

Processing Audio and Creating Music in Python

Matt Segall

Background

- Been playing music since 8
- Went to college for it
- Integral part of my identity
- Curious about ML with audio & music
- Scared a computer might take my job



Coltrane Robot



So, Let's....

1. Check out some libraries to process and analyze audio
2. Explore their applications
3. If time, listen to and discuss some AI-generated music
 - a. Ultimately, maybe answer the question: **Should** musicians be scared for their jobs?

PyDub

- Intuitive audio manipulation!
- Like a very basic DAW?
 - Slice
 - Loop
 - Fade
- Make playlists and DJ
- Basic feature extraction
 - Channels
 - Duration
 - To array

Pydub by [jiaaro](#)

Librosa

- Analysis
 - Detecting rhythmic/beat events
 - Visualizations
- More powerful feature extraction
 - Spectral
 - Rhythmic
 - Decompositions (using sci-kit!)
- Effects
 - Pitch and time shift

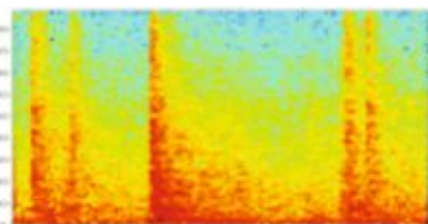


DEMOS

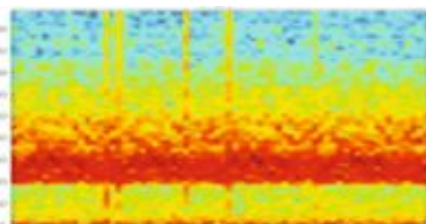
Use Cases

- Wildlife monitoring
 - Size of fish schools
 - Identify bird calls
- Security systems
 - Broken glass
- Genre classification
 - Content-based recommendation systems
- Noise pollution
- What else?

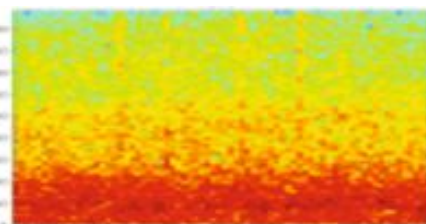




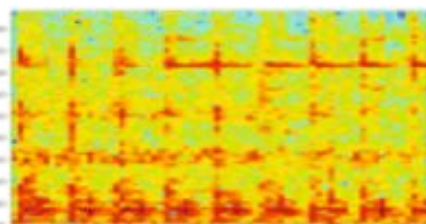
excavator



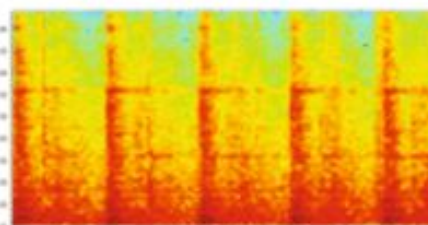
alarm



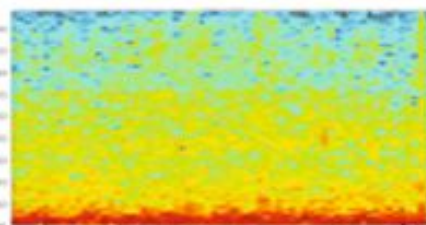
engine sound



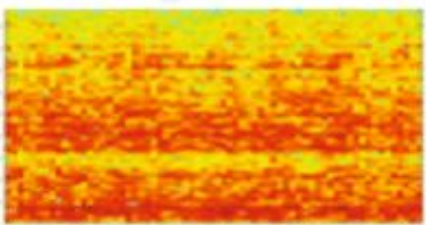
electric hammer



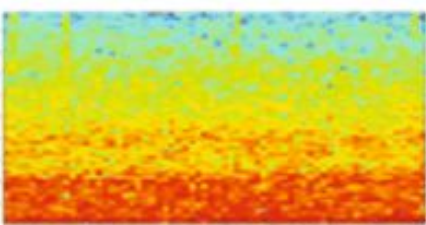
hydraulic hammer



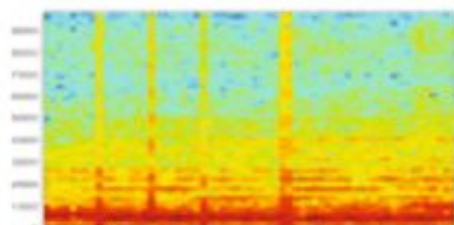
milling machine



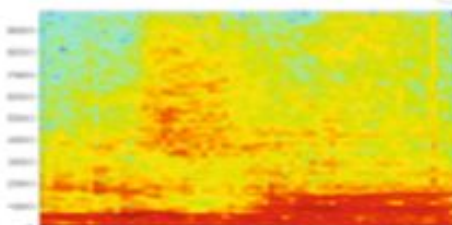
cutting machine



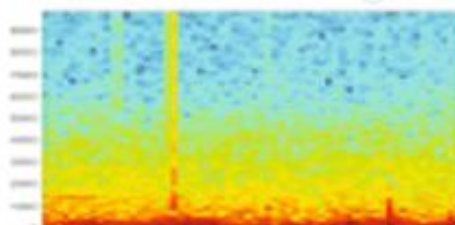
sound of generator



music



talking



wind noise

Similar Techniques Applied To

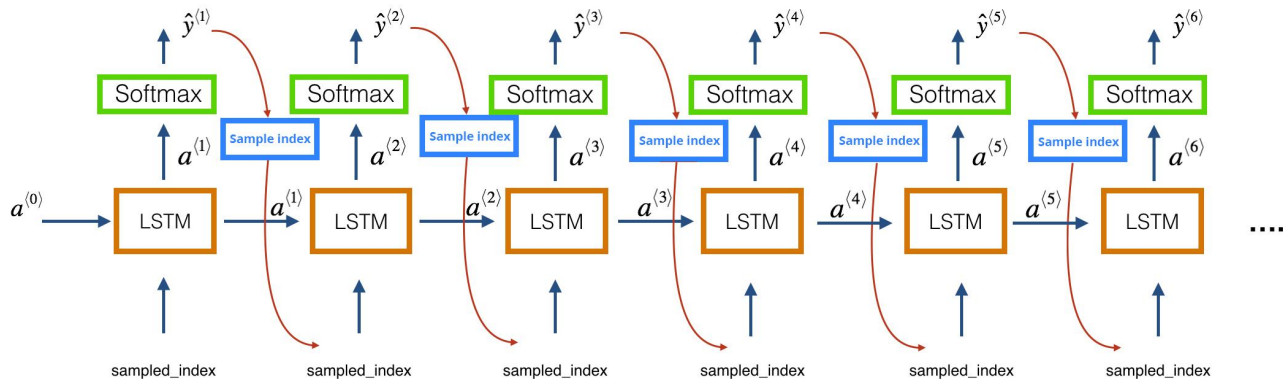
- Seismic Data
- Thermodynamics
- Sonar/radar
- Telecom
- Biomedical engineering
- TONS more...



AI Jazz?

Expertise / Jazz Solos Neural Nets

- Markov models compute a probability distribution for the “next note”
- Neural Nets can do this while taking into account more factors like
 - Higher volume of preceding notes
 - Preceding, current and upcoming harmony
- RNNs allow long term patterns and repetition to persist



Listening Examples

How does the machine represent real music well?

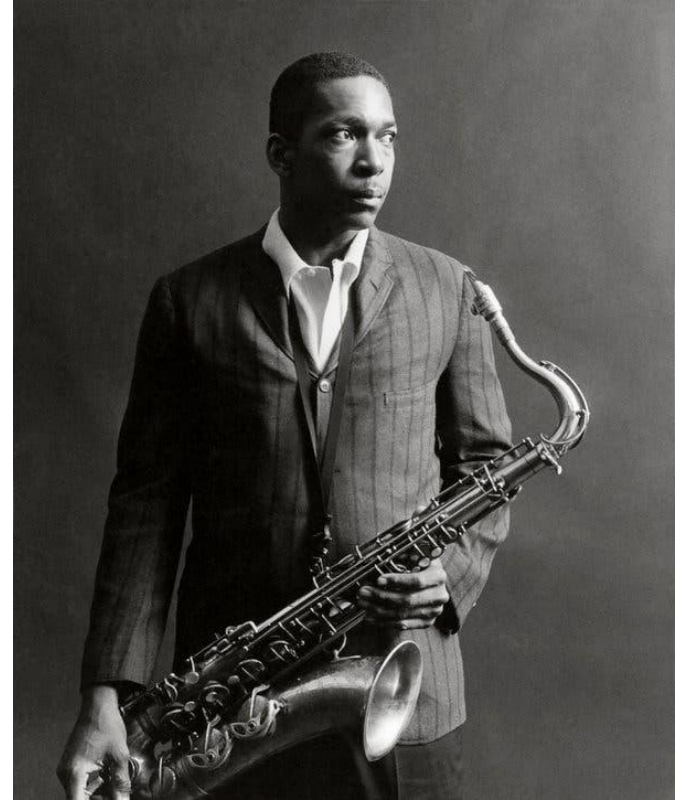


How does the machine represent real music well?

- “Correct” note choice
- Contour/phrasing
- Lots of novel melodic material
 - But NOT things real musicians would play

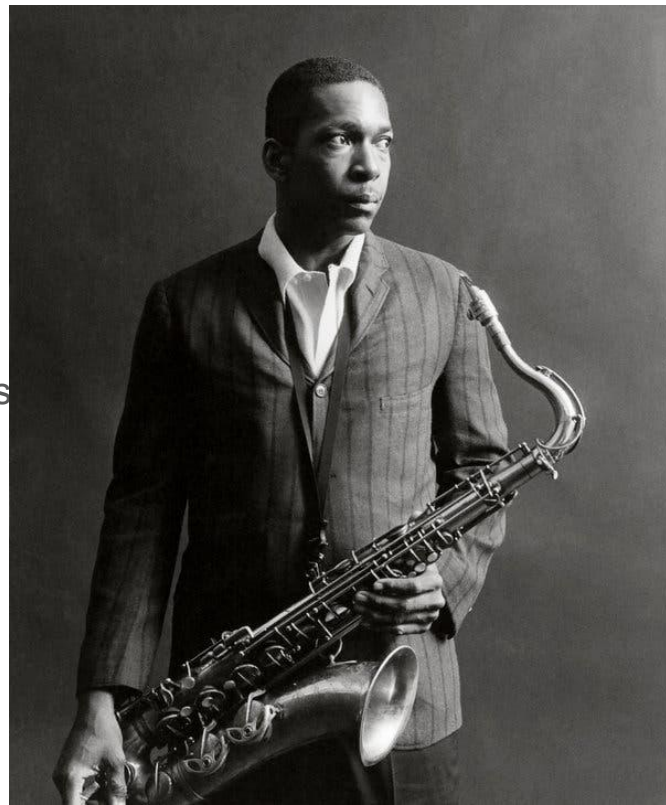


How does the machine NOT represent real music well?



How does the machine NOT represent real music well?

- Singable melodies
- Motivic development
- Groove
 - Especially jazz, is compellingly imperfect
- Expressiveness
 - Machines don't have experiences and feelings to express
 - How I feel when I take a solo
 - Like a chat bot that chats with you just to talk
- A LOT MORE than we can get into
- Smell familiar?
 - NLG



Should we be worried?

