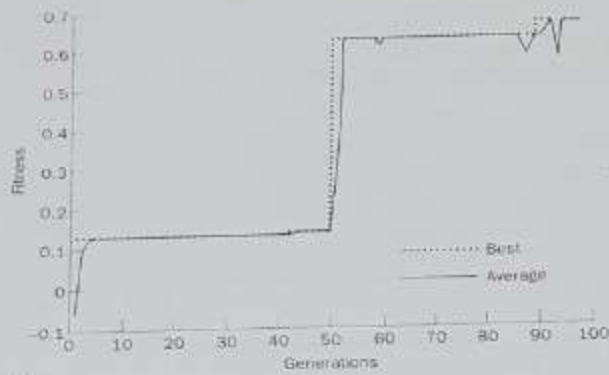
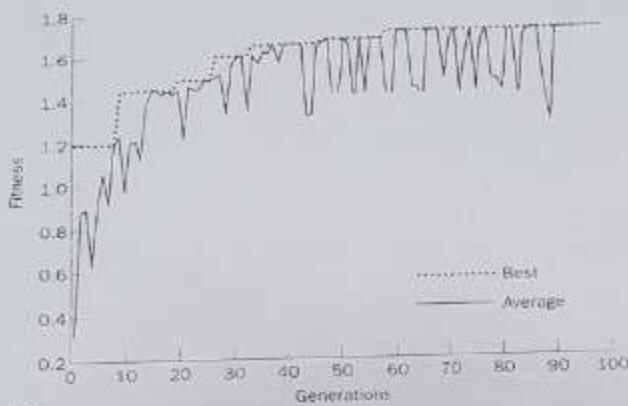


(總分 100)(子題未標示配分，表示均分)

1. (15%) The following two graphs are used the same GA with a different parameter for solving a problem. What is the different parameter? And explain why?

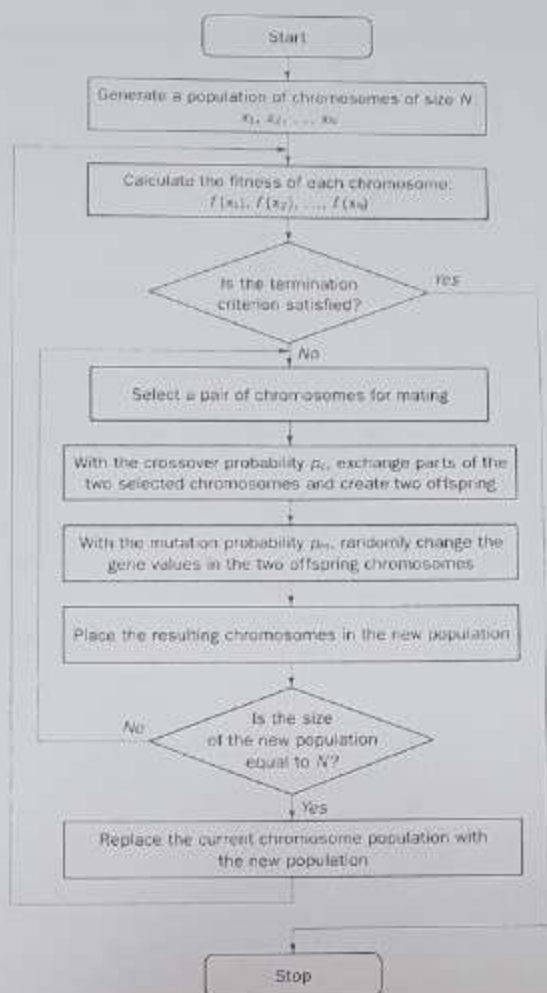


(a)



(b)

2. (20%) (1)(5%) 試說明下列基因演算法運作過程；(2)(10%) 基因演算法中，有哪兩個重要問題要克服？以 10 座電廠分 12 月歲修(每一電廠歲修都只要 1 個月)為例說明。(3)(5%) 突變在基因演算法中，扮演何種角色(功用)？請說明。



3. (10%) 下表 8.1 為比較 4 種技術之優劣，試說明 (1) 何謂“Learning ability”與“Explanation ability”？ (2) 表中“Learning ability”與“Explanation ability”之結果為何如此？

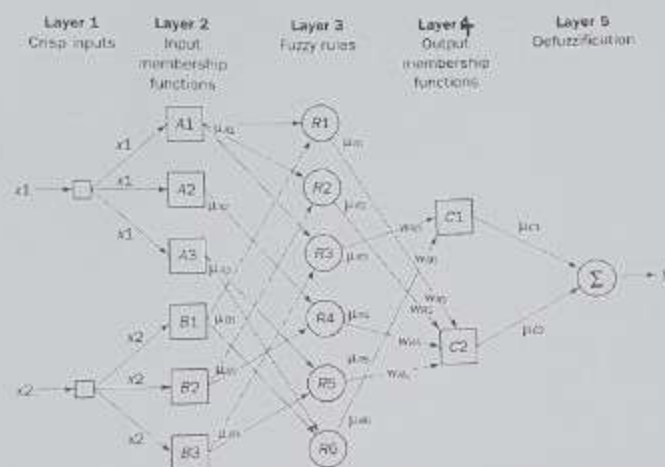
4. (20%) (1) 何謂 **Gini coefficient**？其目的為何？其值範圍為何？越大越好還是越小越好？ (2) 假如今天有人以 **decision tree** 做 **data mining** 時，跟你說他分類效果很好，達到 100%，你會問他什麼問題？確定他的結果是好的？

Table 8.1 Comparison of expert systems (ES), fuzzy systems (FS), neural networks (NN) and genetic algorithms (GA)

	ES	FS	NN	GA
Knowledge representation	○	●	□	■
Uncertainty tolerance	○	●	●	●
Imprecision tolerance	○	●	●	●
Adaptability	□	●	●	●
Learning ability	□	●	●	●
Explanation ability	○	●	□	■
Knowledge discovery and data mining	○	●	●	○
Maintainability	○	○	●	○

The terms used for grading are: □ bad, ■ rather bad, ○ rather good and ● good

5. (15%) 下圖為一 neuro-fuzzy 架構，(1) 請說明每一層功能 (2) 如何學習修正(項目與方法)



6. (20%) 今天小明超商有 5 筆交易，相關資訊如下表，小明希望找出哪些商品是有關聯，定了門檻為：支持度 (support) = 0.3，信心水準 (confidence) = 0.7，請你找出所有符合條件之“關聯法則”。

交易編號	商品編號
1	1, 2, 3, 5, 6
2	2, 4, 6
3	1, 6
4	2, 3, 5
5	2, 5

支持度 =  $\frac{\text{個數}}{\text{總數}}$





1. 主要是在突變率的不同。因為突變率雖然能隨時提高物種的適合度，但並非每次突變都會讓物種的適合度上升，也可能造成下降。整體來看圖(b)的適合度確實突破了 local max (1.2)，但也因突變率增加的因素，導致曲線在上升過程中是密集震盪的。

2.

① Step1: 生成族群大小為  $N$  的基因

Step2: 計算 Step1 的適合度

Step3: 判斷此族群是否達到一定的適合度，即滿足終止條件

Step4: 若未滿足終止條件則隨機選擇一對染色體配對

Step5: 用交叉率 ( $P_c$ ) 和突變率 ( $P_m$ ) 產生另一對新的子代

Step6: Repeat step4-step5 until 新族群數量達到  $N$

Step7: 用新族群取代舊族群

Step8: Repeat step3-step7 until 滿足終止條件為止

②

主要是「計算(適合度)問題」和「編碼問題」， $\text{fitness function} = \text{Min} \{ \text{總發電量} - \text{總使用電量} - \text{產修電量} \}$

假設我們以 1 組 bits 代表一座電廠，那麼一組 bits 的實際 bit 數

要設成多少會比較好算呢？若以題目所述分 12 個月歲修，我們可以

賦予每組 bits 數 12 bit，並以 value 1 表示歲修中，value 0 表示運作中，好

的編碼會使結果更趨於真實狀況；反之下降，而在 GA 運作下 fitness

function 也很重要以此題為例，適合度則是要取 12 月份最小的，否則可

能會使得在某月份電廠的淨保留量  $< 0$  的情況發生。

③

突變的存在是為了避免物種純化，例如  $1010111$  和  $0101010$  在怎麼交

配也無法產生以 1 為首的基因，雖然交配會使 fitness 收敛到 local max

但在某些條件下我們需要更高 fitness 的物種，因此需要突變來產

到 Global Max 維持多樣性 提高 fitness

3.

① Learning ability = 即系統在得到大量樣本後，是否需專家介入又或者能自己從中學習進而辨識出某種情況的能力。

Explanation ability = 即當給予 input data 到得到 output 的過程與結果，是否能讓使用者清楚理解的能力。

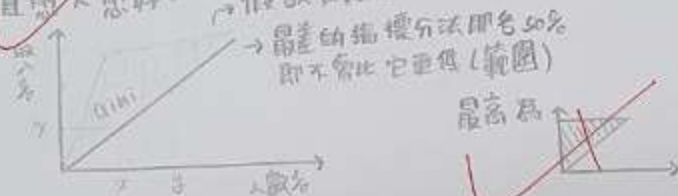
②

Learning ability = 因為 NN 和 GA 不需透過專家的介入，即可透過分析大量樣本數自行調整一些參數(如權重、適合度等等)進行自我學習故為"●"，而 ES 和 FS 則需要透過專家的介入調整 rule 故為"□"。

Explanation ability = ES 和 FS 都有清楚的 rule 規定輸入何種 data 會得到何種 output 故都為"●"，而 NN 和 GA 的解釋都比前二者差得遠，因為此二套系統在輸入 input 後我們並無法得知運作過程而直接看到 output 像黑箱一樣，但 GA 至少還有 fitness function 可觀察，所以 explanation ability 又比 NN 好一些，故 NN 為"□"、GA 為"●"。

4.

(1) Gini coefficient 是用來分析 decision tree 的 predictor 變的好不好的一種函式。其值愈大愈好。→ 假設在經由指標分類可使其中一項佔超過 50% 則為好的指標，即以最少成本即可獲得最大效益。



(2) 問他所建立的 decision tree 或 data mining 有多深，因為建太深反而容易查找，所以身份證字號分類的台灣人。

① 分類樣本數大不大 ② 分類層數深不深

5.

(1)

Layer 1: 取得清晰的輸入，並賦予適當的模糊程度

Layer 2: 將 crisp inputs 對應到相應函數中

Layer 3: 使用 production operators 的方式計算其模糊程度 ANDLayer 4: 使用 OR 的方式計算前項規則，並得到輸出的模糊函數

Layer 5: 逆模糊化是為了取得清晰單一輸出，最常使用的方法為重心法

(2)

藉由修改權重、倒傳遞修改法。

一開始先將各項規則的權重都設成相同，在經過多次訓練後，將出現機率高或確定程度高的規則提高權重；反之降低，進而提高 output 的正確率

最小支持個數 = 支持度  $\times$  全體數 =  $0.3 \times 5 = 1.5$ 

1-2項目	TS	SP
1	{1,3}	2
2	{1,2,4,5}	4
3	{1,4}	2
4	<del>{2}</del>	<del>1</del>
5	{1,4,5}	3
6	{1,2,3}	3

2-3項目	TS	SP
1,6	{1,3}	2
2,3	{1,4}	2
2,5	{1,4,5}	3
2,6	{1,2}	2
3,5	{1,4}	2

3-5項目	TS	SP
2,3,5	{1,4}	2

所有大型項目集 {1} - {2} - {3} - {4} - {5} - {6} - {1,6} - {2,3} - {2,5} - {2,6} - {3,5} - {2,3,5}

關聯法則:

$$\begin{aligned}
 \{1\} \rightarrow \{6\} &= \frac{2}{2} = 1 & \{2\} \rightarrow \{6\} &= \frac{2}{4} = 0.5 \\
 \{6\} \rightarrow \{1\} &= \frac{2}{2} = 1 & \{6\} \rightarrow \{2\} &= \frac{2}{3} = 0.667 \\
 \{2\} \rightarrow \{3\} &= \frac{2}{4} = 0.5 & \{3\} \rightarrow \{5\} &= \frac{2}{2} = 1 \\
 \{3\} \rightarrow \{2\} &= \frac{2}{2} = 1 & \{5\} \rightarrow \{3\} &= \frac{2}{3} = 0.667 \\
 \{2\} \rightarrow \{5\} &= \frac{3}{4} = 0.75 & \{2,3\} \rightarrow \{5\} &= \frac{2}{2} = 1 \\
 \{5\} \rightarrow \{2\} &= \frac{3}{3} = 1 & \{2,5\} \rightarrow \{3\} &= \frac{2}{3} = 0.667 \\
 & & \{3,5\} \rightarrow \{2\} &= \frac{2}{2} = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{Ans: } & \{1\} \rightarrow \{6\} - \{3\} \rightarrow \{5\} \\
 & \{3\} \rightarrow \{2\} - \{2,3\} \rightarrow \{5\} \\
 & \{2\} \rightarrow \{5\} - \{3,5\} \rightarrow \{2\} \\
 & \{5\} \rightarrow \{2\}
 \end{aligned}$$