

MC-Transaction on Biotechnology, 2012, Vol. 4, No. 1, e4

©This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

基於二維小腦模型設計心臟病分類系統

李嘉陵^{1,*}、蔡坤龍²

¹ 銘傳大學健康科技學院生物醫學工程學系(中華民國 台灣 桃園)

² 銘傳大學資訊學院電子工程學系(中華民國 台灣 桃園)

中文摘要

本論文針對葡萄牙的心臟病資料庫，將設計二維小腦模型分類系統以便將病患快速地分出無病或有病之類別。小腦模型分類系統的設計程序包含：(一)隨機選取訓練樣本及測試樣本，(二)規劃二維小腦模型分類系統，(三)訓練二維小腦模型分類系統，(四)測試二維小腦模型分類系統。本論文以區塊交集法可快速建立位址索引矩陣，並以座標計算方式可連結狀態單元及其位址索引。我們嘗試利用兩組不同的心臟病屬性分別作為小腦模型的輸入變數，並採用平均輸出誤差的學習法則來調整小腦模型的記憶細胞權重。我們將採用的二維小腦模型可簡化分類系統的設計，並快速收斂到輸出誤差小於 1% 以內。在測試試驗中，我們找出了兩個與心臟病相關程度較高的屬性作為輸入變數，可使得整體分類正確率達到 81.5% 以及有病捕捉率達到 96.2% 的結果，由此足以證明本分類系統僅以二維小腦模型為架構的有效性，而且，採用相關且病症傾向相同的屬性會有較佳的分類效果。

關鍵字：心臟病症、小腦模型、記憶細胞、分類系統

通訊作者：李嘉陵 [jllee@mail.mcu.edu.tw]

收稿：2012-8-3 修正：2012-8-30 接受：2012-8-31

緒 論

隨著我國社會的現代化、飲食的西式化以及老年人口的增加，造成以心血管所衍生的心臟疾病將成為我國重要的醫療問題。根據衛生署的統計資料顯示，目前國人十大死因之中，心臟疾病已高居第二位。然而目前尚無一種檢查可以快速和正確地評估心臟功能^[1]，因此本論文將嘗試以小腦模型(Cerebellar Model)架構設計心臟病分類系統，以便將病患快速地分出有病或無病之類別，作為醫師診斷的參考，本系統對於國人健康的促進必定有所幫助。關於小腦模型的發展過程如下所

述：Marr於1969年首先提出小腦皮質(Cerebellar Cortex)的生物模型理論^[2]，此模型由內而外依序分別為分子層(Molecular Layer)、細胞層(Purkinje Cell Layer)及粒狀層(Granular Layer)等三層架構，能分層儲存來自於大腦及身體各個感覺器官所傳入的神經訊息。在1975年，Albus依據Marr的小腦生物模型而發展出一套類似小腦皮質運作的小腦數學模型^[3]，能分層儲存所輸入的訊息，和輸出所得訊息，並具備學習功能的類神經網路，經由重複的學習過程來更新各層的記憶細胞權重，使得達到符合期望的輸出，因此小腦模型基本上是屬於記憶體查表法(Lookup Table)的類神經網路(Artificial Neural Network)架構，此特性將使小腦模型具有以下之優點：(一)學習收斂速度快：在小腦模型的架構之中，與學習相關的是規劃儲存訊息的記憶體細胞，使得可以利用小腦模型的實際輸出與期望的輸出所得的誤差回饋，來調整記憶體中各層次的權重，而建立起輸入與輸出的關係，因為在每一次的訓練過程僅需改變局部的記憶體權重，所以收斂速度較一般類神經網路迅速。(二)具有區域概括性(Local Generalization)：每兩個相鄰參考狀態至少共同對映一個以上的記憶單元，亦即每一個狀態至少對映兩個以上的記憶細胞，可分散式儲存訊息。(三)可應用於非線性系統的近似或識別：相較於其他的類神經網路，小腦模型的構造簡單且易於運算處理，並可快速學習許多非線性函數，而且一定收斂於最小平方根誤差^[4]。(四)容易將小腦模型演算法實現在硬體上。由於具備上述之優點，所以近年來普遍獲得各領域研究人員的重視，例如應用在馬達控制^[5]，數位訊號通道等領域^[6]。本論文將以葡萄牙的心臟病資料庫為例，設計二維小腦模型(Cerebellar Model)心臟病分類系統，僅利用資料庫的部分數值屬性資料為輸入，來調整小腦模型各層次的記憶細胞權重，以利找出心臟病資料庫中病症之類別。

材料與方法

(一)材料：資料庫及小腦模型

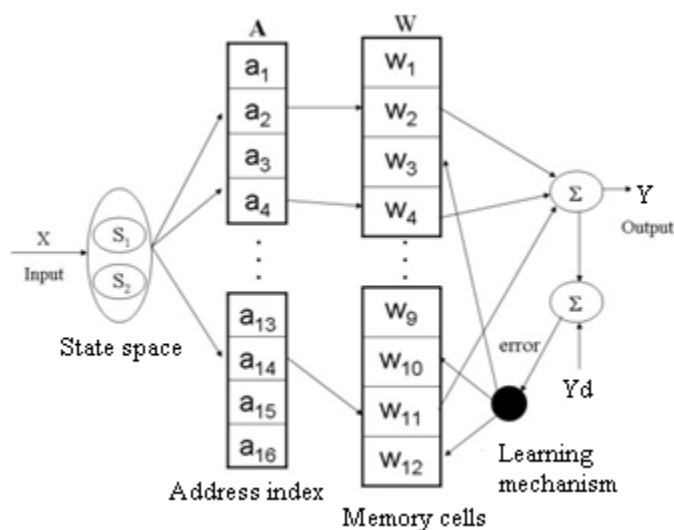
本論文的研究對象是由葡萄牙波多大學(University of Porto)的人工智慧及電腦科學實驗室所提供的心臟病資料庫的病患資料^[7]。此資料庫具有1個目標屬性與13個條件屬性，其中包含7個類別屬性及6個數值屬性，共有270筆資料。該資料庫可區分為患有心臟病(130位)及無心臟病(140位)等兩大類別。本論文僅使用數值屬性，其屬性描述如表一所示。

本論文將提出的小腦模型心臟病分類系統之架構，如圖一所示：其中X表示小腦模型的輸入陣列，包含兩個輸入變數， (S_1, S_2) 即為輸入變數構成的狀態空間(State Space)，A為狀態空間對應記憶細胞(Memory Cells)的位址索引(Address Index)矩陣，W即為記憶細胞的權重， Y_d 與Y分別是期望輸出及小腦模型實際輸出，而兩者之間的誤差(error)回饋至學習機制(Learning Mechanism)以便調整部分記憶細

胞之權重。我們先假設小腦模型有s個輸入變數，而且每一個輸入變數被分割為p個離散區間，則輸入狀態空間可被分割為 $N = p^s$ 個狀態單元(State Elements)。例如有2個輸入變數，每一輸入值被分割為4個小區間，所以會有16個狀態單元，輸入狀態空間如圖二所示。

表一、數值屬性描述

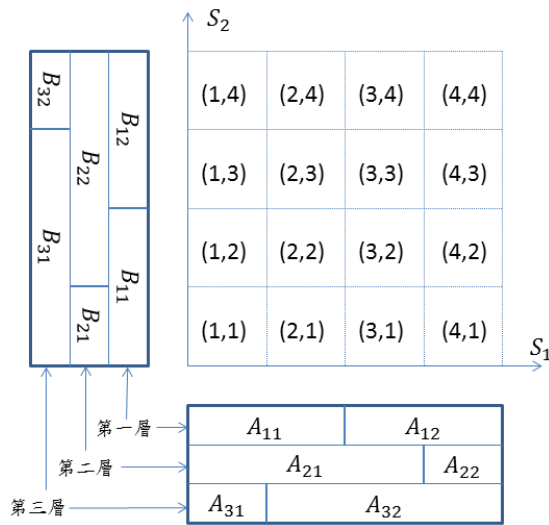
序號	屬性名稱	屬性描述
1	Age	年齡
2	Resting_blood_pressure	收縮血壓
3	Serum_cholesterol	血液中膽固醇含量
4	Maximum_heart_rate_achieved	最大心跳速率
5	Oldpaek	運動心電圖 ST 下降程度
6	Number_of_major_vessels	冠狀動脈被阻塞的數量



圖一、小腦模型架構

在狀態空間中，我們設定以b個狀態單元為一完整區塊(Complete Block)，剩下的則為不完整區塊(Incomplete Block)，假設每一個狀態變數皆可定義m個區塊，包含完整區塊及不完整區塊，此為構成第一層的區塊，然後將第一層的所有區塊向右平移一個狀態單元，在最右邊的一個區塊則轉移至最左邊，則可形成第二層新的區塊型式，以此類推直到第一層區塊型式重複出現為止，假設總共建立了L層的區塊，如圖二中座標軸外側的區塊所示，總共有三層區塊，我們將第一個狀態

變數定義的區塊命名為 A_{ij} ，第二個狀態變數定義的區塊命名為 B_{ij} ，其中 $i=1, 2, 3, \dots, L, j=1, 2, 3, \dots, m$ ，亦即 i 表示第 i 層， j 表示該層的第 j 個區塊，則將不同狀態變數所建立於同一層的區塊分別組合為各層的超立方體(Hypercube) $\Omega = \{A_{ip}B_{iq}, i=1, 2, 3, \dots, L, p, q=1, 2, 3, \dots, m\}$ ，每一個超立方體皆有唯一對應的一個記憶細胞，所以超立方體的個數即為記憶細胞的數目，其數目為 $k = m^2 * L$ ，其中 m 為第一層的區塊數， s 為輸入變數之個數， L 為超立方體的層數。例如有2個輸入狀態變數($s=2$)，每一輸入變數值被分割為4個小區間，因而形成16個狀態單元，以兩個狀態單元為一個區塊，所以在第一層的2個狀態變數可各自形成2個區塊($m=2$)，而且總共可建立3層超立方體($L=3$)，如此則會有12個記憶細胞。接下來須建立每一狀態單元對應到所屬記憶細胞的位址索引矩陣 $A=[a_s]$ ，其中 $a_s=[a_{s1} a_{s2} \dots a_{sk}]^T$ ， $s=1, 2, \dots, N$ ， N 為狀態單元的個數， k 為記憶細胞的個數。當某一狀態單元對應到某些超立方體，則將這些位置記為“1”，其餘位置記為“0”，即可得位址索引矩陣 A ，例如在圖二所示的輸入狀態空間中，與狀態單元座標(1, 1)有對應到的3個超立方體為 $\Omega = \{A_{11}B_{11}, A_{21}B_{21}, A_{31}B_{31}\}$ ，所以在表二中狀態單元(1, 1)與超立方體 $\Omega = \{A_{11}B_{11}, A_{21}B_{21}, A_{31}B_{31}\}$ 的交集位置記為“1”，其餘位置記為“0”，亦即，此狀態單元的位址索引除了 $\{a_{11}, a_{15}, a_{19}\}$ 其值為“1”，其餘皆為“0”，此即為狀態單元(1, 1)的位址索引向量，其餘狀態單元依此原則，可得所有可能輸入的位址索引，如表二所示，其中 Ω 為超立方體的集合。



圖二、輸入狀態空間

至此，每一個輸入狀態可藉由位址索引向量找到對應的記憶細胞內的權重值，因此所有輸入狀態所對應的輸出 Y 可表示為

$$Y = A^T W \tag{1}$$

其中 A 為狀態單元對應記憶細胞的位址索引矩陣， W 即為記憶細胞的權重， T 表

示矩陣的轉置。

表二、狀態單元的位址索引

S/ Ω (座標)	第一層				第二層				第三層			
	A11B11	A11B12	A12B11	A12B12	A21B21	A21B22	A22B21	A22B22	A31B31	A31B32	A32B31	A32B32
S(1) (1,1)	1	0	0	0	1	0	0	0	1	0	0	0
S(2) (1,2)	1	0	0	0	0	1	0	0	1	0	0	0
S(3) (1,3)	0	1	0	0	0	1	0	0	1	0	0	0
S(4) (1,4)	0	1	0	0	0	1	0	0	0	1	0	0
S(5) (2,1)	1	0	0	0	1	0	0	0	0	0	1	0
S(6) (2,2)	1	0	0	0	0	1	0	0	0	0	1	0
S(7) (2,3)	0	1	0	0	0	1	0	0	0	0	1	0
S(8) (2,4)	0	1	0	0	0	1	0	0	0	0	0	1
S(9) (3,1)	0	0	1	0	1	0	0	0	0	0	1	0
S(10) (3,2)	0	0	1	0	0	1	0	0	0	0	1	0
S(11) (3,3)	0	0	0	1	0	1	0	0	0	0	1	0
S(12) (3,4)	0	0	0	1	0	1	0	0	0	0	0	1
S(13) (4,1)	0	0	1	0	0	0	1	0	0	0	1	0
S(14) (4,2)	0	0	1	0	0	0	0	1	0	0	1	0
S(15) (4,3)	0	0	0	1	0	0	0	1	0	0	1	0
S(16) (4,4)	0	0	0	1	0	0	0	1	0	0	0	1

(二)方法：小腦模型分類系統設計程序

本論文的小腦模型心臟病分類系統的設計程序包含：1.隨機選取訓練樣本及測試樣本，2.規劃二維小腦模型分類系統，3.訓練二維小腦模型分類系統，4.測試二維小腦模型分類系統。設計程序分別說明如下：

1. 隨機選取訓練樣本及測試樣本

將無心臟病及有心臟病之樣本分別隨機選取，並組合成為訓練樣本及測試樣本。首先將在資料庫當中的 270 位病患之原始資料以分層（無心臟病/有心臟病）隨機抽樣，進行方式如下：

- (1) 將 140 位無心臟病患者之資料隨機分成兩群 group A: 80% (112 筆資料) 及 group B: 20% (28 筆資料)。
- (2) 將 130 位有心臟病患者之資料隨機分成兩群 group C: 80% (104 筆資料) 及 group D: 20% (26 筆資料)。
- (3) 將 group A 及 group C 的資料組合成為訓練樣本：112 筆無心臟病與 104 筆有心臟病總共有 216 筆資料。
- (4) 將 group B 及 group D 的資料組合成為測試樣本：28 筆無心臟病與 26 筆有心臟病總共有 54 筆資料。

2. 規劃二維小腦模型分類系統

本論文所設計的二維小腦模型如圖一所示，具體的規劃程序如下：

- (1) **選取輸入變數**：在心臟病資料庫中，依屬性與心臟病的相關程度[8]選取四個數值屬性的資料，將其分為 2 組，每組各有 2 個屬性做為小腦模型之輸入狀態變數(S_1, S_2)，其中第一組是由「冠狀動脈被阻塞的數量」屬性與「運動心電圖 ST 下降程度」屬性構成資料樣本 1 (訓練樣本命名為 cmac1，測試樣本命名為 CMAC1)，另一組是由「最大心跳速率」屬性與「收縮血壓」屬性構成資料樣本 2 (訓練樣本命名為 cmac2，測試樣本命名為 CMAC2)。這兩組資料樣本皆包括無心臟病與有心臟病的資料。
- (2) **建構超立方體及記憶細胞**：每一組數值屬性的資料樣本皆由上述的 2 個輸入變數($s=2$)構成狀態平面，如圖二所示，我們依每一輸入變數的範圍將其正規化後均分成四個小區間，因而總共可構成 16 個狀態單元，並以 2 個狀態單元為一個完整區塊，因此在每一層的每個狀態變數皆可形成 2 個區塊($m=2$)，其中第二層及第三層包含一個完整區塊以及一個由 1 個狀態單元構成的不完整區塊，而且總共可形成 3 層($L=3$)超立方體，所以會有 12 個超立方體，因而有 12 個記憶細胞。
- (3) **建立位址索引矩陣**：本論文提出較上一小節的傳統方式更簡易的方法來

決定位址索引矩陣 A，參考圖二的輸入狀態平面圖，將兩個狀態(S_1, S_2)分別定義的同一層區塊(A_{ij} 和 B_{ij})的範圍分別延伸至狀態平面中，其所交集的那些狀態單元與其所對應的超立方體，在表二中給予數值“1”，其餘給予數值“0”，依此原則，即可決定所有狀態單元的位址索引向量，因而構成位址索引矩陣 A。例如在圖二中，由區塊 A_{11} 與區塊 B_{11} 所交集的狀態單元包含座標為 $\{(1,1), (1,2), (2,1), (2,2)\}$ 等四個，因此在表二中由區塊 A_{11} 與 B_{11} 所形成的超立方體和上述四個狀態單元所對應位置的值為“1”，其餘為“0”，其他區塊依此原則即可快速建立如表二的狀態位址索引。有了每一狀態單元所對應的位址索引向量，但在實務上應如何將每一個狀態單元連結到位址索引向量呢？由於本論文將任何輸入狀態變數的範圍分割為四等分，因此只須找出兩個輸入變數在狀態平面的座標(u_i, v_i)，將其代入公式(2)，即可得狀態編號 $s(i)$ 如下：

$$s(i) = 4 * u_i - |4 - v_i| \quad (2)$$

然後據此狀態編號 $s(i)$ 以查表方式可迅速得到任何輸入狀態單元所對應的位址索引向量 a_s^T ，亦即可將每一狀態單元映射至其所對應的記憶細胞的權重值。但以此資料庫而言，並沒有包含小腦模型中所有狀態值，因此為了能適合各種輸入，我們在訓練資料中自行增加了 10 筆訓練資料，以補足原本訓練資料中缺乏的狀態值，如此便可得到調整所有可能狀態的記憶細胞的權重值。至此，本論文規劃的二維小腦模型，總共有 16 種狀態，位址索引矩陣 A 的大小為(16 x 12)，記憶細胞權重 W 的大小為(12 x 1)。

- (4) **決定學習法則**：本論文採用輸出誤差(error)均分法的學習法則，亦即將小腦模型的輸出與期望輸出之間的誤差平均分配到部分所屬的權重調整演算法之中，如下式：

$$w^{(n)} = w^{(n-1)} + \alpha * \frac{a_i(y_{i,d} - a_i^T \cdot w^{(n-1)})}{L}, 1,2,3,\dots,16 \quad (3)$$

其中 $w^{(n)}$ 表示第 n 次學習週期的記憶權重， α 為學習率(Learning Rate)，其值介於(0, 1]之間， $y_{i,d}$ 為某一狀態單元所對應的期望輸出， a_i 為某一狀態單元所對應的位址索引向量，L 為超立方體的層數，在本論文是 3 層(i.e. L=3)，在訓練階段的權重初始值全部給定一個固定值 1。

- (5) **選定期望輸出**：期望輸出是依照狀態平面分割的狀態單元與屬性性質，自行依目的而給予。本論文參考的各種屬性性質的期望輸出如表三所示。

表三、各種屬性性質的期望輸出

屬性性質	期望輸出
望大	170,180,190,...,320
望小	160,150,140,...,10
望目標	320,150,...,160,...,150,320

最後將所有輸入狀態 X 所對應的輸出 Y 可表示為下式：

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_s \\ \vdots \\ y_{16} \end{bmatrix} = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_s^T \\ \vdots \\ a_{16}^T \end{bmatrix}_{16 \times 12} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_s \\ \vdots \\ w_{12} \end{bmatrix}_{12 \times 1} \quad (4)$$

其中 y_s 為第 s 個狀態單元所對應的小腦模型輸出值， a_s 為第 s 個狀態單元所對應的位址索引向量，T 表示矩陣的轉置， w_s 為第 s 個狀態單元所對應的記憶細胞權重。

3. 訓練二維小腦模型分類系統

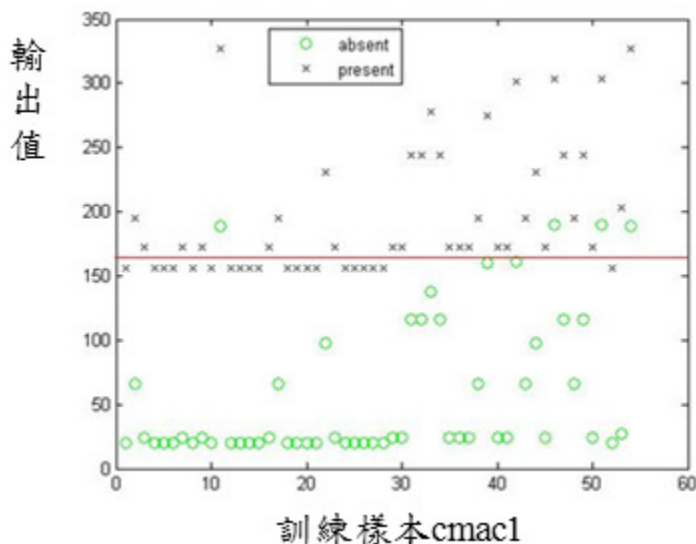
二維小腦模型必須經由學習階段，讓系統自動找出符合期望輸出的記憶細胞權重值，以便能應用於心臟病之分類。本論文的小腦模型學習演算法流程如下：

- (1) 初始化小腦模型
- (2) 讀取並正規化輸入變數
- (3) 計算輸入變數所屬狀態單元
- (4) 計算狀態編號並查出位址索引
- (5) 透過位址索引得到記憶權重及輸出
- (6) 計算輸出誤差是否小於 1%，若否，則繼續下一步；若是，則學習結束
- (7) 以輸出誤差回饋調整記憶權重
- (8) 回到步驟 5

本論文依據上述小腦模型的規劃程序以及小腦模型學習演算法流程，以 MATLAB 撰寫小腦模型分類系統之程式，將讀取訓練樣本的資料，依據平均學習法則，不斷地調整權重值，直到輸出誤差在 1% 以內為止。本論文規劃的第一組訓練資料樣本 cmac1 經由訓練所得記憶體權重為

$$W_1 = [60.9812 \quad 92.1352 \quad 85.0784 \quad 65.6168 \quad 39.9697 \quad 56.0319 \quad 79.3909 \\ 128.4192 \quad 55.0356 \quad 54.4701 \quad 113.7062 \quad 80.5996] \quad (5)$$

經過權重 W_1 所映射的輸出值，如圖三所示，我們可發現在輸出值為 165 之處，明顯的可分成兩群，所以第一組選擇 165 作為分類閾值。

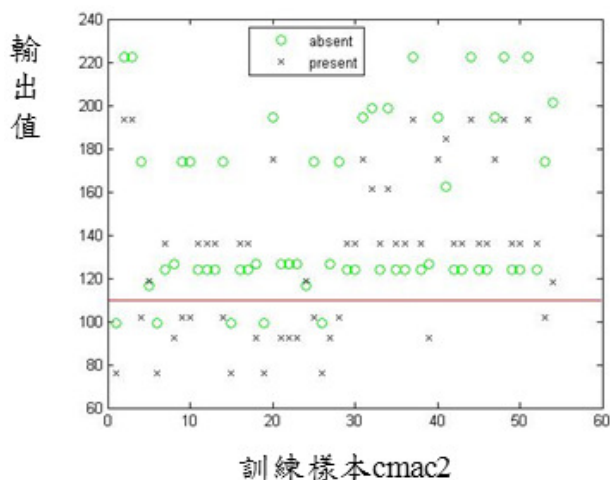


圖三、cmac1 輸出分佈圖

第二組訓練資料樣本 cmac2 經過訓練所得記憶體權重為

$$W_2 = [34.1177 \quad 126.0999 \quad 7.8155 \quad 36.6274 \quad 59.9163 \quad 46.1431 \quad 78.0507 \\ 20.5505 \quad 81.1975 \quad 37.7570 \quad 21.6339 \quad 64.0721] \quad (6)$$

經過權重 W_2 所映射的輸出值，如圖四所示，可看出無病與有病分佈的不同，有病部分幾乎都在 110 以上，所以第二組選擇 110 為閾值。



圖四、cmac2 輸出分佈圖

當經過所有訓練資料調整權重之後，便可讓小腦模型得到每一狀態的特徵權重，即可作為小腦模型的分類測試之用。

4. 測試二維小腦模型分類系統

為衡量本小腦模型分類系統的分類品質，本論文採用下列分類效能指標以檢驗分類之效果：

- (1) 分類正確率(CA)： $CA=(AA+PP)/(AA+AP+PA+PP)$
- (2) 有病捕捉率： $R=PP/(PA+PP)$
- (3) 分類錯誤的成本： $Cost=AP+PA*5$

上述公式中所用到的代號如表四的混亂矩陣所示，

表四、混亂矩陣

	預測無病	預測有病
實際無病	AA	AP
實際有病	PA	PP

其意義說明如下：AA 表示實際無病而且預測亦是無病，PP 表示實際有病而且預測亦是有病，兩者皆得到正確的分類；AP 則是實際無病但預測卻是有病，PA 則是實際有病但預測卻是無病，亦即 AP 與 PA 兩者皆是得到錯誤的分類，因此分類正確率 CA 的意義是正確分類的樣本數佔全部樣本數之比率；有病捕捉率 R 的意義是實際上有病的樣本數中被正確分類的比率，亦可說是有病的成員中被捕捉到多少的比率，有病捕捉率是具有相對重要性的指標，以上兩項指標皆是數值越大越好；分類錯誤的成本 Cost 的意義則是指每次預測錯誤時的一項成本指標，其中以實際有病但預測卻是無病的權重五倍於實際無病但預測卻是有病的權重，也就是實際有病的預測錯誤較為嚴重，此項指標是越小越好。

經由訓練完成後的二維小腦模型已具有每一分割狀態與其特徵權重的直接對應關係，因此固定這些特徵權重值即可進行分類的測試。除了輸入樣本以及權重值固定等不同於小腦模型的訓練階段外，程式的其餘部分都是相同的。第一組測試樣本是 CMAC1，其經由(5)式的記憶體權重，並以 165 作為不同類別之分類閾值。第二組測試樣本是 CMAC2，其經由(6)式的記憶體權重，並以 110 作為不同類別之分類閾值。本論文的測試資料共有 54 筆，其中 28 筆是無病以及 26 筆是有病的資料，下一節將以這些測試資料展示訓練後的二維小腦模型分類系統的分類結果。

結果與討論

依上節所述之測試原則，將第一組測試樣本 CMAC1 輸入至已訓練完成的二維小腦模型中，得到的結果如表五所示，

表五、測試樣本 CMAC1 分類結果

	預測無病	預測有病
實際無病	19	9
實際有病	1	25

並據此數據分別計算分類效能指標，其結果如下：分類正確率 $CA = 81.5\%$ ，有病捕捉率 $R = 96.2\%$ ，分類錯誤的成本 $Cost = 14$ 。同理，將第二組測試樣本 CMAC2 亦輸入至二維小腦模型後得到結果如表六所示，

表六、樣本 CMAC2 分類結果

	預測無病	預測有病
實際無病	17	11
實際有病	2	24

同樣依此數據計算所得的分類效能指標如下：分類正確率 $CA = 75.9\%$ ，有病捕捉率 $R = 92.3\%$ ，分類錯誤的成本 $Cost = 21$ 。從以上兩組的測試結果可知，第一組測試樣本 CMAC1 的三項分類效能指標皆優於第二組測試樣本 CMAC2，亦即選擇第一組測試樣本 CMAC1 的屬性可以達到比較好的分類預測效果；尤其是第一組測試樣本 CMAC1 的有病捕捉率更達到 96.2% 的水準，這是由於測試樣本 CMAC1 在經過狀態映射之後，不同病症類別之間在輸出的分佈上有著明顯的差異，這是因為第一組測試樣本的屬性「冠狀動脈被阻塞的數量」與「運動心電圖 ST 下降程度」皆是心臟病的重要相關屬性因素[8]，而且第一組測試樣本 CMAC1 的屬性是由兩個有病傾向相同的屬性構成。然而第二組測試樣本 CMAC2 的屬性僅有「最大心跳速率」是心臟病比較相關的屬性，但是「收縮血壓」則是比較不相關的屬性，而且第二組測試樣本 CMAC2 的屬性又是由不同傾向的屬性所組成，例如在有病的傾向上，「最大心跳速率」是望小的屬性，「收縮血壓」是望目標的屬性，所以採用相關且傾向相同的屬性會有較佳的分類效果。因此本論文在此測試試驗

中找出了心臟病的兩個相關程度較高的屬性作為輸入變數，使得整體分類正確率可達到 81.5%以及有病捕捉率達到 96.2%的成果，由此足以證明本二維小腦模型分類系統的有效性。若期望設計更佳的小腦模型分類器，可能必須採用更多數目的相關屬性，或是將狀態空間分割成更精細，然而，這樣會使得小腦模型分類器的結構更為複雜，因此，在分類效果與其結構的複雜度之間必須要有所取捨。

結 論

本論文基於葡萄牙的心臟病資料庫已提出二維小腦模型分類系統做為自動判別心臟病患者的病症類型。在二維小腦模型的狀態平面上，本論文以座標方式經簡易公式計算可得狀態單元所對應的位址索引，而且我們提出區塊交集法可快速建立位址索引矩陣，使得在建立位址索引矩陣的錯誤率降低。本論文採用二維小腦模型可使分類系統的設計簡化，並以平均輸出誤差的學習法則來調整小腦模型的記憶細胞權重，可快速收斂到輸出誤差小於 1% 以內。在測試試驗中，我們發現若以心臟病的兩個相關程度較高而且病症傾向相同的屬性為輸入變數，所得結果較佳：整體分類正確率達到 81.5%以及有病捕捉率可達到 96.2%的結果，由此足以證明本論文僅以二維小腦模型設計分類系統的有效性。

參考文獻

- [1] 何逸僊：病理檢驗醫學，2002，91-104。力大圖書。
- [2] Marr DA: A theory of cerebellar cortex. J Physiol 1969, 202: 437-470.
- [3] Albus JS: A new approach to manipulator control: the cerebellar model articulation controller. ASME J-Dynamic Systems, Measurement, and Control. 1975, 97: 220-227.
- [4] Cotter NE and Kim H: A pulse Neural Network Capable of Universal Approximation. IEEE Trans on Neural Network 1992, 3: 642-647.
- [5] 陳珍源：小腦模式在馬達速度控制之實現。中華技術學院學報，2003，26：160-169。
- [6] 林正堅：以模糊小腦模型為基礎之非線性通道等化器。碩士論文，2004。朝陽科技大學資訊工程系。
- [7] <http://www.liacc.up.pt/>

[8] 李嘉陵、蔡坤龍：應用灰色統計於心臟病資料屬性分析。Ming-Chuan Health Tech Journal，2010，1：1-6。

On 2D Cerebellar Model-Based Heart-Disease Classification System

Jia-Ling Lee^{1,*}, Steven Tsai²

¹Department of Biomedical Engineering, School of Health Technology, Ming-Chuan University, (Taoyuan, Taiwan, R.O.C.)

² Department of Electronic Engineering, School of Information Technology, Ming-Chuan University, (Taoyuan, Taiwan, R.O.C.)

Abstract

We will propose a classification system using two dimensional Cerebellar Model (2D CM) for the heart-disease database of Porto university. The purpose of this paper is to classify some classes of those patients in the database into the correct one. The design procedure for the 2D CM classification system is as follows: (1) We select some samples whose heart disease may be absent or present from the database randomly, and compose these samples to act as a training group and an evaluated group. Each group contains both samples with or without heart disease. (2) The classification system is then schemed. (3) Based on the desired output and the learning rule, the weighting memory of CM is tuned from the training group. (4) The classification system is tested from the evaluated group finally. In this paper, we use both the blocks' intersection method to build up address indices rapidly, and the coordinate computing method to connect states with address indices. We use two types of different attributes in the database to act as the input of the classification system, and adopt the mean output error method for the learning rule to tune the weighting memory cell of CM. The proposed framework of 2D CM classification system in this paper is simple and converges fast within 1% output error. In the evaluated trial, we find two important attributes, whose relation level with heart disease is more higher, to jointly screen the potential heart disease. The percentage of accurate classification rate can attain 81.5%, and the percentage of capture rate for those who have heart disease can attain 96.2%. It is demonstrated that the proposed classification scheme in this paper is effective for the 2D CM framework.

Furthermore, we find that the better classification result can be achieved by adopting some attributes with higher relation level and with the same tendency towards heart disease.

Keyword: Heart disease, Cerebellar model, Memory cell, Classification system

Corresponding author: Jia-Ling Lee [jllee@mail.mcu.edu.tw]

Received 3 Aug 2012/Revised 30 Aug 2012/Accepted 31 Aug 2012

MC-Transaction on Biotechnology, 2012, Vol. 4, No. 1, e4

©This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.