IDC410
A course on Image Processing and Machine Learning
(Lecture 07)

Shashikant Dugad,

IISER Mohali

Reading Material

Suggested Books:

1. Neural Networks and Deep Learning by Michael Nielsen
2. Fundamentals of Deep Learning by Nikhil Buduma

Source for this presentation:

https://www.analyticsvidhya.com/blog/2021/10/perceptron-building-block-of-artificial-neural-network/

https://towardsdatascience.com/the-concept-of-artificial-neurons-perceptrons-in-neural-networks-fab22249cbfc

https://www.scaler.com/topics/deep-learning/introduction-to-feed-forward-neural-network/

https://www.turing.com/kb/mathematical-formulation-of-feed-forward-neural-network

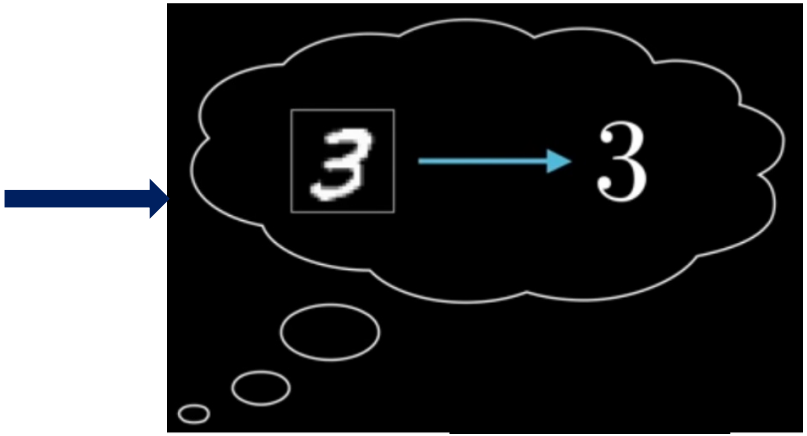
http://machine-learning-for-physicists.org. by Florian Marquardt

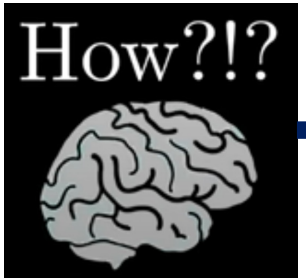
3Blue1Brown (Youtube Videos)

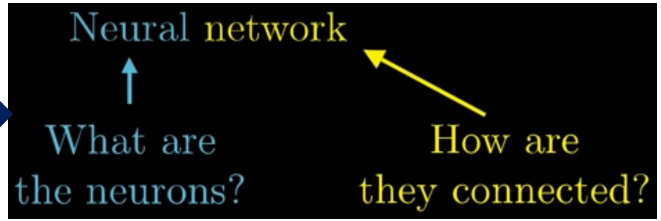
Neurons and Perceprtrons

Pattern Recognition in our brain: How does it happen?

MNSIT Dataset of handwritten numbers: 28 x28 pixels

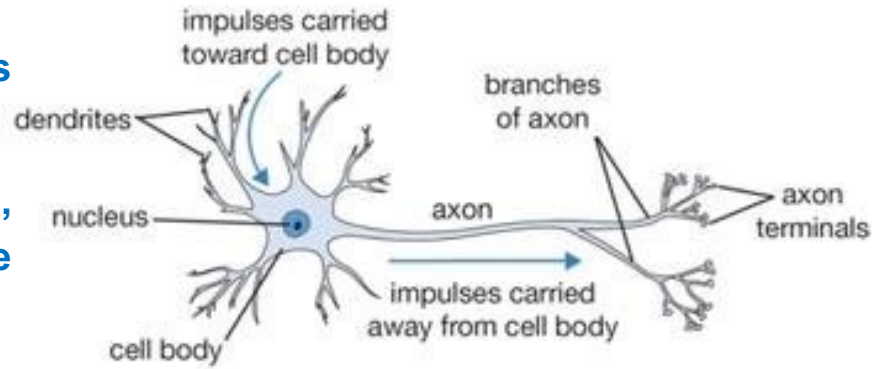






Neurons and Biological Neural Network

- A neuron (nerve cell) is the basic building block of our nervous system.
- A human brain consists of billions of neurons that are interconnected to each other.
- They are responsible for receiving, processing and sending signals from the brain.



- Typical neuron consists of three main parts:
  1. **Dendrites:** They are tree-like branches originating from the cell body. They receive information from the other neurons.
  2. **Soma (Cell Body):** It is the core of a neuron. It is responsible for processing the information received from dendrites.

Neurons and *Biological* Neural Network

- **Typical neuron consists of the three main parts:**
  3. **Axon: It is like a cable through which the neurons send the information.**
     a. **Towards its end, the axon splits up into many branches that make connections with the other neurons through their dendrites.**
     b. **The connection between the axon and other neuron dendrites is called synapses.**
- **The brain consists of a network of billions of neurons!**
- **They communicate by means of electrical and chemical signals through a synapse (junction between two neurons), in which, the information from one neuron is transmitted to other neurons.**
- **The transmission process involves an electrical impulse called 'action potential'.**
- **For the information to be transmitted, the input signals (impulse) should be strong enough to cross a certain threshold barrier, then only a neuron activates and transmits the signal further (output)**

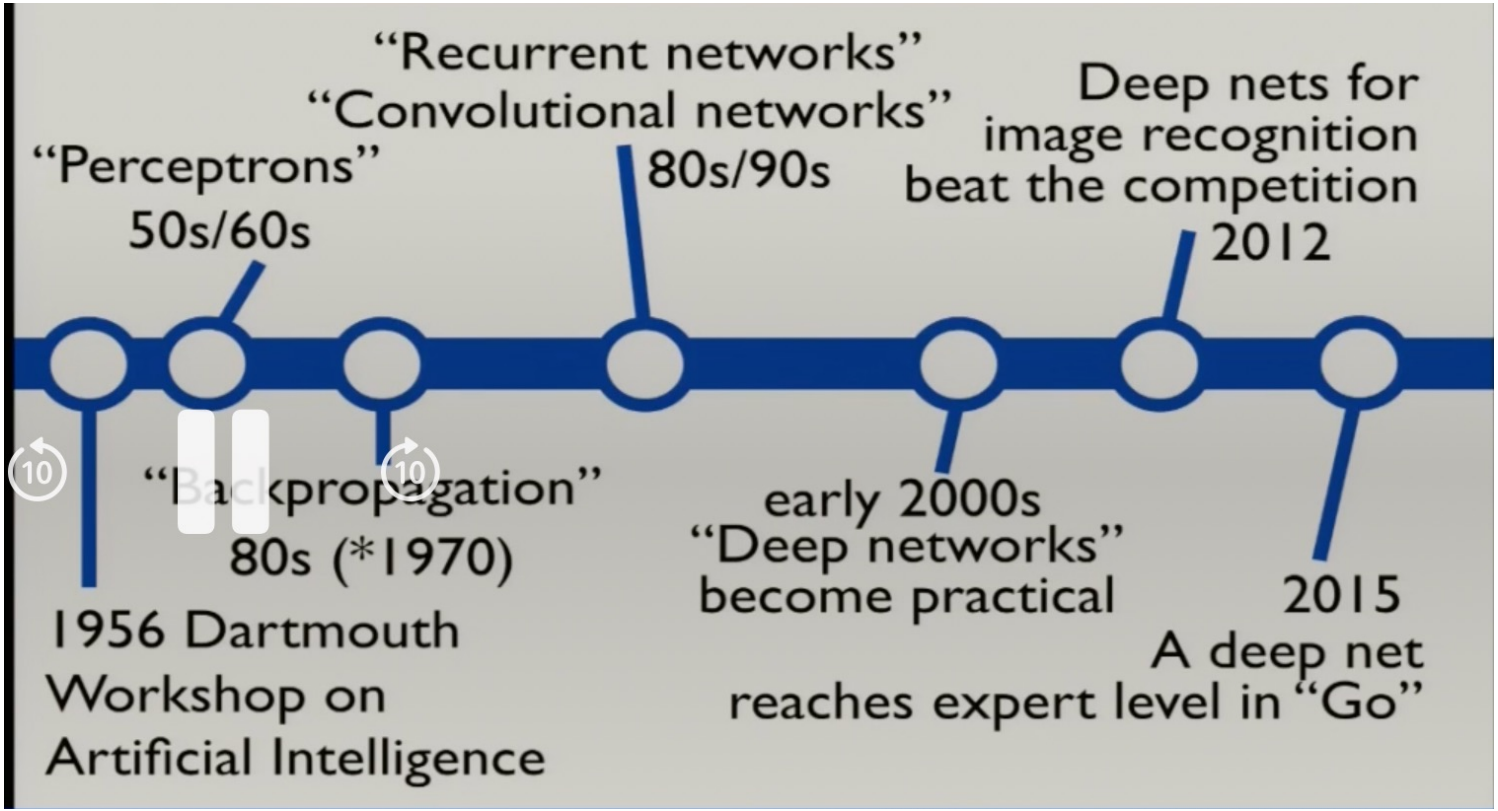
Neurons and *Biological* Neural Network

- **A neuron receives information from other neurons (*through dendrites*) in the form of electrical impulses of varying strength**
- **Neuron processes (integrates) all the electrical impulses it receives from the other neurons and provides one single output on axon**
- **If the resulting summation is larger than a certain threshold value, the neuron 'fires', triggering an action potential that is transmitted to the other connected neurons**

Perceptrons and Artificial Neural Network (ANN)

- **Perceptron is a mathematical model inspired by a biological neurons that are found in the human brain**
  - **Artificial neurons referred as Perceptrons, (Units or Nodes) are the building blocks of a neural network and more advanced neural network structures used in modern machine learning applications**
  - **A concept of perceptron was first introduced by Franck Rosenblatt at Cornell Aeronautical Laboratory in 1957.**
- **Perceptron mimics brain neurons. It processes several inputs with weighted connections and a bias, producing binary outputs by processing through a *decision* making process referred as activation function**
- **Primarily used in binary classification tasks in early stages, perceptrons were instrumental in advancing neural network development and deep learning techniques**

Historical Development of ANN



Perceptrons and Artificial Neural Network (ANN)

- **The Perceptron model in machine learning is characterized by the following key points:**

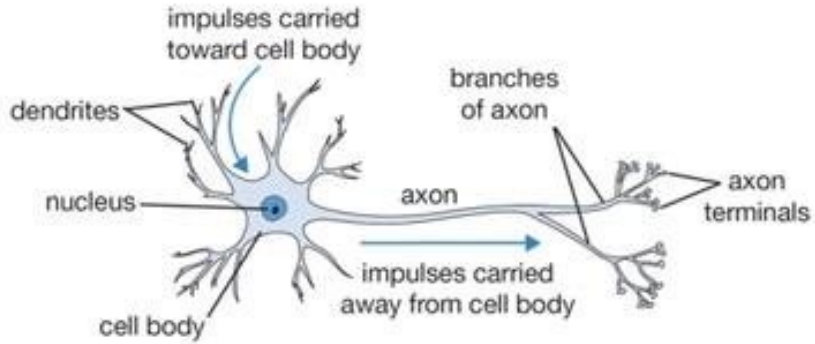
1. **Binary Linear Classifier:** The Perceptron is a type of binary classifier that assigns input data points to one of two possible categories.
2. **Input Processing:** It takes multiple input signals and processes them, each multiplied by a corresponding weight. The inputs are aggregated, and the model produces a single output
3. **Activation Function:** The Perceptron uses an activation function, typically a step function, to determine the output based on the aggregated inputs and weights. If the result exceeds a certain threshold, the output is one category; otherwise, it's the other.
4. **Training Process:** During training, the model adjusts its weights based on the error in its predictions compared to the actual outcomes. This adjustment helps improve the model's accuracy over time.

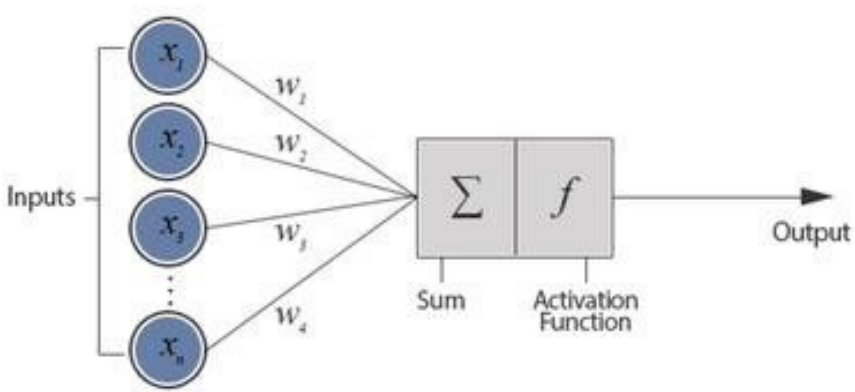
Perceptrons and Artificial Neural Network (ANN)

- **The Perceptron model in machine learning is characterized by the following key points:**
  5. **Single-Layer Model:** **The Perceptron is a single-layer neural network since it has only one layer of output after processing the inputs.**
  6. **Limitations:** **While effective for linearly separable data, the Perceptron has limitations in handling more complex patterns, leading to the development of more sophisticated neural network architectures.**
- **Perceptron consists of one or more input nodes, weights associated with each input node, a summation function to combine the weighted inputs and bias, and finally an activation function to produce a *SINGLE* predicted output**
- **Perceptrons can be stacked together to form more complex architectures, such as single-layer perceptrons or multi-layer perceptrons (MLPs)**

Perceptrons and Artificial Neural Network (ANN)

- **Training a artificial neural network (ANN) containing multi-layer perceptrons (MLPs) involves feeding it with a set of input-output pairs (training data) and adjusting the weights to minimize the difference between the network's predictions and the actual outputs. This process, often referred to as supervised learning**
- **Weights are obtained through techniques like gradient descent to iteratively update the weights based on the error between the predicted and actual outputs.**



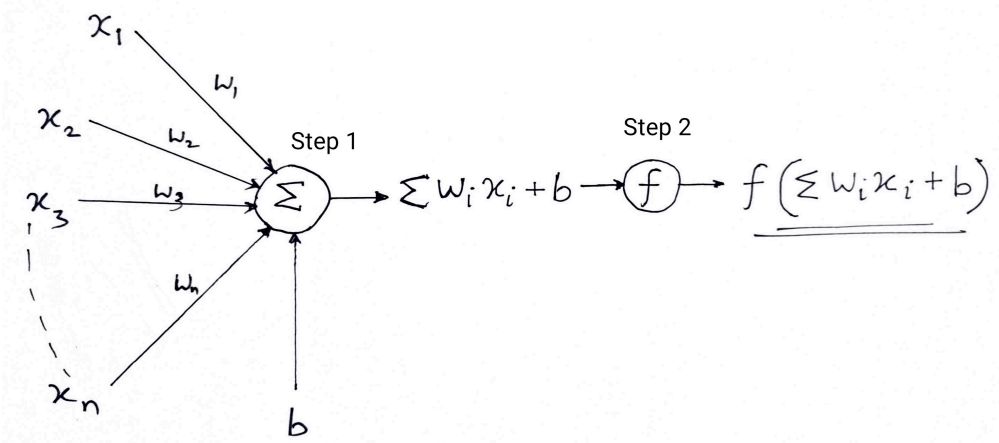


Components of Perceptrons

- **Rosenblatt's initial perceptron model was basically a binary classifier consisting of 3 main parts:**
  1. **Input nodes or input layer:** The input layer takes the initial data into the system for further processing. Each input node is associated with a numerical value. It can take any real value
  2. **Weights and bias:** Weight parameters represent the strength of the connection between units. Higher is the weight, stronger is the influence of the associated input neuron to decide the output. Bias plays the same as the intercept in a linear equation
  3. **Activation function:** The activation function determines whether the neuron will fire or not. At its simplest, the activation function is a step function, but based on the scenario, different activation functions can be used.

Processing a Perceptrons

1. A ***weighted sum*** of all the input values going to a given perceptron is obtained by multiplying with their respective weights and adding them together
2. Additionally, ***a bias term b*** is added to this sum
3. In the second step, an activation function f is applied over the above sum (including bias) to obtain the output Y = f(Sum)
4. Depending upon the scenario and the activation function used, the Output is either binary {1, 0} or a continuous value



Activation Function for a Neuron and Perceptrons

1. A biological neuron only fires when a certain threshold is exceeded
2. Similarly, the artificial neuron will also fire, only when the sum of the inputs (weighted sum) exceeds a certain threshold value, say 0
3. This was the reason for choosing a ***Unit Step (Threshold) function*** as an ***activation function***, originally used by Rosenblatt
4. However, the step function is discontinuous at 0, causing problems in mathematical computations
5. A smoother version of the above function such as the ***sigmoid*** function, the ***Hyperbolic tangent***(tan*h*) function are preferred as they produce the output between -1 and 1.
6. Both ***sigmoid*** and ***tanh*** functions suffer from vanishing gradients problems
7. ***ReLU*** and ***Leaky ReLU*** are the most popularly used activation functions. They are comparatively stable over deep networks.

$$f(x) = \begin{cases} 0 \text{ if } 0 > x \\ 1 \text{ if } x \geq 0 \end{cases}$$