

Announcements

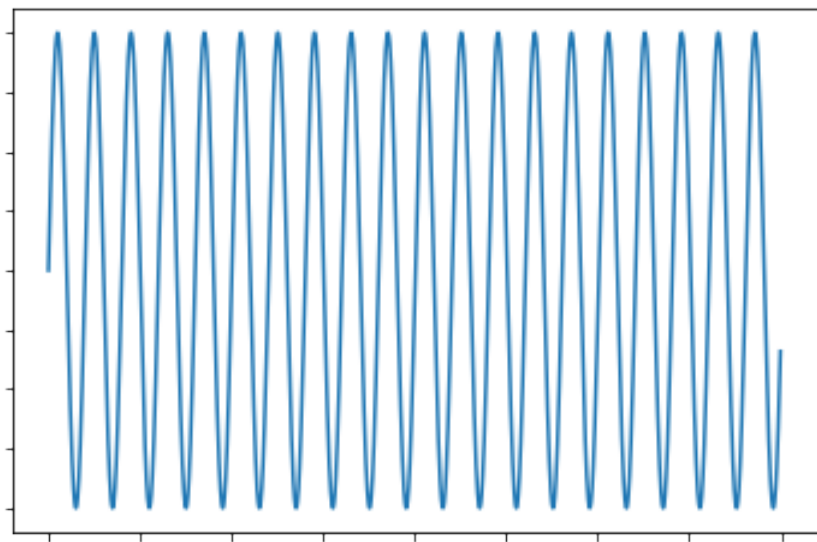
- Readings (Week 6 due Sunday 2/16)
 - Audio (online today!)
- **Assignment 02**
 - Due: Friday 2/21/2020
 - Email me: 1 iPython notebook on Text **or/and** Audio
- **Today** Audio
 - From a simple tone to music
 - Adding noise and echo
 - From a speech to text

Audio Processing

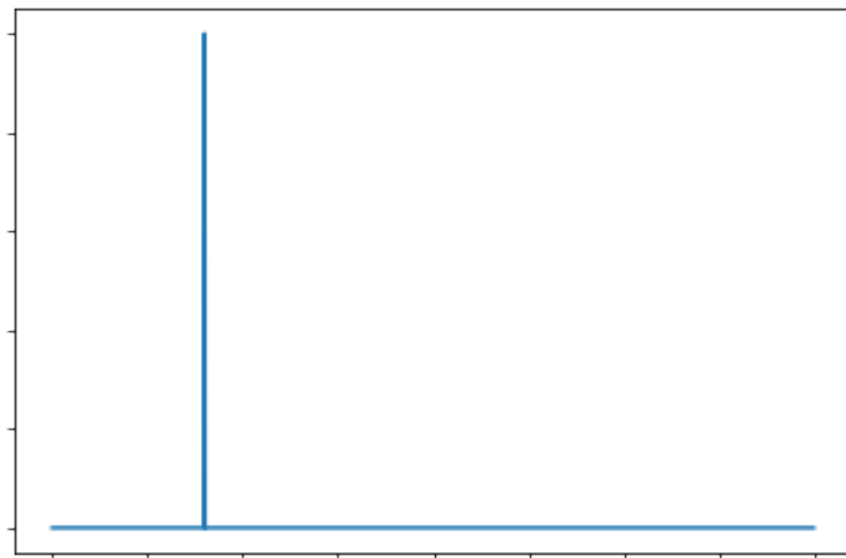


Audio processing

Time

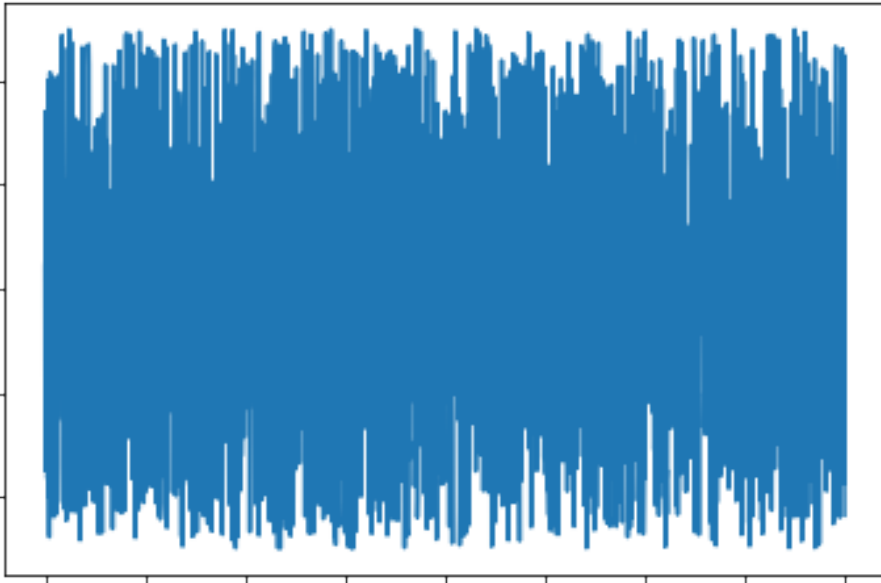


Frequency

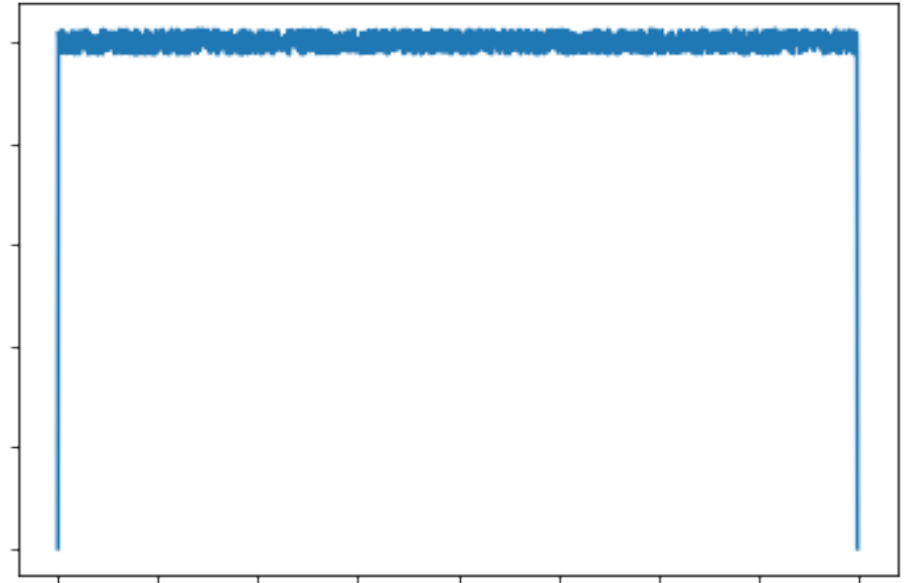


Audio processing

Time



Frequency



Digital Signal Processing

DSP is most strongly associated with linear transformations such as:

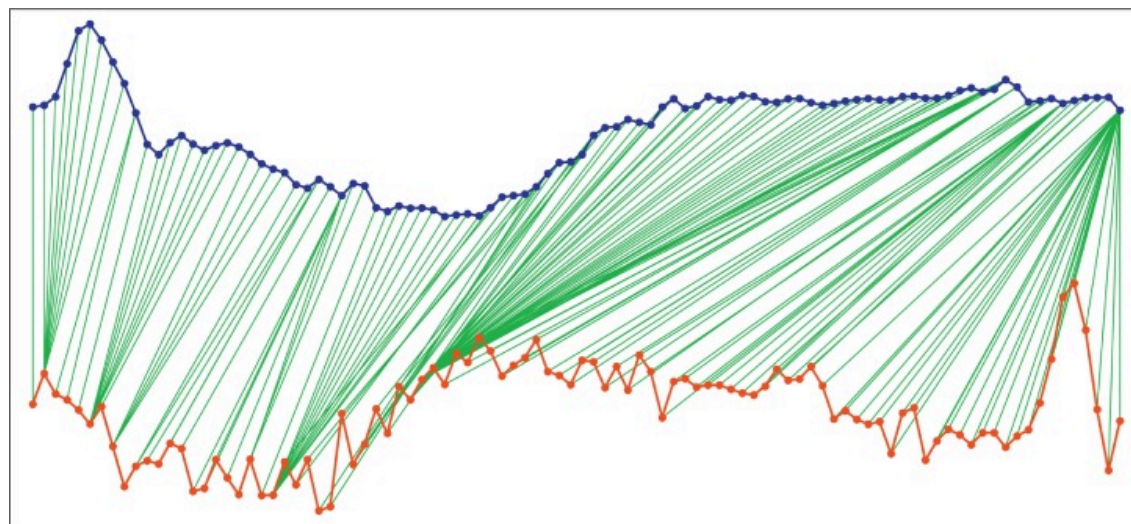
- **Filtering** (changing the relative magnitude of different frequencies)
- **Reverberation** (adding time delayed copies)

Other common DSP techniques for audio signals include:

- **Amplitude compression** (making everything loud)
- **Lossy data compression** (making files smaller and less accurate)
- Manipulations to make speech more intelligible or music sound better

Speech Recognition History

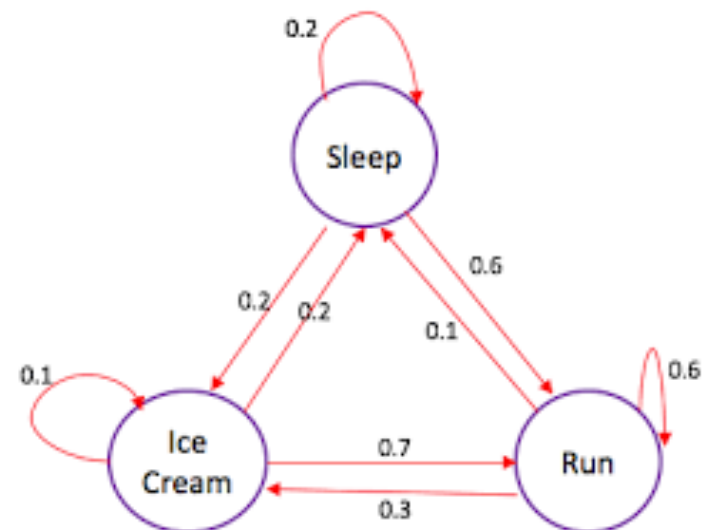
1960s Dynamic Time Warping (frames). 200 words (with isolated words)



Speech Recognition History

1960s Dynamic Time Warping (frames). 200 words (with isolated words)

1970s Markov Chains to encode language and syntax. Continuous speech at one hour of processing per minute.



Speech Recognition History

1960s Dynamic Time Warping (frames). 200 words

1970s Markov Chains to encode language and syntax of processing per minute.

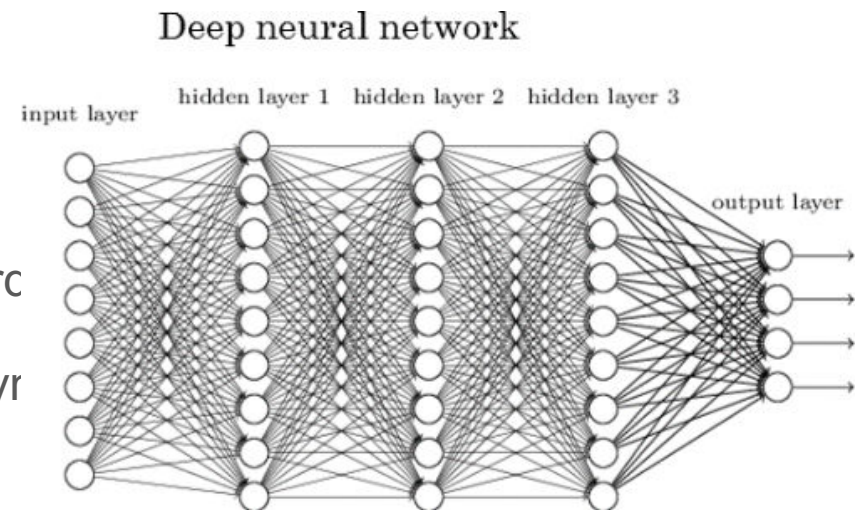
1980s n-grams and faster computers: 20,000 words, still not realtime.

1990s commercial success with limited vocabulary and special microprocessors

2000s labeled data gathering and the NSA.

Poor accuracy prevented widespread use.

2010s Deep learning, Siri, Alexa, Cortana, Google.



Why is speech recognition so hard?

Large vocabulary (proper nouns and jargon)

Speaker variability (accents!)

Word confusability

Context-dependency

Noise / reverberation

Conversational versus human-machine speech

Today

- `making_sounds.ipynb`
 - Create a sound (numpy array)
 - Listen to it
 - Create noise
 - Add sound and noise
 - Create a short song?
- `audio_modification.ipynb`
 - Listen to an audio file
 - Add some noise, save, download and listen
 - Remove the long pauses from the speech (speaking recognition)
 - Use a pre-built audio classification API, and explore the effect of noise
- Start your audio project in `A02_Audio_YourFirstName_YourLastName`