

Homework 1

工科所碩二 r08525116 吳承哲

1-5 : d e d c d

6-10 : X e b e b

11-15: c d c d c

16-20: b b c d d

1. d

判斷是否適合使用機器學習的三要點：

(1) exists some underlying pattern to be learned.

(2) no programmable (easy) definition.

(3) somehow there is data about pattern.

(a) 不適合。去除掉物理等因素，這些中獎號碼的產生是公正且隨機，則不會有適合的模型能夠增進預測的效果。

(b) 不適合。因為有詳細的規則能去計算。

(c) 不適合。因為有詳細的規則能去計算。

(d) 適合。可藉由告訴機器哪種芒果是好的，哪種是不好，讓他判斷芒果的品質。

2. e

(a) 丟硬幣的結果跟是否為垃圾郵件的結果無相關性，丟硬幣是一個簡單的機率模型，此模型沒辦法幫助機器去學習並且增進。

(b) 可以簡單定義模型。

(c) 可以簡單定義模型。

(d) 可以找到規則去定義模型。

3. d

Handwritten mathematical derivation showing the relationship between R^2 and ρ^2 :

$$R^2 = \max_n \|x_n\|^2$$
$$\max_n \left\| \frac{x_n}{4} \right\|^2 = \frac{1}{16} \max_n \|x_n\|^2 = \frac{1}{16} R^2$$
$$\rho = \min_n y_n \cdot \frac{w_f^T x_n}{\|w_f\|}$$
$$\min_n y_n \cdot \frac{w_f^T x_n}{\|w_f\|} \cdot \frac{1}{4} = \frac{1}{4} \rho$$
$$\frac{R^2}{\rho^2} \Rightarrow \frac{\frac{1}{16} R^2}{\frac{1}{4^2} \rho^2} = \frac{R^2}{\rho^2} \quad \#$$

4. c

$$\begin{aligned} \rho^2 &= \max_n \|X_n\|^2 \\ \rho &= \min_n y_n \frac{W^T}{\|W\|} X_n = \min_n \frac{|W^T X_n|}{\|W\|} \\ \hat{\rho} &= \min_{n \in \{1, 2, \dots, N\}} \frac{|W^T X_n|}{\|W\| \|X_n\|} = \frac{\rho}{\|X_n\|} \\ \frac{R^2}{\rho^2} &= \frac{\|X_n\|^2}{\rho^2} = \left(\frac{\|X_n\|}{\rho} \right)^2 = (\hat{\rho}^{-1})^2 = \hat{\rho}^{-2} \end{aligned}$$

5. d

$$\begin{aligned} W_{t+1} &\leftarrow W_t + \eta y_n X_n \\ &\quad \begin{matrix} > 0 & & < 0 \end{matrix} \\ y_n X_n W_{t+1} &\leftarrow y_n X_n W_t + \eta (y_n X_n)^2 \\ y_n X_n W_t + \eta y_n X_n y_n X_n &> 0 \\ \eta (y_n X_n)^2 &> -y_n X_n W_t \\ \eta &> \frac{-y_n X_n W_t}{y_n^2 X_n^2} \\ \eta &> \frac{-y_n X_n W_t}{\|X_n\|^2} \quad \# \end{aligned}$$

6. 不會

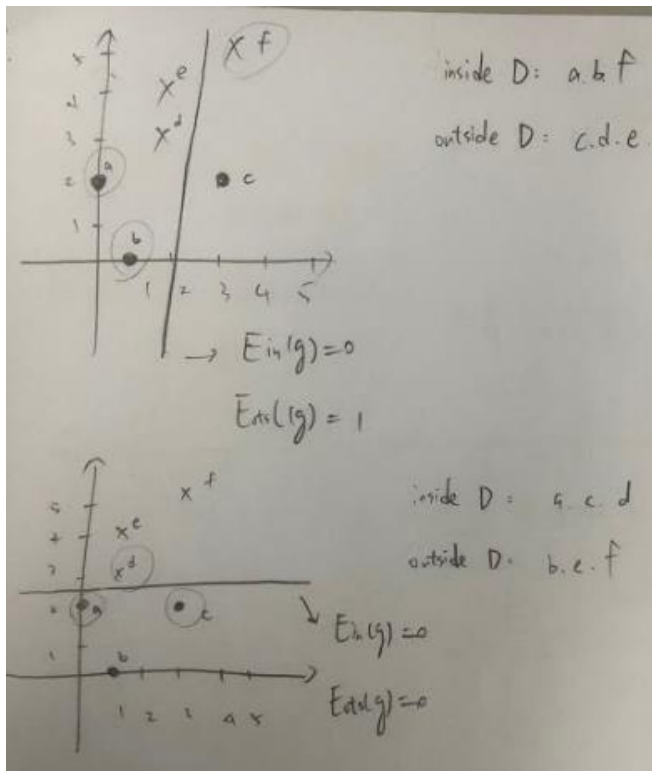
7. e

(e) 照題目中 **self-practicing** 的定義，機器藉由跟自己下棋，每盤棋的結果會根據 **judge environment** 來告訴他這是好的還是不好的，而不是告訴他正確的下法，機器根據一盤又一盤的經驗去學習下棋技巧，因此 **self-practicing** 是屬於強化學習。

8. b

輸出 **y** 是一個很大的多類別問題，很複雜，為 **structured learning**
使用的資料中部分有標準答案，部分沒有，為 **semi-supervised learning**
資料為一整個資料集，整批進去學習，為 **batch learning**
答案唯一連串的行為，屬於比較抽象的特徵，為 **raw features**

9. e



10. b

10. $P \leq z e^{-z^2 N}$

$\log P \leq \log z + \log e^{-z^2 N}$

$\log P \leq \log z + (-z^2 N)$

$z^2 N \leq \log z - \log P$

$N \leq \frac{1}{z^2} \log \frac{z}{P}$

$P[|V-M| < \epsilon] \geq 1 - \delta$

$P[|V-M| > \epsilon] \leq \delta$

$N \leq \frac{1}{z^2} \log \frac{z}{\delta}$

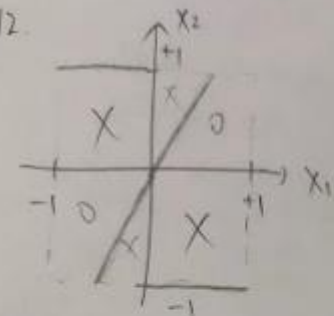
11. c

$f(x) = \text{sign}(x)$ $[-1, +1]$. A ∞ samples
 $h_1(x) = \text{sign}(2x_1 - x_2)$ $\text{sign}(x) > 0 \rightarrow 1$
 $h_2(x) = \text{sign}(x_2)$ $\leq 0 \rightarrow -1$
 $E_{in}(h_2) = 0$
 $f_{in} = h_2(x)$
 $\text{sign}(x_1) = \text{sign}(x_2)$ $(\frac{1}{2})^5$

$2x_1 - x_2 = x_2$
 $2x_1 = 2x_2$
 $x_1 = x_2$

12. d

12.



5 examples

0 wrong: $(\frac{3}{8})^5$

1 wrong: $\frac{1}{8} \cdot \frac{1}{2} \cdot (\frac{3}{8})^3 \cdot \frac{5!}{3!} = \frac{3843}{2^{15}} \#$

2 wrong: $(\frac{1}{8})^2 \cdot (\frac{1}{2})^2 \cdot \frac{3}{8} \cdot \frac{5!}{2!2!}$

$(\frac{3}{8})^5 + \left[\frac{1}{8} \cdot \frac{1}{2} \cdot (\frac{3}{8})^3 \cdot \frac{5!}{3!} \right] + \left[(\frac{1}{8})^2 \cdot (\frac{1}{2})^2 \cdot \frac{3}{8} \cdot \frac{5!}{2!2!} \right]$

13. c

$$\begin{aligned}
 & P_D(\text{BAD D}) \\
 &= P_D(\text{BAD D for } h_1 \text{ or BAD D for } h_2 \text{ or } \dots \text{ or BAD D for } h_n) \\
 &\leq P_D(\text{BAD D for } h_1) + P_D(\text{BAD D for } h_2) + \dots + P_D(\text{BAD D for } h_n) \\
 &\leq z e^{-z \epsilon^2 N} + z e^{-z \epsilon^2 N} + \dots + z e^{-z \epsilon^2 N} \\
 &= z M e^{-z \epsilon^2 N} \\
 &H = \{h_1, h_2, \dots, h_{2d}\} \\
 &P(\text{BAD D for } H) \leq C \cdot z e^{-z \epsilon^2 N} \\
 &C = M = z d.
 \end{aligned}$$

14. d

14.

	green 3	green 1	orange 2	green 2	green 4	green 5
A	o	o	o	✓	✓	o
B	✓	o	o	✓	✓	o
C	o	o	✓	o	o	o
D	✓	o	o	✓	o	✓

15. c

green 1 2 3 4 5 6

A 0 ✓ 0 ✓ 0 ✓

B 0 ✓ ✓ ✓ 0 0

C 0 0 0 0 0 ✓

D 0 ✓ ✓ 0 ✓

$$0 + \left(\frac{3}{4}\right)^5 + \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 + \left(\frac{1}{4}\right)^5 + \left(\frac{1}{2}\right)^5 = \frac{275}{1024}$$

\downarrow \downarrow
 与2重叠 与2重叠

$$\frac{275}{1024} - \frac{1}{1024} = \frac{274}{1024} \quad \#$$

\downarrow
 6A重叠