

(總分 130)

1. (10%) In 1950, Alan Turing proposed the concept of a universal machine (Turing machine), and proceeded the famous Turing Test. What is the Turing Test? What are the remark qualities that make it really universal?

$$P(A \cap B_1) = 0.04$$

$$P(A \cap B_2) = 0.16$$

$$0.1$$

$$0.1$$

$$0.1$$

$$0.5$$

$$P(A|B_i) = \frac{P(A \cap B_i)}{P(B_i)}$$



(10%) 傳統 rule-based 專家系統問題有哪些？如何改善？

(10%) 何謂 Bayesian rule (事後機率)？

$$p(A|B) = \frac{p(B|A) \times p(A)}{p(B)}$$

這是事前機率

為何使用事後機率推論？

(10%) If A and B are two sets, A is included in B (with probability $P(A) = 0.5$) and B can be divided into 5 subsets (mutually exclusive) with the probabilities in the following Table 1. Data in the table is correct or not? Why?

Table 1.

Probability	Hypothesis				
	i=1	i=2	i=3	i=4	i=5
$P(B_i)$	0.2	0.2	0.2	0.2	0.2
$P(A B_i)$	0.2	0.8	0.5	0.5	0.5

(10%) 警察抓到四個嫌疑犯：甲、乙、丙、丁(涵蓋所有可能)，他們由前科紀錄猜測可能犯案機率分別為 0.3、0.4、0.7、0.5；警方又依序找到四項證物(彼此獨立無關)，分別是四位嫌犯留下之機率如下表所示：

留下之機率	甲 0.3	乙 0.4	丙 0.7	丁 0.5
證物 1	0.2	0.5	0.2	0.2
證物 2	0.5	0.4	0	0.4
證物 3	0	0.2	0.1	0
證物 4	0.4	0.1	0.2	0.5

請問：誰最有可能犯案？(以 Bayesian reasoning 求解，每一過程(當證物是陸續發現，每一證物出現後，誰最有可能，要說明)均要呈現，只寫答案不給分)

相關公式：

Handwritten calculations for Bayesian reasoning:

Initial probabilities: 甲 0.3, 乙 0.4, 丙 0.7, 丁 0.5

After Evidence 1:

$$0.3 \times 0.2 = 0.06$$

$$0.4 \times 0.5 = 0.2$$

$$0.7 \times 0.2 = 0.14$$

$$0.5 \times 0.2 = 0.1$$

Sum: $0.06 + 0.2 + 0.14 + 0.1 = 0.5$

After Evidence 2:

$$0.06 \times 0.5 = 0.03$$

$$0.2 \times 0.4 = 0.08$$

$$0.14 \times 0 = 0$$

$$0.1 \times 0.4 = 0.04$$

Sum: $0.03 + 0.08 + 0 + 0.04 = 0.15$

After Evidence 3:

$$0.03 \times 0 = 0$$

$$0.08 \times 0.2 = 0.016$$

$$0 \times 0.1 = 0$$

$$0.04 \times 0.1 = 0.004$$

Sum: $0 + 0.016 + 0 + 0.004 = 0.02$

After Evidence 4:

$$0 \times 0.4 = 0$$

$$0.016 \times 0.1 = 0.0016$$

$$0 \times 0.2 = 0$$

$$0.004 \times 0.5 = 0.002$$

Sum: $0 + 0.0016 + 0 + 0.002 = 0.0036$

Final probabilities: 甲 0.0036, 乙 0.0016, 丙 0.002, 丁 0.0036

Conclusion: 甲 and 丁 are the most likely suspects.

$$p(A) = \sum_{i=1}^n p(A|B_i) \times p(B_i)$$

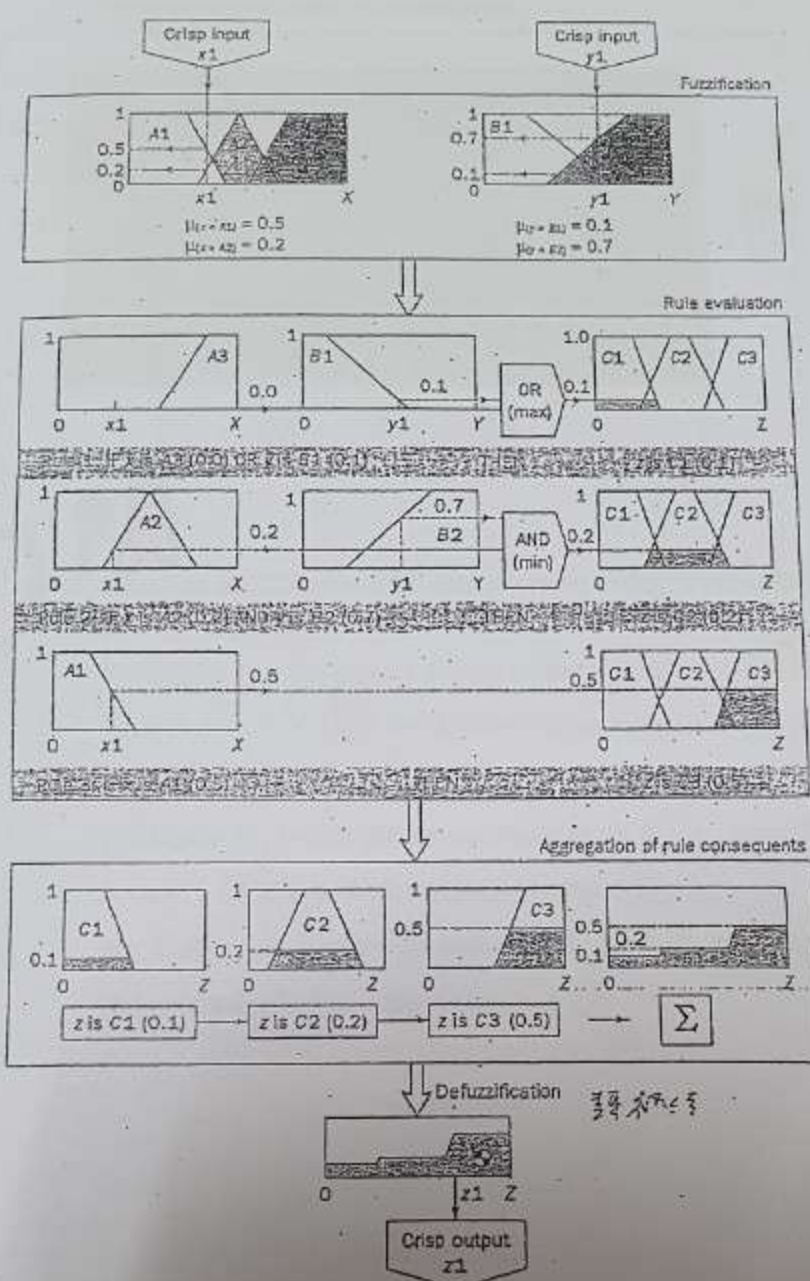
$$p(H_i|E) = \frac{p(E|H_i) \times p(H_i)}{\sum_{k=1}^m p(E|H_k) \times p(H_k)}$$

$$p(H_i|E_1, E_2, \dots, E_n) = \frac{p(E_1, E_2, E_3, \dots, E_n|H_i) \times p(H_i)}{\sum_{k=1}^m p(E_1, E_2, \dots, E_n|H_k) \times p(H_k)}$$

6. (10%) 何謂(1)linguistic variables 及對應之(2)fuzzy set? 以天氣溫度為例，作一說明。

7. (10%) 模糊推論(fuzzy inference) 的優點為何? 困難點為何?

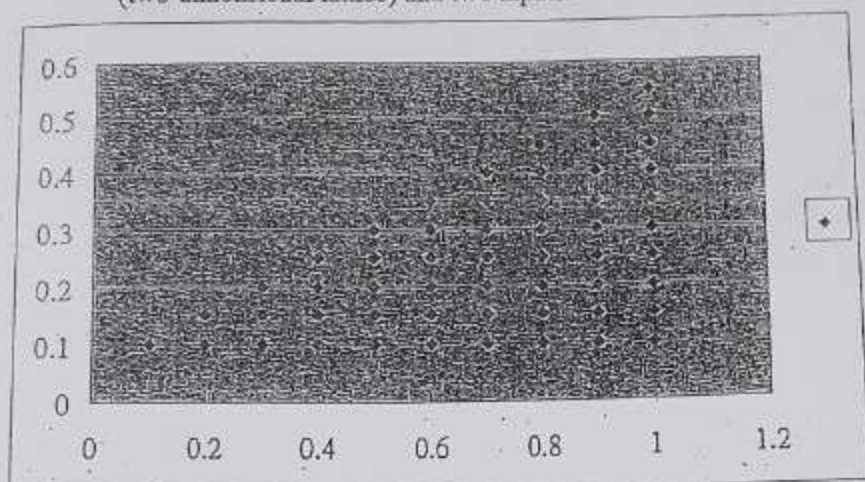
8. (10%) 下圖為 Mamdani 模糊推論方式，請說明每一步驟：



9. (10%) 類神經網以何種方式達到學習目的？以 MLP 為例，說明其學習過程(越詳細越好，不用導公式，只需說明原理)。

10. (10%) If a set of data with distribution is shown as follows. What is the weights distribution for a Self-organization Feature Map NN with 25 outputs

(two-dimensional lattice) and two inputs.



11. (10 %) Based on the NN, select the correct corresponding number for each curve in the following graph.

For Figure (a), the twenty hidden neurons is curve (1); ten hidden neurons is curve (2); five hidden neurons is curve (3); two hidden neurons is curve (4).

For Figure (b), twenty hidden neurons is curve (5); ten hidden neurons is curve (6); five hidden neurons is curve (7).

For Figure (c), network trained with 'perfect' examples is (8); Network trained with 'noisy' examples (9).

hidden $\theta_3 = 42$, R error θ_3
 perfect = R error θ_3

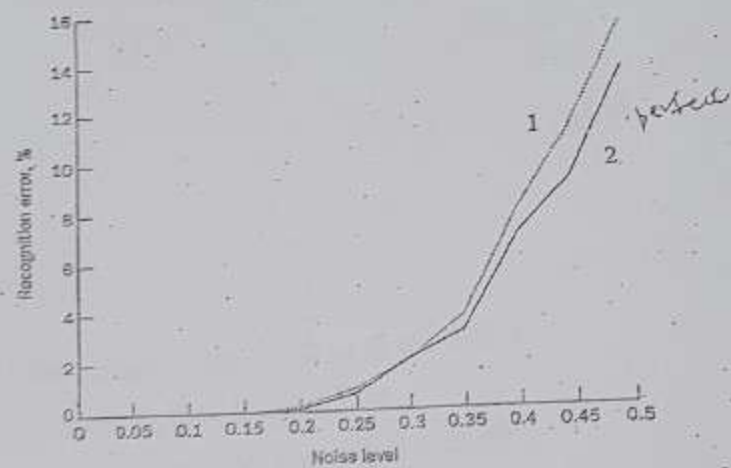
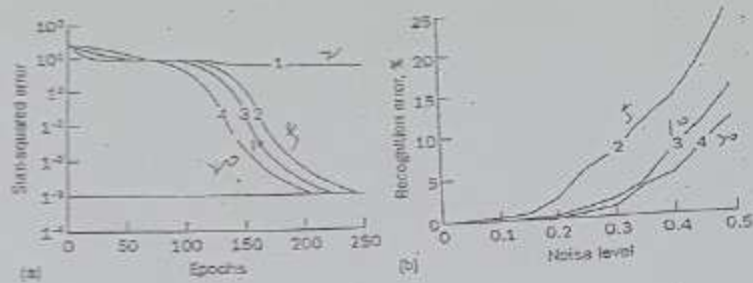


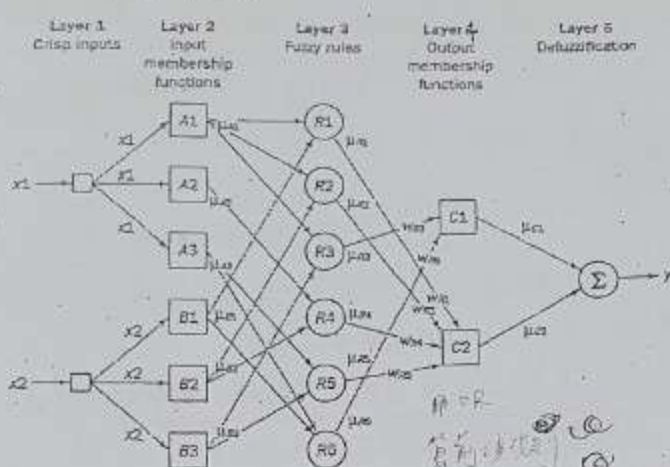
Figure 9.22 Performance evaluation of the digit recognition network trained with 'noisy'

(c)

OK
 12. (10%) 試說明下列基因演算法流程 6A 流程

13. ① 每一層的作用為何? ② 各層主要為何? ③ 哪層資料可修改?

108 期末
第 5 題 answer



1. 模糊度
2. 模糊度
3. 模糊度
4. 模糊度
5. 模糊度

模糊度
模糊度
模糊度
模糊度
模糊度

模糊度
模糊度
模糊度
模糊度
模糊度

模糊度

- (1)
- Layer 1: 取得清晰的輸入，賦予適當的模糊程度
 - Layer 2: 將 Crisp inputs 對應到歸屬函數中
 - Layer 3: 使用 product operation 的方式計算其模糊程度
 - Layer 4: 使用 OR 的方式計算前項規則，然後得到輸出的歸屬函數
 - Layer 5: 逆模糊化是為了取得清晰的單一結果，最常使用的方式為重心法

(2)

Layer 1: $x_i = y_i$

Layer 2:
$$\mu_{A_i}(x_i) = \begin{cases} 0 & \text{if } x_i \leq A - \frac{b}{2} \\ 1 - \frac{2|x_i - A|}{b} & \text{if } A - \frac{b}{2} \leq x_i \leq A + \frac{b}{2} \\ 0 & \text{if } x_i \geq A + \frac{b}{2} \end{cases}$$

(3)

A, b 全數可修改
會使 triangle membership function 發生變化

Layer 3: $g_i = \mu_{A_i}(x_i) \wedge \mu_{B_i}(x_i) \rightarrow g_{R1} = \mu_{A1} \wedge \mu_{B1} = \mu_{R1}$

Layer 4: $g_i = \mu_{A_i}(x_i) \vee \mu_{B_i}(x_i) \rightarrow g_{C1} = \mu_{R1} \vee \mu_{R2} = \mu_{C1}$

Layer 5:
$$\frac{\mu_{A1}x_{C1} + \mu_{A2}x_{C2}}{\mu_{A1} + \mu_{A2}}$$

語義值: 氣溫高(15°C ~ 24°C), 氣溫低(15°C以下)

(2) 語意值: 很冷 - 冷 - 適中 - 熱 - 很熱

1. Turing Test:
測試機器(電腦)的智力水平。表現可以在認為任何程度達到人類水平。表現
remark: 1. 人類及非人類客觀的測試標準
2. 測試本身是相當複雜的
受測者可以自由問任何問題

- a. 不用精準的物理模型
- b. 可以將人類的專業經驗應用在多系統中

所需規則很難用語意式模糊規則表達, 所以無法處理所有的問題
模糊規則需要較長時間調整

b. 很難將模糊規則賦予適當的歸屬函數

Fuzzification:

取得清晰的輸入 x_i, y_i , 賦予適當的模糊集程度

Rule evaluation:

取得模糊規則的輸出: $\mu(x=A_1)=0.5, \mu(x=A_2)=0.2, \mu(y=B_1)=0.1, \mu(y=B_2)=0$
將模糊輸出當成規則的前項。如果前項數越多, 就利用模糊操作 (OR, AND) 的方式取得評估前項結果的單一數值。其單一數值可應用在歸屬函數中。

Aggregation of rule consequents:

聚合是所有規則輸出的單一化過程。

Defuzzification:

解逆模糊化的結是為了取清晰結果的單一值
最常使用到的方式是重心技術 (COG), 其

計算: $\alpha \times x_i + \text{error} = \text{後果}$

① 學習方式是透過 Back propagation 演算法

不斷輸出修改權重直到誤差值
直到中止條件為止

Step 1: Initialization 初始化

變數取得權重初值與^{內標值}臨界值

Step 2: (Activation) 激活

計算隱藏層輸出層神經元的實際輸出

Step 3: (Weight training)

計算隱藏層輸出層神經元的誤差梯度 權重校正 更新權重

Step 4: (Iteration) 迭代

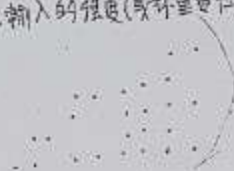
回到步驟二 重複其過程直到滿足誤差中止條件為止

9. (1) 神經元透過連結組態 每個連結都有權重。權重是 ANN 中長期記憶的基本方式。
它們表達了神經元輸入的強度(或重要性)。ANN 透過不斷修正權重進行學習。

⑩

10

X



修改

1, ⑪

小約在上, perfect 在下

① 4 4

② 3 3

③ 2 2

④ 1 1

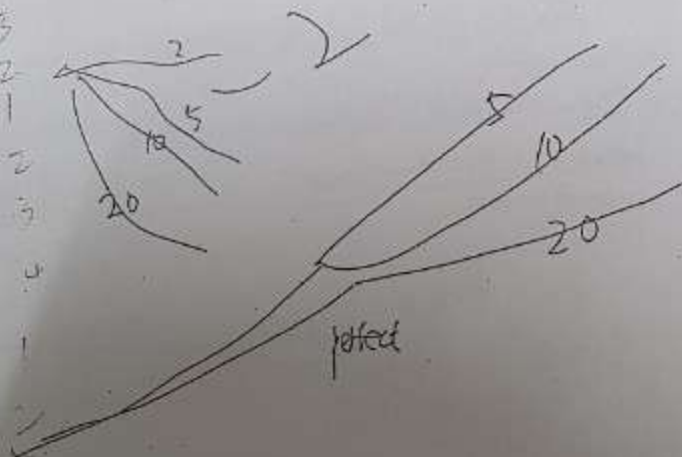
⑤ 4 2

⑥ 3 3

⑦ 4 4

⑧ 1 1

⑨ 2 2



- Step 1: 亂數取得群體 $N = X_1, X_2, X_3, \dots, X_N$
- Step 2: 計算群體的適應度 $f(X_1), f(X_2), f(X_3), \dots, f(X_N)$
- Step 3: 選取高適應度的群體個體
- Step 4: 根據 P_c, P_m 透過交配或突變數得後代個體
- Step 5: 將後代個體加入新群體中
- Step 6: 重回 Step 3 直到產生新群體
- Step 7: 將新群體取代初始群體
- Step 8: 重回 Step 2 重複其過程直至停止條件為止。典型疊代為 50~5

① 每一層功用為何?

Layer 1: 取得清晰的輸入，賦予適當的模糊集程度 (Crisp inputs)

Layer 2: 將 Crisp inputs 對映到歸屬函數中 (input membership functions)

Layer 3: 使用 product operation 的方式計算其模糊規則 (Fuzzy rules)

Layer 4: 使用 OR 的方式計算前項規則，然後得到輸出的歸屬函數 (Output membership functions)

Layer 5: 逆模糊化是為了取得清晰的單一結果。最常使用到的方式 (Defuzzification)

② 各權重為何?

Layer 1: $X_i^{(1)} = y_i^{(1)}$

Layer 2:
$$\begin{cases} 0 & \text{if } X_i^{(2)} \leq a - \frac{b}{2} \\ 1 - \frac{2|X_i^{(2)} - a|}{b} & \text{if } a - \frac{b}{2} \leq X_i^{(2)} \leq a + \frac{b}{2} \\ 0 & \text{if } X_i^{(2)} \geq a + \frac{b}{2} \end{cases}$$

Layer 3: $\mu = X_{i1}^{(3)} \times X_{i2}^{(3)} \times X_{i3}^{(3)} \times \dots \times X_{in}^{(3)} = \mu_{R1} \mu_{R2} \dots \mu_{Rn} = R_1$

Layer 4: $\mu^{(4)} = X_{i1}^{(4)} \oplus X_{i2}^{(4)} \oplus X_{i3}^{(4)} \times \dots \times X_{in}^{(4)} = \mu_{R3} \oplus \mu_{R6} = R_1$

Layer 5: $y = \frac{\mu_{R1} \times \mu_{R2} \times \mu_{R3} + \mu_{R4} \times \mu_{R5} \times \mu_{R6}}{\mu_{R1} + \mu_{R2} + \mu_{R3} + \mu_{R4} + \mu_{R5} + \mu_{R6}}$

③ 哪個參數可修改?



a, b 參數可修
它會使得 triangle membership function 產生變化