附錄 B 類神經網路

● 類神經網路簡介

類神經網路是一個可計算之多層網路，它使用大量簡單而具有平行處理能力之人工神經元來模擬人類學習行為，因此本章節將針對生物及人工神經元的構造、類神經網路的組成架構和基本理論，以及類神經網路中之倒傳遞網路(Backpropagation Network)詳加介紹。

● 生物神經元模型

生物神經網路是由巨量的神經細胞(neuron，又稱神經元)所組成，形成一個高度連結網狀的神經網路，資訊的處理工作即透過上述之連結來進行。以人腦而言，人腦大約由 10^{11} 個神經元所組成，而每一個神經元約有 10^{3} 根連結與其他神經元相連，所以人腦中約有 10^{14} 根連結，因此人腦可以儲存大量而複雜的知識。神經元構造如圖附 B.1 所示，其主要構造如下：

1.神經核(soma):神經細胞的核心，為一呈核狀的處理機構。
2.神經軸(axon):神經細胞呈軸索狀的輸送機構。
3.神經樹(dendrites):神經細胞呈樹枝狀的輸出入機構。
4.神經節(synapse):神經細胞神經樹上呈點狀的連結機構。

當神經細胞透過神經節與神經樹從其它神經元輸入脈波訊號後，經神經核處理，產生一個新的脈波訊號，這個訊號再經過神經軸傳送到神經樹，再透過神經節與神經樹成為其它神經元的輸入脈波訊

號，如果脈波訊號是經過興奮神經節(excitatory synapse)，則會增加脈波訊號的速率(pulse rate)，如果脈波訊號是經過抑制神經節(inhibitory synapse)，則會減少脈波訊號的速率。因此，脈波訊號的速率是同時取決於輸入脈波訊號的速率，以及神經節的強度。而神經節的強度可視為神經網路儲存資訊之所在，神經網路的學習即在調整神經節的強度(葉怡成，1994)。

- 人工神經元模型

根據神經細胞的結構與功能，從 40 年代開始，先後提出的神經元模型有數百種之多，其中對於腦模型、自動機、人工智慧有重大影響的是 1943 年由美國心理學家 McCulloch 和數學家 Pitts 共同提出的形式神經元模型，同常稱之為 MP 模型。(謝明富，1999)

此一模型（見圖附 B.2 所示）具有將輸入變數與輸出變數間，複雜的內在對映關係充份呈現的功能，其為解決非線性動態問題的最佳工具之一。此模型由許多人工神經元所組成，神經元又稱為節點(node)或臨界值元件(threshold element)，其輸入端輸入各自之訊息，藉由各自權重加權總和後傳入節點，透過閥值的過濾，繼而經由轉換函數轉換後輸出，其數學表示式如下：

$$Y_j = f\left(net_j\right)，net_j = \sum_i W_{ij} X_i - b_j \qquad i = 1,2,\ldots,n \quad \ldots\ldots\ldots（附 B.1）$$

其中：

Y_j:為模仿生物神經元模型的**輸出訊號**。

f:為模仿生物神經元模型的**轉換函數**(transfer function)，將

輸入值之加權乘積和轉換成處理單元輸出值。

W_{ij}:為模仿生物神經元模型的神經節強度，又稱連結加權值。

X_i:為模仿生物神經元模型的輸入訊號。

b_j:為模仿生物神經元模型的閥值(bias)。

net_j:為輸入值之加權乘積和。

n：為輸入訊號個數

類神經網路常用之轉換函數有下列四種（見圖附 B.3 至附 B.6(b)）：

1. 位階臨界轉換函數(Step Threshold Transfer Function)：臨界函數的輸出只隨輸入值的正副號所改變。
2. 線性轉換函數(Linear Transfer Function)：函數輸入值與輸出值呈線性關係。
3. 非線性轉換函數(Nonlinear Transfer Function)：函數輸出的最大值與最小值限制在一個特定的範圍內。
4. 臨界 S 型轉換函數(Sigmoid Threshold Transfer Function)：是最常被應用之函數，因為此種函數型態具有可微分且連續等性質，此特質使網路可以應用到非線性的學習領域中。常用函數有雙彎曲函數(式附 B.2)與雙曲線正切函數(式附 B.3)兩種。

$$f(x)=\frac{1}{1+e^{-x}} \quad \text{（附 B.2）}$$

$$f(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \quad \text{（附 B.3）}$$

● 類神經網路基本理論

類神經網路的架構乃仿照人類腦部神經網路的模型而發展出來的，一個類神經網路是由許多個人工神經元連結組成，並可以組成各種網路模式。(葉怡成，1994)

類神經網路的總體運作模式有兩種：(1)學習過程(Learning)－網路依學習演算法，從範例中學習，以調整網路權重值的過程，與(2)回想過程(Recalling)－網路依學習演算法，以輸入資料決定網路輸出資料的過程。

此外，學習過程的學習演算法又可分為三類：

(1)監督式學習

從問題領域中取得訓練範例(有輸入變數值，也有輸出變數)，並從中學習輸入變數與輸出變數的內在對映規則，以應用於新的案例(只有輸入變數值，而需推論輸出變數值的應用)。感知機網路、倒傳遞網路、機率神經網路、學習向量量化網路與反傳遞網路等五種類神經網路皆屬監督式學習。

(2)無監督式學習

從問題領域中取得訓練範例(只有輸入變數值)，並從中學習範例的內在集群規則，以應用於新的案例(有輸入變數值，而需推論它與那些訓練範例屬同一集群的應用)。自組織映射圖網路、自適應共振理論網路等兩種類神經網路皆屬無監督式學習。

(3)聯想式學習

從問題領域中取得訓練範例(狀態變數值),並從中學習範例的內在記憶規則，以應用於新的案例(只有不完整的狀態變數值，而需推論其完整的狀態變數值的應用)。霍普菲爾網路、雙向聯想記憶網路等兩種類神經網路皆屬聯想式學習。

● 倒傳遞神經網路（Bp）

在監督式學習模式中，由 Rumelhart 等學者於 1986 年所發展的誤差向後推導或稱倒傳遞學習演算法（簡稱 BP）是被廣泛使用的一種學習演算法，因為其具有學習及回想的功能，故可進行定率預測。一般倒傳遞網路可分為三部份(如圖附 B.7 所示),輸入層用以接受外在環境的訊息，其神經元數目則依問題而定；輸出層用以輸出訊息給外在環境，其神經元數目同樣依問題而定；隱藏層將輸入與輸出層各處理單元間的相互關係充份地表現出來，其神經元數目並無標準可決定。倒傳遞網路模式學習訓練方式由所探討問題中取得相當數量之訓練樣本，並從樣本中應用向前餽入與誤差向後推導兩步驟推求輸入變數與輸出變數的內在對映規則，再應用回想功能，進行新案例之輸出變數值推估。下列為倒傳遞網路（Bp）之學習與回想過程建立步驟：

一、倒傳遞網路學習過程的建立:

(1) 令輸入層、隱藏層與輸出層節點分別以 i、j、k 為下標符號，建立一輸入層、隱藏層與輸出層節點數分別為 m、n、o 個，以均勻分佈隨機亂數設定加權值矩陣 W_{ij}、W_{jk}與閥值 b_j、b_k的初始值。

(2) 輸入一個訓練案例的輸入向量 $\vec{X}=[x_i]\quad i=1,2,....m$ 與目標輸出向量 $\vec{T}=[T_k]\quad k=1,2,....o$ 。

(3)計算推估輸出向量 $\vec{O}$ 。

(a)輸入層向量

$\vec{X}=[x_i]\quad i=1,2,....m$

(b)計算隱藏層輸出向量 $\vec{H}$

$$H_j = f_h(net_j)\text{，} net_j=\sum_{i=1}^{m} x_i w_{ij} - b_j \qquad j=1,2,3.....n \text{.......（附 B.4）}$$

其中 $f_h(x)$ 為隱藏層轉換函數，可依問題型態挑選適當轉換函數。

(c)計算輸出層輸出向量 $\vec{O}$

$$O_k = f_o(net_k)\text{，} net_k=\sum_{j=1}^{n} H_j w_{jk} - b_k \qquad k=1,2,3.....o \text{......（附 B.5）}$$

其中 $f_o(x)$ 為輸出層轉換函數，可依問題型態挑選適當轉換函數。

(4)計算加權值矩陣修正量 Δw，及閥值修正量 Δb 。

因為監督式學習的目的在降低網路輸出單元目標輸出值與推論輸出值之差距，所以一般以能量函數（又稱誤差函數）表示學習的品質：

$$E=\frac{1}{2}\sum_{k=1}^{o}(O_k - T_k)^2 \text{ ....（附 B.6）}$$

因此網路的學習過程即為使能量函數最小化的過程，通常以最陡坡降法來使能量函數最小化，即每輸入一個訓練案例，網路即小幅調整加權值的大小，調整的幅度和誤差函數對該加權值的敏感程度成正比，即與誤差函數對加權值的偏微分值大小成正比：

$$\Delta w = -\eta \frac{\partial E}{\partial w} \quad \text{(附 B.7)}$$

式中：

Δw：加權值調整的幅度。

η：學習速率(learning rate)，控制每次以最陡坡降法最小化誤差函數的步幅，其範圍為0≤η≤1，可視所需狀況自行設定範圍容許內之值。

經由一連串的代入及演算可推得以下各值：

(a)輸出層

$$\Delta w_{jk} = \eta (T_k - O_k) \cdot df_o(net_k) \cdot H_j \quad \text{(附 B.8)}$$

$$\Delta b_k = -\eta (T_k - O_k) \cdot df_o(net_k) \quad \text{(附 B.9)}$$

其中 $df_o(x)$ 為輸出層轉換函數之一階導函數，舉例說明：若以雙彎曲函數 $f_o(x) = \frac{1}{1+e^{-x}}$ 為輸出層之轉換函數，其一階導函數則為 $df_o(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ 。

(b)隱藏層

$$\Delta w_{ij} = \eta (T_k - O_k) \cdot df_o(net_k) \cdot w_{jk} \cdot df_h(net_j) \cdot x_i \quad \text{(附 B.10)}$$

$$\Delta b_j = -\eta(T_k - O_k)\cdot df_o(net_k)\cdot w_{jk}\cdot df_h(net_j) \quad \text{(附 B.11)}$$

其中 $df_h(x)$ 為隱藏層轉換函數之一階導函數。

(5)更新加權值矩陣與閥值

$$w_{new} = w_{old} + \Delta w \quad \text{(附 B.12)}$$

$$b_{new} = b_{old} + \Delta b \quad \text{(附 B.13)}$$

(6)重覆步驟 2 至步驟 6，直至收斂，即誤差不再有明顯變化且符合所要求的精度。

二、倒傳遞網路回想過程的建立:

(1)讀入加權值矩陣與閥值矩陣。

(2)輸入一個欲作檢定或驗證的案例之輸入向量 $\vec{X} = [x_i] \quad i = 1,2,....m$。

(3)計算推估輸出向量 $\vec{O}$。

(a)輸入層向量

$$\vec{X} = [x_i] \quad i = 1,2,....m$$

(b)計算隱藏層輸出向量 $\vec{H}$

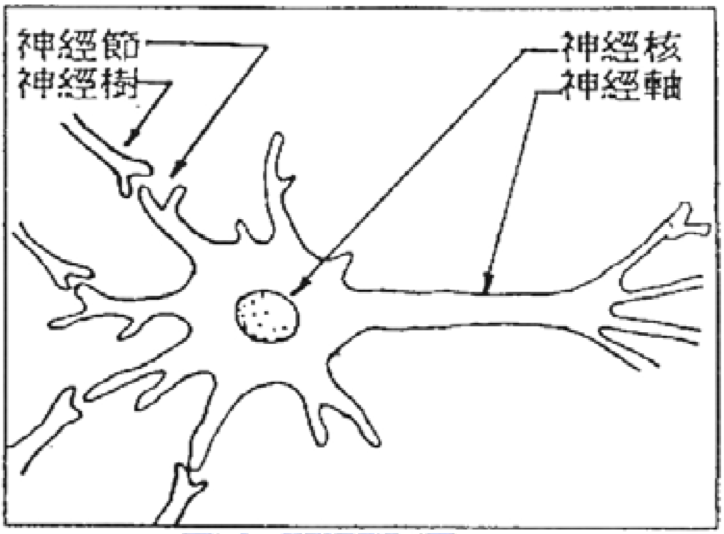
$$H_j = f_h(net_j)\text{，}net_j = \sum_{i=1}^{m} x_i w_{ij} - b_j \quad j = 1,2,3.....n \quad \text{(附 B.4)}$$

其中 $f_h(x)$ 為隱藏層轉換函數，可依問題型態挑選適當轉換函數。

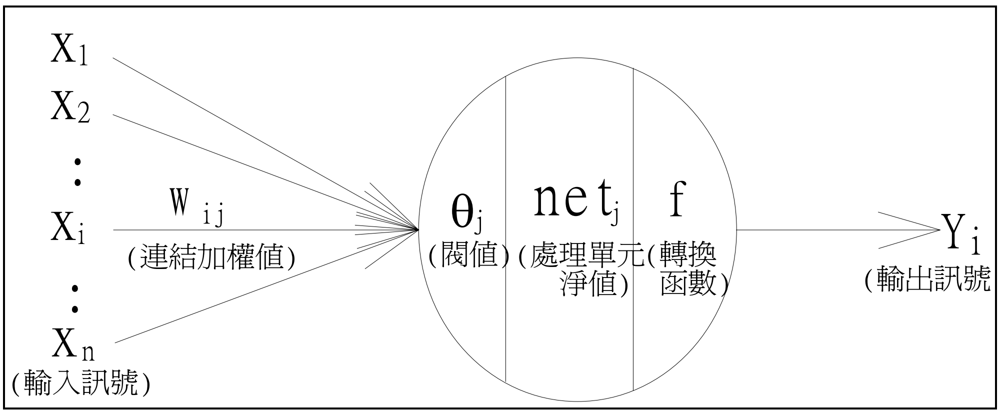
(c)計算輸出層輸出向量 $\vec{O}$

$$O_k = f_o(net_k) \text{，} net_k = \sum_{j=1}^{n} H_j w_{jk} - b_k \qquad k = 1,2,3.....o \text{......（附 B.5）}$$

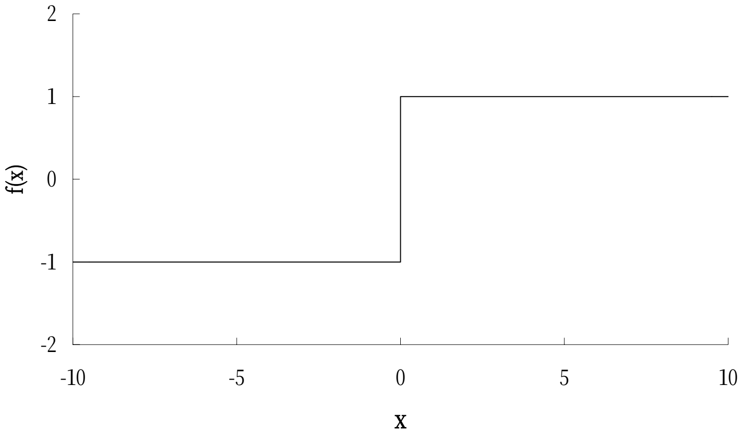
其中 $f_o(x)$ 為輸出層轉換函數，可依問題型態挑選適當轉換函數。



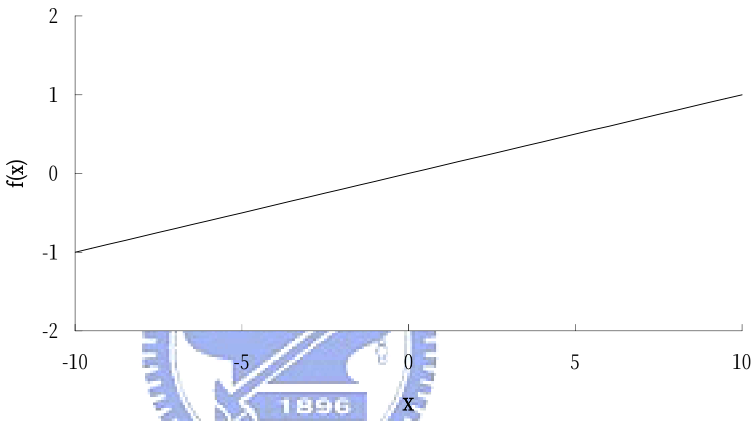
圖附 B.1 生物神經元模型（葉怡成，1993）



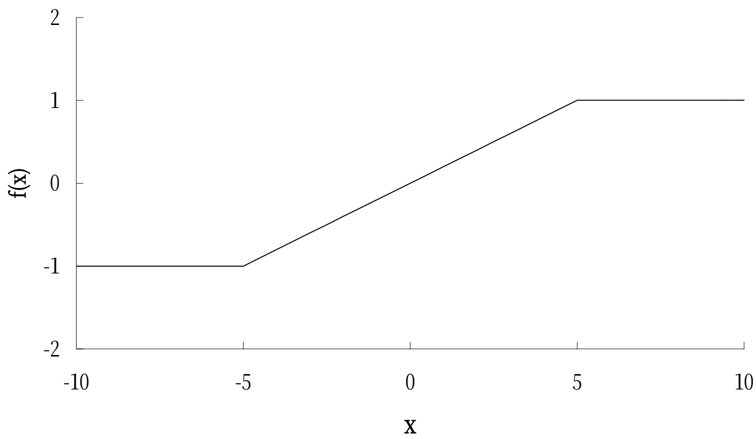
圖附 B.2 人工神經元模型



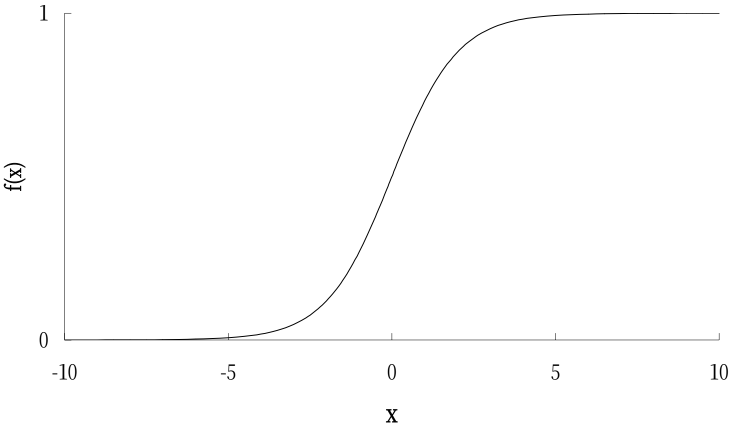
圖附 B.3 位階臨界轉換函數圖



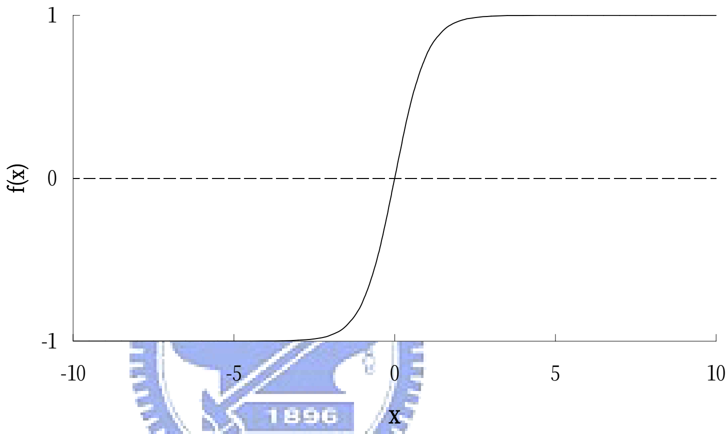
圖附 B.4 線性轉換函數圖



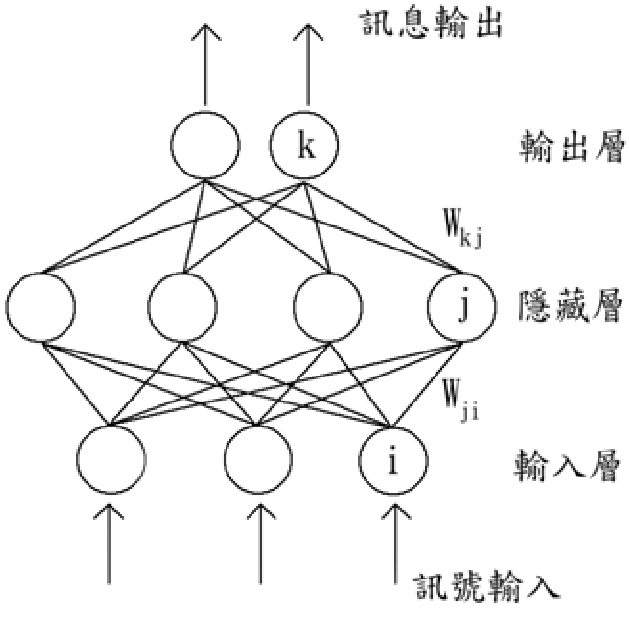
圖附 B.5 非線性轉換函數圖



圖附 B.6(a) 雙彎曲轉換函數圖



圖附 B.6(b) 雙曲線正切轉換函數圖



圖附 B.7 倒傳遞網路模型