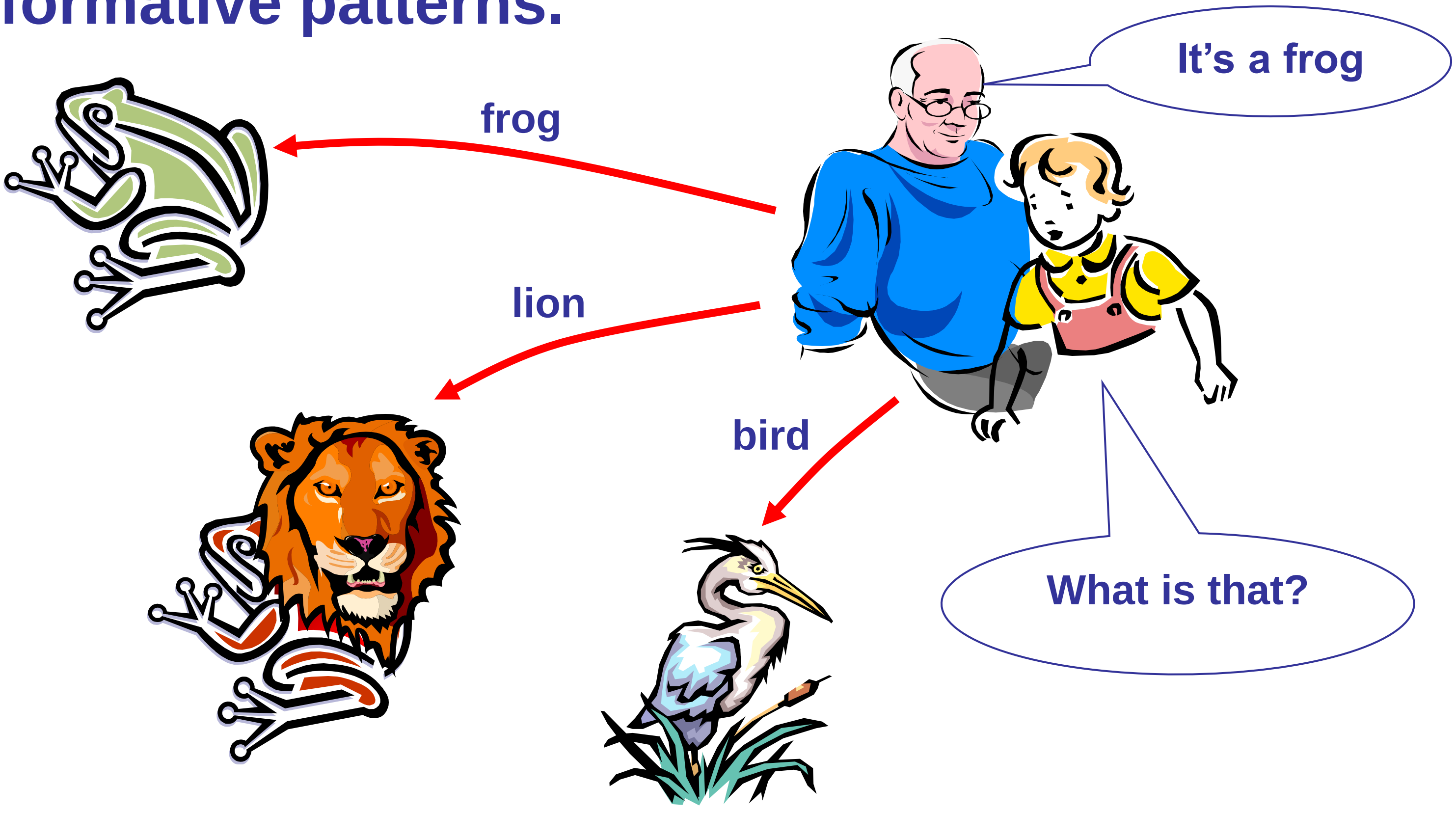
INTRODUCTION TO ARTIFICIAL NEURAL NETWORKS (ANN)

Outline

- Definition, why and how are neural networks being used in solving problems
- Human biological neuron
- Artificial Neuron
- Applications of ANN
- Comparison of ANN vs conventional AI methods

The idea of ANNs..?

- **NNs learn relationship between cause and effect or organize large volumes of data into orderly and informative patterns.**



Neural networks to the rescue…

- **Neural network:** *information processing paradigm inspired by biological nervous systems, such as our brain*
- Structure: large number of highly interconnected processing elements (*neurons*) working together
- Like people, they learn *from experience* (by example)

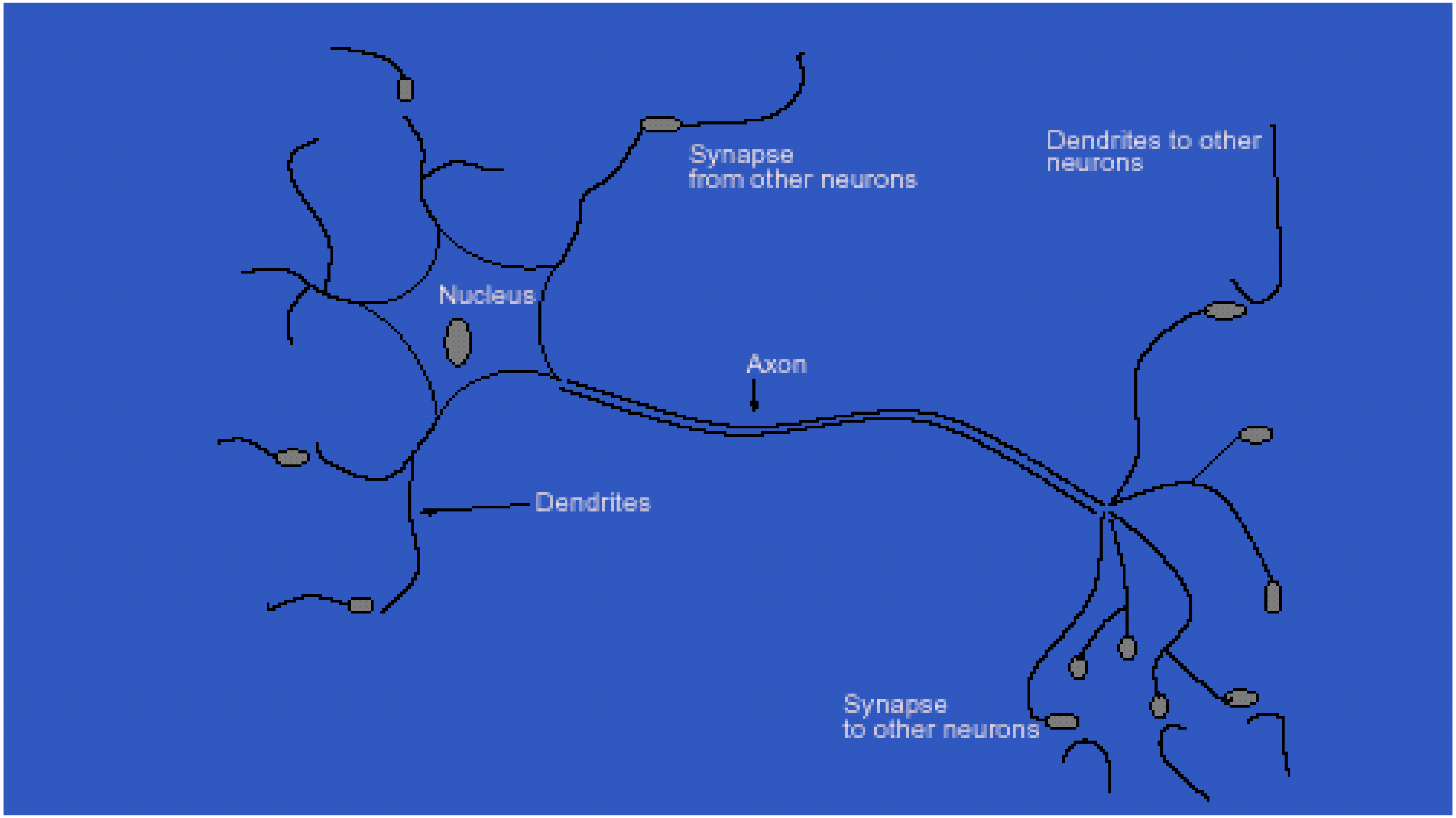
Definition of ANN

"Data processing system consisting of a large number of simple, highly interconnected processing elements (artificial neurons) in an architecture inspired by the structure of the cerebral cortex of the brain"

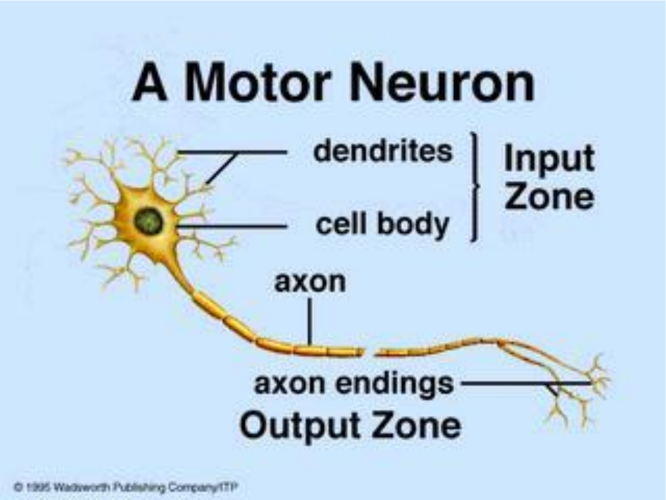
(Tsoukalas & Uhrig, 1997).

Inspiration from Neurobiology

Human Biological Neuron



Biological Neural Networks



- **A biological neuron has three types of main components; <u>dendrites</u>, <u>soma</u> (or cell body) and <u>axon</u>.**
- **Dendrites receives signals from other neurons.**
- **Each input signal flowing through dendrite passes through a synapse or synaptic junction. The junction is an infinitesimal gap in the dendrite, which is filled with neurotransmitter fluid that either accelerates or retards the flow of the signal. The adjustment of the impedance or conductance of the synaptic gap by the neurotransmitter fluid contributes to the memory or intelligence of the brain.**
- **The soma, sums the incoming signals. When sufficient input is received, the cell fires; that is it transmit a signal over its axon to other cells.**

Comparison of Brains and Traditional Computers

- 200 billion neurons, 32 trillion synapses
- Element size: 10^{-6} m
- Energy use: 25W
- Processing speed: 100 Hz
- Parallel, Distributed
- Fault Tolerant
- Learns: Yes
- Intelligent/Conscious: Usually

- 1 billion bytes RAM but trillions of bytes on disk
- Element size: 10^{-9} m
- Energy watt: 30-90W (CPU)
- Processing speed: 10^{9} Hz
- Serial, Centralized
- Generally not Fault Tolerant
- Learns: Some
- Intelligent/Conscious: Generally No

Artificial Neurons

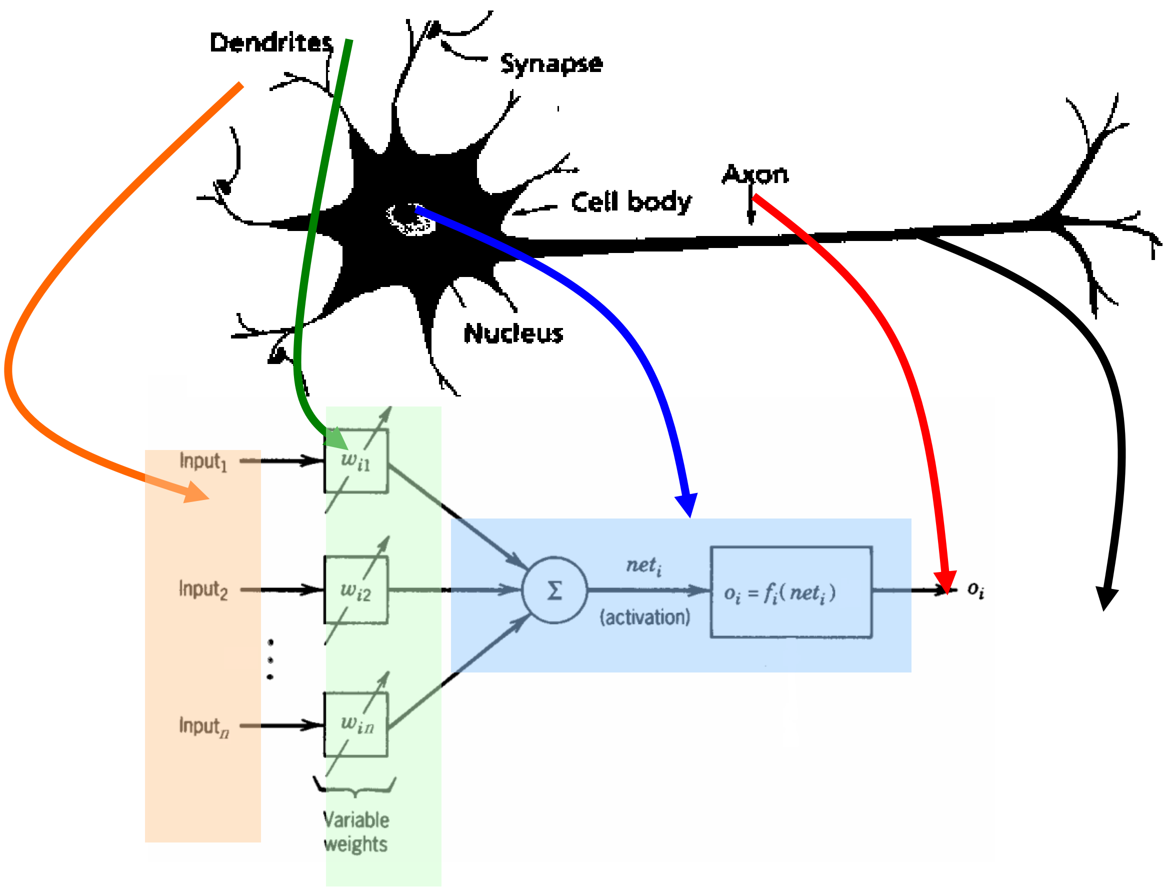
- ANN is an information processing system that has certain performance characteristics in common with biological nets.
- Several key features of the processing elements of ANN are suggested by the properties of biological neurons:

1. The processing element receives many signals.
2. Signals may be modified by a weight at the receiving synapse.
3. The processing element sums the weighted inputs.
4. Under appropriate circumstances (sufficient input), the neuron transmits a single output.
5. The output from a particular neuron may go to many other neurons.

Artificial Neurons

- From experience: examples / training data
- Strength of connection between the neurons is stored as a weight-value for the specific connection.
- Learning the solution to a problem = changing the connection weights

A physical neuron



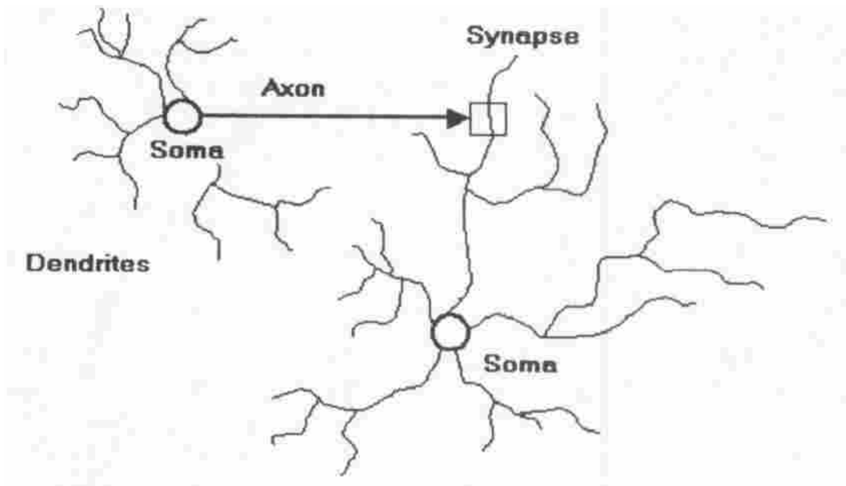
An artificial neuron

Artificial Neurons

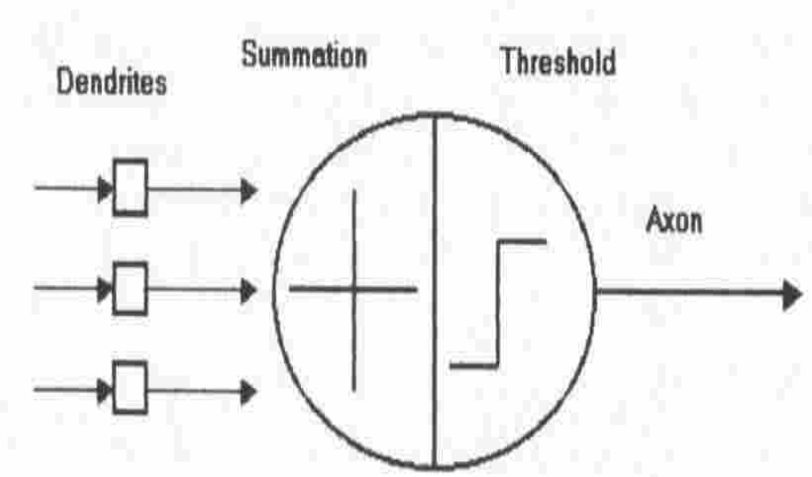
- **ANNs have been developed as generalizations of mathematical models of neural biology, based on the assumptions that:**

1. **Information processing occurs at many simple elements called neurons.**
2. **Signals are passed between neurons over connection links.**
3. **Each connection link has an associated weight, which, in typical neural net, multiplies the signal transmitted.**
4. **Each neuron applies an activation function to its net input to determine its output signal.**

Artificial Neuron

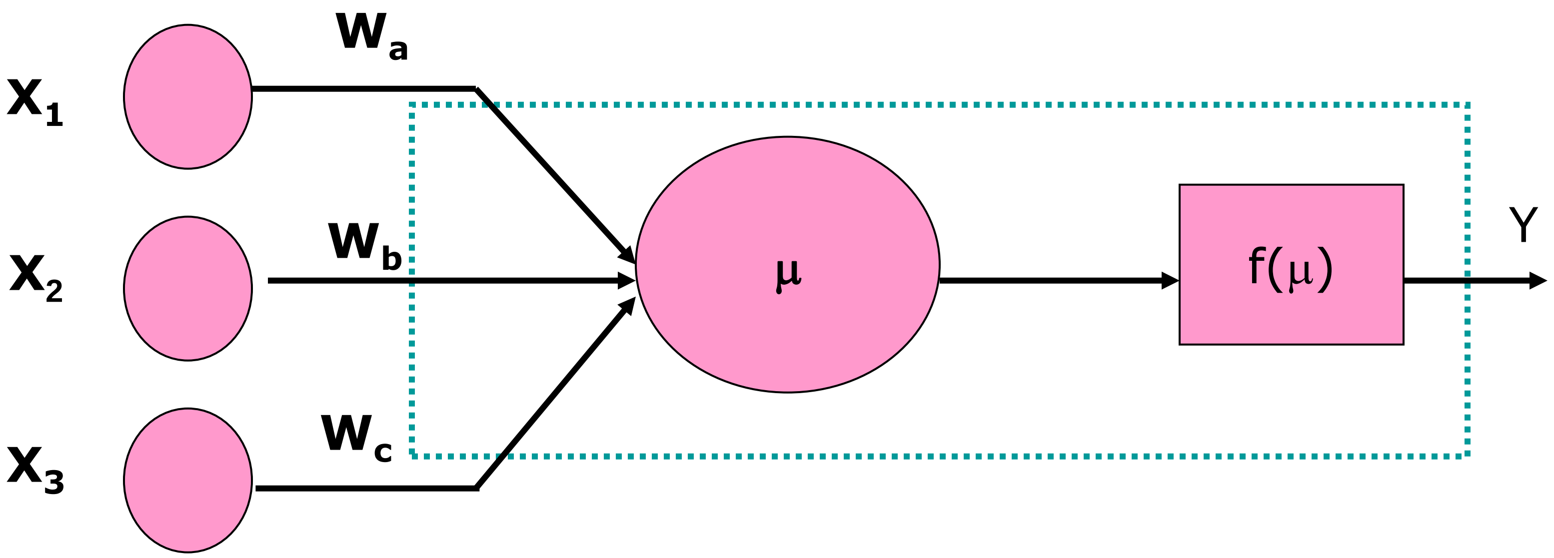


Four basic components of a human biological neuron



The components of a basic artificial neuron

Model Of A Neuron



Input units (dendrite)

Connection weights (synapse)

Summing function

computation

(soma)

(axon)

- A neural net consists of a large number of simple processing elements called neurons, units, cells or nodes.

- Each neuron is connected to other neurons by means of directed communication links, each with associated weight.

- The weight represent information being used by the net to solve a problem.

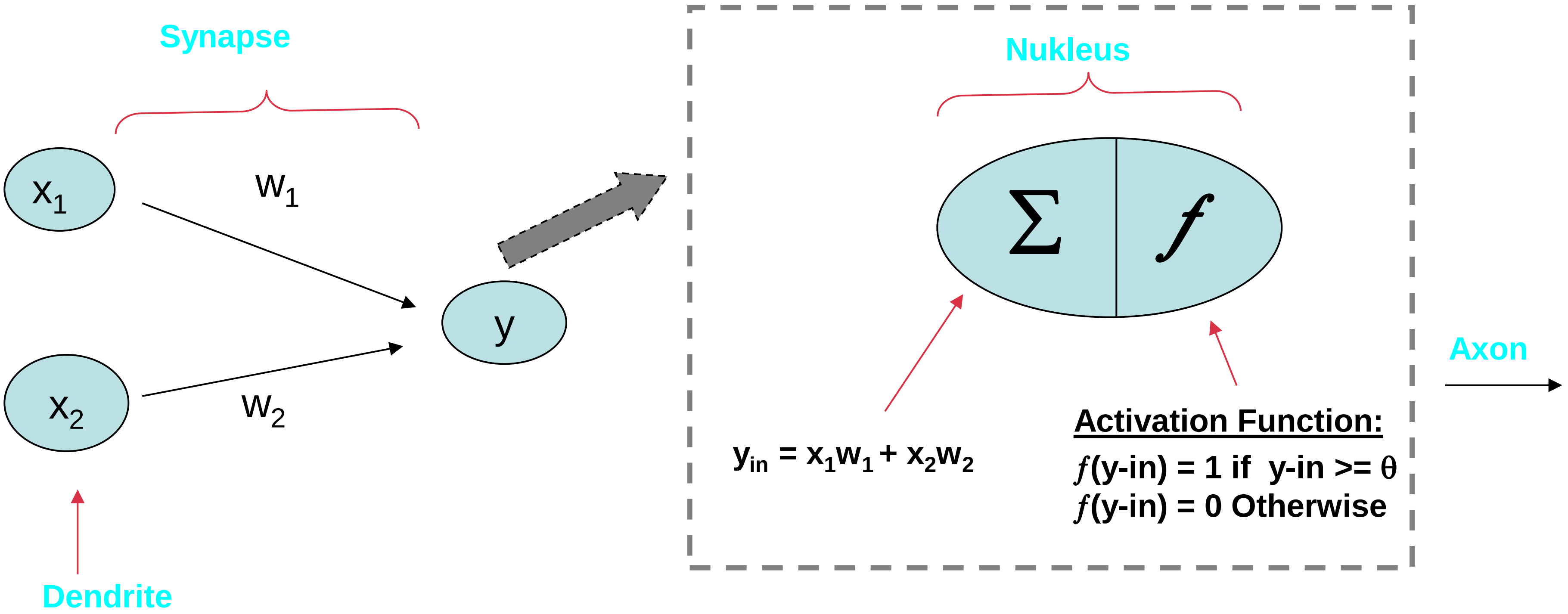
- Each neuron has an internal state, called its activation or activity level, which is a function of the inputs it has received. Typically, a neuron sends its activation as a signal to several other neurons.

- It is important to note that a neuron can send only one signal at a time, although that signal is broadcast to several other neurons.

- Neural networks are configured for a specific application, such as pattern recognition or data classification, through a learning process
- In a biological system, learning involves adjustments to the synaptic connections between neurons

➔ same for artificial neural networks (ANNs)

Artificial Neural Network (Thresh Hold Function)



-A neuron receives input, determines the strength or the weight of the input, calculates the total weighted input, and compares the total weighted with a value (threshold)

-The value is in the range of 0 and 1

- If the total weighted input greater than or equal the threshold value, the neuron will produce the output, and if the total weighted input less than the threshold value, no output will be produced

Some Other Activation Functions- Signum/Quantizer, Logistic/ Squashing (Page-194)

History

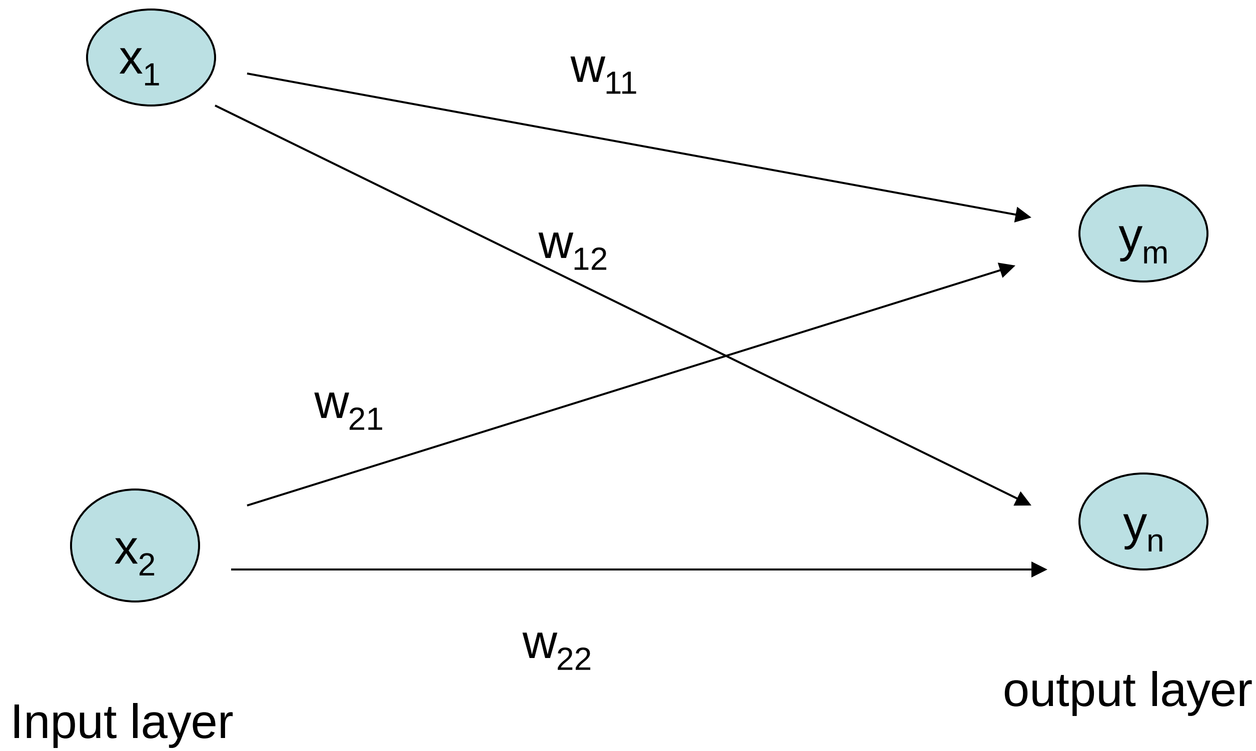
- 1943 McCulloch-Pitts neurons
- 1949 Hebb’s law
- 1958 Perceptron (Rosenblatt)
- 1960 Adaline, better learning rule (Widrow, Huff)
- 1969 Limitations (Minsky, Papert)
- 1972 Kohonen nets, associative memory

- 1977 Brain State in a Box (Anderson)
- 1982 Hopfield net, constraint satisfaction
- 1985 ART (Carpenter, Grossfield)
- 1986 Backpropagation (Rumelhart, Hinton, McClelland)
- 1988 Neocognitron, character recognition (Fukushima)

Characterization

- Architecture
  - a pattern of connections between neurons
    - Single Layer Feedforward
    - Multilayer Feedforward
    - Recurrent
- Strategy / Learning Algorithm
  - a method of determining the connection weights
    - Supervised
    - Unsupervised
    - Reinforcement
- Activation Function
  - Function to compute output signal from input signal

Single Layer Feedforward NN



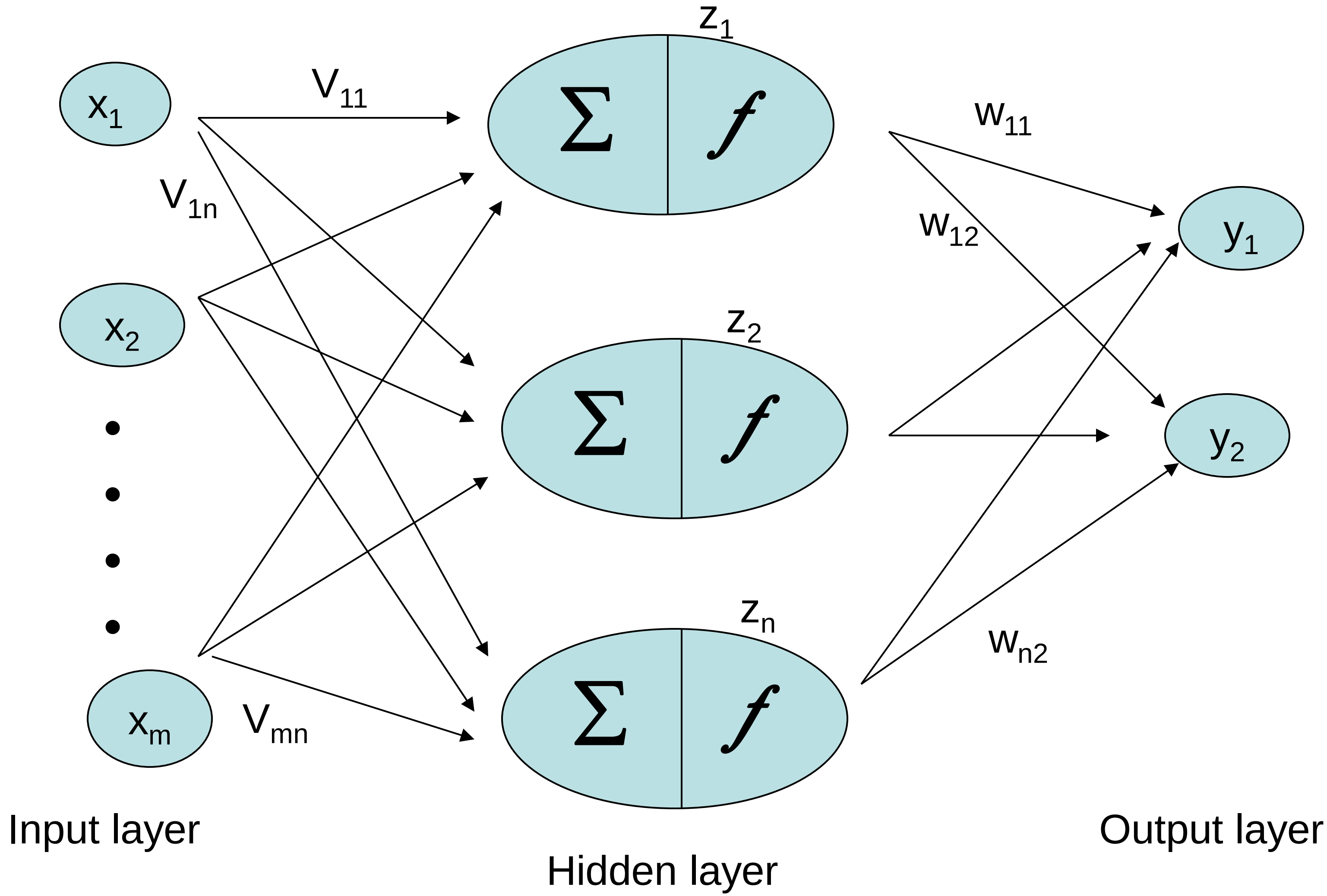
Single Layer Perceptron (SLP) drawbacks & solutions

Main Drawback:
It can only solve problems that are linearly separable. It fails on simple non-linear problems like the XOR gate.
complex problems can't solve

Primary Solution:
Use a Multi-Layer Perceptron (MLP)
with hidden layers and non-linear activation functions (e.g., Sigmoid, ReLU).
This allows the network to learn complex, non-linear decision boundaries. The Backpropagation algorithm is used to train these networks.

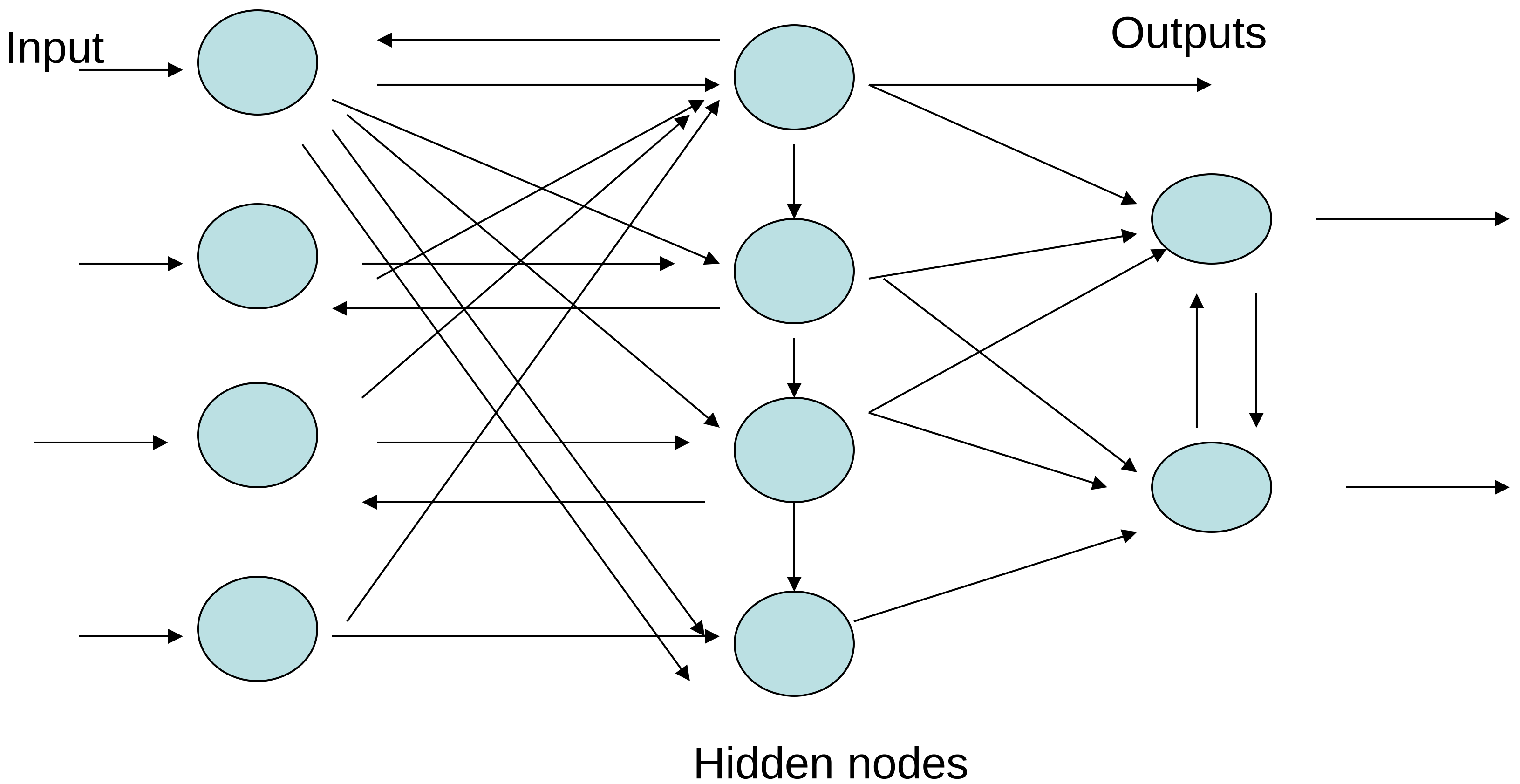
Example: **ADALINE, Hopfield, Perceptron etc.**

Multilayer Neural Network



Example: **MADALINE, Back Propagation etc.**

Recurrent NN



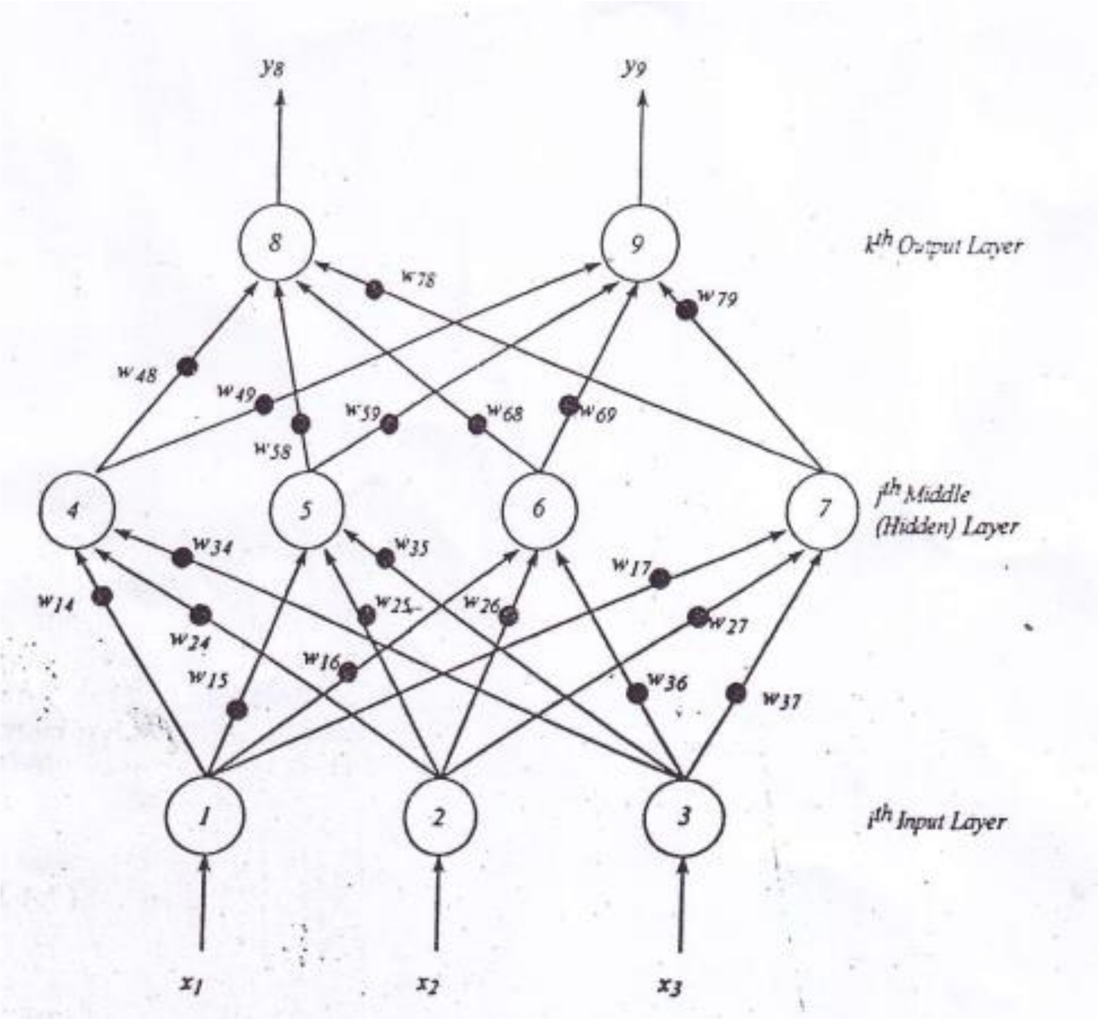
Contoh: **ART, BAM, BSB, Boltzman Machine, Cauchy Machine, RNN**

Important Terms

- Heteroassociative
- Autoassociative
- Fully Connected
- Feedforward Network

Page-197

Simple Feedforward Neural Network

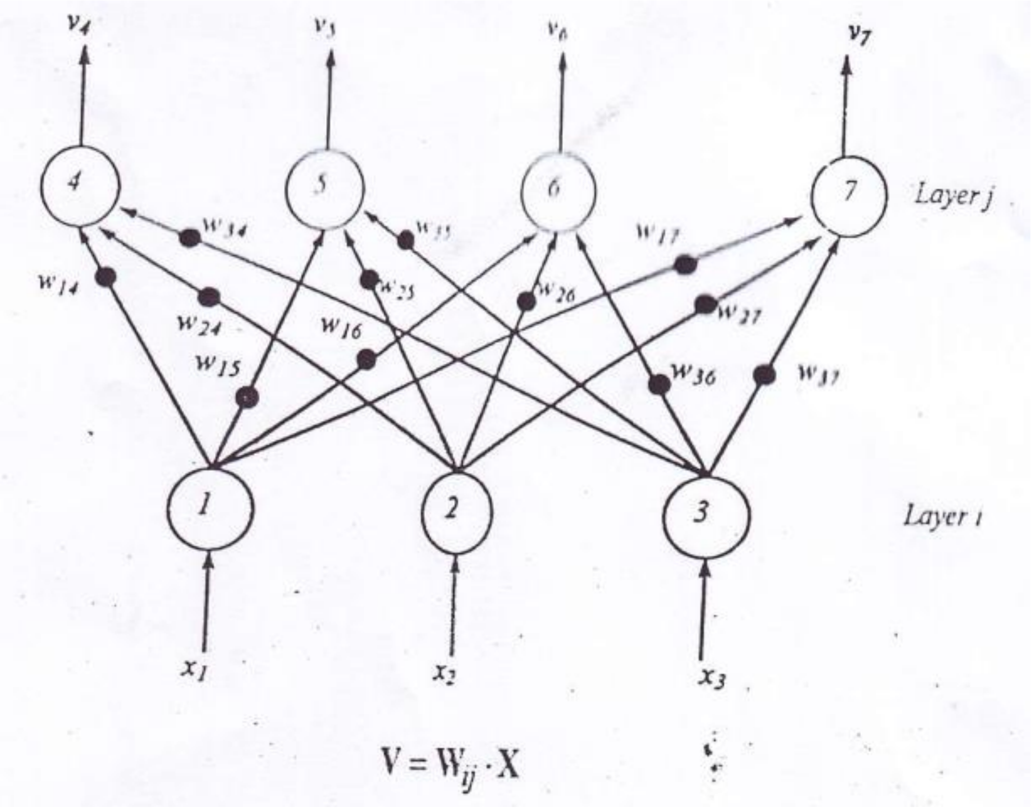


Vector and Matrix Notation

$$\begin{bmatrix} v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = \begin{pmatrix} w_{14} & w_{24} & w_{34} \\ w_{15} & w_{25} & w_{35} \\ w_{16} & w_{26} & w_{26} \\ w_{17} & w_{27} & w_{37} \end{pmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$V_j = Wij \cdot Xi$$

Lower Portion of NN

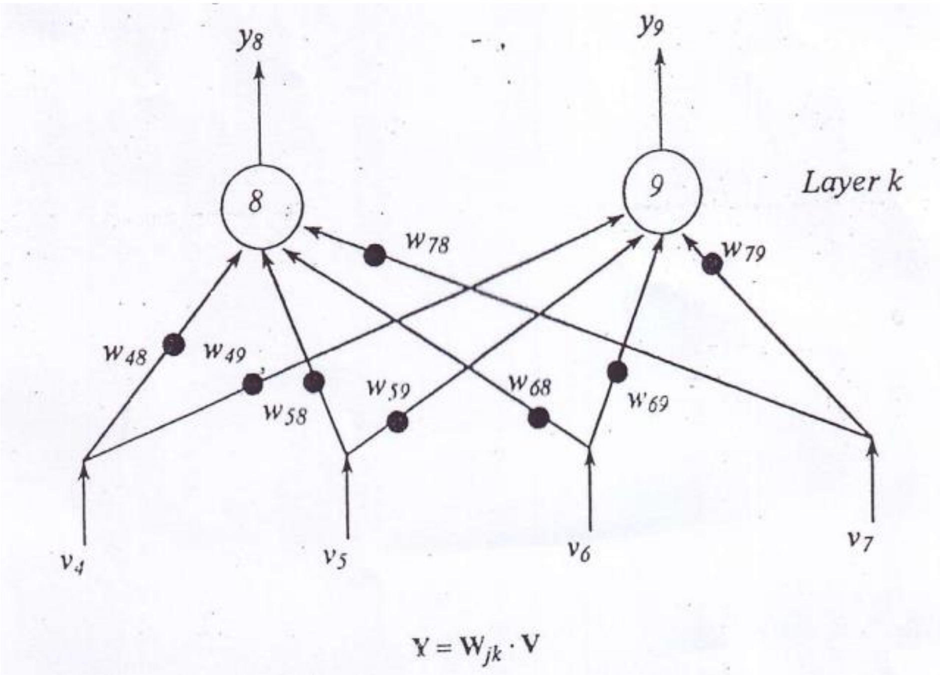


Vector and Matrix Notation

$$\begin{pmatrix} y_8 \\ y_9 \end{pmatrix} = \begin{pmatrix} w_{48} & w_{58} & w_{68} & w_{78} \\ w_{49} & w_{59} & w_{69} & w_{79} \end{pmatrix} \cdot \begin{pmatrix} v_4 \\ v_5 \\ v_6 \\ v_7 \end{pmatrix}$$

$$Y_k = Wjk \cdot Vi$$

Upper Portion of NN



Vector and Matrix Notation

$$\begin{pmatrix} y_8 \\ y_9 \end{pmatrix} = \begin{pmatrix} w_{48} & w_{58} & w_{68} & w_{78} \\ w_{49} & w_{59} & w_{69} & w_{79} \end{pmatrix} \cdot \begin{pmatrix} w_{14} & w_{24} & w_{34} \\ w_{15} & w_{25} & w_{35} \\ w_{16} & w_{26} & w_{26} \\ w_{17} & w_{27} & w_{37} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$Y_k = Wi_j \cdot W_{jk} \cdot Xi = W_{ijk} \cdot Xi$$

Major Function

- ❑ Learning
  - Process of adapting the connection weight in an ANN to produce the desired output vector in response to a stimulus vector presented in the input buffer
- ❑ Recall
  - Process of accepting an input stimulus and producing an output response in accordance with the network weight structure

Strategy / Learning Algorithm

Supervised Learning

Labeled Data (Input+Output both are available)

Analogy: A student learning with a textbook that has the questions and the answers.

- Learning is performed by presenting pattern with target
- During learning, produced output is compared with the desired output
  - The difference between both output is used to modify learning weights according to the learning algorithm
- Recognizing hand-written digits, pattern recognition and etc.
- Neural Network models: perceptron, feed-forward, radial basis function, support vector machine.

Unsupervised Learning

Unlabeled Data (Only Inputs)

Analogy: Giving a child a bunch of different fruits without names and asking them to sort them into groups.

- Targets are not provided
- Appropriate for clustering task
  - Find similar groups of documents in the web, content addressable memory, clustering.
- Neural Network models: Kohonen, self organizing maps, Hopfield networks.

Graded/ Reinforcement Learning

- Output is graded as good or bad on a numerical scale and the connection weights are adjusted in accordance with the grade

No static dataset. An agent interacts with an environment.

Learning by interaction and feedback from actions.

Sequential Decision Making. Learn a policy of actions to maximize cumulative reward.

Analogy: Training a dog with treats (rewards) and scolds (penalties). The dog learns which tricks lead to the best outcomes.

Features of ANN

- Learn by example
- Constitute a distributed, associative memory
- Fault-tolerant
- Capable of pattern recognition

Where can neural network systems help…

- when we can't formulate an algorithmic solution.
- when we ***can*** get lots of examples of the behavior we require.

  ‘learning from experience’

- when we need to pick out the structure from existing data.

Who is interested?...

- Electrical Engineers – signal processing, control theory
- Computer Engineers – robotics
- Computer Scientists – artificial intelligence, pattern recognition
- Mathematicians – modelling tool when explicit relationships are unknown

Problem Domains

- Storing and recalling patterns
- Classifying patterns
- Mapping inputs onto outputs
- Grouping similar patterns
- Finding solutions to constrained optimization problems

ANN Applications

Applications of ANNs

- Signal processing
- Pattern recognition, e.g. handwritten characters or face identification.
- Diagnosis or mapping symptoms to a medical case.
- Speech recognition
- Human Emotion Detection
- Educational Loan Forecasting

Advantages Of NN

NON-LINEARITY

It can model non-linear systems

INPUT-OUTPUT MAPPING

It can derive a relationship between a set of input & output responses

ADAPTIVITY

The ability to learn allows the network to adapt to changes in the surrounding environment

EVIDENTIAL RESPONSE

It can provide a confidence level to a given solution

Advantages Of NN

CONTEXTUAL INFORMATION

Knowledge is presented by the structure of the network. Every neuron in the network is potentially affected by the global activity of all other neurons in the network. Consequently, contextual information is dealt with naturally in the network.

FAULT TOLERANCE

Distributed nature of the NN gives it fault tolerant capabilities

NEUROBIOLOGY ANALOGY

Models the architecture of the brain

Comparison of ANN with conventional AI methods

CHARACTERISTICS	TRADITIONAL COMPUTING (including Expert Systems)	ARTIFICIAL NEURAL NETWORKS
Processing style	Sequential	Parallel
Functions	Logically (left brained) via Rules Concepts Calculations	Gestault (right brained) via Images Pictures Controls
Learning Method	by rules (didactically)	by example (Socratically)
Applications	Accounting, word processing, math, inventory, digital communications	Sensor processing, speech recognition, pattern recognition, text recognition

The END