CS 451 - Computational Intelligence

Unit # 9-1

Acknowledgement

- The slides of this lecture have been taken from the lecture slides of CS307 – “Introduction to Artificial Intelligence” and CSE652 – “Knowledge Discovery and Data mining” by Dr. Sajjad Haider.

Artificial Neural Networks

- Neural networks are biologically motivated computing structures that are conceptually modeled after the brain.

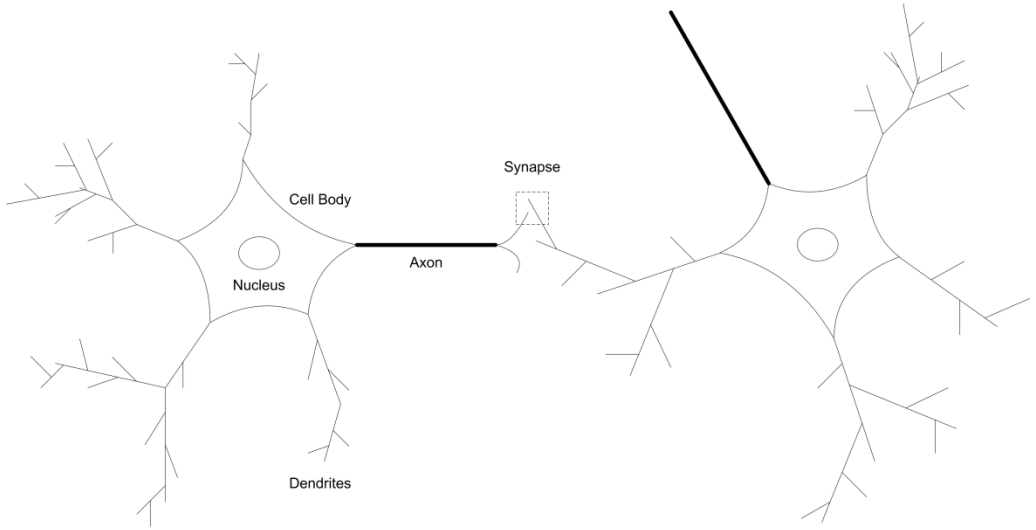
How Our Brain Works?

- While neural networks are modeled after our understanding of the way in which our brain works, surprisingly little is known about how our brains actually function.
- Through various types of inspection, we can see our brain in operation, but because of the massive number of neurons and interconnections between these neurons, how it works remains a mystery (though many theories exist).

Neural Network

- The neural network is made up of a highly connected network of individual computing elements (mimicking neurons) that collectively can be used to solve interesting and difficult problems.

Neurons Inside Our Body



- We are born with about 100 billion neurons
- A neuron may connect to as many as 100,000 other neurons
- Signals “move” via electrochemical signals
- The synapses release a chemical transmitter – the sum of which can cause a threshold to be reached – causing the neuron to “fire”

Definition of Neurons (from Wikipedia)

- **Neurons** are responsive cells in the nervous system that process and transmit information by chemical signals within the neuron.
- A number of different types of neurons exist: sensory neurons respond to touch, sound, light and numerous other stimuli affecting cells of the sensory organs that then send signals to the spinal cord and brain.
- Motor neurons receive signals from the brain and spinal cord and cause muscle contractions and affect glands.
- Inter-neurons connect neurons to other neurons within the brain and spinal cord.
- **Neurons respond to stimulus and communicate the presence of that stimuli to the central nervous system, which processes that information and sends responses to other parts of the body for action.**

Birth of Artificial Neural Networks

- The story of neural networks is interesting because, like AI itself, it's one of grand visions, eventual disappointment, and finally, silent adoption.
- In 1943, McCulloch and Pitts developed a neural network model based on their understanding of neurology, but the models were typically limited to formal logic simulations (simulating binary operations).

Dark Period

- In 1969, the growing popularity of neural networks was brought to a halt.
- Marvin Minsky and Seymour Papert wrote a book entitled “Perceptrons” in which limitations of single-layer perceptrons were discussed.
- The result was severe reductions in neural network research funding, and a corresponding reduction in the effort applied to the field.

Revival

- In 1974, Paul Werbos developed the **backpropagation** algorithm, which permitted successful learning in multilayer neural networks.
- Since the 1970s, research and successful results in neural network design have attracted scientists back to the field.

Artificial Neural Network

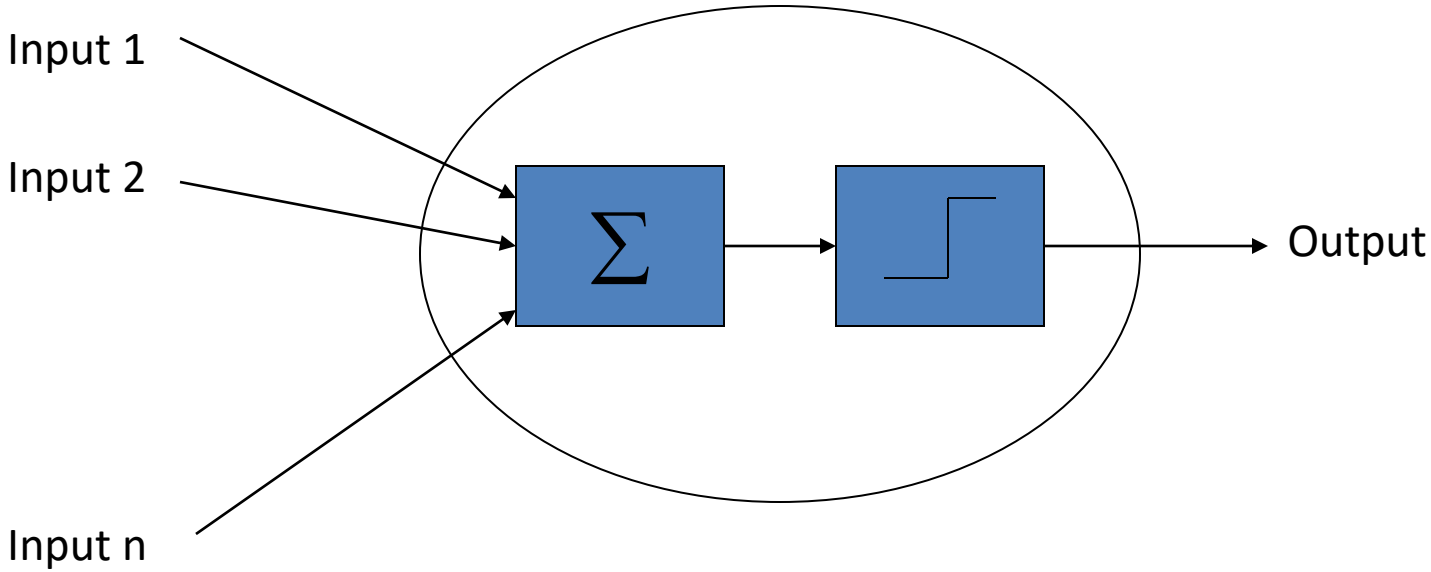
Artificial Neural Network

- The neural network is made up of a highly connected network of individual computing elements (mimicking neurons) that collectively can be used to solve interesting and difficult problems.

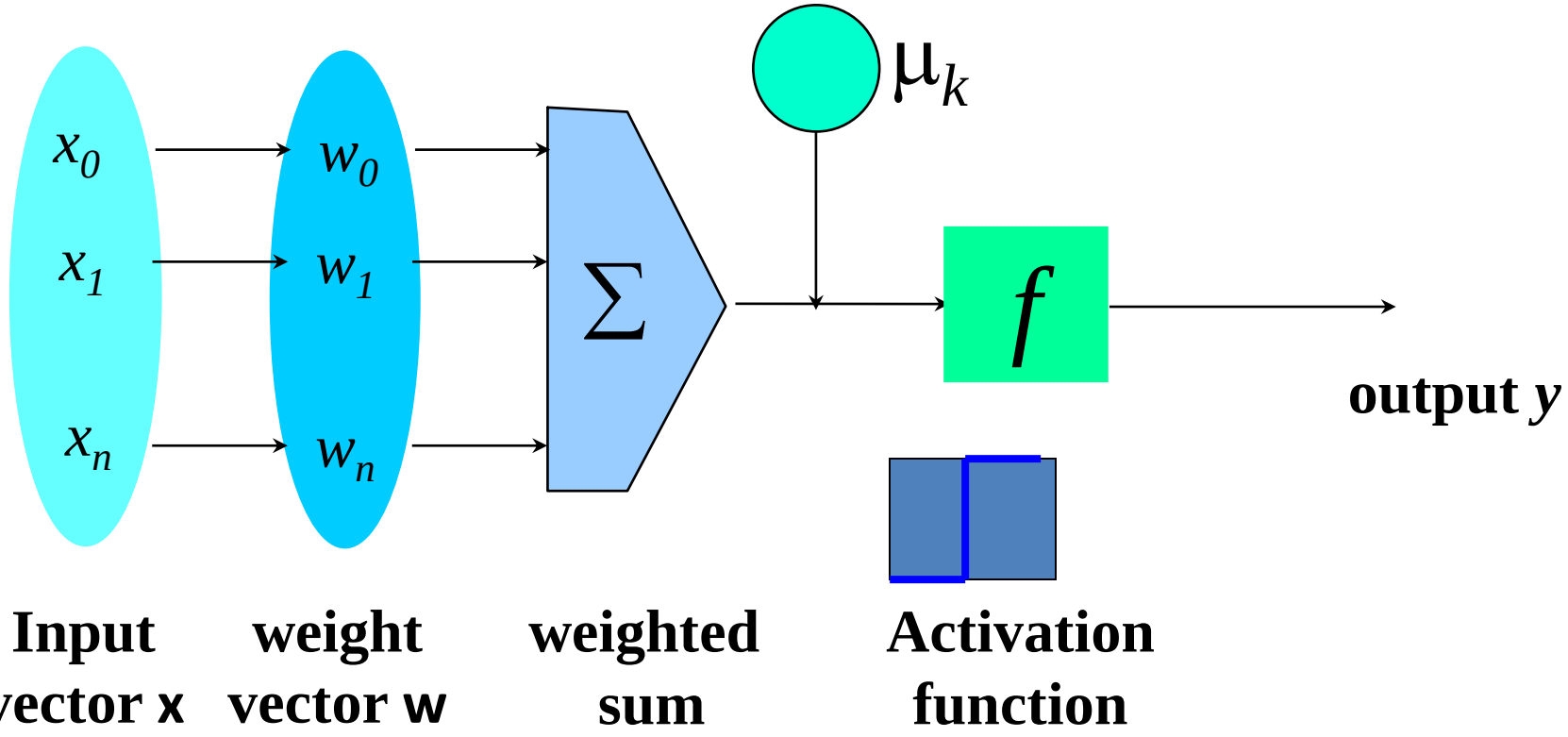
Artificial Neuron

- An artificial neuron is an information-processing unit that is fundamental to the operation of an ANN.
- It consists of three basic elements:
  - A set of input links
  - An Output
  - An adder for summing the input signals weighted by the respective synaptic strengths.
  - An activation function for limiting the amplitude of the output of a neuron.

McCulloch-Pitts Neuron



A Neuron (= a perceptron)

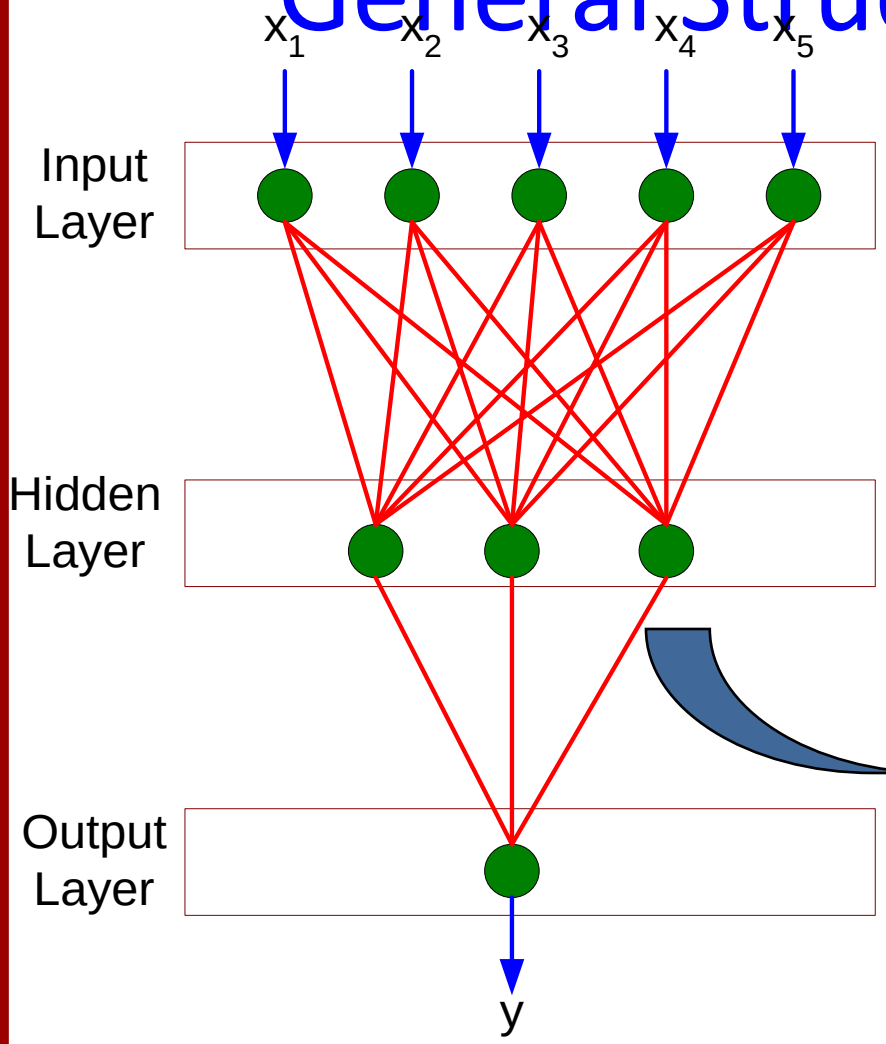


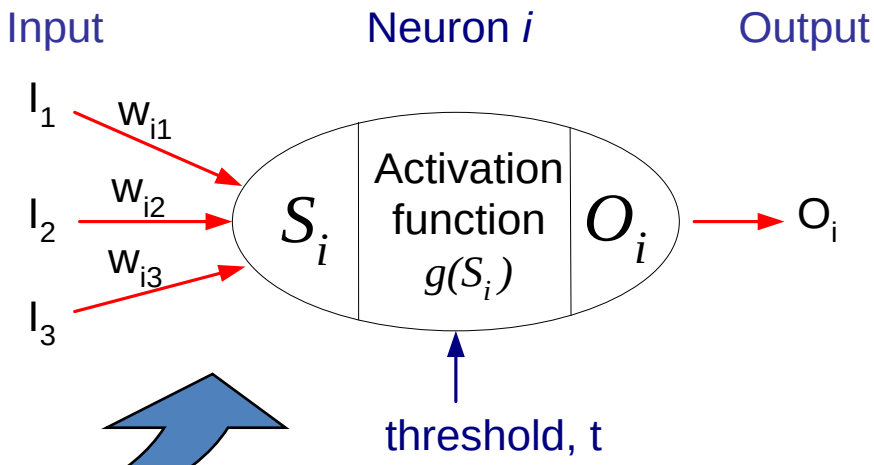
- The *n*-dimensional input vector **x** is mapped into variable y by means of the scalar product and a nonlinear function mapping

Multilayer Feed-Forward Networks

- Multilayer feed-forward networks are one of the most important and most popular classes of ANNs in real-world applications.
- They are commonly referred to as multilayer perceptrons.

General Structure of ANN





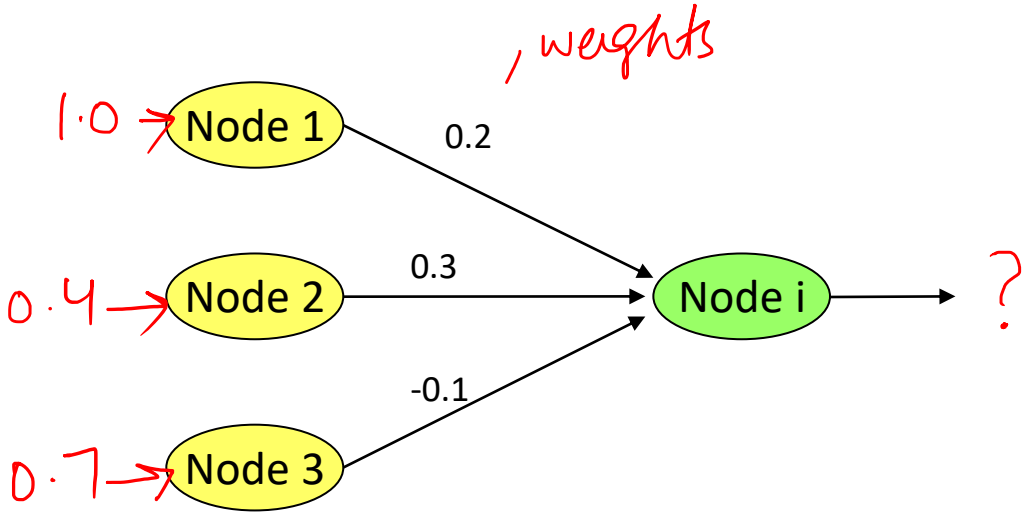
Training ANN means learning the weights of the neurons

Characteristics of MLP

- A multilayer perceptron has three distinctive characteristics
  - The network contains one or more layer of hidden neurons that are not a part of the input or output of the network. These hidden nodes enable the network to learn complex and highly nonlinear tasks by extracting progressively more meaningful features from the input patterns.
  - The model of each neuron in the network includes usually a nonlinear activation function, sigmoid or hyperbolic.
  - The network exhibits a high degree of connectivity from one layer to the next one.

Feed-forwarding Data

- Input = {1.0, 0.4, 0.7}
- Output of Node I = ??



Input (Node i) = $(1 \times 0.2) + (0.4 \times 0.3) + (0.7 \times -0.1)$

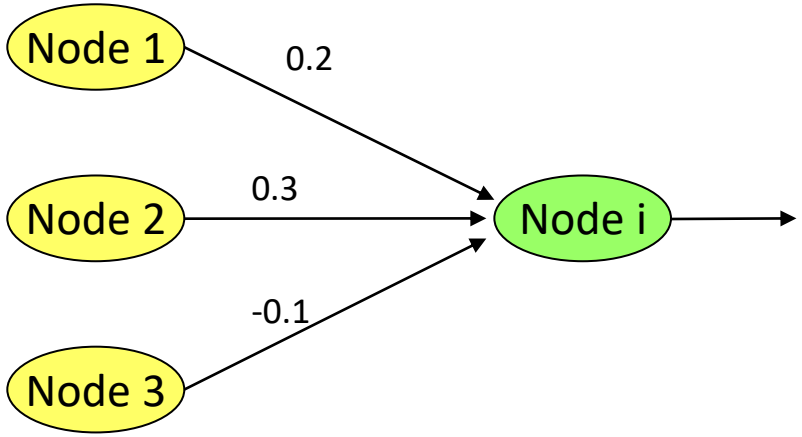
$= 0.25$

Sigmoid $\frac{1}{1+e^{-x}}$ $f(0.25) = \frac{1}{1+e^{-0.25}}$

Output of Node i = 0.56

Feed-forwarding Data

- Input = {1.0, 0.4, 0.7}

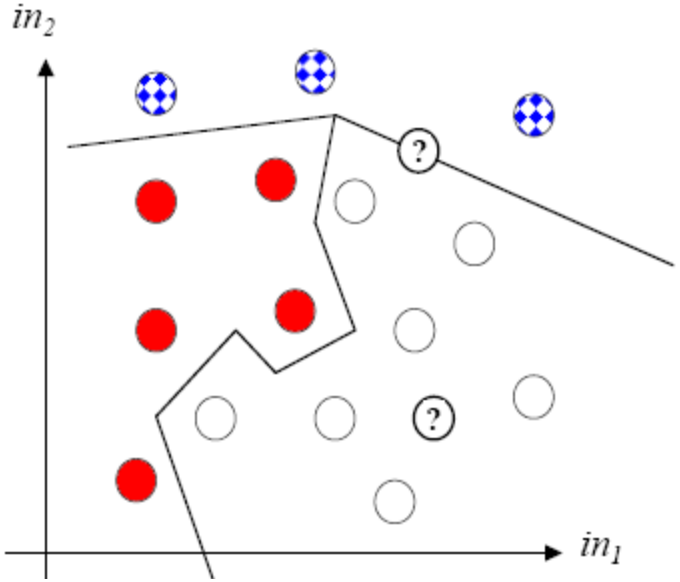


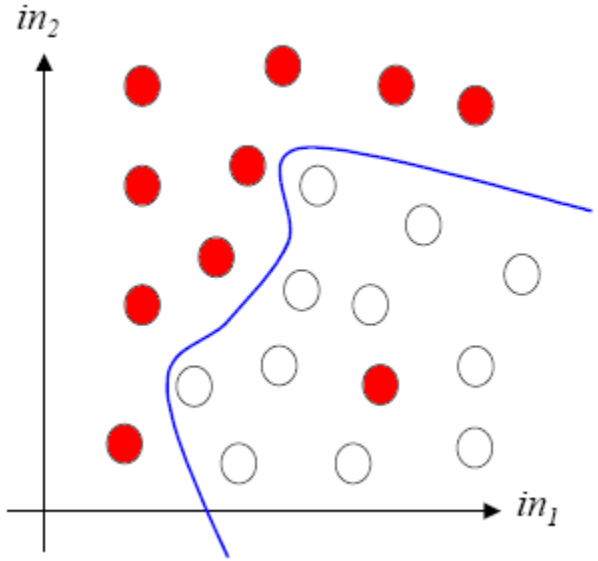
- Input to node i = 0.2x1.0 + 0.3x0.4 - 0.1x0.7 = 0.25
- Now apply the sigmoid function: f(0.25) = 0.562
- The output of Node I is 0.562

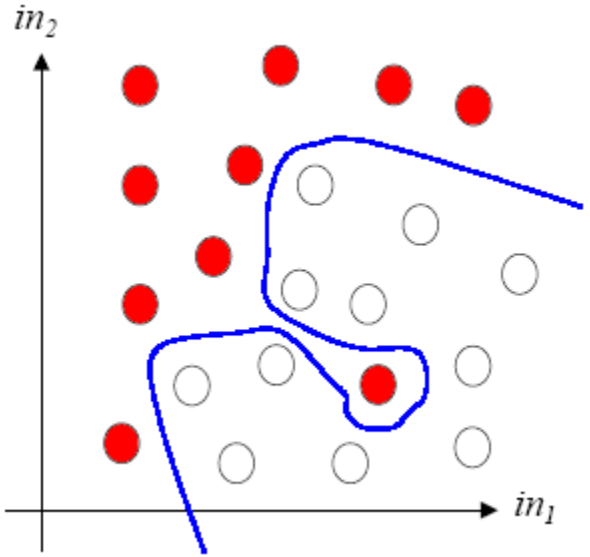
How A Multi-Layer Neural Network Works?

- The **inputs** to the network correspond to the attributes measured for each training tuple.
- Inputs are fed simultaneously into the units making up the **input layer.**
- They are then weighted and fed simultaneously to a **hidden layer.** The number of hidden layers is arbitrary, although usually only one
- The weighted outputs of the last hidden layer are input to units making up the **output layer**, which emits the network's prediction
- From a statistical point of view, networks perform **nonlinear regression**: Given enough hidden units and enough training samples, they can closely approximate any function.

Decision Boundaries

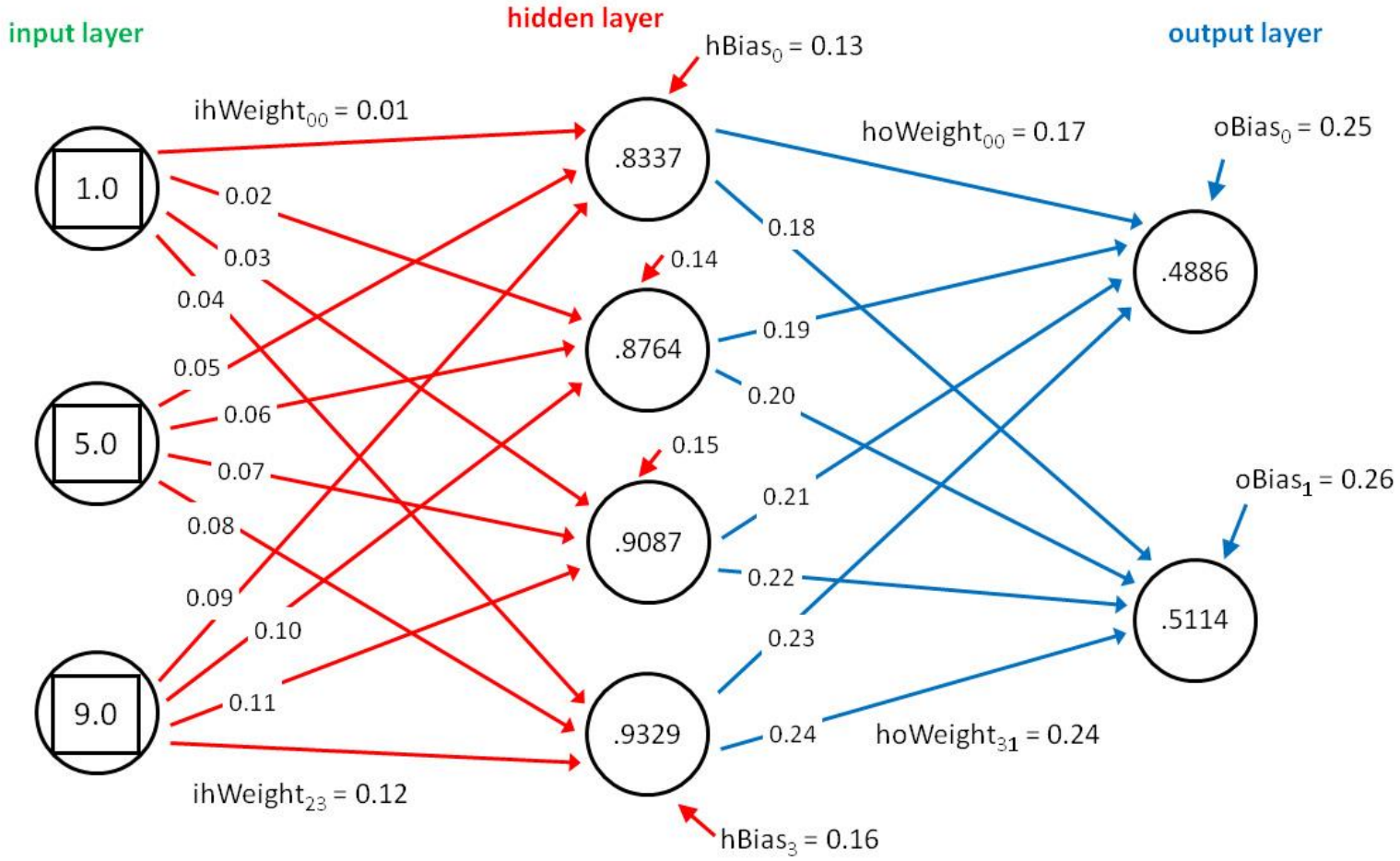






Learning a Neural Network

Learning a Neural Network



https://visualstudiomagazine.com/articles/2014/11/01/use-python-with-your-neural-networks.aspx

Learning a Neural Network

- Topology of Neural Network
- Activation Function
- Learning – Backpropagation

Deciding the Structure of a Neural Network

Specification of ANN

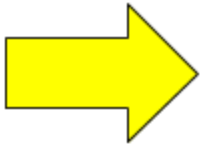
- The number of input attributes found within individual instances determines the number of input layer nodes.
- The user specifies the number of hidden layers as well as the number of nodes within a specific hidden layer.

Input Format

- The input to individual neural network nodes should be numeric and fall in the closed interval range [0,1].
- We need a way to numerically represent categorical data.
  - Attribute Color: {Red, Green, Blue, Yellow}
- We also need a conversion method for numerical data falling outside the [0,1] range.
  - Values: 100, 200, 300, 400

One-hot Encoding

Color
Red
Red
Yellow
Green
Yellow



Red	Yellow	Green
1	0	0
1	0	0
0	1	0
0	0	1

Architecture of NN?

- How many neurons are required in the input layer?

Name	Give Birth	Can Fly	Live in Water	Have Legs	Class
human	yes	no	no	yes	mammals
python	no	no	no	no	non-mammals
salmon	no	no	yes	no	non-mammals
whale	yes	no	yes	no	mammals
frog	no	no	sometimes	yes	non-mammals
komodo	no	no	no	yes	non-mammals
bat	yes	yes	no	yes	mammals
pigeon	no	yes	no	yes	non-mammals
cat	yes	no	no	yes	mammals
leopard shark	yes	no	yes	no	non-mammals
turtle	no	no	sometimes	yes	non-mammals
penguin	no	no	sometimes	yes	non-mammals
porcupine	yes	no	no	yes	mammals
eel	no	no	yes	no	non-mammals
salamander	no	no	sometimes	yes	non-mammals
gila monster	no	no	no	yes	non-mammals
platypus	no	no	no	yes	mammals
owl	no	yes	no	yes	non-mammals
dolphin	yes	no	yes	no	mammals
eagle	no	yes	no	yes	non-mammals

Architecture of NN?

- How many neurons are required in the input layer?

Outlook	Temperature	Humidity	Windy	Class
sunny	hot	high	false	N
sunny	hot	high	true	N
overcast	hot	high	false	P
rain	mild	high	false	P
rain	cool	normal	false	P
rain	cool	normal	true	N
overcast	cool	normal	true	P
sunny	mild	high	false	N
sunny	cool	normal	false	P
rain	mild	normal	false	P
sunny	mild	normal	true	P
overcast	mild	high	true	P
overcast	hot	normal	false	P
rain	mild	high	true	N

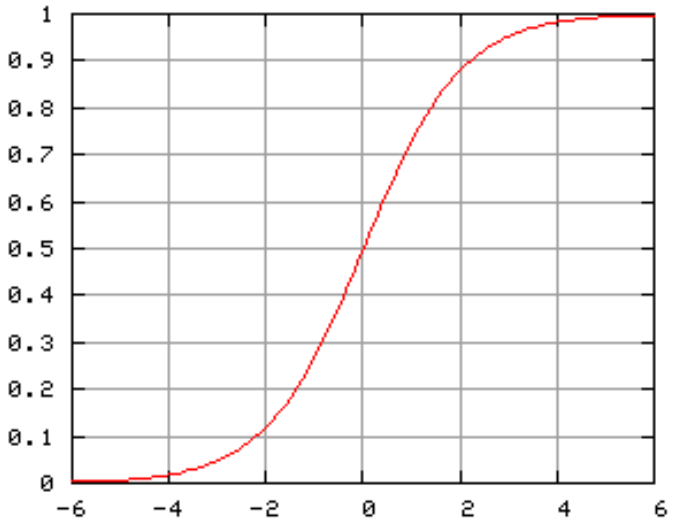
Output Format

- The nodes of the input layer pass input attribute values to the hidden layer unchanged.
- A hidden or output layer node takes input from the connected nodes of the previous layer, combines the previous layer node values into a single value, and uses the new value as input to an evaluation function.
- The output of the evaluation function is a number in the closed interval [0, 1].

Activation Function

Sigmoid Function

- The first criterion of an evaluation function is that the function must output values in the [0, 1] interval range.
- A second criterion is that the function should output a value close to 1 when sufficiently excited.
- The sigmoid function meets both criterion and is often used for node evaluation.
  - $f(x) = 1 / (1 + e^{-x})$

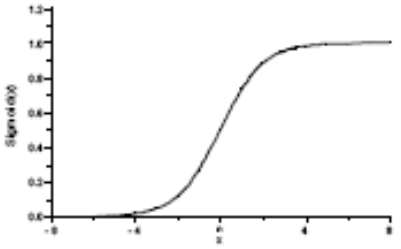


Transfer Functions

Sigmoid Functions These are smooth (differentiable) and monotonically increasing.

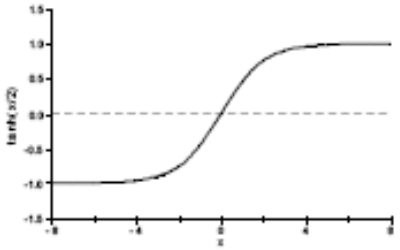
The logistic function

$$\text{Sigmoid}(x) = \frac{1}{1+e^{-x}}$$



Hyperbolic tangent

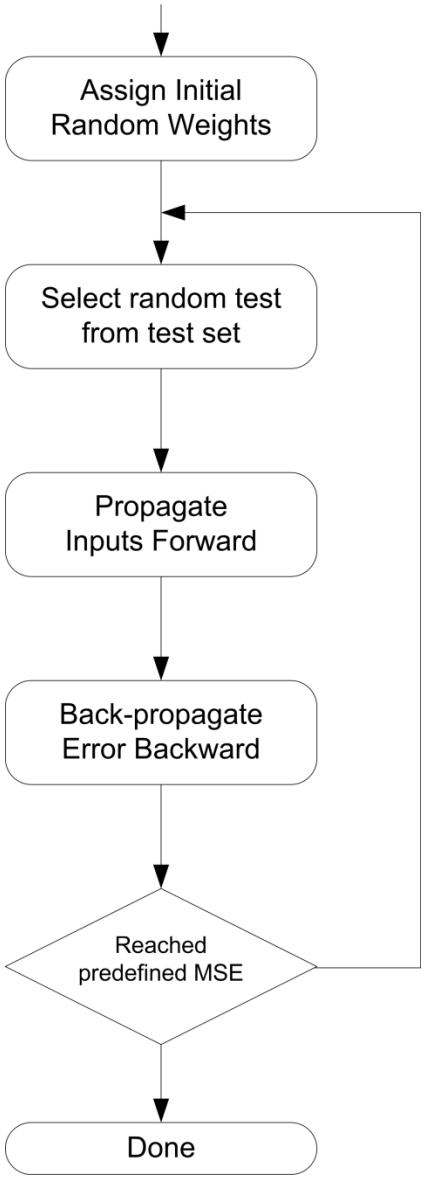
$$\tanh\left(\frac{x}{2}\right) = \frac{1-e^{-x}}{1+e^{-x}}$$



Learning a Neural Network

Working of ANN

- Learning is accomplished by modifying network connection weights while a set of input instances is repeatedly passed through the network.
- Once trained, an unknown instance passing through the network is classified according to the value(s) seen at the output layer.



Resources

- https://www.coursera.org/learn/machine-learning/lecture/du981/backpropagation-intuition
- https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/
- https://scikit-neuralnetwork.readthedocs.io/en/latest/index.html