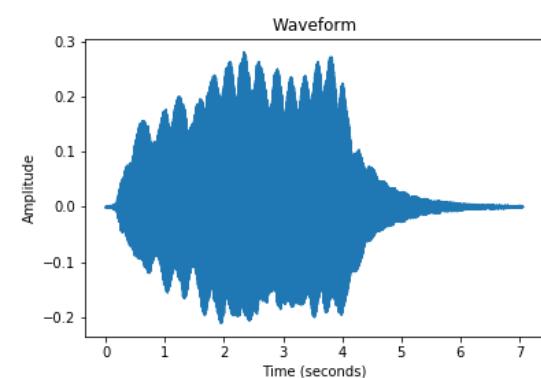
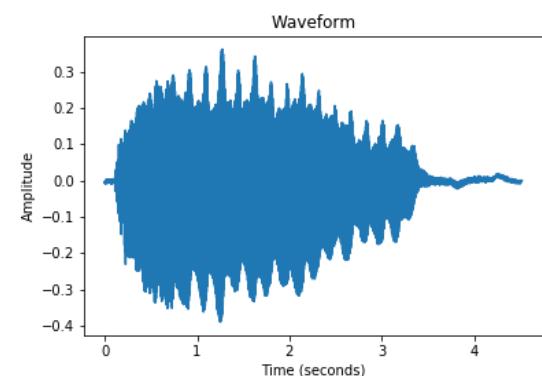
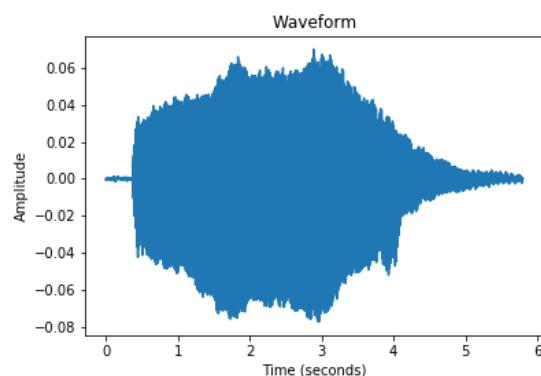
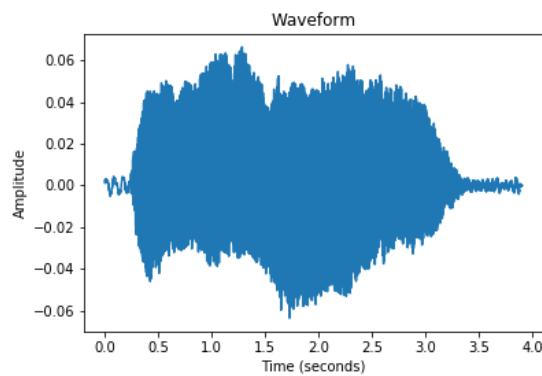
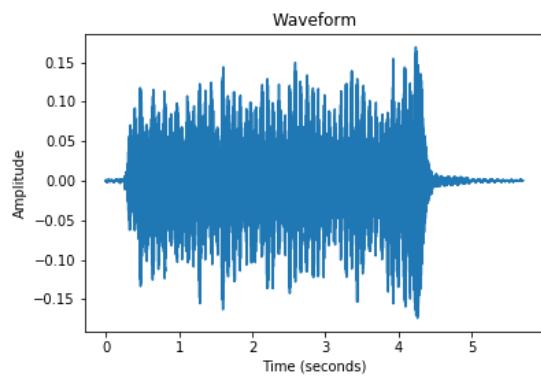


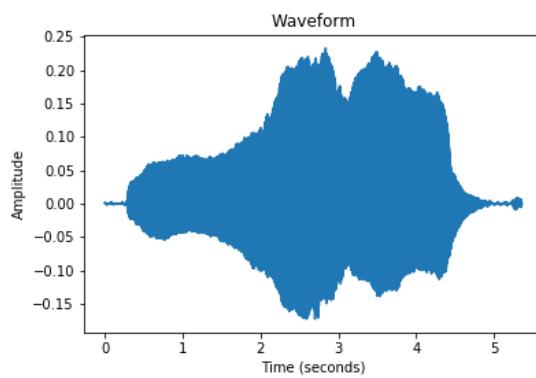
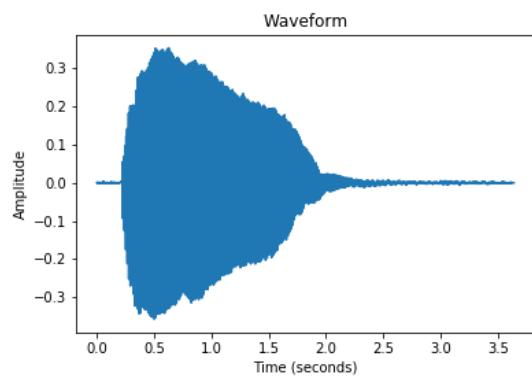
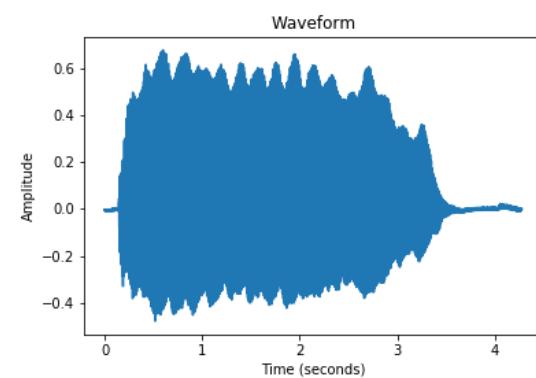
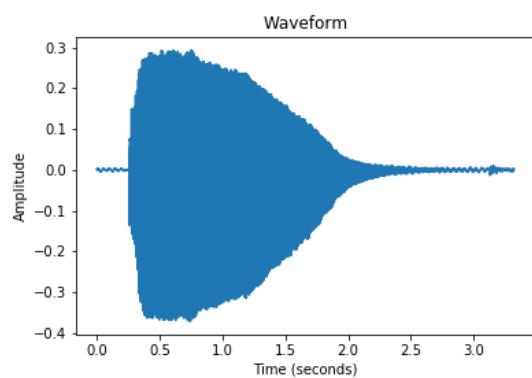
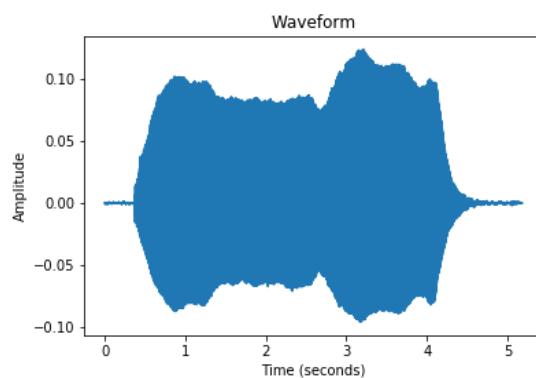
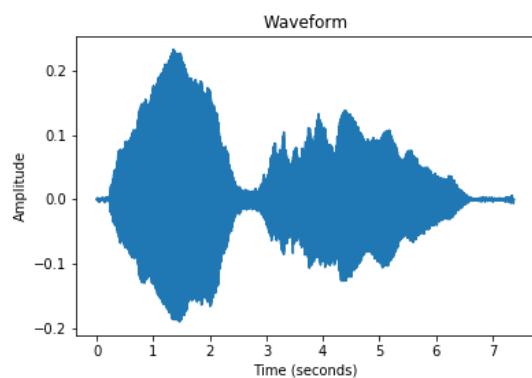
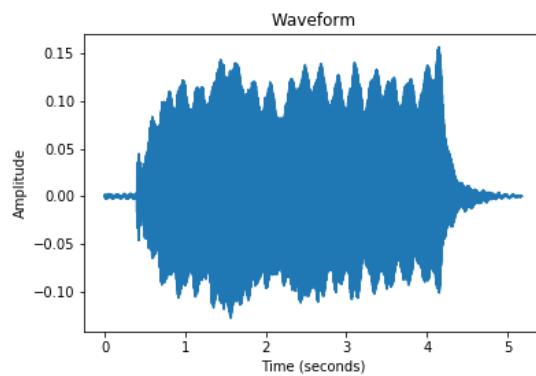
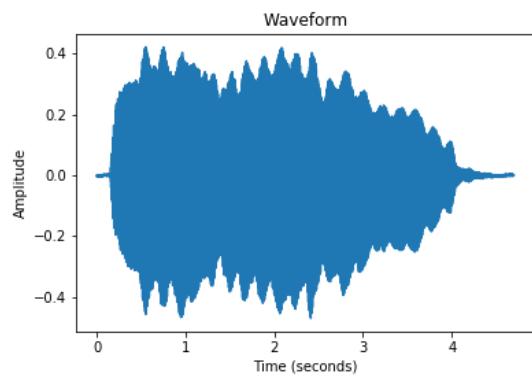


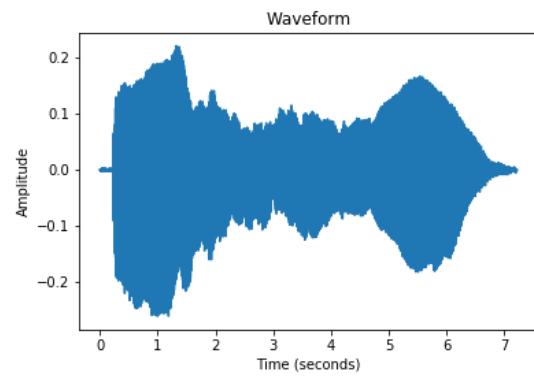
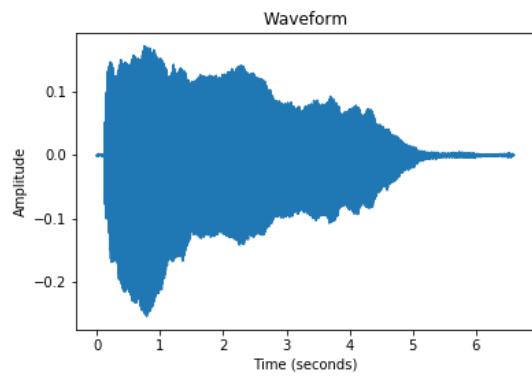
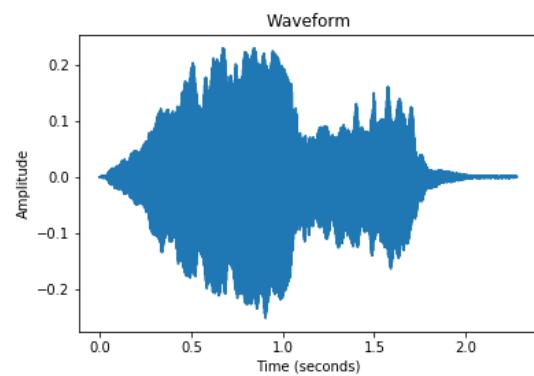
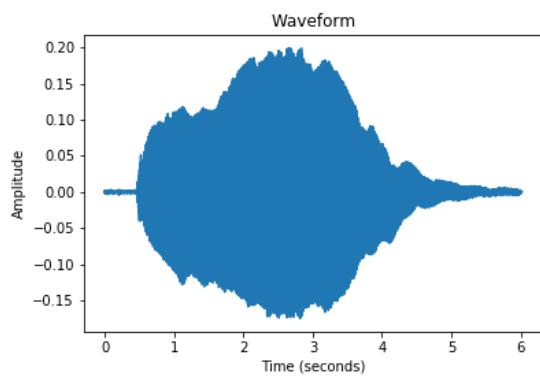
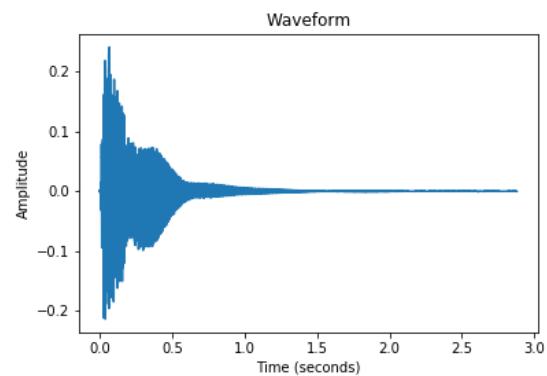
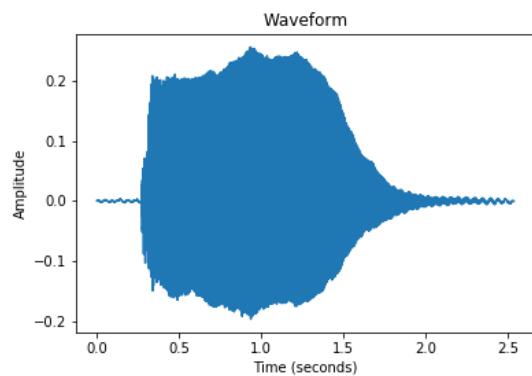
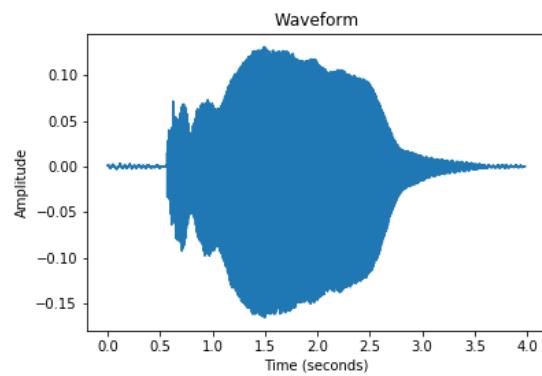
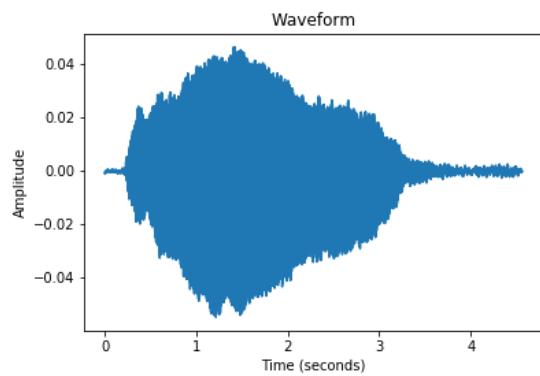


## Cello

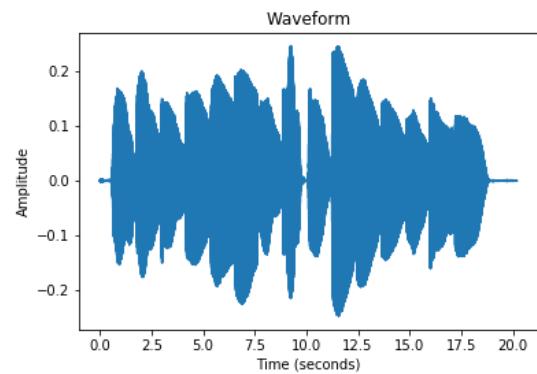
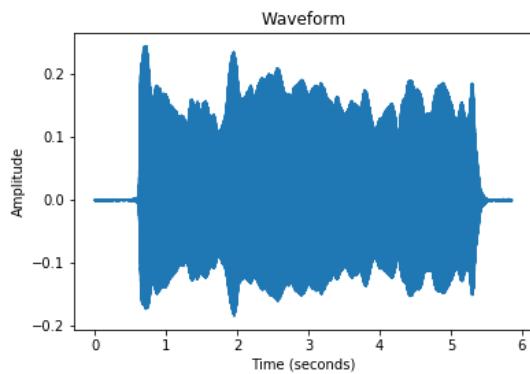
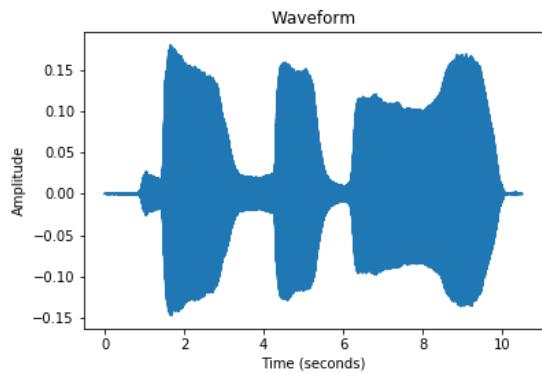
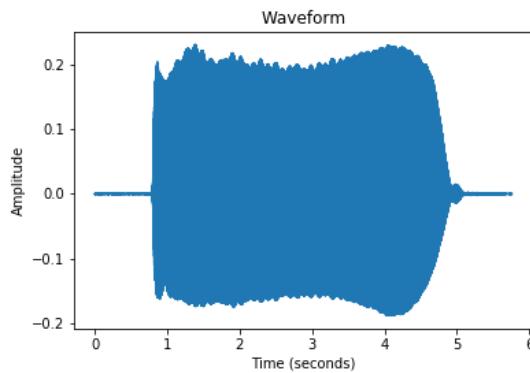
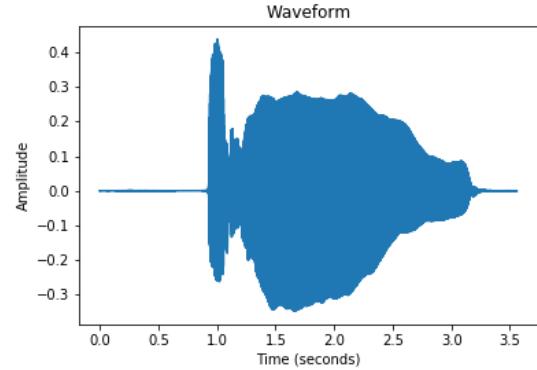
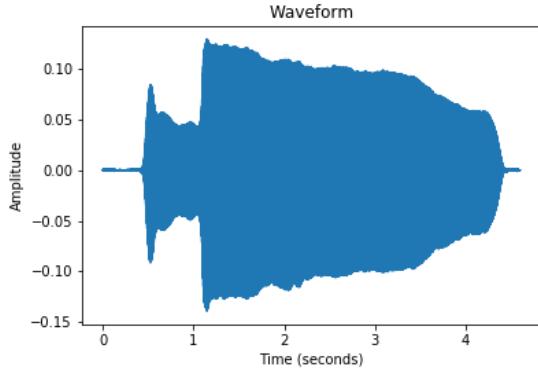


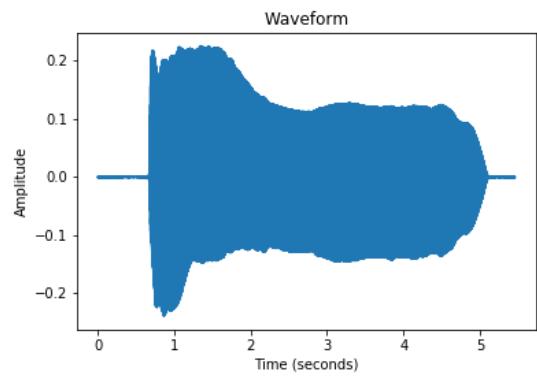
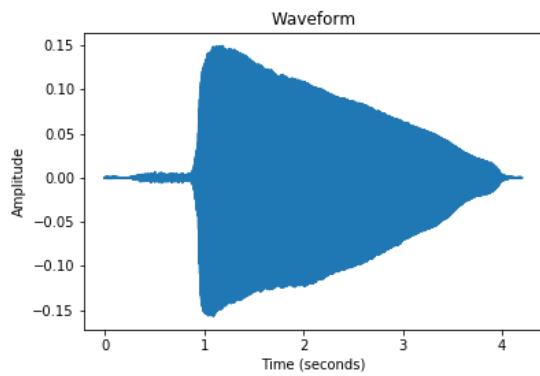
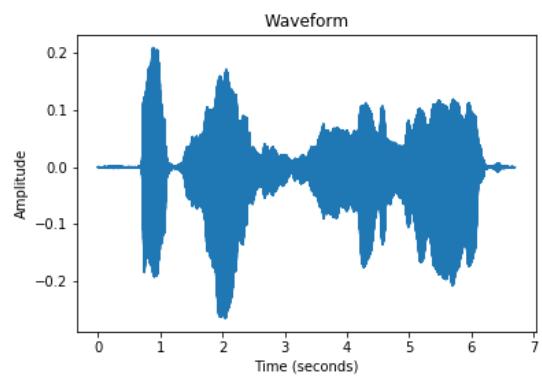
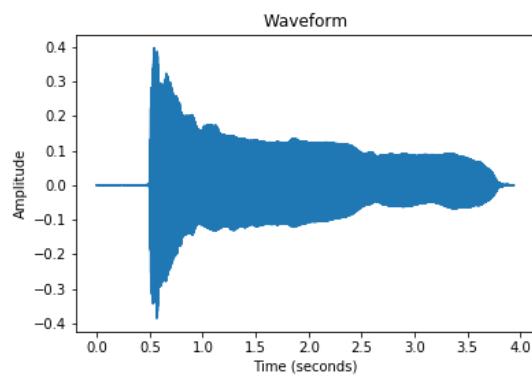
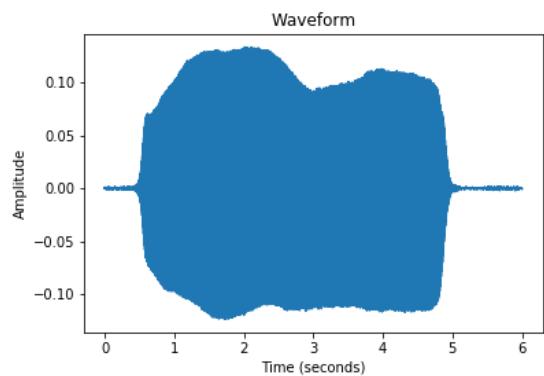
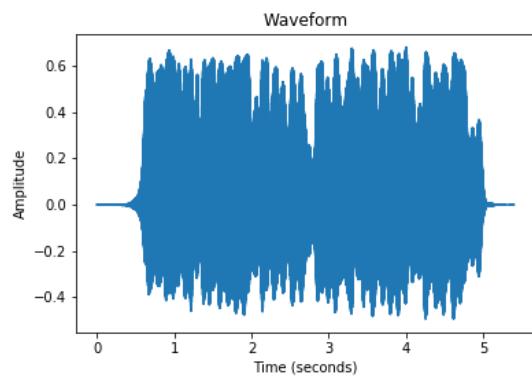
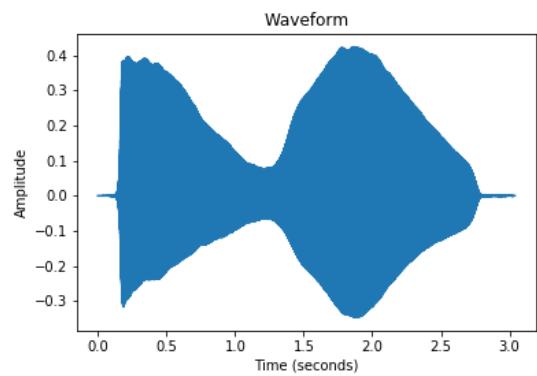
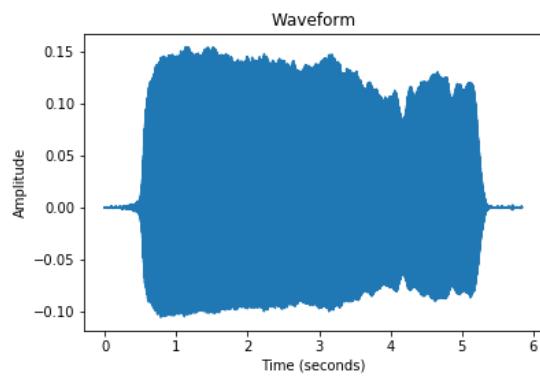


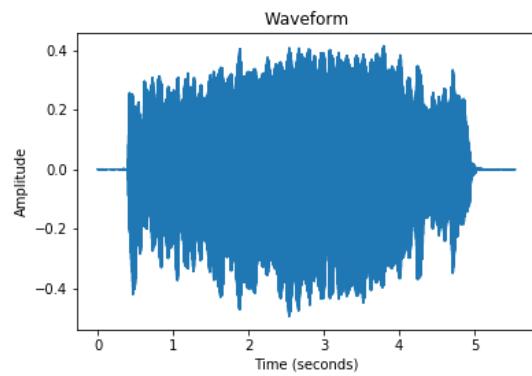
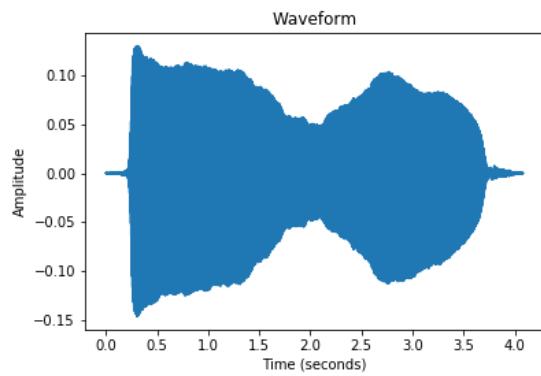
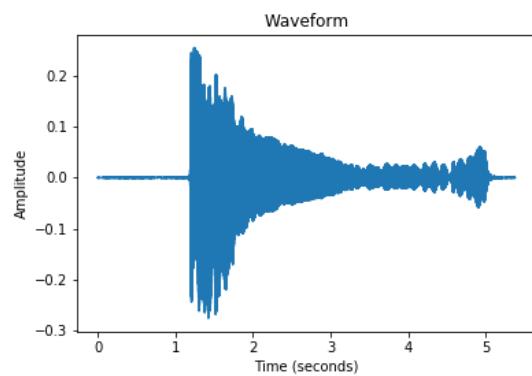
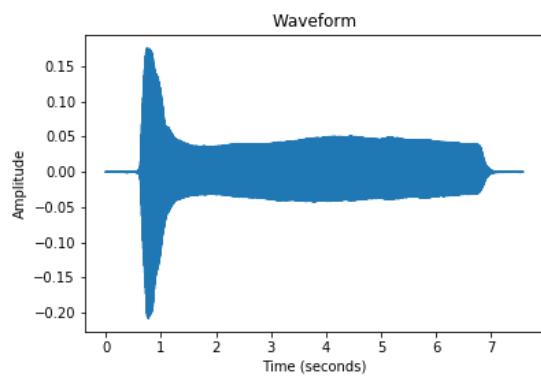
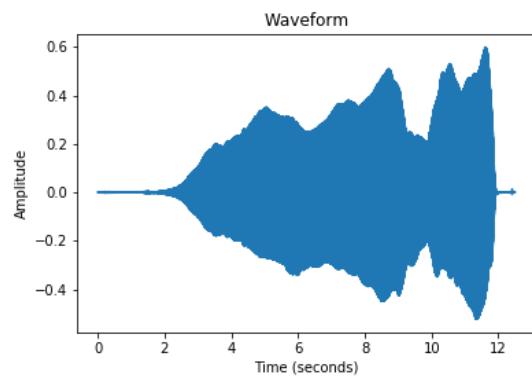
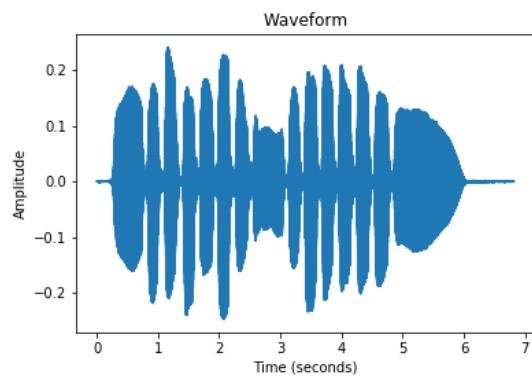
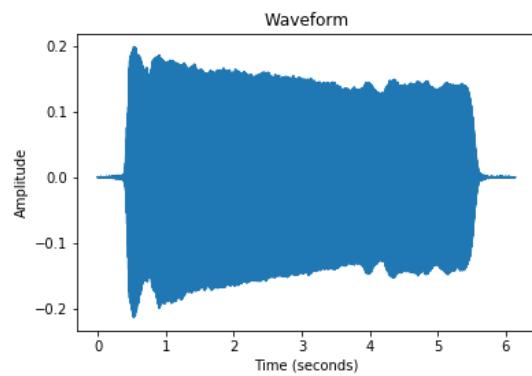
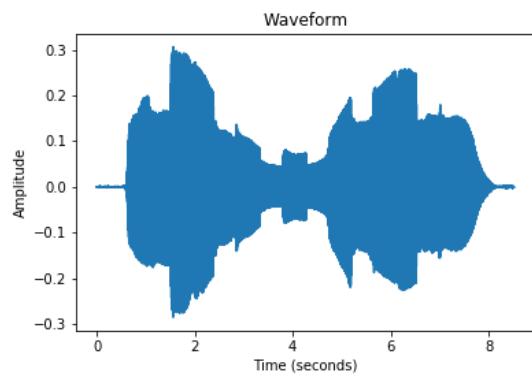


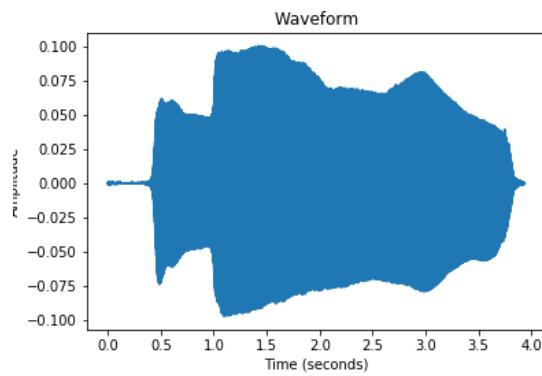
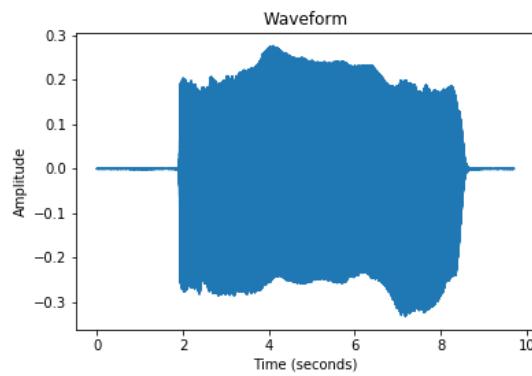
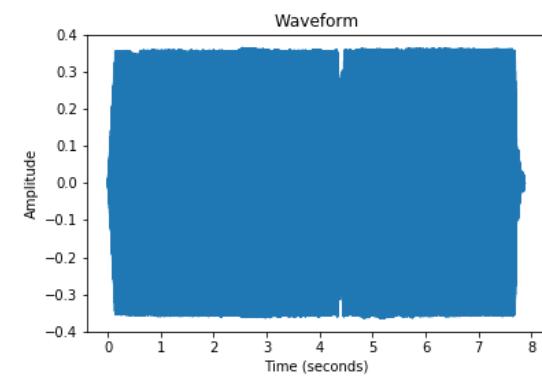
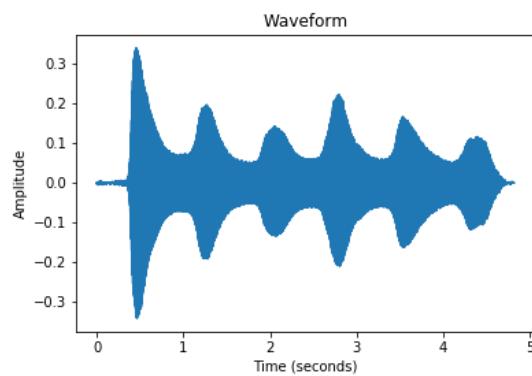
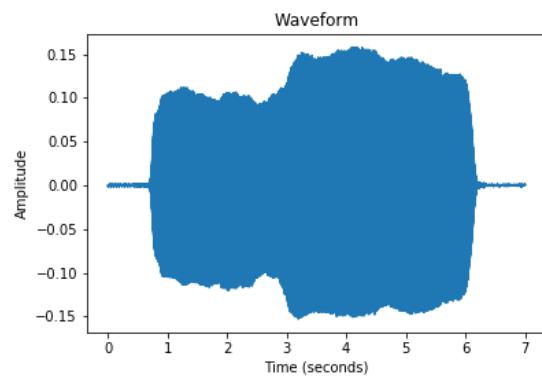
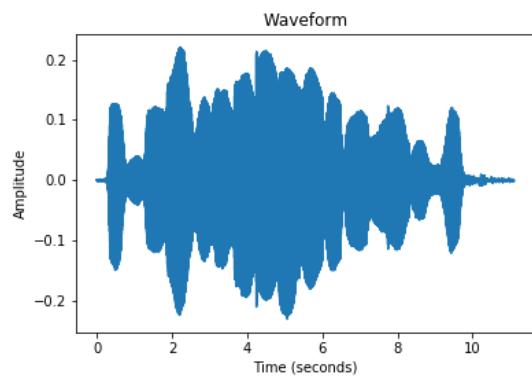
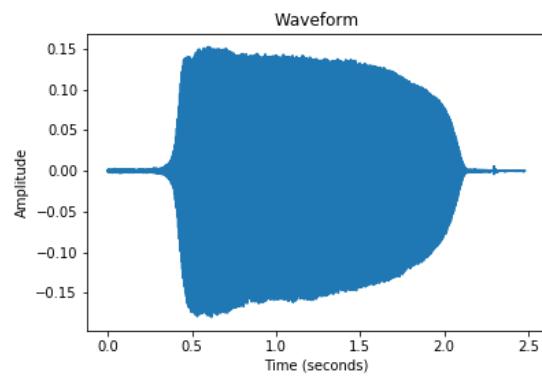
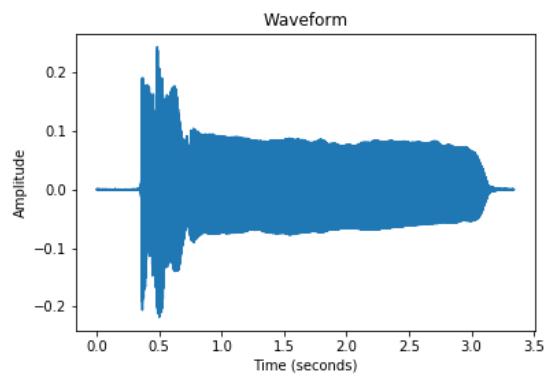


## Clarinet

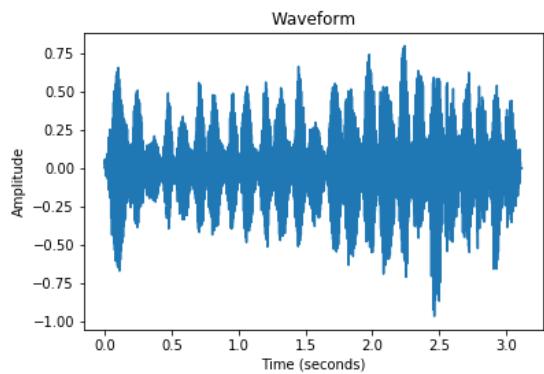
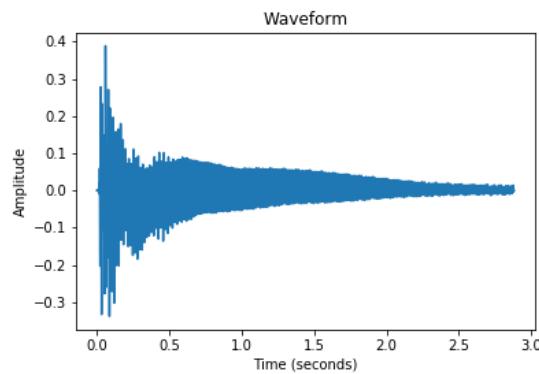
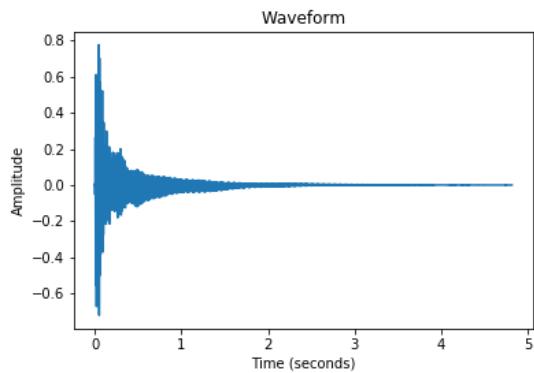
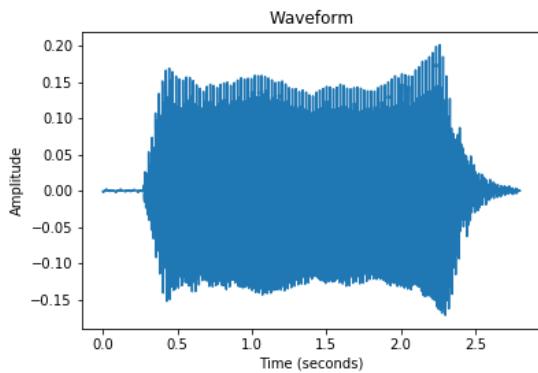
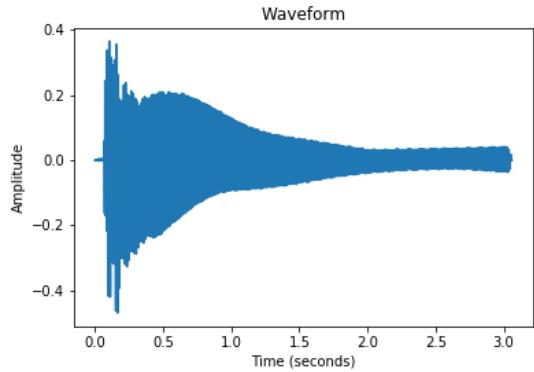
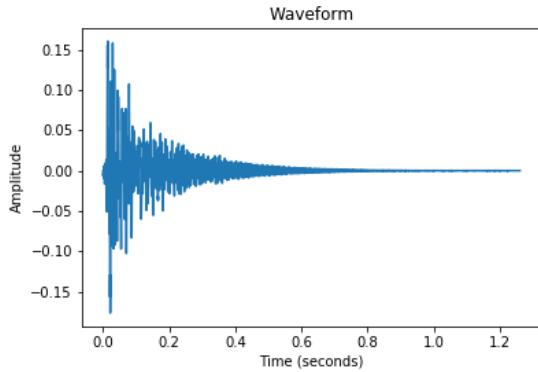


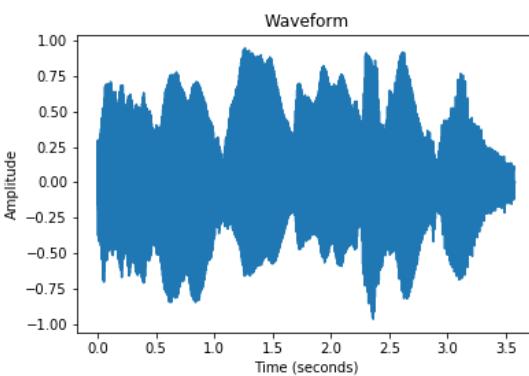
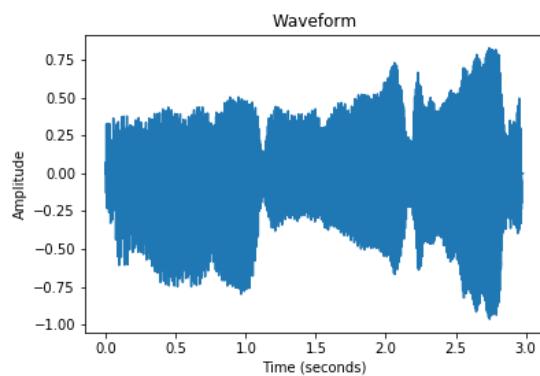
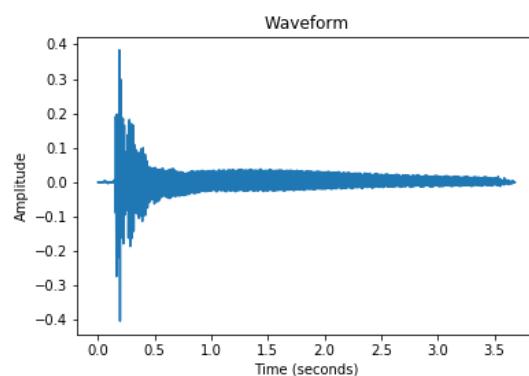
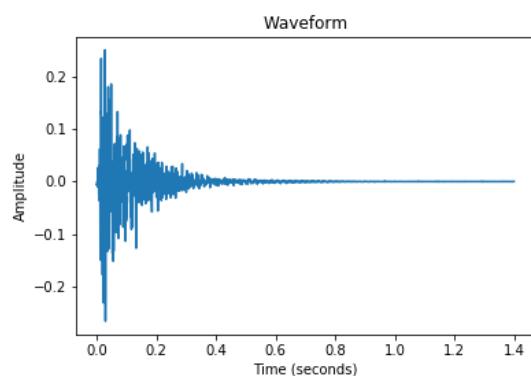
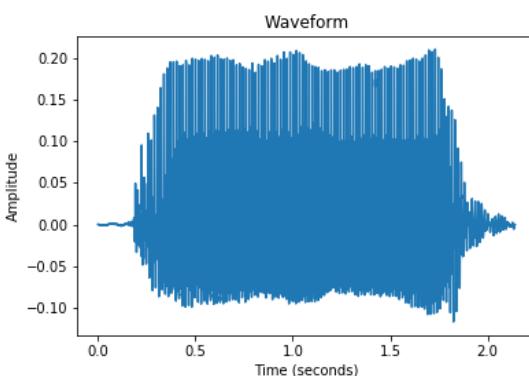
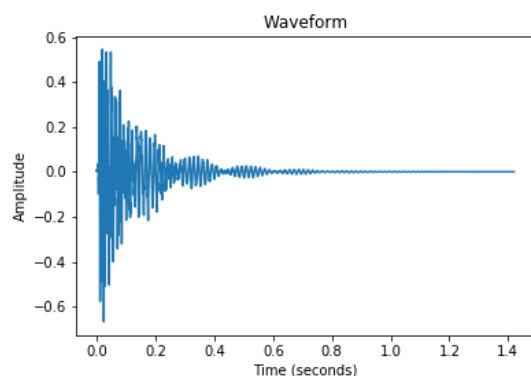
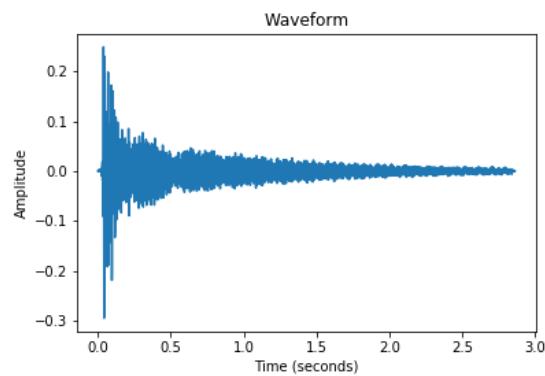
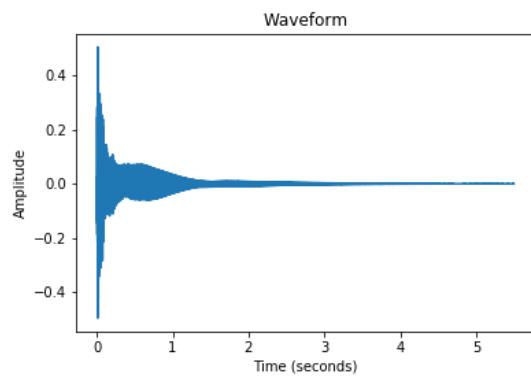


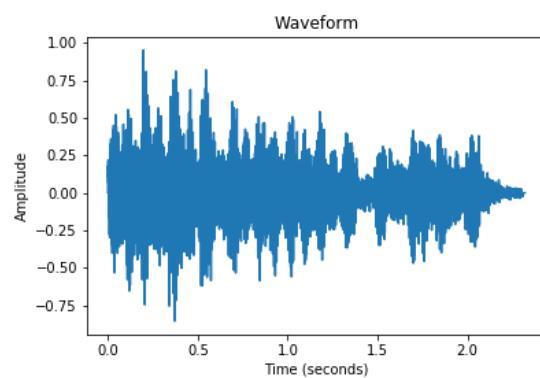
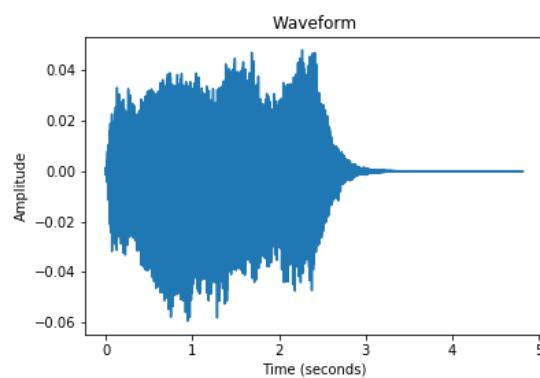
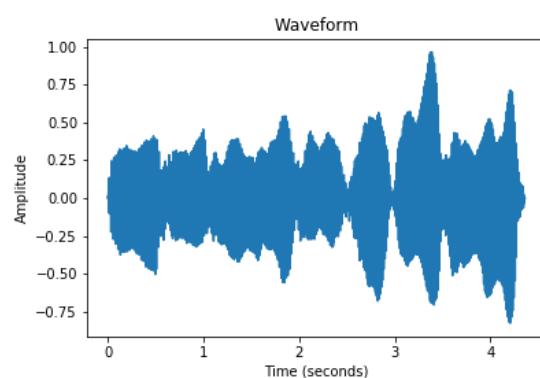
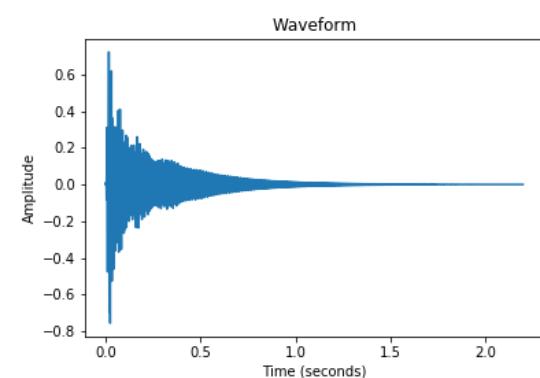
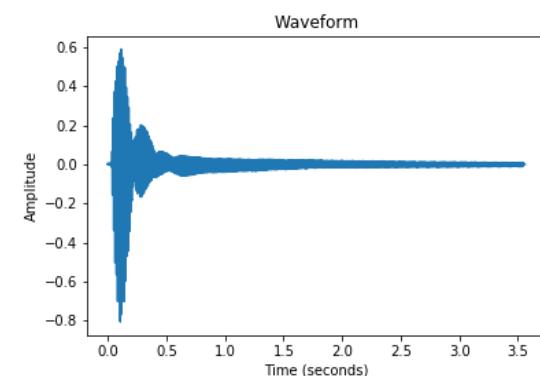
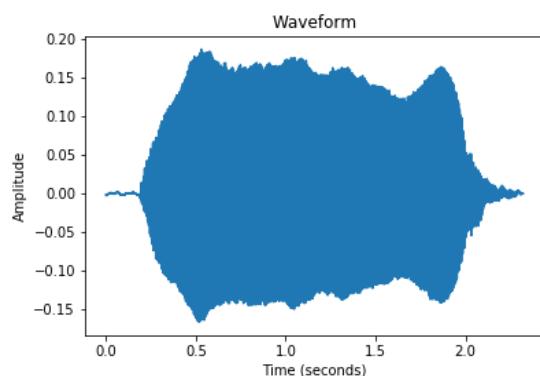
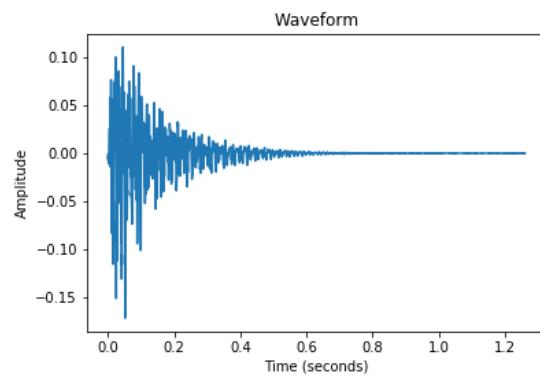
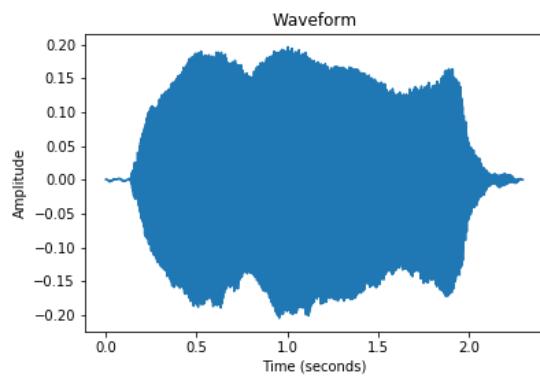


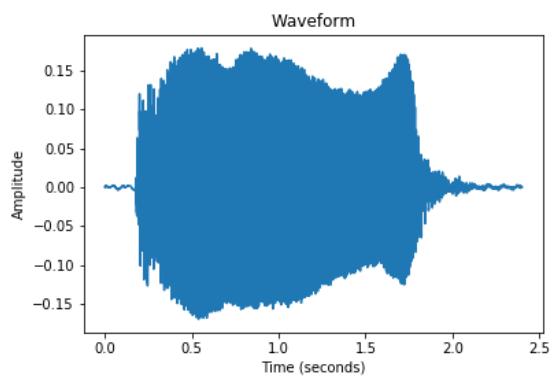
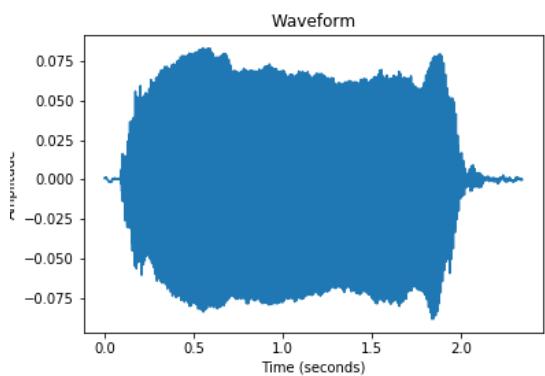
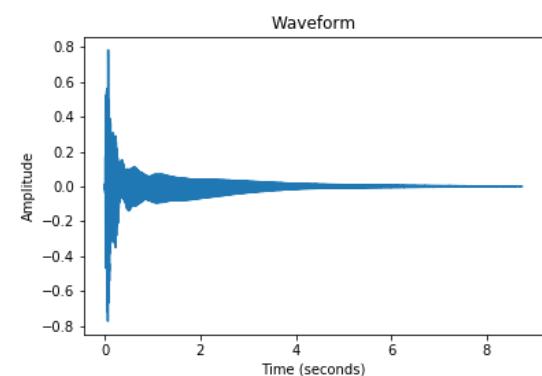
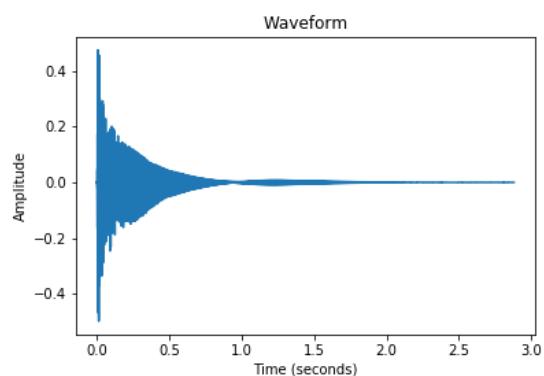
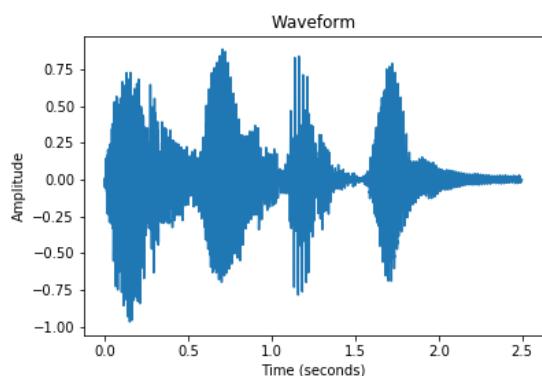
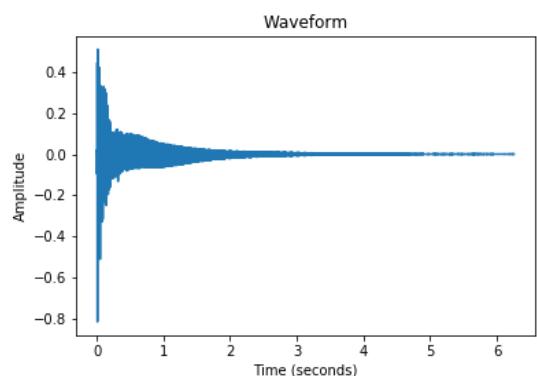
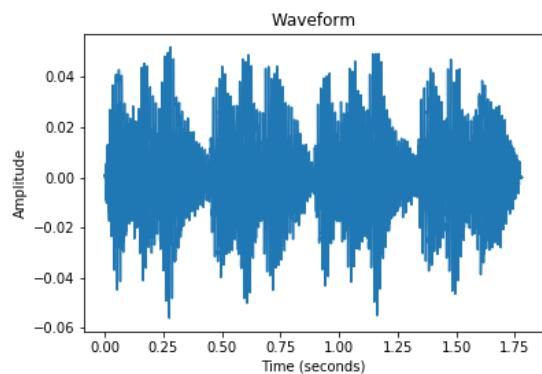
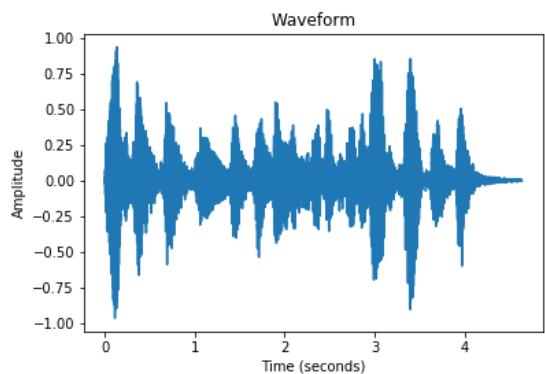


## Double Bass

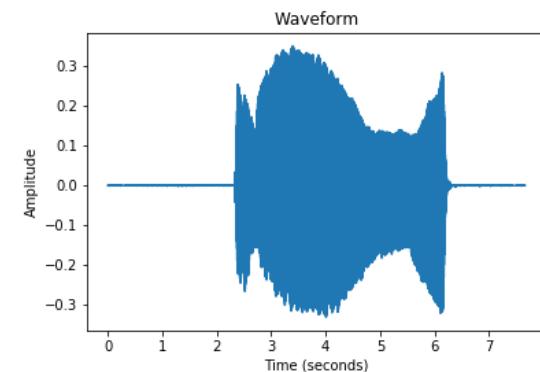
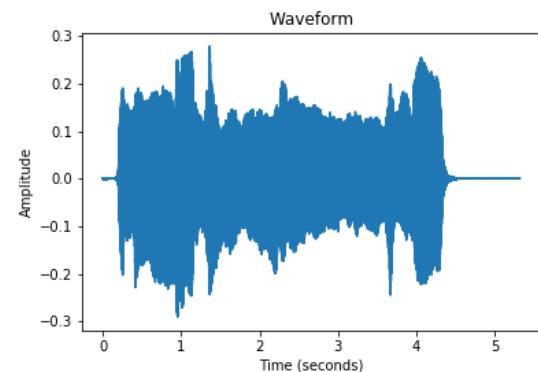
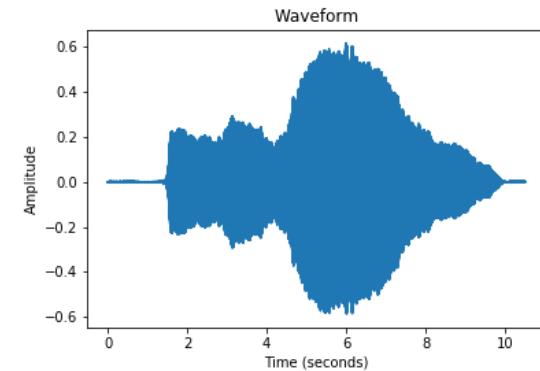
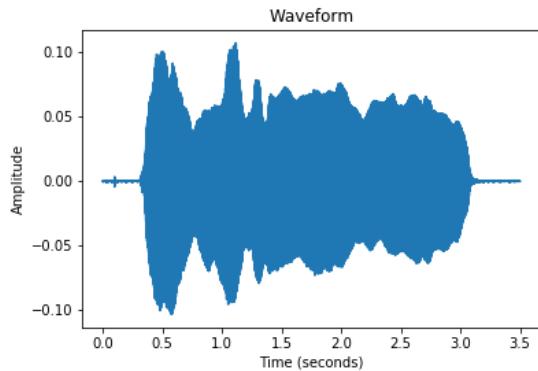
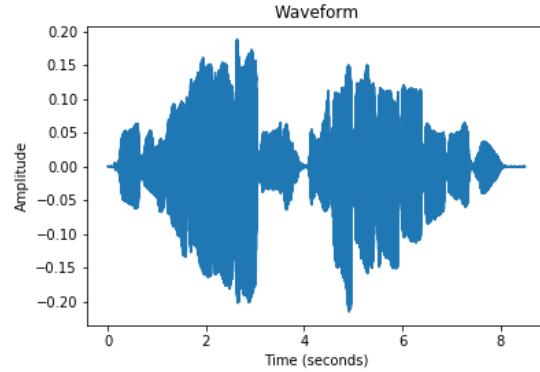
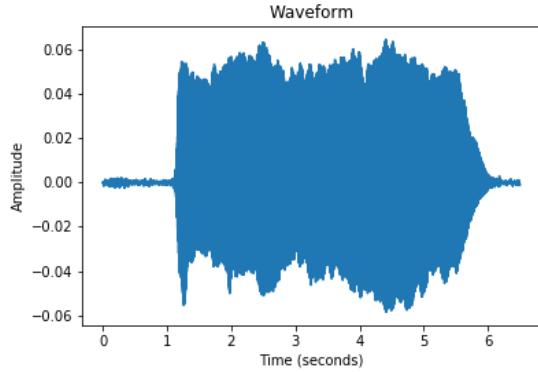


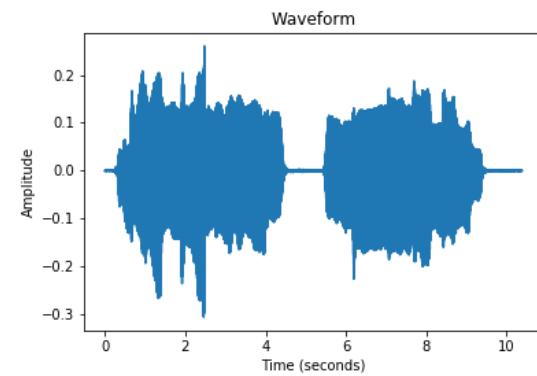
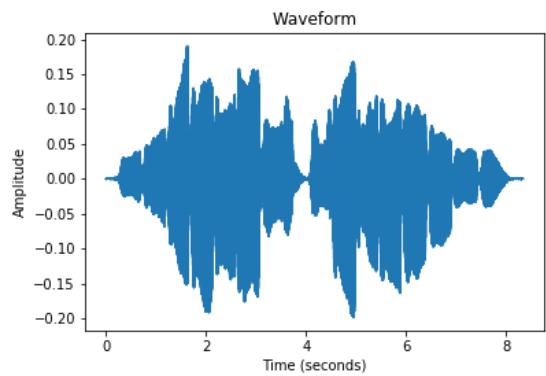
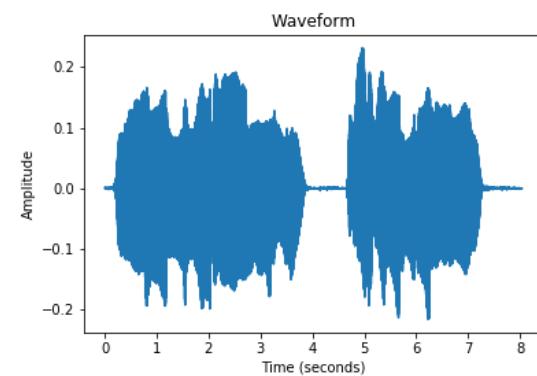
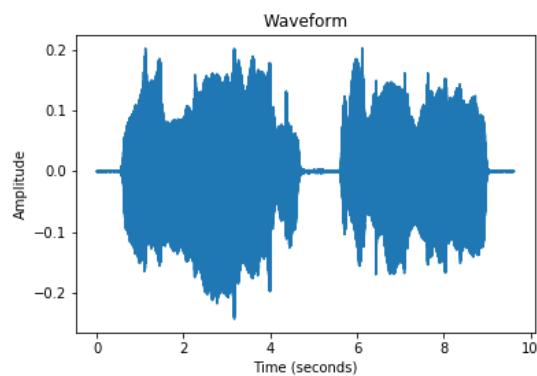
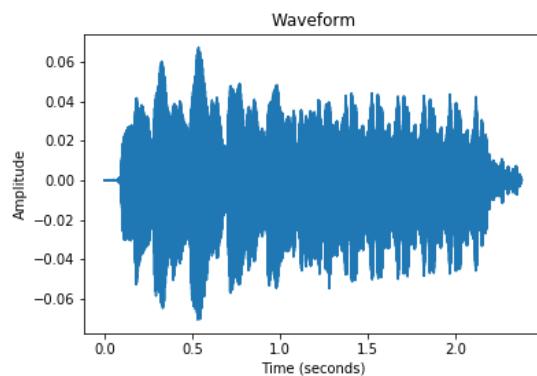
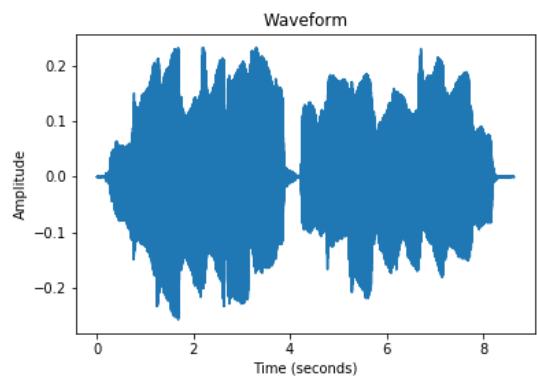
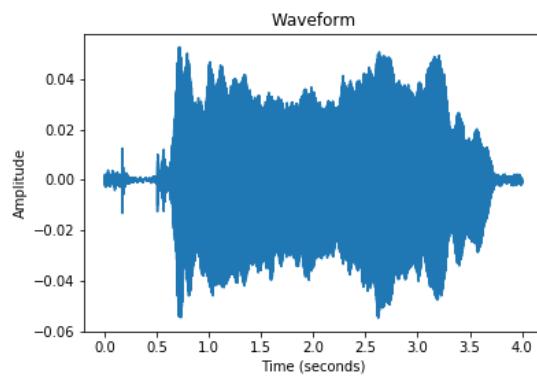
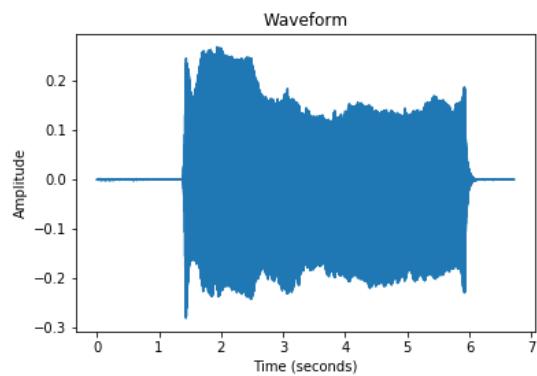


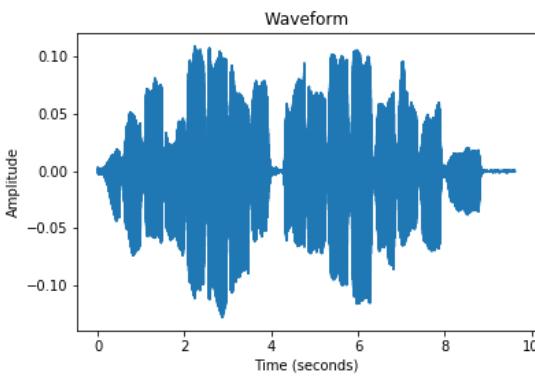
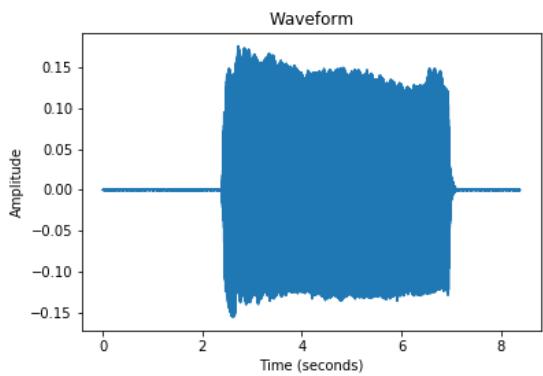
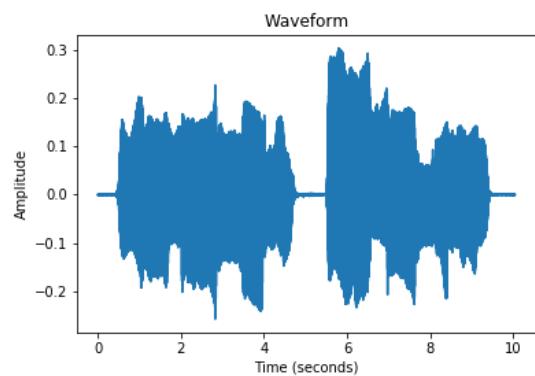
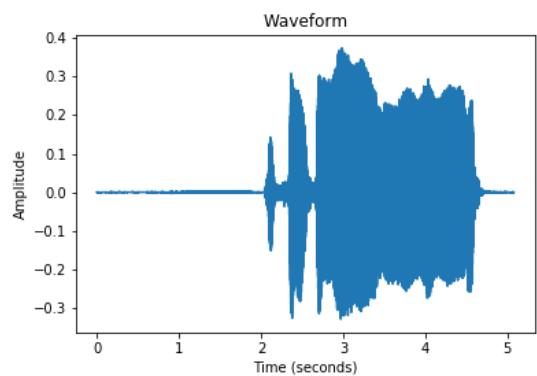
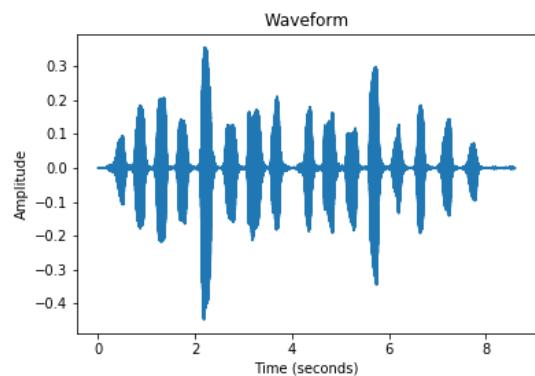
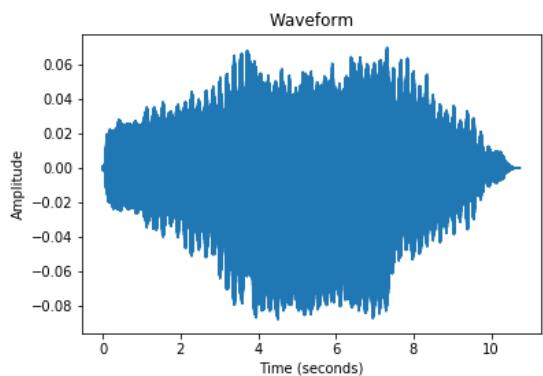
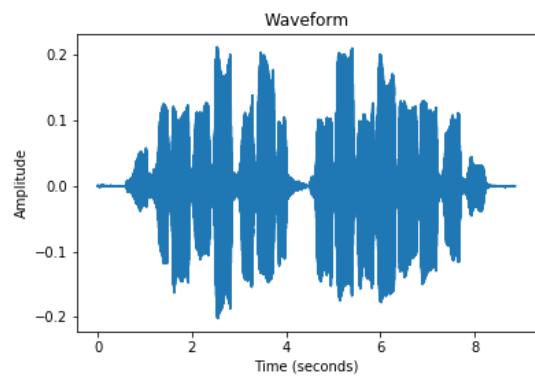
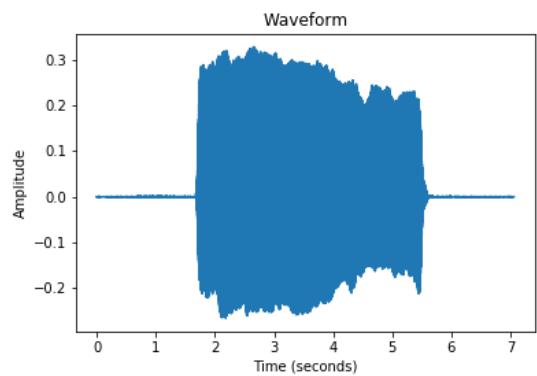


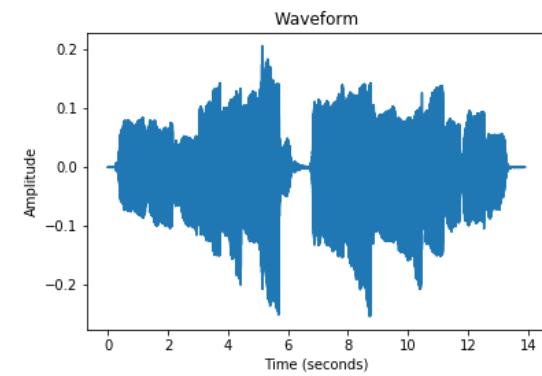
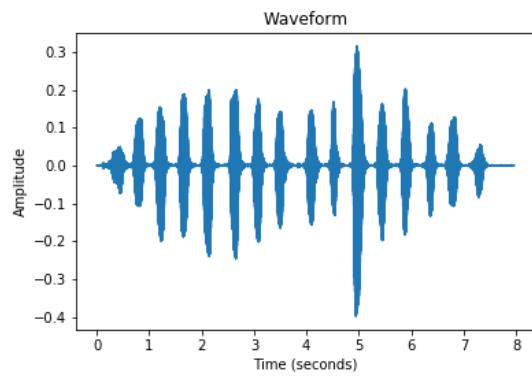
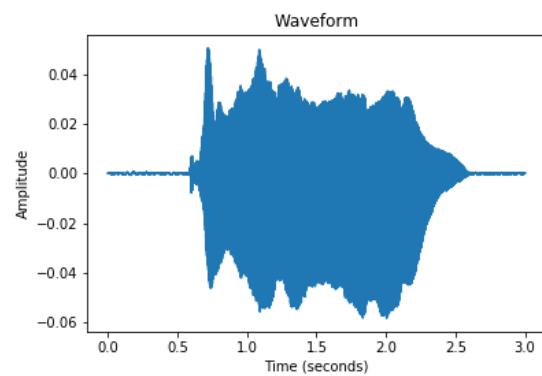
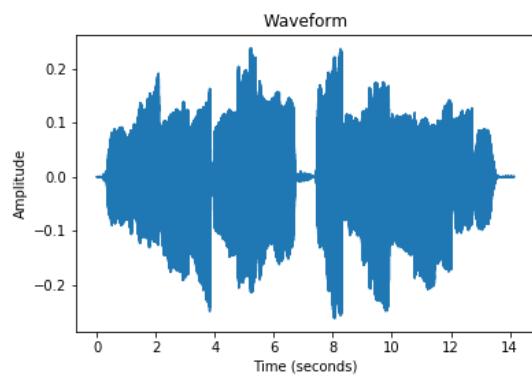
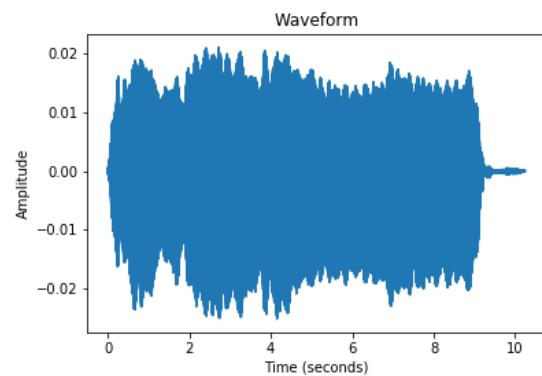
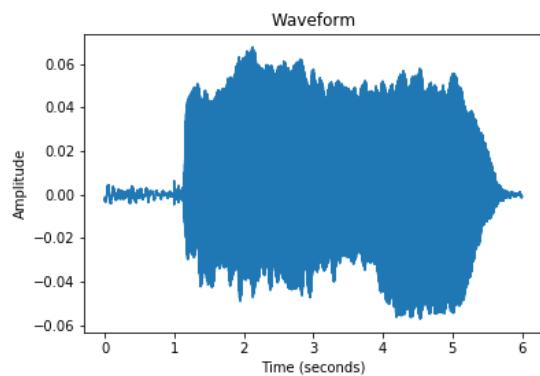
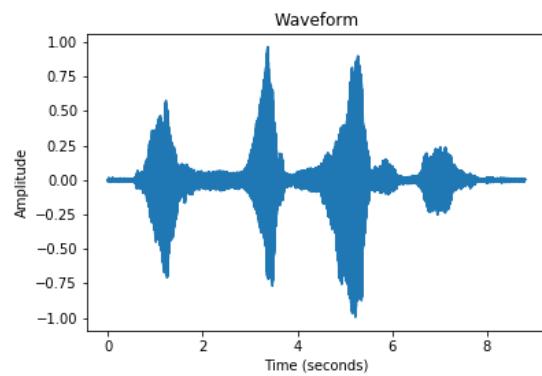
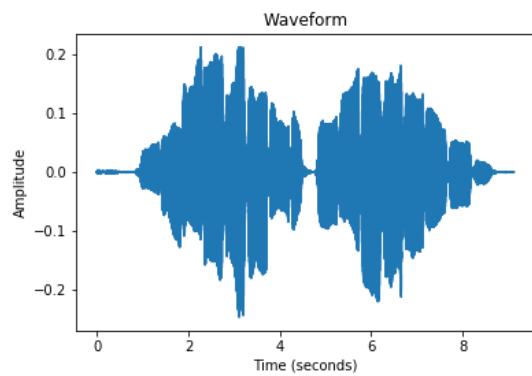


## Flute

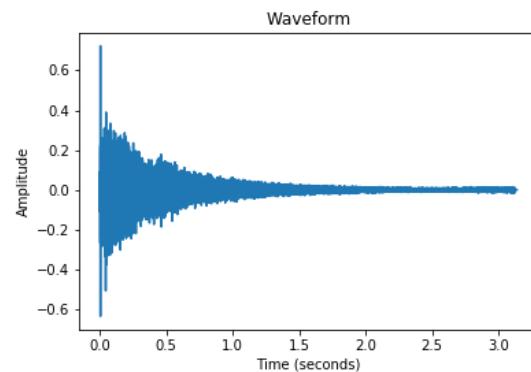
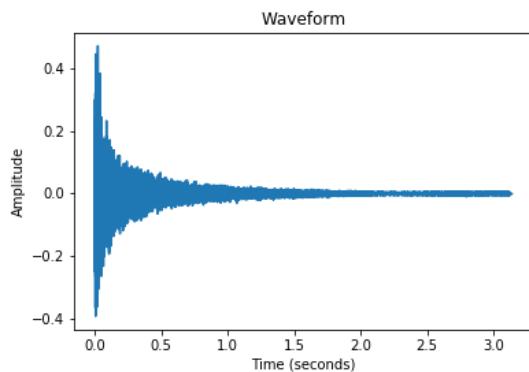
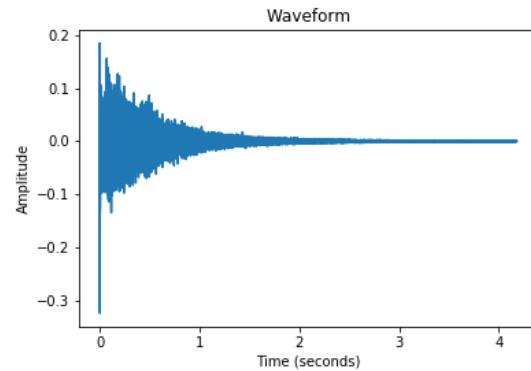
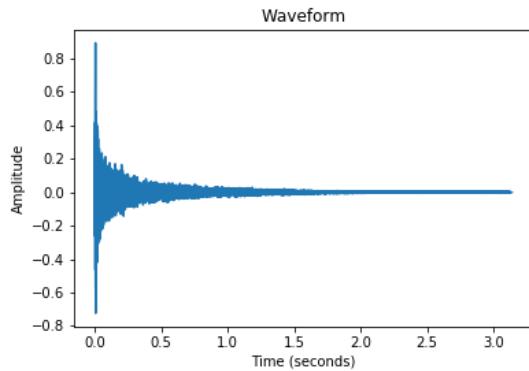
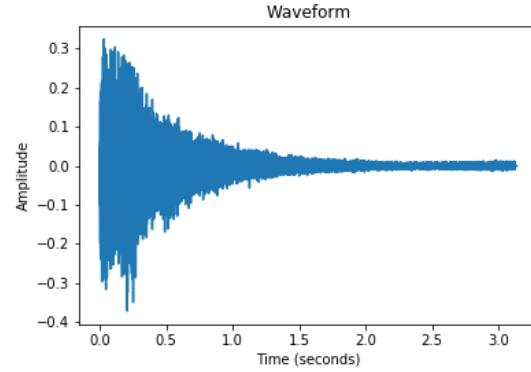
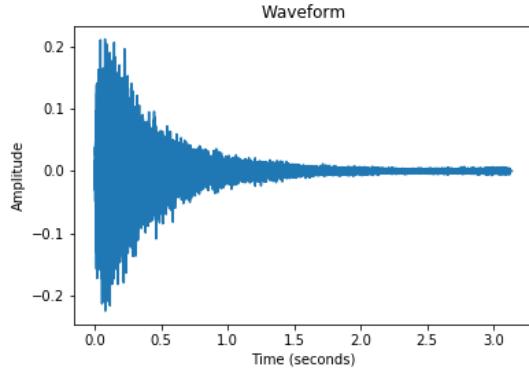


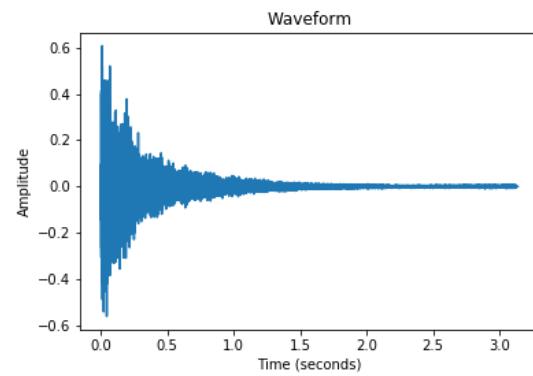
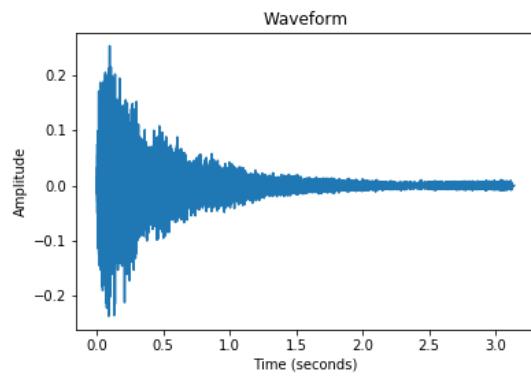
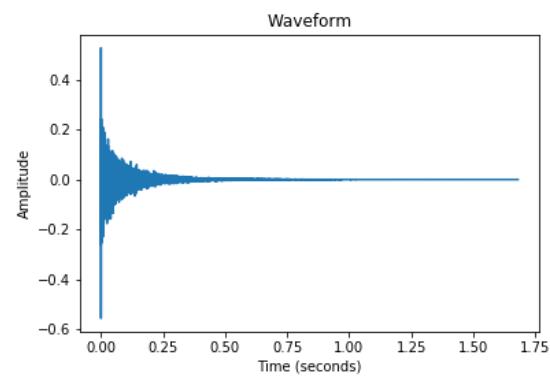
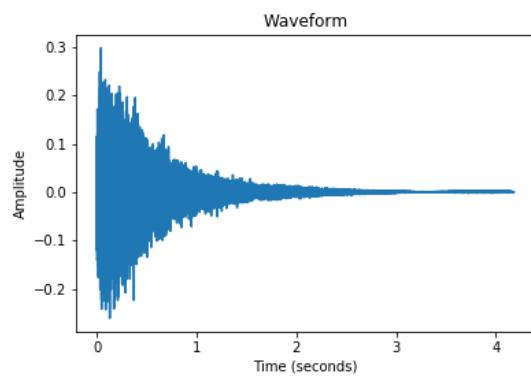
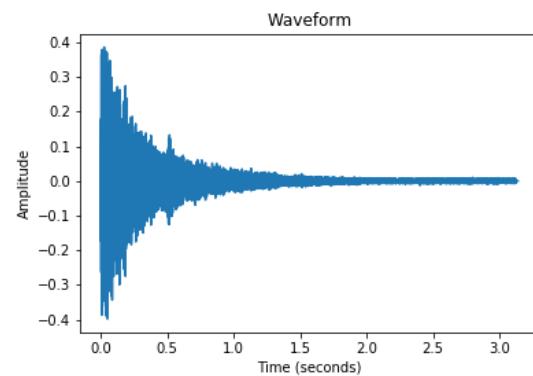
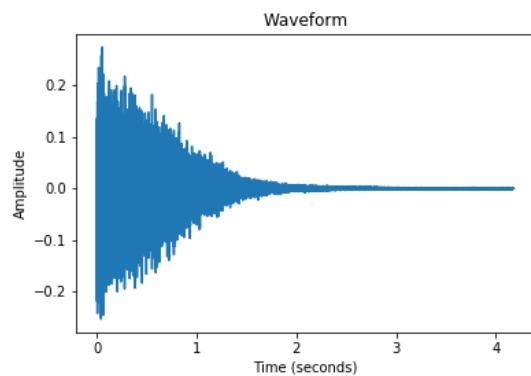
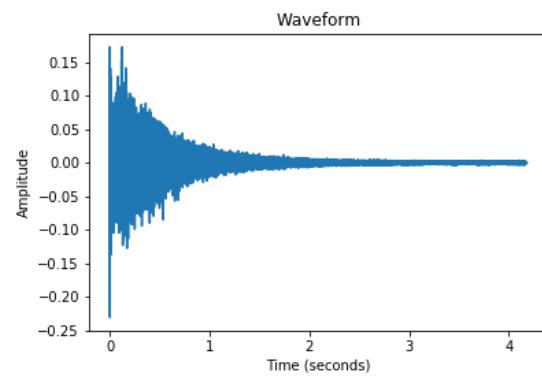
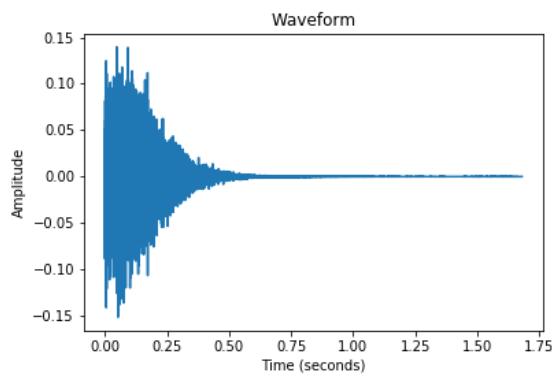


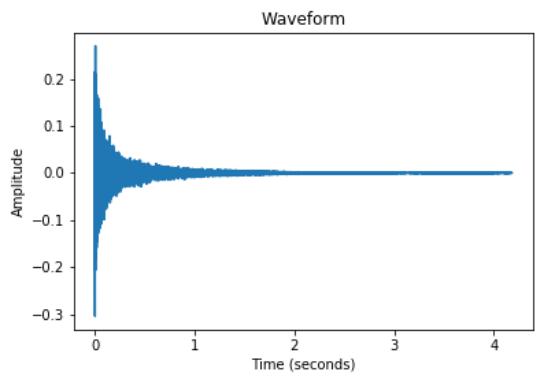
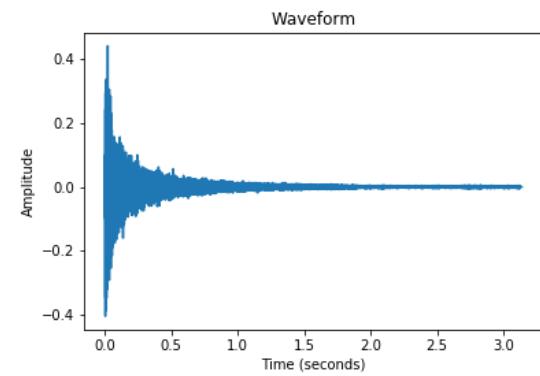
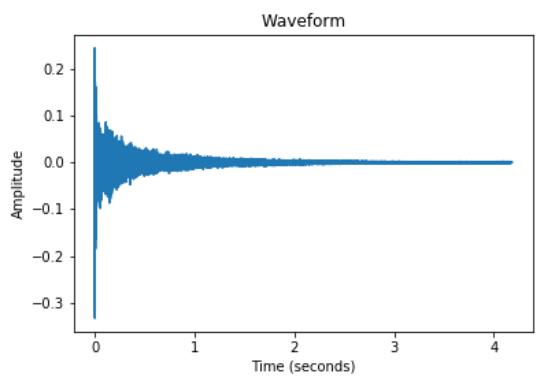
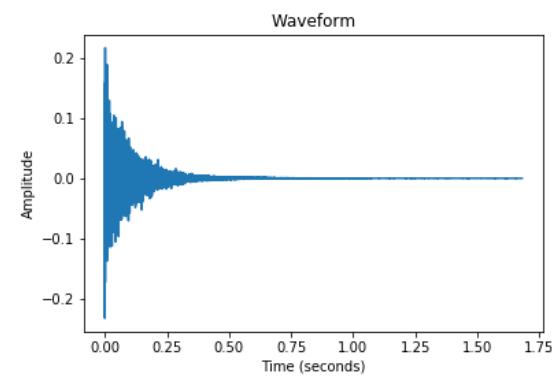
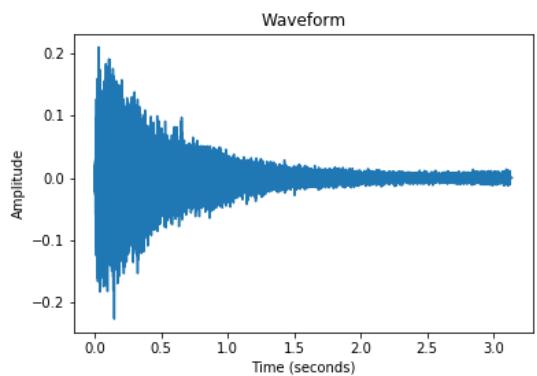
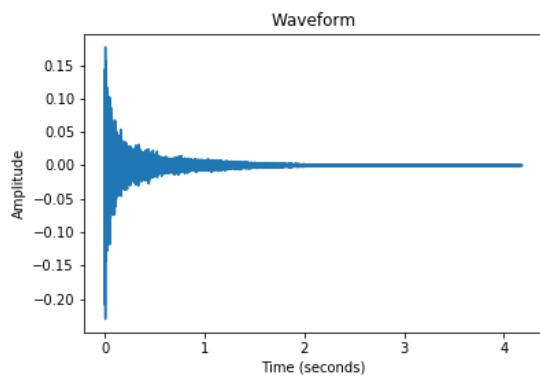
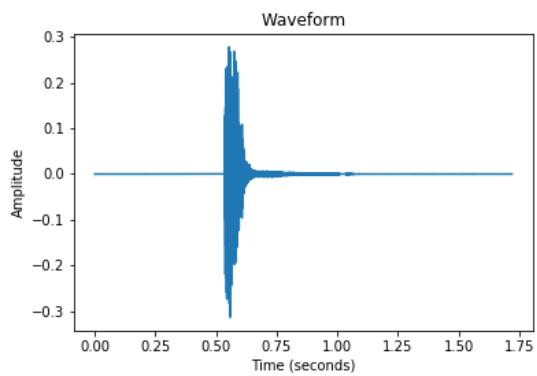
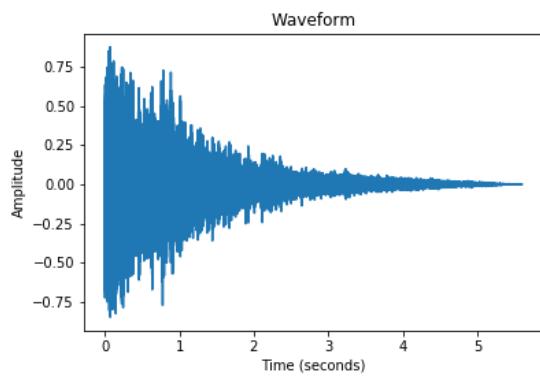


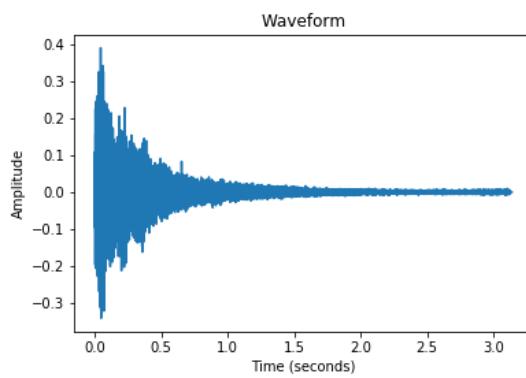
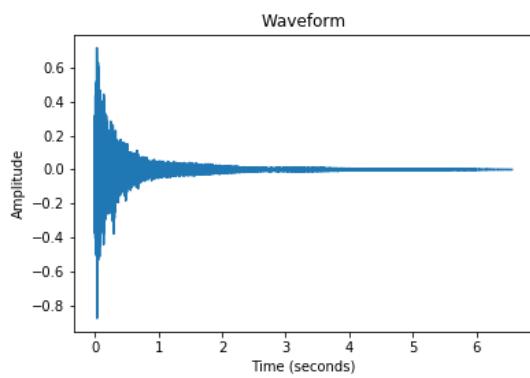
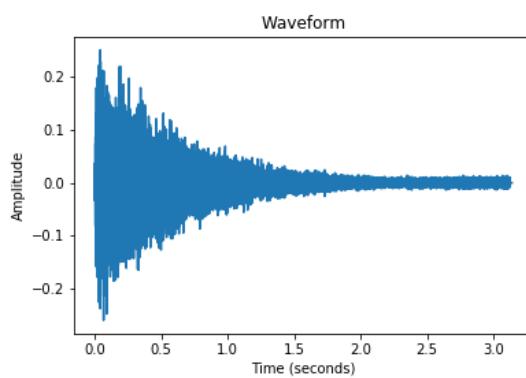
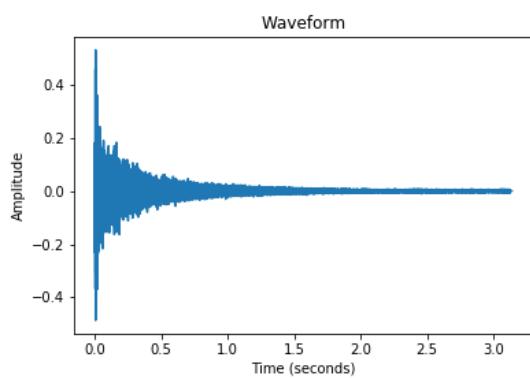
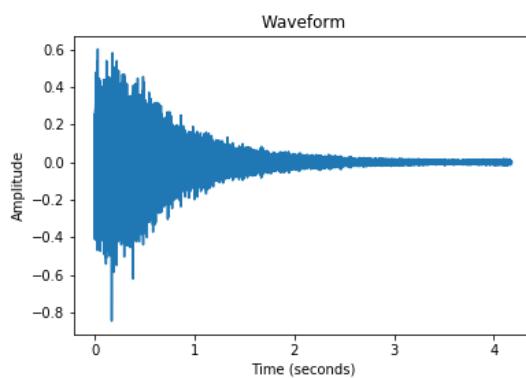
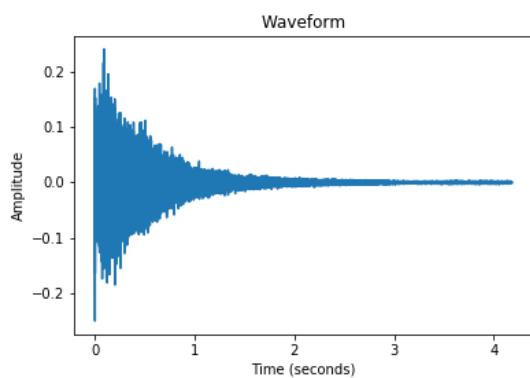
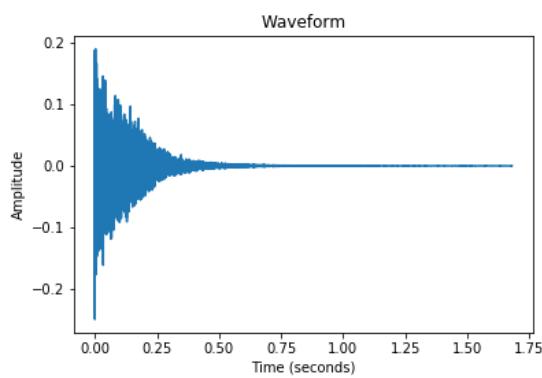
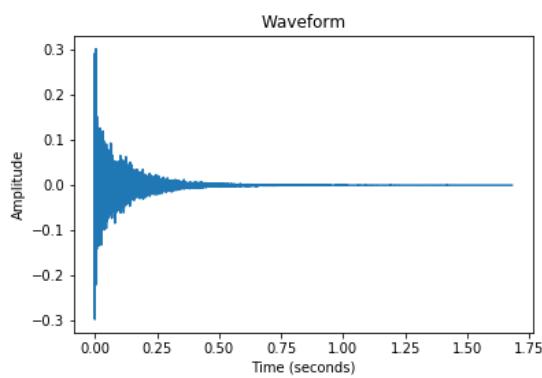


## Hi-hat

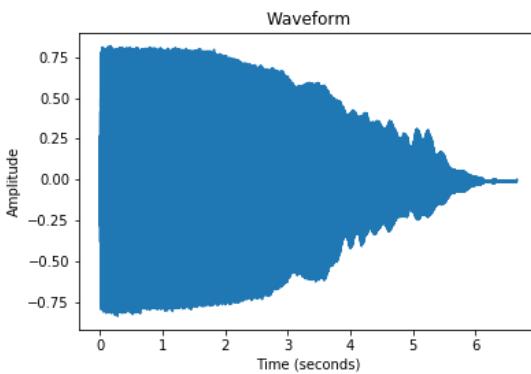
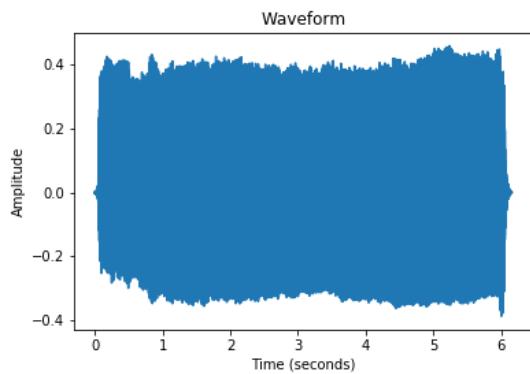
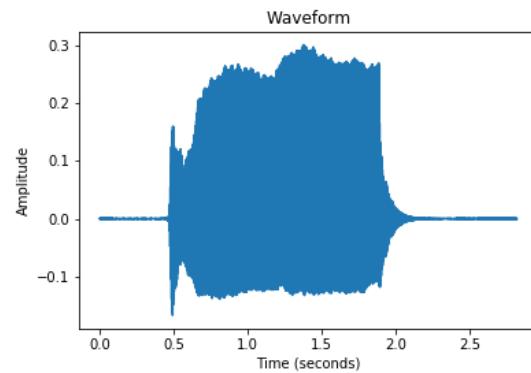
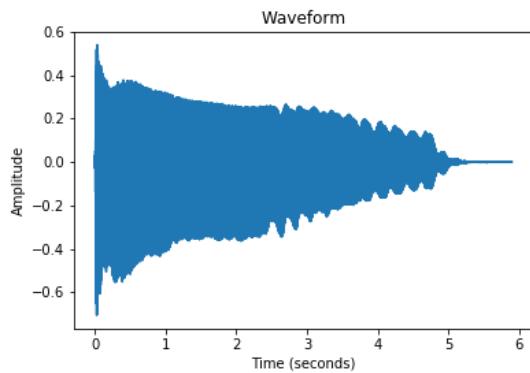
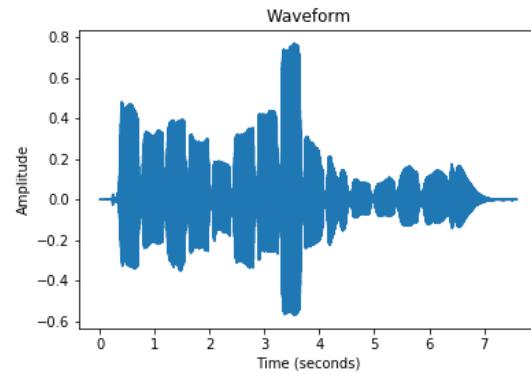
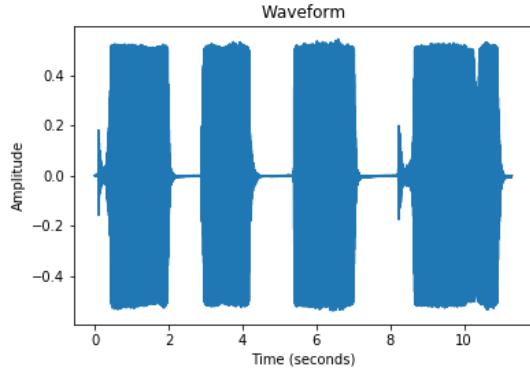


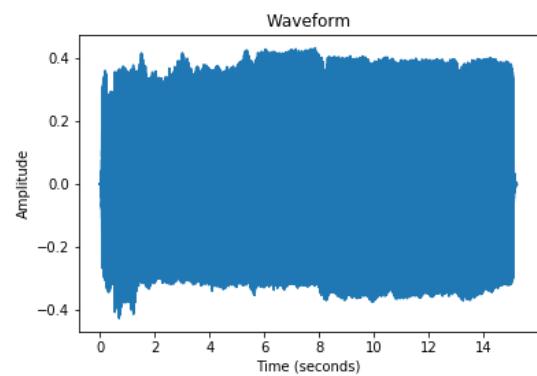
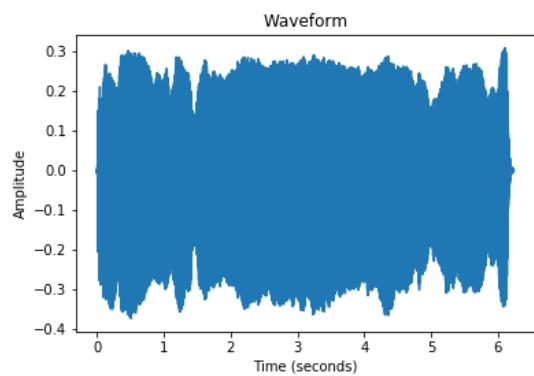
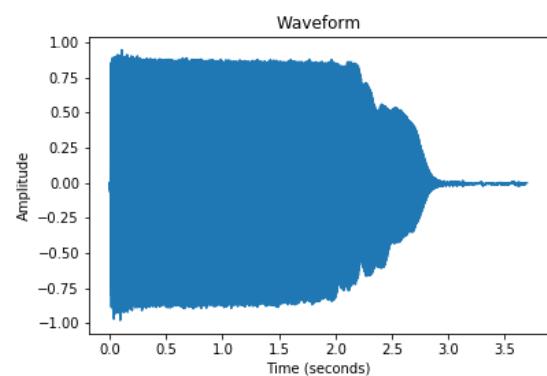
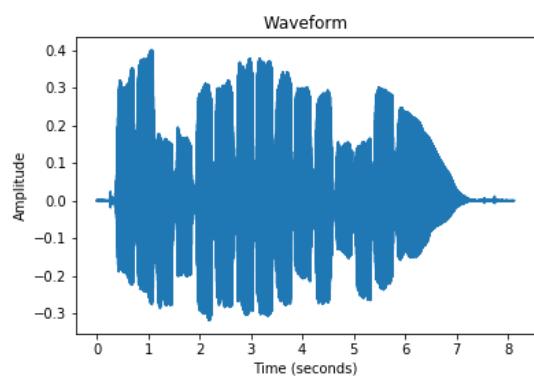
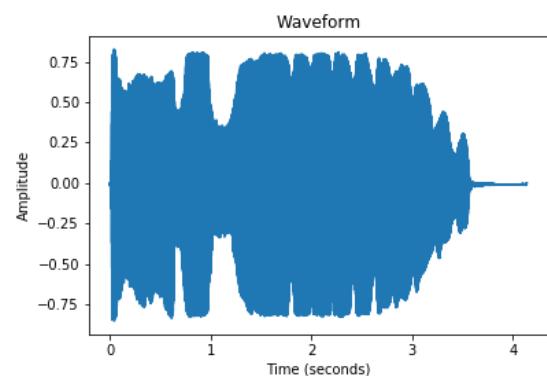
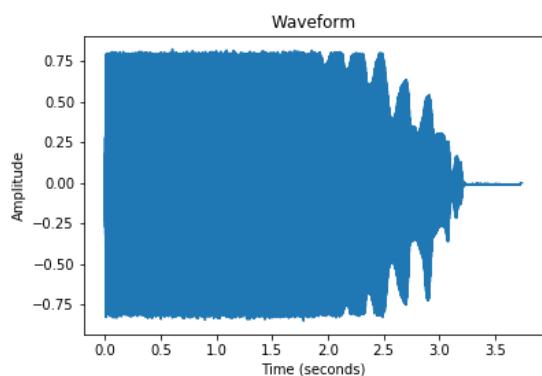
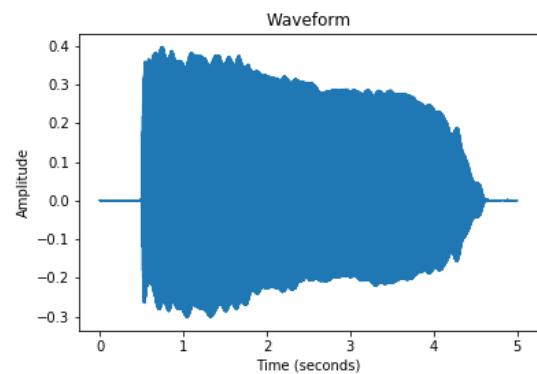
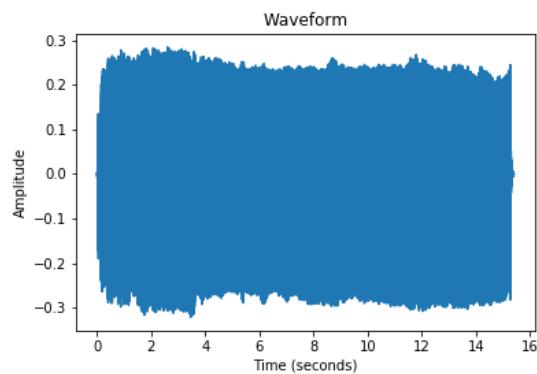


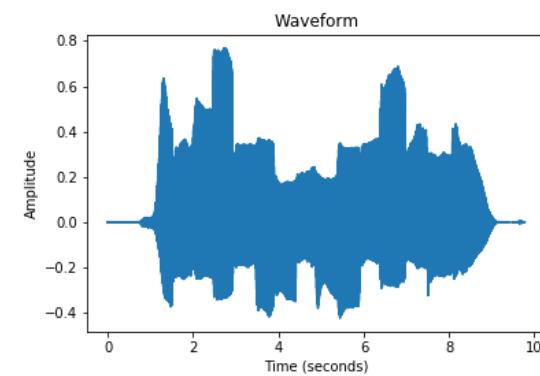
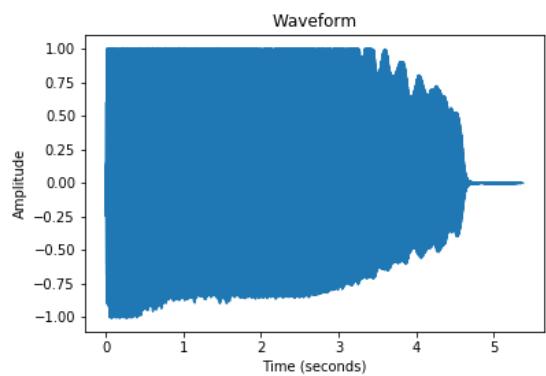
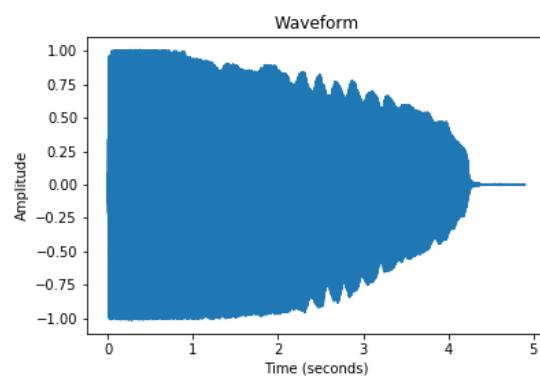
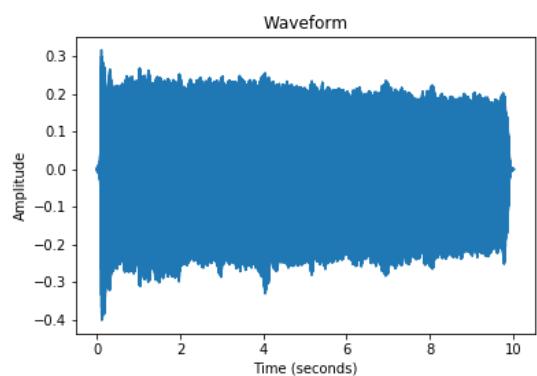
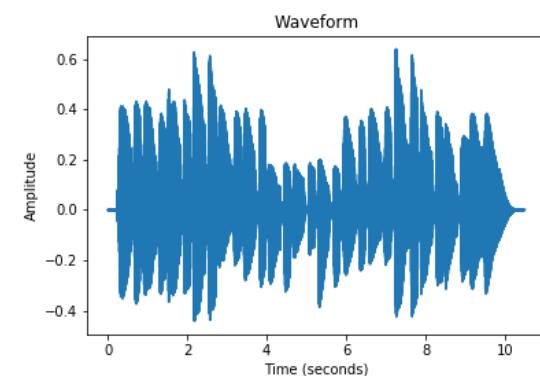
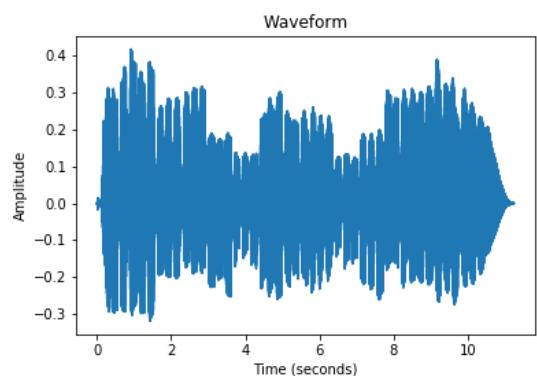
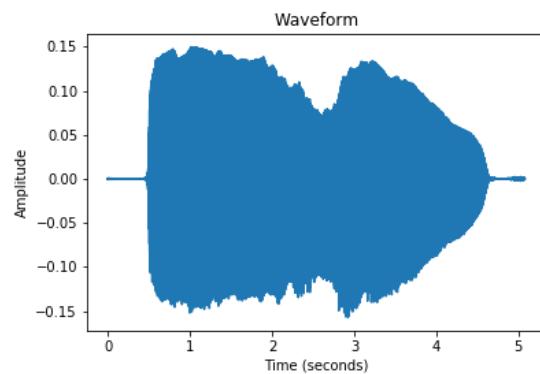
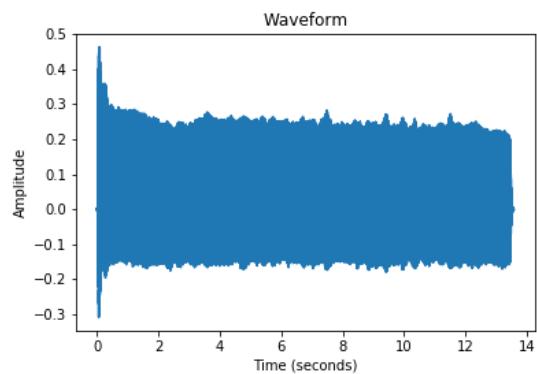


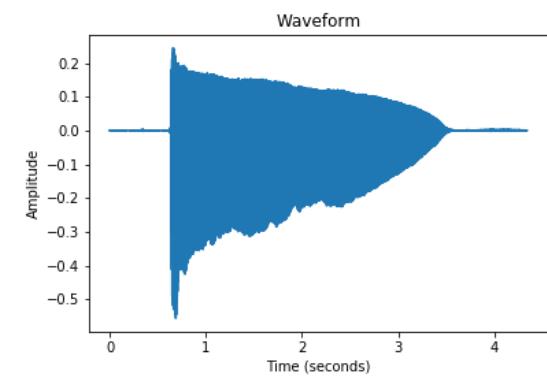
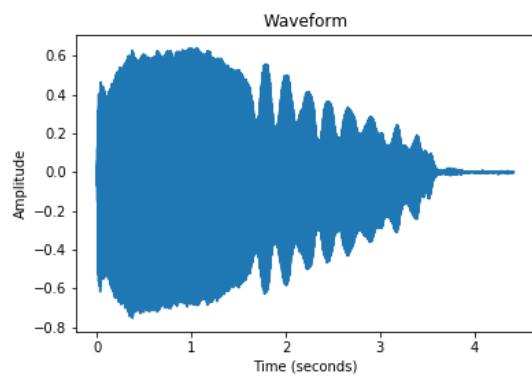
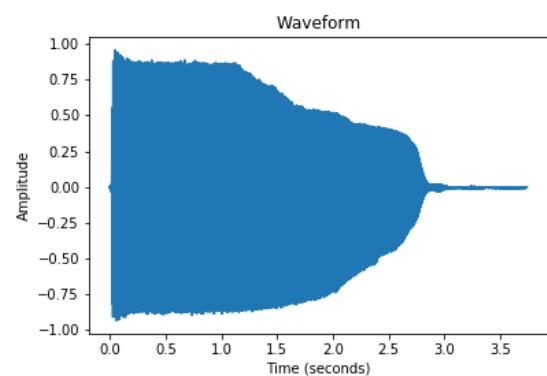
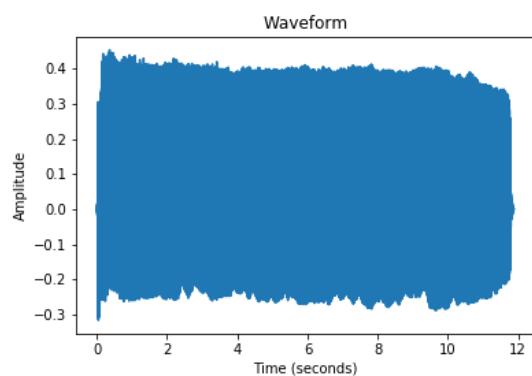
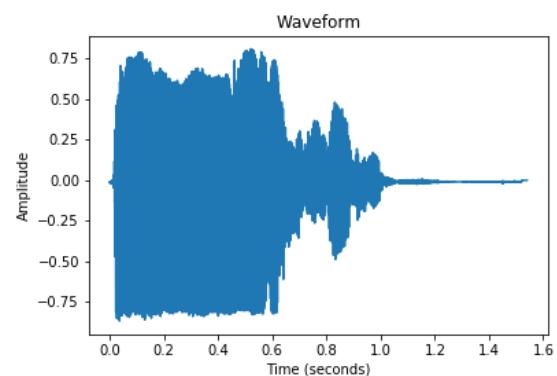
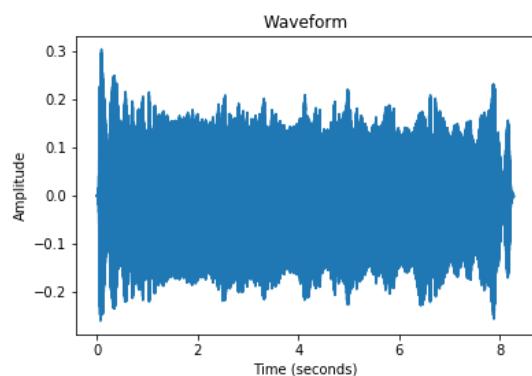
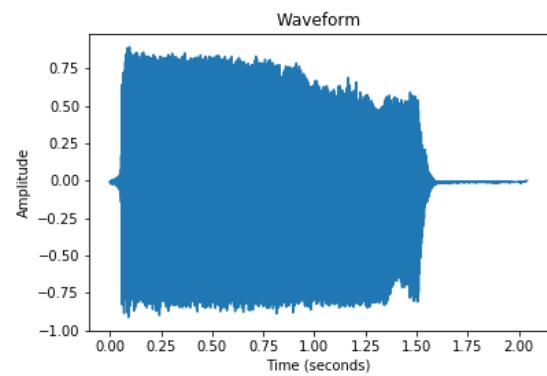
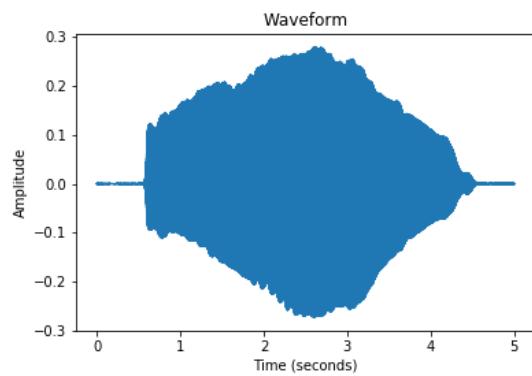


## Saxophone

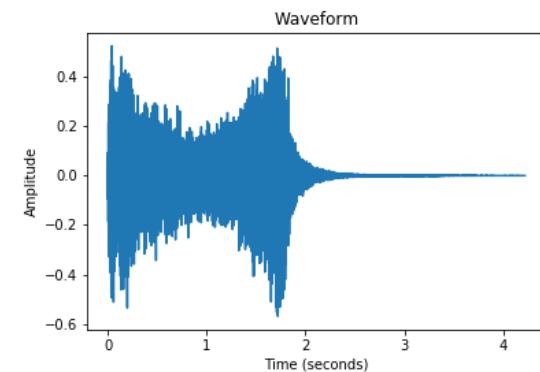
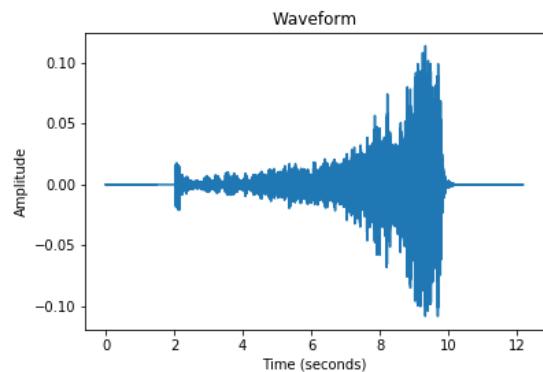
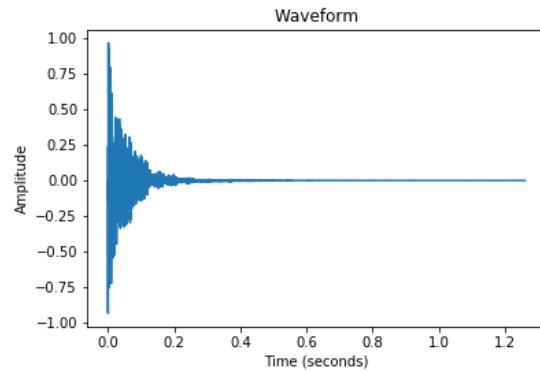
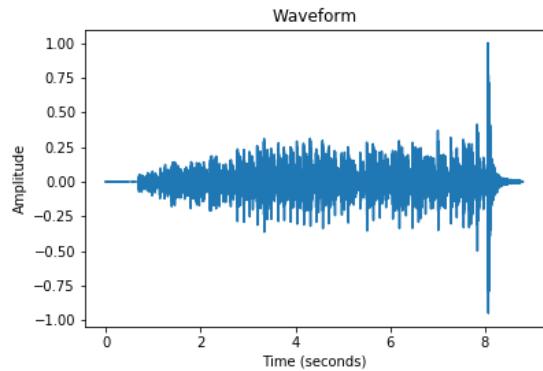
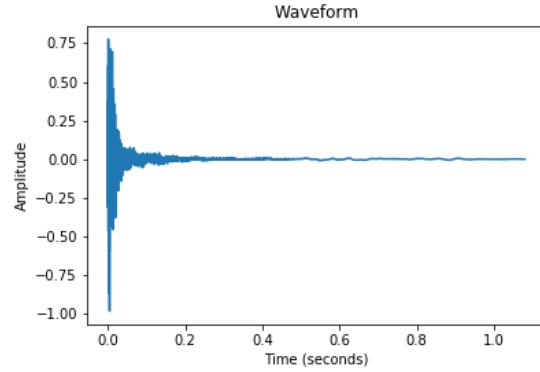
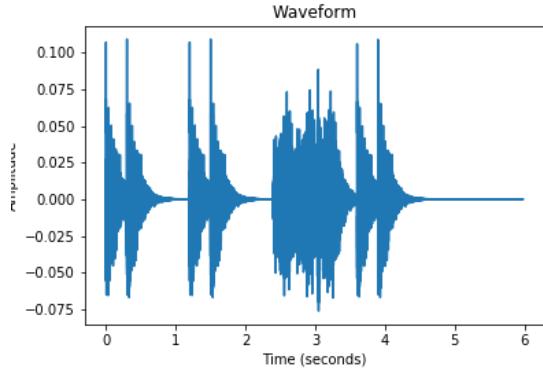


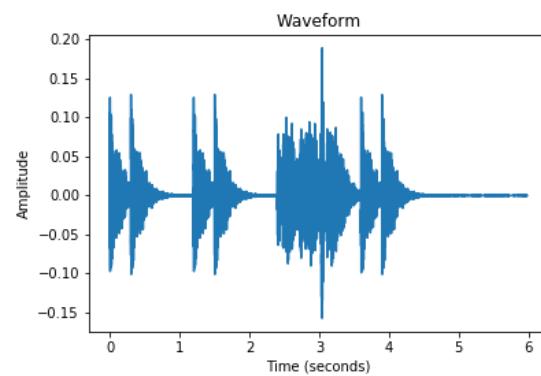
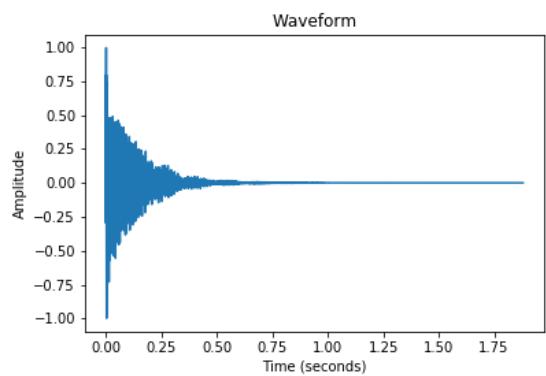
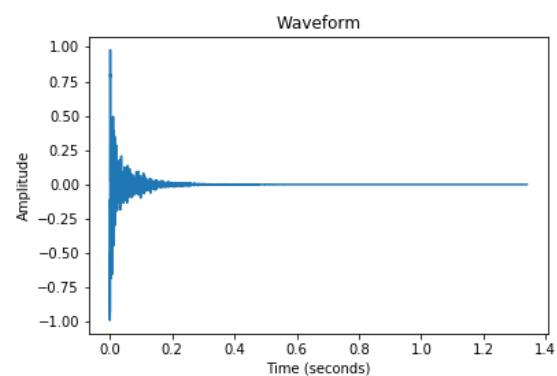
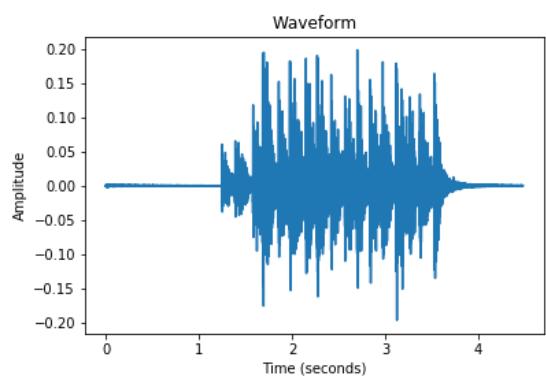
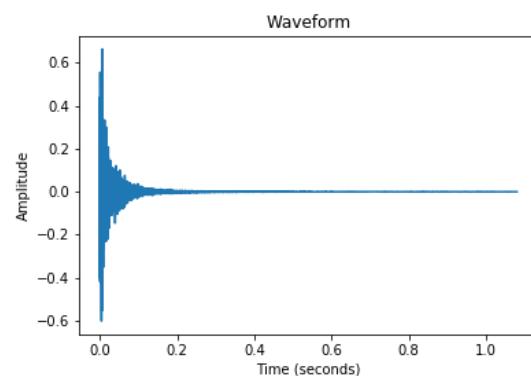
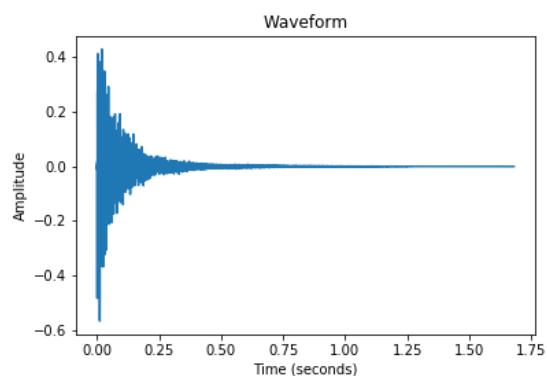
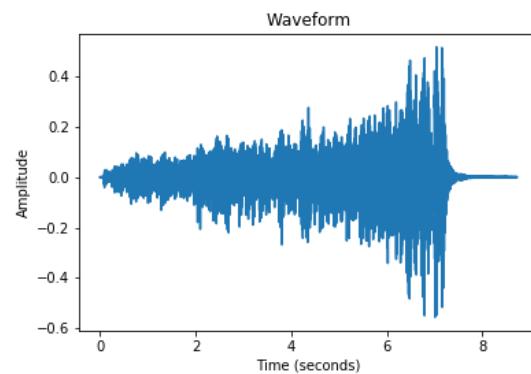
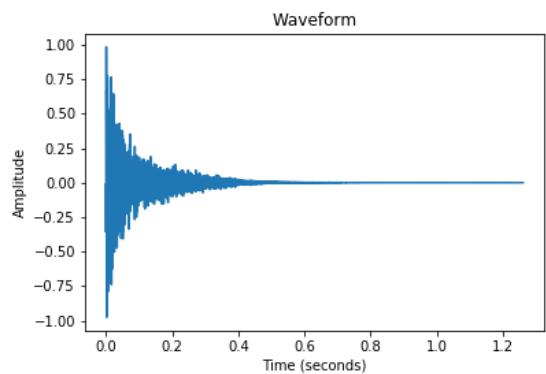


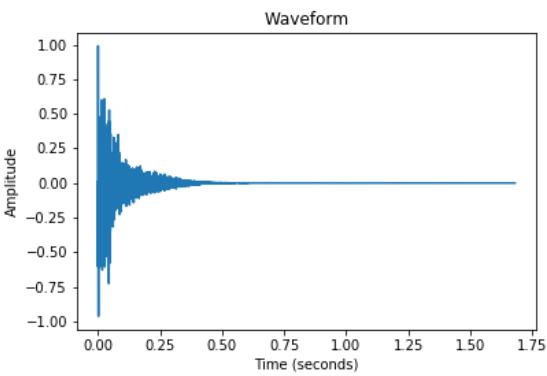
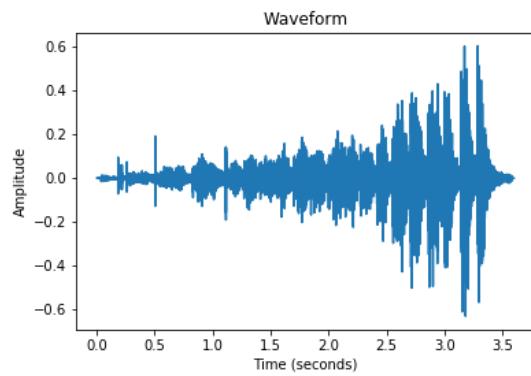
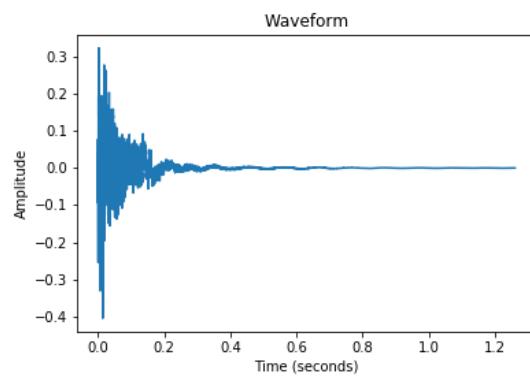
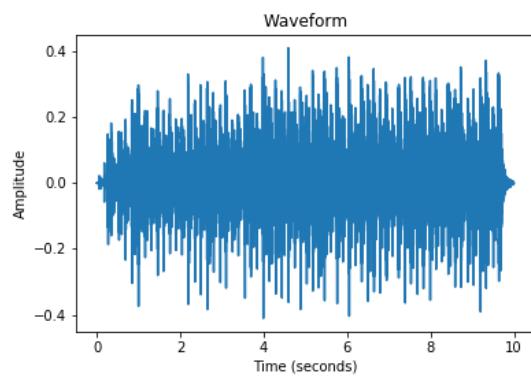
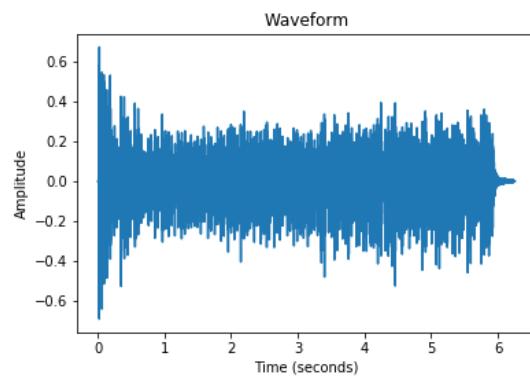
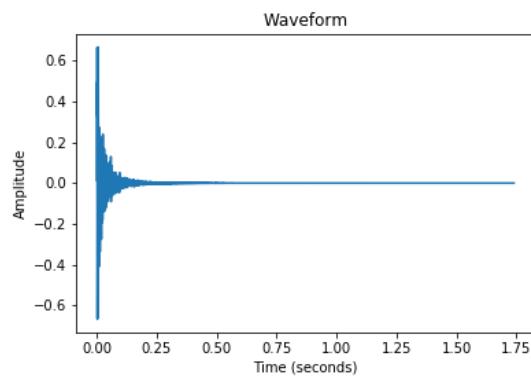
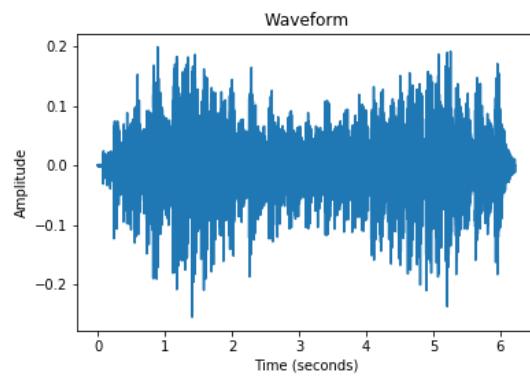
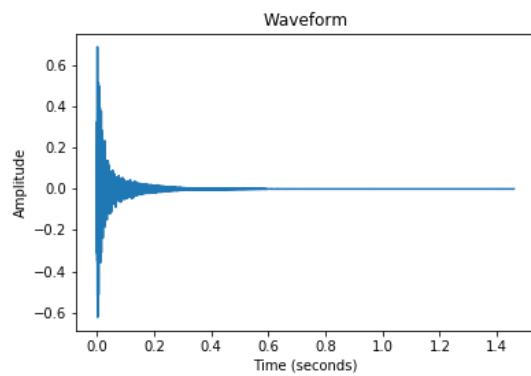


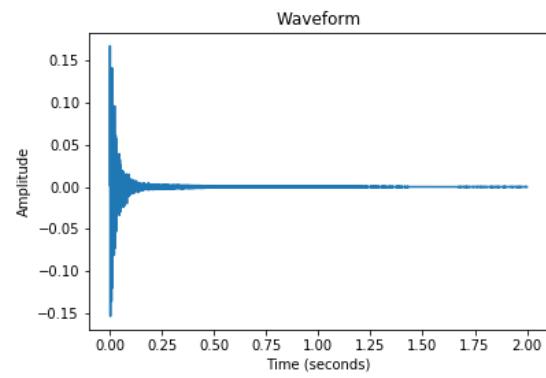
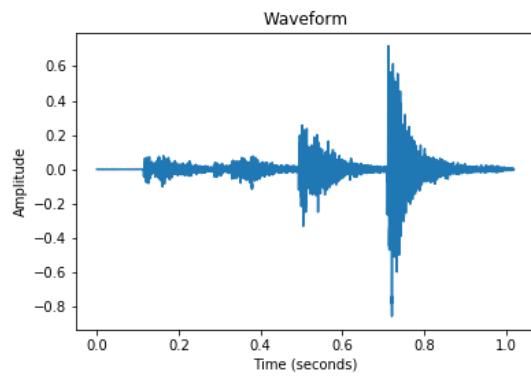
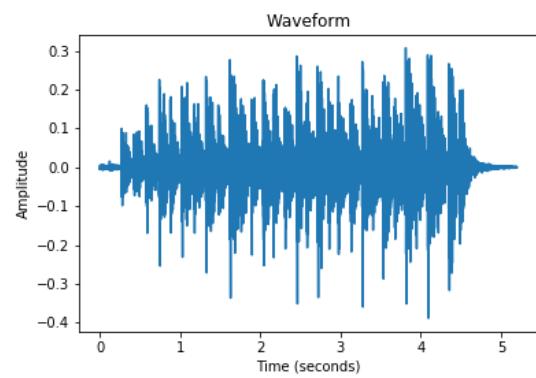
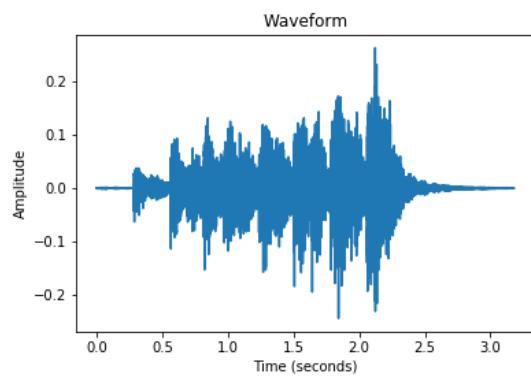
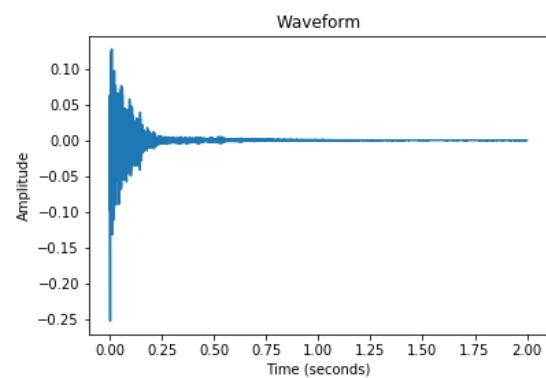
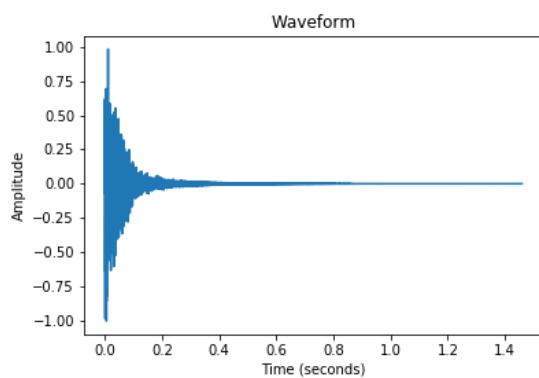
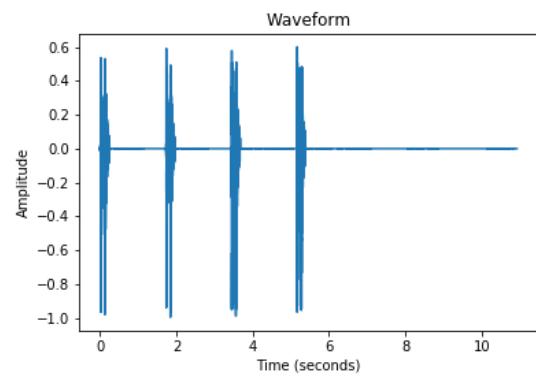
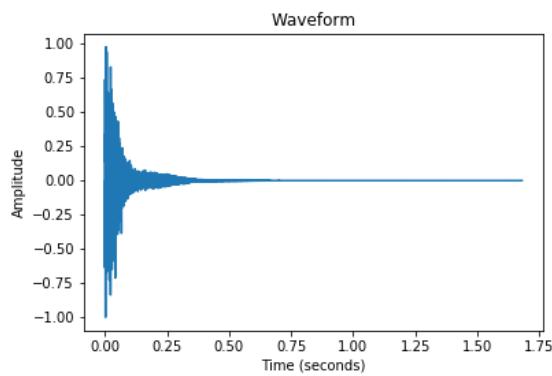


## Snare Drum

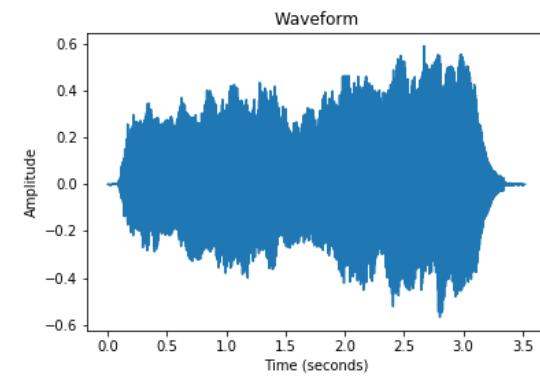
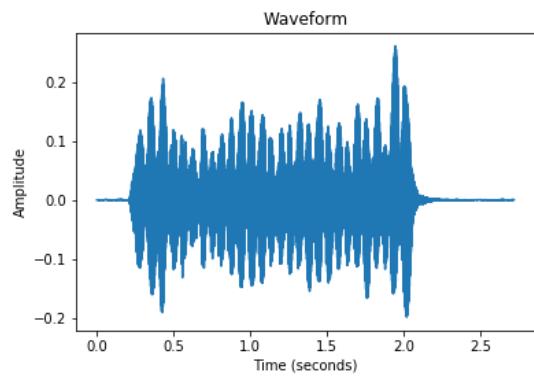
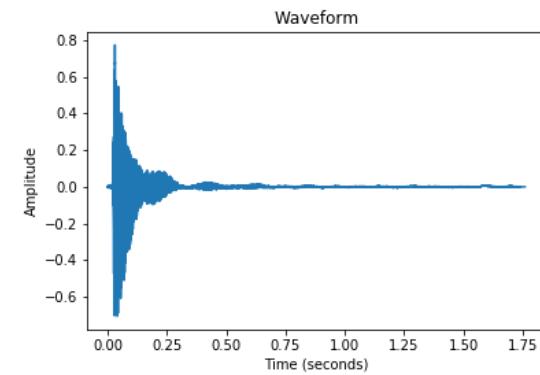
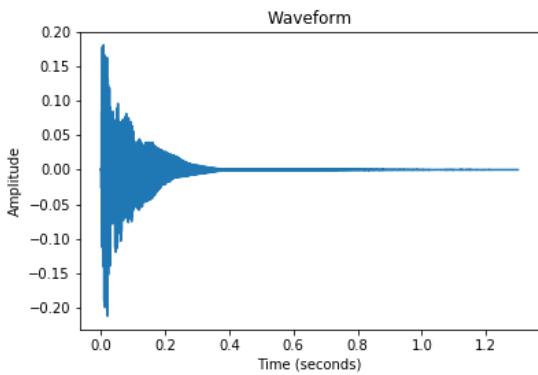
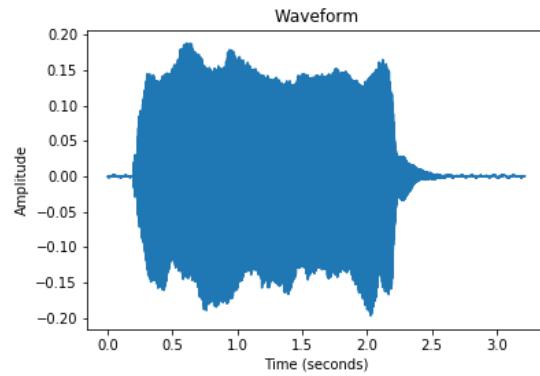
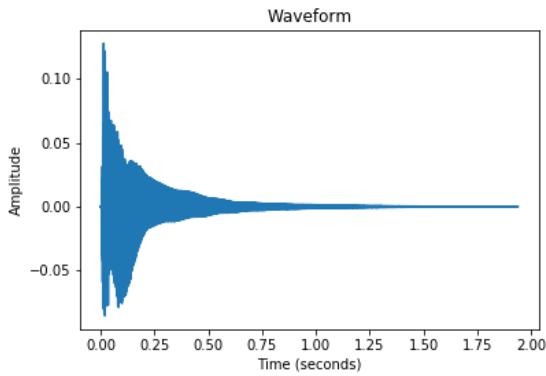


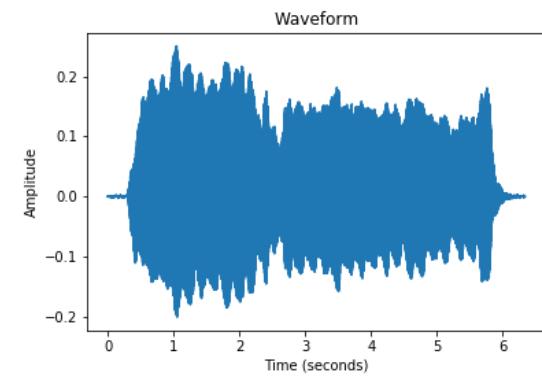
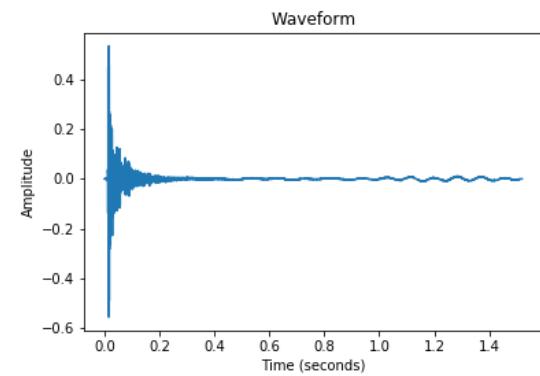
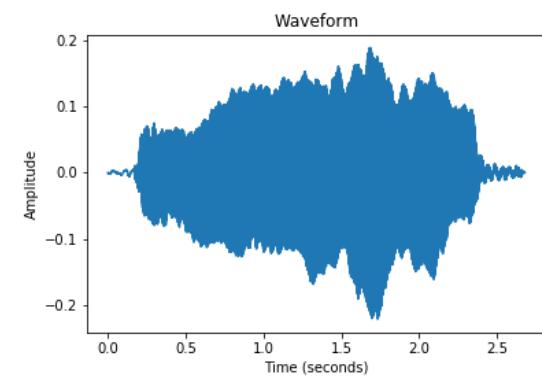
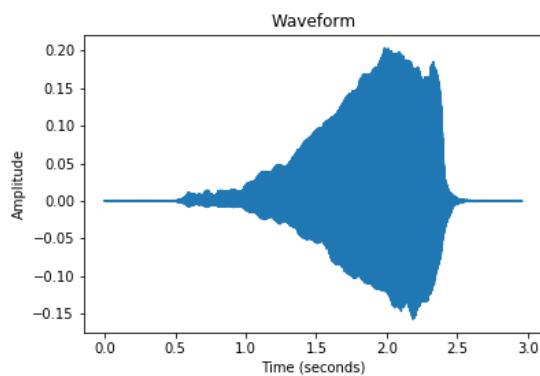
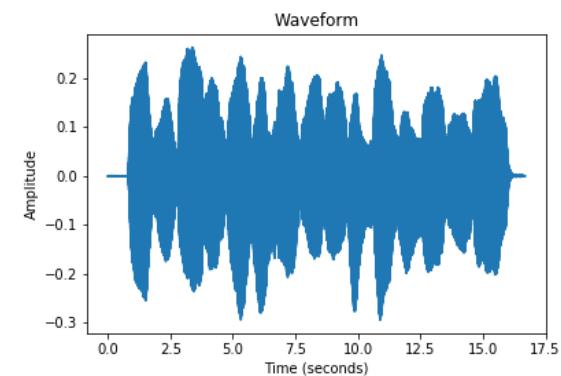
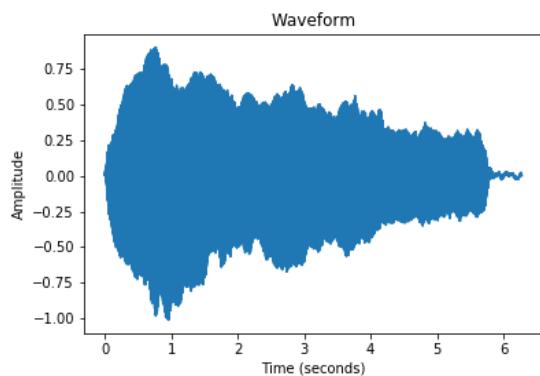
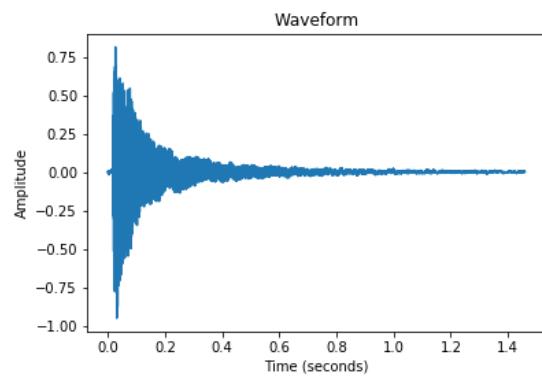
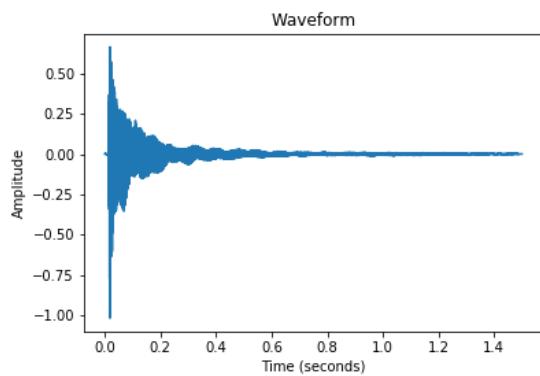


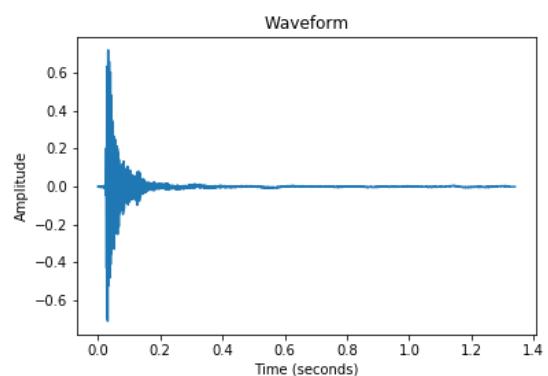
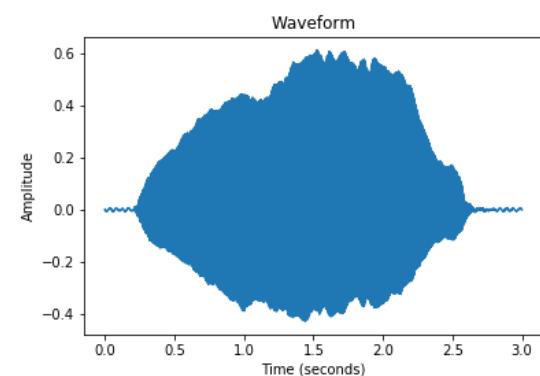
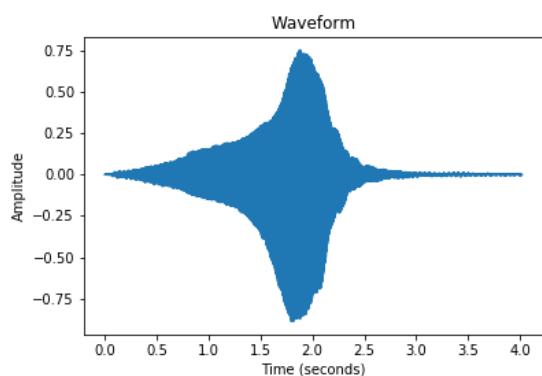
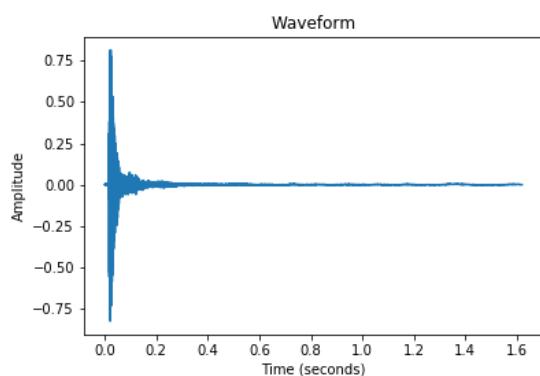
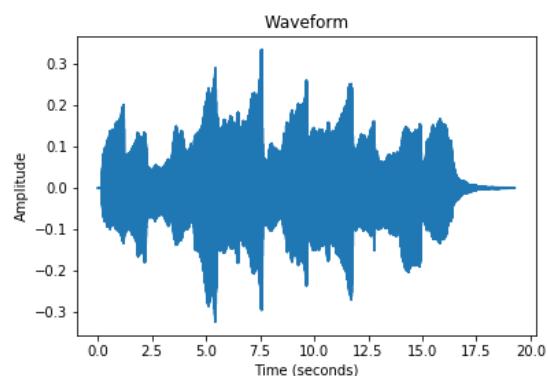
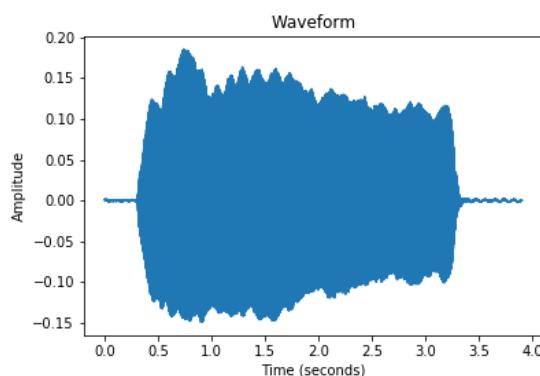
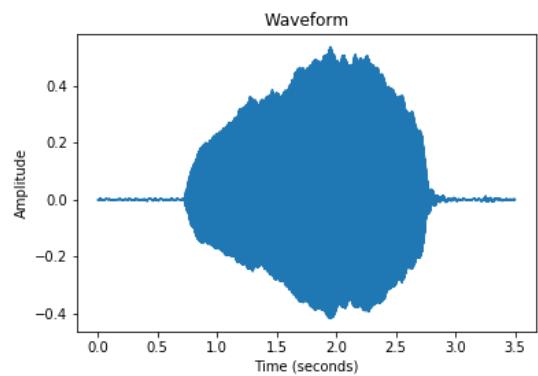
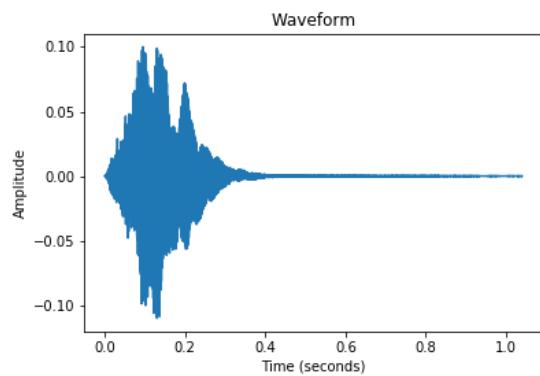


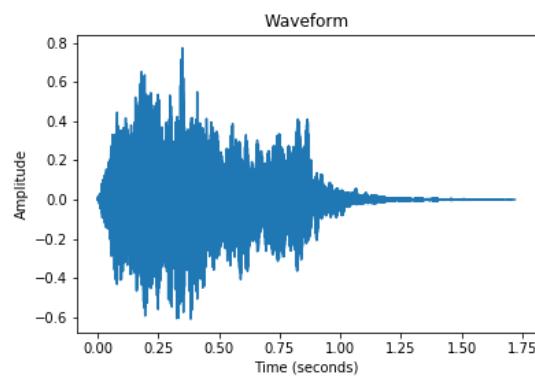
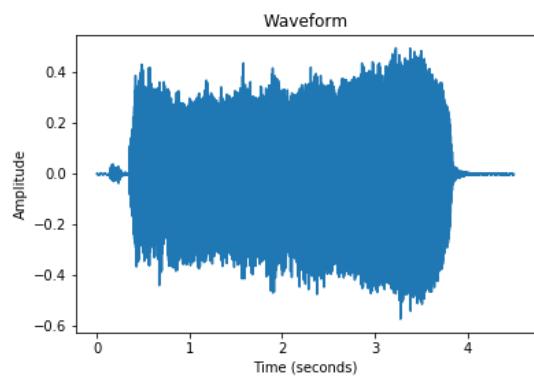
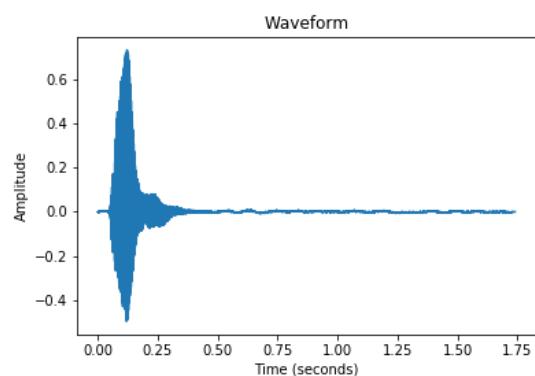
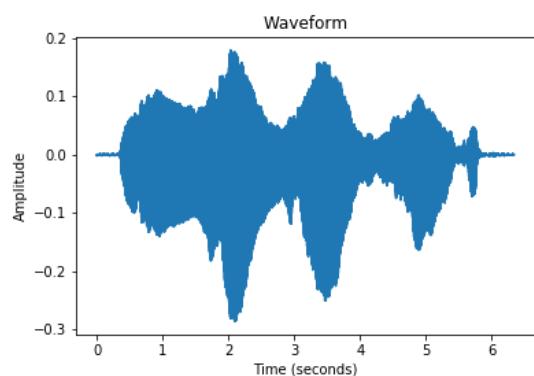
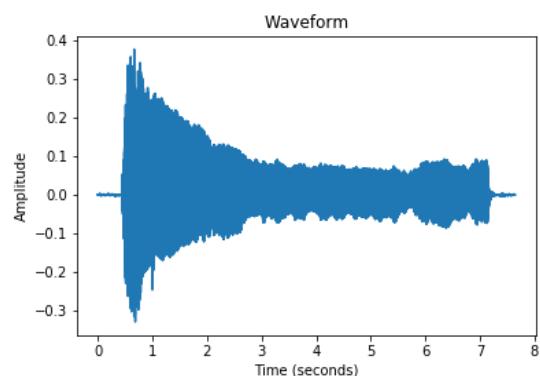
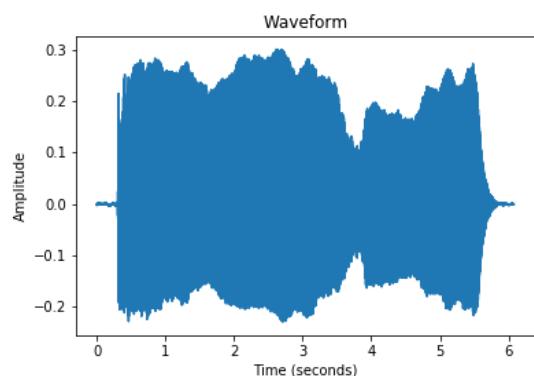
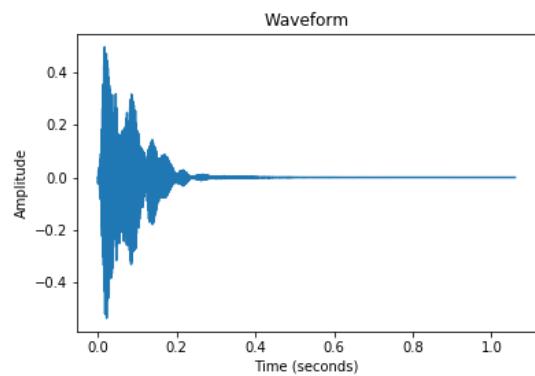
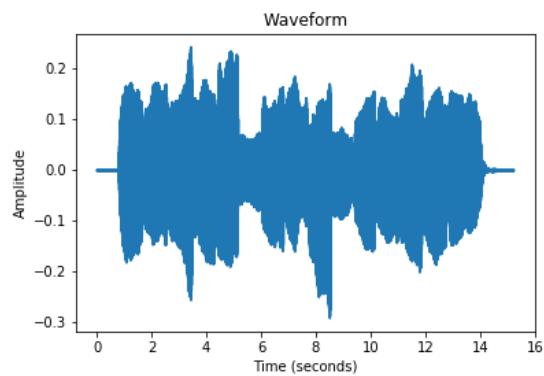


## Violin/Fiddle





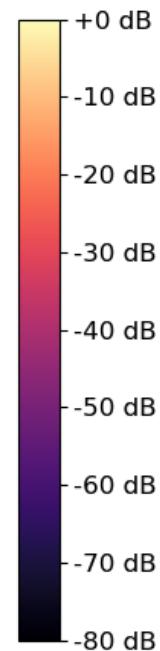




## Mel Spectrograms

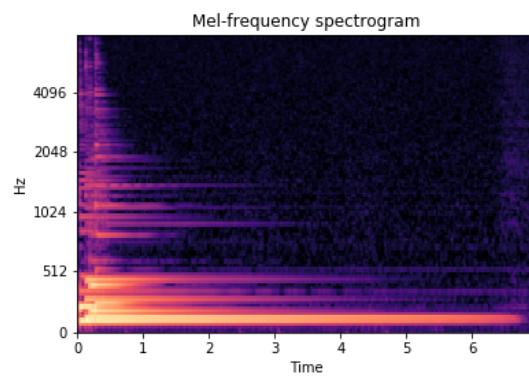
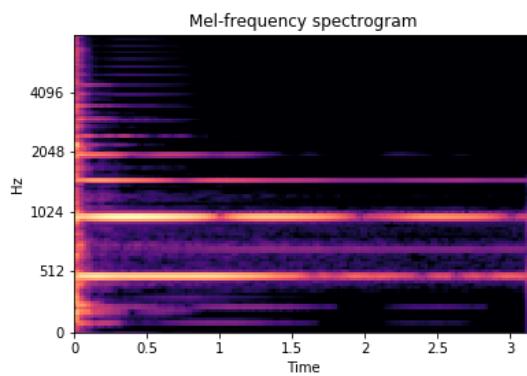
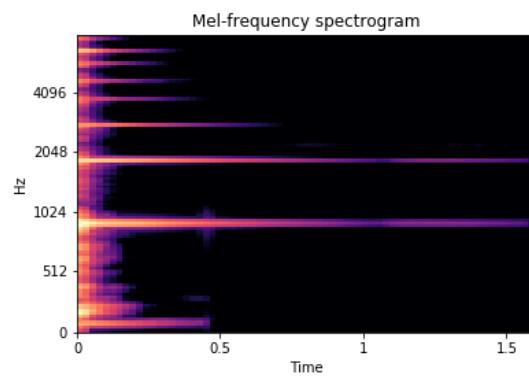
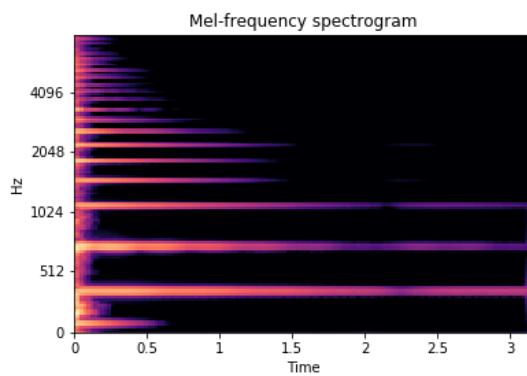
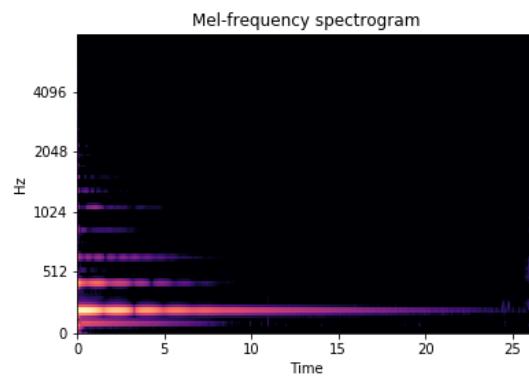
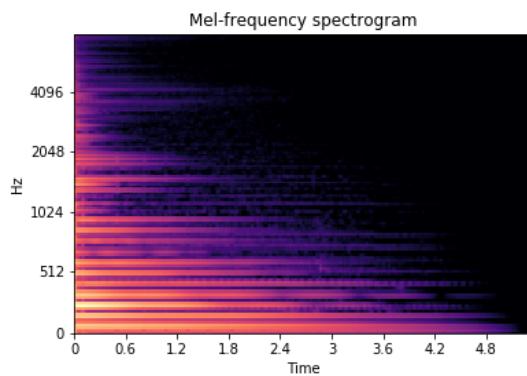
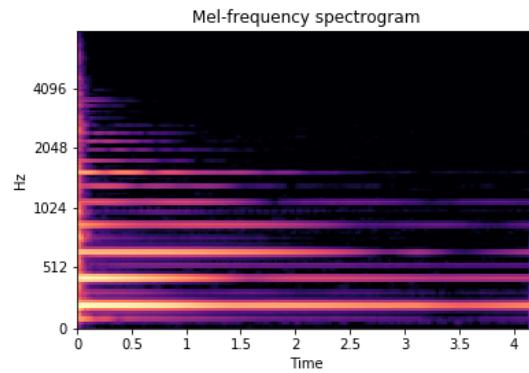
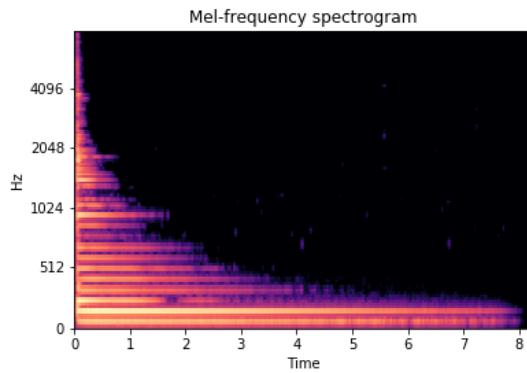
The generated Mel Spectrograms for each class are shown below. They are shown in order such that the first spectrogram corresponds to the first waveform shown above and so on. As with the waveforms, out of the 30 audio samples per class, the first 15 spectrograms are shown on the left side and the remaining 15 spectrograms are shown on the right side.

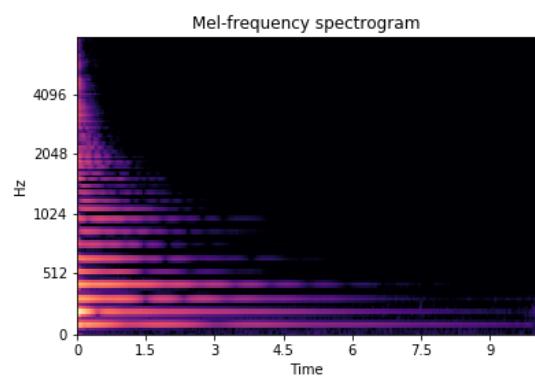
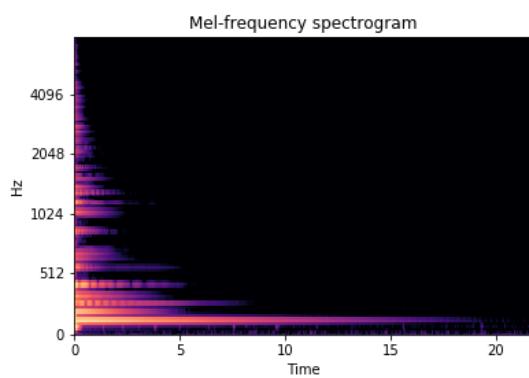
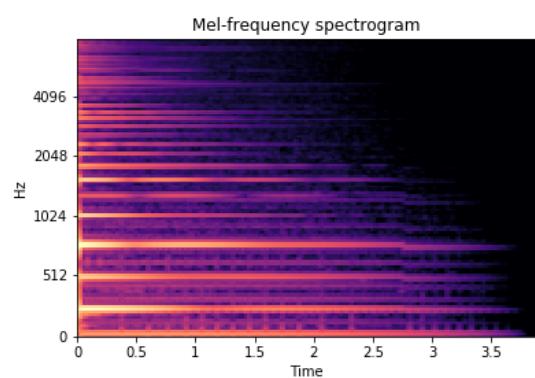
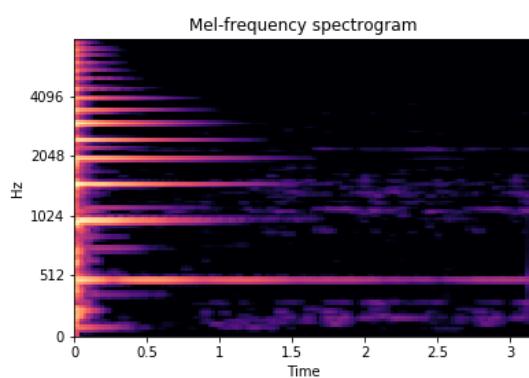
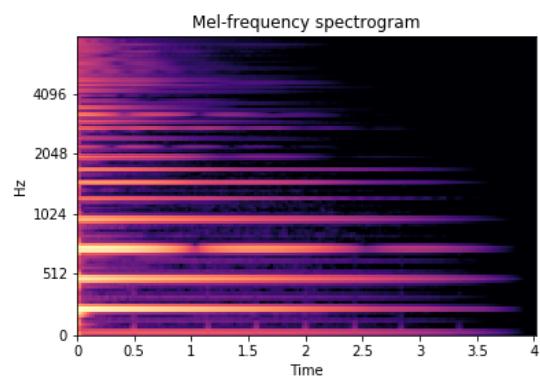
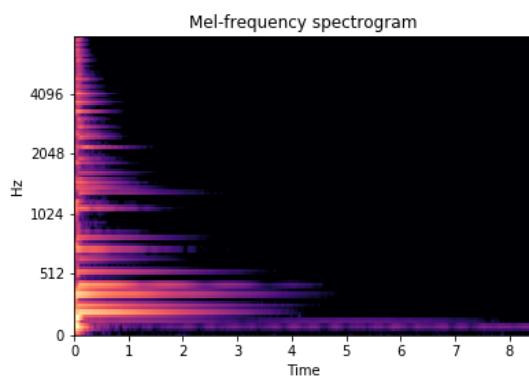
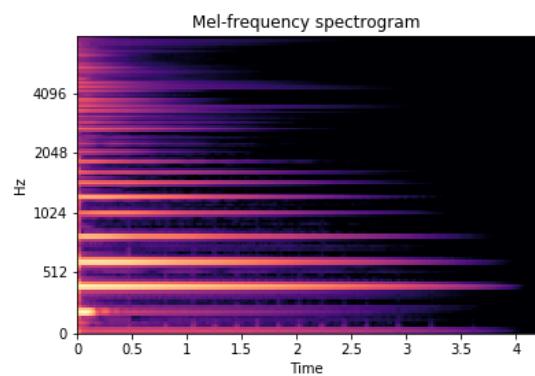
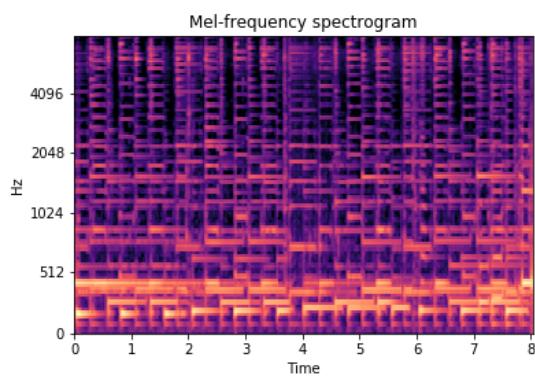
The color bar scale for all of the Mel Spectrogram's is shown to the right:

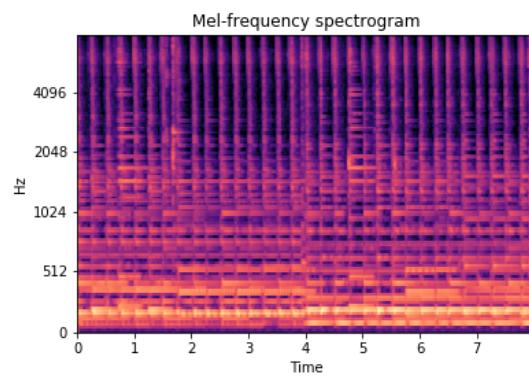
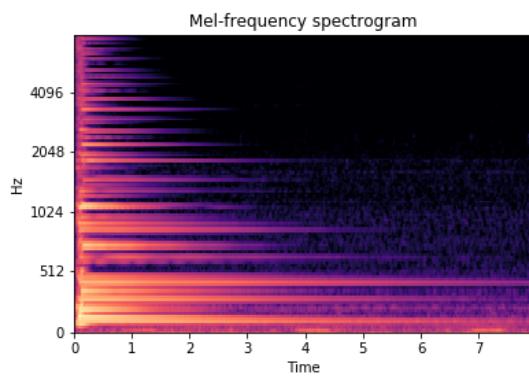
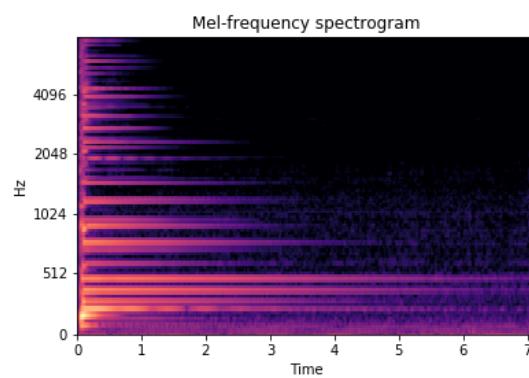
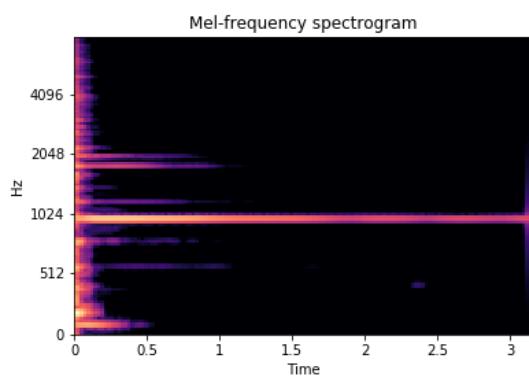
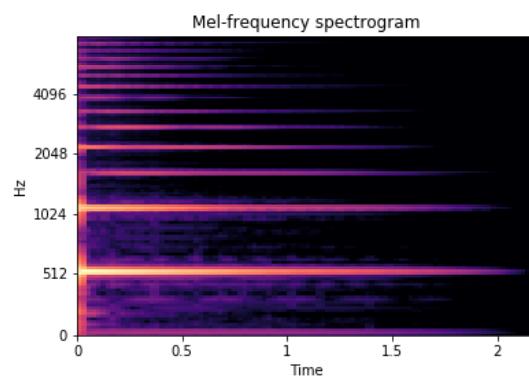
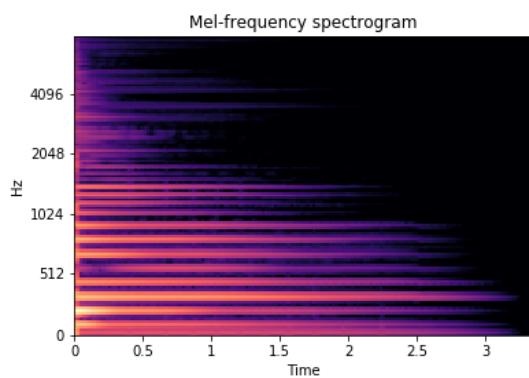
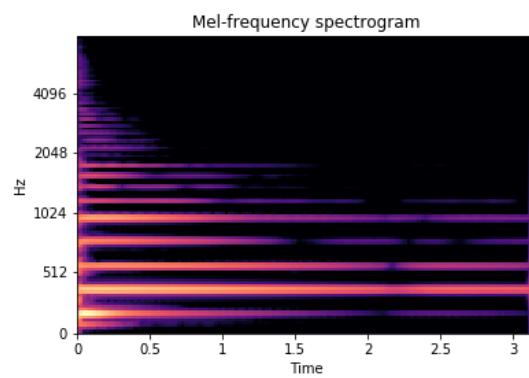
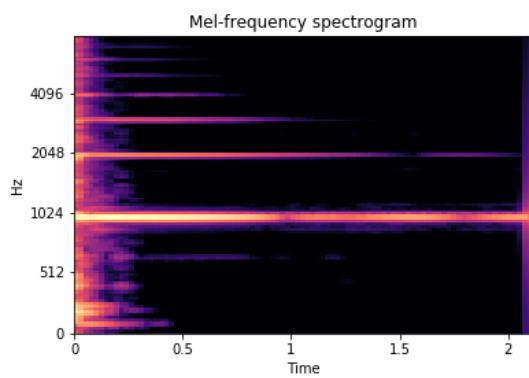


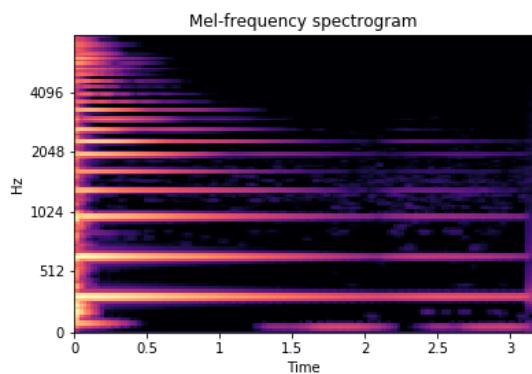
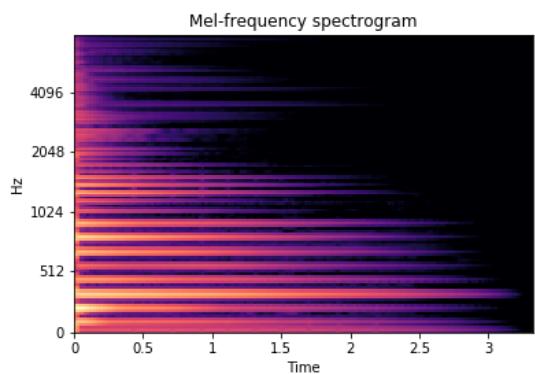
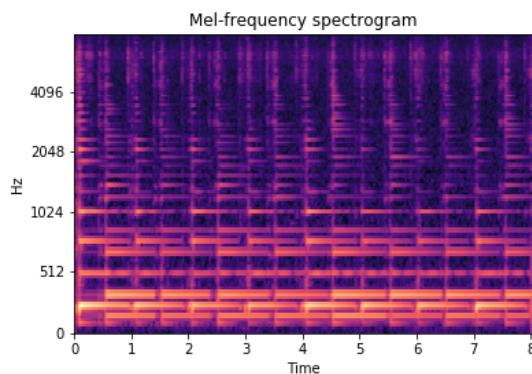
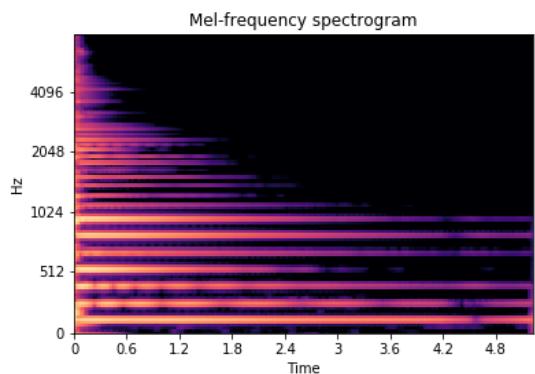
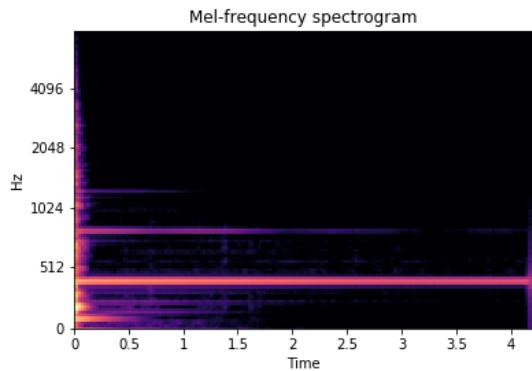
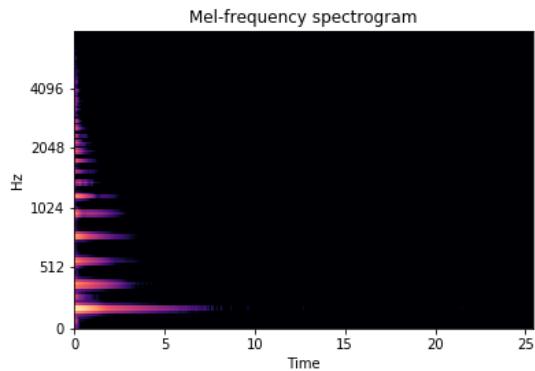
Color bar scale

## Acoustic Guitar

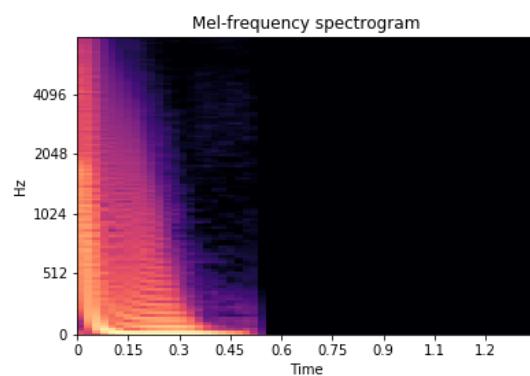
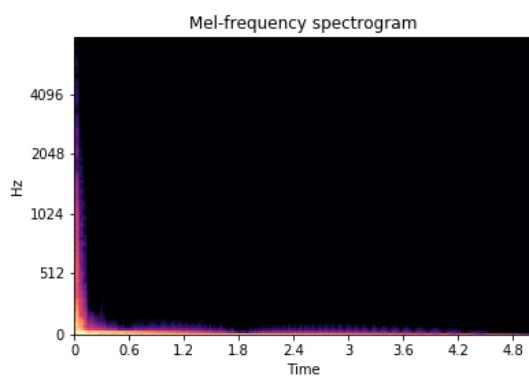
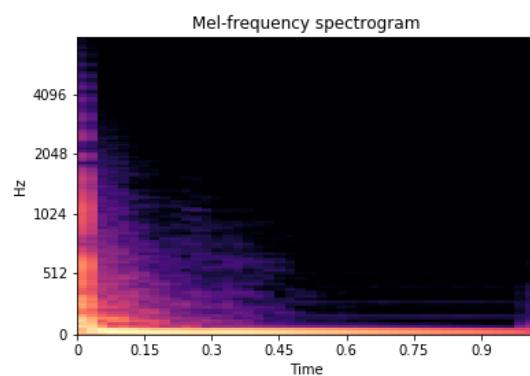
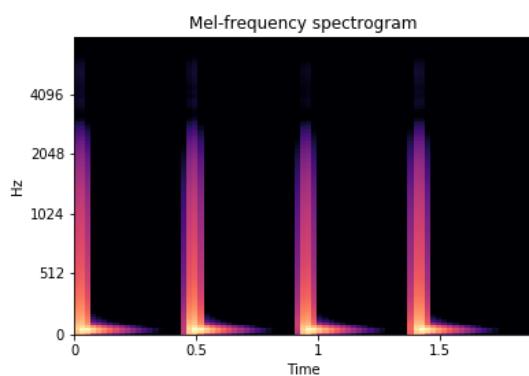
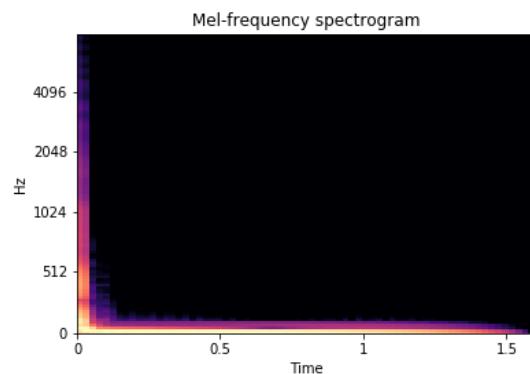
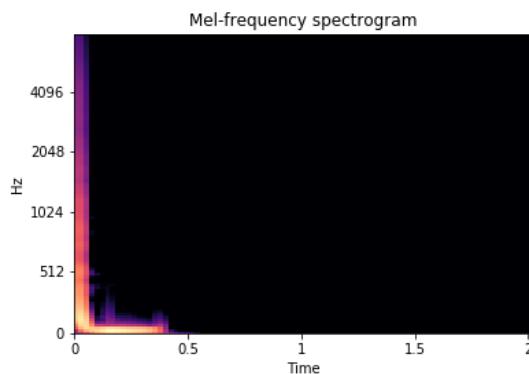
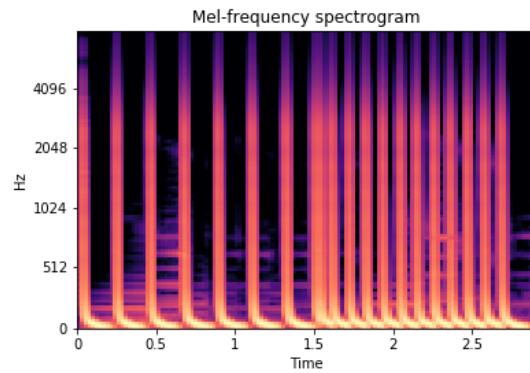
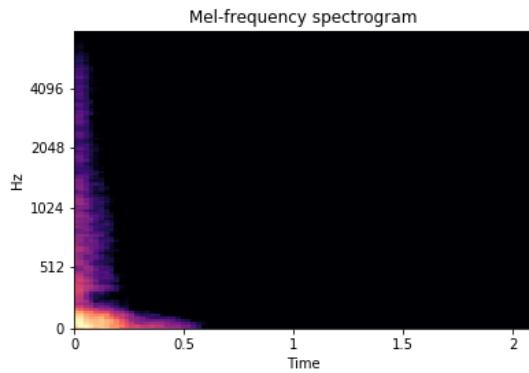


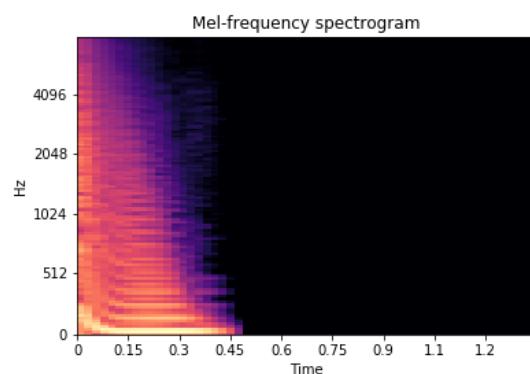
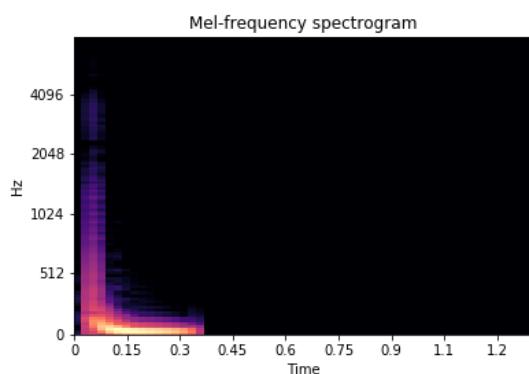
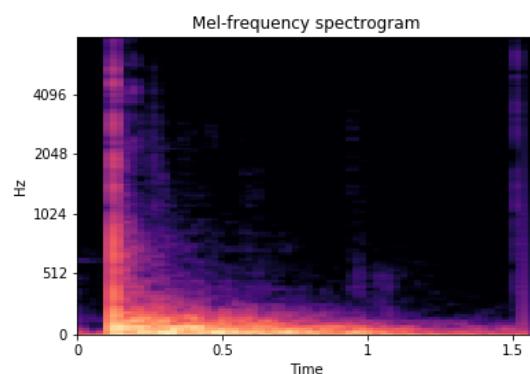
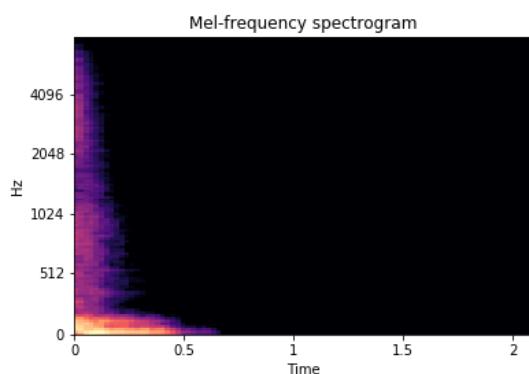
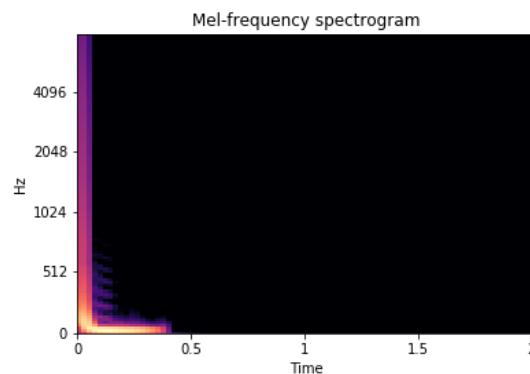
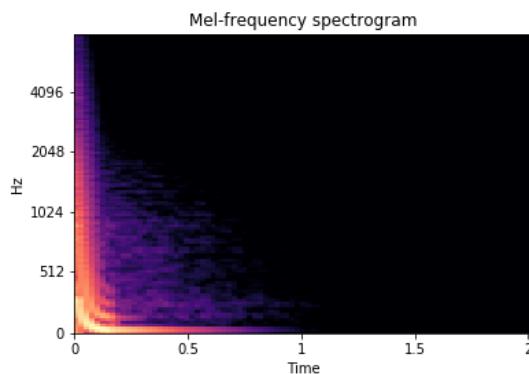
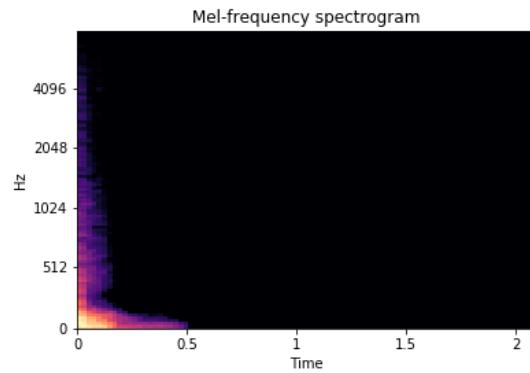
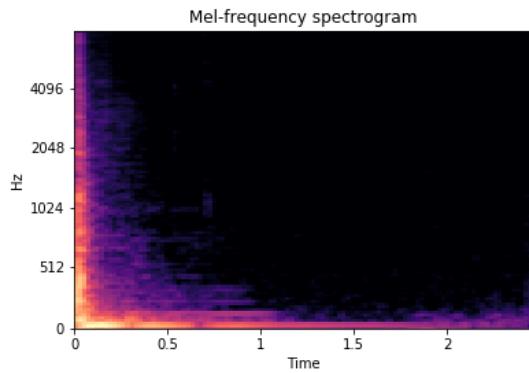


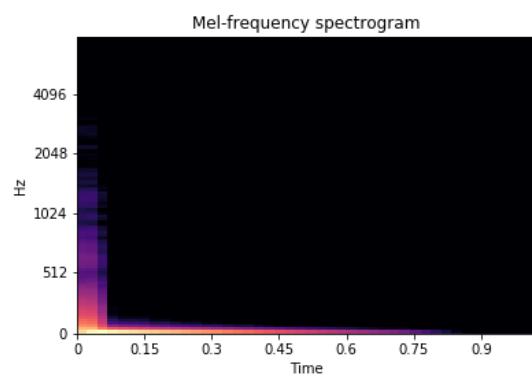
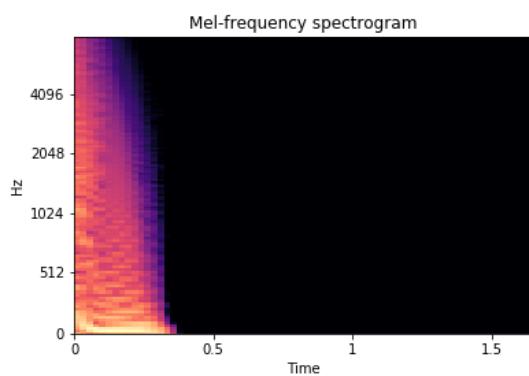
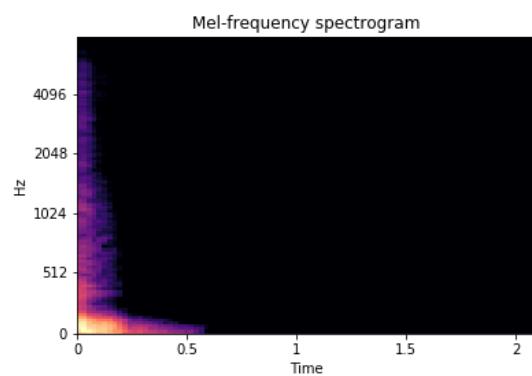
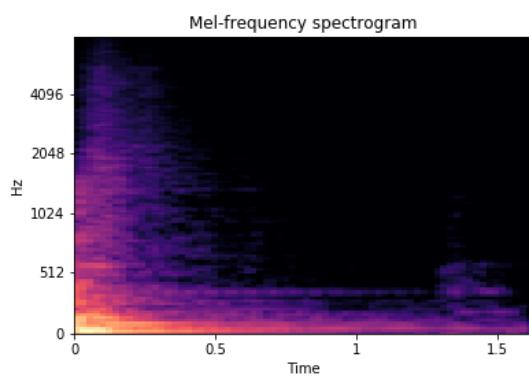
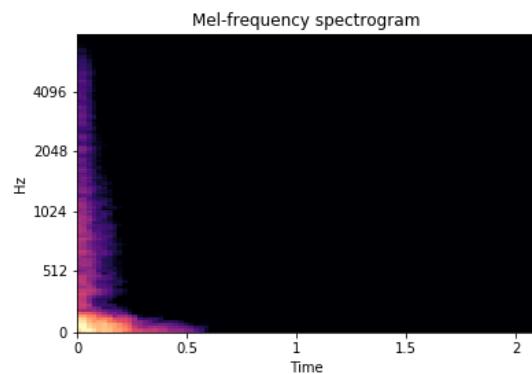
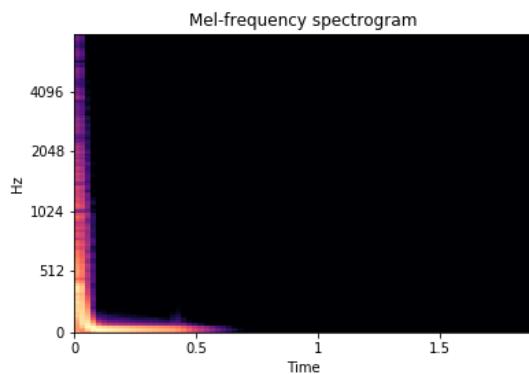
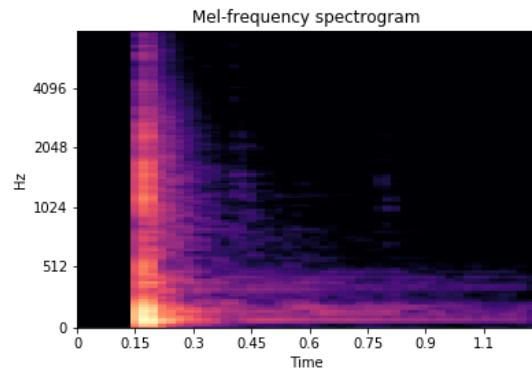
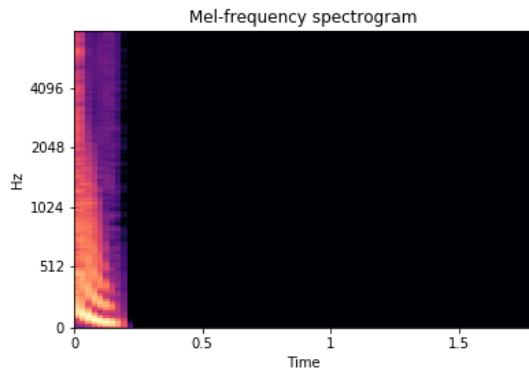


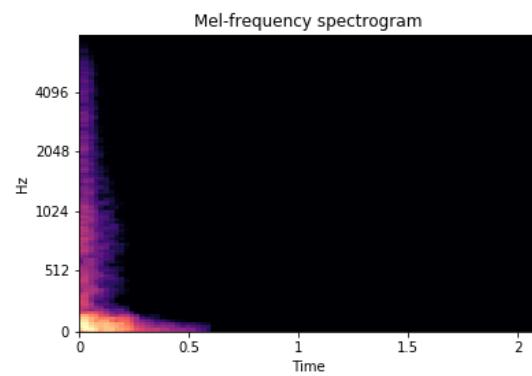
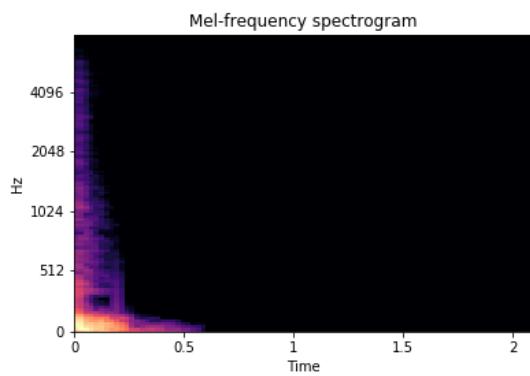
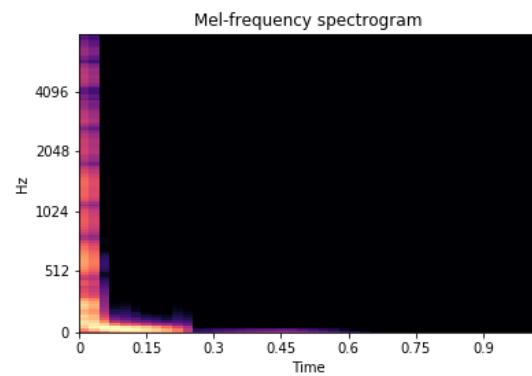
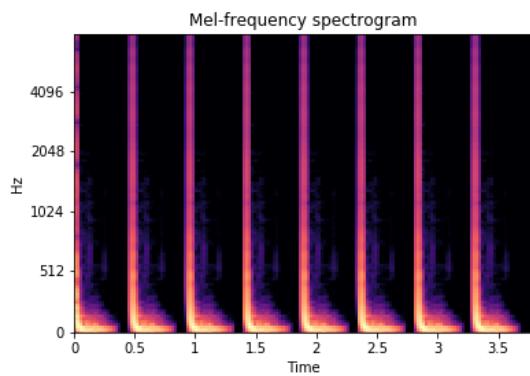
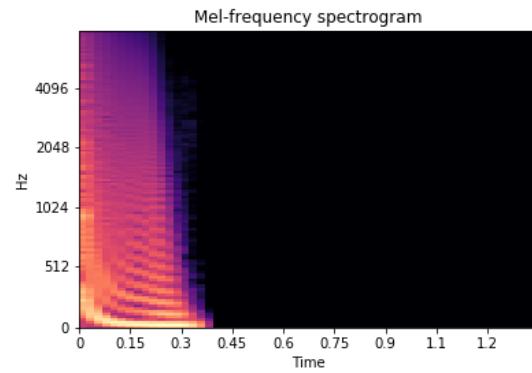
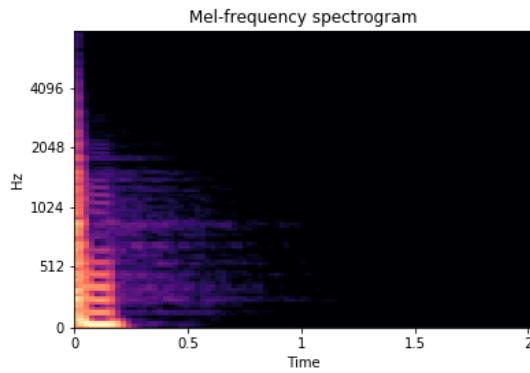


## Bass Drum

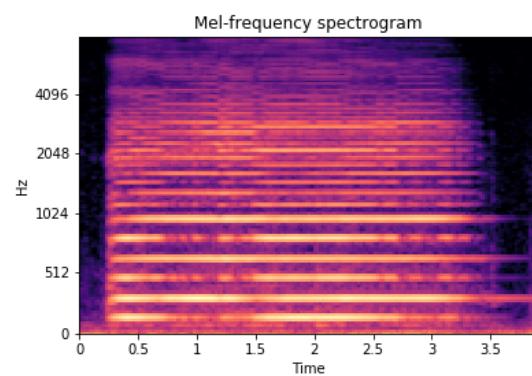
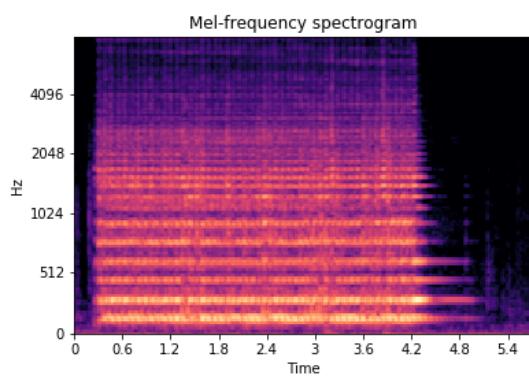
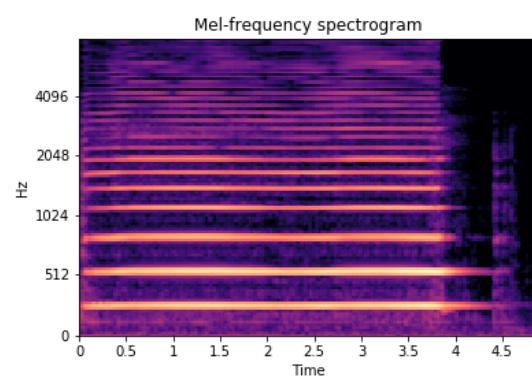
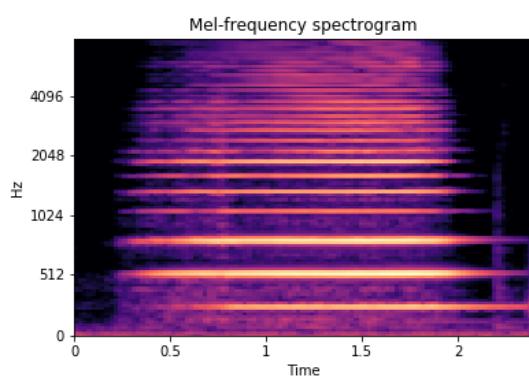
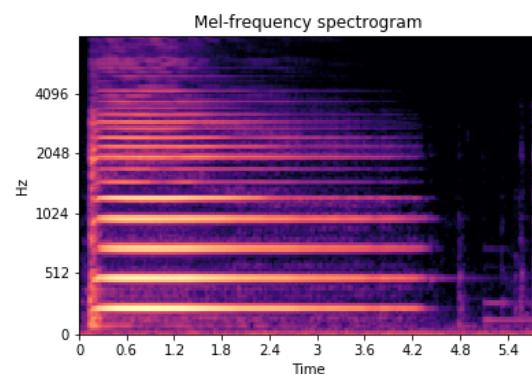
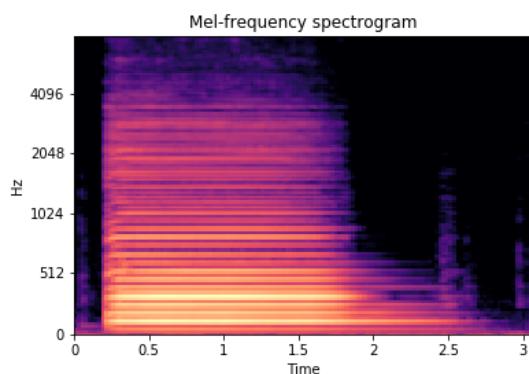
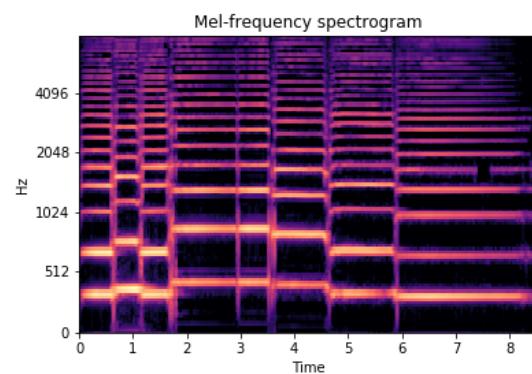
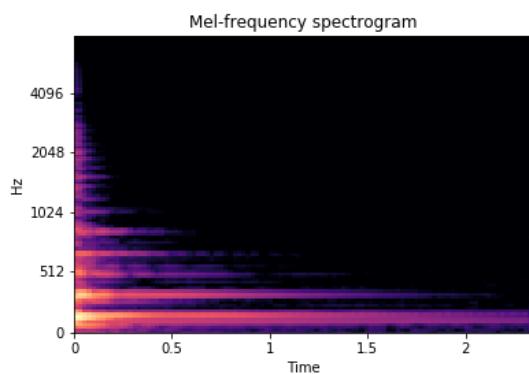


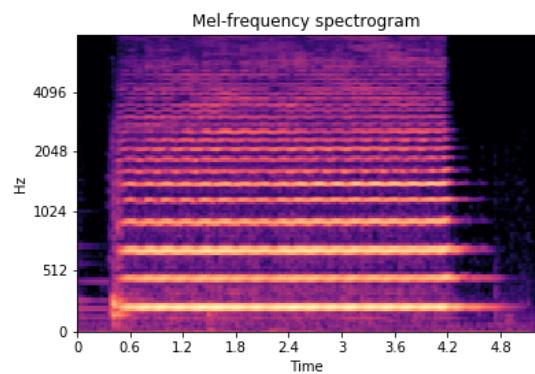
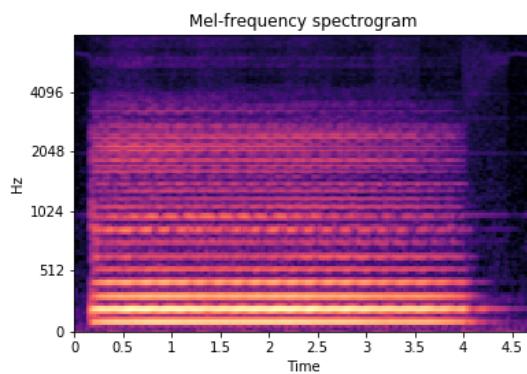
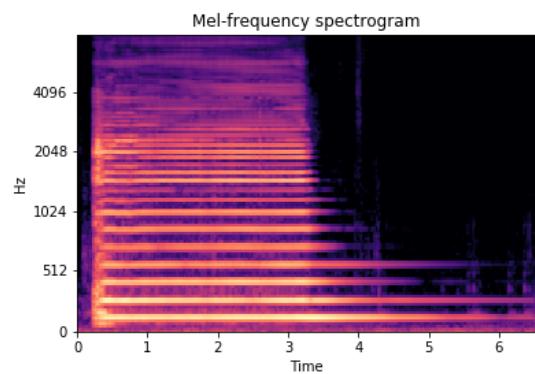
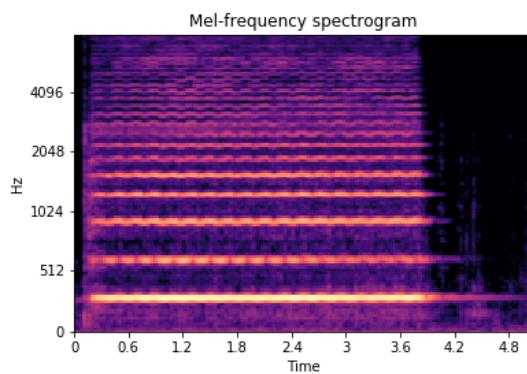
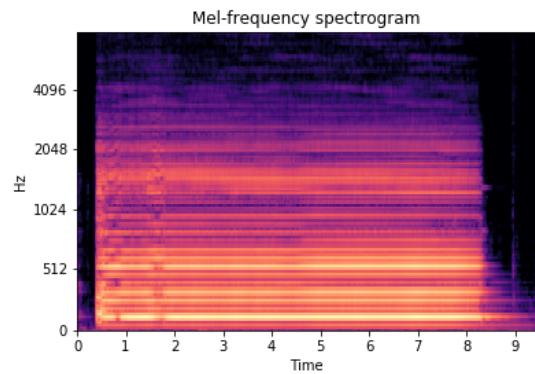
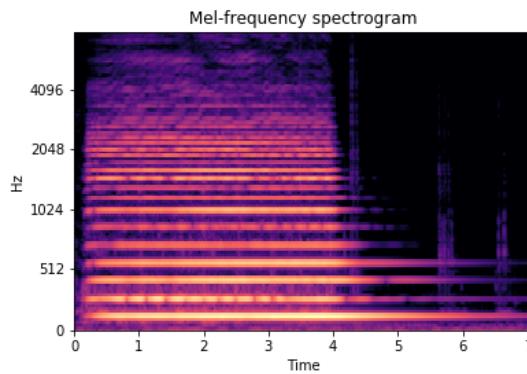
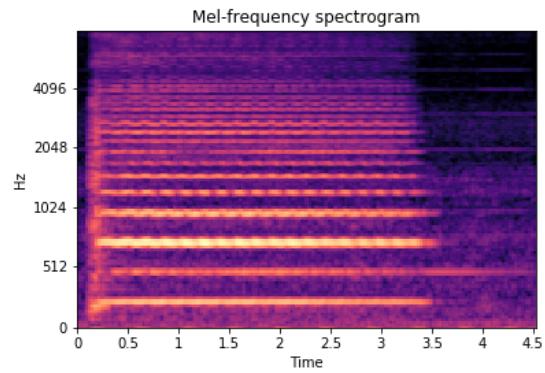
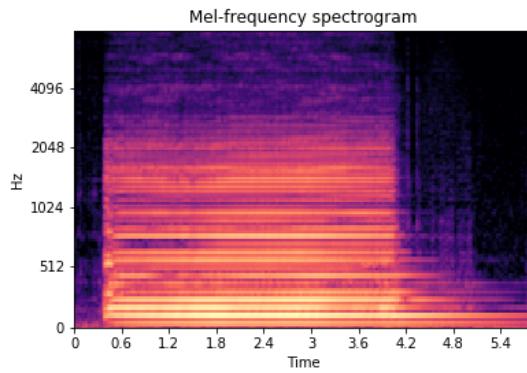


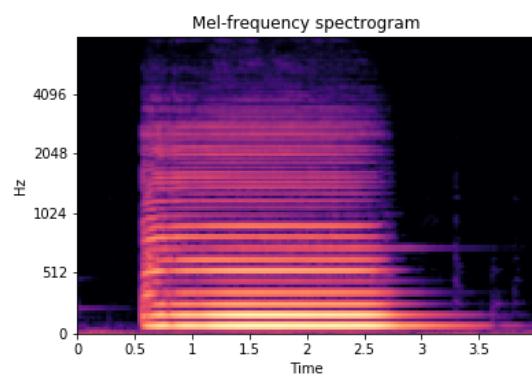
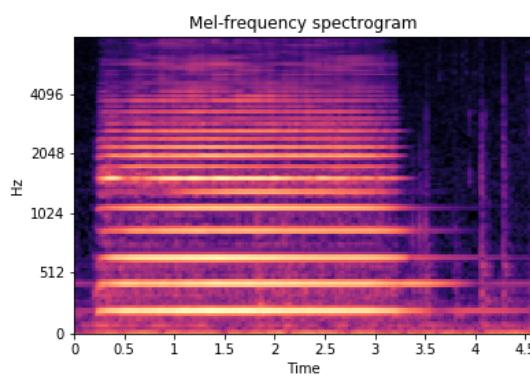
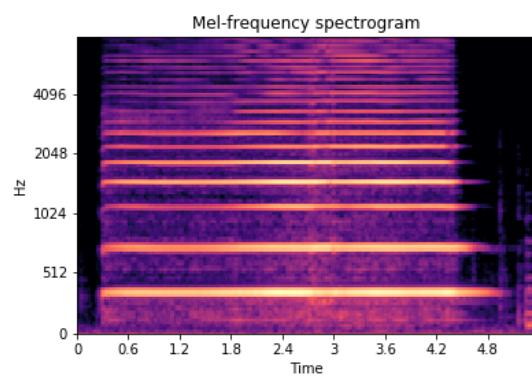
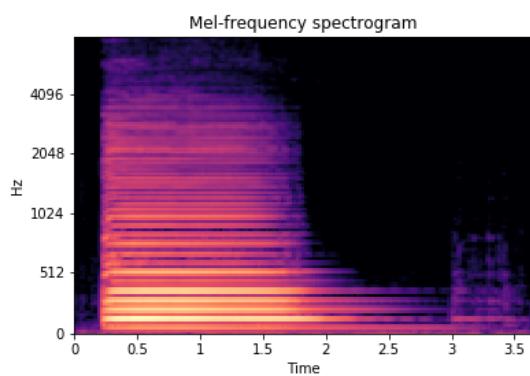
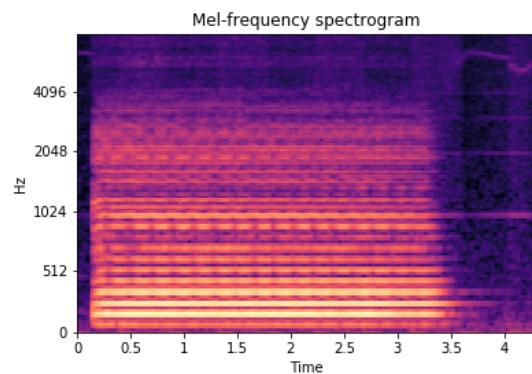
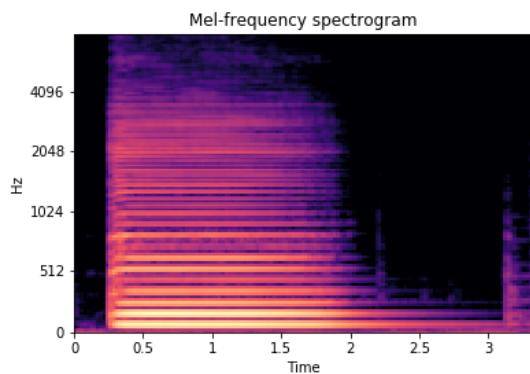
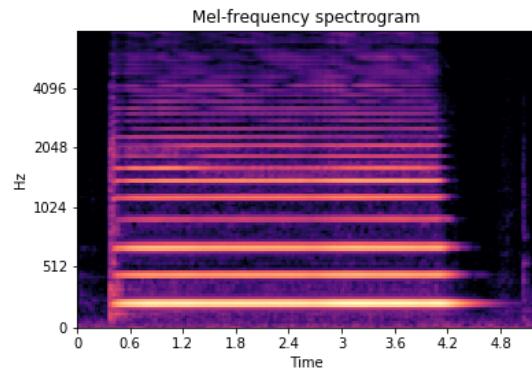
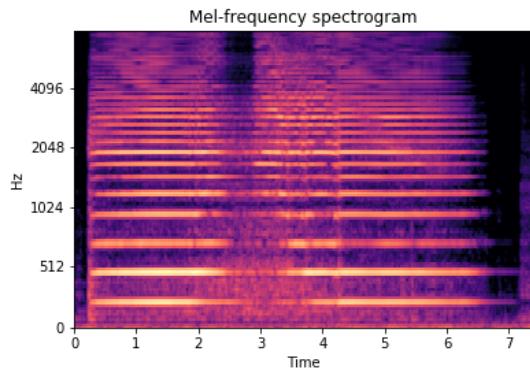


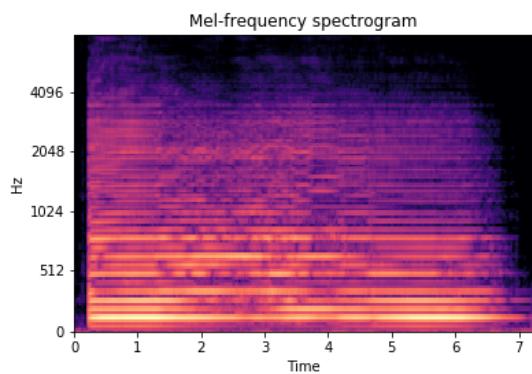
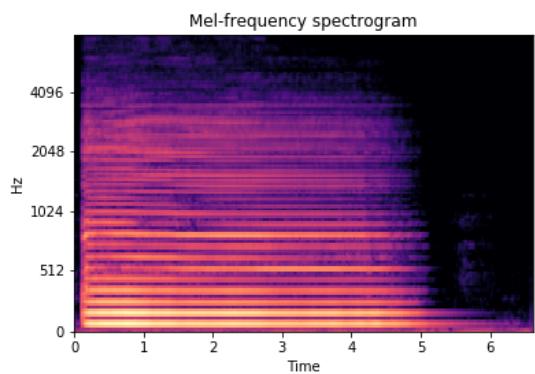
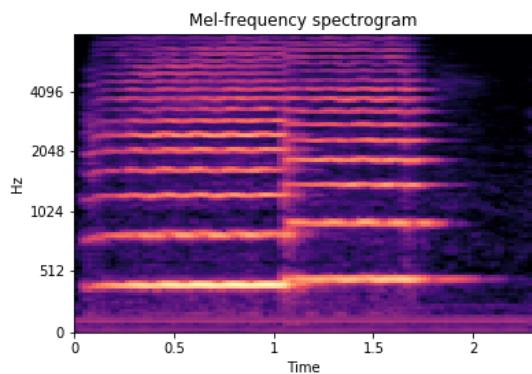
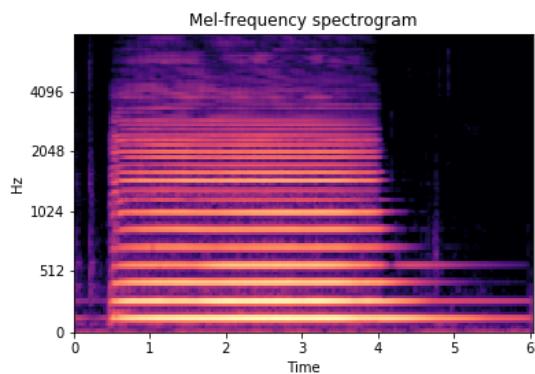
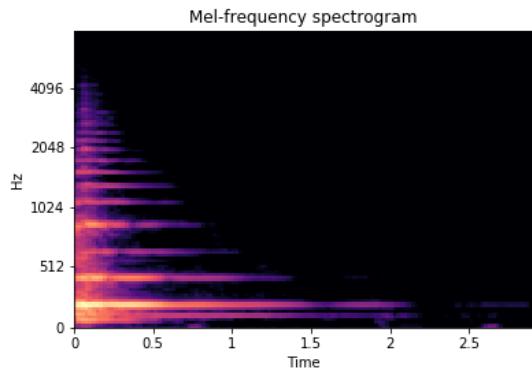
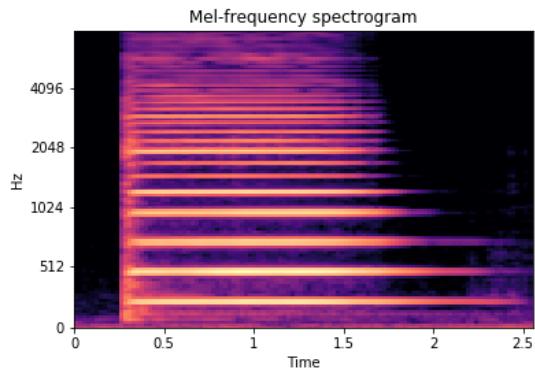


## Cello

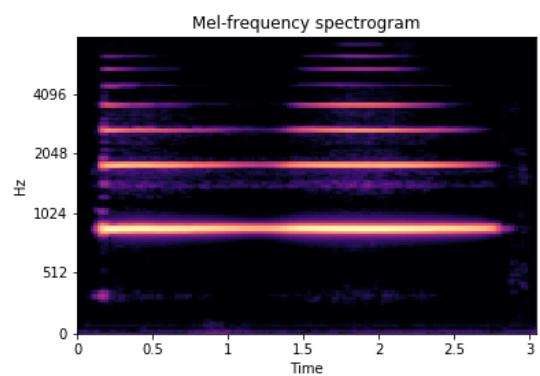
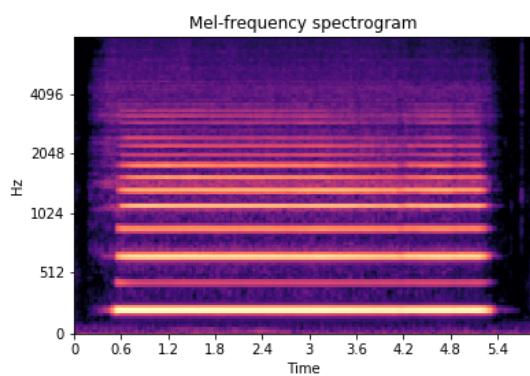
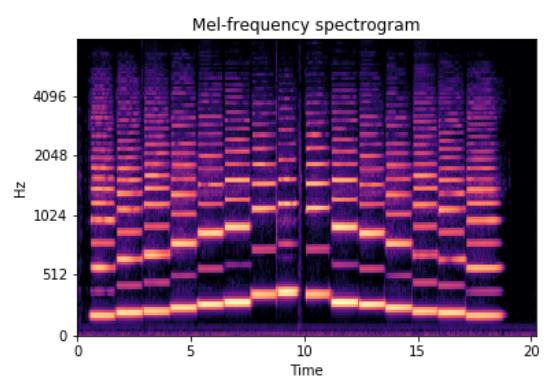
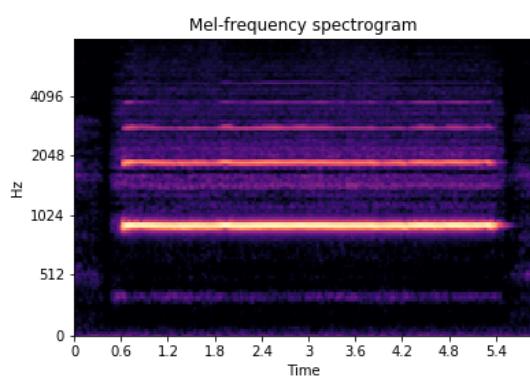
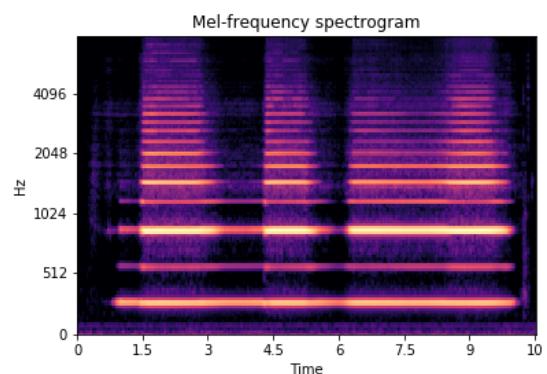
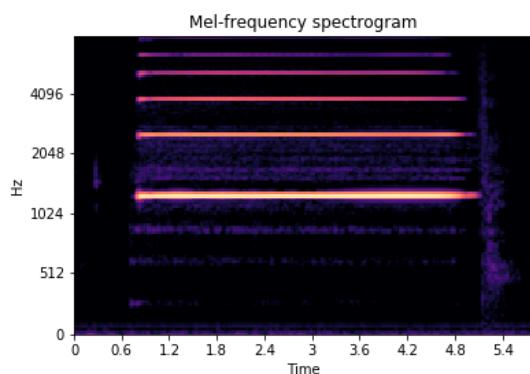
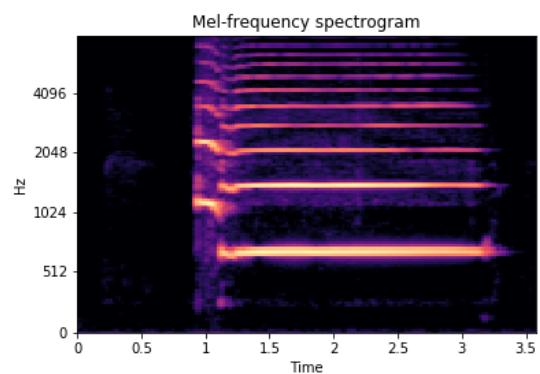
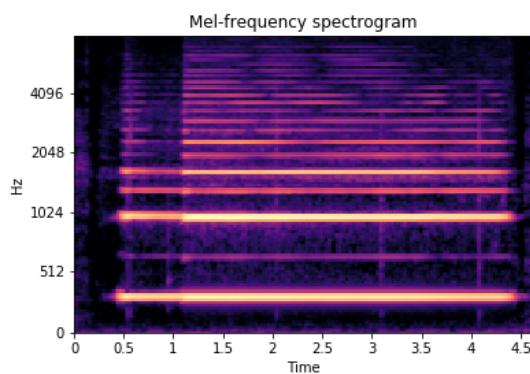


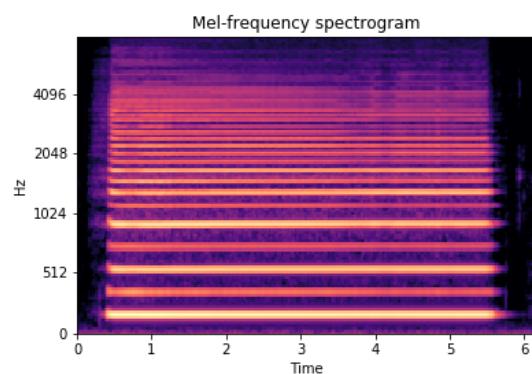
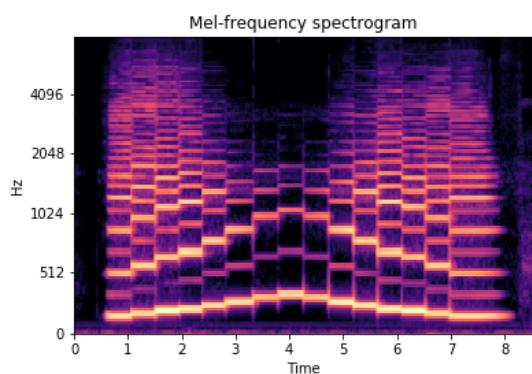
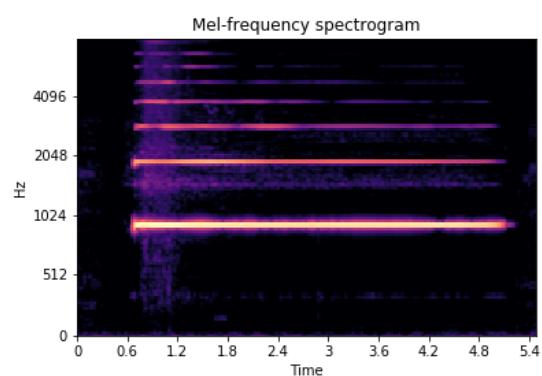
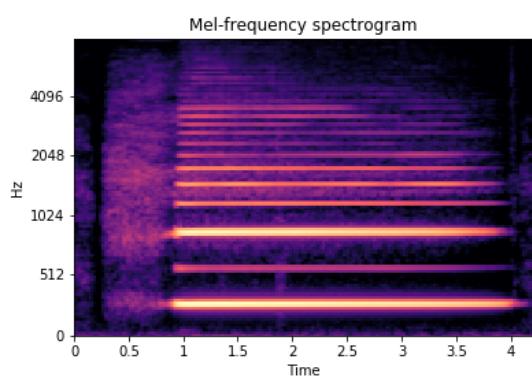
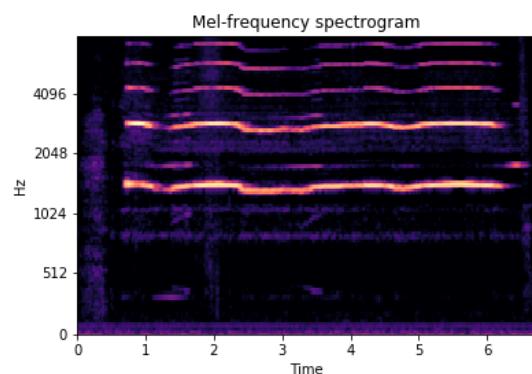
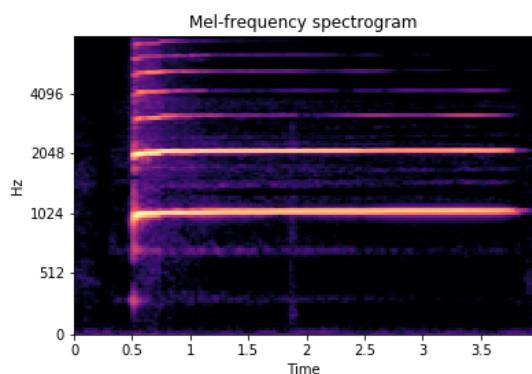
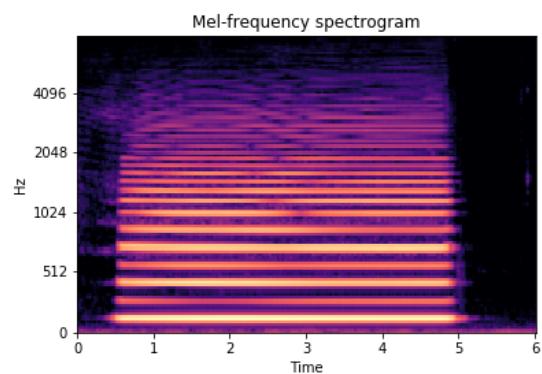
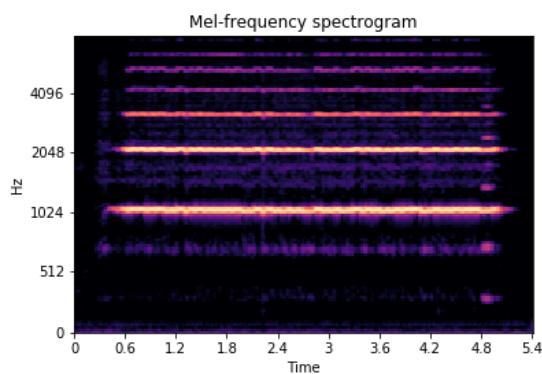


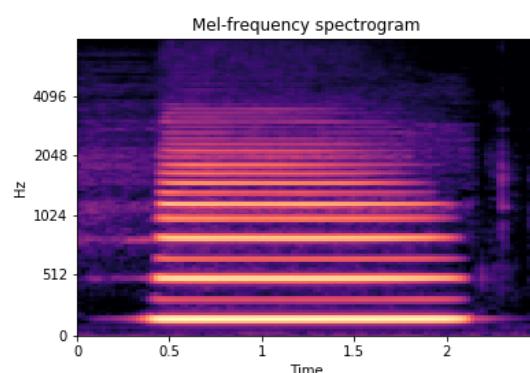
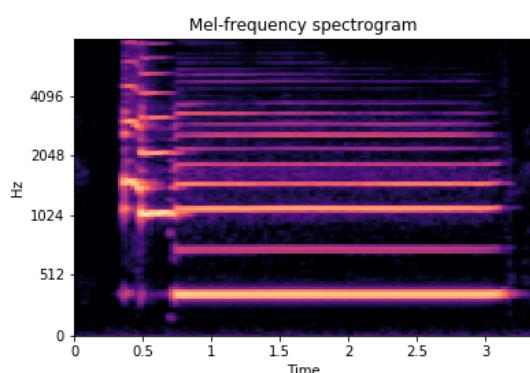
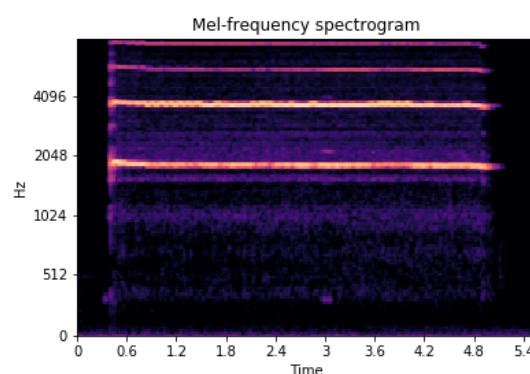
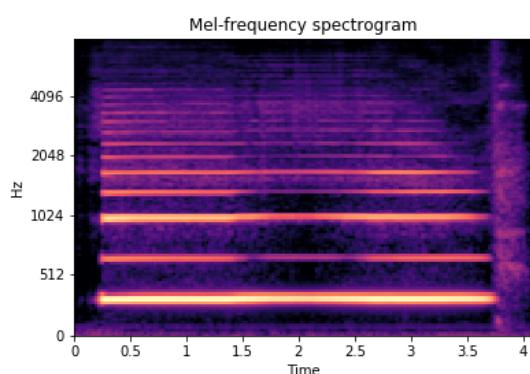
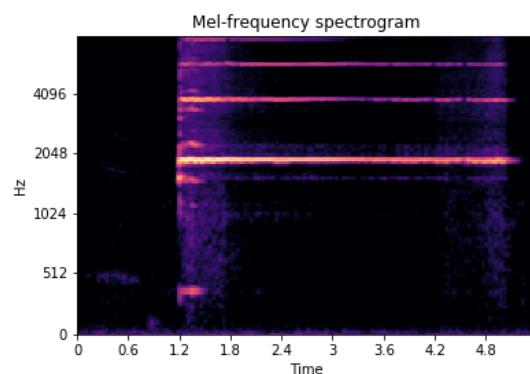
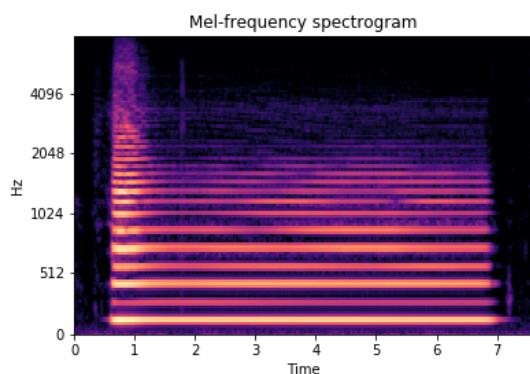
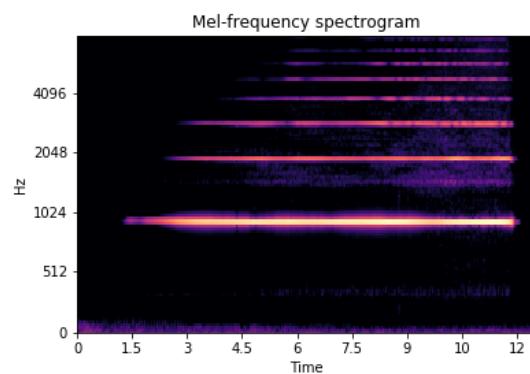
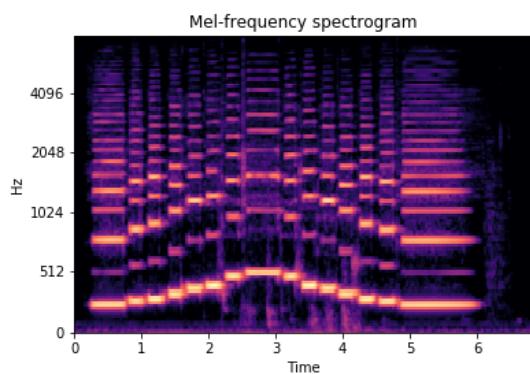


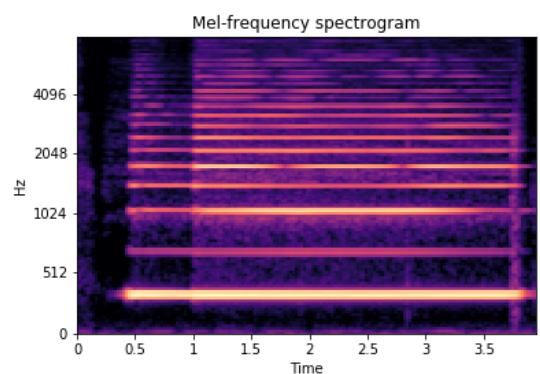
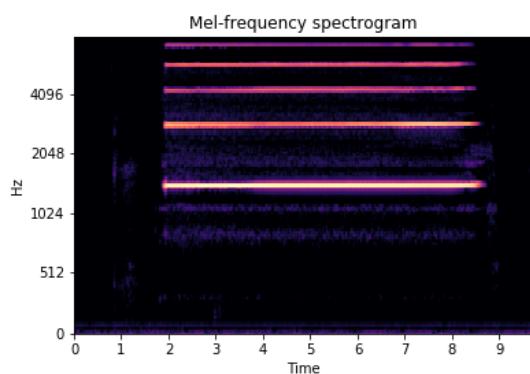
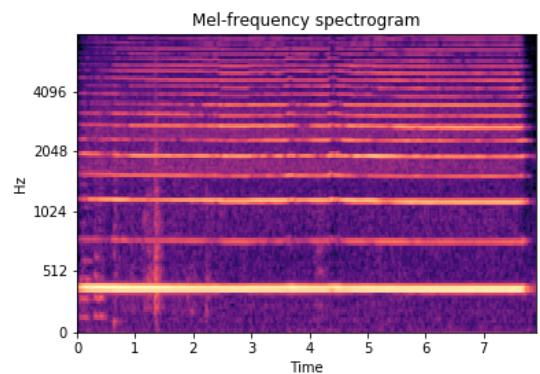
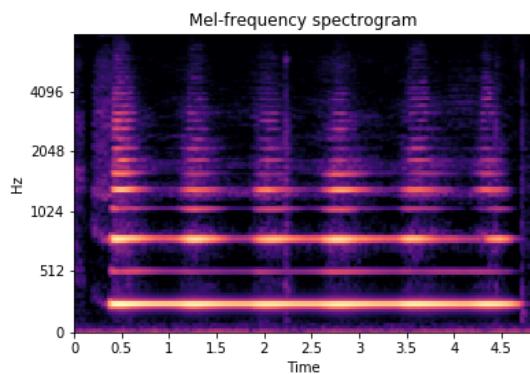
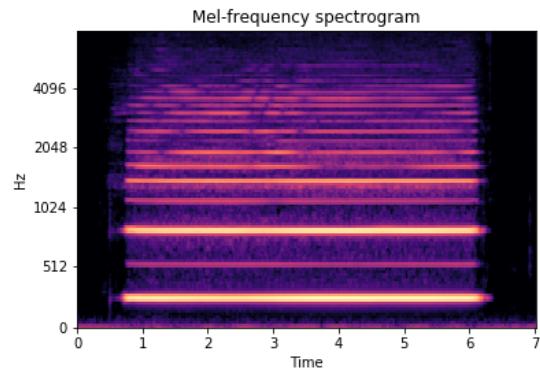
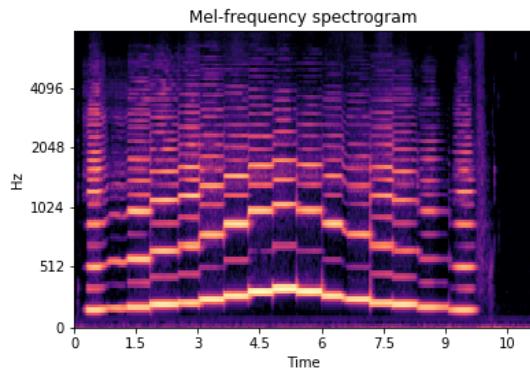


## Clarinet

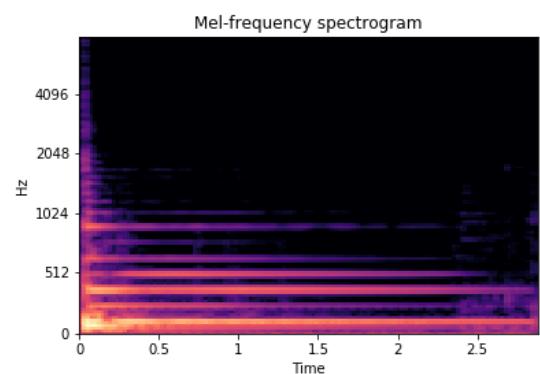
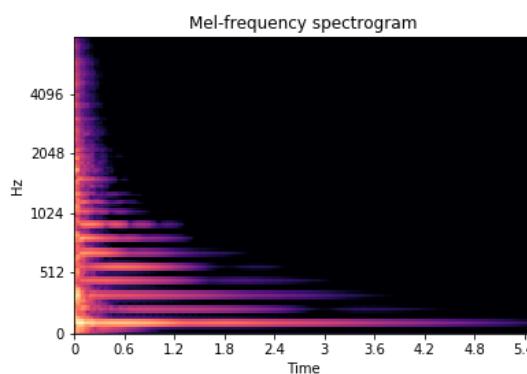
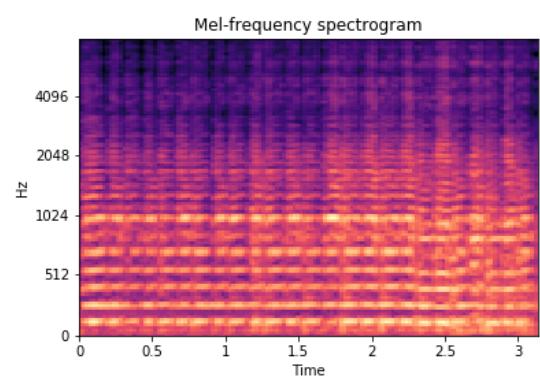
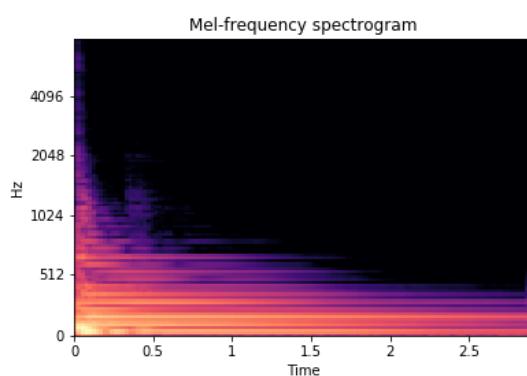
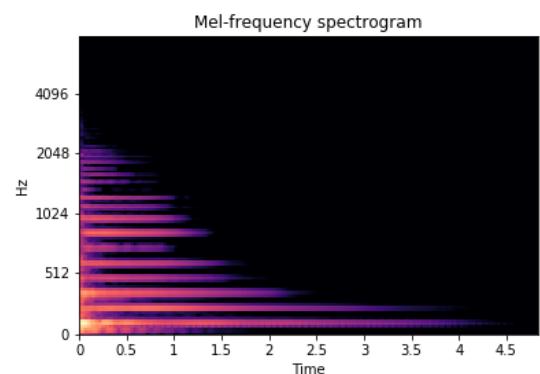
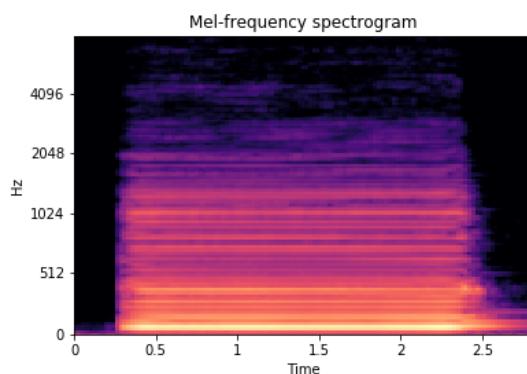
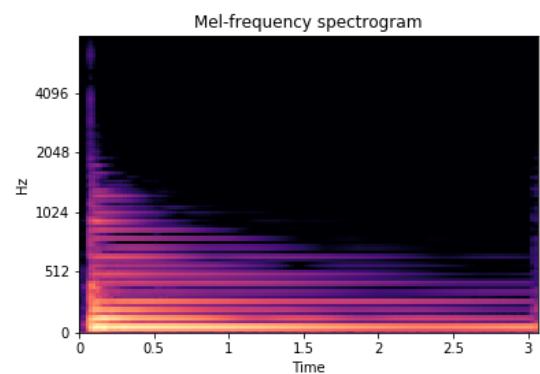
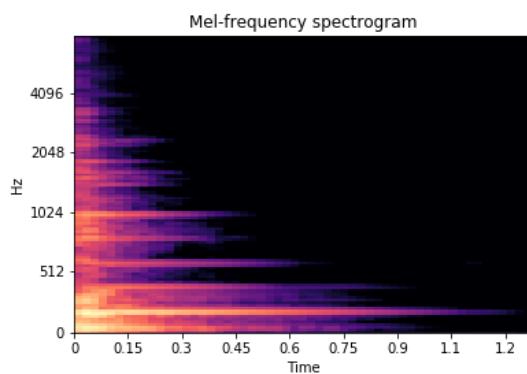


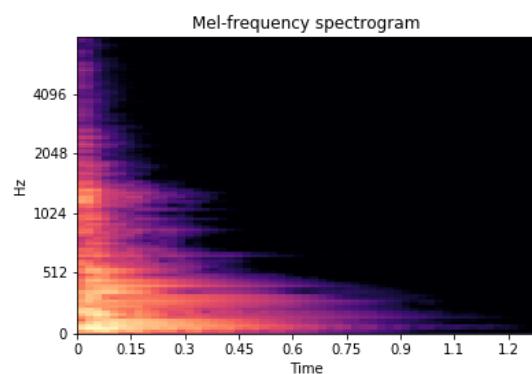
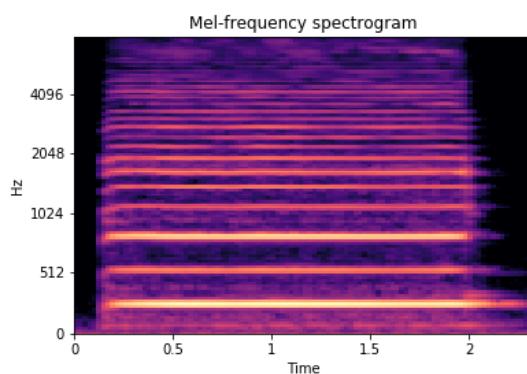
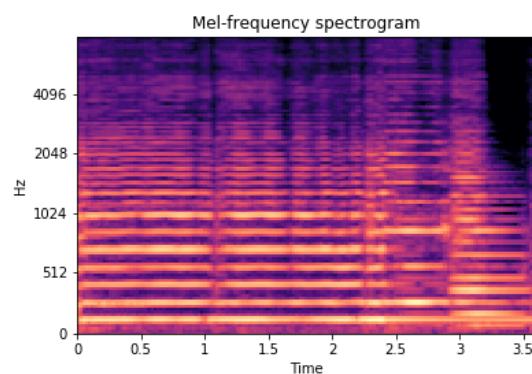
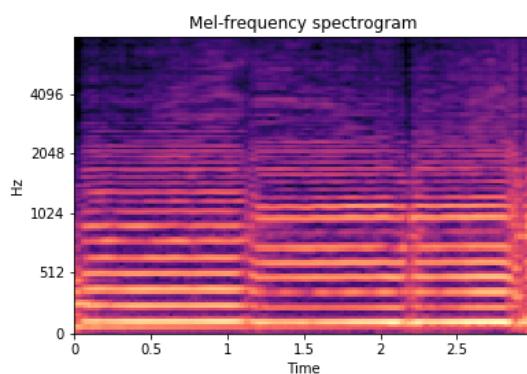
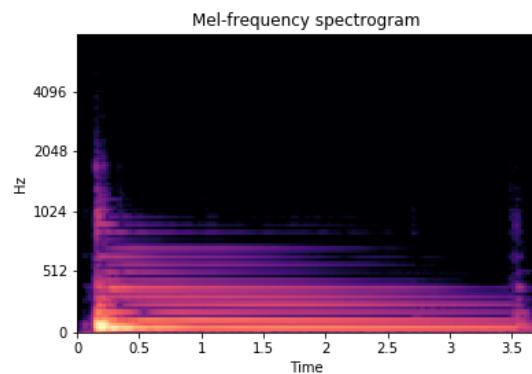
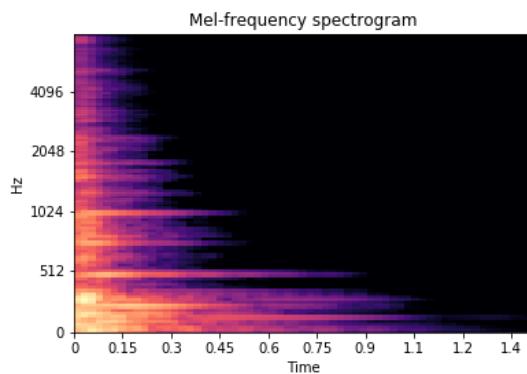
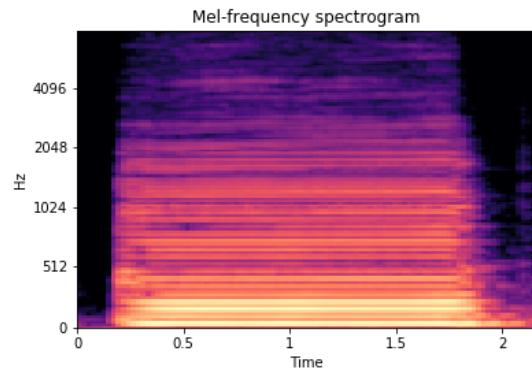
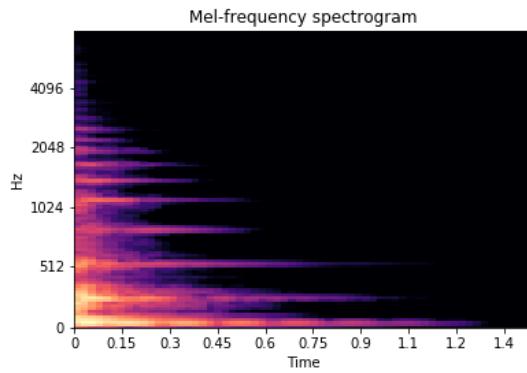


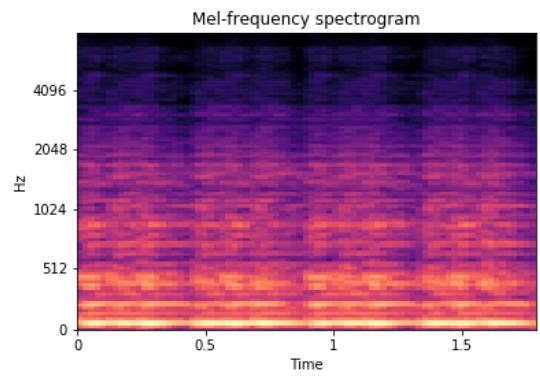
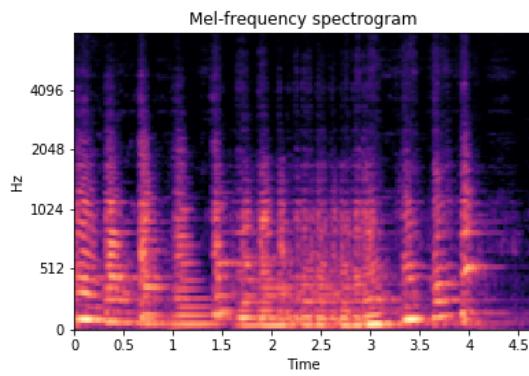
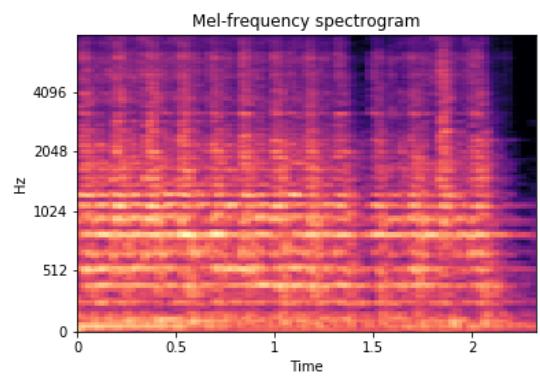
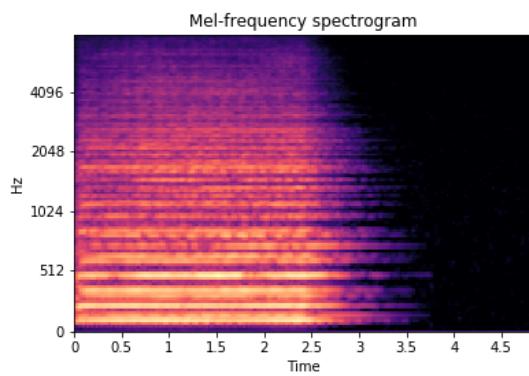
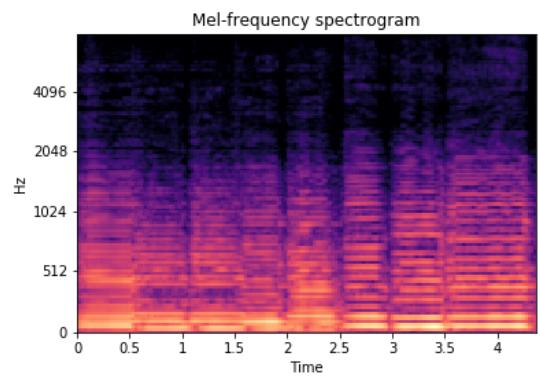
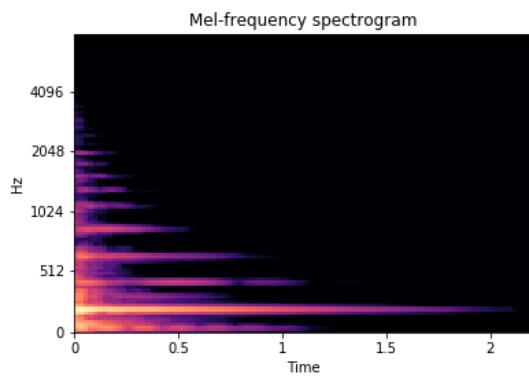
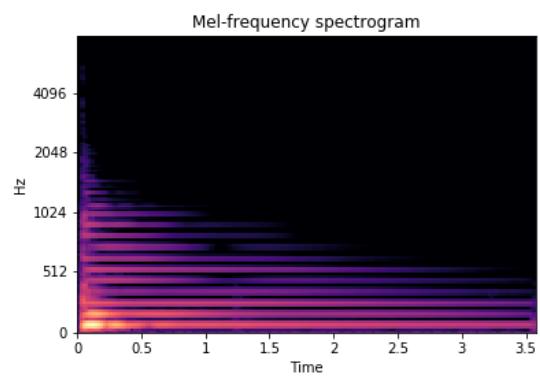
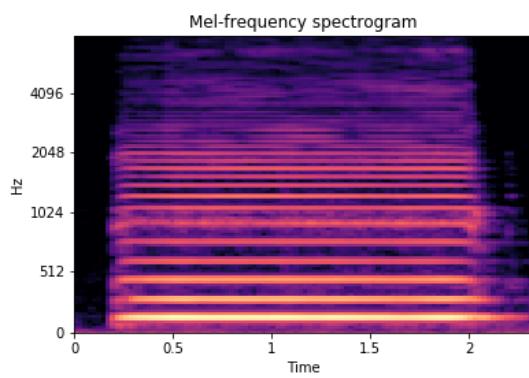


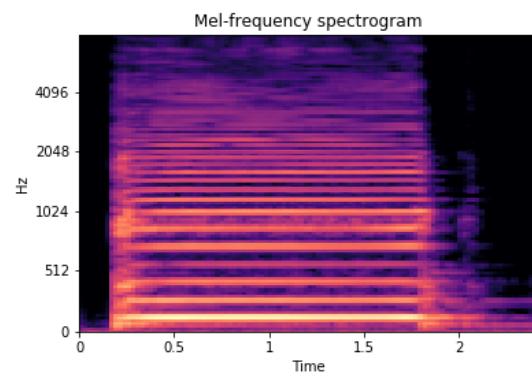
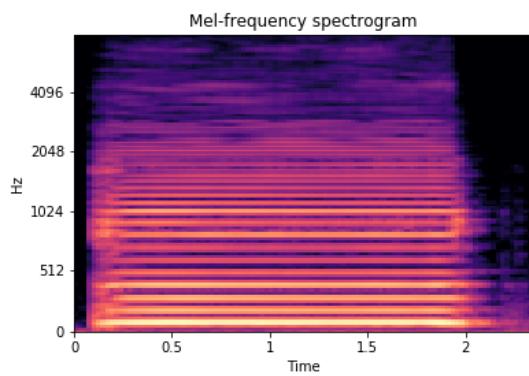
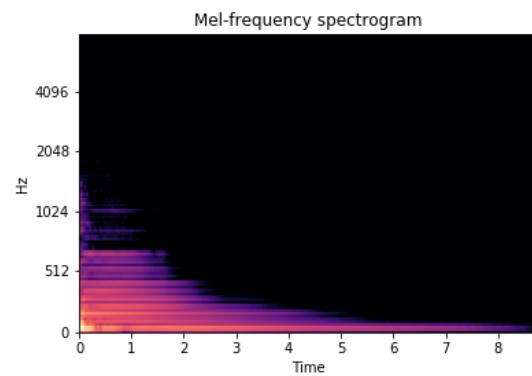
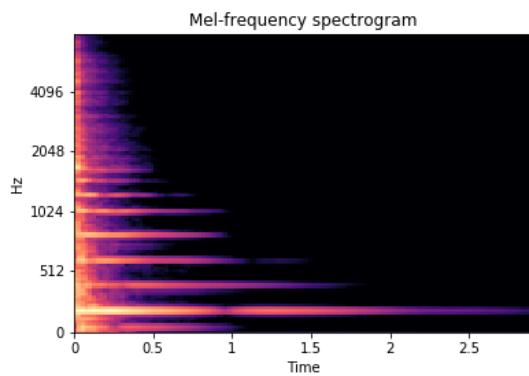
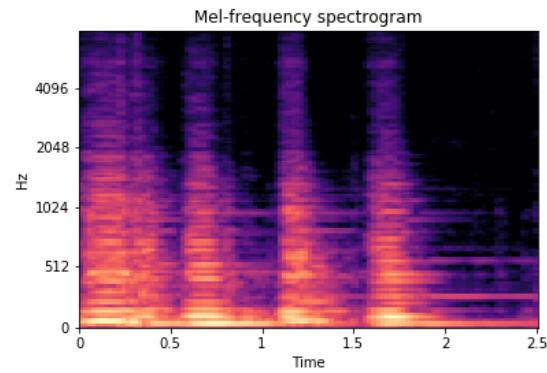
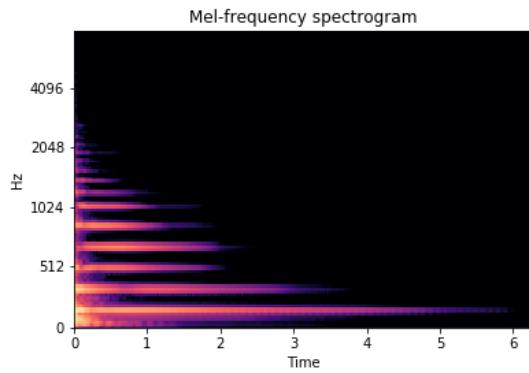


## Double Bass

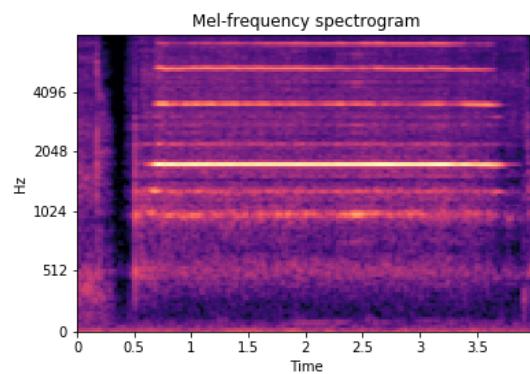
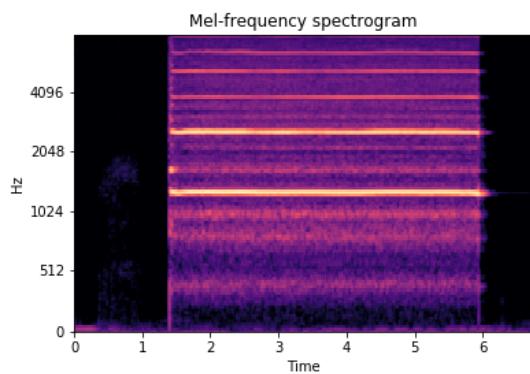
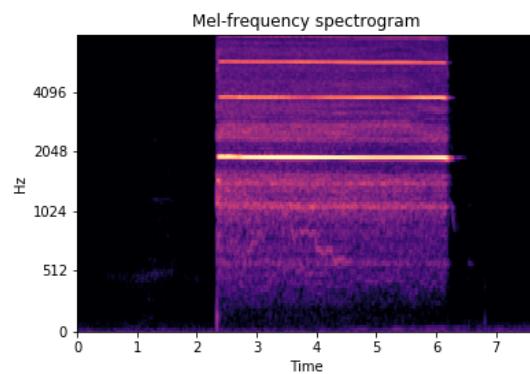
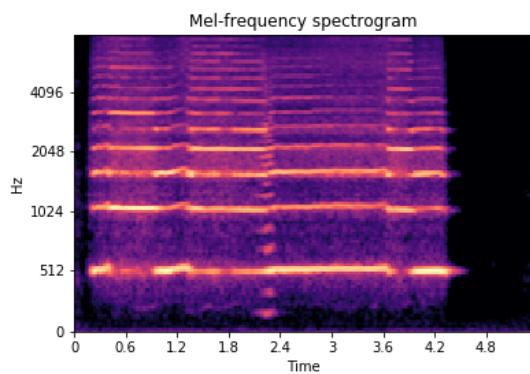
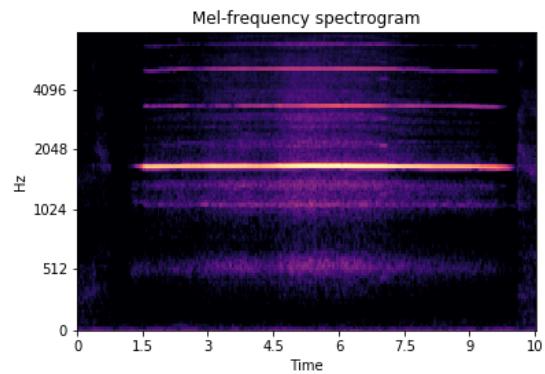
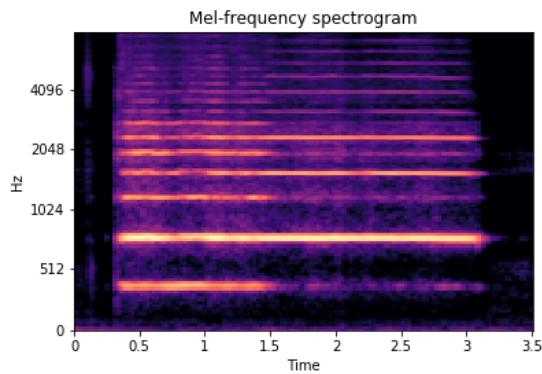
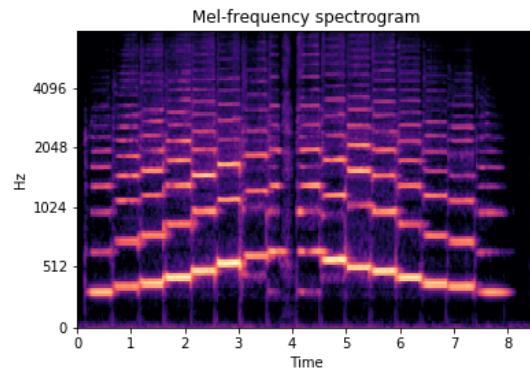
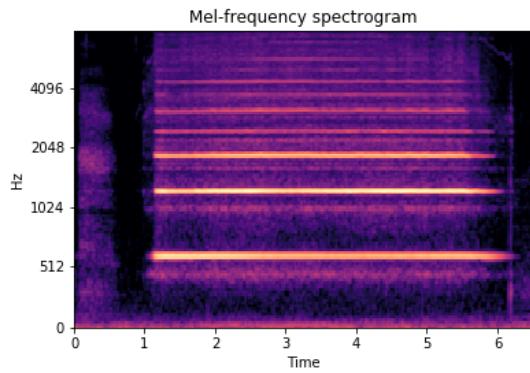


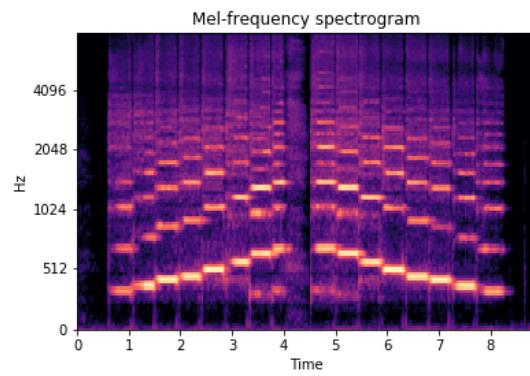
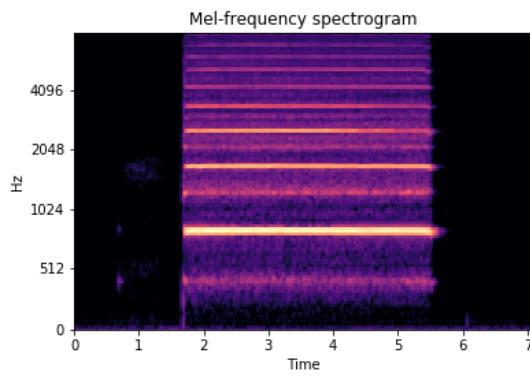
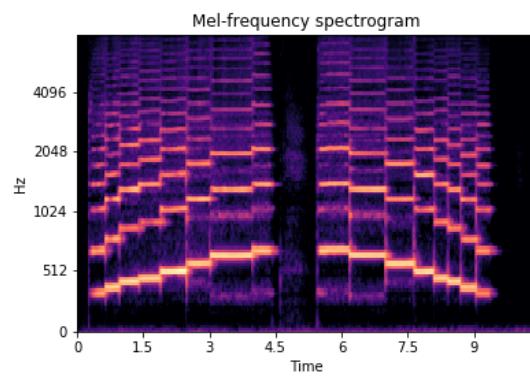
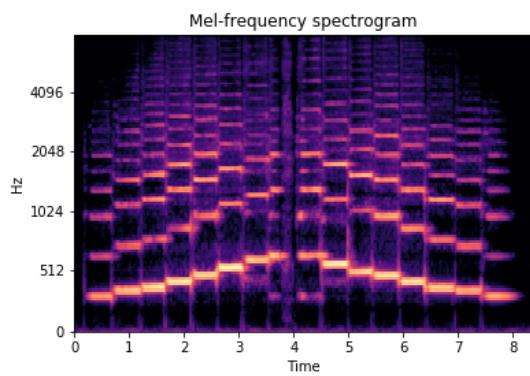
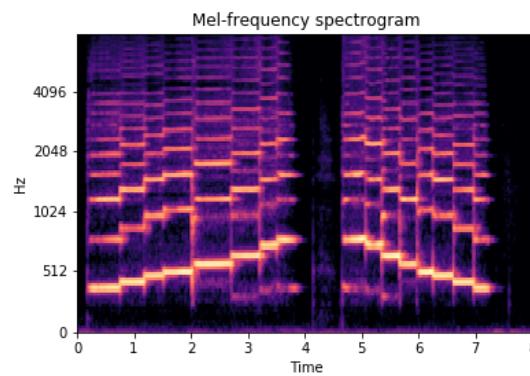
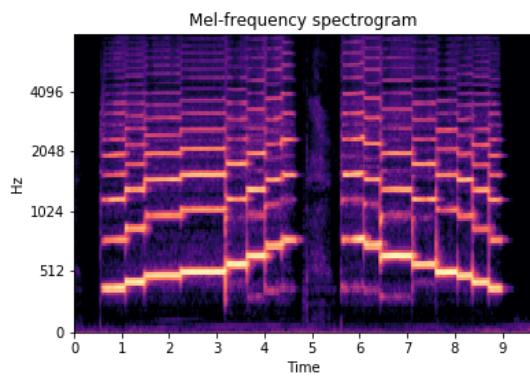
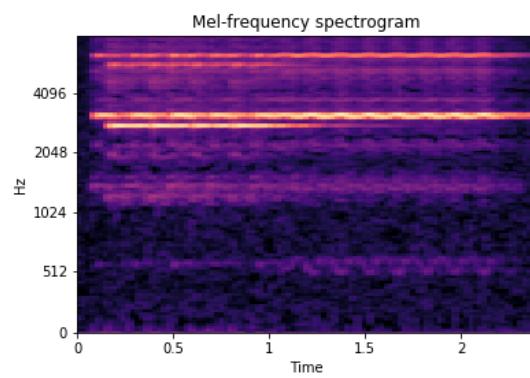
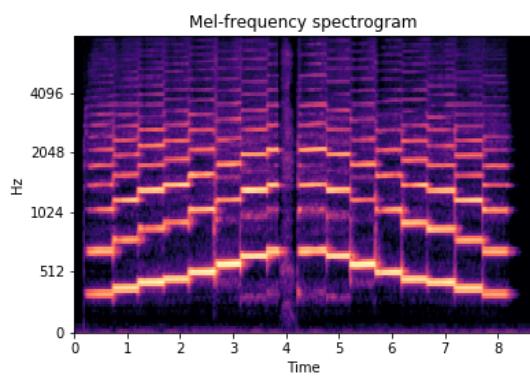


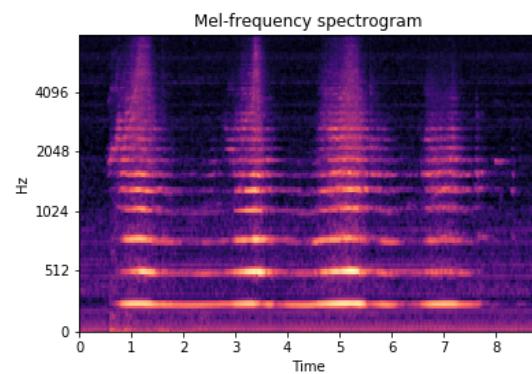
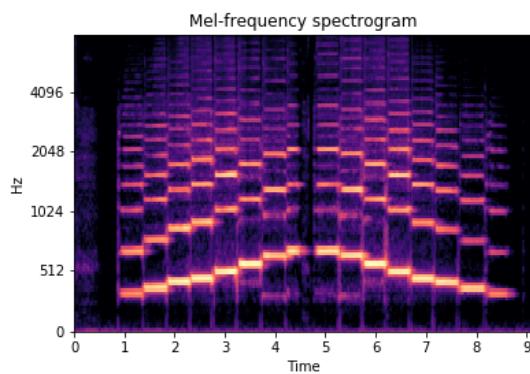
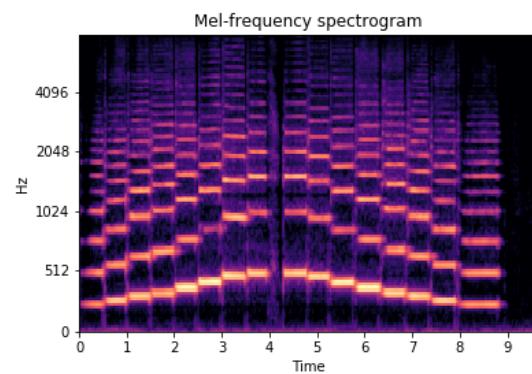
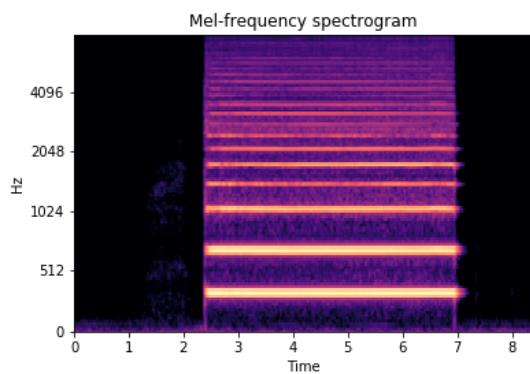
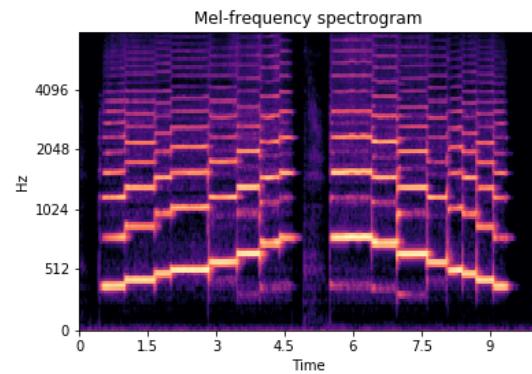
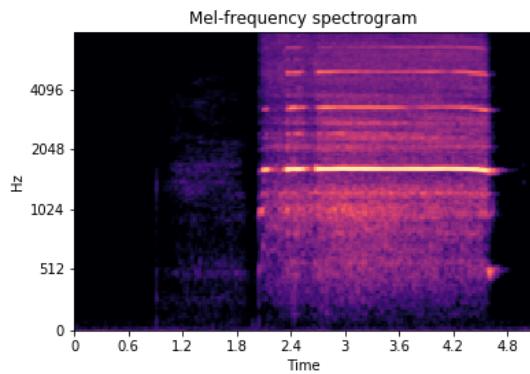
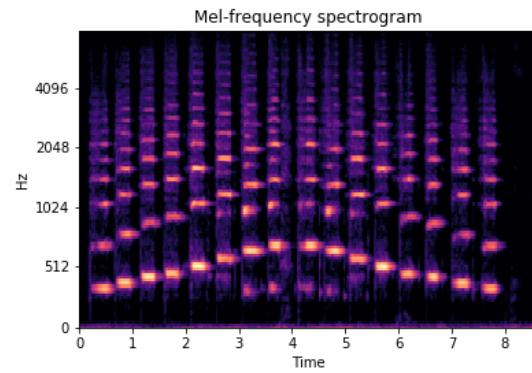
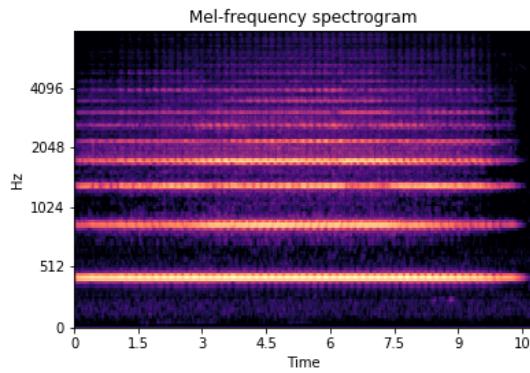


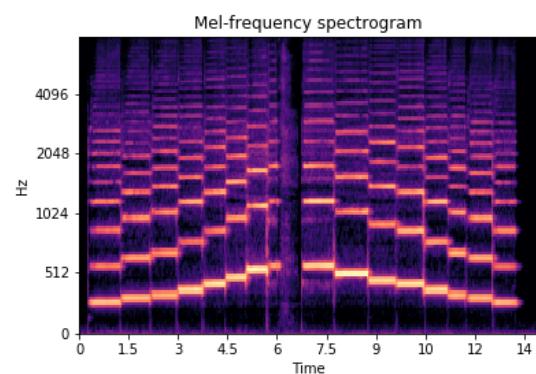
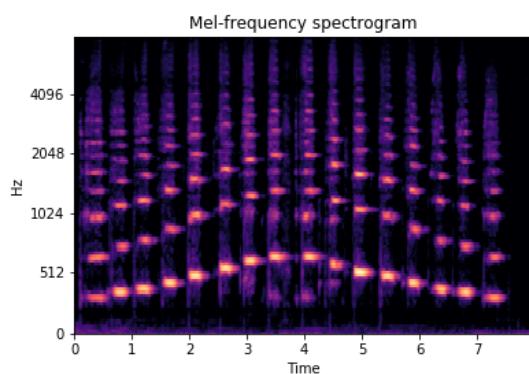
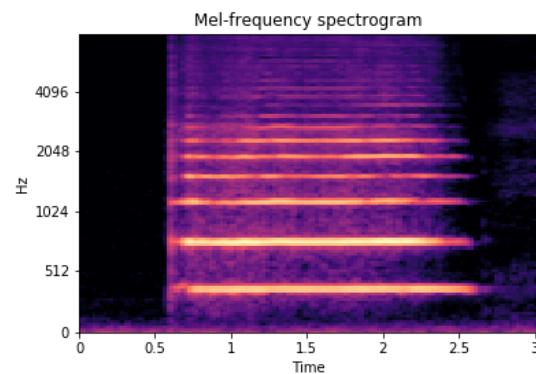
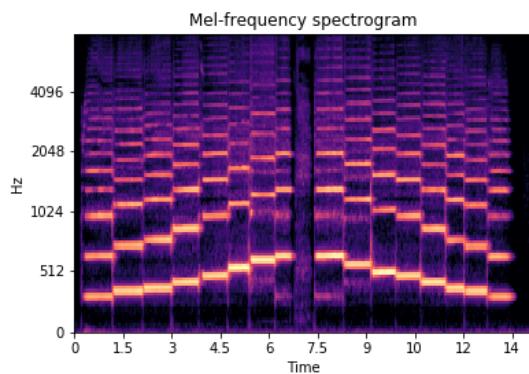
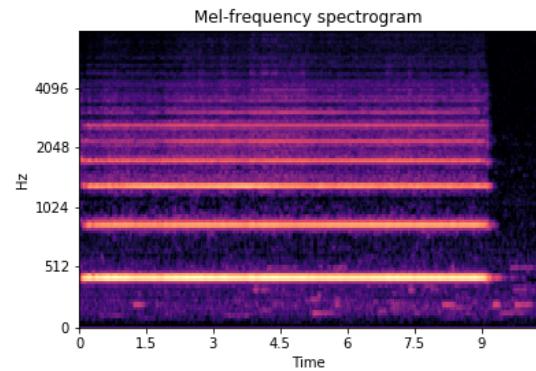
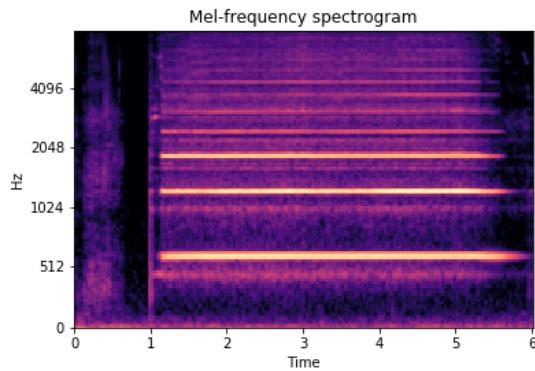


## Flute

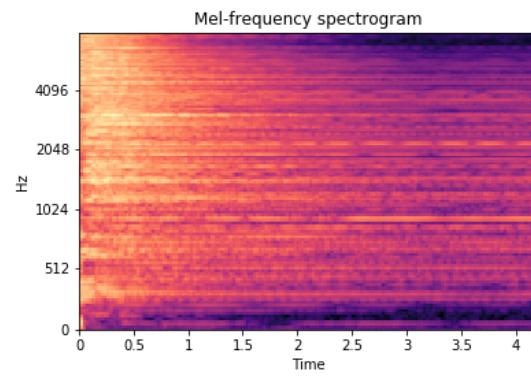
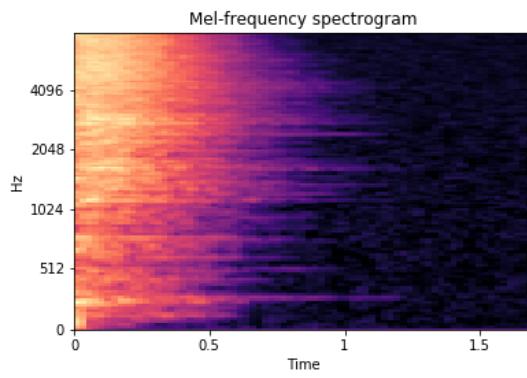
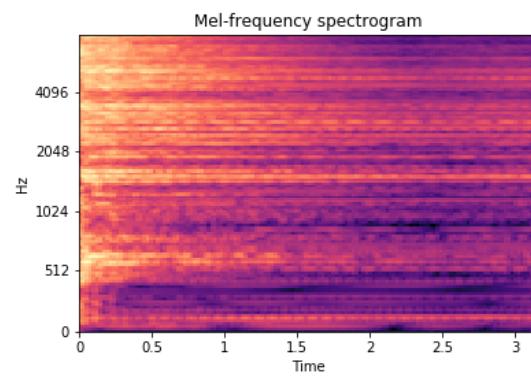
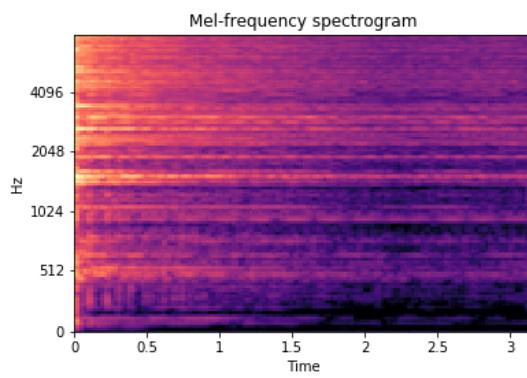
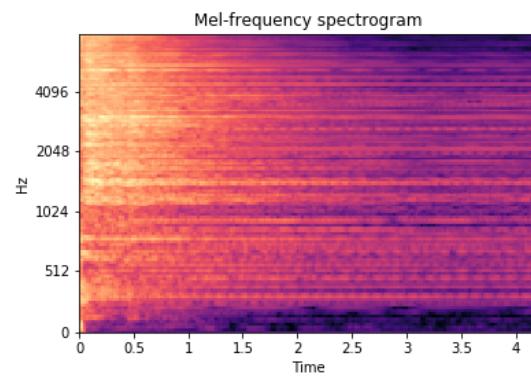
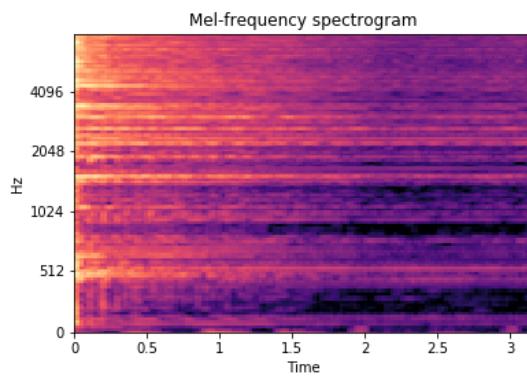
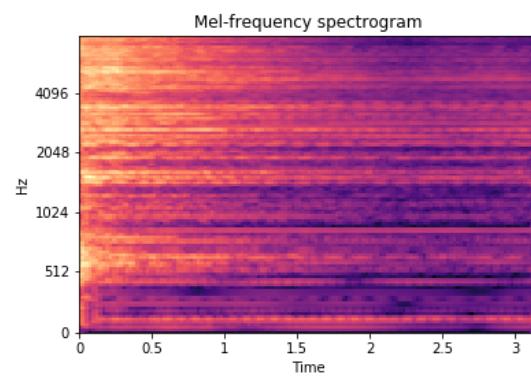
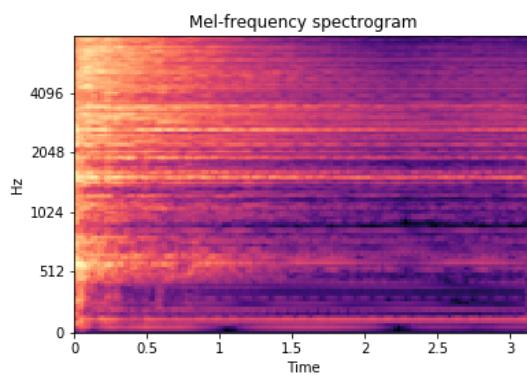


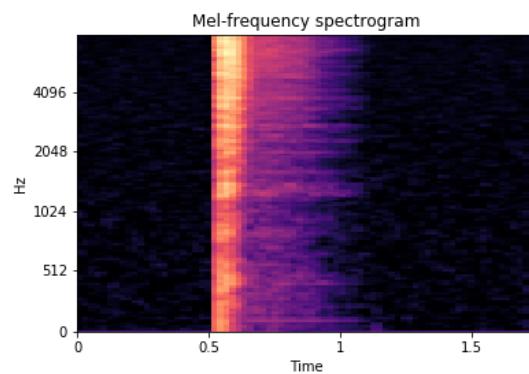
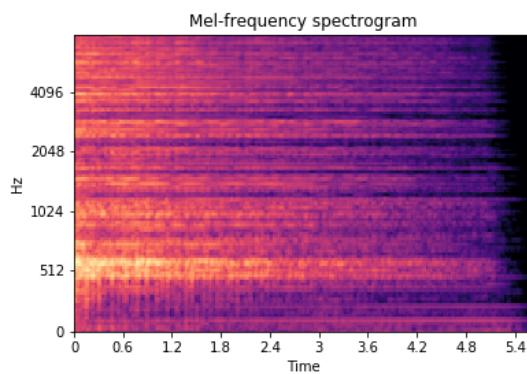
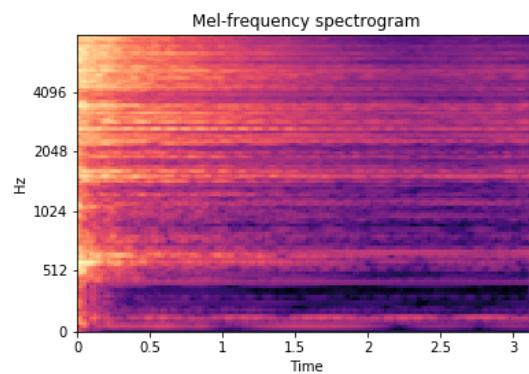
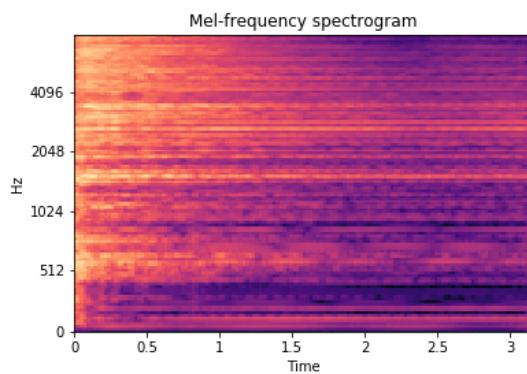
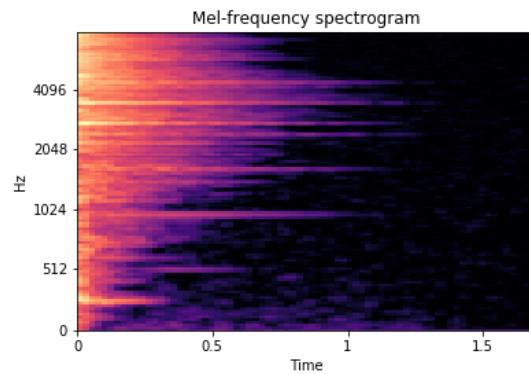
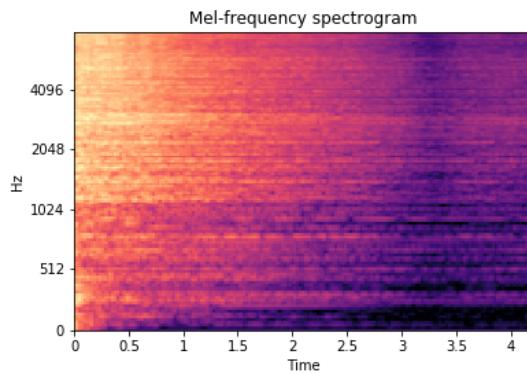
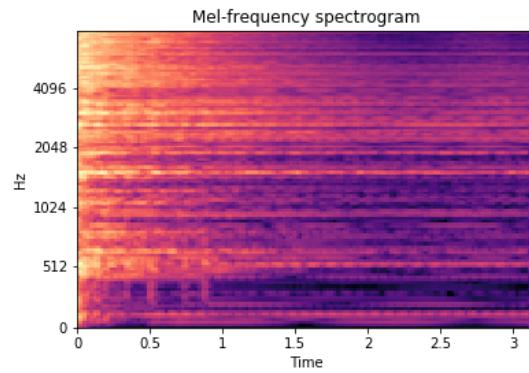
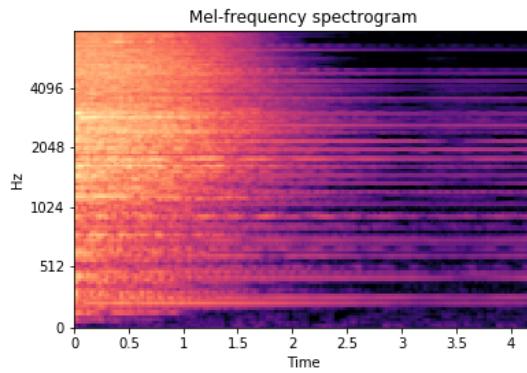


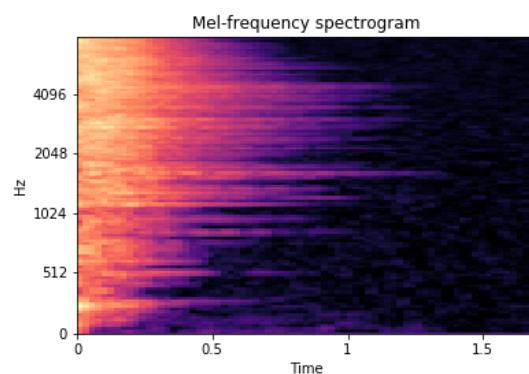
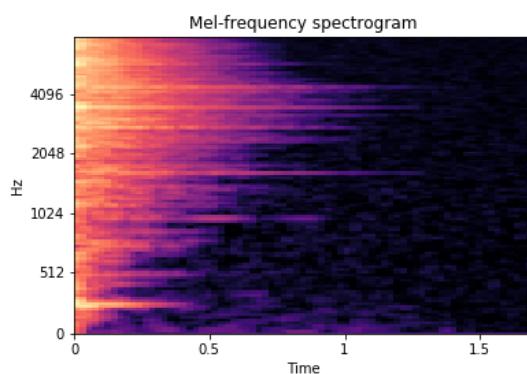
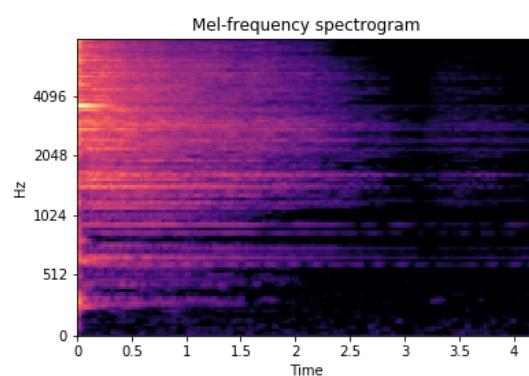
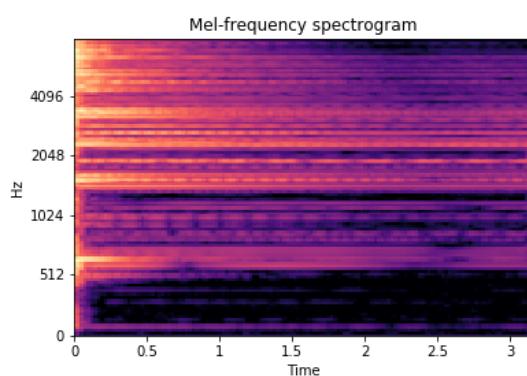
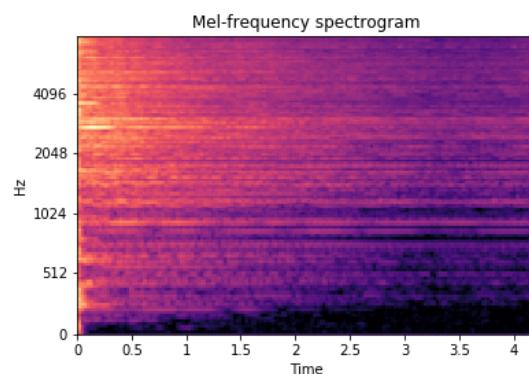
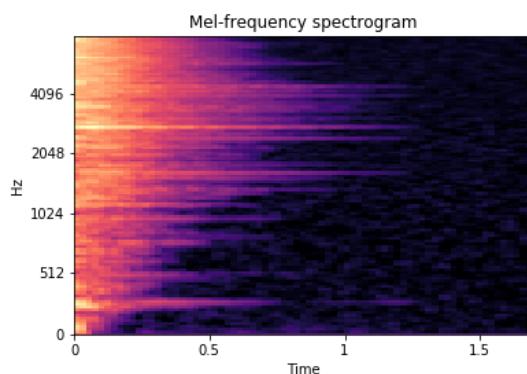
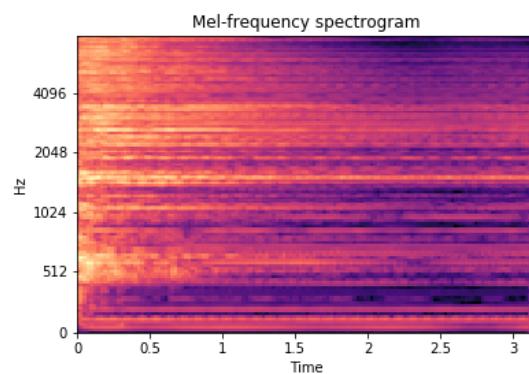
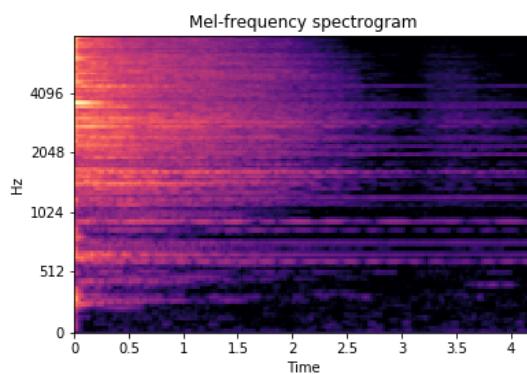


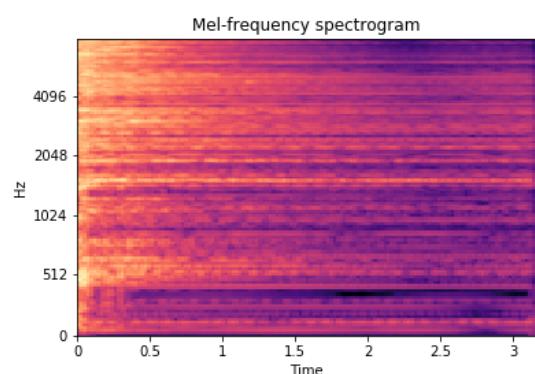
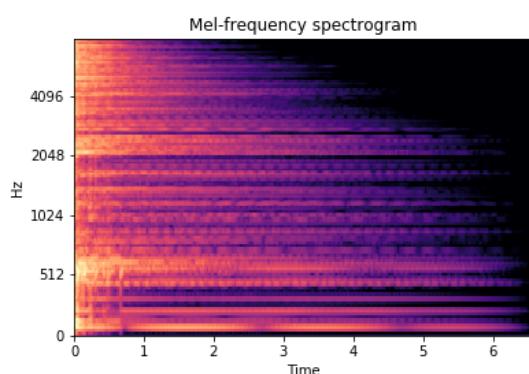
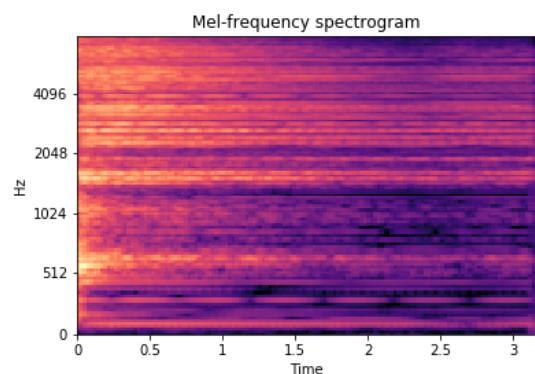
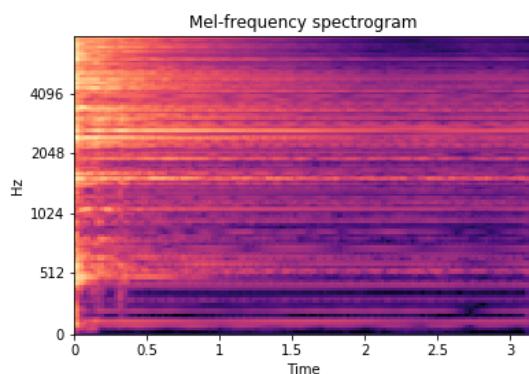
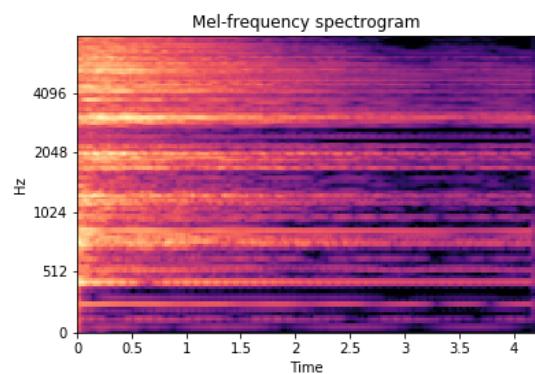
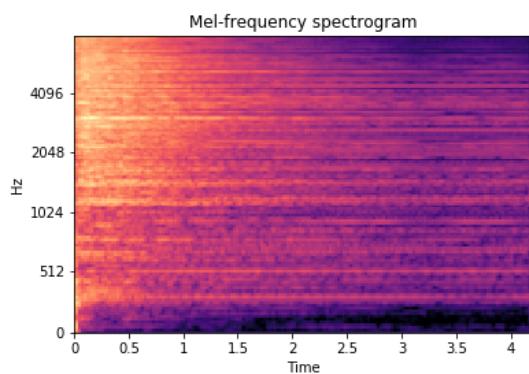


## Hi-hat

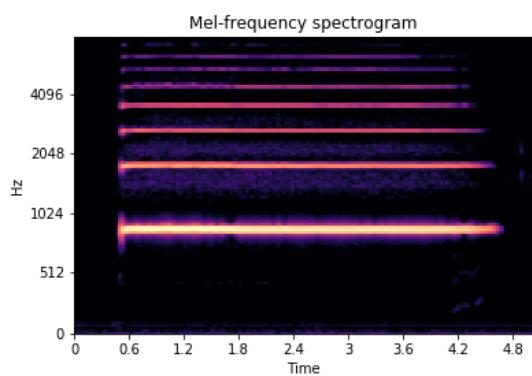
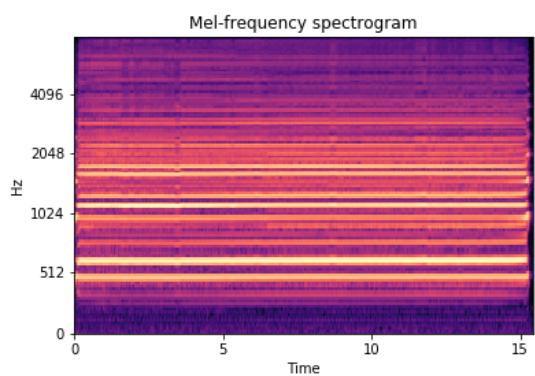
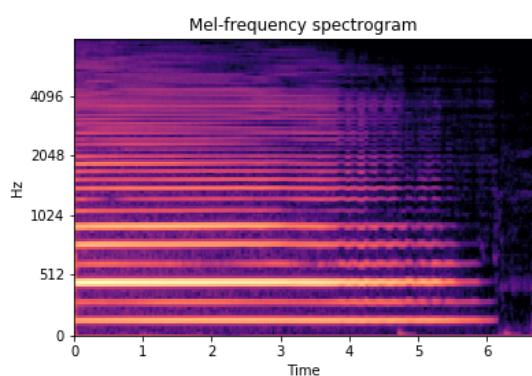
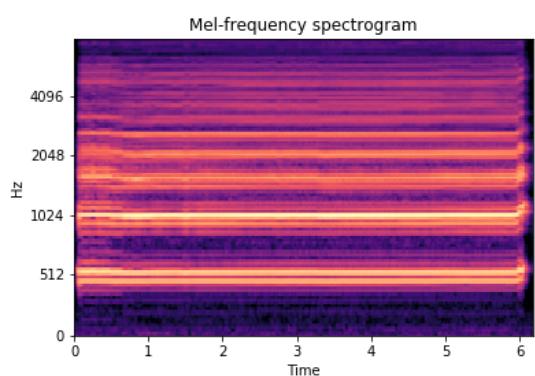
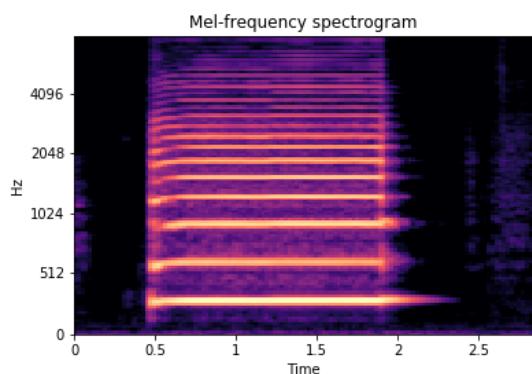
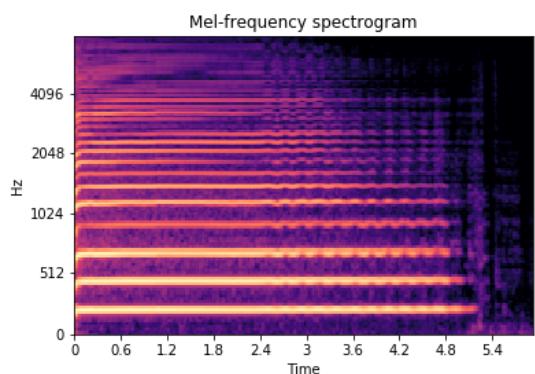
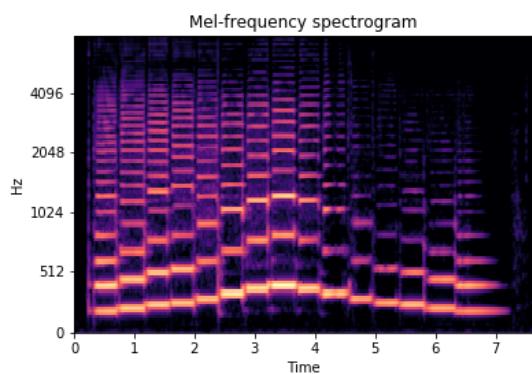
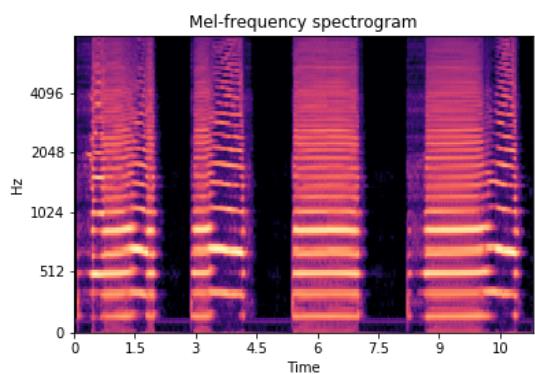


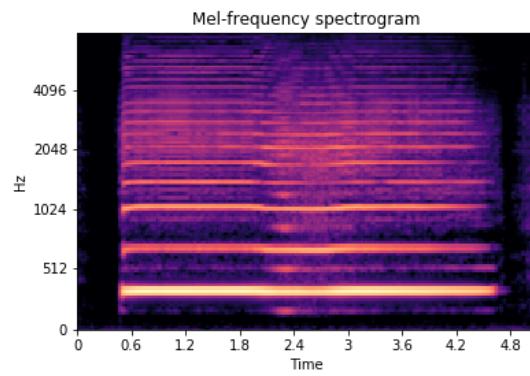
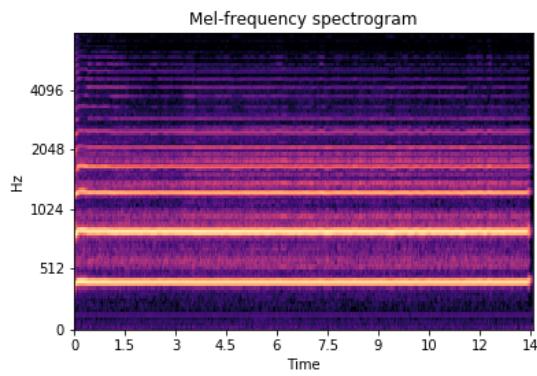
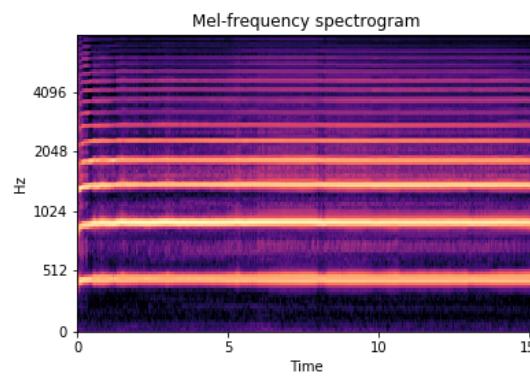
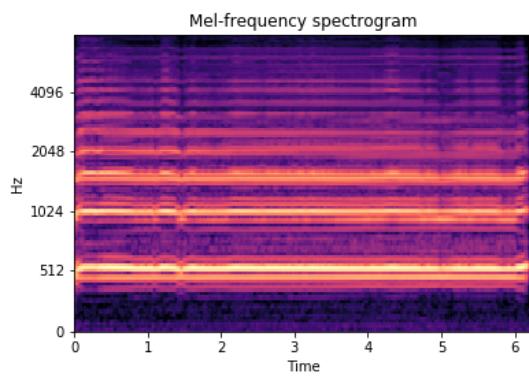
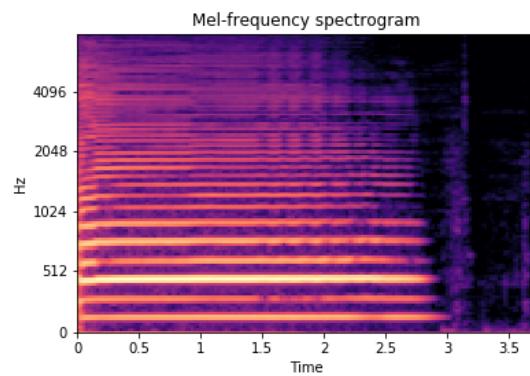
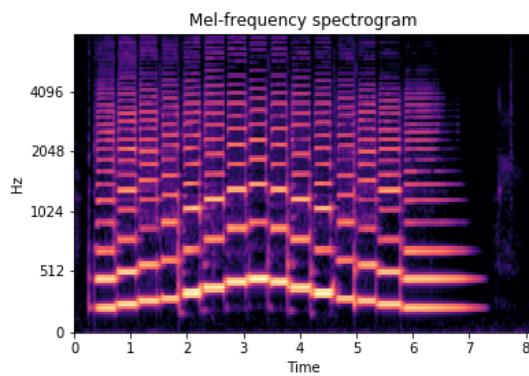
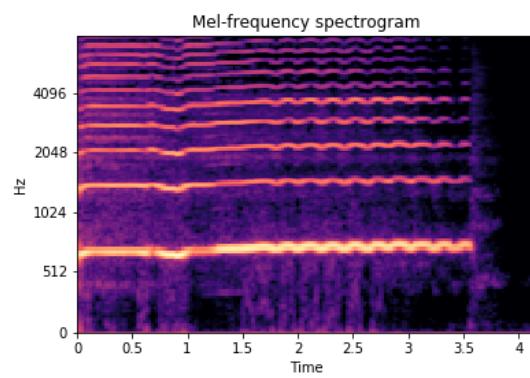
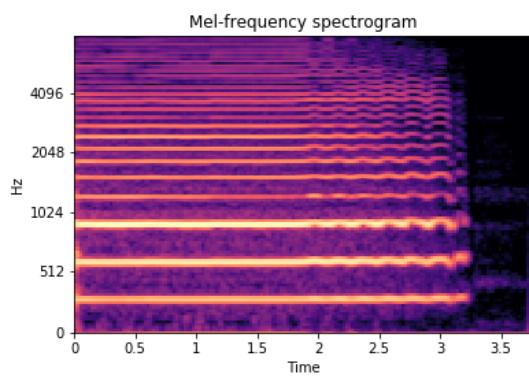


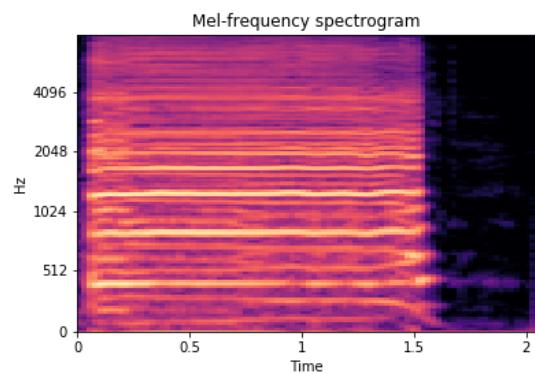
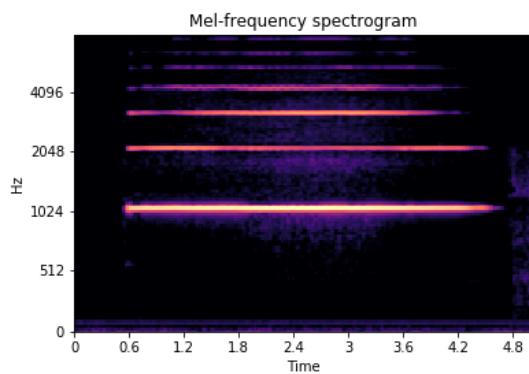
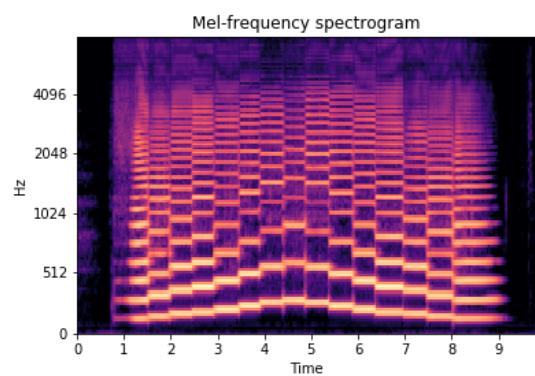
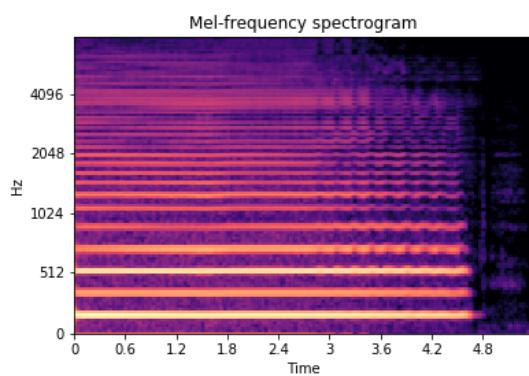
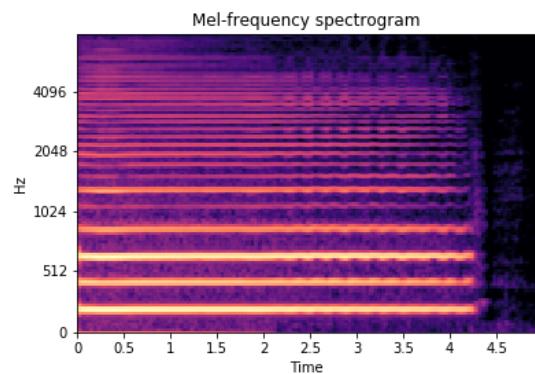
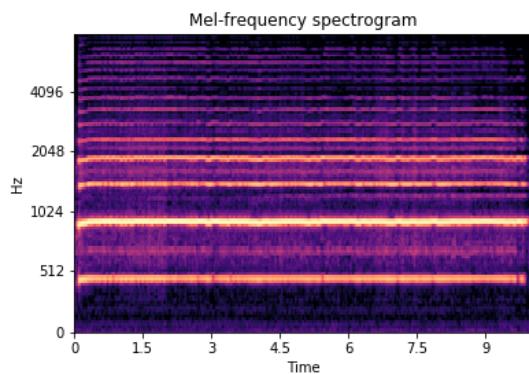
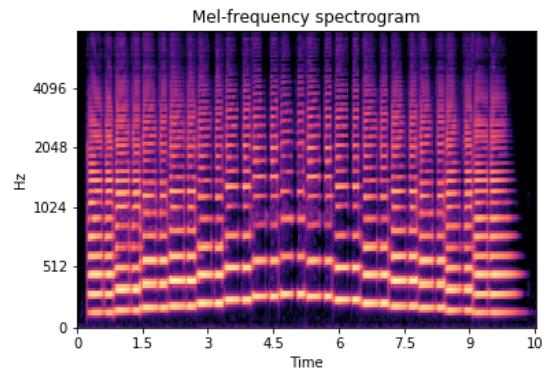
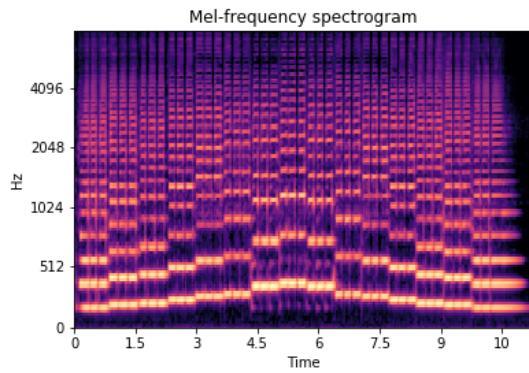


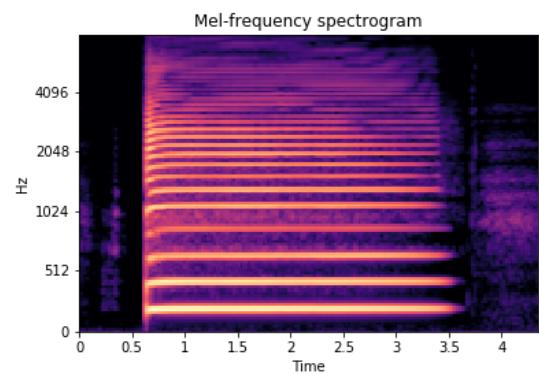
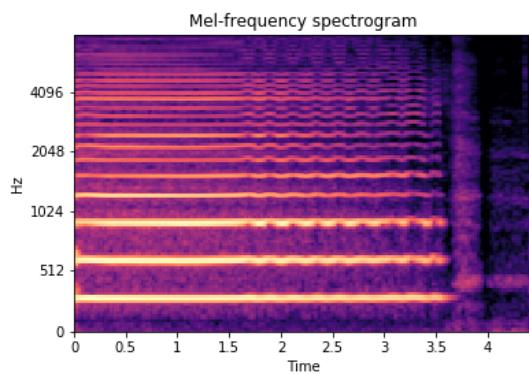
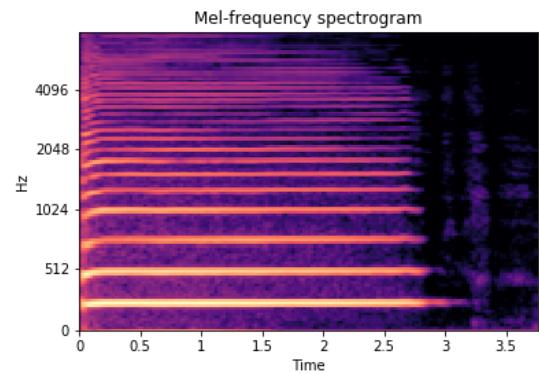
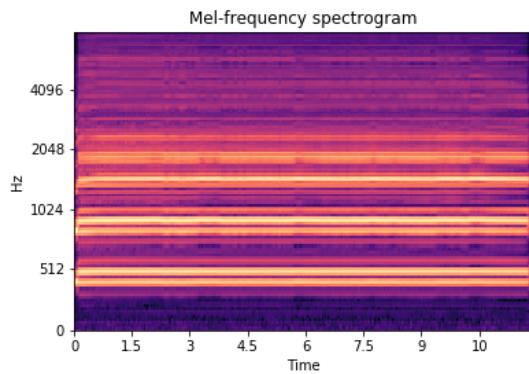
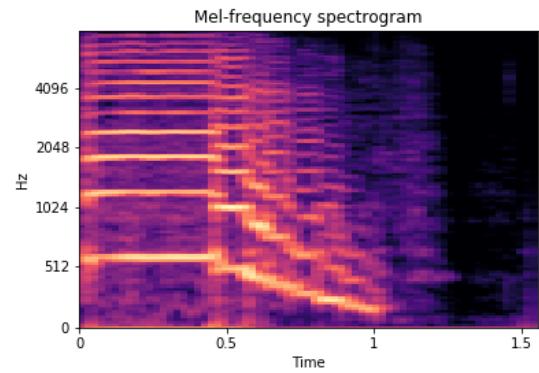
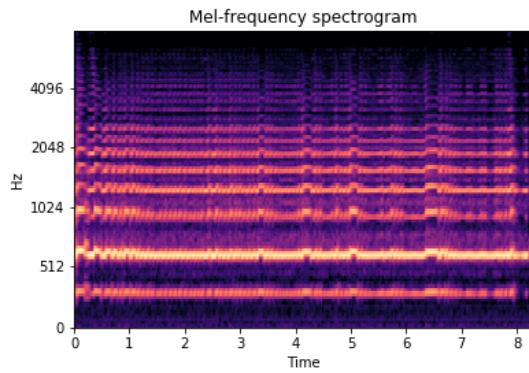


## Saxophone

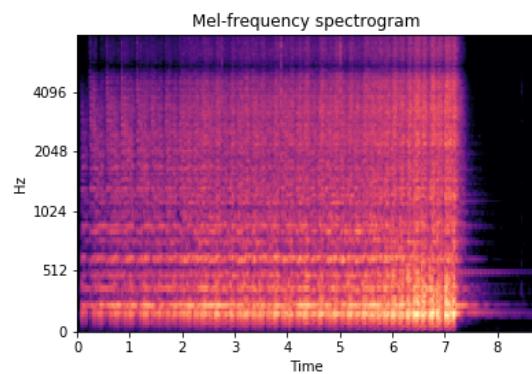
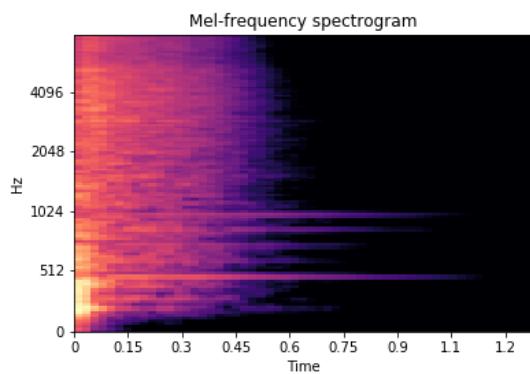
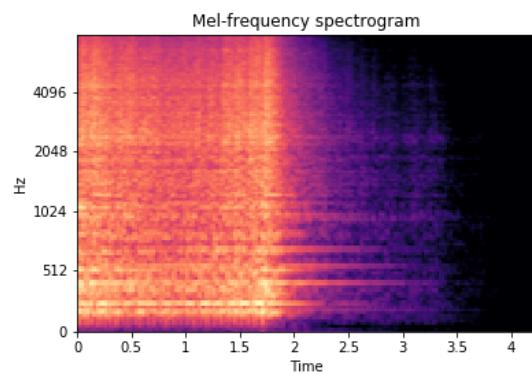
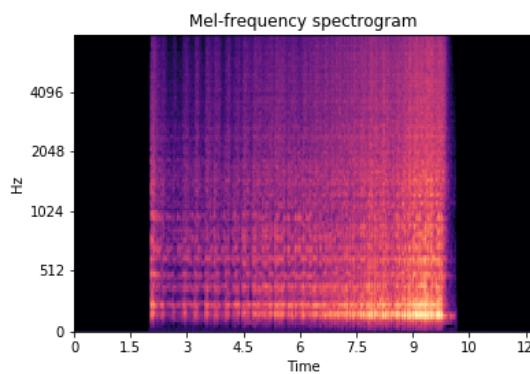
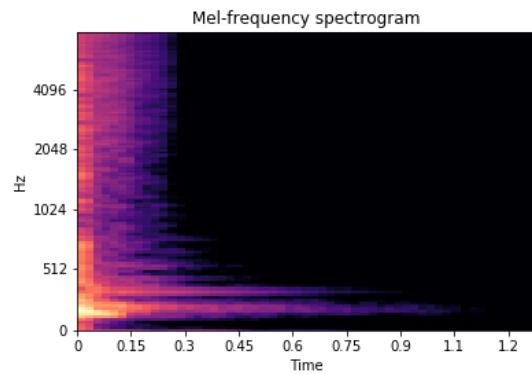
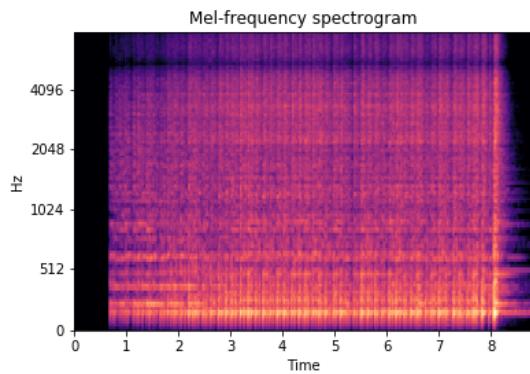
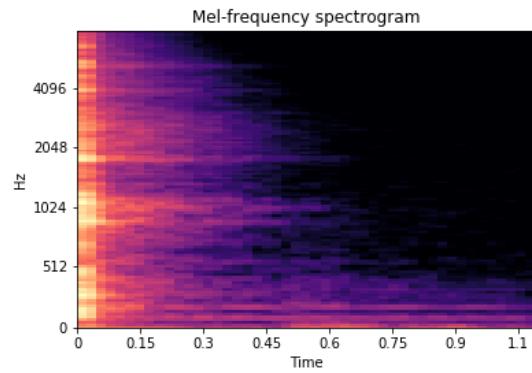
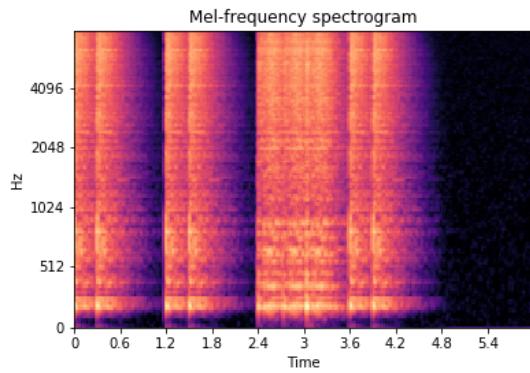


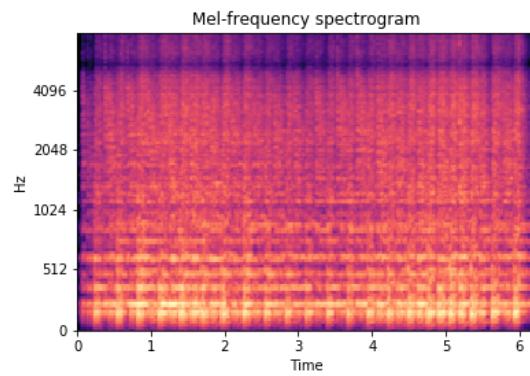
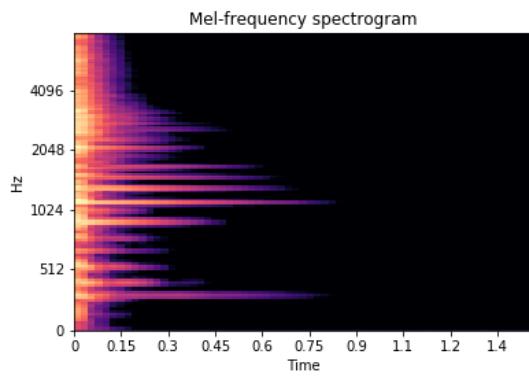
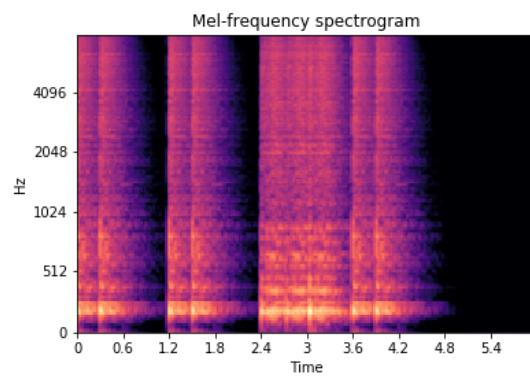
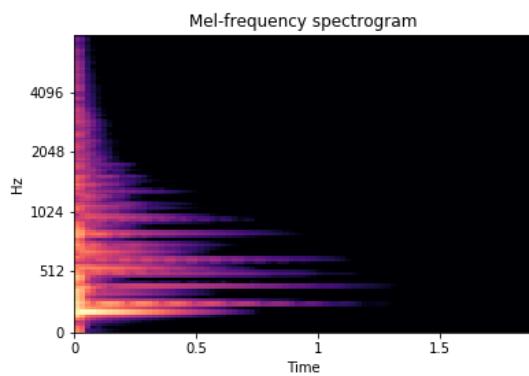
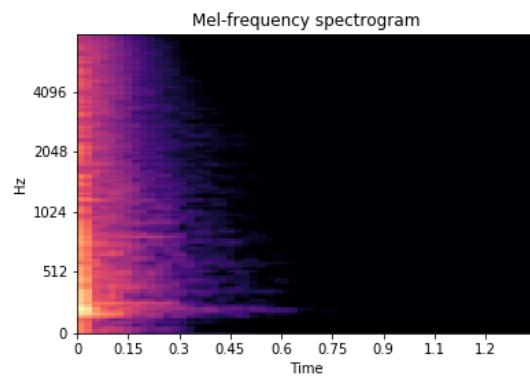
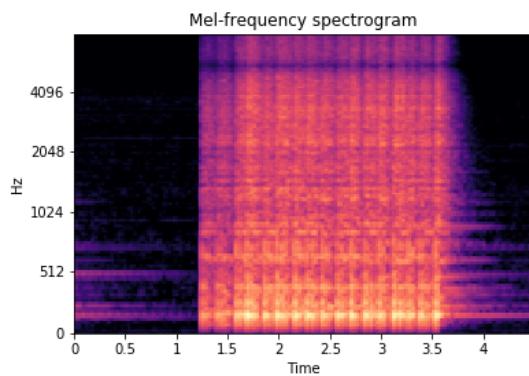
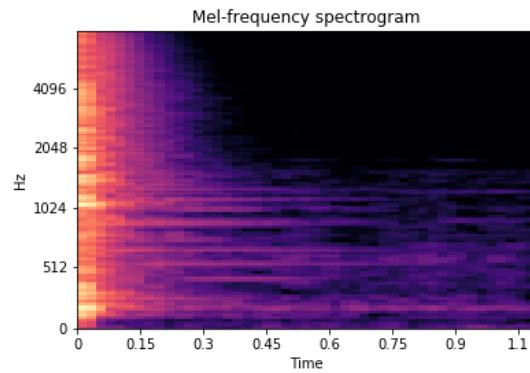
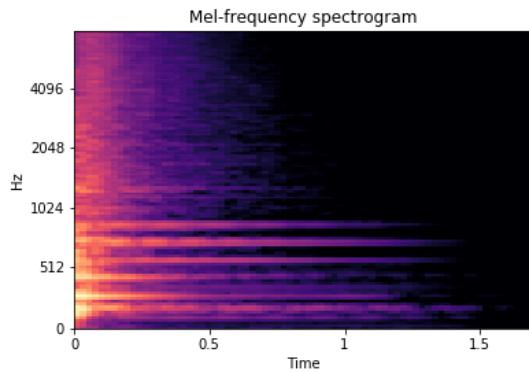


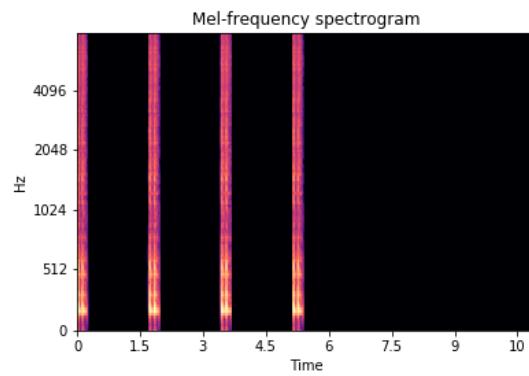
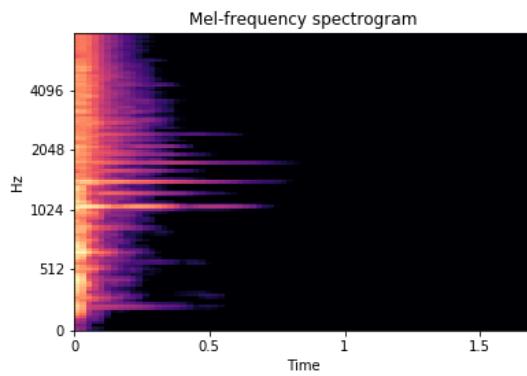
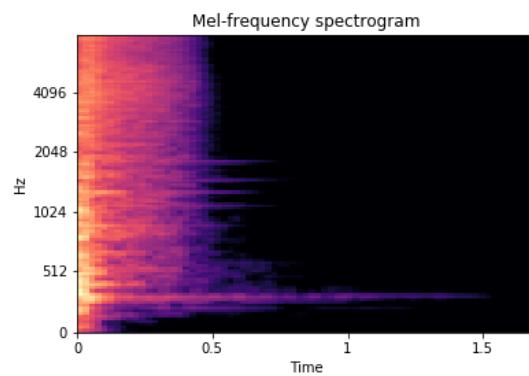
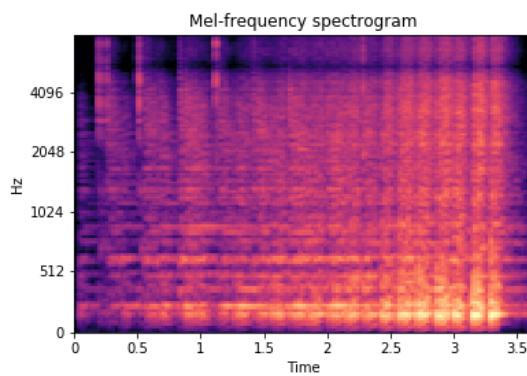
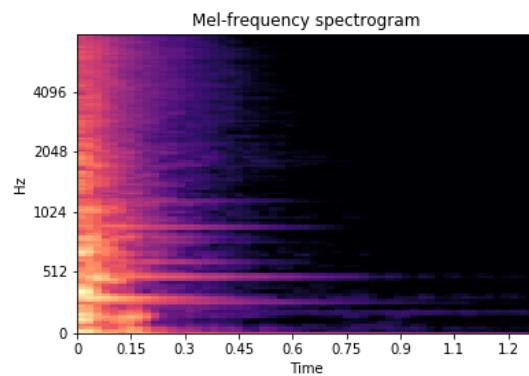
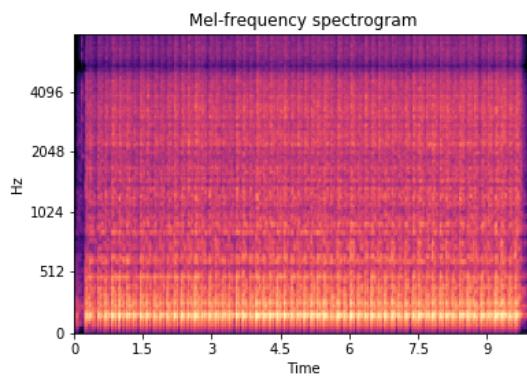
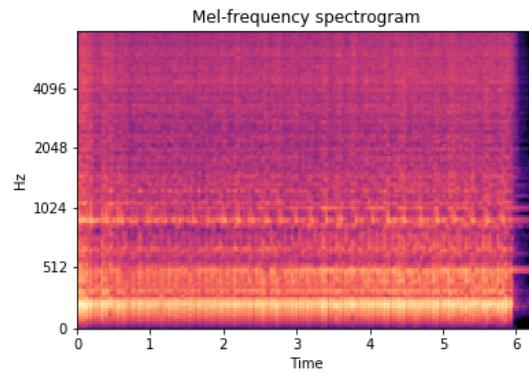
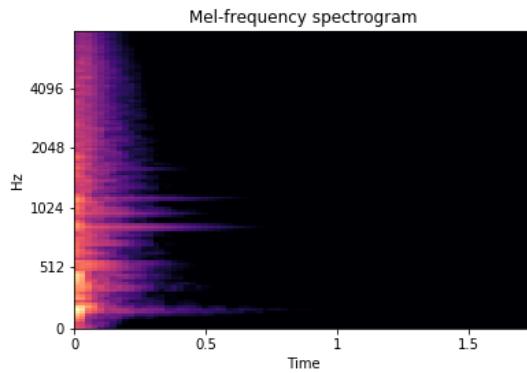


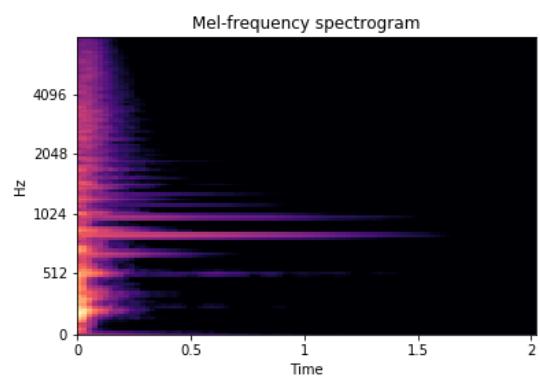
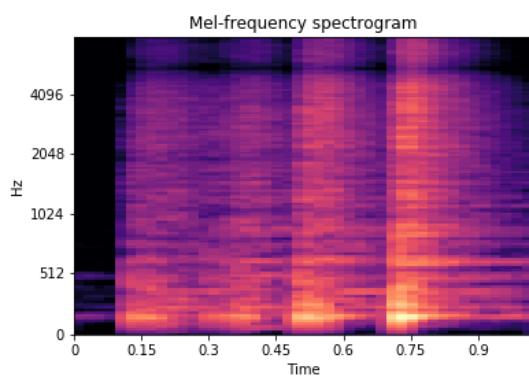
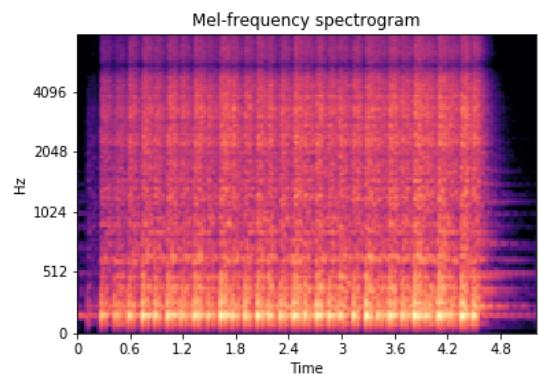
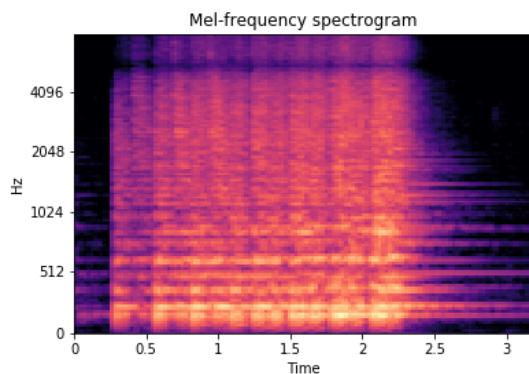
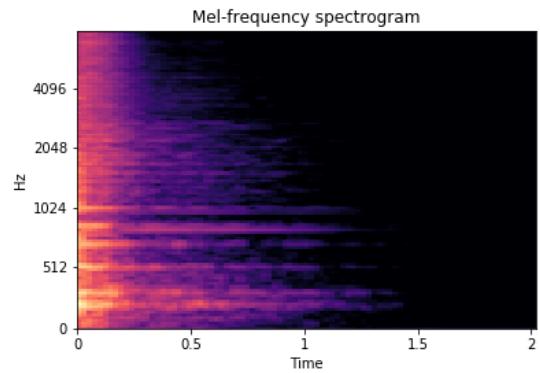
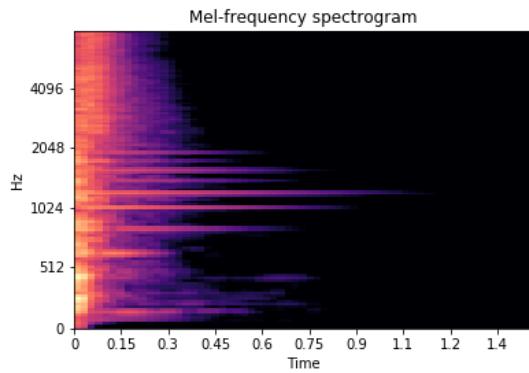


## Snare Drum









## Violin/Fiddle

