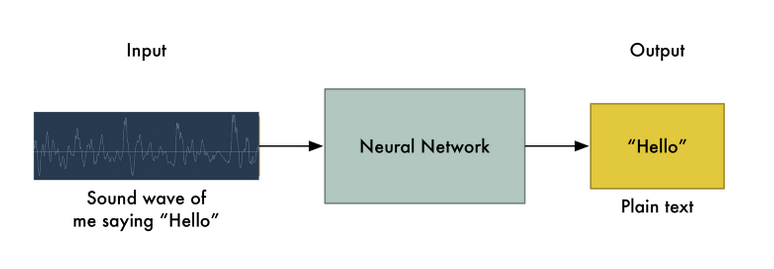
**Feature Extraction Techniques for an Audio Signal**

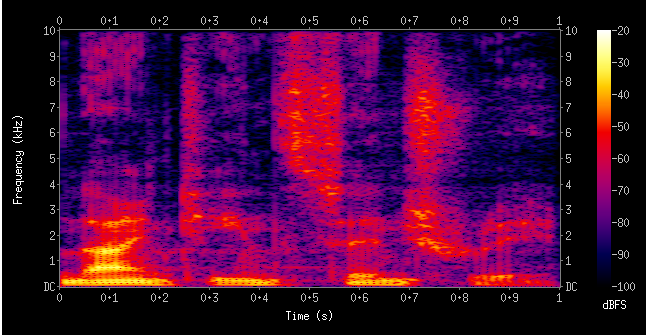


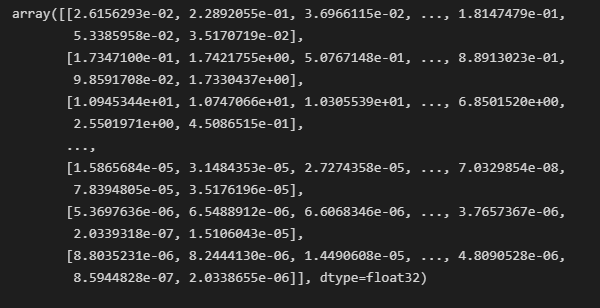
**2 approaches:**

* **Spectrogram :** frequency (spectral) analysis

The spectral analysis of the speech signal is carried out by finding Discrete Fourier Transform (DFT) of the samples in the frame.

**t’s a 2D plot between time and frequency where each point in the plot represents the amplitude of a particular frequency at a particular time in terms of intensity of color**

****

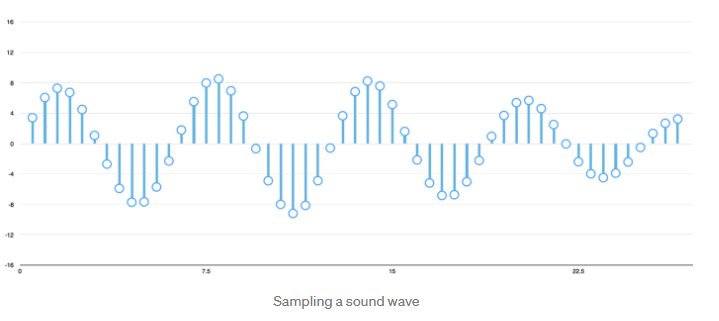


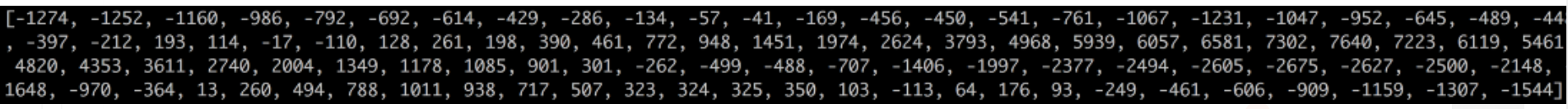
# Turning Sounds into Bits :

# Sound waves are one-dimensional. At every moment in time, they have a single value based on the height of the wave,

# To turn this sound wave into numbers, we just record of the height of the wave at equally-spaced points:

# for speech recognition, a sampling rate of 16khz (16,000 samples per second) is enough to cover the frequency range of human speech.





The big problem is that speech varies in speed: Automatically aligning audio files of various lengths to a fixed-length piece of text

* Preprocessing in addition to a deep neural network

**Procedure:**

words=[“اه”, “لا”, “ممكن”, “عمرى”, “مار مار”]

Dataset : 1 000 audio samples

Each word : 200 audio samples

Data splitting — Train,Cross-Validation and Test

**VGG16 model**

is a convolution neural net (CNN )

