-Artificial Neural Network(ANN) -

Chapter 1 Introduction

朝陽科技大學
資訊管理系
李麗華 教授

Background

- 自古人類即對”人”為什麼能有這麼多能力很好奇，人們很好奇大腦的功能，人類也不斷的從生物的角度來觀察人的智慧如何產生及如何運作。

  你大頭啦~~ http://www.dls.ym.edu.tw/neuroscience/nsdivide_c.htm#cns

  真是神經”析析” http://www.dls.ym.edu.tw/neuroscience/alz_c.html

- 在電腦早期出現時(1940年代)即有許多學者不斷提出想要模擬人類思考或推理的學理，最有名的就像是AI領域及ANN。

人類的好奇與探索花絮

1. 解剖高手達文西(Leonardo da Vinci):

達文西他=不但是義大利的建築師、雕刻家、發明家、 工程師、畫家，他也是位解剖學者，達文西共解剖了30具人體，依解剖內容繪製超過200篇畫作，其中包含腦的頭骨形態及腦部不同的交叉截面圖(橫斷面、縱切面、正切面）。他的手稿亦包含論人的記憶、智力等。

（資料來源: http://zh.wikipedia.org/
http://zh.wikipedia.org/w/index.php?title=%E5%88%97%E5%A5%A5%E7%BA%B3%E5%A4%9A%C2%B7%E8%BE%BE%E8%8A%AC%E5%A5%87&variant=zh-tw）

人類的好奇與探索花絮

2. 愛因斯坦大腦人人愛

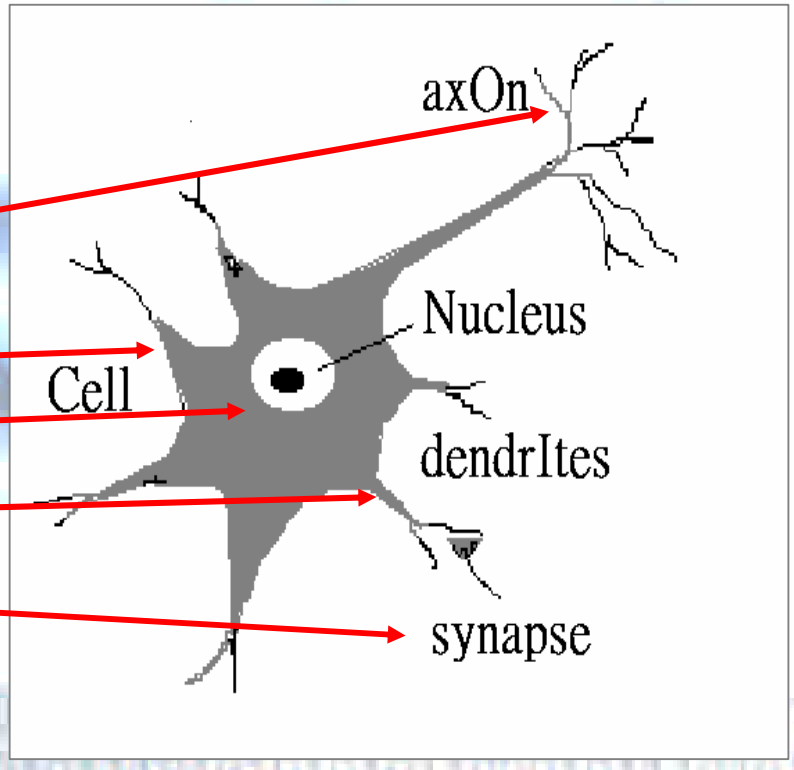
- 1985年，由美國神經科學家戴蒙(Marian Diamond)領銜，發表了第一篇研究愛因斯坦大腦的報告。
- 報告指出，愛因斯坦的左頂葉，神經元與神經膠細胞的比例小於常人。
- 根據過去的研究，哺乳類神經元與神經膠細胞比例，從小鼠到人有逐步降低的趨勢，有些學者因而推測，神經元執行的功能越複雜，越需要神經膠細胞的支持。
- 第二篇研究論文發表於1996年，由神經學安德森(Britt Anderson)助理教授提出，愛因斯坦的大腦皮質中，**神經元密度較高**。這表示愛因斯坦**大腦皮質神經元有較佳的傳訊效率**，因而可以解釋愛因斯坦的超卓天才。

（資料來源: http://tw.epochtimes.com/bt/4/12/31/n765455.htm）

現在人類對大腦與神精的了解已有許多進展
http://www.dls.ym.edu.tw/neuroscience/introb_c.html#bb

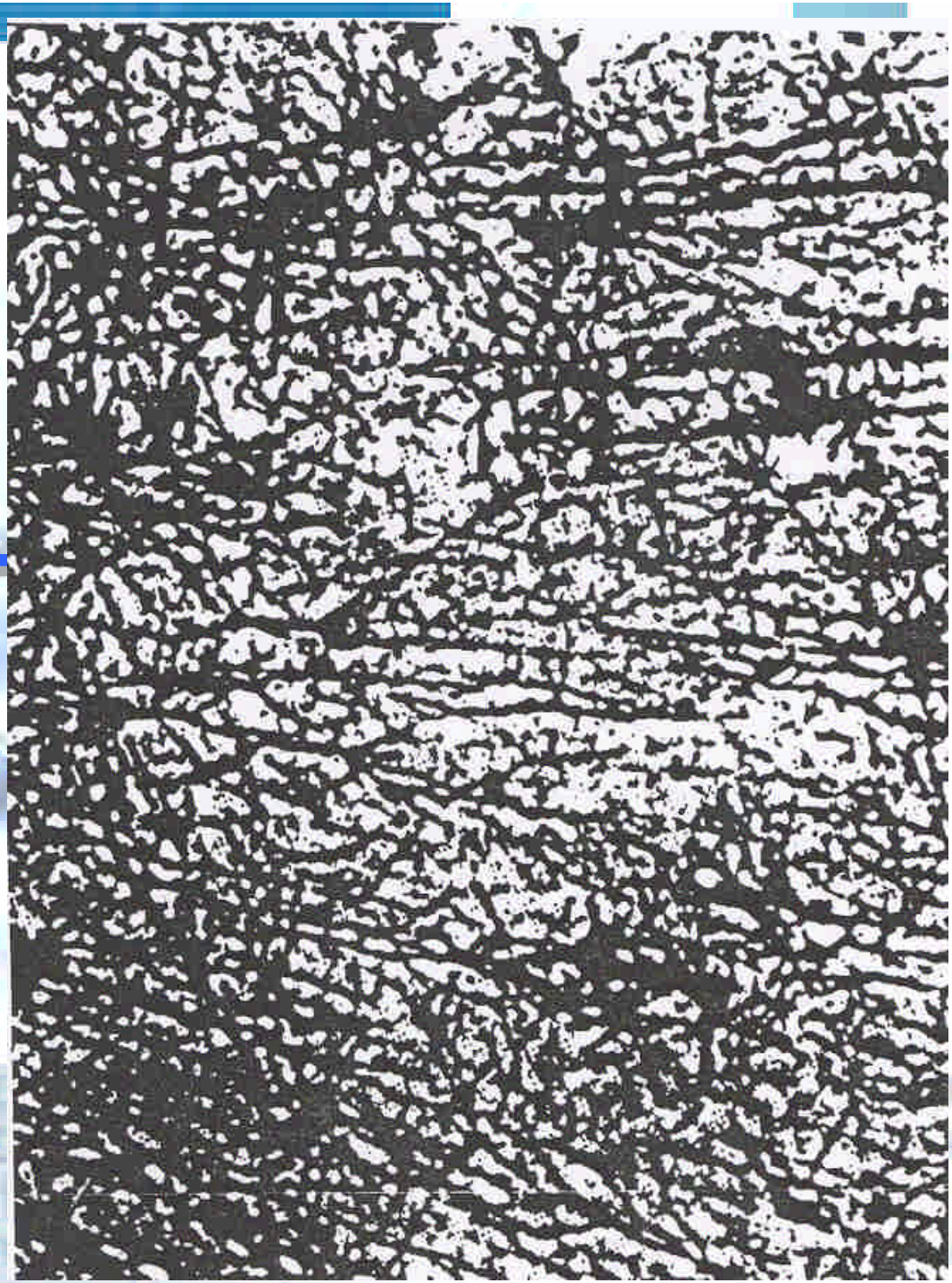
Biological Neural

- 人類自古即對解開人類的大腦及如何思考有極大的興趣。早期學者欲透過人腦的結構及運作方式，思考人腦的奧妙
- **人腦：**大約有1000億個神經細胞（neuron），每個神經細胞約有1000個神經連結
  (共約有100萬億根的神經連結)
- **神經細胞主要元素:**
  - Axon(神經軸)
  - Cell body(neuron)神經元
  - Nucleus(神經核)
  - Dandrites(神經樹)
  - synapse(神經節)

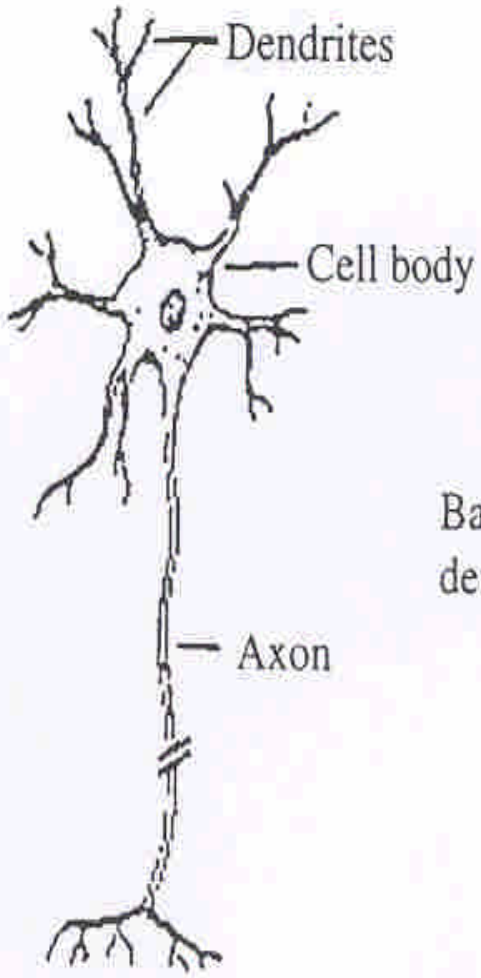


- **作用方式:** 可傳導化學物質
  Two Status：Excitatory／Inhibitory

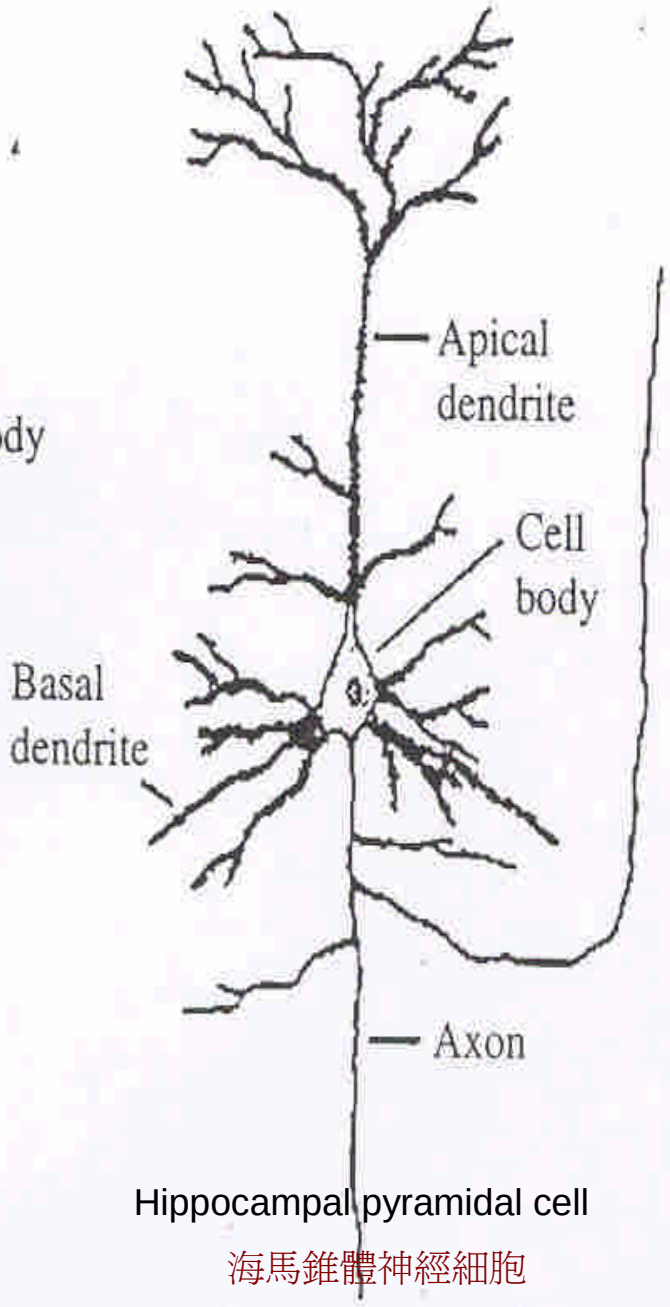
Question:
What is this?



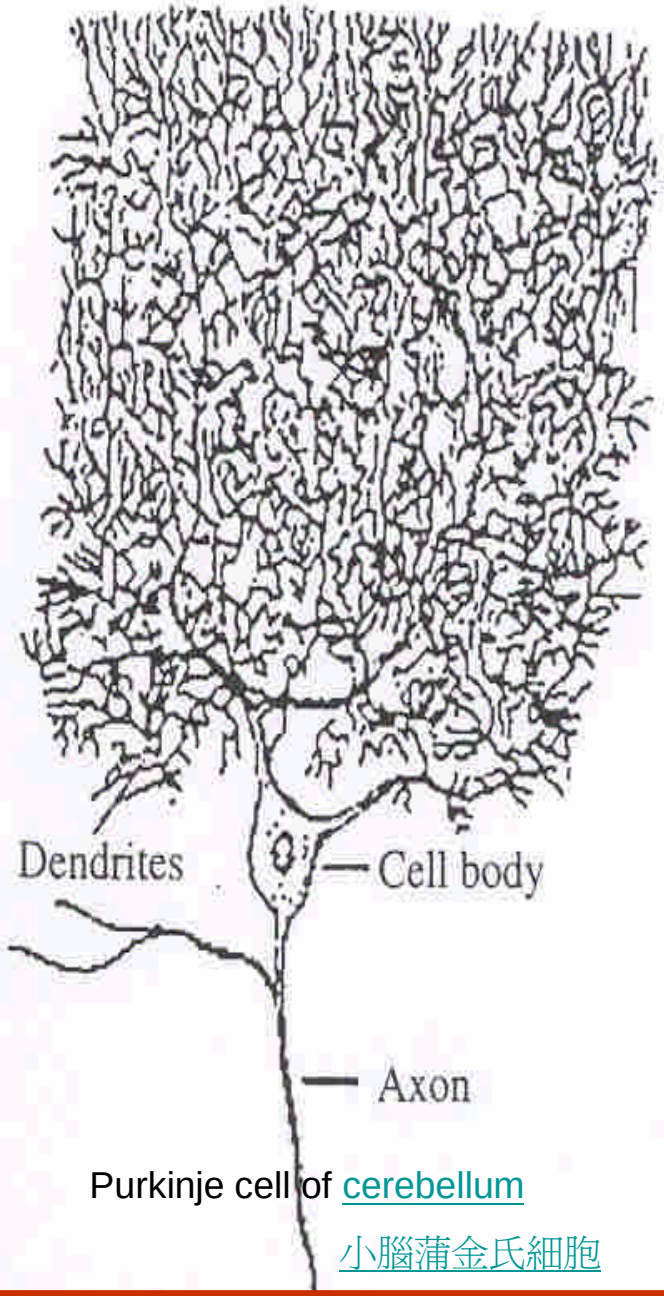
A Slice of Neurons



Spinal motor neuron
脊髓運動神經



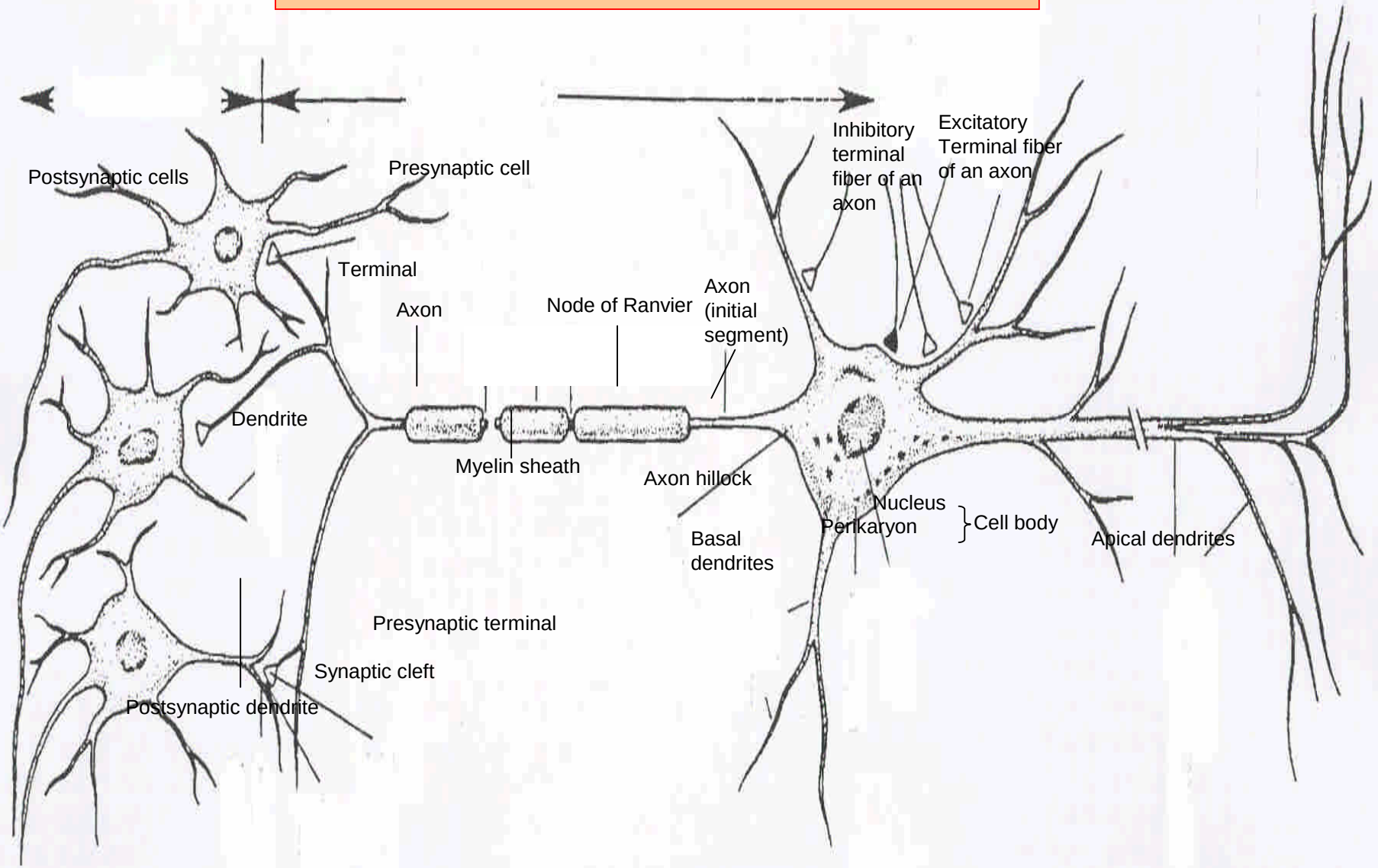
Hippocampal pyramidal cell
海馬錐體神經細胞



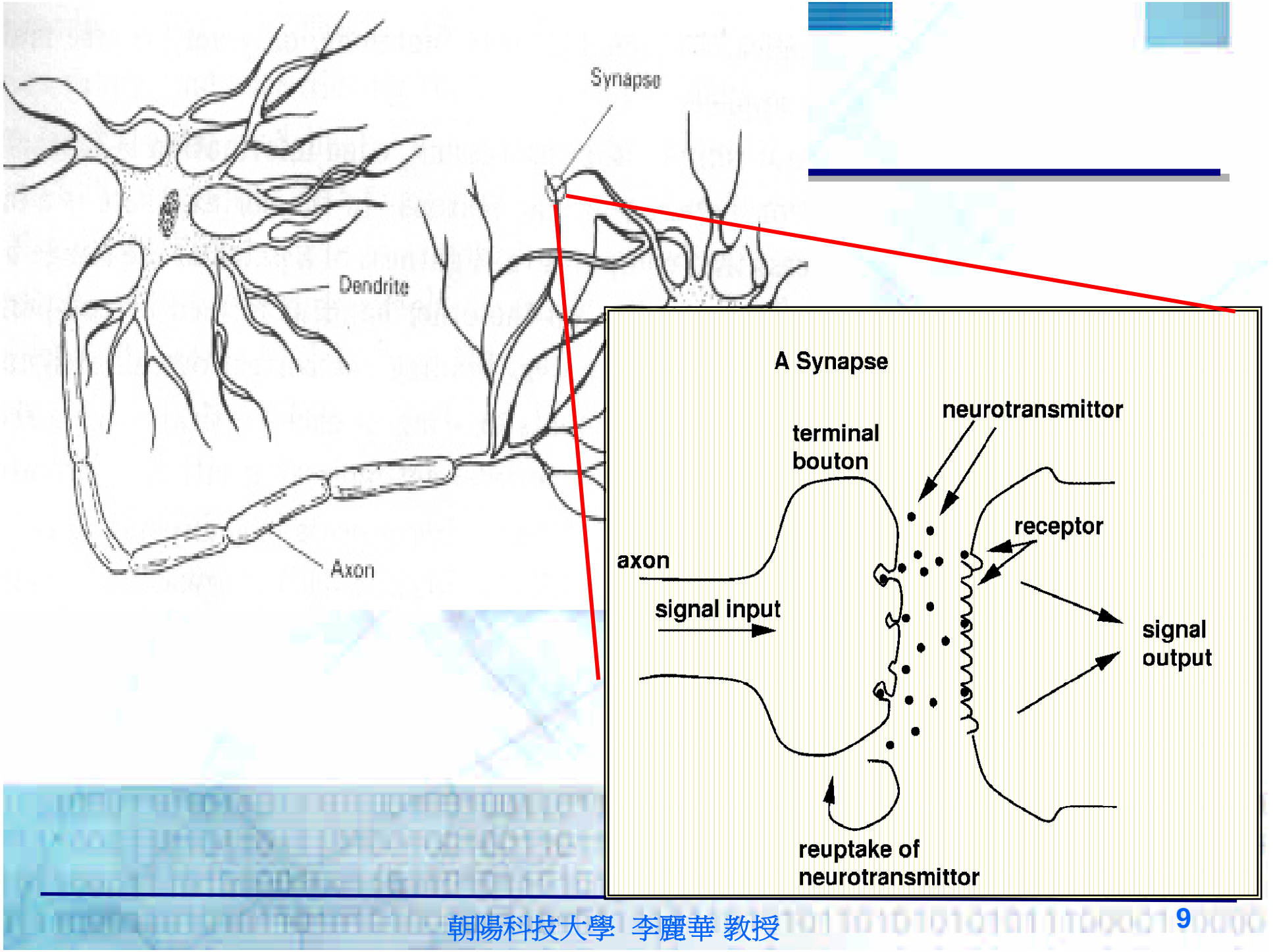
Purkinje cell of cerebellum
小腦蒲金氏細胞

Examples of multipolar cells [Kan91]

照過來……瞧瞧我神經CC的內在美



Expanded view of single neuron morphology [Kan91]


Synapse
Dendrite
Axon
A Synapse
neurotransmittor
terminal
bouton
receptor
axon
signal input
signal
output
reuptake of
neurotransmittor

Introduction to ANN

Def: A structure (network) composed of a number of interconnected units (artificial neurons). Each unit has an I/O characteristic and implements a local computation or function.

The output of any unit is determined by its I/O characteristic, its interconnection to other units, and (possibly) external inputs.

Although “hand crafting” of the network is possible, the network usually develops an overall functionality through one or more forms of training.

Introduction to ANN

中文的ANN定義:

類神經網路是一種計算系統，包括軟體與硬體，它使用大量簡單的相連人工神經來模仿生物神經網路的能力.

人工神經是生物神經的簡單模擬，它從外界環境或者其它人工神經取得資訊，並加以簡單運算，並將其結果輸出到外界環境或者其它人工神經元.

ANN History

1. Creation Age(孕育期)(before1956)
2. Birth Age(誕生期)(1957-1968)
3. Dark Age(挫折期)(1969-1981)
4. Reborn Age(重生期)(1982-1986)
5. Mature Age(成熟期)(1987~present)

Creation Age

- In 1943, McCulloch & Pitts first proposed the neural mathematical model, earlier than the first computer were created. Their premises(assumptions) are:

  1. Neuron has two status, i.e., all or none. (excitatory or inhibitory)
  2. A neuron is triggered by certain amount of accumulated synapses and the trigger action has nothing to do with the previous status.
  3. Only synapse can be delayed.
  4. When under inhibitory status a neuron can not be trigged.
  5. The neuron net structure is not changed.

Creation Age

- In 1949, **Hebb** proposed the Hebb learning rule:
  1. Information resides in Synapes
  2. learning rule
  3. symmetrical weights
  4. When a group of weak neurons is triggered, the strength of connection between them is increased. (即鄰近neuron的訊號可以累積增強weights的值)

Birth Age

- In 1957, **Rosenblatt** proposed the first network model, i.e., Perceptron.（即感知器；當時僅有單層架構）
- In 1960, Widrow porposed another model, i.e., Adaline.(這是屬於連續值的線性網路, 已開始採用 learning rule與當時人工智慧邏輯推理很不同)

Dark Age

- In 1969, Minsky & Papert proved that Perceptron has limited learning usage, because this model cannot solve the XOR problem.

(註) 因當時AI較紅,電腦速度慢, 由於被Minsky的證明困住等原因，ANN研究因此而限於低潮) 但此時期仍有: 芬蘭的Kohonen所提出的Self Organization Map(SOM)與 Grossberg學者等提出Adaptive Resonance Theory(ART) model.

Reborn Age

- In 1982, **Hopfield** proposed the **Hopfield Network** and the **Auto－associative Memory Models**.
- In 1985, Hopfield proposed another **Hopfield & Tank Network** to solve the Traveling Salesman Problem.
- After these researches, the ANN models were again treated as a promising research area .

Reborn Age

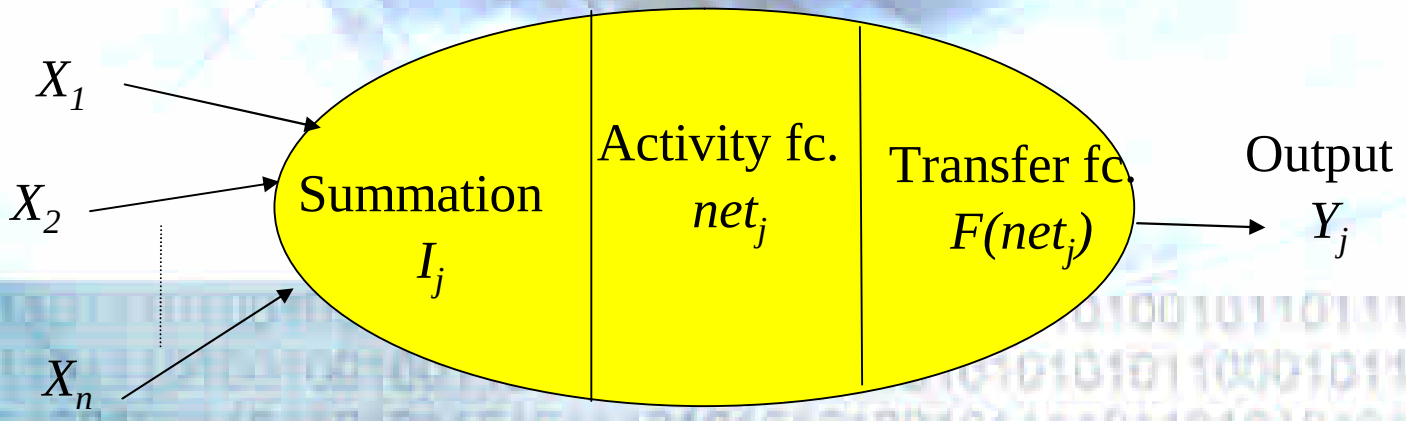
- In 1986, Rumelhart et al. introduced the BPN in their book: “Parallel Distributed Processing” in which generalized delta rule are explained. In addition, they explain how BPN can solve the XOR problem.
- Until 1990, BPN had become one of the most popular and highly utilized ANN model.

Mature Age

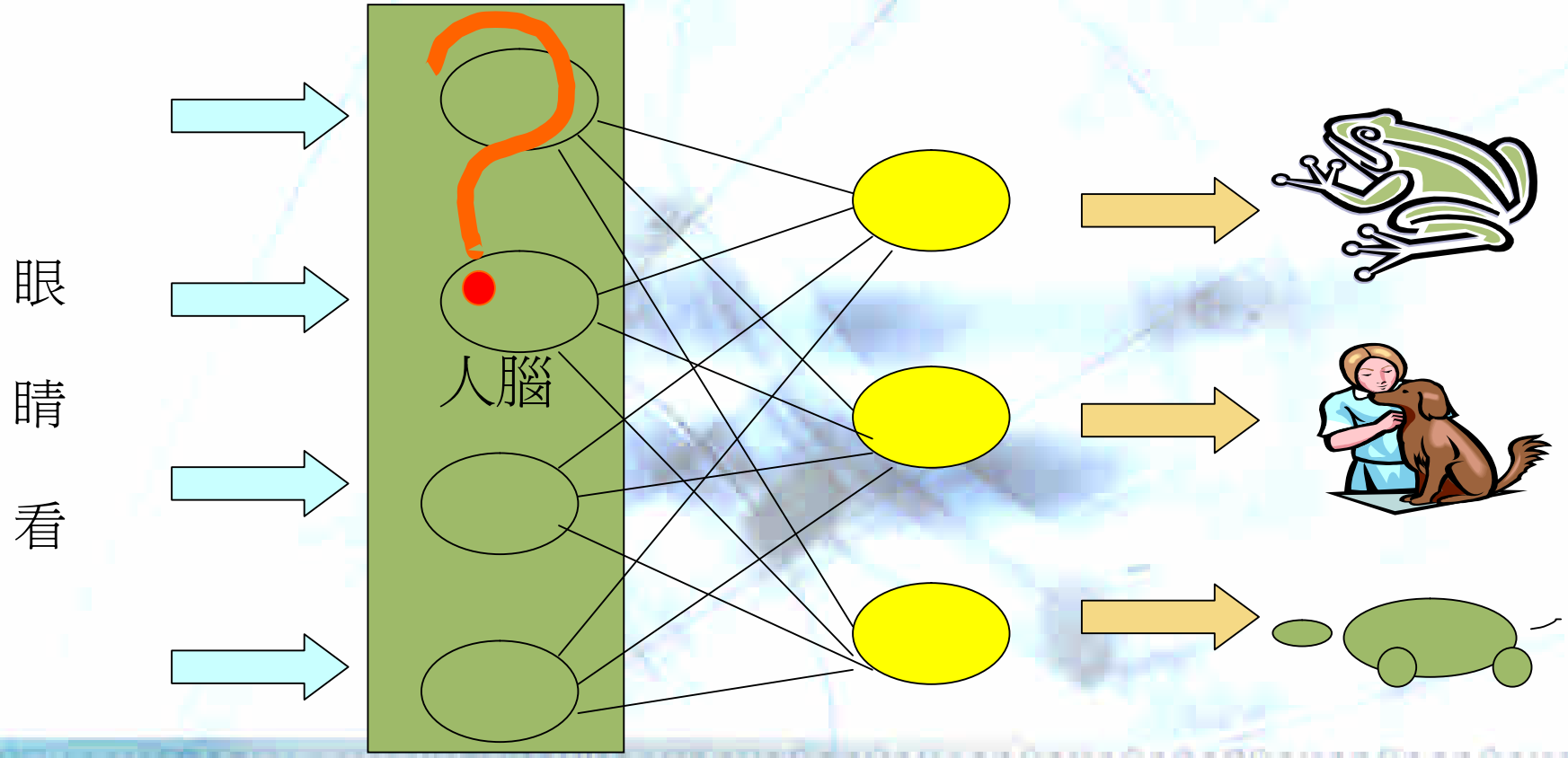
- Up to now, the ANN models has been widely studied and many models has been proposed.
- Conferences and Journals are created for ANN studies, such as ICNN(International Conference on NN, IJCNN(International Joint Conference on NN, held by IEEE & INNS).
- Besides, many tools and software，such as SNNS and MatLab, are been developed for making applying NN easier．

The Node Characteristics of ANN

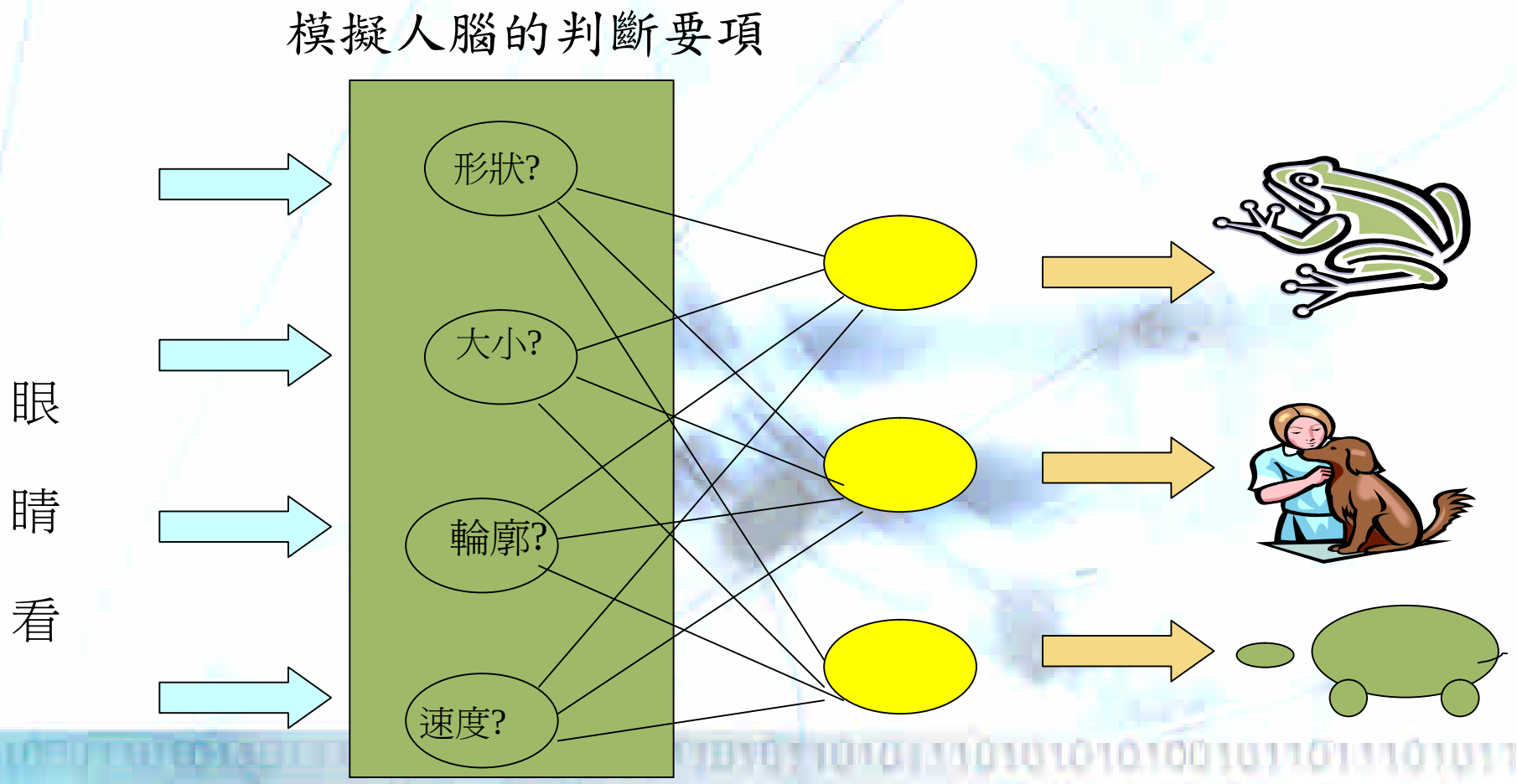
1. Input: training sets(or training patterns), $X=[X_1, X_2, \ldots, X_n]$.
2. Output: computed output $Y=[Y_1,Y_2,\ldots,Y_j]$, testing sets $T=[T_1,T_2,\ldots,T_j]$
3. Connections: Weights, W_{ij}.
4. Processing Element(PE): Summation function, Activity function, & Transfer function.



視覺應用案例



視覺應用案例



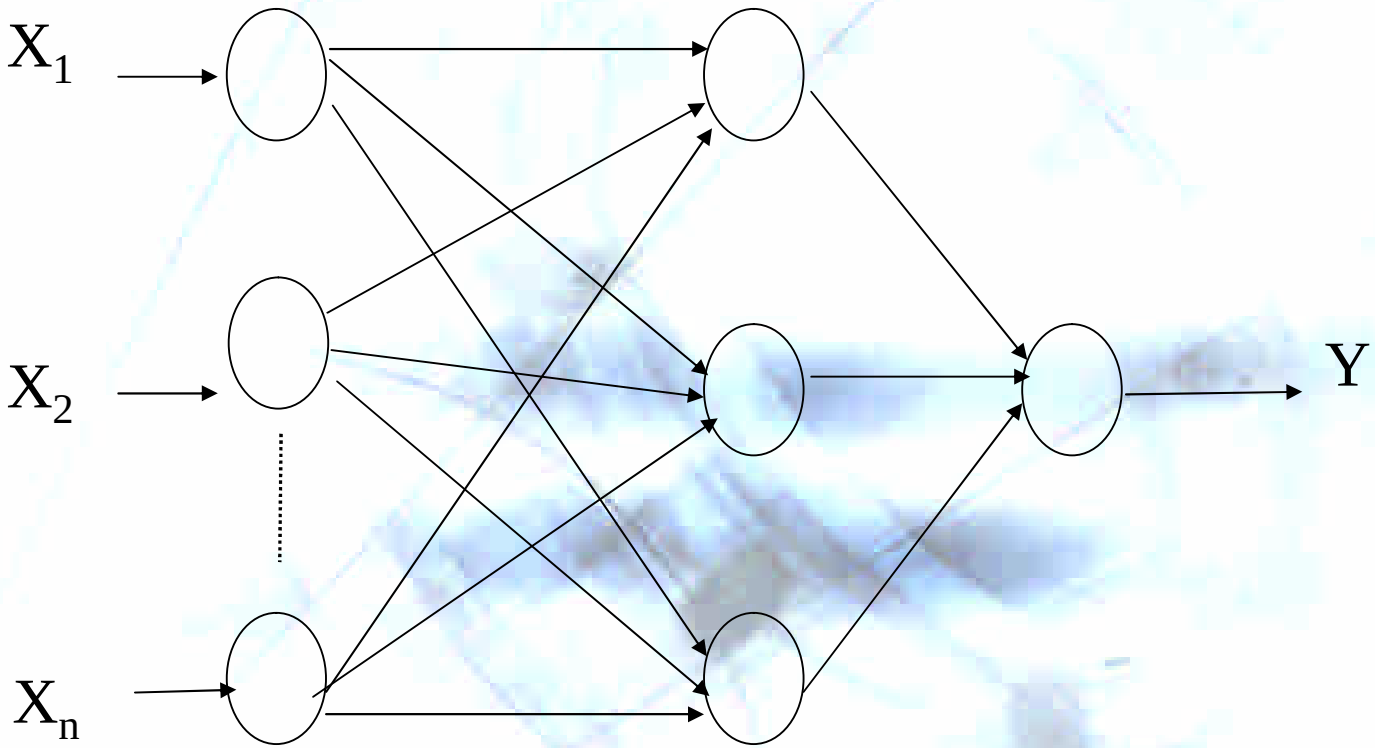
Types of ANN

- According to Learning Type:
  - Supervise Learning: 運用已知的一組輸入範例及預期答案來訓練網路,例如: Perceptron, BPN, PNN, LVQ, CPN
  - Unsupervise Learning: 不斷透過輸入的範例來做學習與修正網路, 例:SOM, ART
  - Associative memory Learning: 直接訓練並記憶所訓練過的所有對照資料or 圖形, 例: Hopfield, Bidirectional Associative Memory(BAM), Hopfield-Tank
  - Optimization Application:找尋最佳解, 例: ANN, HTN

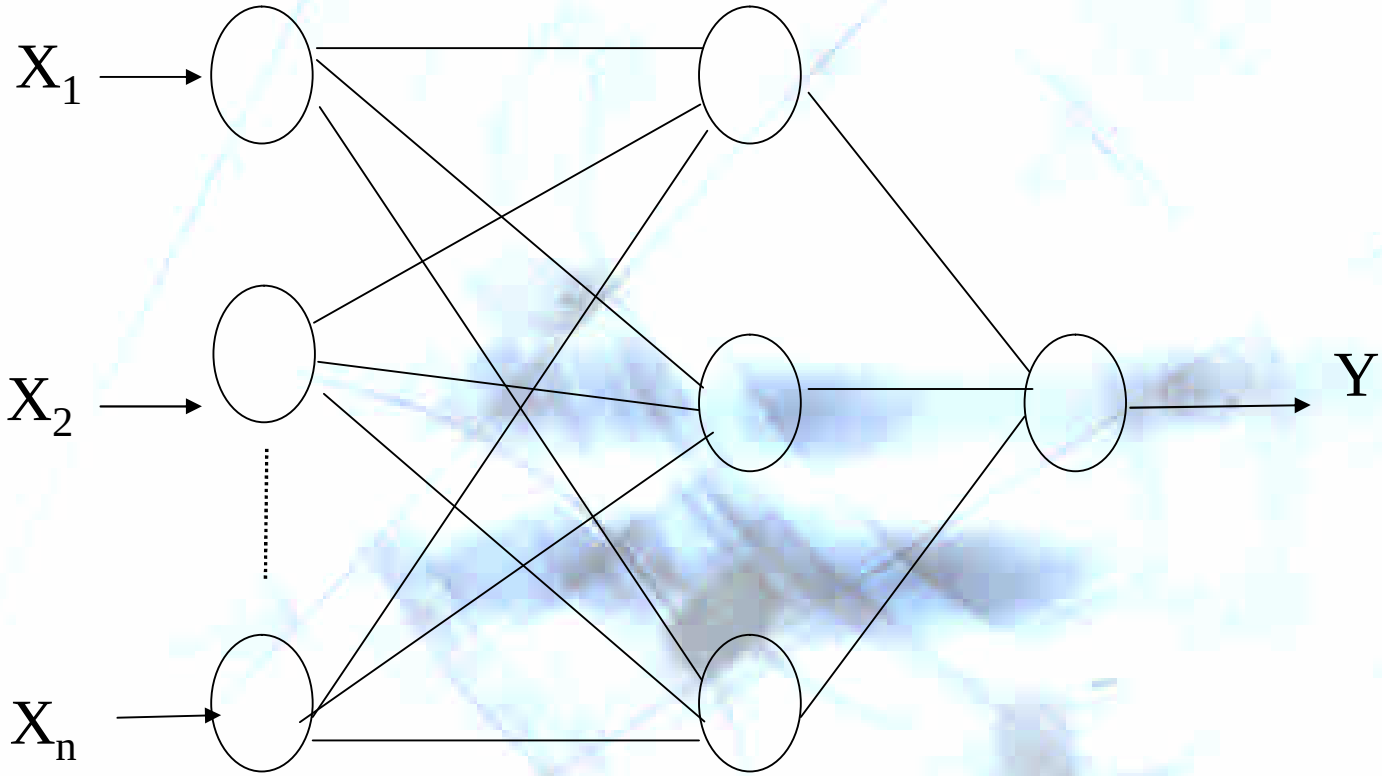
Types of ANN

- According to Network Structure:
  - Feedforward (one-way)
  - Feedforward (Two-way)
  - Feedback

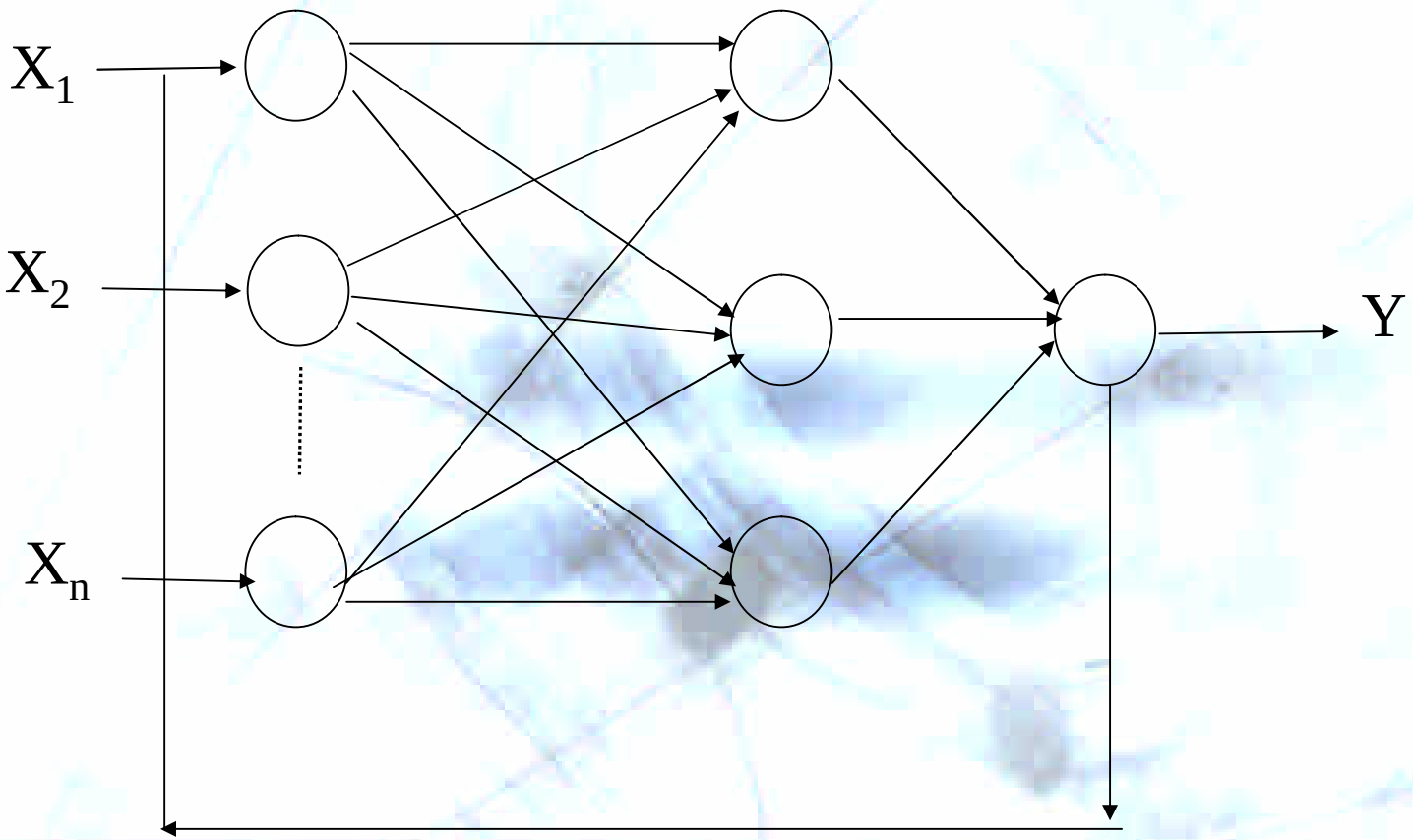
Feedforward (one-way)

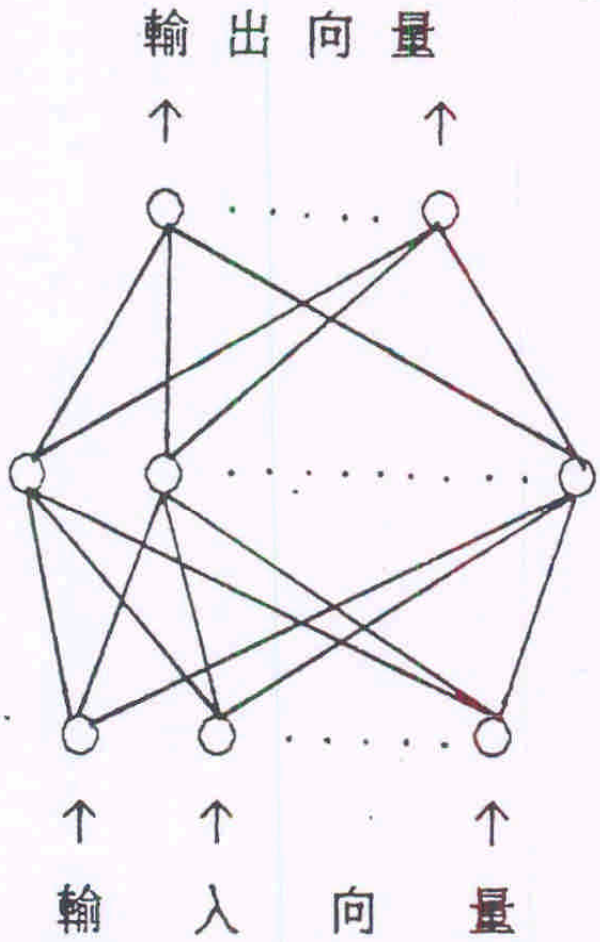


Feedforward (two-way)

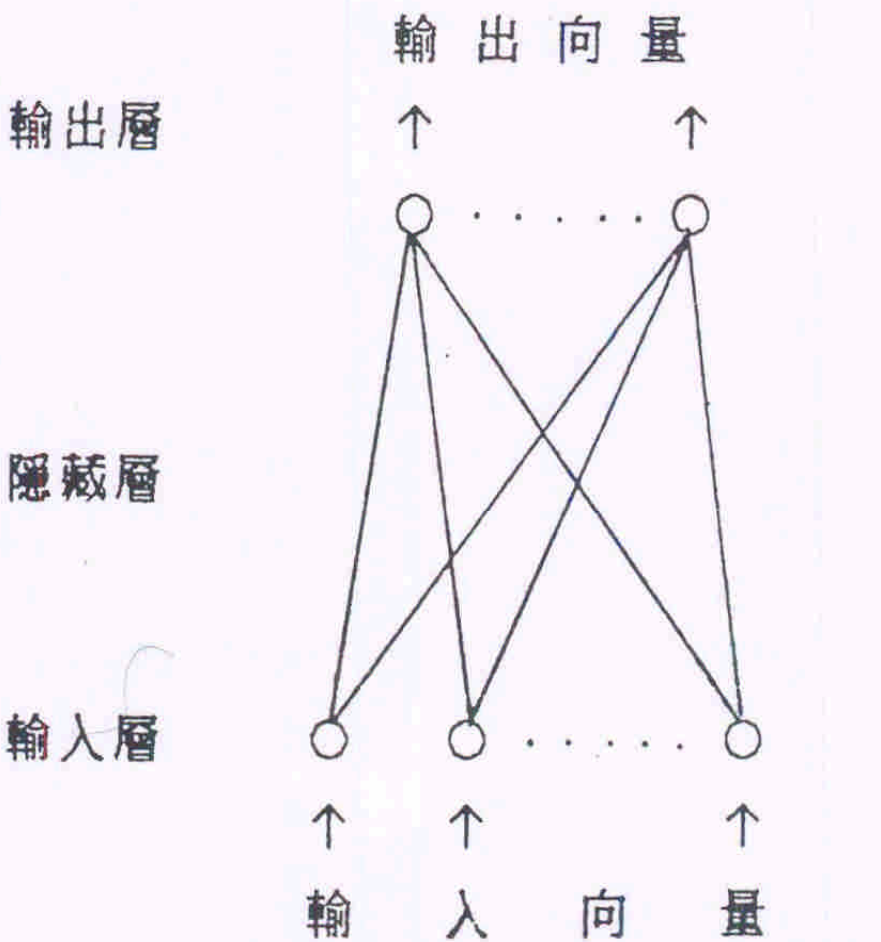


Feedback (one-way)





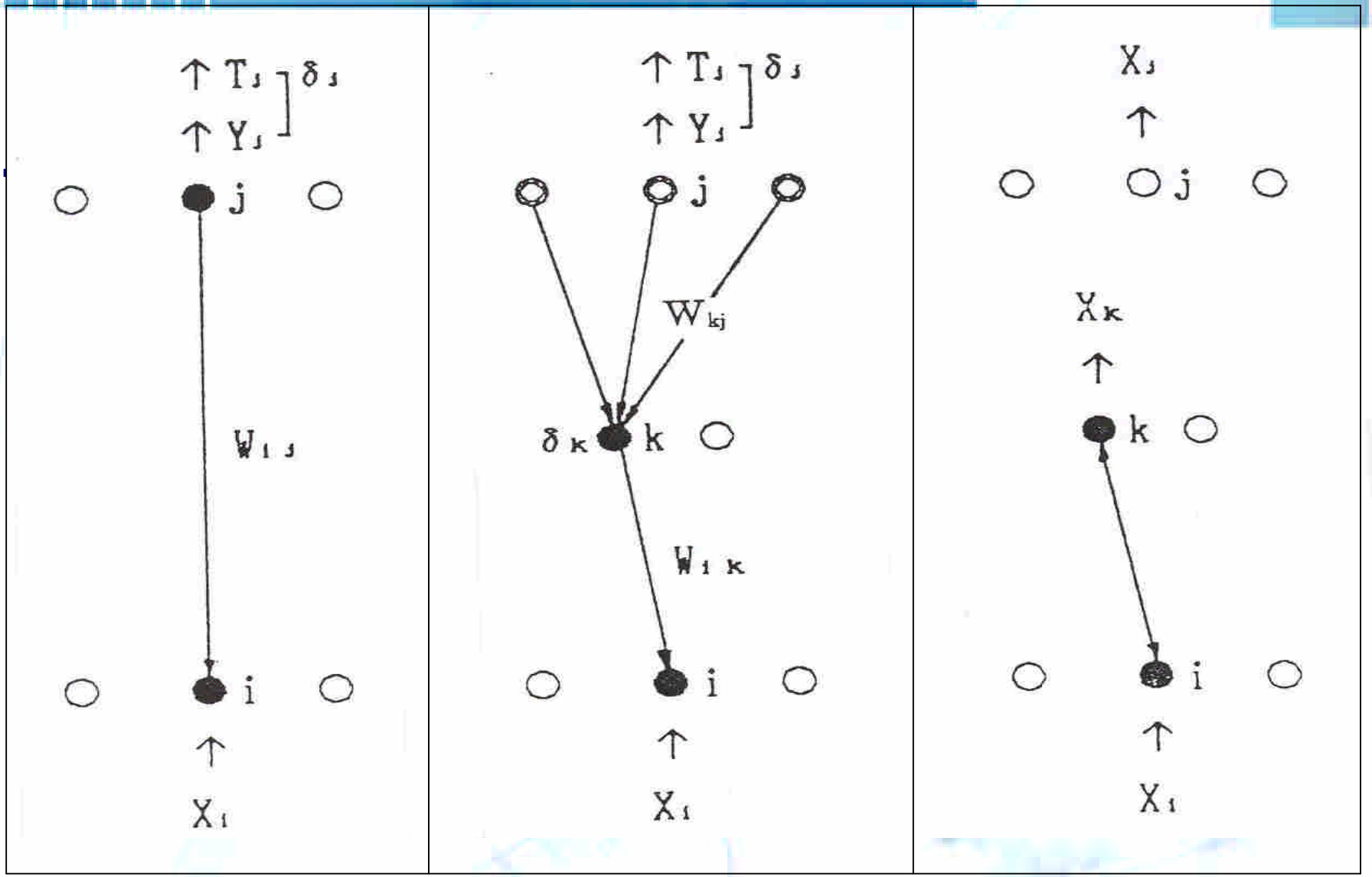
(1) 有隱藏層的感知機架構



(2) 無隱藏層的感知機架構

監督式 **Supervised**

非監督式 **Unsupervised**



(a)差距法則　(b)通用差距法則　(c)波茲曼機學習法則

Learning Type of Supervised Network

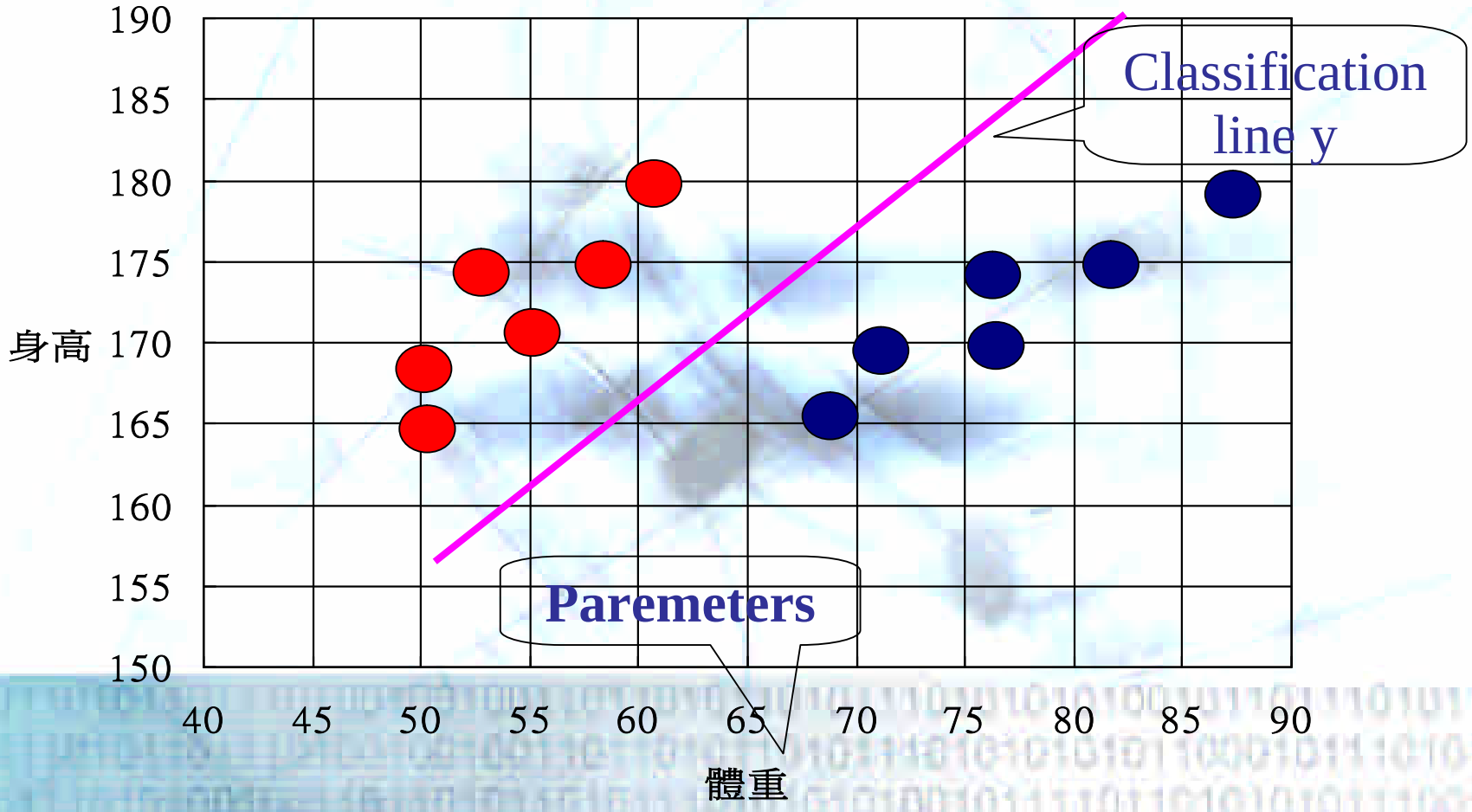
Problem Solving Area

- Classification
- Clustering
- Prediction
- Memorizing
- Learning

- **Optimization**
- **Control**
- **Recognition**
- **Decision-making**

Classification problem

EX: A simple case of classification problem

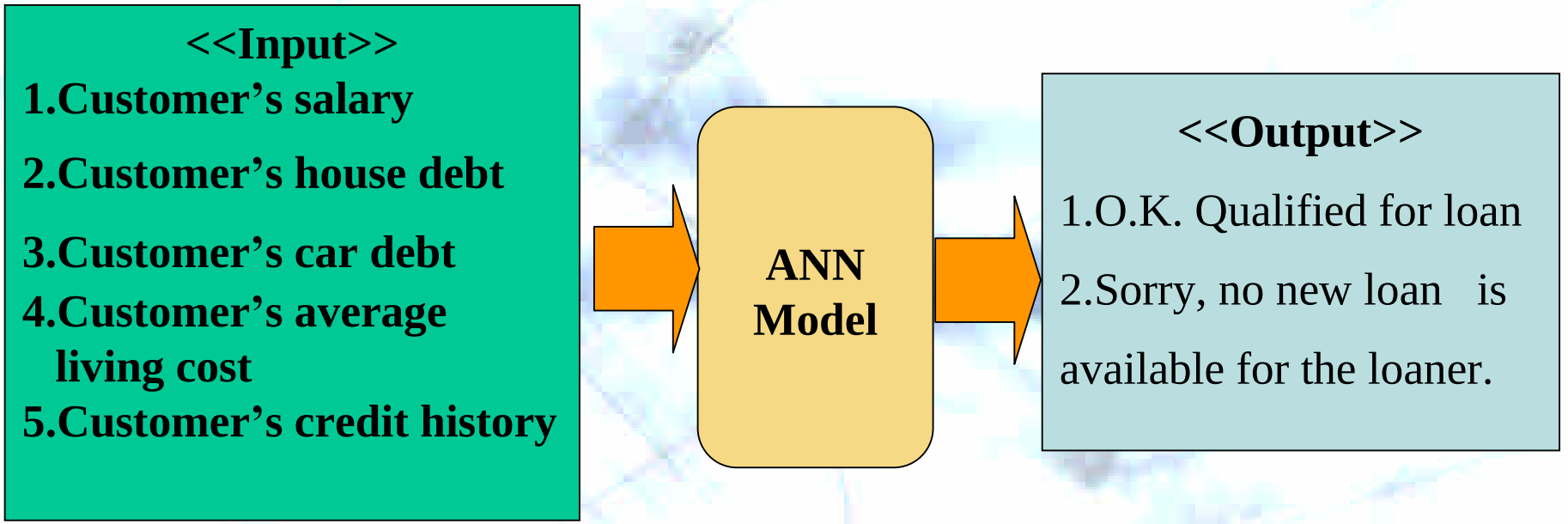


Application Area

- 信用評估
- 管制圖判讀
- 生產製程變數預測
- 顧客篩選
- 銷售預測
- 製程監控
- 石油探勘
- 汽車診斷
- 工廠排程
- 資源配置
- 股價預測
- 投資決策
- 稅務稽查
- 代款審核
- 債券分級
- 醫學診斷
- 氣象預測
- 儀器分析
- 目標追蹤
- 電腦音樂
- 專家系統

Example using ANN model

- The credit prediction for a customer who applys for loan.



The comparison of ANN with Regression(1/2)

- **Variable prediction vs. Regression Analysis**

$$Y = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + \ldots + a_n X_n$$

  - **For regression, the main task is to find out the parameters** $a_0, a_1, a_2, a_3, \ldots a_n$. Therefore, Regression can be used to do the classification or prediction.
  - However, if the problem is belongs to non-linear type, then it will become difficult to solve. ANN is good for the nonlinear problem, especially when the problem is very complex.

The comparison of ANN with Regression(2/2)

- **ANN vs. Regression→ The ANN advantages:**
  - **Can solve non-linear problem**
  - **Parameters can be modified easily**
  - **Easy to construct the model**
  - **Accepts any type of input**
- **ANN vs. Regression → The disadvantage:**
  - **Takes time to find the global minimum(the best solution)**
  - **May be over learning**
  - **Accepts any type of input**

The comparison of ANN with Time Series

- **Time Series**

$$X_t = a_0 + a_1 X_{t-1} + a_2 X_{t-2} + a_3 X_{t-3} + \ldots + a_p X_{t-p}$$

  - **Based on the (time) history values to predict future results.**
  - **EX:**
    - **prediction of stock market,**
    - **prediction of volcano eruption**

The comparison of ANN with Decision Making

- **Decision Making**

$$f_1 = f_1(X_1, X_2, X_3, \ldots X_n) = a_{10} + a_{11}X_1 + a_{12}X_2 + \ldots + a_{1n}X_n$$
$$f_2 = f_2(X_1, X_2, X_3, \ldots X_n) = a_{20} + a_{21}X_1 + a_{22}X_2 + \ldots + a_{2n}X_n$$
$$\vdots$$
$$f_m = f_m(X_1, X_2, X_3, \ldots X_n) = a_{10} + a_{m1}X_1 + a_{m2}X_2 + \ldots + a_{mn}X_n$$

  - By Applying the same inputs to find out which f_i has the best outcome. The decision is made based on the best outcome.
  - EX: Credit evaluation, Scheduling, Strategic decision

類神經網路模式列表(1/3)

分類	項目 / 模式	主要研發者	研發年代	主要應用	主要特點
監督式	感知機	F. Rosenblatt	1957	•字母識別 •目前已多改為multilayer network做應用	•無法解XOR問題 •模式最簡單 •發展最早
	倒傳遞網路	P. Werbos D. Parker D. umelhart	1974-1985	樣本識別 分類問題 函數合成 適應控制 雜訊過濾 資料壓縮 專家系統	•可解XOR問題 •應用最普遍 •成功案例多 •學習精度高 •學習速度慢 •回想速度快 •理論簡明

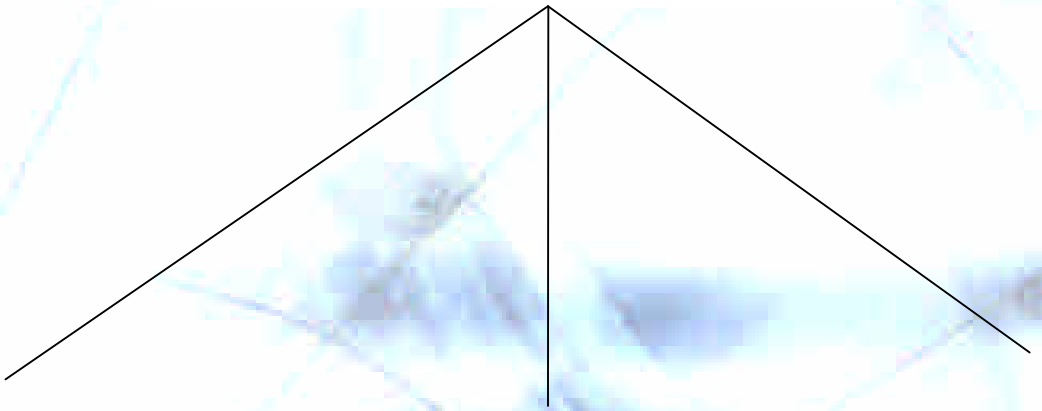
類神經網路模式列表(2/3)

分類	項目 模式	主要研發者	研發 年代	主要應用	主要特點
監督式	機率神經網路	D. F. Specht	1988	•樣本識別 分類問題	•學習速度快 •回想速度慢 •理論簡明
	學習向量量化網路	T. Kohonen	1988	•樣本識別 •分類問題	•學習速度快 •回想速度快 •理論簡明

類神經網路模式列表(3/3)

分類	項目 模式	主要研發者	研發年代	主要應用	主要特點
聯想式	自組織映射圖	T. Kohonen	1980	•聚類問題 •拓撲映射	•具有臨近區域觀念 •學習速度快
	自適應共振理論	G. A. Carpenter S. Grossberg	1976-1986	•樣本識別 •聚類問題	•網路具穩定性 •網路具可塑性 •學習速度快 •具警戒值觀念
最適式	霍普菲爾坦克網路	J. Hopfield D. Tank	1985	•組合最適化問題	•較不易陷入局部最小值
	退火神經網路	D. E. Ven den Bout T. K. Miller	1988	•組合最適化問題	

Typical Learning Methods for Basic Learning Strategies



Supervised

Delta Rule
Backpropagation
Hebbian
Stochastic

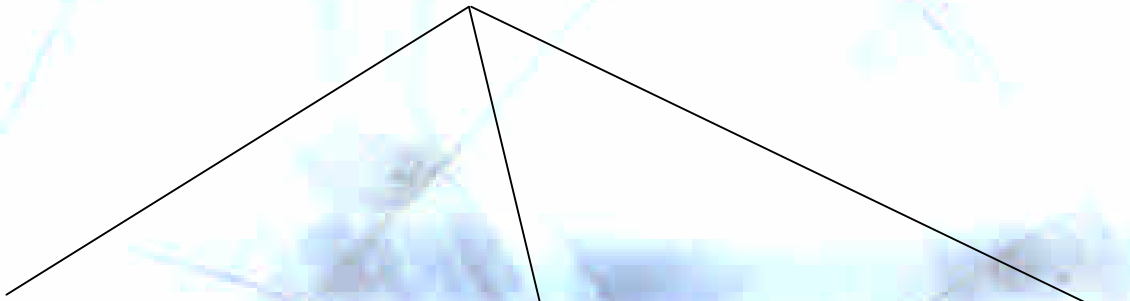
Reinforcement

Learning Automata

Unsupervised

Competitive
Hebbian

Categories of Network Types by Broad Learning Method



Supervised

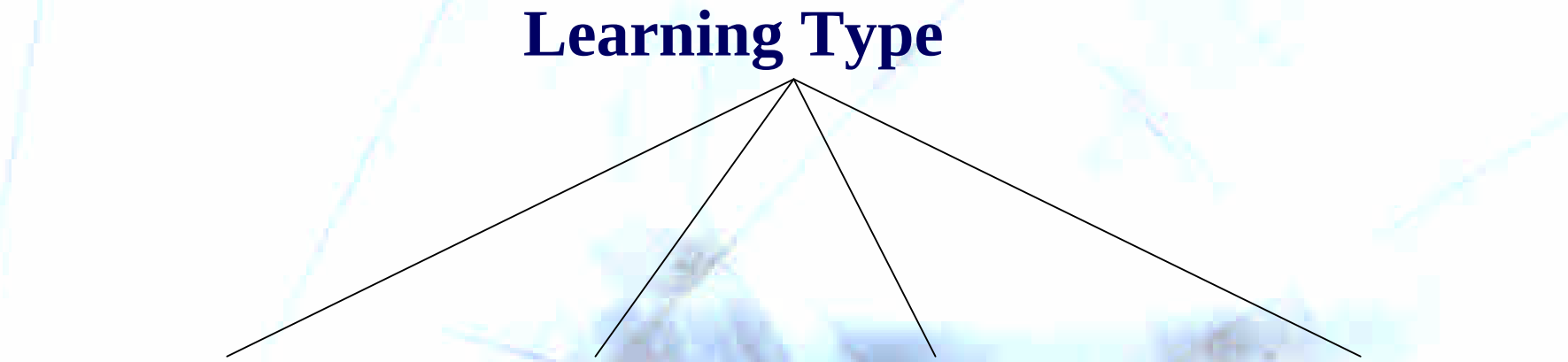
ADALINE
Boltzmann
Cascade Correlation
GRNN
Hopfield
LVQ
MLFF with BP
PNN
RBF
RNN
RCE

Reinforcement

Unsupervised

ART
Hopfield
LVQ
Neocognitron
SOFM

Categories of Network Types by Learning Type



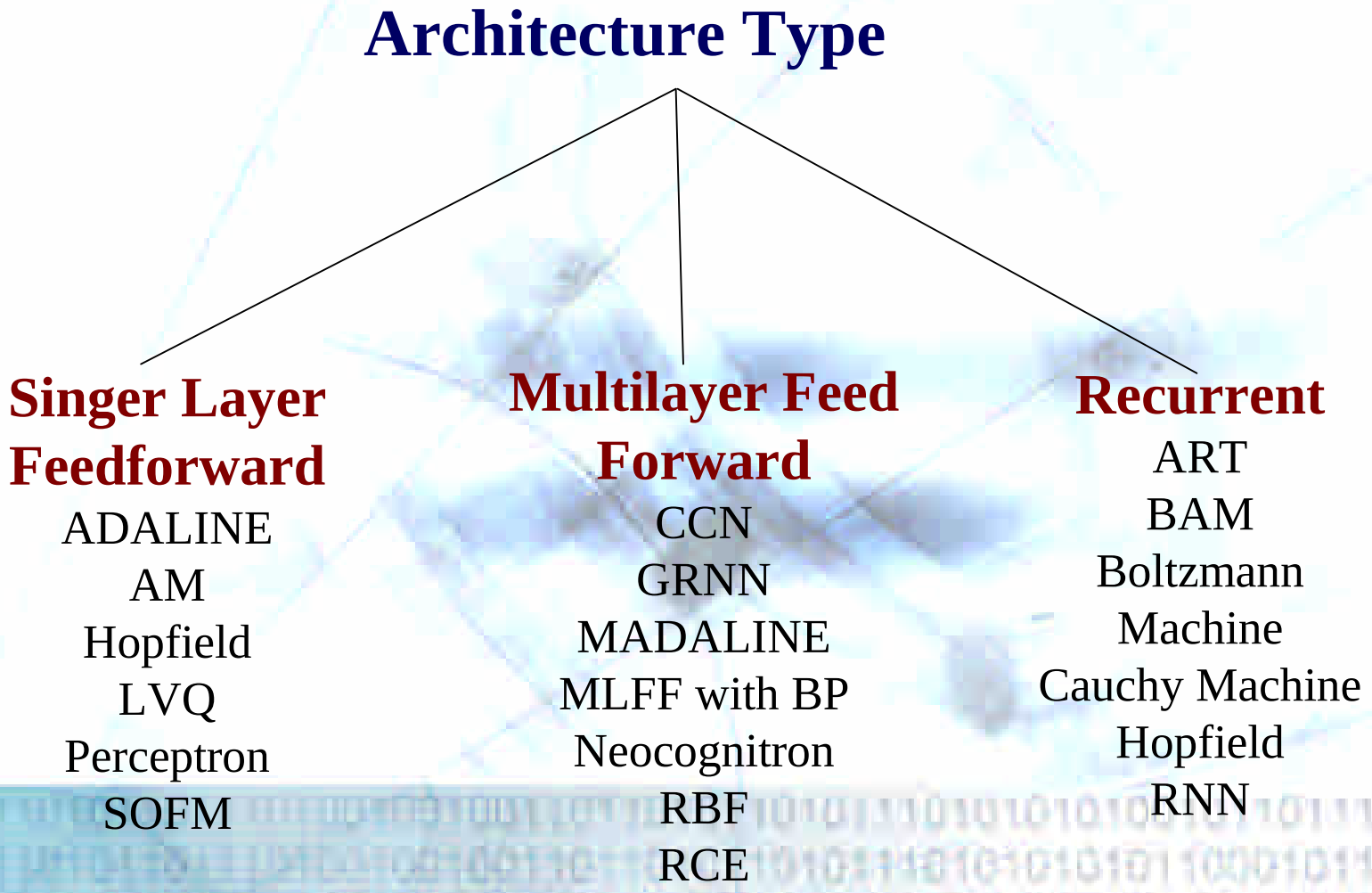
Error Correction
ADALINE
CCN
GRNN
Hopfield
MLFF with BP
Perceptron
RBF
RNN

Hebbian
AM
BSB
BAM
Hopfield
Neocognitron

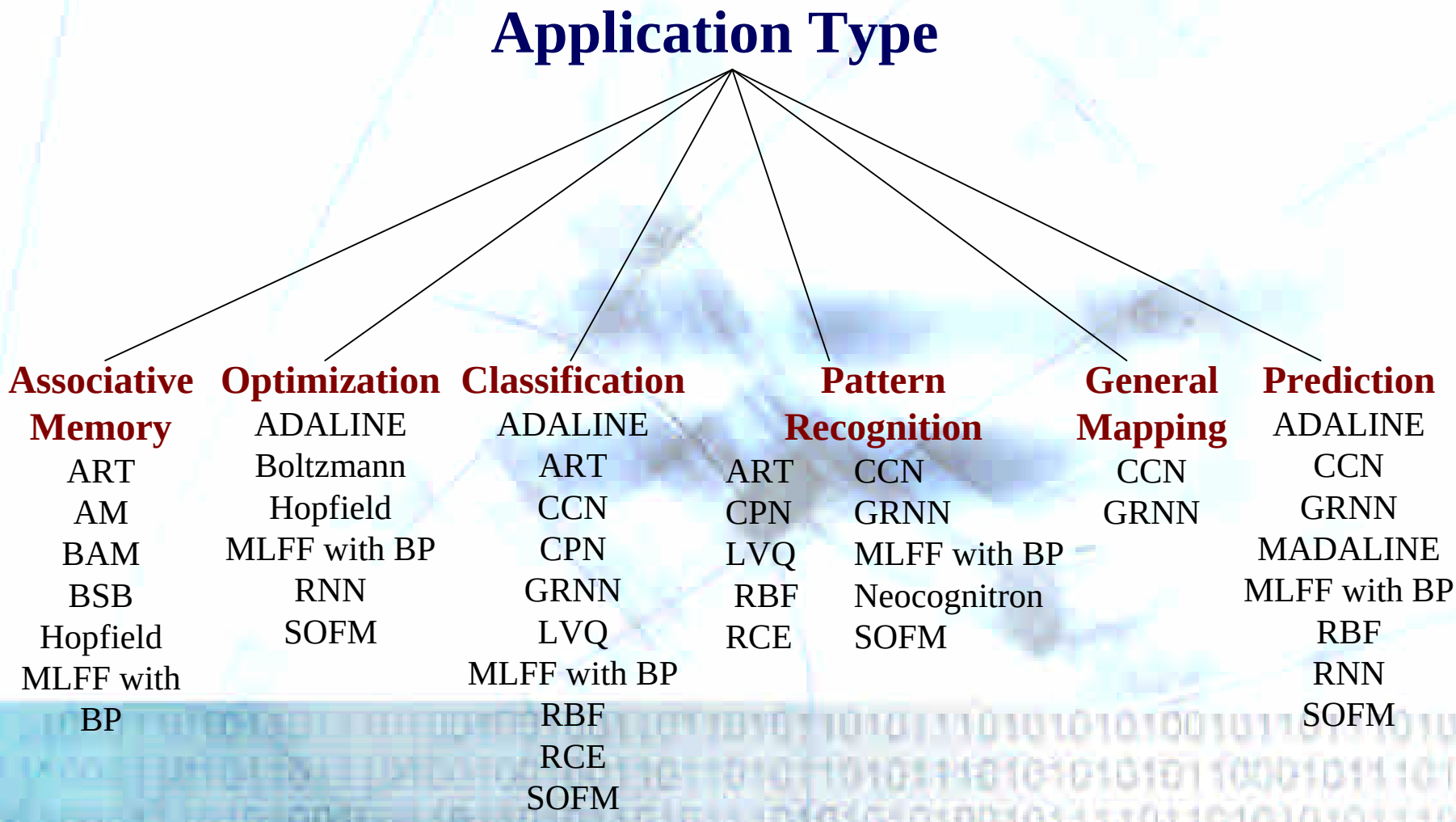
Competitive
ART
CPN
LVQ
SOFM

Stochastic
Boltzmann Machine
Cauchy Machine

Categories of Network Types by Architectural Type



Categories of Network Types by Application Type



Neural Network Taxonomies

- Perceptron
- Hopfield
- ADALINE (Adaptive Linear Neural Element)
- MADALINE (Multilayer ADALINE)
- BPN (Back Propagation Network)
- ART (Adaptive Resonant Theory)
- AM (Associative Memories)
- BAM (Bidirectional Associative Memory)
- Boltzmann Machine
- CCN (Cascade Correlation)
- CPN (Counter Propagation Network)
- GRNN (Generalized Regression Neural Network)

Neural Network Taxonomies

- LVQ (Learning Vector Quantization)
- MLFF with BP (Multilayer Feedforward Backpropagation)
- PNN (Probabilistic Neural Network)
- RBF (Radial Basic Function)
- RNN (Recurrent Neural Networks)
- SOM (Self-Organizing Map) or SOFM (Self-Organizing Feature Map)
- NLN (Neurologic Networks)