Introduction to Machine Learning

Artificial Neural Networks

Dr Raymond Lee
Associate Professor
Division of Science and Technology
BNU-HKBU United International College

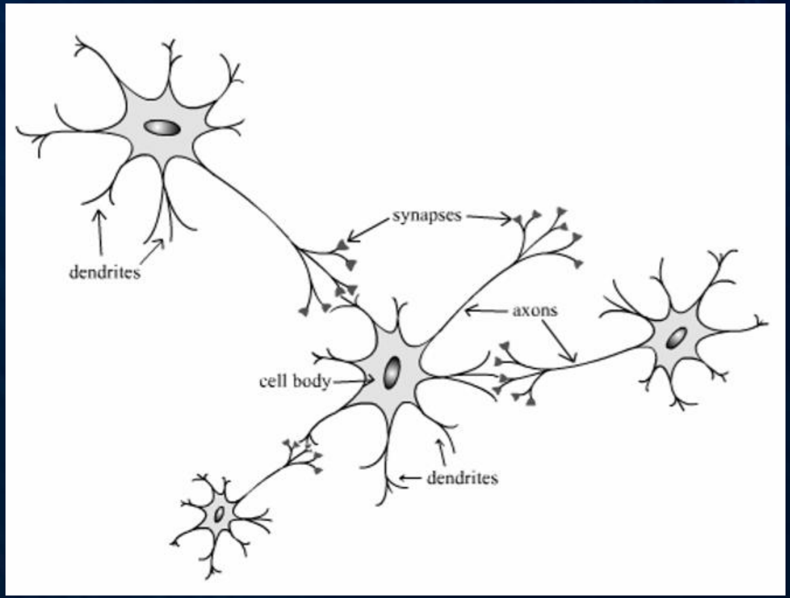
RoadMap

- **Basic concepts – Biological Neural Networks**
- A Neuron Model
- Classification of Neural Networks
- Auto-associative Network
- Hopfield Network
- Feedforward Backpropagation Network
- Summary

Neural Networks

Biological Neural Network – OUR BRAIN

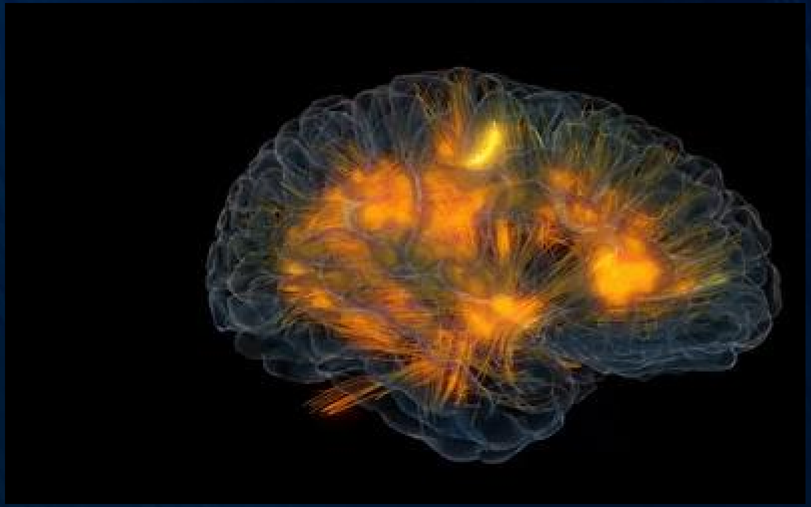
- Artificial Neural Networks, ANN (or *Neural Networks* in short) is one of the major components of (microscopic) AI which focuses on the study and modeling of intelligent system with the mimic of one of the most important organ of humans – The Brain.
- The first scientist to work in the area of "brain science" was an Italian physician Camilo Golgi (1843-1926) who invented the ***stain method*** to investigate the neural activities inside the brain. By using this method, he proposed that the brain is made up of ***syncytium*** – a sponge-like tissue that are ***activated*** by the staining operation.
- Based on his discovery, a Spanish neuroanatomist Santiago Ramón y Cajal (1852-1934) proposed an innovative idea that "These staining tissues were not sponge-like elements, but rather the collections of the brain cells called "neurons", which were interlinked together to form a complex group – "neural networks" (as shown).
- As shown in the figure, each neural cell (neuron) consists of:
  ① Nucleus – central body of the neuron;
  ② Axon – prolonged filament which connects to other neurons;
  ③ Dendrites – tree-like structures which branch from the neuron;
  ④ Synapse – the axon tips (junctions) that make contact with other neurons by attaching to the dendrites of these neighboring neurons.
- However, how its biological neural network works still a mystery, until 1943.



Biological Neural Networks
(Human brain consists of around 10^{11} neurons)

Neural Networks

Integrate-and-Fire operations in Biological Neural Network



Information processing in brain
(Integrate-and-fire operations in neural network)

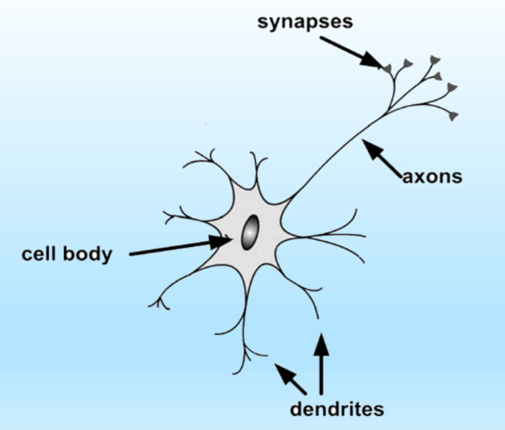
- Before 1943, almost all neuroscientists believed that the sole purpose of the neurons was to process energy and how it works (e.g. process information, store memory) still a mystery.
- In 1943, neurophysiologist Warren McCulloch and mathematician Walter Pitts published an influential paper "A logical calculus of the ideas immanent in nervous activity" which trigger the birth of Artificial Neural Networks (ANN).
- In the paper, they proposed that the main function of neural activities was to ***process information*** – not storing energy. They maintained that the functions of the neurons were just like "logical switches". The transmission of signals from one neuron to another at synapses is the result of complex chemical process in which specific transmitter substances are released from the sending points of the junctions. If the potential reaches a certain "threshold", a pulse will be generated down the axon, known as "firing" (as shown).
- More importantly, in the paper, they demonstrated how their proposed network (now called Artificial Neural Network) could be used to perform basic logical operations such as AND, OR and NOT operations.
- In fact, this breakthrough not only solved the "century mystery" of now biological neural network works, but also provides a solid foundation for the development of digital computing technology.
- Although we now know that the neural activities in our brains are quite different from logical switches such as transistors in digital computer, which are rather like nonlinear (and even chaotic) integrate-and-fire operators for the information transmission and processing, anyway the discovery in 1943 coined the so-called First Golden Age of Artificial Neural Networks.

RoadMap

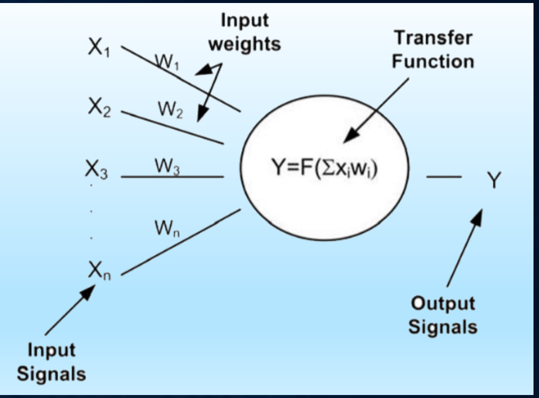
- **Basic concepts – Biological Neural Networks**
- A Neuron Model
- Classification of Neural Networks
- Auto-associative Network
- Hopfield Network
- Feedforward Backpropagation Network
- Recurrent Neural Networks
- Convolutional Neural Networks
- Summary

Neural Networks

A Neuron Model



Biological neuron



A neuron model

- As an direct analogue of the biological neuron, the schematic diagram of neuron structure can be interpreted as a computational model in which the synapses are represented by weights that modulate the effect of the associated input signals, with the formulation given by:

  $$y = f\left[\sum_{i=1}^{n} w_i x_i \right] \tag{7.1}$$

  Where x is the input signals, w's are the weights and y is the output.

- The nonlinear characteristics exhibited by the neuron is represented by a transfer function $f(x)$ such as a binary or bipolar sigmoid function, given by:

  Binary sigmoid function : $$f(x) = \frac{1}{1+e^{-\sigma x}} \tag{7.2}$$

  Bipolar sigmoid function : $$f(x) = \frac{1-e^{-\sigma x}}{1+e^{-\sigma x}} \tag{7.3}$$

  Where σ is the steepness parameter to control the curvature of the transfer function.

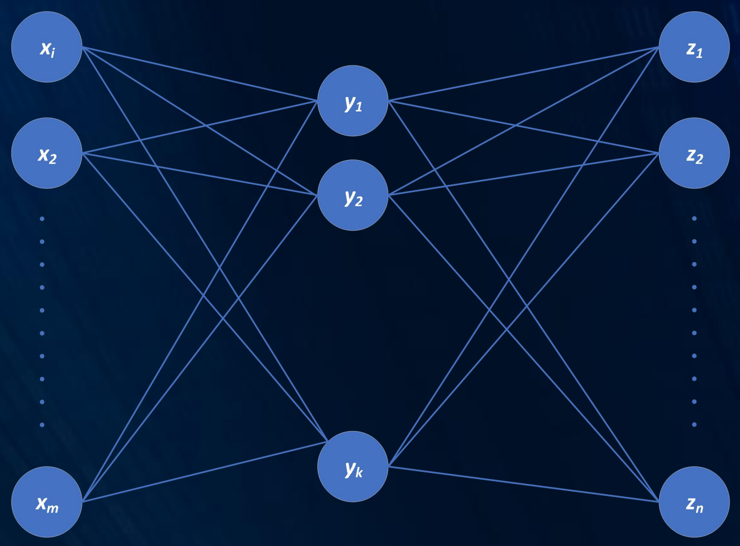
- The learning capability of an artificial neuron is achieved by adjusting the weights in accordance with a predefined learning algorithm, usually in the form of:

  $$\Delta w_j = \alpha \sigma x_j \tag{7.4}$$

  Where α is the "learning rate" and σ is the "learning momentum".

Neural Networks

Typical single-layer neural network model



Typical single hidden layer
Neural Network Model

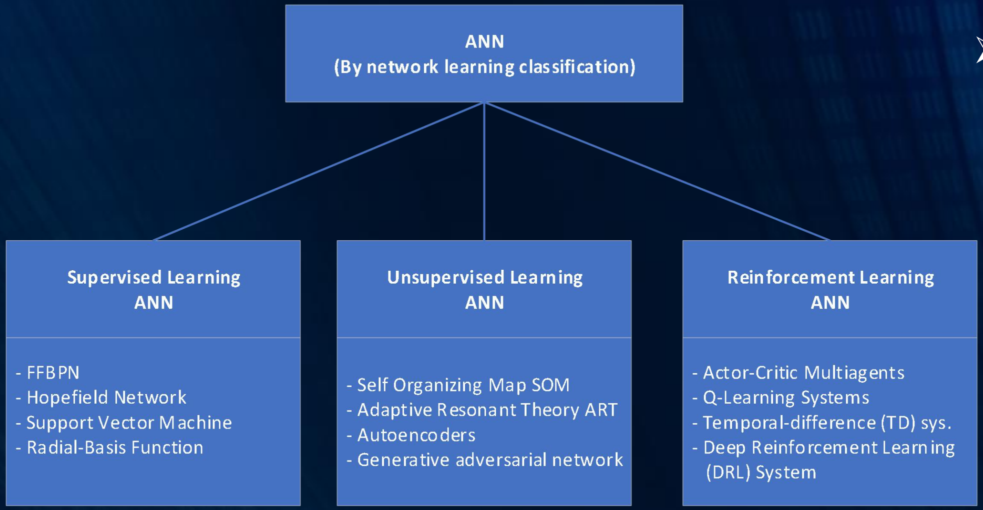
- Typical Artificial Neural Network (ANN) consist of intermediate layer(s) known as "Hidden Layers" to facilitate the nonlinear computational capabilities of the network system.
- Classical ANNs, such as the Feedforward Neural Network (FFNN) (as shown), allow signals (information) to flow from the input units to the output units, in a forward direction.
- Other basic ANNs include the classical Kohonen Self-organizing Map (SOM) and Learning Vector Quantization (LVQ) – based on competition, and the Adaptive Resonance Theory (ART) and of course our main theme – Feedforward Backpropagation Neural Network (FFBPN).
- ANNs can be regarded as multivariate nonlinear analytical tools, and are known to be superior at recognizing patterns from noisy, complex data, and estimating their nonlinear relationships.
- Many studies revealed that ANNs have the distinguished capability to learn the underlying mechanics of time series problems ranging from prediction of stocks and foreign exchange rates in various financial markets to the weather forecast.

RoadMap

- **Basic concepts – Biological Neural Networks**
- A Neuron Model
- Classification of Neural Networks
- Auto-associative Network
- Hopfield Network
- Feedforward Backpropagation Network
- Summary

Neural Networks

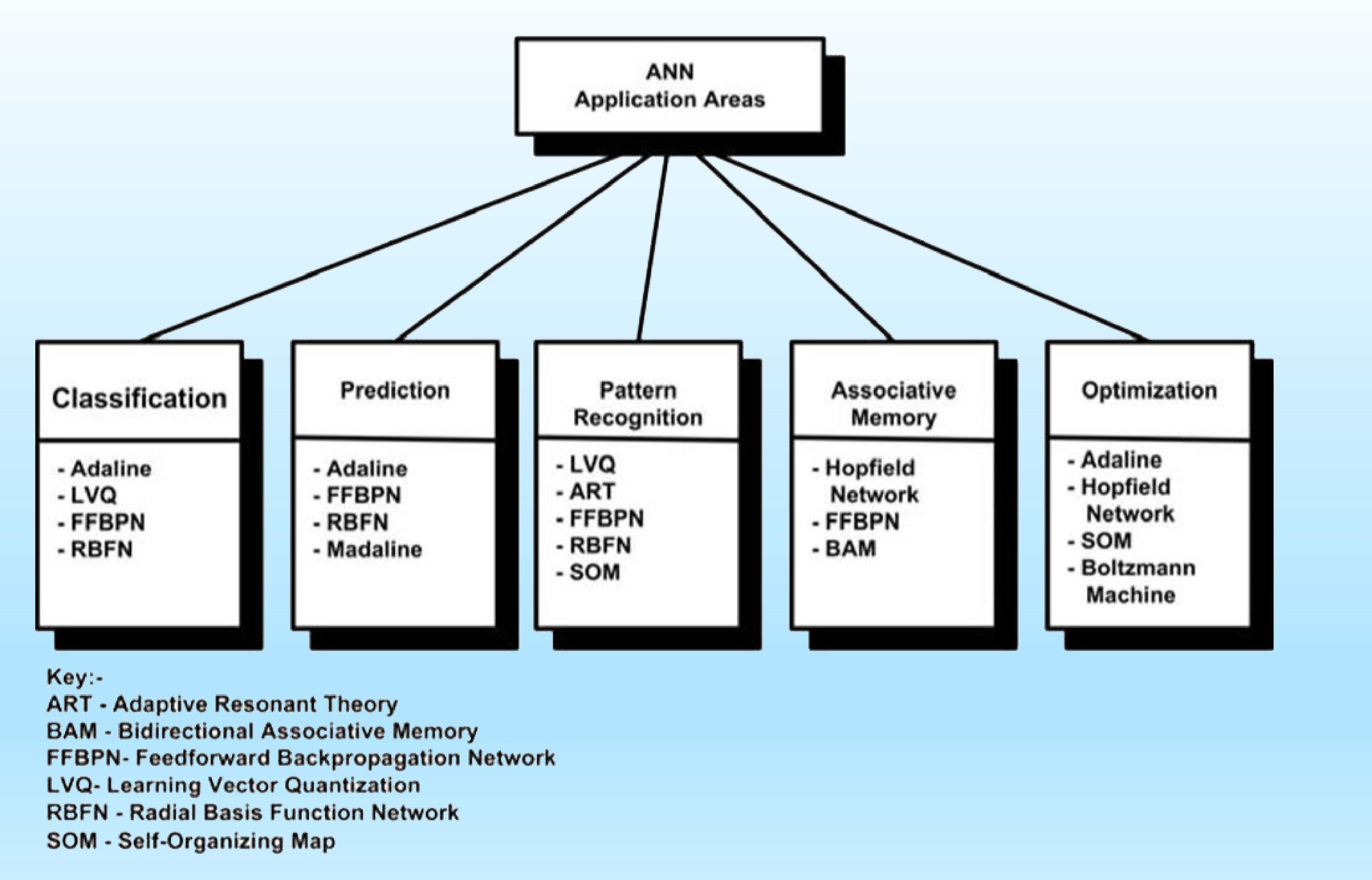
Classification by Machine Learning Technique



- After half a century of development of neural networks, numerous neural networks have been proposed.
- In fact, there are over 20 different types of commonly used ANNs.
- Basically speaking, ANNs are commonly classified by: 1) Machine Learning Technique; and 2) Areas of application.
- In terms of Machine Learning Techniques, ANNs can be categorized into THREE main categories:
  ① Supervised-Learning Neural Networks – network learning (training) based on input-output (target output) pairs. Typical examples includes: Feedforward Backpropagation Neural Network (FFBPN), Hopefield Network, Support Vector Machine (SVM), Radial-Basis Function (RBF) Network, etc.
  ② Unsupervised-Learning Neural Networks - neural networks that do not need any supervised learning and training strategies, including all kinds of self-organizing, self-clustering, and learning networks such as SOM, ART (Adaptive Resonant Theory), etc.
  ③ Reinforcement-Learning Neural Networks - different from Supervised-Learning (SL) with well-defined input-output pairs, Reinforcement-Learning (RL) train the neural network with the adoption of feedback signals namely Reinforcement-Signal (RS). For the right behavior, the network will respond with a positive RS to "award" the RL network; while for the wrong behavior, the network will respond with a negative RS to "punish" the RL network. This method is particularity useful tackle with optimization problem without exact target solutions such as trading strategy optimization, we will discuss it in details in the advanced topics on multiagent-based intelligent trading strategies in chapter 14.

Neural Networks

Classification by Areas of Application



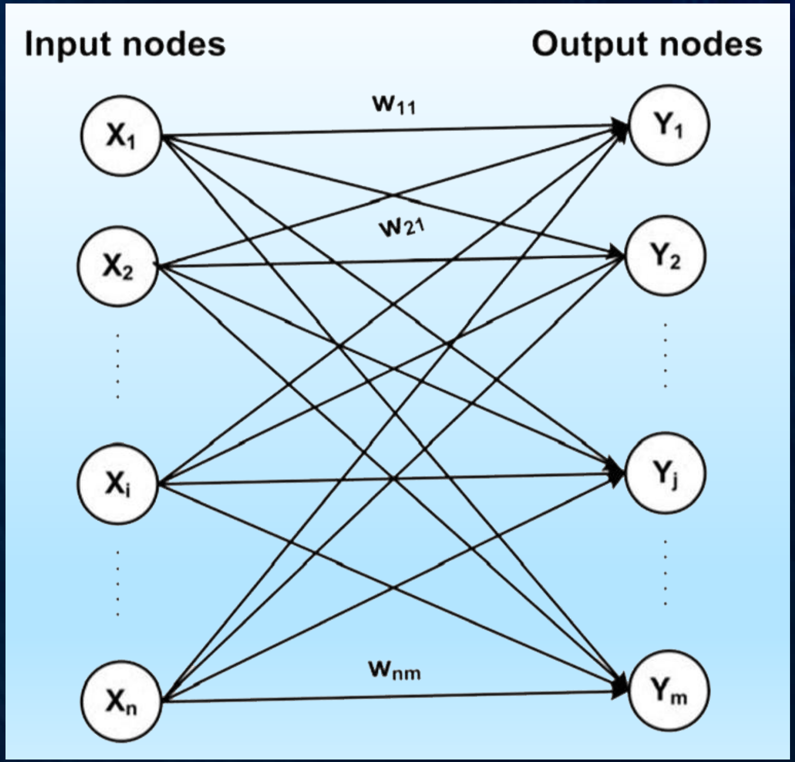
- In terms of Areas of Application, ANNs can be categorized into FIVE major types.
  1. Classification
  2. Prediction
  3. Pattern Recognition
  4. Associative Memory
  5. Optimization
- In this course, we will discuss THREE basic and commonly used ANNs, they are:
  1. Associative Network
  2. Hopfield Network
  3. Feedforward Backpropagation Network (FFBPN)

RoadMap

- **Basic concepts – Biological Neural Networks**
- A Neuron Model
- Classification of Neural Networks
- Auto-associative Network
- Hopfield Network
- Feedforward Backpropagation Network
- Summary

Neural Network Models #1

Auto-associative Network – Network Architecture



Associative Network

- ***Associative Learning*** is one of the major characteristics of human behavior.
- It is one of the most fundamental of human intellectual behaviors.
- It is also widely used by humans and machines for pattern recognition such as visual pattern identification and recognition such as the recalling of human faces, voices and music.
- More importantly, it relates to the association of knowledge and memory recalling.
- Generally speaking, an associative network is a single-layered neural network used to store a set of patterns for pattern association (or what we call "recalling").
- The training of the associative network is conducted by the iterative presentation of the stored patterns for weights updated according to the training algorithm.
- Once the training is completed, the network can be used to associate not only the stored pattern, but also the correct stored pattern upon the presentation of an incomplete or noisy query pattern.
- Basically, there are two major kinds of associative networks:
  ① ***Auto-associative networks*** - in which the input (and query) patterns are of the same type (and nature) as the associated pattern; and
  ② ***Hetero-associative networks*** - in which the input (and query) patterns are of totally different types (and nature) from the associated patterns.

Neural Network Models #1

Auto-associative Network – Network Training Algorithm

Training Algorithm (4.1) – Associative Network (Fausett 1994; Patterson 1996)

Step 1: Network weight initialization

For all i, j — *where* $i \in [1 \ldots n], j \in [1 \ldots m]$, *n, m are the total numbers of neuron input nodes and output nodes respectively.*

Set $w_{i,j} = 0$

Step 2: For each training pair (x', y'), perform the following operations:

Step 2.1: Set the activation values for the input nodes as the values of the training input

i.e., $x_i = x'_i \quad i \in [1 \ldots n]$

Step 2.2: Set the activation values for the output nodes as the values of the target output

i.e., $y_j = y'_j \quad j \in [1 \ldots m]$

Step 2.3: Update ALL the weights in the network

i.e., $w_{ij}(new) = w_{ij}(old) + x_i y_j$

$i \in [1 \ldots n], j \in [1 \ldots m]$

- For network training, the Hebb Rule is commonly used.
- Also, both binary and bipolar vectors can be used in the associative networks, and the training vectors will be a set of training input and target output pairs (x', y').
- Besides the simple Hebb Rule for adjustments to weights, other methods such as the Delta Rule can also be adopted. In that case, the weight adjustment formula (in Step 2.3) will be replaced by:

$$w_{ij}(new) = w_{ij}(old) + \alpha(y'_j - y_j)x_i, \; i \in [1..n], j \in [1..m] \qquad (7.5)$$

where α is the learning rate of the network.

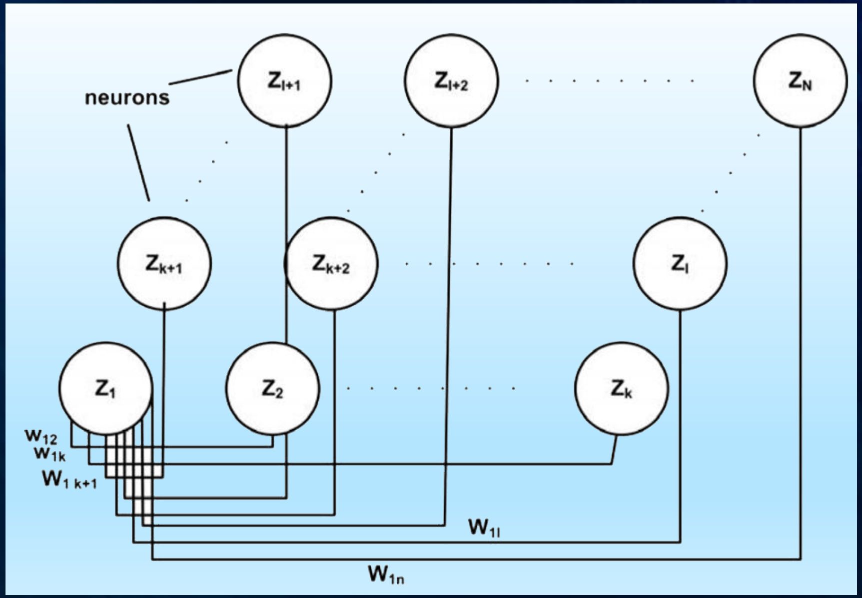
- While the architecture of the associative network is too simple to be used on complex patterns such as human face association and character recognition, it does provide an innovative means of applying neural networks for pattern association, memory storing, and recalling.
- Inspired by these discoveries in neuroscience and neuro-physiology, the author (Lee 2004) proposed a chaotic neural associative network – the Lee-associator – to model the chaotic and progressive human memory recalling mechanisms.

RoadMap

- **Basic concepts – Biological Neural Networks**
- A Neuron Model
- Classification of Neural Networks
- Auto-associative Network
- Hopfield Network
- Feedforward Backpropagation Network
- Summary

Neural Network Models #2

Hopfield Network – Network Architecture



Discrete Hopfield Network

- In 1984, John Hopfield published his influential paper "Neurons with graded response have collective computational properties like those of two-state neurons" (Hopfield 1984).
- He described how a simple recurrent auto-associative network can be used for content-addressable memory systems.
- Which can also be used for pattern recognition and to tackle complex optimization problems such as the typical Traveling Salesman Problem (TSP).
- Basically, the architecture of the Hopfield network is similar to a classical auto-associative network, but with three basic differences:
  ① The Hopfield network is a recurrent network in the sense that the output nodes in one time step are fed as the input in the next time step;
  ② In the classical associative network, all the neurons will update their activations at the same time, but in the Hopfield network only one neuron will be chosen to update its activation at a time and will then "broadcast" its new state to other members of the network;
  ③ Each neuron will keep on receiving the "stimulus" from the external signal during the whole process.

Neural Network Models #2

Hopfield Network – Network Training Algorithm

Training Algorithm (4.2) – Discrete Hopfield Network (Fausett 1994; Patterson 1996)

Step1: Store all the (binary) patterns into the network (using the Hebb Rule). For each pattern $\mathbf{x}'_p = (x'_{p1}, x'_{p2}, \ldots, x'_{pm})$, where p is the pattern number and m is the total number of patterns to store, calculate:

$$w_{ij} = \sum_{p=1}^{m} \left[2x'_{pi} - 1\right]\left[2x'_{pj} - 1\right] \quad \textit{for } i \neq j$$

otherwise $w_{ij} = 0$

Step 2: If neuron activations have not yet converged, do the following:

Step 2.1: Set the initial activation values for the network as the values of the external input vector $\mathbf{x}'$:

i.e., $z_i = x'_i, \quad i \in [1 \,..\, n]$

Step 2.2: For each unit X_i performs Steps 2.2.1–2.2.3 (update unit in random order):

Step 2.2.1: Calculate the network input:

$$z_{in_i} = x_i + \sum_{j=1}^{n} z_j w_{ji}$$

Step 2.2.2. Determine the activation value:

$$z_i = \begin{cases} 1 & \textit{if } z_{in_i} > \theta_i \\ z_i & \textit{if } z_{in_i} = \theta_i \\ 0 & \textit{if } z_{in_i} < \theta_i \end{cases}$$

Step 2.2.3. "Broadcast" the new value of z_i to all neurons in the network

Step 3: Check for the convergence condition.

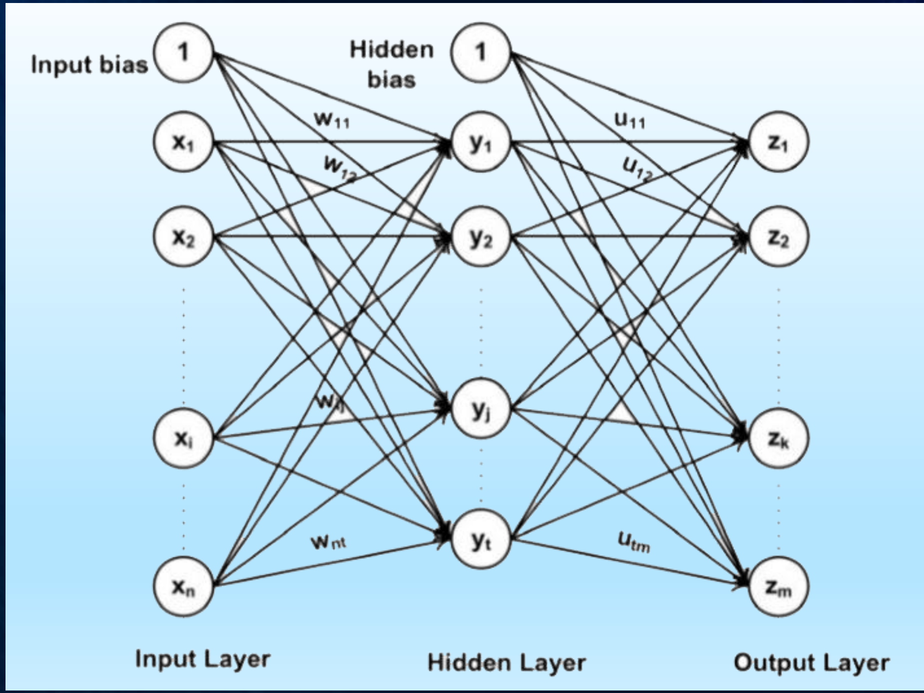
- The next figure shown the training algorithm of a classical Discrete Hopfield Network.
- One of the important points for the Hopfield network is that it demonstrates how a simple auto-associative network can be modified to produce a powerful memory storage and retrieval device.
- In fact, the vast application areas of Hopfield networks also triggered the rebirth of ANNs and the exploration of how neural networks can be applied to complex problems in daily operations.
- Hopfield networks also provide a model for understanding human memory, and important component in Artificial Intelligence.

RoadMap

- **Basic concepts – Biological Neural Networks**
- A Neuron Model
- Classification of Neural Networks
- Auto-associative Network
- Hopfield Network
- Feedforward Backpropagation Network
- Summary

Neural Network Models #3

Feedforward Backpropagation Network (FFBPN) – Network Architecture



Feedforward Backpropagation Network (FFBPN)

- Different from the pervious two neural networks, FFBPNs provide a multilayer network architecture.
- A typical FFBPN consists of an input layer, a hidden layer, and an output layer. Although the FFBPN can consist of several hidden layers, in most of the cases one hidden layer is usually sufficient.
- The network training of FFBPNs consists of three main processes:
  ① The "feedforward" process of network training;
  ② The "error evaluation" process to calculate the errors between the calculated output values and the target output values; and
  ③ The "backpropagation" process of the errors for weight adjustments.
- Similar to most other networks, training stops when the errors are bound within the tolerance level.
- Note:
  - In the network architecture, w's denote the network weights between the input and hidden layer, and u's denote the network weights between the hidden and output layer.
  - The total numbers of neurons in the input, hidden, and output layers are n, t, and m, respectively.
  - For the activation functions, normally a sigmoid function is adopted.

Neural Network Models #3

Feedforward Backpropagation Network (FFBPN) – Network Training Algorithm

Training Algorithm (4.3) – FFBPN (Fausett, 1994; Patterson, 1996)

Step 1: Network weight initialization.
Set all network weights w_{ij}, u_{jk} to a small random number between 0 and 1.

Step 2: While error $\geq$ threshold value, do the following:

Step 2.1: For each training pair (**x**, **z**) do Steps 2.1.1 to 2.1.6.

Feedforward Procedure

Step 2.1.1 Calculate the input state of each hidden node:

$$y_{in_j} = \sum_{i=0}^{n} x_i w_{ij} \quad \text{where } x_0 \text{ is the input bias}$$

Step 2.1.2 Calculate the activation value for the hidden node:

$$y_j = f_y(y_{in_j}) \quad \text{where } f_y(\,) \text{ is the activation function}$$

Step 2.1.3 Calculate the input state of each output node:

$$z_{in_k} = \sum_{j=0}^{t} y_j u_{jk} \quad \text{where } y_0 \text{ is the hidden bias}$$

Step 2.1.4 Calculate the activation value for the output node:

$$z_k = f_z(z_{in_k}) \quad \text{where } f_z(\,) \text{ is the activation function}$$

Backpropagation Procedure

Step 2.1.5 For each output node:

(a) Calculate the error with the target value

$$\zeta_k = (z_k' - z_k) f_z'(z_{in_k}) \quad \text{where } f'(\,) \text{ is } df_z/dz$$

(b) Calculate the correction errors

$$\Delta u_{jk} = \alpha \zeta_k y_j \quad \text{where } \alpha \text{ is the learning rate}$$

Step 2.1.5 For each hidden node:

(a) Calculate the accumulated errors in the hidden node

$$\lambda_{in_j} = \sum_{k=1}^{m} \zeta_k u_{jk}$$

(b) Calculate the correction errors in hidden node

$$\lambda_j = \lambda_{in_j} f_y'(y_{in_k}) \quad \text{where } f'(\,) \text{ is } df_y/dy$$

(c) Calculate the weight adjustments

$$\Delta w_{ij} = \alpha \lambda_j x_i \quad \text{where } f'(\,) \text{ is } df_y/dy$$

Step 2.1.6 Update all weights for the two layers (simultaneously):

$$w_{ij}(new) = w_{ij}(old) + \square w_{ij}$$

$$u_{jk}(new) = u_{jk}(old) + \square u_{jk}$$

Step 2.2: Check the stopping criteria.

- The next figure shown the training algorithm of a typical FFBPN.
- An FFBPN can model various kinds of problems such as weather prediction and stock forecasting, pattern recognition problems such as character recognition, classification, and even optimization problems.
- For Quantum Finance stock (or forex) prediction, time-series financial data (e.g. Daily O, H, L, C, V) are normalized, together with QPLs evaluated in the previous chapters and used as input signals for network training, for the calculation of next-day forecast H/L.
- However, the FFBPN has certain intrinsic problems and limitations such as the trapping in local minima and the difficulty in the choice of optimal parameter settings (and input vectors); the slow rate of convergence is also another major consideration.
- One possible solution is the integration with Chaotic Neural Oscillator technique, which will be discussed in the coming lectures.

RoadMap

- **Basic concepts – Biological Neural Networks**
- A Neuron Model
- Classification of Neural Networks
- Auto-associative Network
- Hopfield Network
- Feedforward Backpropagation Network
- Summary

Summary
Artificial Neural Networks

- In this section, we have explored an overview of neural networks, their basic structure, and mechanisms.
- In fact, current trends in the research and development of neural networks are focused on three major areas:
  ① The R&D of Deep Neural Networks on DMML such as:
    ① Recurrent Neural Networks (RNN)
    ② Convolutional Neural Networks
  ② The integration of other AI techniques to remedy some intrinsic limitations of the classical neural networks, including the integration of fuzzy logic – fuzzy–neuro systems – and the integration with GAs (genetic algorithms) to overcome the problem of parameter selection and fine tuning of networks;
  ③ The investigation of the neural dynamics, especially in the study of neural oscillators and neural oscillatory models; and
  ④ The investigation and study of the chaotic neural dynamics of chaotic neural networks to model complex AI problems.

Next

Convolution Neural Networks