# Types Audio Features for ML

Valerio Velardo

# Why audio features?

- Description of sound

# Why audio features?

- Description of sound
- Different features capture different aspects of sound

# Why audio features?

- Description of sound
- Different features capture different aspects of sound
- Build intelligent audio systems

# Audio feature categorisation

- Level of abstraction
- Temporal scope
- Music aspect
- Signal domain
- ML approach

# Level of abstraction

## High-level

Examples: instrumentation, key, chords, melody, rhythm, tempo, lyrics, genre, mood

## Mid-level

Examples: pitch- and beat-related descriptors, such as note onsets, fluctuation patterns, MFCCs


## Low-level

Examples: amplitude envelope, energy, spectral centroid, spectral flux, zero-crossing rate

Knees, P., & Schedl, M. (2016). *Music similarity and retrieval: an introduction to audio-and web-based strategies*

# Temporal scope

- Instantaneous (~50ms)
- Segment-level (seconds)
- Global

# Music aspect

- Beat
- Timbre
- Pitch
- Harmony
- ...

# Signal domain

# Signal domain

- Time domain

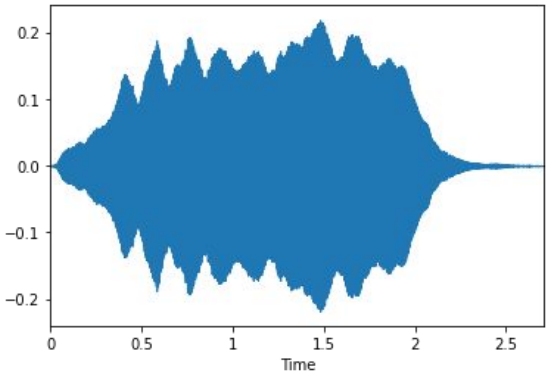
# Signal domain

- Time domain

Amplitude envelope
Root-mean square energy
Zero crossing rate

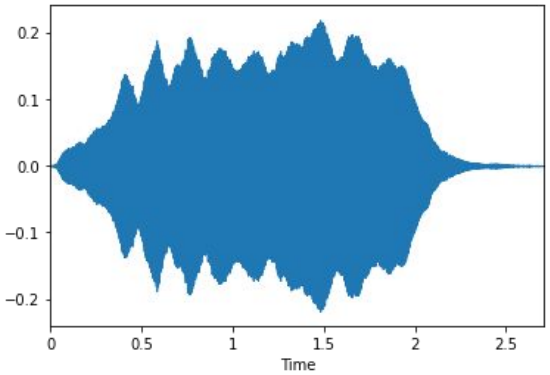
# Signal domain

- Time domain



# Signal domain

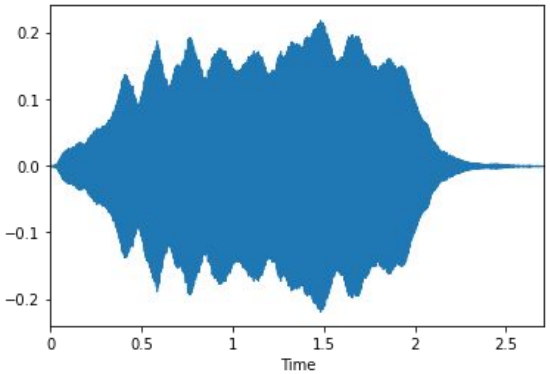
- Time domain
- Frequency domain



# Signal domain

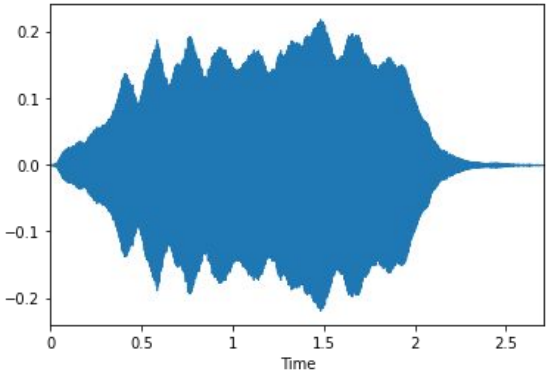
- Time domain
- Frequency domain

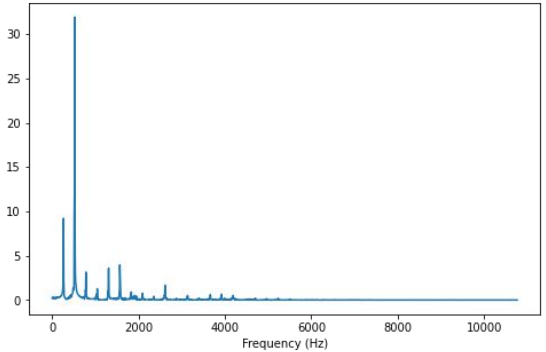
Band energy ratio
Spectral centroid
Spectral flux



# Signal domain

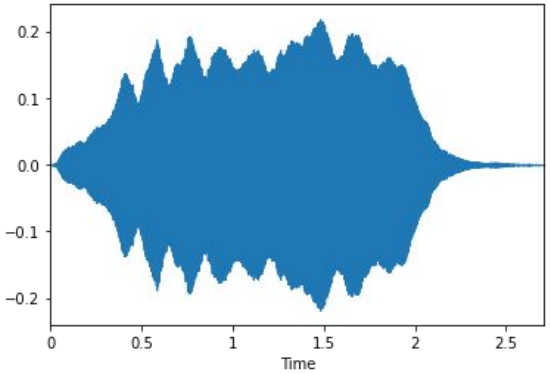
- Time domain
- Frequency domain

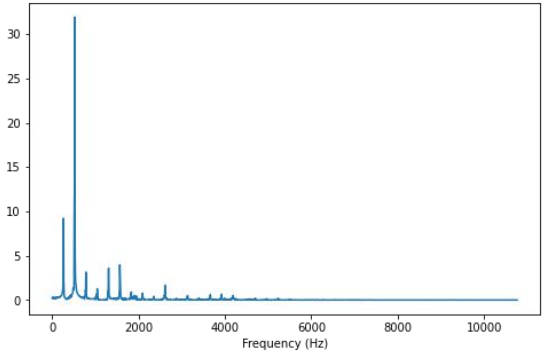


# Signal domain

- Time domain
- Frequency domain
- Time-frequency representation

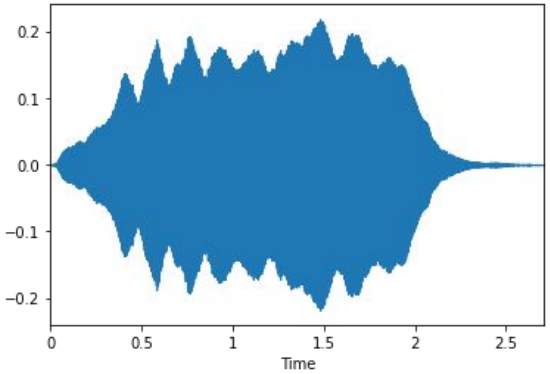


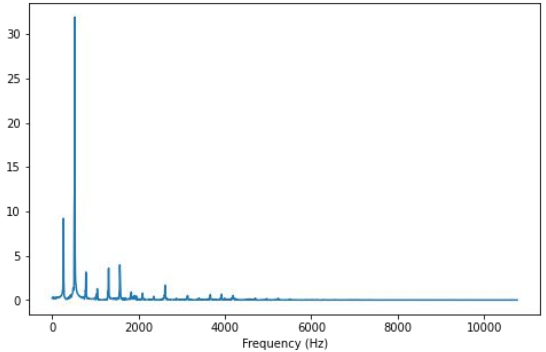



# Signal domain

- Time domain
- Frequency domain
- Time-frequency representation

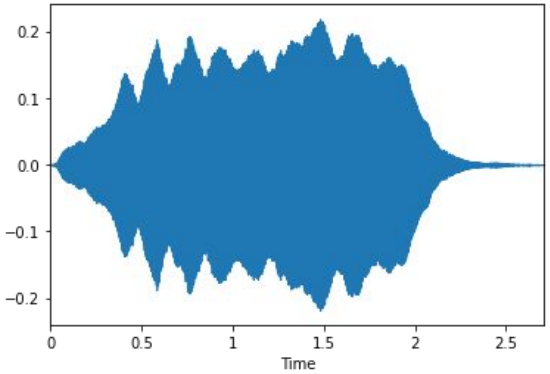
Spectrogram
Mel-spectrogram
Constant-Q transform

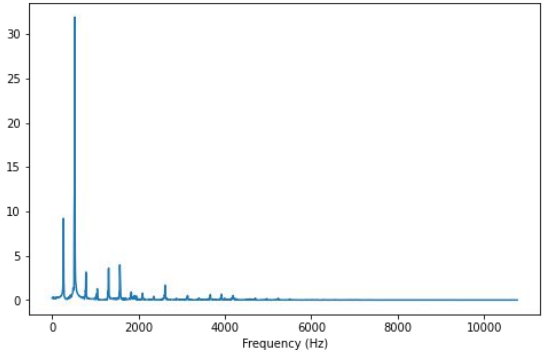


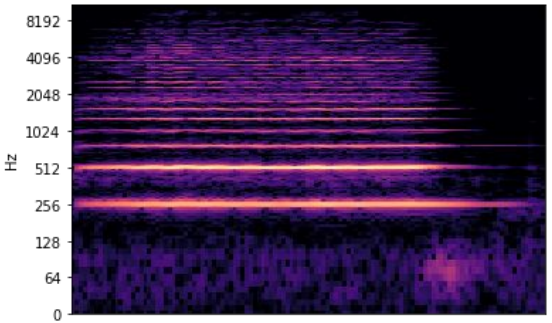



# Signal domain

- Time domain
- Frequency domain
- Time-frequency representation



# Machine learning approach

- Traditional machine learning
- Deep learning

# Traditional ML

Amplitude envelope
Root-mean square energy
Zero crossing rate
Band energy ratio
Spectral centroid
Spectral flux
Spectral spread
Spectral roll-off
...

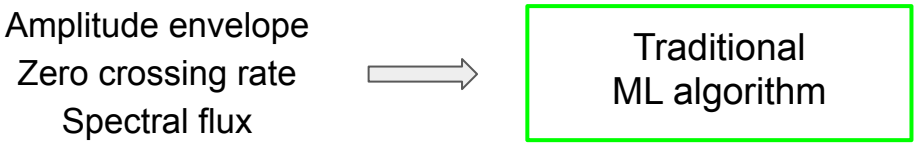
# Traditional ML

Amplitude envelope
Root-mean square energy
Zero crossing rate
Band energy ratio
Spectral centroid
Spectral flux
Spectral spread
Spectral roll-off
...

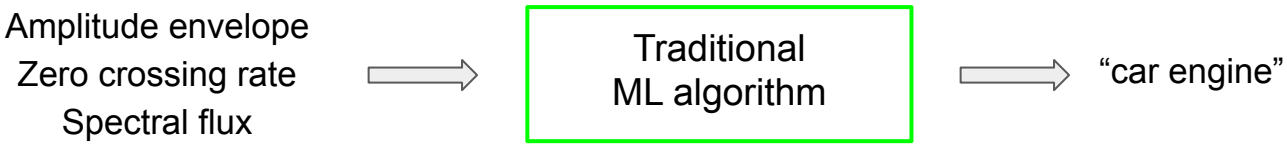
# Traditional ML

Amplitude envelope
Zero crossing rate
Spectral flux

# Traditional ML

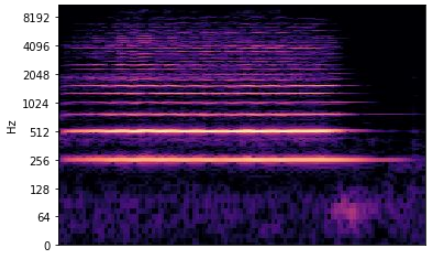


# Traditional ML

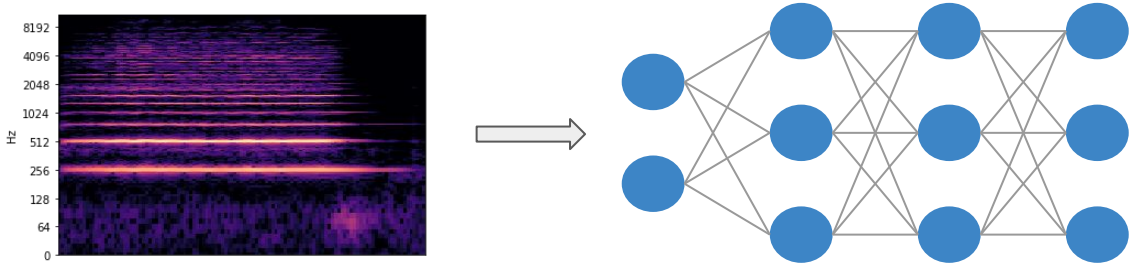


# Deep Learning

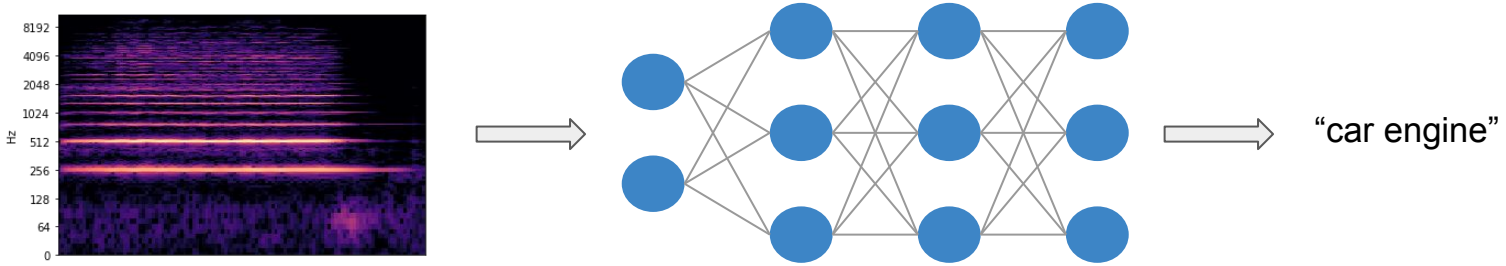
# Deep Learning



# Deep Learning



# Deep Learning

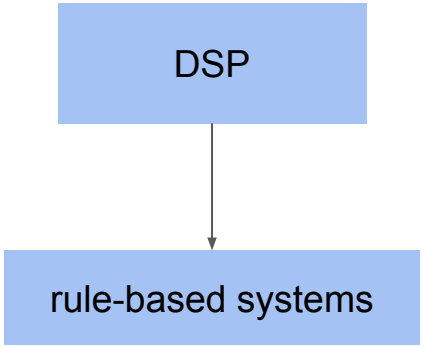


# Types of intelligent audio systems

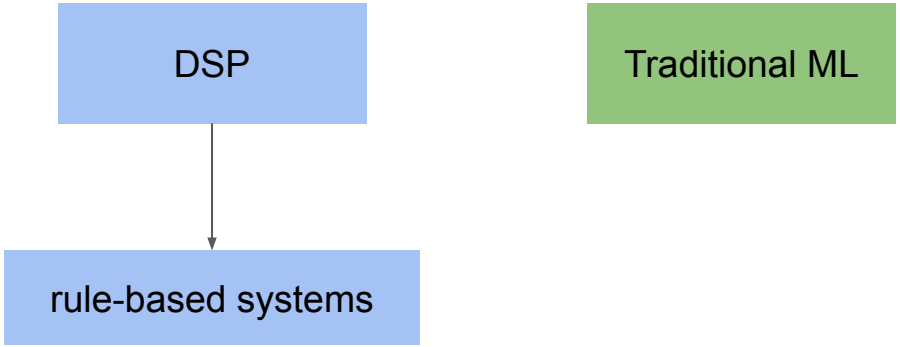
# Types of intelligent audio systems

DSP

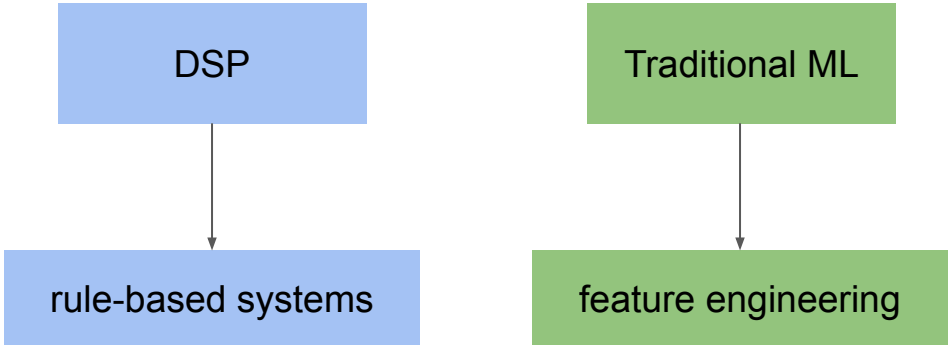
# Types of intelligent audio systems



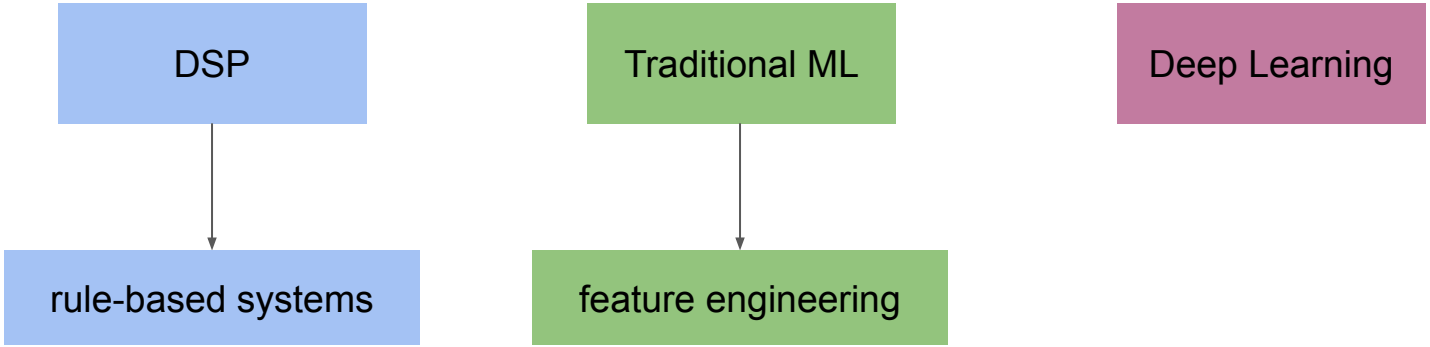
# Types of intelligent audio systems



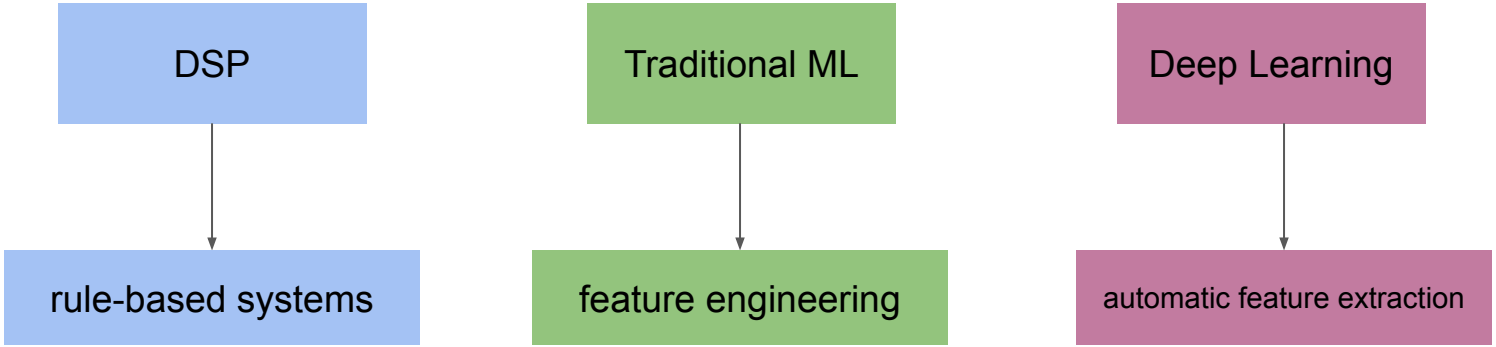
# Types of intelligent audio systems



# Types of intelligent audio systems



# Types of intelligent audio systems



# What’s up next?

- Feature extraction pipeline

# Join the community!

thesoundofai.slack.com