# Filters Implementation in Data Augmentation on Music Emotion Prediction

Duanning Wang
May 1st, 2024

# Introduction (background)

DATA LIMITATION !

emotion

Valence

Arousal

**Downstreams industries**:
Music Recommendation Systems,
Music Therapy Applications,
Artist and Repertoire (A&R) Development
…

# Augmentation

- Pitch shift
- Add background noise
- Reverb
- ***<u>Filters</u>***
  - Lowpass
  - Highpass
  - Bandpass
  - Peaking
  - Low shelf
  - High shelf

  Lowpass / Highpass filter types:
  - Butterworth
  - Chebyshev (type I)
  - Chebyshev (type II)
  - Elliptic

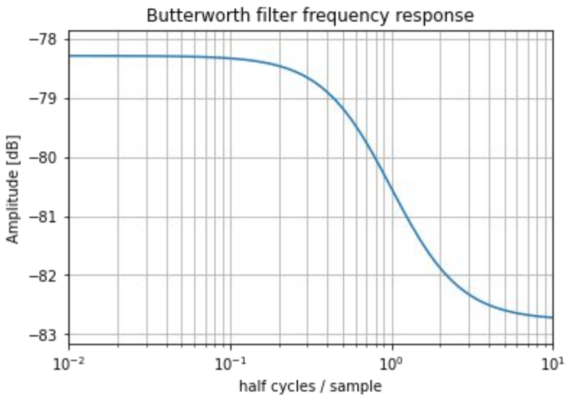
**Dataset: PMEmo**
includes 794 songs with both static and dynamic emotion labels (valence & arousal), chorus clips in MP3 format

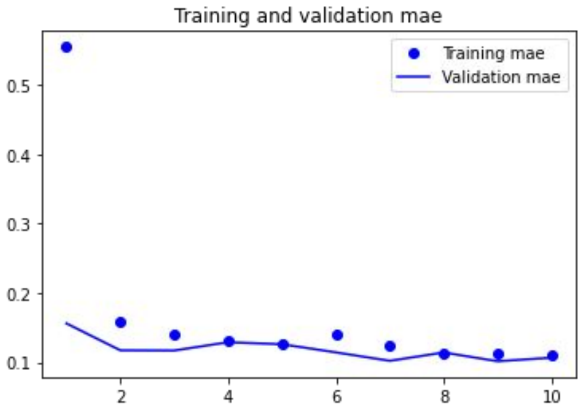
# Baseline results

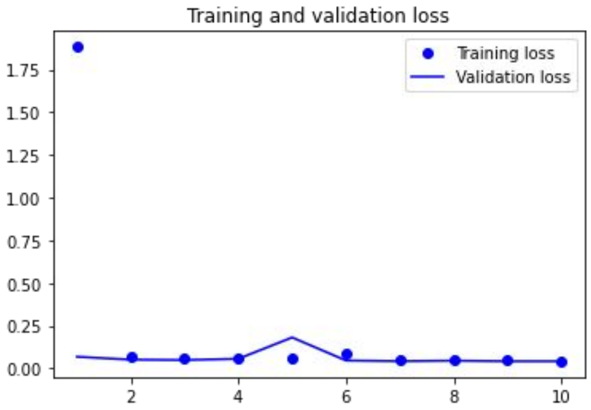
| | Test Loss<br>(Mean squared error) | Test MAE<br>(Mean absolute error) |
|---|---|---|
| *Original* | 0.036 | 0.160 |
| Pitch shift | 0.032 | 0.146 |
| Background noise | 0.034 | 0.157 |
| Reverb | 0.031 | 0.145 |

# Filters (lowpass)

Butterworth filter: (order = 4, frequency gain drop = 1000Hz)



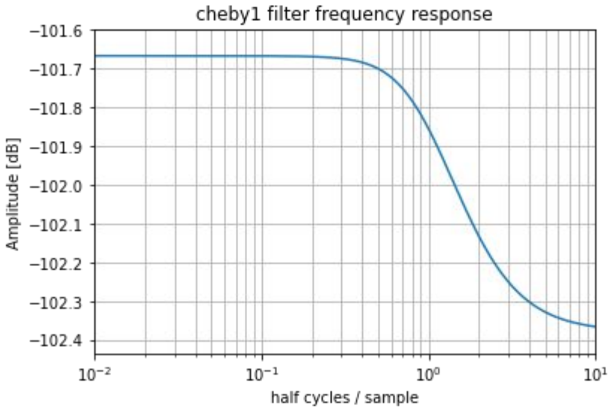



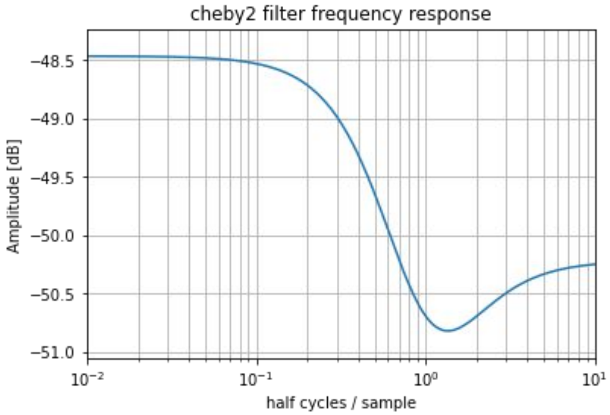




Test Loss: 0.04165396839380264, Test MAE: 0.10684800893068314

# Filters (lowpass)

Chebyshev (type I) filter: (order = 4, maximum ripple allowed below unity gain in the passband = 5dB)
Chebyshev (type II) filter: (order = 4, minimum attenuation required in the stop band = 50dB)



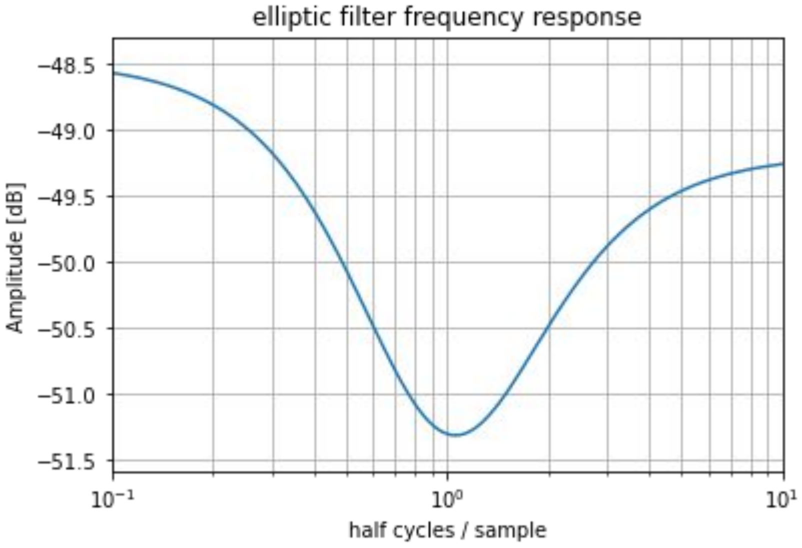




Chebyshev I:
Test Loss: 0.05766607075929642,
Test MAE: 0.12264931201934814

Chebyshev II:
Test Loss: 0.06217707321047783,
Test MAE: 0.12897001206874847

# Filters (lowpass)

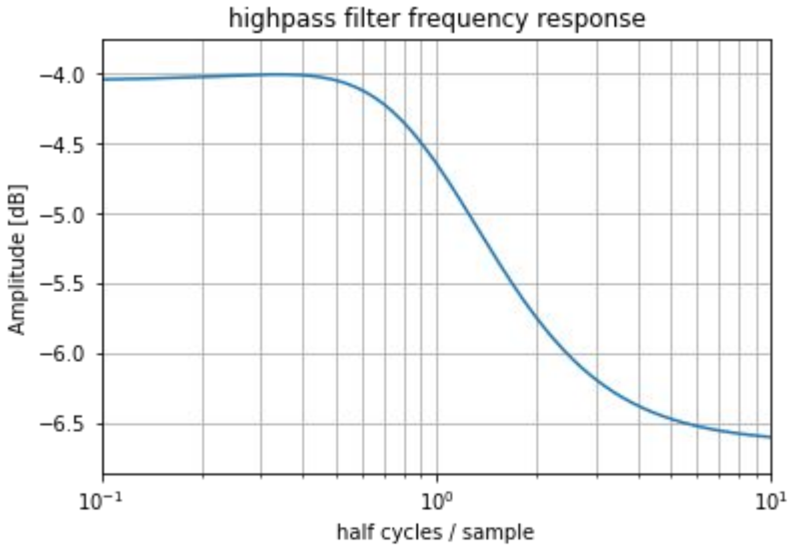
Elliptic filter: (order = 4, maximum ripple allowed below unity gain in the passband = 5dB, minimum attenuation required in the stop band = 50dB)




Test Loss: 0.0508585125207901,
Test MAE: 0.12048794329166412

# Filters (highpass)

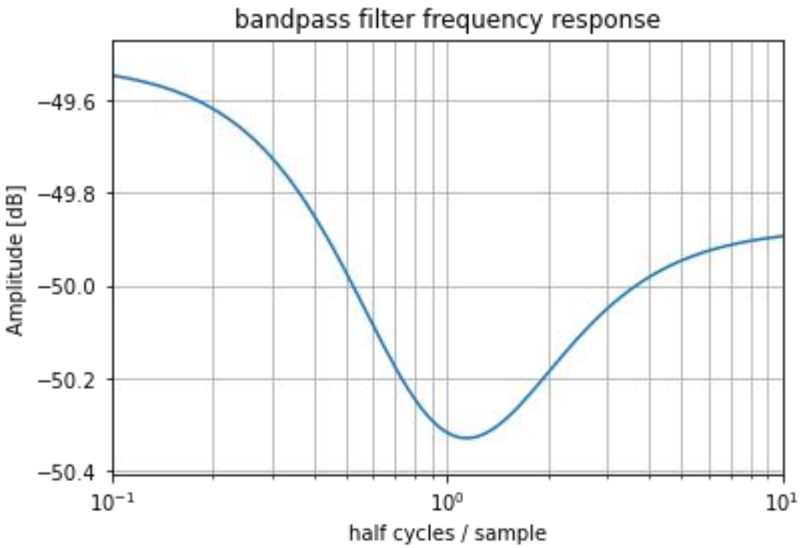
Elliptic filter (same parameters)




Test Loss: 0.05013992264866829,
Test MAE: 0.1088569238781929

# Filters (bandpass)

Elliptic filter



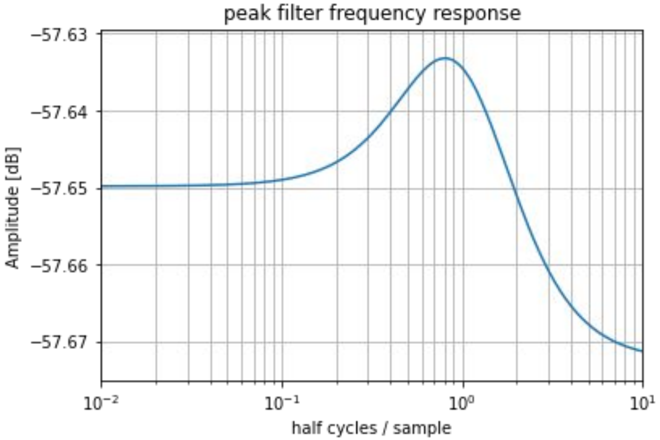

Test Loss: 0.05656576529145241,
Test MAE: 0.11583404242992401

# Filters (peaking)

(Frequency to be retained in a signal = 200Hz,

Quality factor -

characterizes peak filter -3 dB bandwidth bw relative to its center frequency, Q = w0/bw =30)




Test Loss:
0.050447650253772736,
Test MAE:
0.11729744076728821

# Filter (high shelf & low shelf)

min_center_freq=100, max_center_freq=1000

High shelf:

$$H(s) = A\frac{As^2 + \frac{\sqrt{A}}{Q}s + 1}{s^2 + \frac{\sqrt{A}}{Q}s + A}$$

Test Loss:
0.08168621361255646,
Test MAE:
0.1257268637418747

Low shelf:

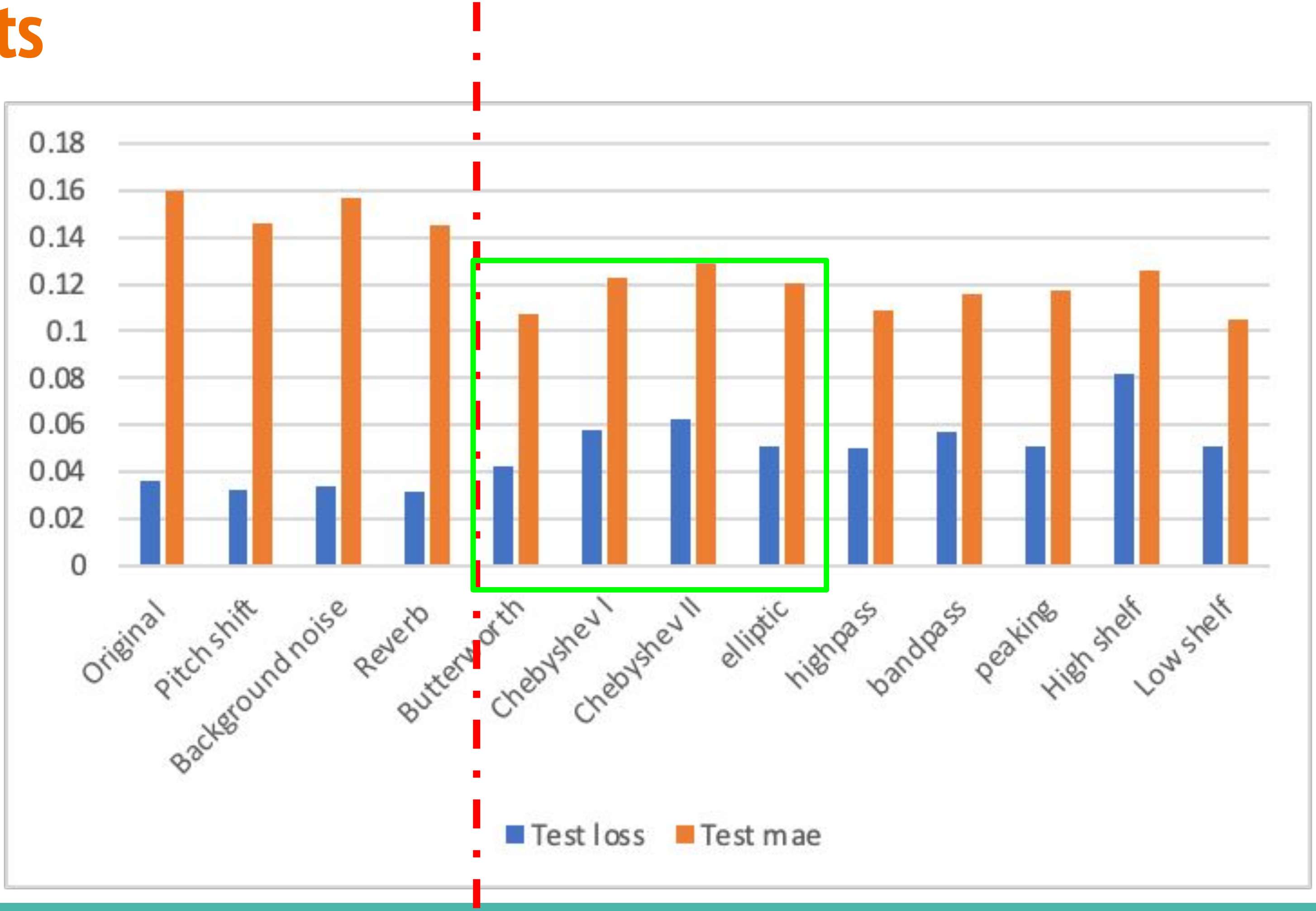
$$H(s) = A\frac{s^2 + \frac{\sqrt{A}}{Q}s + A}{As^2 + \frac{\sqrt{A}}{Q}s + 1}$$

Test Loss:
0.05099983885884285,
Test MAE:
0.10496504604816437

# results

| | | Test loss | Test mae |
|---|---|---|---|
| lowpass | butterworth | 0.04165396839380264 | 0.10684800893068314 |
| | Chebyshev I | 0.05766607075929642 | 0.12264931201934814 |
| | Chebyshev II | 0.06217707321047783 | 0.12897001206874847 |
| | elliptic | 0.0508585125207901 | 0.12048794329166412 |
| | highpass | 0.05013992264866829 | 0.1088569238781929 |
| | bandpass | 0.05656576529145241 | 0.11583404242992401 |
| | peaking | 0.05044765025377273 | 0.11729744076728821 |
| | High shelf | 0.08168621361255646 | 0.1257268637418747 |
| | Low shelf | 0.0509983885884285 | 0.10496504604816437 |

# results


0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Original
Pitch shift
Background noise
Reverb
Butterworth
Chebyshev I
Chebyshev II
elliptic
highpass
bandpass
peaking
High shelf
Low shelf
Test loss
Test mae

# conclusion

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

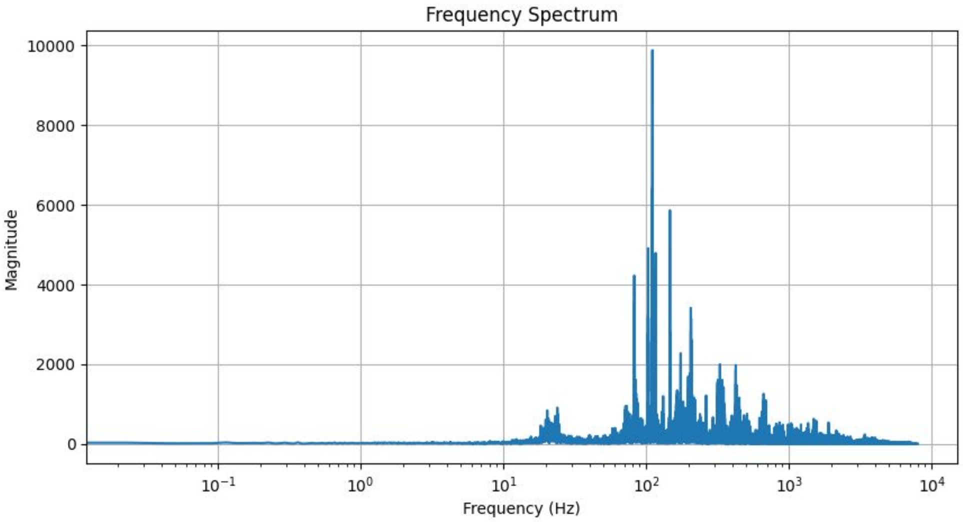
$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MSE is more sensitive to outliers. MSE might be more suitable for applications where it is critical to penalize large errors more heavily and when the data is relatively free of outliers.

1. **Using filters to do the augmentation leads to some notable outliers, but the overall performances are better.**
2. All of filters contributes to the improvement. Butterworth lowpass filter performs the best.
3. The results of applying low-pass filters have no commonality. The results depend on the specific situations.

# Future Thoughts

- The filters can be designed based on audio's characteristics dynamically.




- Combination of different filters.

# Thank you!