# Deep Learning

Reference: https://www.youtube.com/watch?v=T0r-uCXvDzQ

Deep Learning is a young sub-field within Neural Networks which is a sub-field of Machine Learning. Neural network has only been in existence for 30 years and only in the last 8 years there have been good and viable design patterns constructed and they mostly are in the field of AI. Deep is the concept of the depth of the actual hidden layers that form part of the neural network.

Machine Learning

Neural Networks

Deep Learning

Some of the different models and purpose of deep learning used to resolve simple to complex problems are:

* Classification Models
* Segmentation Models
* Image Capturing Models
* Image Captioning Models
* Generative Models
* Sequence Models

# Deep dive into Shazam

Reference: (https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting-and-recognition)

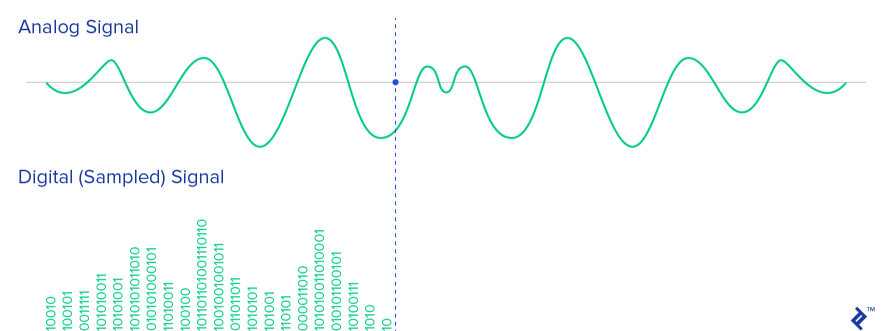
Purpose:

Today, Shazam is the most popular and sophisticated AI tool that identifies a song. Now we all hear songs in the Radio or on the go, but cannot remember always what song it is? Shazam comes handy in identifying a song from a whole range of music collection within a music database.

There are many other programs and libraries in different programming languages written to achieve the analog to digital conversion of sound. Shazam is Java based and almost matches the objective of this dissertation.

Analog to Digital conversion:

As input, Shazam converts the Analog signals to Digital. As Sound is the vibration created by air pressure and displacement through a medium, within the human ears, this vibration is transmitted to hair cells through the eardrum which then produces electrical impulses that are then fed to the brain through our auditory nervous system. Recording a song fairly produces a similar electrical signal. But, the sound wave is a continuous signal and the microphone also captures a recording as a continuous analog voltage signal, which is particularly not useful in the digital world. Therefore, before this can be processed by the computer, this needs to be translated into discrete signal which can be stored digitally. This is achieved by capturing the amplitude (the peaks or troughs at a given point in time).



As audible human frequency is 20 – 20KHz, the sampling Rate is always set to twice the frequency of 44100 Hz. Other settings that can be useful to analyse sound channels (1 for mono), sample size (16 bit).

The conversion of analog to digital involves quantization of the input and therefore introduces a small error. The converters, to overcome this, perform multiple conversion on very small pieces of signal, a process called sampling.

Audio Recording:

The byte array contains the signal recorded in the time domain (amplitude – change over time)